GENERAL MATHEMATICS

UNITS 3 & 4

CAMBRIDGE SENIOR MATHEMATICS
FOR QUEENSLAND

PETER JONES | MICHAEL EVANS | KAY LIPSON | KYLE STAGGARD

Consultants: Ray Minns | Steve Sisson

INCLUDES INTERACTIVE TEXTBOOK POWERED BY CAMBRIDGE HOTMATHS

CAMBRIDGE UNIVERSITY PRESS

ISBN 978-1-108-46154-2

Cambridge University Press

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India
79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108461542

First published 2019
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Cover and text designed by Sardine Design
Typeset by diacriTech
Printed in China by C & C Offset Printing Co. Ltd.

A catalogue record for this book is available from the National Library of Australia at www.nla.gov.au

ISBN 978-1-108-46154-2 Paperback

Additional resources for this publication at www.cambridge.edu.au/GO

Contents

ISBN 978-1-108-46154-2

Cambridge University Press

ISBN 978-1-108-46154-2

ISBN 978-1-108-46154-2

ISBN 978-1-108-46154-2

ISBN 978-1-108-46154-2

Cambridge University Press

Online appendices

Appendix 1 Review of computation and practical arithmetic

A copy of the review chapter of computation in the Units 1 & 2 textbook is provided in the Interactive Textbook.

Appendix 2 Online guides to using technology

These online guides are accessed through the Interactive Textbook or PDF Textbook

Online assessment and examination practice items

These items, accessed via the Interactive Textbook and Online Teaching Suite, are listed at the ends of the revision chapters 6 and 13.

Note: *A printable copy of a formula sheet is available in the Interactive Textbook*

ISBN 978-1-108-46154-2

About the lead authors and consultants

Peter Jones is Emeritus Professor of Statistics at Swinburne University. He has worked as a consultant to the Australian Curriculum, Assessment and Reporting Authority (ACARA) on the development of the Australian Curriculum, General Mathematics. He has worked in curriculum development and teacher professional development and has written examinations and has been a chief examiner for many years. He has been a writer of textbooks for the Years 11 and 12 General Mathematics courses for twenty-five years. His textbooks have become the most popular senior mathematics textbooks in Australia.

Michael Evans was a consultant to ACARA on the writing of the Australian Curriculum on which the new Queensland syllabus is based. He is a consultant with the Australian Mathematical Sciences Institute, and is coordinating author of the ICE-EM 7–10 series also published by Cambridge. He has also been active in the Australian Mathematics Trust, being involved with the writing of enrichment material and competition questions. He has many years' experience as a Chief Examiner and Chairperson of examination panels.

Ray Minns is Head of Mathematics at Northpine Christian College, Dakabin.

Steve Sisson is Head of Mathematics at Redeemer Lutheran College, Rochedale.

ISBN 978-1-108-46154-2

Introduction and overview

Cambridge Senior Mathematics for Queensland General Mathematics Units 3&4 has been written for the QCAA syllabus to be implemented in Year 12 from 2020. As well as covering all the subject matter of the Queensland General Mathematics syllabus, the package addresses its objectives, assessment, underpinning factors, formula sheet, and pedagogical and conceptual frameworks.

Its four components—the print textbook, the downloadable PDF textbook, the online Interactive Textbook and the Online Teaching Resource*—contain a huge range of resources, including worked solutions, available to schools in a single package at one convenient price. There are no extra subscriptions or per-student charges to pay.
**The Online Teaching Resource is included with class adoptions, conditions apply.*

▶ Overview of the print textbook (shown on the page opposite)

1. The print book comes with linked digital resources and access to some previous years' resources.
2. Chapter **outcomes** are listed at the beginning of each chapter under the syllabus units and topics.
3. Each section and most exercises begin at the top of the page to make them easy to find and access.
4. Step-by-step **worked examples** with precise explanations and **video** versions encourage independent learning, and are linked to exercises.
5. Important concepts are formatted in boxes for easy reference.
6. **Degree of difficulty categories** are indicated for exercises and are featured in the **revision chapters**. **Degree of difficulty** classification of questions: in the exercises, questions are classified as ***simple familiar*** **SF**, ***complex familiar*** **CF**, and ***complex unfamiliar*** **CU** questions. The revision chapters described below also contain model questions for each of these categories, and tests are also provided in the teacher resources, made up of such categorised model questions.
7. **Problem-solving and modelling questions** are included in Unit 3. QCAA guidelines have been followed.
8. Two **revision chapters** are provided, one for each unit, with multiple-choice, short-answer and extended response questions. Each exercise covers a chapter and is divided into degree of difficulty categories, with problem-solving and modelling questions and investigations in Unit 3.
9. Technology is supported via **scientific calculator** guidance, **spreadsheets** and **Desmos widgets**.
10. **Spreadsheet activities** are integrated throughout the text, with accompanying Excel files in the Interactive Textbook.
11. Chapter reviews contain a **chapter summary** and **multiple-choice, short-answer** and **extended-response questions.**

ISBN 978-1-108-46154-2

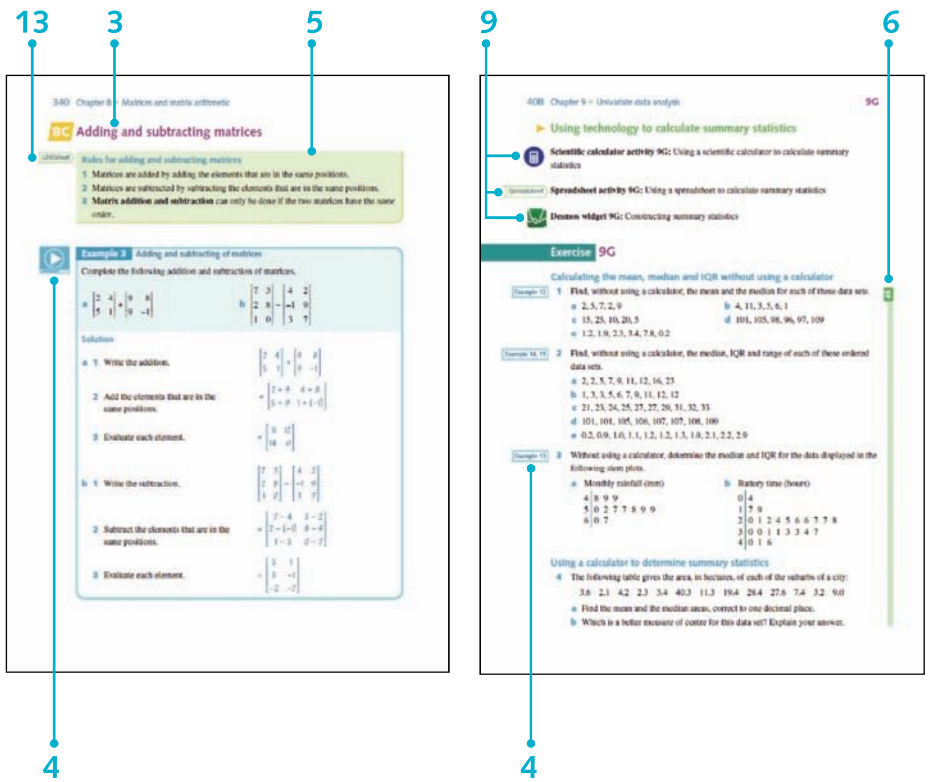
12 A comprehensive **glossary** is included.

13 Additional linked resources in the Interactive Textbook and Online Teaching Suite are indicated in the text, such as:

- videos of worked examples and certain concepts
- skillsheets
- Desmos widgets
- spreadsheet activities
- calculator activities.

PRINT TEXTBOOK

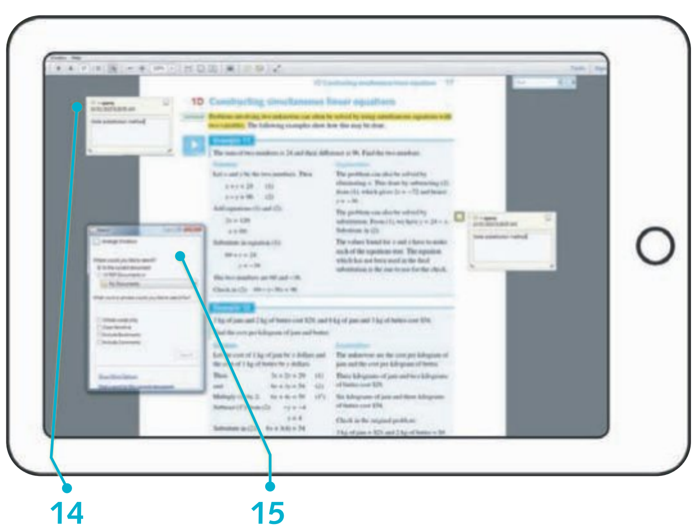
Numbers refer to the descriptions in the overview.



Overview of the downloadable PDF textbook

14 The convenience of a downloadable PDF textbook has been retained for times when users cannot go online.

15 **PDF annotation** and **search features** are enabled.



ISBN 978-1-108-46154-2

Cambridge University Press

Overview of the Interactive Textbook (shown on the page opposite)

The **Interactive Textbook** (ITB), an online HTML version of the print textbook powered by the HOTmaths platform, is included with the print book or available as a separate digital-only product.

16 The material is formatted for on-screen use, with a convenient and easy-to-use navigation system and links to all resources.

17 The new **Workspaces** enable students to enter working and answers online and to save them. Input is by typing, with the help of a symbol palette, handwriting and drawing on tablets, or by uploading images of writing or drawing.

18 The new **self-assessment tools** enable students to check answers, mark their own work, and rate their confidence level in their work. This helps develop responsibility for learning, and communicates progress and performance to the teacher. Student accounts can be linked to the learning management system used by the teacher in the Online Teaching Suite.

19 Examples have **video versions** to encourage independent learning.

20 **Worked solutions** are included and can be enabled or disabled in the student accounts by the teacher.

21 Interactive **Desmos widgets** demonstrate key concepts and enable students to visualise the mathematics.

22 The **Desmos scientific calculator** and geometry tool is also available for students to use for their own calculations and exploration.

23 **Quick quizzes** containing automarked multiple-choice questions enable students to check their understanding.

24 **Definitions** pop up for key terms in the text, and are also provided in a **dictionary**.

25 Messages from teacher assign tasks and tests.

26 **Practice assessment tasks and exam-style papers** are provided in downloadable PDF and Word files.

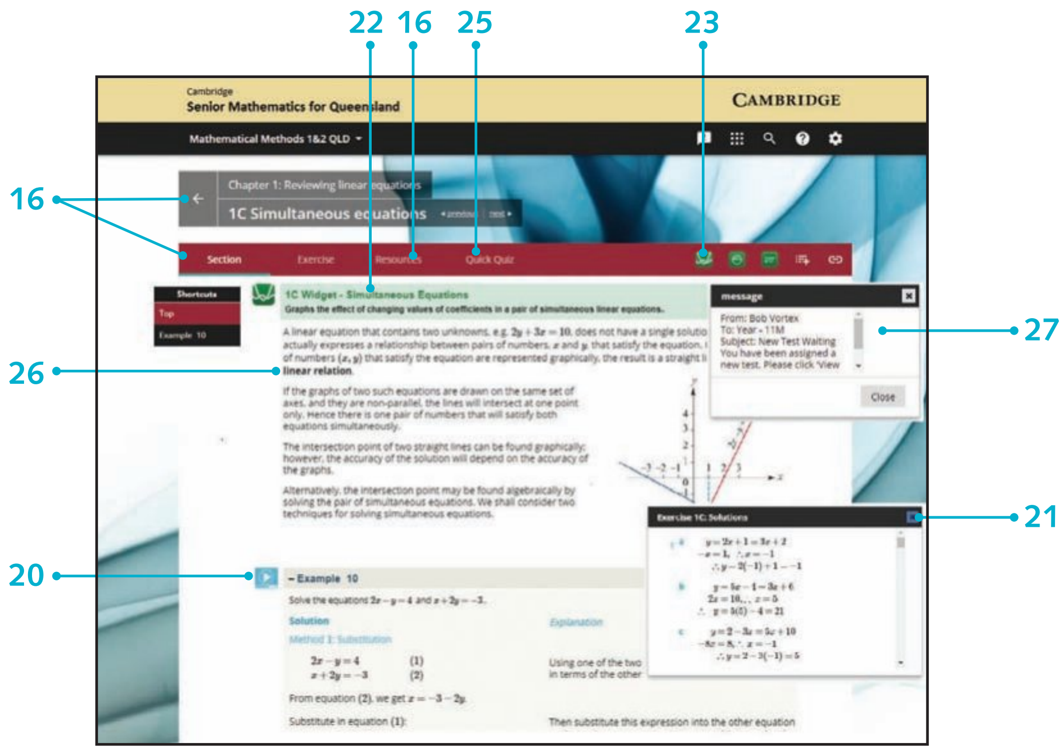
27 **Spreadsheets** are provided in Excel format.

28 **Calculator** guides are provided as PDFs.

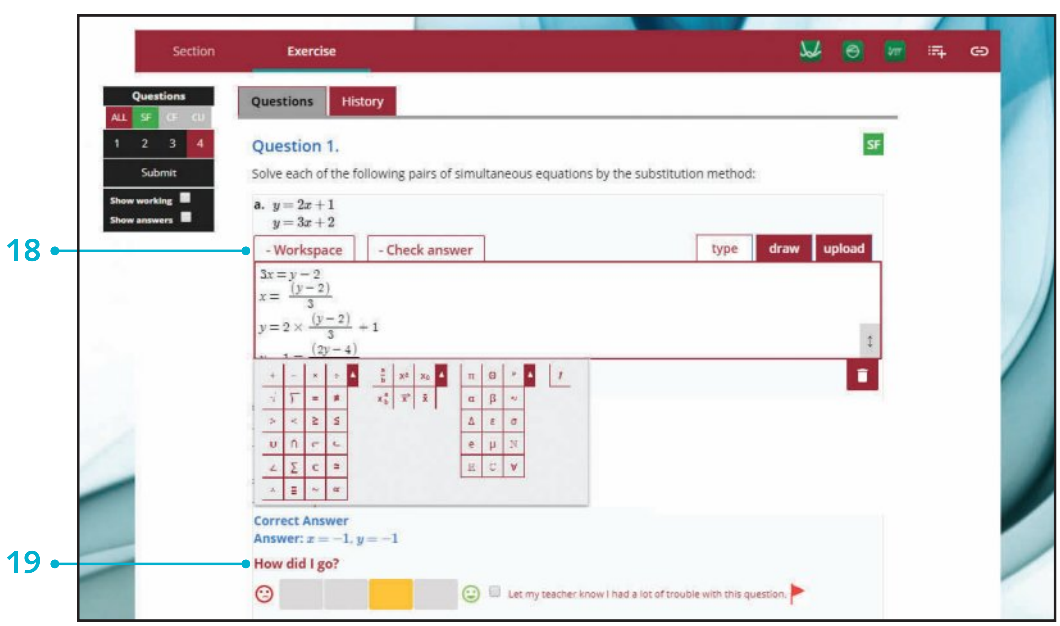
ISBN 978-1-108-46154-2

INTERACTIVE TEXTBOOK POWERED BY THE HOTmaths PLATFORM

A selection of features is shown. Numbers refer to the descriptions on the opposite page. HOTmaths platform features are updated regularly.



WORKSPACES AND SELF-ASSESSMENT



ISBN 978-1-108-46154-2

Cambridge University Press

Overview of the Online Teaching Suite Powered by the HOTmaths platform (shown below)

The Online Teaching Suite is automatically enabled with a teacher account and is integrated with the teacher's copy of the Interactive Textbook. All the assets and resources are in one place for easy access. The features include:

29 The HOTmaths **learning management system** with class and student analytics and reports, and communication tools.

30 Teacher's view of a student's working, scores and self-assessment, which they can comment upon.

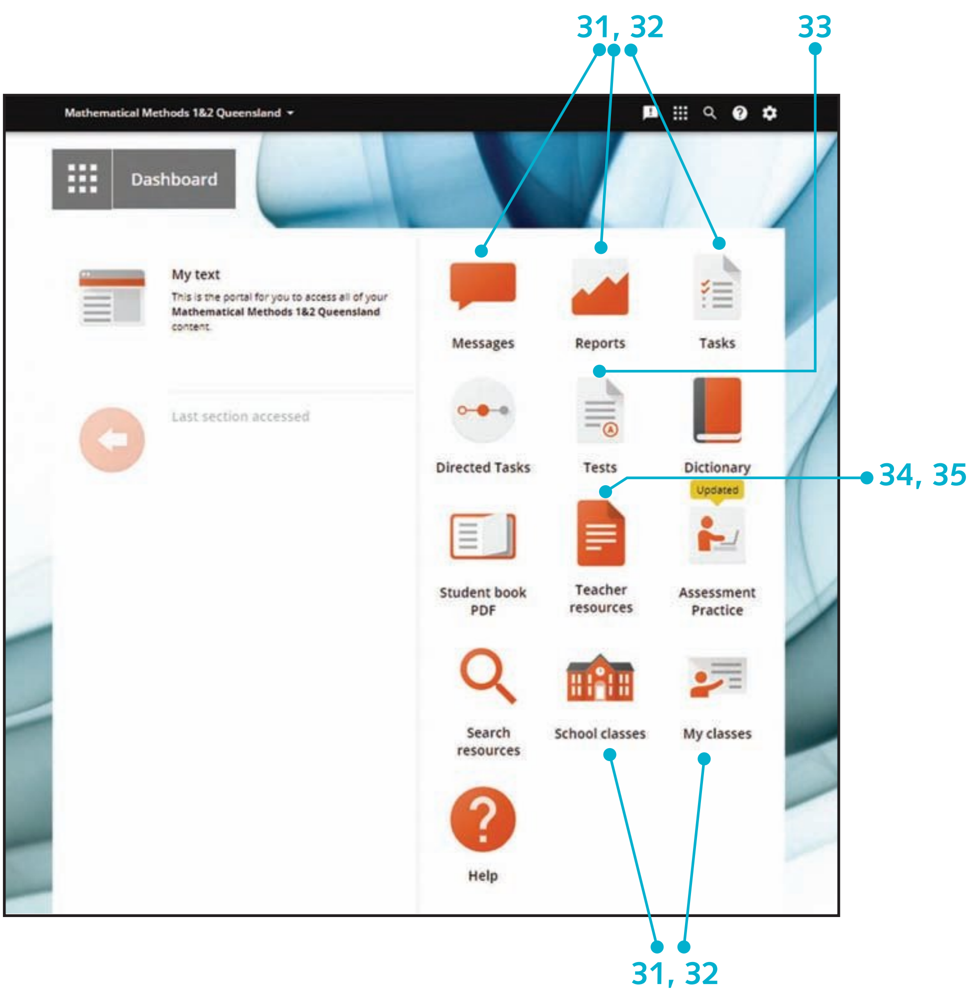
31 A HOTmaths-style **test generator**.

32 Chapter test **worksheets** and **assessment tasks** and exam **practice papers**.

33 Editable **curriculum grids** and **teaching programs**.

ONLINE TEACHING SUITE POWERED BY THE HOTmaths PLATFORM

Numbers refer to the descriptions above. Screenshots are taken from the Mathematical Methods textbook in this series. HOTmaths platform features are updated regularly



ISBN 978-1-108-46154-2

Acknowledgements

The author and publisher wish to thank the following sources for permission to reproduce material:

Cover: © Getty Images / DuxX

Images: © Getty Images / Sorin Rechitan, Chapter 1 Opener / Michael Willson / AFL Photos, 1A (1) / FatCamera, 1A (2) / Tetra Images, 1B (1) / kali9, 1C (1) / oxygen, 1C (2) / Yuichiro Chino, 1C (3) / Matthew Gollan, 1D (1) / James Lauritx, 7F (1) / MirageC, Chapter 2 Opener / Sean Gladwell, 2A (1) / Maskot, 2A (2), 2B (5) / fotog, 2B (1) / Johner Images, 2B (3) / luoman, 2B (3) / Thomas_EyeDesign, 2B (4) / Maica, 2B (6) / Thinkstock, 2B (7) / Sam Edwards, 2D (1) / Jamie Grill, 2D (2) / M King, Chapter 3 Opener / Mint Images, 3A (1) / Murray Severn, 3A (2) / Creativ Studio Heinemann, 3A (3) / Westend61, 3B (1), 3F (2), 7B (1), 8A (2), 7C (2) / R Productions Ltd., 3B (2) / Marco Secchi, 3B (3) / Michael H, 3B (4), 4F (2), 5A (1) / Francesco Carta fotografo, 3B (5) / Phil Copp, 3D (1) / Benjamin Ang, 3D (2) / Hero Images, 3D (3) / courtneyk, 3E (1) / Compassionate Eye Foundation, 3E (2) / pamspix, 3F (1) / Hero Images, 3F (3) / P Greenway, Chapter 4 Opener / K Nonnittayanan, 4A (1) / shunli zha, 4A (2) / C Anderson Productions Pty Ltd., 4B (1) / amittus, 4C (1) / SoStock, 4D (1) / C Pundej, 4E (1) / Viorika, 4E (2) / www.boelke-art.de, 4F (1) / webphotographer, 4G (1) / Luisella Sem, Chpater 5 Opener / James Crawley, 3B (3) / K Nafatni, 5C (1) / WIN-Initiative, Chapter 6 Opener / Tobias Titz, Chapter 7 Opener / oxygen, 7A (1) / pawel.gaul, 7B (2) / V Changyencham, 7B (3) / Peter Parks, 7C (1) / F Carta Fotografo, 7D (1) / AFP Contributor, 7D (2) / Parinda Yatha, 8A (1) / Chris Whitehea, 8C (1) / Tim Platt, Chapter 9 Opener / A Okek, 9A (1) / TexPhoto, 9A (2) / S Mekvorawuth, 9A (3) / K Phanthong, 9B (1) / Hill Street Studios, 9D (1) / Saul Gravy, Chapter 10 Opener / Douglas Klug, 10A (2) / Juice Images, 10B (1) / Doug Wenjie, 10C (1) / Nikada, 10D (1) / M Purdie, 10D (2) / Andrew Merry, 10E (1) / Jacobs Stock Photography Ltd., 10F (2) / P Openshaw, Chapter 11 Opener / R Kneschke, 11B (1) / Kali9, 11B (2) / Hero Images, 11B (3) / alengo, 11C (1) / V Khamsong, 11C (2) / Karn Bulsuk, 11C (3) / Mandy Disher, Chapter 12 Opener / Echo, 12A (1) / A Bergman, 12A (2) / Mint Images, 12A (3) / Chris Ryan, 12B (1) / F Cirou, 12B (2) / Hinterhaus Productions, 12C (1) / borchee, Chapter 13 Opener / Coal photography / Alexander Legaree, 13B (1) / Image Source, 13C (1) / Sam Edwards, 13C (2) / lacer, 13C (3); Copyright © 2019 The Puzzle Museum - J. Dalgety, 10G (1).

Every effort has been made to trace and acknowledge copyright. The publisher apologises for any accidental infringement and welcomes information that would redress this situation.

ISBN 978-1-108-46154-2

Cambridge University Press

1 Investigating associations between two variables

UNIT 3 BIVARIATE DATA, SEQUENCES AND CHANGE, AND EARTH GEOMETRY

Topic 1 Bivariate data analysis

- How do we define bivariate data?
- How do we construct two-way tables?
- How do we interpret and identify patterns in two-way tables using percentages?
- How do we construct a scatterplot?
- How do we describe an association between two numerical variables in terms of direction, form and strength?
- How do we calculate and interpret the correlation coefficient?
- How do we differentiate between association and causation?

ISBN 978-1-108-46154-2

Introduction

Much of the analysis that is carried out in statistics is not concerned with a single variable, but rather with the relationships between two or more **variables**. Questions such as 'Is the new treatment for headache more effective than the old treatment?', 'Are females more likely to vote for the Greens party than males?' or 'Are younger people more knowledgeable about environmental issues than older people?' are concerned with understanding the association between two variables.

To answer these questions requires two items of data for each subject. So, for example, for the first question we need to know for each person in the study which treatment they used and how effective the treatment has been for them.

When there are data from two variables for the same subject, this is called **bivariate data**.

1A Bivariate data – classifying the variables

To determine how to answer questions involving two variables requires the variables to be clearly defined.

Categorical and numerical variables

You will recall from General Mathematics in Year 11 we defined two classifications of variables, *categorical* and *numerical* variables:

- **Categorical variables** generate data values that are names or labels, such as *sex* (male, female) or *coffee size* (small, medium, large).
- **Numerical variables** generate data values that are numbers, usually resulting from counting or measuring, such as *number of brothers* (0, 1, 2, …) or *hand span* (cm).

Example 1 Identifying variables as categorical or numerical

Identify each of the following variables as either categorical or numerical:

a weight (kg)

b favourite colour

c support for same sex marriage (yes, no)

d number of pine trees per acre of forest

e attitude to lowering the driving age (strongly agree, agree, no opinion, disagree, strongly disagree)

ISBN 978-1-108-46154-2

Cambridge University Press

Solution

a Weight	*Numerical – the data values arise from measuring*
b Favourite colour	*Categorical – the data values are labels such as red or blue*
c Support for same sex marriage	*Categorical – the data values are labels*
d Number of pine trees per acre of forest	*Numerical – the data values arise from counting*
e Attitude to lowering the driving age	*Categorical – the data values are labels*

The first step in investigating the association between two variables is to classify each variable as either categorical or numerical. What can we say about the types of variables involved in the questions previously posed?

'Is the new treatment for headache more effective than the old treatment?'

To investigate this question requires firstly a definition of 'effective'. Suppose that the effectiveness of the treatment is to be measured by the time it takes for the headache to be relieved, measured in minutes. Then the two variables in this question are *type of treatment*, a categorical variable taking the values 'new' and 'old', and *time taken for the headache to be relieved*, a numerical variable. Thus, investigation of a question like this can be classified as *investigating the association between a categorical variable and a numerical variable*.

'Are females more likely to vote for the Greens party than males?'

This question involves two variables, *sex*, which is a categorical variable taking the values 'male' and 'female', and *vote for the Greens*, which also is a categorical variable taking the values 'yes' and 'no'. Investigation of a question like this can be classified as *investigating the association between two categorical variables*.

'Are younger people more knowledgeable about environmental issues than older people?'

If the *age* of the respondent is measured in years, then *age* is a numerical variable. Suppose *knowledge of environmental issues* is measured with a sequence of questions and the respondent is given a score out of 100, then *knowledge of environmental issues* is also a numerical variable. Investigation of a question like this can be classified as *investigating the association between two numerical variables*.

Numerical variables can be divided into ***continuous variables*** which represent a quantity that is measured rather than counted, for example the weights of people in kilograms, and ***discrete variables*** which represent a quantity that is determined by counting, for example, the number of people waiting in a queue.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 2 **Identifying associations as categorical or numerical**

For each of the following questions, determine if they involve investigating associations between:

- one numerical variable and one categorical variable or
- two categorical variables or
- two numerical variables.

a Are younger people (age measured in years) more likely to believe in astrology (measured as yes or no) than older people?

b Do students spend more hours studying each week get higher test scores?

c Are people who have a driver's licence more likely to be in favour of lowering the driving age?

Solution

a One numerical variable (age) and one categorical variable (belief in astrology)

b Two numerical variables (hours studied per week and test score)

c Two categorical variables (have a driver's licence and support for lowering the driving age)

Identifying response and explanatory variables

The second step in investigating the association between two variables is to determine which of the two variables is the explanatory variable and which is the response variable. We use the explanatory variable to explain the associated changes in the response variable. For example:

- 'Is the new treatment for headache more effective than the old treatment?'
  - *Type of treatment* is the **explanatory variable** as it may explain any changes in *time taken for the headache to be relieved*.
  - *Time taken for the headache to be relieved* is the **response variable** as changes could occur in response to the *type of treatment* used.
- 'Are females more likely to vote for the Greens party than males?'
  - *Sex* is the explanatory variable as it may explain any difference in *vote for the Greens*.
  - *Vote for the Greens* is the response variable as differences could occur in response to the *sex* of the voter.
- 'Are younger people more knowledgeable about environmental issues than older people?'
  - *Age* is the explanatory variable as it may explain any changes in *knowledge of environmental issues*.
  - *Knowledge of environmental issues* is the response variable as changes could occur in response to the *age* of the person.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 3 Identifying response and explanatory variables

We wish to investigate the question, 'Does the time it takes a student to travel to school depend on their mode of transport?' The variables here are *time* and *mode of transport*. Which is the response variable (RV) and which is the explanatory variable (EV)?

Solution

In asking the question in this way, we are suggesting that a student's *mode of transport* might explain the differences we observe in the time it takes students to travel to school.

EV: mode of transport
RV: time

Example 4 Identifying response and explanatory variables

Can we predict people's height (in cm) from their wrist circumference (cm)? The variables in this investigation are *height* and *wrist measurement*. Which is the response variable (RV) and which is the explanatory variable (EV)?

Solution

Since we wish to predict height from wrist circumference, we are using *wrist measurement* as the predictor or explanatory variable. *Height* is then the response variable.

EV: wrist measurement
RV: height

It is important to note that, in Example 4, we could have asked the question the other way around; that is, 'Can we predict people's wrist measurement from their height?' In that case *height* would be the explanatory variable, and *wrist measurement* would be the response variable. The way we ask our statistical question is an important factor when there is no obvious explanatory variable.

Response and explanatory variables

When investigating the association between two variables, the *explanatory variable (EV)* is the variable we expect to *explain* or *predict* the value of the *response variable (RV)*.

Note: The explanatory variable is sometimes called the independent variable (IV) and the response variable the dependent variable (DV).

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 1A

Identifying variables as categorical or numerical

Example 1

1 Identify each of the following variables as either categorical or numerical.

SF

a income (low, medium, high)
b favourite TV show
c time taken to drive to work in minutes
d emotional intelligence, as measured on a standardised psychological test on a scale from 1–100
e self-assessed state of health (1 = excellent, 2 = good, 3 = satisfactory, 4 = poor, 5 = very poor)
f temperature (°C)
g weekly salary ($)
h weekly salary (1 = below average, 2 = average, 3 = above average)
i weekly salary (less than $500, $500–$999, $1000–$1999, more than $2000)

Identifying associations as categorical or numerical

Example 2

2 For each of the following questions, determine if they involve investigating associations between:

- one numerical and one categorical variable or
- two categorical variables or
- two numerical variables.

a Are males and females equally likely to be in favour of same sex marriage?
b Do Year 11 students watch more hours of television each week than Year 12 students?
c Do countries with higher household incomes ($) tend to have lower infant mortality rates (deaths/1000 births)?
d Is there a relationship between attitude to gun control and country of birth?

Identifying response and explanatory variables

Example 3, 4

For each of the following situations identify the explanatory variable (EV) and the response variable (RV). In each situation the variable names are *italicised.*

3 **a** We wish to investigate whether a fish's *toxicity* can be predicted from its *colour*.
b The relationship between *weight loss* and *type of diet* is to be investigated.
c We wish to investigate the relationship between a used car's *age* and its *price*.
d It is suggested that the *cost* of heating in a house depends on the type of *fuel* used.
e The relationship between the *house price* and its *location* is to be investigated.

4 The following pairs of variables are related. Which is likely to be the explanatory variable? The variable names are italicised.

a *exercise level* and *age*

b *years of education* and *salary level*

c *comfort level* and *temperature*

d *time of year* and *incidence of hay fever*

e *age group* and *musical taste*

f *AFL team supported* and *state of residence*

5 For each of the following pairs of variables, determine:

- which are numerical and which are categorical, and
- which is the explanatory variable and which is the response variable.

a *sex* and *attitude to lowering the legal drinking age*

b *hours of study per week* and *hours spent per week using social media for Year* 12 *students*

c *gestation time* and *birth weight of babies*

d *sex* and *hours spent per week using social media for Year* 12 *students*

e *voting preference* (Liberal, Labor, Greens, other) and *support for tax cuts*

ISBN 978-1-108-46154-2

1B Investigating associations between two categorical variables

To begin our analysis of data arising from two categorical variables we will introduce a table used to summarise bivariate data.

Two-way frequency table

It has been suggested that males and females have differing attitudes to gun control; that is, that support for gun control depends on the sex of the person. How might we investigate this relationship? Suppose we ask a sample of three people about their *attitude to gun control*, and we also record their *sex*. The resulting data for the three people might look like this:

Subject no.	Sex	Attitude to gun control
1	Female	For
2	Male	For
3	Male	Against

The first thing to note is that these two variables, *attitude to gun control* (for or against) and *sex* (male or female), are both categorical variables. Categorical data are usually presented in the form of a frequency table.

Suppose we continue until we have interviewed a sample of 100 people, and we find that there are 58 males and 42 females. We can present this result in a frequency table as shown to the right.

Sex	Frequency
Male	58
Female	42
Total	100

From this table, we can see that there were more males than females in our sample.

Suppose also when we record the attitude to gun control, we might have 62 'for' and 38 'against' gun control. Again, we could present these results in a frequency table as shown to the right.

Attitude to gun control	Frequency
For	62
Against	38
Total	100

From this table, we can see that more people in the sample were for gun control than against gun control. However, we cannot tell from the information contained in the tables whether *attitude to gun control* depends on the *sex* of the person. To do this we need to construct a **two-way frequency table**, which gives *both* the *attitude to gun control* and the *sex* for each person in the sample.

ISBN 978-1-108-46154-2

Cambridge University Press

We begin by counting the number of people in the sample who are:

- male and for gun control
- female and for gun control
- male and against gun control
- female and against gun control.

Suppose again from our sample of 100 people we find the following frequencies:

- 32 males are for gun control
- 30 females are for gun control
- 26 males are against gun control
- 12 females are against gun control.

Explanatory and response variables in two-way frequency tables

Before we set up the two-way frequency table, we need to decide which is the explanatory variable and which is the response variable of the two variables. Since we think that a person's attitude to gun control might depend on their sex, but not the other way around, then:

- *sex* is the explanatory variable (EV)
- *attitude to gun control* is the response variable (RV).

In two-way frequency tables, it is conventional to let the categories of the *response variable* label the *rows* of the table and the categories of the *explanatory variable* label the *columns* of the table. Following this convention, we can create the following two-way frequency table.

	Sex	
Attitude to gun control	Male	Female
For	32	30
Against	26	12

To complete the table, it is usual to calculate the row and column sums, as shown below.

	Sex			
Attitude to gun control	Male	Female	Total	
For	32	30	62	Row sum
Against	26	12	38	Row sum
Total	58	42	100	
	Column sum	Column sum		

The shaded regions in the table are called the *cells* of the table. It is the numbers in these cells that we look at when investigating the relationship between the two variables.

ISBN 978-1-108-46154-2

Example 5 Constructing a two-way frequency table

The following data were obtained when a sample of 10 Year 9 students were asked if they intended to go to university. The sex of the student was also recorded.

Student No.	Sex	Intends to go to university	Student No.	Sex	Intends to go to university
1	Female	Yes	6	Male	Yes
2	Male	Yes	7	Female	Yes
3	Female	No	8	Male	No
4	Female	Yes	9	Female	No
5	Male	No	10	Female	Yes

Create a two-way frequency table from these data.

Solution

1 We first need to identify the explanatory variable and the response variable.

It is possible that a student's intention to go to university may depend on their sex, but not the other way around. Thus, sex is the explanatory variable and intends to go to university is the response variable.

2 Create the table showing the values of *sex* labelling the columns, and *intends to go to university* labelling the rows.

	Sex	
Intends to go to university	Male	Female
Yes		
No		

3 Consider Student 1, who is female and indicated yes to go to university. Place a mark in the corresponding cell of the table.

	Sex	
Intends to go to university	Male	Female
Yes		I
No		

4 Go through the data set one person at a time, placing a mark in the appropriate cell for each person.

	Sex	
Intends to go to university	Male	Female
Yes	II	IIII
No	II	II

ISBN 978-1-108-46154-2

Cambridge University Press

5 Finally, tally the marks in each cell, and then calculate the row and columns sums. Make sure the total adds to the number of students in the sample.

	Sex		
Intends to go to university	Male	Female	Total
Yes	2	4	6
No	2	2	4
Total	4	6	10

Consider again the two-way frequency table created to investigate the association between sex and attitude to gun control. This table tells us that more males are in favour of gun control than females. But is this just due to the fact that there were more males in the sample, or are males really more in favour of gun control than females? To help us answer this question we need to express the frequencies in each cell as **percentage frequencies**.

Two-way percentage frequency table

When the two-way frequency table has been constructed so that the values of the explanatory variable label the columns, then we calculate *column percentages* to help us investigate the association. This will give us the percentage of males and the percentage of females for and against gun control, which can then be compared.

Column percentages are determined by dividing each of the cell frequencies by the relevant column sums. Thus, the percentage of:

- males who are for gun control is: $\frac{32}{58} \times 100 = 55.2\%$
- males who are against gun control is: $\frac{26}{58} \times 100 = 44.8\%$
- females who are for gun control is: $\frac{30}{42} \times 100 = 71.4\%$
- females who are against gun control is: $\frac{12}{42} \times 100 = 28.6\%$

Note: Unless small percentages are involved, it is usual to round percentages to one decimal place in tables.

	Sex	
Attitude	Male	Female
For	55.2%	71.4%
Against	44.8%	28.6%
Total	100.0%	100.0%

ISBN 978-1-108-46154-2

Using percentages to identify relationships between variables

Calculating the values in the table as percentages enables us to compare the attitudes of males and females on an equal footing. From the table, we see that 55.2% of males in the sample were for gun control compared to 71.4% of the females. This means that the females in the sample were more supportive of gun control than the males. This reverses what the frequencies showed. It is easy to be misled if you just compare frequencies in a two-way frequency table.

The fact that the percentage of 'males for gun control' differs from the percentage of 'females for gun control' indicates that a person's attitude to gun control *depends* on their sex. Thus, we can say that the variables *attitude to gun control* and *sex* are *associated*.

If the variables *attitude to gun control* and *sex* were *not associated*, we would expect approximately equal percentages of males and females to be 'for' gun control.

We could have also arrived at this conclusion by focusing our attention on the percentages 'against' gun control. We might report our findings as follows.

Report

In this sample of 100 people, we see a higher percentage of females were for gun control than males: 71.4% to 55.2%. This indicates that a person's attitude to gun control is associated with their sex.

Note: Finding a single row in the two-way frequency distribution in which percentages are clearly different is sufficient to identify a relationship between the variables.

We will now consider a two-way percentage frequency table that shows no evidence of a relationship. Consider the following table that summarises responses to the question 'Should mobile phones be banned in cinemas?' These responses were obtained from 100 students in Year 10 and Year 12 – we are interested in investigating whether there is an association between these variables.

	Year level	
Should mobile phones be banned in cinemas?	Year 10	Year 12
Yes	87.9%	86.8%
No	12.1%	13.2%
Total	100.0%	100.0%

When we look across the first row of the table, we see that the percentages in favour are very similar. In this case, we might report our findings as follows.

Report

In this sample of 100 Year 10 and Year 12 students, we see that the percentage of Year 10 and Year 12 students in support of banning mobile phones in cinemas is similar: 87.9% to 86.8%. This indicates that a person's support for banning mobile phones in cinemas is not associated with their year level.

ISBN 978-1-108-46154-2

Example 6 Interpreting a two-way percentage frequency table

Are males and females in Year 9 equally likely to indicate an intention to go to university? Data from interviews with 200 Year 9 students are summarised in the following table. Write a brief report addressing this question and quoting appropriate percentages.

	Sex		
Intend going to university	Male	Female	Total
Yes	50	54	104
No	55	41	96
Total	105	95	200

Solution

1. Determine the column percentages as follows:

   $\% \text{ of males} = \frac{50}{105} \times 100 = 47.6\%$

2. Complete the table as shown.

	Sex	
Intend going to university	*Male*	*Female*
Yes	*47.6%*	*56.8%*
No	*52.4%*	*43.2%*
Total	*100.0%*	*100.0%*

3. Select an appropriate row to compare the male and female percentages. We can see from the row indicated that a greater proportion of females than males (56.8% compared with 47.6%) were intending to go to university.

	Sex	
Intend going to university	*Male*	*Female*
Yes	*47.6%*	*56.8%*
No	*52.4%*	*43.2%*
Total	*100.0%*	*100.0%*

4. Construct a report.

Report

In this sample of 200 Year 9 students, a greater proportion of females than males (56.8% compared with 47.6%) were intending to go to university. There is an association between sex and intention to go to university.

ISBN 978-1-108-46154-2

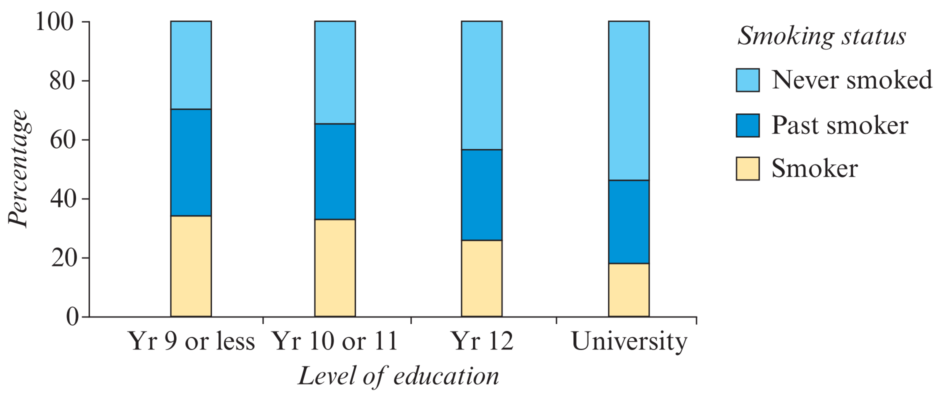
Cambridge University Press

Two-way frequency tables for categorical variables with more than two categories

The two-way percentage frequency table below displays the smoking status for 500 adults (smoker, past smoker, never smoked) by highest level of education (Year 9 or less, Year 10 or 11, Year 12, university).

	Level of education (%)			
Smoking status	Year 9 or less	Year 10 or 11	Year 12	University
Smoker	34.0	31.7	26.5	18.4
Past smoker	36.0	33.8	30.9	28.0
Never smoked	30.0	34.5	42.6	53.6
Total	100.0	100.0	100.0	100.0

The following segmented bar chart has been constructed from this table to help us with the analysis. Each column represents a column from the purple-shaded part of the table.



From the table and the segmented bar chart it is clear that there is an association between level of education and smoking status, with those with higher levels of education more likely to have never smoked. We could report this finding as follows, quoting exact percentages from the table.

Report

In this sample of 500 adults there is an association between level of education and smoking status as the percentage of adults who have never smoked tends to increase with level of education. The rate is lowest at 30.0% for those with an education level of Year 9, increases to 34.5% for those with Year 10 and 11, 42.6% for those with Year 12, and is highest at 53.6% for those who attended university.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 7 Identifying and describing associations between two categorical variables from a two-way table

A survey was conducted with 1000 males under 50 years old. As part of this survey, they were asked to rate their interest in sport as 'high', 'medium' or 'low'. Their age group was also recorded as 'under 18','19–25', '26–35' and '36–50'. The results are displayed in the table.

	Age group (%)			
Interest in sport	Under 18 years	19–25 years	26–35 years	36–50 years
High	56.5	50.2	40.7	35.0
Medium	30.1	34.4	36.8	45.8
Low	13.4	13.4	22.5	20.3
Total	100.0	100.0	100.0	100.0

a Which is the explanatory variable, *interest in sport* or *age group*?

b Is there an association between *interest in sport* and *age group*? Write a brief response quoting appropriate percentages.

Solution

a Age is a possible explanation for the level of interest in sport, but interest in sport cannot explain age.

Age group is the EV.

b If we look across all rows, we can see that the percentages are different for each age group. Select one row to compare and discuss – here we have chosen 'high'.

There is an association between the level of interest in sport and age. A high level of interest in sport is seen to decrease steadily across the age categories from 56.5% for under 18 years, 50.2% for 19–25 years, 40.7% for 26–35 years to, at its lowest, 35% for 36–50 years.

ISBN 978-1-108-46154-2

Exercise 1B

Constructing a two-way frequency table

Example 5

1 The following data were obtained when a sample of 20 Year 12 students were asked if they *intended to go to university*. The *sex* of the student was also recorded. SF

Student No.	Sex	Intend to go to university	Student No.	Sex	Intend to go to university
1	F	Yes	11	F	Yes
2	M	Yes	13	M	Yes
3	F	No	13	F	No
4	F	Yes	14	F	Yes
5	M	No	15	M	No
6	M	Yes	16	M	Yes
7	F	Yes	17	F	Yes
8	M	No	18	M	No
9	F	No	19	F	No
10	F	Yes	20	F	Yes

a Identify which variable is the explanatory variable and which is the response variable.

b Create a two-way frequency table from the data, with the values of the explanatory variable labelling the columns.

2 The following data were obtained when a sample of 30 adults were asked if they supported *reducing university fees*. They were also classified by their *age group* 17–18 years, 19–25 years, or 26 years or more. The results are given in the table below.

Age group	Reduce fees	Age group	Reduce fees	Age group	Reduce fees
17–18	Yes	26 or more	Yes	26 or more	No
19–25	Yes	17–18	Yes	19–25	Yes
26 or more	No	19–25	Yes	17–18	No
17–18	Yes	17–18	Yes	26 or more	Yes
19–25	Yes	17–18	Yes	17–18	No
26 or more	Yes	26 or more	No	26 or more	Yes
17–18	Yes	19–25	Yes	19–25	Yes
19–25	No	26 or more	Yes	17–18	Yes
26 or more	No	17–18	No	19–25	No
19–25	No	17–18	Yes	26 or more	Yes

ISBN 978-1-108-46154-2

a Identify which variable is the explanatory variable and which is the response variable.

b Create a two-way frequency table from these data, with the values of the explanatory variable labelling the columns.

c Calculate the column percentages for the table.

Using two-way tables to identify and describe associations between two categorical variables

Example 7

3 A survey was conducted with 242 university students. For this survey, data were collected on the students' *enrolment status* (full-time, part-time) and whether or not each *drinks alcohol* ('Yes' or 'No'). Their responses are summarised in the table below.

Drinks alcohol	Enrolment status (%)	
	Full-time	Part-time
Yes	80.5	81.8
No	19.5	18.2
Total	100.0	100.0

a Which variable is the explanatory variable?

b Is there an association between drinking alcohol and enrolment status? Write a brief report quoting appropriate percentages.

4 It has been suggested that females might be more satisfied with their lives overall than males. Data were collected from a sample of 360 adults and are summarised in the two-way frequency table below.

	Sex of respondent		
Are you satisfied with your life overall?	Female	Male	Total
Yes	153	155	308
No	24	28	52
Total	177	183	360

a Identify which variable is the explanatory variable and which is the response variable.

b Does the data support the suggestion that females are more satisfied with their lives than males? Write a brief report quoting appropriate percentages.

ISBN 978-1-108-46154-2

Example 6

5 The table below was constructed from data collected to see if *handedness* (left, right) was associated with *sex* (male, female).

	Sex%	
Handedness	Male	Female
Left	22	16
Right	222	147

a Which variable is the response variable?
b Convert the table to percentages by calculating column percentages.
c Is *handedness* associated with *sex*? Write a brief report using appropriate percentages.

6 A survey was conducted with 59 male and 51 female university students to determine whether, each day, they exercised, 'regularly', 'sometimes' or 'rarely'. Their responses are summarised in the table below.

	Sex%	
Exercised	Male	Female
Rarely	28.8	39.2
Sometimes	52.5	54.9
Regularly	18.6	5.9
Total	99.9	100.0

a Which is the explanatory variable?
b What percentage of females exercised sometimes?
c Is there an association between how regularly these students exercised and their sex? Write a brief report quoting appropriate percentages

7 Are those people who are satisfied with their job more likely to be satisfied with their life? Data collected from a survey of 110 adults are summarised in the two-way frequency table below.

	Satisfaction with job		
Satisfaction with life	Dissatisfied	Satisfied	Total
Dissatisfied	36	14	50
Satisfied	12	48	60
Total	48	62	110

a Identify which variable is the explanatory variable and which is the response variable.
b Does the data support the contention that people who are satisfied with their job are more likely to be satisfied with their life? Write a brief report quoting appropriate percentages.

ISBN 978-1-108-46154-2

8 It was suggested that students in Dr Evan's mathematics class would achieve higher grades than students in Dr Smith's mathematics class. Write a brief report on the association between teacher and grade, based on the data summarised in the following table.

	Class		
Exam grade	Dr Evans	Dr Smith	Total
Fail	2	3	5
Pass	11	20	31
Credit or above	5	9	14
Total	18	32	50

9 Researchers predicted that using a special pillow would be more effective in curing snoring than treatment with drugs. Discuss the association between outcome of treatment and type of treatment shown in the following table.

	Type of treatment		
Outcome of treatment	Drug	Pillow	Total
Complete cure	4	10	14
Partial cure	11	12	23
No improvement	26	10	36
Total	41	32	73

ISBN 978-1-108-46154-2

Cambridge University Press

1C Displaying bivariate data from two numerical variables – the scatterplot

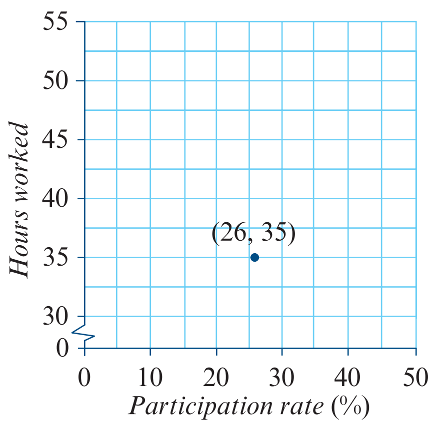
In this and the following sections of this chapter, we will start to look at techniques for investigating and understanding the relationship between *two numerical variables*.

The first step in investigating the association between two numerical variables is to construct a scatterplot. We will illustrate the process by constructing a scatterplot to display average *hours worked* (the RV) against university *participation rate* (the EV) in nine countries. The data are shown below.

Participation rate (%)	26	20	36	1	25	9	30	3	55
Hours worked	35	43	38	50	40	50	40	53	35

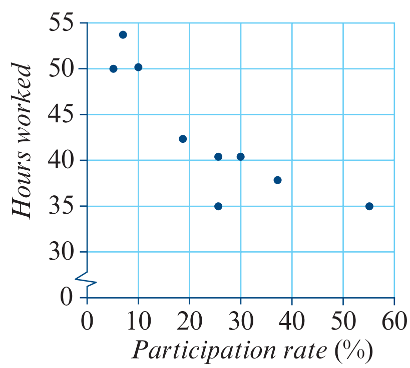
▶ Constructing a scatterplot

In a **scatterplot**, each point represents a single case; in this instance, a country. The horizontal or *x*-coordinate of the point represents the university participation rate (the EV). The vertical or *y*-coordinate represents the average hours worked (the RV).



The scatterplot opposite shows the point for a country for which the university participation rate is 26% and average hours worked is 35.

The points for each of the remaining countries are then plotted, as shown opposite.



Which axis?

When constructing a scatterplot, it is conventional to use the *vertical* or *y-axis* for the response variable (*RV*) and the *horizontal* or *x-axis* for the explanatory variable (*EV*).

Note: Following this convention will become very important when we begin fitting lines to scatterplots in the next chapter, so it is a good habit to get into from the start.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 8 Constructing a scatterplot using graph paper

The following table gives the time in minutes it takes for a headache to respond to medication together with the dose of the medication received by a group of 10 patients. Construct a scatterplot of these data.

Patient	1	2	3	4	5	6	7	8	9	10
Dose (mg)	0.5	1.2	4.0	5.3	2.6	3.7	5.1	1.7	0.3	4.0
Response time (mins)	65	35	15	10	22	16	10	18	70	20

Solution

1 Which variable will be on which axes? It is likely that the response time may be explained by the drug dosage.

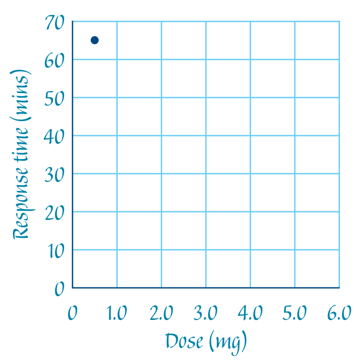
Dose is the EV — this will label the horizontal axis. Response time is the RV — this will label the vertical axis.

2 Detemine the scales for each axis.

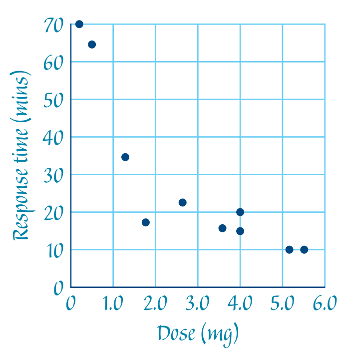
Dose ranges from 0.3 mg to 5.3 mg. A horizontal scale from 0 to 6 with intervals of 1 mg would be suitable.

Response time ranges from 10 mins to 70 mins. A vertical scale from 0 mins to 70 mins with intervals of 10 mins would be suitable.

3 Set up the axes, and then plot the data from the first patient (0.5, 65).



4 Complete the graph, adding all ten data points.



ISBN 978-1-108-46154-2

Cambridge University Press

Example 9 Constructing a scatterplot using Excel

Use the data in the following table to construct a scatterplot of the height (RV) of 20 adults against arm span, which is the distance between a person's fingertips on each hand when the arms are outspread (EV).

Subject	Arm span	Height	Subject	Arm span	Height
1	150	158	11	177	173
2	157	160	12	177	176
3	159	162	13	178	178
4	160	157	14	184	180
5	161	160	15	188	189
6	161	162	16	188	187
7	165	166	17	188	181
8	170	170	18	188	192
9	170	167	19	194	193
10	173	176	20	200	186

Solution

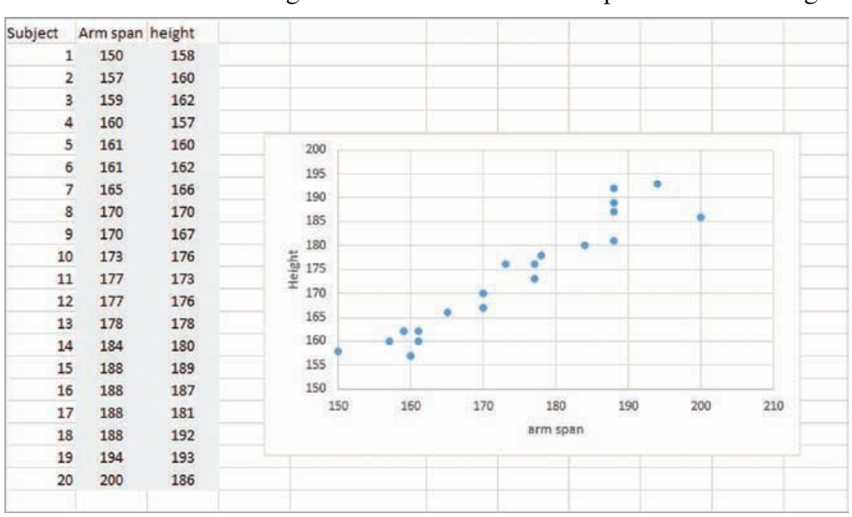
1 Enter the data into two columns B and C as shown below. Make sure that you save the data for use in later examples.

2 Select both columns (including heading) and on the **Insert** tab, in the **Charts** group, click **Scatter**.

3 Double click on each scale separately and edit to cover the range of the data.

4 Axis labels can be added using the Add Chart Elements option when editing the scale.

Spreadsheet



Subject	Arm span	height
1	150	158
2	157	160
3	159	162
4	160	157
5	161	160
6	161	162
7	165	166
8	170	170
9	170	167
10	173	176
11	177	173
12	177	176
13	178	178
14	184	180
15	188	189
16	188	187
17	188	181
18	188	192
19	194	193
20	200	186

ISBN 978-1-108-46154-2

Cambridge University Press

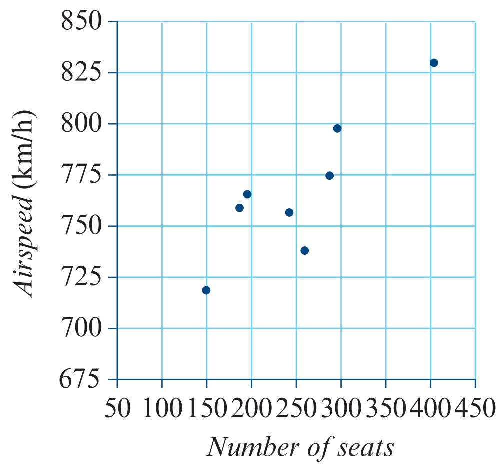
Exercise 1C

Save any scatterplots constructed in this section for use in later exercises.

The elements of a scatterplot

SF

1 The scatterplot below has been constructed to investigate the association between the airspeed (in km/h) of commercial aircraft and the number of passenger seats.



Use the scatterplot to answer the following questions.

a Which is explanatory variable?

b What type of variable is airspeed?

c How many aircraft were investigated?

d What was the airspeed of the aircraft that could seat 300 passengers?

Constructing a scatterplot using graph paper

Example 8

2 The table below shows the maximum and minimum temperatures in Toowoomba during one six-day period.

Day	1	2	3	4	5	6
Minimum temperature (°C)	17.7	19.8	23.3	22.4	22.0	25.6
Maximum temperature (°C)	29.4	34.0	34.5	35.0	36.9	36.4

Use a sheet of graph paper to complete the following.

- Construct a set of axes.
- Use a scale on the x-axis starting at 16°C and ending at 27°C, with increments of 1°C.
- Use a scale on the y-axis starting at 26°C and ending at 38°C, with increments of 1°C.
- Plot the maximum temperature against the minimum temperature for each of the six days.

ISBN 978-1-108-46154-2

Cambridge University Press

3 The price and age of several secondhand dirt bikes is listed in the table.

Age (years)	Price ($)	Age (years)	Price ($)
7	4800	11	1650
10	5700	4	6900
7	3900	3	9600
9	1950	8	6500
8	4275	4	8400
9	3300	1	11 400
9	3900	7	6600

a Determine which is the explanatory variable and which is the response variable.
b Construct a scatterplot of the price of the dirt bikes against their age.

Constructing a scatterplot using Excel

Example 9

4 The proprietor of a hairdressing salon recorded the amount spent advertising in the local paper, and the volume of business undertaken for each month for a year, with the following results.

Month	Advertising ($)	Volume of business ($)
1	3500	28 350
2	4500	30 210
3	4000	28 140
4	5000	27 330
5	2500	15 660
6	1500	9300
7	3500	24 180
8	3000	21 090
9	5500	34 500
10	6000	38 610
11	5500	31 680
12	4500	29 550

a Determine which is the explanatory variable and which is the response variable.
b Construct a scatterplot of the volume of business conducted against the amount spent on advertising.

ISBN 978-1-108-46154-2

Cambridge University Press

5 The table below shows the number of runs scored and the number of balls faced by batsmen in a one-day international cricket match. Identify the RV and the EV. Construct a scatterplot of this data.

Balls faced	29	16	19	62	13	40	16	9	28	26	6
Runs scored	27	8	21	47	3	15	13	2	15	10	2

6 The table below shows the changing diameter of a metal ball as it is heated. Identify the RV and the EV. Construct a scatterplot of this data.

Temperature (°C)	0	10	50	75	100	150
Diameter (cm)	2.00	2.02	2.11	2.14	2.21	2.28

7 The table below shows the number of people in a movie theatre at 5-minute intervals after the advertisements started. Identify the RV and the EV. Construct a scatterplot of this data.

Number in theatre	87	102	118	123	135	137
Time (minutes)	0	5	10	15	20	25

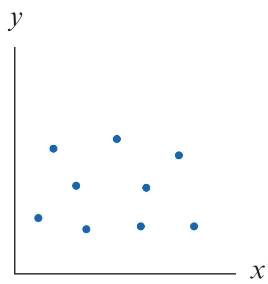
ISBN 978-1-108-46154-2

Cambridge University Press

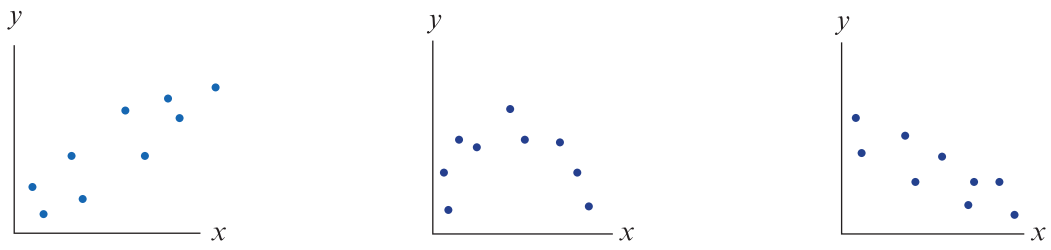
1D Interpreting a scatterplot

What features do we look for in a scatterplot that will help us to identify and describe any relationships present? First, we look to see if there is a *clear pattern* in the scatterplot.

In the example below, there is *no clear pattern* in the points. The points are *randomly scattered* across the plot, so we conclude that there is *no relationship*.



For the three examples below, there is a *clear* (but different) *pattern* in each set of points, so we conclude that there is a *relationship* in each case.



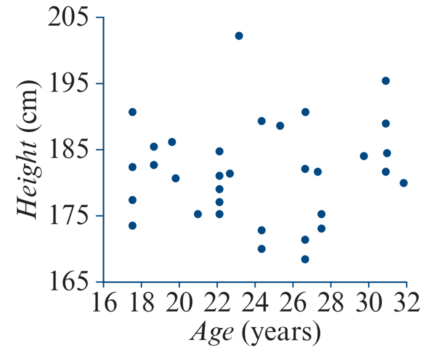
After finding a clear pattern, we need to be able to describe these relationships clearly, as they are obviously different. To do this, there are several things we look for in the pattern of points:

- direction and **outliers**
- form
- strength.

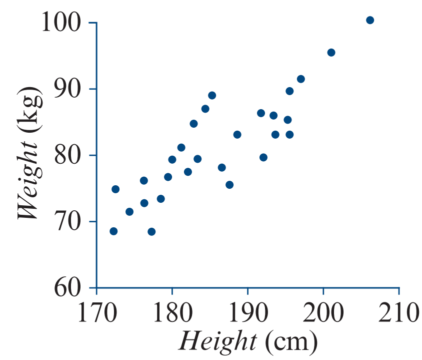
We will consider each of these attributes of the scatterplot in turn.

▶ Direction of an association and outliers

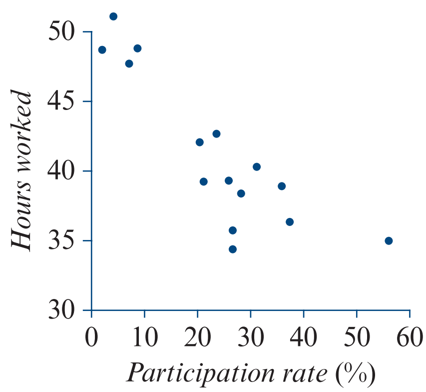
The scatterplot of height against age for a group of footballers (shown opposite) is just a *random* scatter of points. This suggests that there is *no association* between the variables *height* and *age* for this group of footballers. However, there is an *outlier*; the footballer who is 201 cm tall.



In contrast, there is a *clear pattern* in the scatterplot of weight against height for these footballers (shown opposite). The two *variables* are *associated*. Furthermore, the points *drift upwards* as you move across the plot. When this happens, we say that there is a *positive association* between the variables. Tall players tend to be heavy and vice versa. In this scatterplot, there are *no outliers*.



Likewise, the scatterplot of working hours against university participation rates for 15 countries shows a *clear pattern*. The two *variable*s are *associated*. However, in this case the points *drift downwards* as you move across the plot. When this happens, we say that there is a *negative association* between the variables. Countries with high working hours tend to have low university participation rates and vice versa. In this scatterplot, there are *no outliers*.



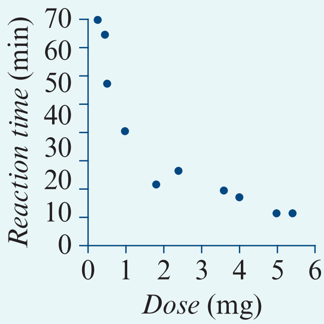
ISBN 978-1-108-46154-2

Cambridge University Press

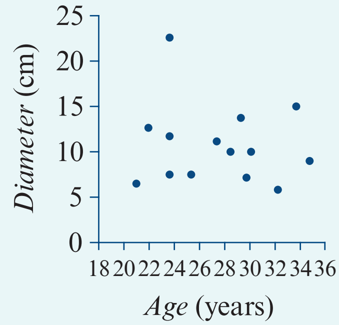
Example 10 Classifying the direction of an association

Classify each of the following scatterplots as exhibiting positive, negative or no association. Where there is an association, describe the direction of the association in terms of the variables in the scatterplot and interpret what this means in terms of the variables involved.

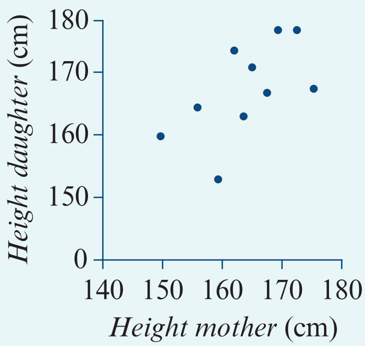
a



b



c



Solution

a There is a *clear pattern* in the scatterplot. The points in the scatterplot drift *downwards* from left to right.

The direction of the association is negative. Reaction times tend to decrease as the drug dose increases.

b There is no pattern in the scatterplot of *diameter* against *age*.

There is no association between diameter and age.

c There is a *clear pattern* in the scatterplot. The points in the scatterplot drift *upwards* from left to right.

The direction of the association is positive. Taller mothers tend to have taller daughters.

In general terms, we can interpret the *direction of an association* as follows.

Direction of an association

- Two variables have a *positive association* when the value of the response variable tends to increase as the value of the explanatory variable increases.
- Two variables have a *negative association* when the value of response variable tends to decrease as the value of the explanatory variable increases.
- Two variables have *no association* when there is no consistent change in the value of the response variable when the values of the explanatory variable increases.

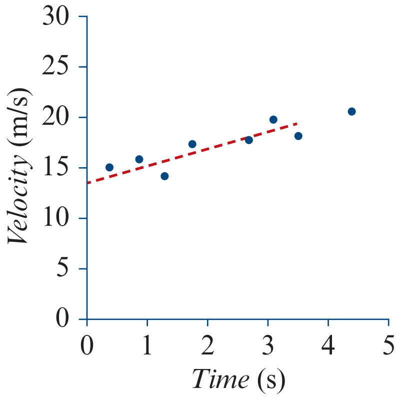
ISBN 978-1-108-46154-2

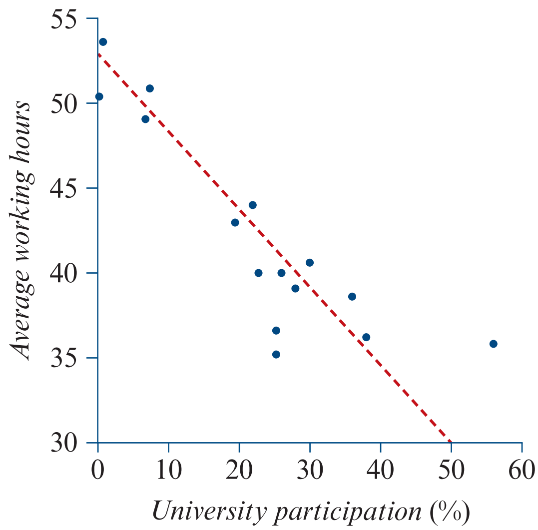
Cambridge University Press

Form of an association

What we are looking for is whether the pattern in the points has a *linear form*. If the points in a scatterplot appear to be as random fluctuations around a *straight line*, then we say that the scatterplot has a *linear form*, then we say that the variables are *linearly associated*.

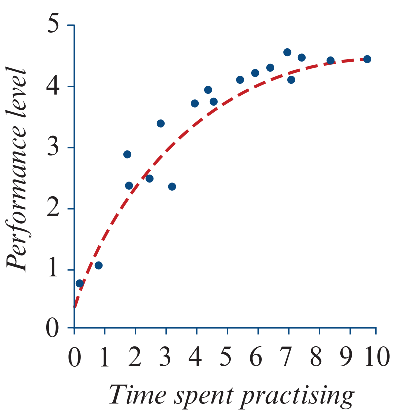
For example, both of the scatterplots below can be described as having a *linear form*; that is, the scatter in the points can be thought of as random fluctuations around a straight line. We can say that the associations between the variables involved are linear. (The dotted lines have been added to the graphs to make it easier to see the linear form.)





By contrast, consider the scatterplot below, plotting performance level against time spent practising a task. There is an association between performance level and time spent practising, but it is clearly non-linear.

The scatterplot shows that while level of performance on a task increases with practice, there comes a time when the performance level will no longer improve substantially with extra practice.



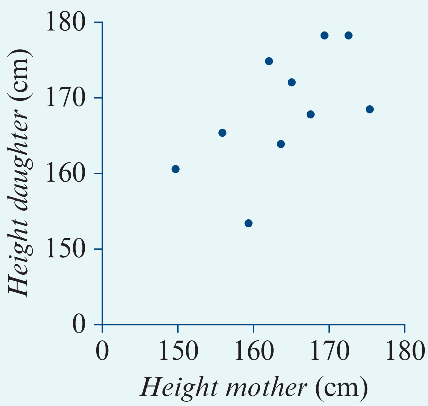
ISBN 978-1-108-46154-2

Cambridge University Press

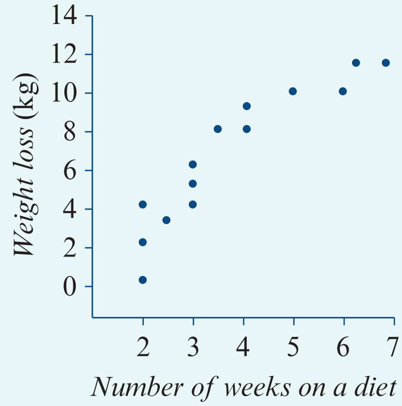
Example 11 Classifying the form of an association

Classify the *form* of the association in each scatterplot as linear or non-linear.

a



b



Solution

a There is a *clear pattern*.
The points in the scatterplot can be imagined to be scattered around a *straight line*.

The association is linear.

b There is as a *clear pattern*.
The points in the scatterplot can be imagined to be scattered around a *curved line* rather than a straight line.

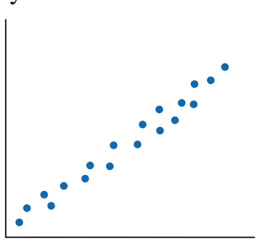
The association is non-linear.

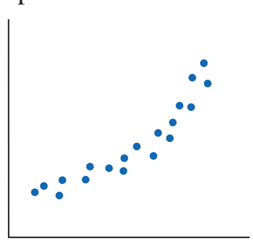
Strength of an association

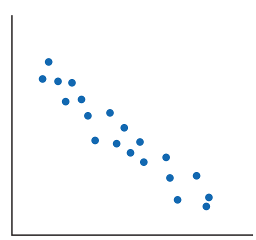
The strength of an association is the measure of how much scatter there is in the scatterplot.

Strong association

When there is a *strong association* between the variables, a pattern is clearly seen. There is only a small amount of scatter in the plot.


Strong positive association


Strong positive association

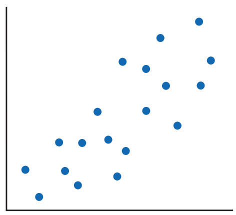

Strong negative association

ISBN 978-1-108-46154-2

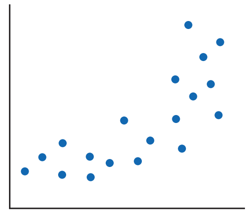
Cambridge University Press

Moderate association

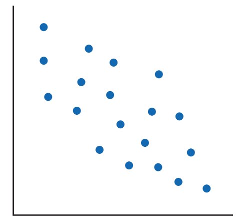
As the amount of scatter in the plot increases, the pattern becomes less clear. This indicates that the association is less strong. In the examples below, we might say that there is a *moderate association* between the variables.



Moderate positive association



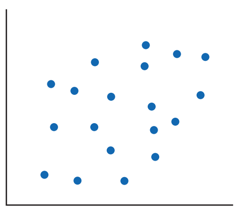
Moderate positive association



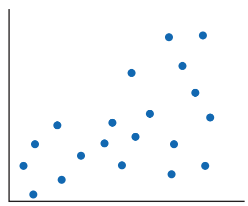
Moderate negative association

Weak association

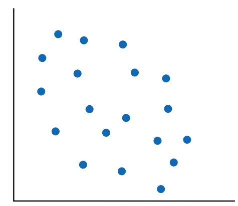
As the amount of scatter increases further, the pattern becomes even less clear. This indicates that any association between the variables is weak. The scatterplots below are examples of *weak association* between the variables.



Weak positive association



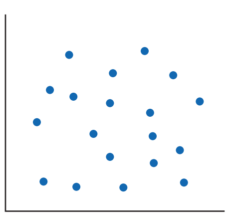
Weak positive association



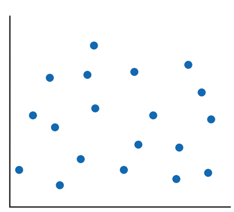
Weak negative association

No association

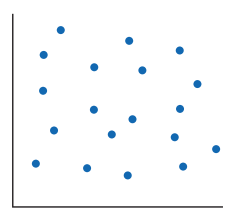
Finally, when all we have is scatter, as seen in the scatterplots below, no pattern can be seen. In this situation we say that there is *no association* between the variables.



No association



No association



No association

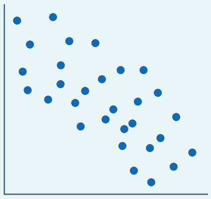
The scatterplots above should help you to get a feel for the strength of an association from the amount of scatter. At the moment, you only need be able to estimate the strength of an association as strong, moderate, weak or none, by comparing it with the standard scatterplots given above. In the next section, you will learn about a statistic, the correlation coefficient, that can be used to give a value to the strength of linear association.

ISBN 978-1-108-46154-2

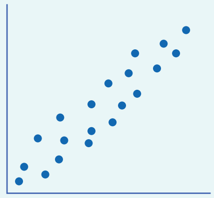
Example 12 Assessing the strength of an association

Assess the strength of the relationship in each of the following scatterplots as no association, weak association, moderate association or strong association.

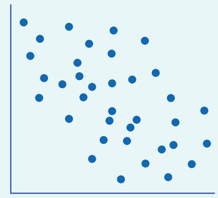
a



b



c



Solution

Compare each of these scatterplots to the previous examples to classify the associations as weak, moderate, strong or none.

a *moderate*
b *strong*
c *weak*

Exercise 1D

Assessing the direction of an association from the variables

Example 10

CF

1 For each of the following pairs of variables, indicate whether you expect an association to exist between the variables. If associated, say whether you would expect the variables to be positively or negatively associated.

- **a** *intelligence* and *height*
- **b** *level of education* and *salary level*
- **c** *salary* and *tax paid*
- **d** *frustration* and *aggression*
- **e** *population density* and *distance from the city centre*
- **f** *time using social media* and *time spent studying*

ISBN 978-1-108-46154-2

Cambridge University Press

Using a scatterplot to classify the direction, form and strength of an association

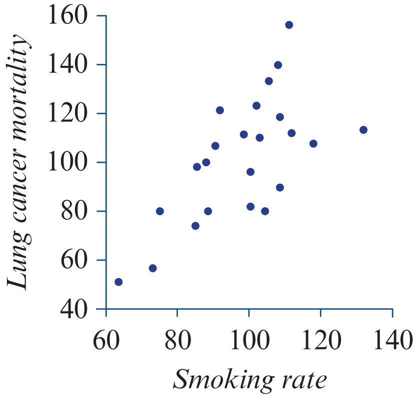
Example 10–12

2 For each of the following scatterplots, state whether the variables appear to be related. If the variables appear to be related:

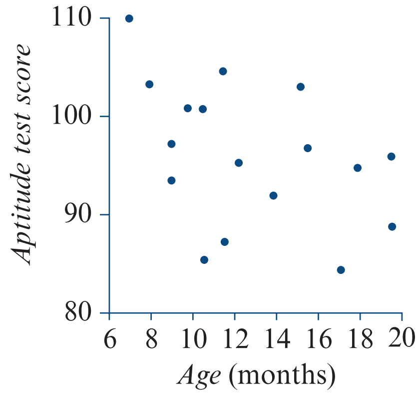
a state whether the association is positive or negative
b classify the association as linear or non-linear
c classify the strength of the association as weak, moderate, strong or no association.

SF

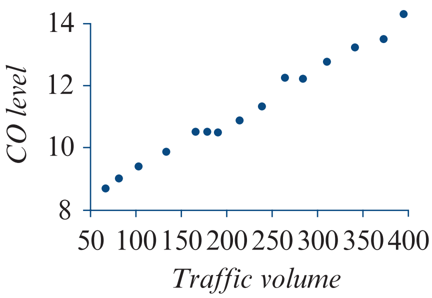
i



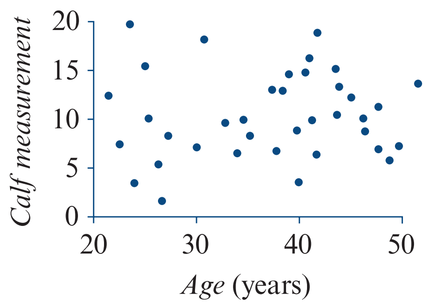
ii



iii



iv



3 The table below shows the maximum and minimum temperatures in Toowoomba during one six-day period.

Day	1	2	3	4	5	6
Minimum temperature (°C)	17.7	19.8	23.3	22.4	22.0	22.0
Maximum temperature (°C)	29.4	34.0	34.5	35.0	36.9	36.4

Use the scatterplot constructed in Exercise 1C, question 2 to complete the following.

a State whether the association is positive or negative.
b Is the association linear or non-linear?
c Classify the strength of the association as weak, moderate, strong or no association.

ISBN 978-1-108-46154-2

Cambridge University Press

4 The proprietor of a hairdressing salon recorded the amount spent advertising in the local paper, and the volume of business undertaken for each month for a year, with the following results.

Month	Advertising ($)	Volume of business ($)
1	5000	28 350
2	6000	30 210
3	4000	28 140
4	5000	27 330
5	2500	15 660
6	1500	9300
7	4000	24 180
8	3000	21 090
9	5500	34 500
10	7000	38 610
11	5500	31 680
12	4500	29 550

Use the scatterplot constructed in Exercise 1C, question 4 to complete the following.

a State whether the association is positive or negative.

b Is the association linear or non-linear?

c Classify the strength of the association as weak, moderate, strong or no association.

5 The price and age of several secondhand dirt bikes are listed in the table.

Age (years)	Price ($)	Age (years)	Price ($)
7	4800	11	1650
10	5700	4	6900
7	3900	3	9600
9	1950	8	6500
8	4275	4	8400
9	3300	1	11 400
9	3900	7	6600

Use the scatterplot constructed in Exercise 1C, question 3 to complete the following.

a State whether the association is positive or negative.

b Is the association linear or non-linear?

c Classify the strength of the association as weak, moderate, strong or no association.

ISBN 978-1-108-46154-2

6 The following table gives the age at marriage of sixteen couples.

Wife's age	Husband's age	Wife's age	Husband's age
26	29	23	27
22	26	27	27
29	43	31	36
17	21	20	20
27	33	27	26
22	24	35	25
21	22	20	25
23	28	21	19

a Plot the age at marriage of the wife against the age at marriage of the husband.
b State whether the association is positive or negative. Identify any outliers.
c Is the association linear or non-linear?
d Classify the strength of the association as weak, moderate, strong or no association.

7 A sample of 12 adult males gave the following data for height and weight.

Height (cm)	Weight (kg)
190	77
185	74
183	73
188	77
176	70
163	54
178	65
185	95
185	65
190	76
160	75
185	79

a Plot the weight against height for these adult males.
b State whether the association is positive or negative. Identify any outliers.
c Is the association linear or non-linear?
d Classify the strength of the association as weak, moderate, strong or no association.

ISBN 978-1-108-46154-2

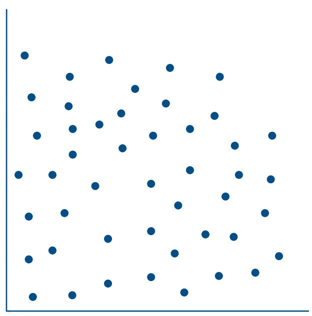
Cambridge University Press

1E A measure of strength for a linear relationship – the correlation coefficient

The strength of a linear association is an indication of how closely the points in the scatterplot fit a straight line. If the points in the scatterplot lie exactly on a straight line, we say that there is a perfect linear association. If there is no fit at all, we say there is no association. In general, we have an imperfect fit, as seen in all of the scatterplots to date.

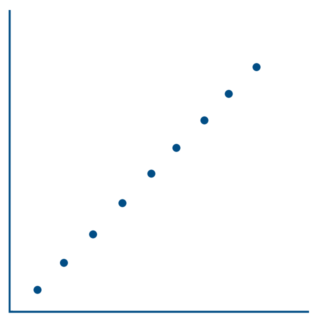
To measure the **strength of a linear relationship**, a statistician called Carl Pearson developed a **correlation coefficient**, ***r***, which has the following properties.

- If there is *no linear* association, $r = 0$.



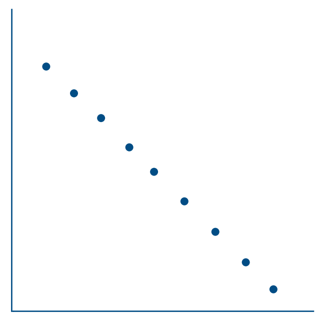
$r = 0$

- If there is a *perfect positive linear* association, $r = +1$.



$r = +1$

- If there is a *perfect negative linear* association, $r = -1$.

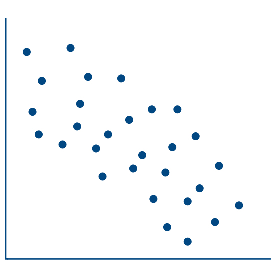


$r = -1$

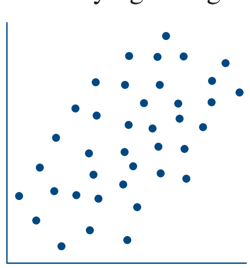
Pearson's correlation coefficient:

- measures the strength of a linear relationship, with larger values indicating stronger relationships
- has a value between –1 and +1
- is positive if the direction of the linear relationship is positive
- is negative if the direction of the linear relationship is negative.

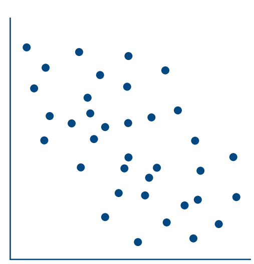
If there is a less than perfect linear association, then the correlation coefficient, r, has a value between –1 and +1, or $-1 < r < +1$. The scatterplots below show approximate values of r for linear associations of varying strengths.



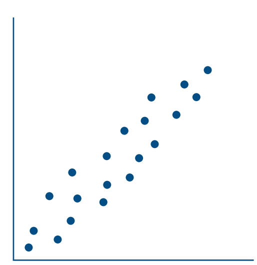
$r = -0.7$



$r = +0.5$



$r = -0.3$



$r = +0.9$

General Mathematics Units 3 & 4 ISBN 978-1-108-46154-2

Calculating the correlation coefficient

Skillsheet The correlation coefficient measures the extent to which two variables tend to vary together. It is based on the *covariance* of the two variables.

You are already familiar with concepts of the sample variance and sample standard deviation as measures of variability of a single variable. Recall, for a variable, X, the definitions of variance:

$$s_x^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

and standard deviation:

$$s_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

The sample covariance between two variables X and Y is defined as:

$$s_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n - 1}$$

You can see from the formula that if the values of X and Y tend to increase together, the value of the covariance will also increase. In fact, sample covariance may have any positive or negative value, and this is a problem with its interpretation. A large covariance can mean a strong association between variables, but it is influenced by the scale of the variables. Larger X or Y values give larger values for the covariance. To allow for this we divide the covariance by the standard deviation for each of the variables, giving us the formula for the *correlation coefficient*.

The correlation coefficient

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

where:

- r_{xy} = sample correlation between X and Y
- s_{xy} = sample covariance between X and Y
- s_x = sample standard deviation of X
- s_y = sample standard deviation of Y

The advantage of the correlation coefficient is that while covariance can take on any numerical value, the correlation is limited to values between −1 and +1, and it can be compared between any variables regardless of the unit of measurement used for each variable.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 13 Calculating the correlation coefficient from covariance and standard deviation

Suppose the variance of X is 9, the variance of Y is 16, and the covariance of X and Y is 9. Find the value of r, the correlation between X and Y.

Solution

1 Write down the formula for r.

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

2 From the information in the question, we can determine the values of the standard deviations of x and y.

$Var(X) = 9$, so $s_x = \sqrt{9} = 3$
$Var(Y) = 16$, so $s_y = \sqrt{16} = 4$

3 Write down the value of the covariance.

$s_{xy} = 9$

4 Substitute in the formula to determine the value of r, the correlation coefficient.

$$r_{xy} = \frac{9}{(3 \times 4)} = 0.75$$

Example 14 Calculating the correlation coefficient from first principles

Use the formula to calculate the correlation coefficient, r, for the following data.

x	1	3	5	4	7
y	2	5	7	2	9

$\bar{x} = 4,\ s_x = 2.236$
$\bar{y} = 5,\ s_y = 3.082$

Give the answer correct to two decimal places.

Solution

1 Write down the values of the means, standard deviations and n.

$\bar{x} = 4 \quad s_x = 2.236$
$\bar{y} = 5 \quad s_y = 3.082 \quad n = 5$

2 Set up a table like that shown opposite to calculate $\sum (x - \bar{x})(y - \bar{y})$.

x	$(x - \bar{x})$	y	$(y - \bar{y})$	$(x - \bar{x}) \times (y - \bar{y})$
1	−3	2	−3	9
3	−1	5	0	0
5	1	7	2	2
4	0	2	−3	0
7	3	9	4	12
Sum	0		0	23

$$\therefore \sum (x - \bar{x})(y - \bar{y}) = 23$$

3 Determine the value of the covariance s_{xy}.

$$s_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n - 1} = \frac{23}{4} = 5.75$$

4 Determine the value of correlation coefficient r.

$$r = \frac{5.75}{2.236 \times 3.082} = 0.83 \text{ (to two decimal places)}$$

ISBN 978-1-108-46154-2

Calculator activity 1E Using a calculator to find the correlation coefficient

Task

Use a calculator to find the correlation coefficient for the following set of bivariate data.

X	1	3	5	4	7
Y	2	5	7	2	9

Casio fx82

Change the mode to statistics with two random variables.

Press [Mode] > [2] > [2] and you should get a table like this:

	x	y
1		
2		
3		

Insert the first x data value by pressing [1] then [=]. Continue inserting the rest of the x data values similarly until they are all entered.

Use the arrow keys (on the blue circular button near the screen) to navigate to the y column and insert the y data values similarly, ensuring that each observation aligns correctly.

Leave table by pressing [AC]

Press [Shift] > [1] [STAT] > [5] [Reg] > [3] [r] > [=] to get the following result:

r
0.8342976876

TI-30XB

Press [data] to move to a data entry table.

Insert the first x data value by pressing [1] then [▼] and continue inserting the rest of the x data values. Move to the second column by pressing [▶] and then put in the y data values, ensuring that they match up with their respective x values. When finished, press [⏻ Clear] to exit the table.

Press 2nd > [data] [stat] > [2] > [enter] [enter] [enter]

Use the arrow keys on the top right circular button to navigate to the correlation coefficient r.

Press [enter] [enter] and you should get:

r
0.8342976876

 ISBN 978-1-108-46154-2 Cambridge University Press

Sharp

Enter statistics mode for (x, y) observations by pressing **Mode** **1** **1**

Insert the first (x, y) observation by pressing **1** > **STO** [(x, y)] > **2** > **M+** [DATA] and continue following the same process for the other observations.

Press **ALPHA** > **÷** [r] > **=** and you should get:

r =
0.8342976876

The CORREL function in Excel can be used to find the correlation coefficient between two variables, as shown in the following example.

Example 15 Calculating the correlation coefficient using Excel

Use the data from Example 9 to calculate the correlation coefficient between *height* and *arm span* for a group of 20 adults.

Note: The data can be seen in the spreadsheet below.

Solution

1. Enter the data into two columns B and C as shown below.
2. In an empty cell enter the formula '=CORREL(B2:B21,C2:C21)'.
3. Press Enter and the value of the correlation will be shown in that cell.

Spreadsheet

E10 =CORREL(B2:B21,C2:C21)

	A	B	C	D	E
1	Subject	Arm span	height		
2	1	150	158		
3	2	157	160		
4	3	159	162		
5	4	160	157		
6	5	161	160		
7	6	161	162		
8	7	165	166		
9	8	170	170		
10	9	170	167		0.950667531
11	10	173	176		
12	11	177	173		
13	12	177	176		
14	13	178	178		
15	14	184	180		
16	15	188	189		
17	16	188	187		
18	17	188	181		
19	18	188	192		
20	19	194	193		
21	20	200	186		

4. *The correlation coefficient is 0.95 (to two decimal places).*

ISBN 978-1-108-46154-2

Guidelines for classifying the strength of a linear relationship using the correlation coefficient

The correlation coefficient, r, can be used to classify the strength of a linear association as follows.

strong positive linear association	$0.75 \leqslant r \leqslant 1$
moderate positive linear association	$0.50 \leqslant r < 0.75$
weak positive linear association	$0.25 \leqslant r < 0.50$
no linear association	$-0.25 < r < 0.25$
weak negative linear association	$-0.50 < r \leqslant -0.25$
moderate negative linear association	$-0.75 < r \leqslant -0.50$
strong negative linear association	$-1 \leqslant r \leqslant -0.75$

Example 16 Classifying the strength of a linear relationship using the correlation coefficient

Classify the strength of each of the following values of the correlation coefficient according to the previous table:

a $r = -0.08$ **b** $r = 0.80$ **c** $r = 0.56$ **d** $r = -0.3$

Solution

a $r = -0.08$ lies in the interval $-0.25 < r < 0.25$ — *No linear association*

b $r = 0.80$ lies in the interval $0.75 \leqslant r \leqslant 1$ — *Strong positive linear association*

c $r = 0.56$ lies in the interval $0.50 \leqslant r < 0.75$ — *Moderate positive linear association*

d $r = -0.3$ lies in the interval $-0.50 < r \leqslant -0.25$ — *Weak negative linear association*

Warning

If you use the value of the *correlation coefficient* as a measure of the strength of an association, you are implicitly assuming that:

1. the variables are *numeric*
2. the association is *linear*
3. there are *no outliers* in the data.

The *correlation coefficient* can give a *misleading* indication of the strength of the linear association if there are outliers present.

ISBN 978-1-108-46154-2

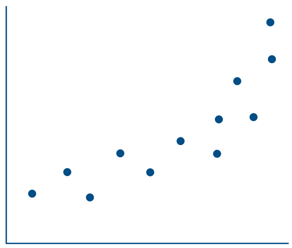
Cambridge University Press

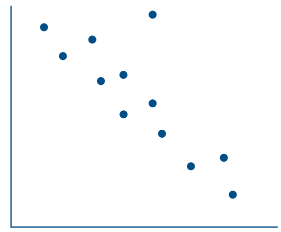
Exercise 1E

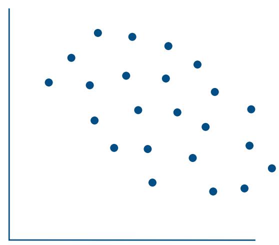
Basic ideas

CF

1 The scatterplots of three sets of related variables are shown.


Scatterplot A


Scatterplot B


Scatterplot C

a For each scatterplot, describe the association in terms of strength, direction, form and outliers (if any).

b For which of these scatterplots would it be inappropriate to use the correlation coefficient, r, to give a measure of the strength of the association between the variables? Give reasons.

Example 13

SF

2 Suppose the variance of X is 5, the variance of Y is 6, and the covariance of X and Y is 3. Find the value of r, the correlation between X and Y.

3 If $s_x = 3.55$, $s_y = 1.79$ and $s_{xy} = 4.75$, find the correlation coefficient r.

4 If $r = 0.75$, $s_x = 0.85$ and $s_y = 0.45$, what is the covariance of X and Y?

Calculating r using the formula

Example 14

5 Use the formula to calculate the correlation coefficient, r, correct to two decimal places.

x	2	3	6	3	6
y	1	6	5	4	9

$\bar{x} = 4,\ s_x = 1.871$

$\bar{y} = 5,\ s_y = 2.915$

6 Use the formula to calculate the correlation coefficient, r, correct to three decimal places.

x	1	2	4	6	7
y	0	2	3	7	10

$s_x = 2.5495,\ s_y = 4.0373$

ISBN 978-1-108-46154-2

Cambridge University Press

7 **a** The table below shows the *maximum* and *minimum* temperatures during a heat-wave. The *maximum* and *minimum* temperature each day are linearly associated. Use your calculator to show that $r = 0.818$, correct to three decimal places.

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday
Maximum (°C)	29.4	34.0	34.5	35.0	36.9	36.4
Minimum (°C)	17.7	19.8	23.3	22.4	22.0	22.0

b This table shows the number of runs scored and balls faced by batsmen in a cricket match. *Runs scored* and *balls faced* are linearly associated. Use your calculator to show that $r = 0.8782$, correct to four decimal places.

Batsman	1	2	3	4	5	6	7	8	9	10	11
Runs scored	27	8	21	47	3	15	13	2	15	10	2
Balls faced	29	16	19	62	13	40	16	9	28	26	6

c This table shows the hours worked and university participation rate (%) in six countries. *Hours worked* and university *participation rate* are linearly associated. Use your calculator to show that $r = -0.6727$, correct to four decimal places.

Country	Australia	Britain	Canada	France	Sweden	US
Hours worked	35.0	43.0	38.2	39.8	35.6	34.8
Participation rate (%)	26	20	36	25	37	55

Example 16

8 The price and age of several secondhand dirt bikes is listed in the table.

a Calculate the value of the correlation coefficient, giving your answer correct to three decimal places.

b Classify the strength of the association according the table on page 41.

Age (years)	Price ($)
7	4800
10	5700
7	3900
9	1950
8	4275
9	3300
9	3900
11	1650
4	6900
3	9600
8	6500
4	8400
1	11 400
7	6600

 ISBN 978-1-108-46154-2

9 The following table gives the age at marriage of sixteen couples.

Wife's age	Husband's age
26	29
22	26
29	43
17	21
27	33
22	24
21	22
23	28
23	27
27	27
31	36
20	20
27	26
35	25
20	25
21	19

a Calculate the value of the correlation coefficient, giving your answer correct to three decimal places.

b Classify the strength of the association according the table on page 41.

10 A sample of 12 adult males gave the following data for height and weight.

Height (cm)	Weight (kg)
190	77
185	74
183	73
188	77
176	70
163	54
178	65
185	95
185	65
190	76
160	75
185	79

a Calculate the value of the correlation coefficient, giving your answer correct to three decimal places.

b Classify the strength of the association according the table on page 41.

ISBN 978-1-108-46154-2

Cambridge University Press

1F The coefficient of determination

Introduction

If two variables are associated, it is possible to estimate the value of one variable from that of the other. For example, people's weights and heights are associated. Thus, given a person's height, we can roughly predict their weight. The degree to which we can make such predictions depends on the value of r. If there is a perfect linear association ($r = 1$) between two variables, we can make an exact prediction.

For example, when you buy cheese by the gram there is an exact association between the weight of the cheese and the amount you pay ($r = 1$). At the other end of the scale, there is no association between an adult's height and their IQ ($r \approx 0$). So knowing an adult's height will not enable you to predict their IQ any better than guessing.

The coefficient of determination

The *degree* to which one variable can be predicted from another linearly related variable is given by a statistic called the **coefficient of determination**.

The coefficient of determination is denoted as R^2 and *calculated* by squaring the correlation coefficient:

$$\text{coefficient of determination, } R^2 = r^2$$

Calculating the coefficient of determination

If the correlation between weight and height is $r = 0.8$, then:

$$R^2 = \text{coefficient of determination} = 0.8^2 = 0.64 \text{ or } 0.64 \times 100 = 64\%$$

Note: We have converted the coefficient of determination from a decimal into a percentage (64%) as this is the most useful form when we come to interpreting the coefficient of determination.

Interpreting the coefficient of determination

We now know how to calculate the coefficient of determination, but what does it tell us?

Interpreting the coefficient of determination

The coefficient of determination (as a percentage) tells us the *variation in the response variable* that is *explained* by the *variation in the explanatory variable*.

But what does this mean in practical terms?

Let's take the relationship between weight and height that we just considered. Here the coefficient of determination is 0.64 (or 64%).

The coefficient of determination tells us that 64% of the variation in people's weights is explained by the variation in their heights.

ISBN 978-1-108-46154-2

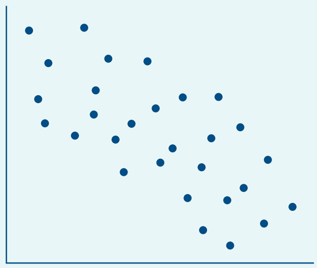
What do we mean by 'explained'?

If we take a group of people, their weights and heights will vary. One explanation is that taller people tend to be heavier and shorter people tend to be lighter. The coefficient of determination tells us that 64% of the variation in people's weights can be explained by the variation in their heights. The rest of the variation (36%) in their weights will be explained by other factors, such as sex, lifestyle, build.

Example 17 Calculating the correlation coefficient from the coefficient of determination

For the relationship described by this scatterplot, the coefficient of determination = 0.5210.

Determine the value of the correlation coefficient, r.



Solution

1. The coefficient of determination = R^2. Use the value of the coefficient of determination to set up an equation for R^2. Then solve.

   $R^2 = 0.5210$

   $\therefore\ r = \pm\sqrt{0.5210} = \pm 0.7218$

2. There are two solutions, one positive and the other negative. Use the scatterplot to decide which applies.

   Scatterplot indicates a negative association.

3. Write down your answer.

   $\therefore\ r = -0.7218$

Example 18 Calculating and interpreting the coefficient of determination

Carbon monoxide (CO) levels in the air and traffic volume are linearly related, with:

$$r_{\text{CO level, traffic volume}} = +0.985$$

Determine the value of the coefficient of determination, write it in percentage terms and interpret. In the relationship, *traffic volume* is the explanatory variable.

Solution

The coefficient of determination is:

$R^2 = (0.985)^2 = 0.970\ldots$ or $0.970 \times 100 = 97.0\%$

Therefore, 97% of the variation in carbon monoxide levels in the air can be explained by the variation in traffic volume.

ISBN 978-1-108-46154-2

Clearly, traffic volume is a very good predictor of carbon monoxide levels in the air. Thus, knowing the traffic volume enables us to predict carbon monoxide levels with a high degree of accuracy. This contrasts with the next example which concerns predicting mathematical ability from verbal ability.

Example 19 Calculating and interpreting the coefficient of determination

Scores on tests of verbal and mathematical ability are linearly related with:

$r_{\text{mathematical, verbal}} = +0.275$

Determine the value of the coefficient of determination, write it in percentage terms and interpret. In this relationship, *verbal ability* is the explanatory variable.

Solution

The coefficient of determination is:

$R^2 = (0.275)^2 = 0.0756\ldots$ or $0.076 \times 100 = 7.6\%$

Therefore, only 7.6% of the variation observed in scores on the mathematical ability test can be explained by the variation in scores obtained on the verbal ability test.

Clearly, scores on the verbal ability test are not good predictors of the scores on the mathematical ability test; 92.4% of the variation in mathematical ability is explained by other factors.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 1F

Calculating the coefficient of determination from r

SF

1 For each of the following values of r, calculate the value of the coefficient of determination and convert to a percentage (correct to one decimal place).

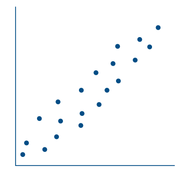
a $r = 0.675$ **b** $r = 0.345$ **c** $r = -0.567$
d $r = -0.673$ **e** $r = 0.124$

Calculating r from the coefficient of determination given a scatterplot

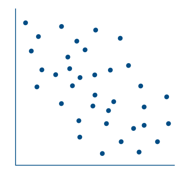
Note: The scatterplots have been included in Question 2 to help you decide the sign of r.

Example 17

2 **a** For the relationship described by the scatterplot shown, the coefficient of determination, $R^2 = 0.8215$. Determine the value of the correlation coefficient, r (correct to three decimal places).



b For the relationship described by the scatterplot shown, the coefficient of determination $R^2 = 0.1243$. Determine the value of the correlation coefficient, r (correct to three decimal places).



Calculating and interpreting the coefficient of determination

CF

Example 18, 19

3 For each of the following, determine the value of the coefficient of determination, write it in percentage terms, and interpret.

a Scores on hearing tests and age (EV) are linearly related, with $r_{\text{hearing, age}} = -0.611$.

b Mortality rates and smoking rates (EV) are linearly related, with $r_{\text{mortality, smoking}} = 0.716$.

c Life expectancy and birth rates (EV) are linearly related, with $r_{\text{life expectancy, birth rate}} = -0.807$.

d Daily maximum (RV) and minimum temperatures are linearly related, with $r_{\text{max, min}} = 0.818$

e Runs scored (RV) and balls faced by a batsman are linearly related, with $r_{\text{runs, balls}} = 0.8782$

4 A study was conducted where a group of 100 adults were asked to record their weekly income, their weekly expenditure on food, and their weekly expenditure on leisure. The researcher determined the following correlation coefficients:

$r_{\text{income, food}} = 0.6 \quad r_{\text{income, leisure}} = 0.5$

a Calculate and interpret the coefficient of determination relating income and expenditure on food.

b Calculate and interpret the coefficient of determination relating income and expenditure on leisure.

c Write a sentence comparing the explanatory power of income in understanding expenditure on food and expenditure on leisure.

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Key ideas and chapter summary

AS

Bivariate data	**Bivariate data** are data in which each observation involves recording information about two variables for the same person or thing. An example would be the heights and weights of the children in a preschool.
Categorical and numerical variables	**Categorical variables** generate data values which are labels, **numerical variables** generate values which are numbers.
Explanatory and response variables	The **explanatory variable** (EV) may explain the associated changes in the **response variable** (RV).
Two-way frequency tables	A **two-way frequency table** summarises **bivariate data** obtained from two categorical variables. In a two-way frequency table, the columns are labelled with the values of the EV, and the rows are labelled with the values of the RV.
Identifying associations between two variables	Associations between two categorical variables are identified by comparing percentages in a two-way **percentage frequency** table. Associations between two numerical variables are identified using a scatterplot.
Scatterplots	A **scatterplot** is used to identify the relationship between two numerical variables. In a scatterplot, the EV is plotted on the horizontal axis, and the RV is plotted on the vertical axis.
Correlation coefficient r	The **correlation coefficient**, r, gives a measure of the **strength of a linear relationship** between two numerical variables.
Coefficient of determination R^2	The **coefficient of determination**, R^2, gives the percentage of variation in the RV that can be explained by the variation in the EV.

Skills check

Having completed this chapter you should be able to:

- interpret the information contained in a two-way frequency table
- identify, where appropriate, the response and explanatory variable in a pair of related variables
- identify associations in tabulated data by forming and comparing appropriate percentages

ISBN 978-1-108-46154-2

- construct a scatterplot
- use a scatterplot to describe an association between two numerical variables in terms of:
  - direction (positive or negative association) and possible outliers
  - form (linear or non-linear)
  - strength (weak, moderate, strong)
- calculate and interpret the correlation coefficient, r
- know the three key assumptions made when using Pearson's correlation coefficient r as a measure of the strength of the linear association between two variables; that is:
  - the variables are numerical
  - the association is linear
  - there are no clear outliers
- calculate and interpret the coefficient of determination.

Multiple-choice questions

Use the information in the following frequency table to answer Questions 1 to 4.

	Sex	
Plays sport	Male	Female
Yes	68	79
No	34	
Total	102	175

1 The variables *plays sport* and *sex* are:

A both categorical variables
B a categorical and a numerical variable, respectively
C a numerical and a categorical variable, respectively
D both numerical variables
E neither numerical nor categorical variables

2 The number of females who do not play sport is:

A 21 **B** 45 **C** 79 **D** 96 **E** 175

3 The percentage of males who do not play sport is:

A 19.4% **B** 33.3% **C** 34.0% **D** 66.7% **E** 68.0%

4 The variables *plays sport* and *sex* appear to be associated because:

A more females play sport than males
B fewer males play sport than females
C a higher percentage of females play sport compared to males
D a higher percentage of males play sport compared to females
E both males and females play a lot of sport

ISBN 978-1-108-46154-2

Use the information in the following frequency table to answer Questions 5 and 6.
The results of a survey conducted with 578 secondary students are summarised in the following table.

	How important is it to obey rules?			
How exciting is your life?	Very important	Important	Not important	Total
Very exciting	50	84	44	178
Routine	103	169	56	328
Dull	25	42	5	72
Total	178	295	105	578

5 The percentage of students who think it is important to obey rules and who find life dull is closest to:
A 14% **B** 58% **C** 7% **D** 51% **E** 12%

6 From this table, we can conclude there is a relationship between how exciting respondents find their lives and how important they think it is to obey rules because:
A many more students find life to be very exciting than routine or dull
B the percentage of students who find life very exciting is highest for those who think it is not important to obey rules (42% compared to 28% for the other two categories)
C the percentage of students who think it is not important to obey rules is highest for those who think life is very exciting (42%), followed by 53% who think life is routine, and 5% who say life is dull
D the percentage of students who find life routine is 57%, which is higher than those who find it exciting (31%) or dull 12%
E there is no relationship evident in these data between these variables

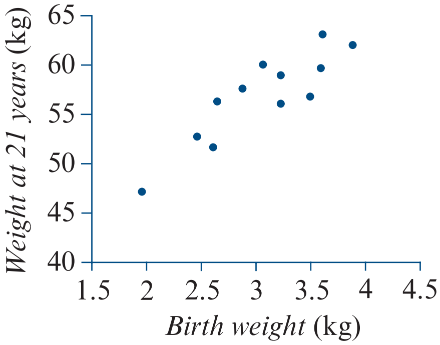
7 The association between weight at age 21 (in kg) and weight at birth (in kg) is to be investigated. The variables *weight at age 21* and *weight at birth* are:
A both categorical variables
B a categorical and a numerical variable, respectively
C a numerical and a categorical variable, respectively
D both numerical variables
E neither numerical nor categorical variables

ISBN 978-1-108-46154-2

Review

8 The scatterplot shows the weights of 12 women at birth and at the age of 21. The association is best described as:

A weak, positive and linear
B weak, negative and linear
C moderate, positive and non-linear
D strong, positive and non-linear
E strong, positive and linear



9 The variables *response time* to a drug and *drug dosage* are linearly associated, with $r = -0.9$. From this information, we can conclude that:

A response times are –0.9 times the drug dosage
B response times decrease with decreased drug dosage
C response times decrease with increased drug dosage
D response times increase with increased drug dosage
E response times are 81% of the drug dosage

10 The birth weight and weight at age 21 of eight women are given in the table below.

Birth weight (kg)	1.9	2.4	2.6	2.7	2.9	3.2	3.4	3.6
Weight at 21 (kg)	47.6	53.1	52.2	56.2	57.6	59.9	55.3	56.7

The value of the correlation coefficient is closest to:

A 0.536 **B** 0.6182 **C** 0.7863 **D** 0.8232 **E** 0.8954

11 If $s_x = 1.41$, $s_y = 2.56$ and $s_{xy} = -1.30$, then the correlation coefficient r is:

A −0.13 **B** 0.13 **C** −0.36 **D** 0.36 **E** 0.71

12 The value of correlation coefficient is $r = -0.7685$. The value of the corresponding coefficient of determination is closest to:

A −0.77 **B** −0.59 **C** 0.23 **D** 0.59 **E** 0.77

13 The correlation coefficient between heart weight and body weight in a group of mice is $r = 0.765$.

Using body weight as the explanatory variable, we can conclude that:

A 58.5% of the variation in heart weight is explained by the variation in body weights.
B 76.5% of the variation in heart weight is explained by the variation in body weights.
C heart weight is 58.5% of body weight
D heart weight is 76.5% of body weight
E 58.5% of the mice had heavy hearts

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Short-answer questions

SF

1 For each of the following questions, determine if they involve investigating associations between:

- One numerical and one categorical variable or
- Two categorical variables or
- Two numerical variables

a Are females more likely to believe in Astrology than males?
b Do males tend to be heavier than females (weight measured in kg)?
c Are students who work more hours per week in paid employment less likely to get high marks in their mathematics exam?

2 For each of the pairs of variables in Question 1 identify which is the explanatory variable and which is the response variable.

Use the information in the following frequency table to answer questions 3–5

Are males and females in Year 12 equally likely to indicate an intention to go to university? Data from interviews with 300 Year 12 students are summarised in the following table.

Intend going to university	Male	Female	Total
Yes	94	123	208
No	46	37	92
Total	140	160	300

3 Why is this two-way frequency table an appropriate representation of these data?

4 Construct a percentaged frequency table.

CF

5 Use the percentaged two-way frequency table to describe the relationship between sex and intention to go to university.

SF

6 A retailer recorded the number of ice-creams sold and the day's maximum temperature over 8 consecutive Saturdays one summer. Use the data in the table to construct a scatterplot of these data, and describe the features of the scatterplot.

Temperature (°C)	22	25	36	34	21	28	41	31
Number of ice-creams sold	145	155	200	198	150	179	230	180

7 **a** Use the data in Question 6 to determine the value of the correlation coefficient between Temperature and Number of ice-creams sold.
b Classify the strength of this relationship

ISBN 978-1-108-46154-2

Cambridge University Press

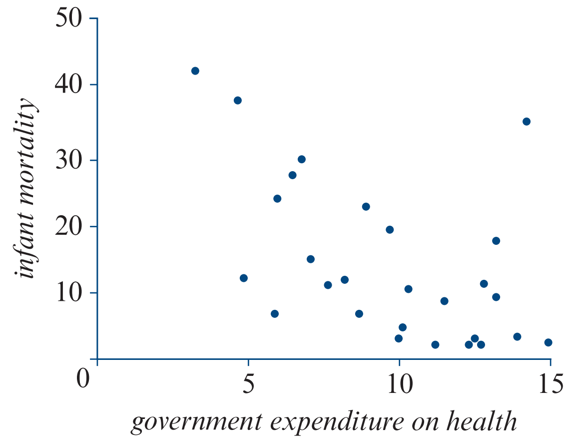
Review

8 The scatterplot on the right shows the Government expenditure on health against the infant mortality for a group of countries

CF

a Describe the relationship in this scatterplot.

b Is it appropriate to calculate the correlation coefficient for these data?



9 If the variance of X is 16, the variance of Y is 25, and the covariance of X and Y is 20 find the value of r, the correlation between X and Y.

10 If $r = 0.67$, $s_x = 3.8$, and $s_y = 1.7$, what is the covariance between X and Y?

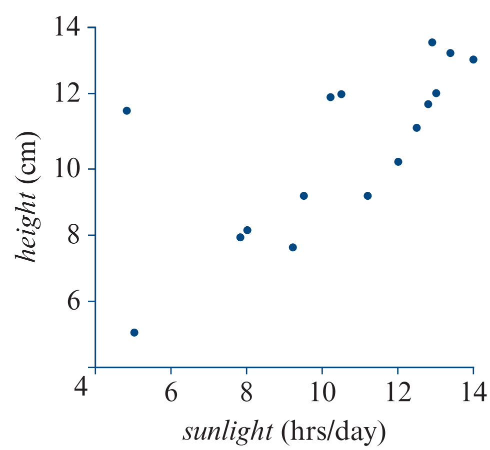
11 Suppose that the correlation between age and scores on a hearing test are linearly related, and that $r = -0.77$.

a Determine the value of the coefficient of determination, R^2.

b Interpret R^2 in terms of the variables age and score on the hearing test.

12 Miller conducted study to investigate the height of his seedlings and the average number of hours of daily sunshine each plant received over a 14-day period. He planted his seedlings in 10 different locations in his garden, and his results are shown in the following scatterplot.

CU



a Describe the relationship in the scatterplot.

b Miller checks his data and realises he has made a mistake is recording the data. The actual height of the plant which received an average of 4.8 hrs/day was not the 11.5 cm he had used in his analysis. He decides to remove this data value. How would the value of the correlation coefficient calculated with the outlier excluded compare to that calculated with the outlier included?

ISBN 978-1-108-46154-2

Cambridge University Press

Extended-response questions

1 One thousand drivers who had an accident during the past year were classified according to their age and the number of accidents they had.

Number of accidents	Age < 30	Age ⩾ 30
At most one accident	130	170
More than one accident	470	230
Total	600	400

a What are the variables shown in the table? Are they categorical or numerical?
b Determine the response and explanatory variables.
c How many drivers under the age of 30 had more than one accident?
d Convert the table values to percentages by calculating the column percentages.
e Use these percentages to comment on the statement: 'Of drivers who had an accident in the past year, younger drivers (age < 30) are more likely than older drivers (age ⩾ 30) to have had more than one accident.'

2 The data below are the hourly pay rates (in dollars per hour) of 10 production-line workers along with their years of experience on initial appointment.

Rate ($/h)	15.90	15.70	16.10	16.00	16.79	16.45	17.00	17.65	18.10	18.75
Experience (years)	1.25	1.50	2.00	2.00	2.75	4.00	5.00	6.00	8.00	12.00

a Construct a scatterplot of the data, with the variable *rate* plotted on the vertical axis and the variable *experience* on the horizontal axis. Why has the vertical axis been used for the variable *rate*?
b Comment on direction, outliers, form and strength of any association revealed.
c Determine the value of the correlation coefficient correct (r) to three decimal places.
d Determine the value of the coefficient of determination (R^2) and interpret.

3 As part of the General Social Survey conducted in the US, respondents were asked to say whether they found life *exciting, pretty routine* or *dull*. Their marital status was also recorded as married, widowed, divorced, separated or never married. The results are organised into a table as shown.

Attitude to life	Marital status (%)				
	Married	Widowed	Divorced	Separated	Never
Exciting	47.6	33.8	46.7	45.9	52.3
Pretty routine	48.7	54.3	47.6	44.6	44.4
Dull	3.7	11.9	5.7	9.5	3.3
Total	100.0	100.0	100.0	100.0	100.0

ISBN 978-1-108-46154-2

Cambridge University Press

Review

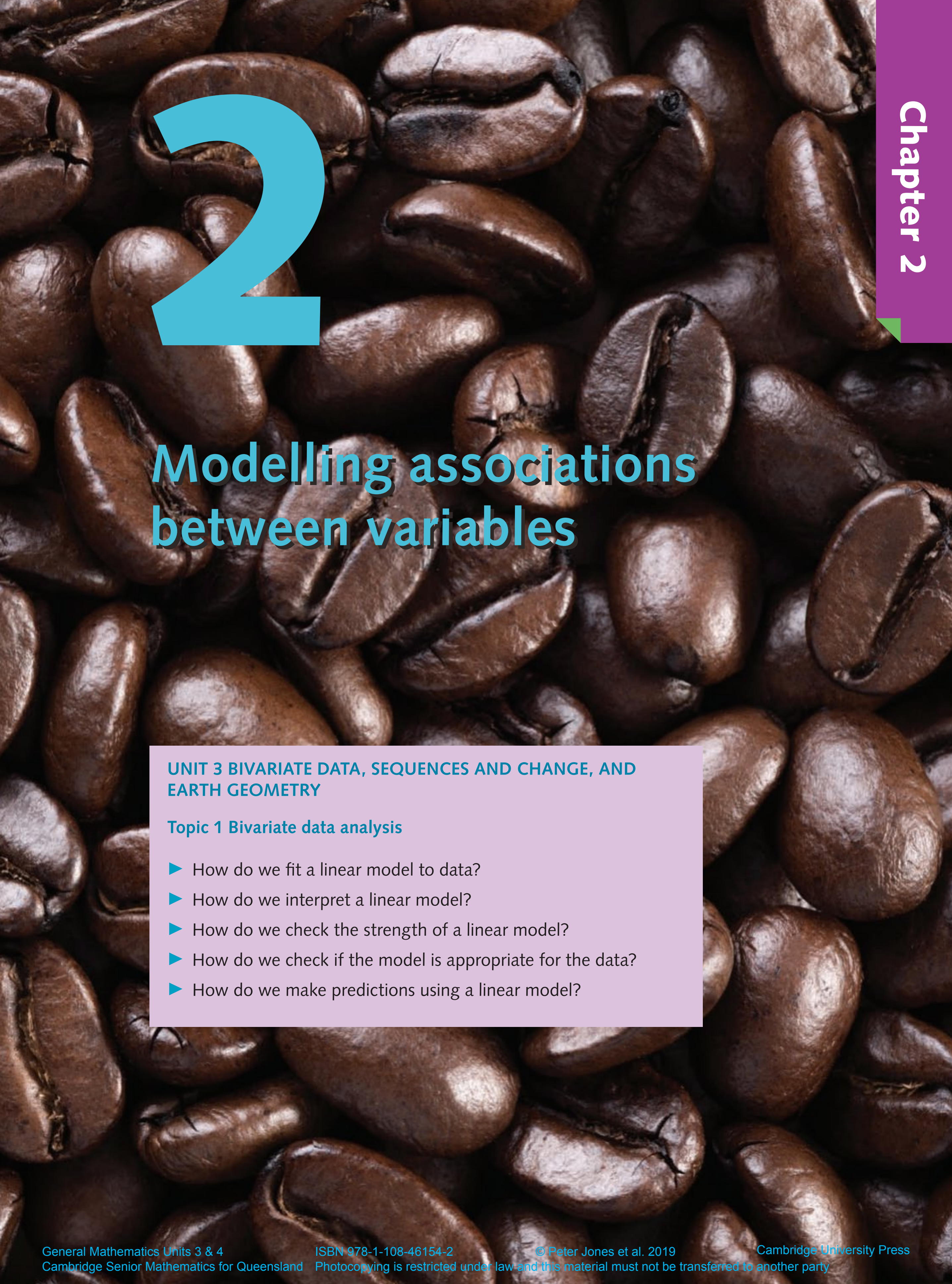
a What percentage of widowed people found life ‘dull’?
b What percentage of people who were never married found life ‘exciting’?
c What is the likely explanatory variable in this investigation?
d Does the information you have been given support the contention that a person’s *attitude to life* is related to their marital status? Justify your argument by quoting appropriate percentages.

4 The manager of a large manufacturing plant was worried about the quality of the products coming from the plant. She noticed that some workers made far more mistakes than others in assembling the gizmos and was concerned that they were trying to work too quickly. The manager collected information from a random sample of 15 workers, recording the time taken to assemble a gizmo and the number of mistakes made.

Time taken (mins)	Number of mistakes
45	7
45	2
44	6
37	7
38	6
55	3
56	4
42	6
57	1
47	3
41	3
52	5
51	7
36	6
61	0

a Which is the explanatory variable, and which is the response variable?
b Construct a scatterplot of these data. Use the scatterplot to describe the relationship between the time taken and the number of mistakes made.
c Is it appropriate to calculate Pearson’s r for these data? Why?
d Determine the value of Pearson’s r for these data.
e Give the coefficient of determination (R^2) and interpret.

ISBN 978-1-108-46154-2



Chapter 2

2

Modelling associations between variables

UNIT 3 BIVARIATE DATA, SEQUENCES AND CHANGE, AND EARTH GEOMETRY

Topic 1 Bivariate data analysis

- How do we fit a linear model to data?
- How do we interpret a linear model?
- How do we check the strength of a linear model?
- How do we check if the model is appropriate for the data?
- How do we make predictions using a linear model?

ISBN 978-1-108-46154-2

Cambridge University Press

Introduction

In Chapter 1, we learned how to identify and describe the relationships between two categorical variables, using two-way percentage frequency tables, and between two numerical variables, using scatterplots and the correlation coefficient. In this chapter, we will extend these ideas further, beginning with fitting a line to bivariate numerical data.

2A Fitting a linear model to numerical data

Once we identify a linear association between two numerical variables, we can fit a linear model to the data and find its equation. This equation gives us a better understanding of the nature of the relationship between the two variables, and we can also use the model to make predictions based on this understanding of the relationship.

The process of modelling an association with a straight line is known as **linear regression** and the resulting line is often called the *regression line*.

The equation of a line relating two variables x and y is of the form

$$y = a + bx$$

where a and b are constants. When the equation is written in this form:

- a represents the coordinate of the point where the line crosses the y-axis (the y-intercept)
- b represents the slope of the line.

In order to summarise any particular (x, y) data set, numerical values for a and b are needed that will ensure the line passes close to the data. There are several ways in which the values of a and b can be found.

The easiest way to fit a line to bivariate data is to construct a scatterplot and draw the line 'by eye'. We do this by placing a ruler on the scatterplot so that it seems to follow the general trend of the data. You can then use the ruler to draw a straight line. Unfortunately, unless the points are very tightly clustered around a straight line, the results you get by using this method will differ a lot from person to person.

The most common approach to fitting a straight line to data is to use the **least squares method**. This method assumes that the variables are linearly related, and works best when there are no clear outliers in the data.

 ISBN 978-1-108-46154-2 Cambridge University Press

Some terminology

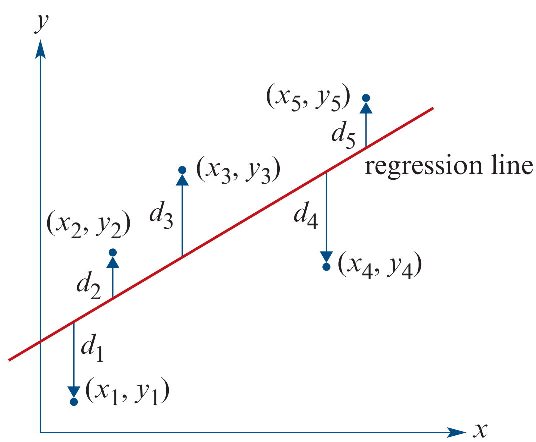
To explain the least squares method, we need to define several terms.

The scatterplot shows five data points, (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4) and (x_5, y_5).

A regression line (not necessarily the least squares line) has also been drawn on the scatterplot.

The vertical distances d_1, d_2, d_3, d_4 and d_5 of each of the data points from the regression line are also shown.

These vertical distances, d, are known as **residuals**.



The least squares regression line

The least squares regression line is the line where the sum of the squares of the residuals is as small as possible; that is, it minimises:

$$\text{the sum of the squares of the residuals } = d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2$$

Why do we minimise the sum of the *squares* of the residuals and not the sum of the residuals? This is because the sum of the residuals for the least squares regression line is always zero. Some residuals are positive and some negative, and in the end they add to zero. Squaring the residuals solves this problem.

The least squares regression line

The *least squares regression line* is the line that *minimises* the *sum of the squares* of the *residuals*.

The *assumptions* for fitting a least squares regression line to data are the same as for using the correlation coefficient, r. These are that:

- the data are numerical
- the association is linear
- there are no clear outliers.

ISBN 978-1-108-46154-2

Cambridge University Press

How do we determine the equation of the least squares regression line?

To determine exactly the equation of the least squares regression line we need to determine the values of the intercept (a) and the slope (b) that define the line. Calculus can be used to give us rules for these values:

The equation of the least squares regression line

The equation of the least squares regression line is given by $y = a + bx$, where:

the **slope** (b) is given by $b = \dfrac{rs_y}{s_x}$

and

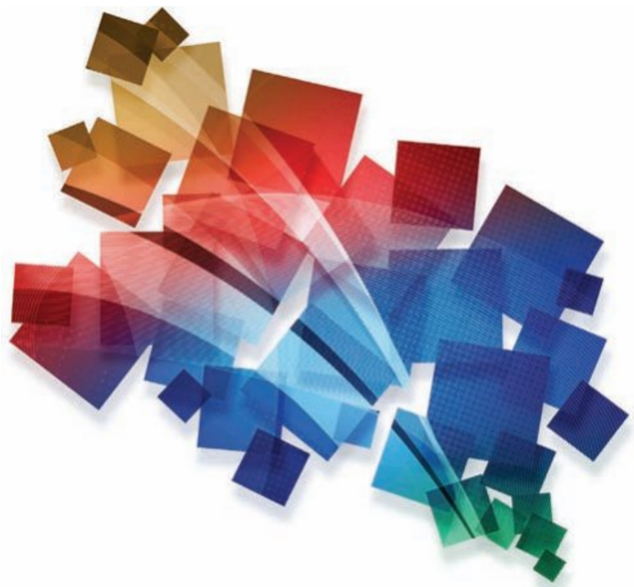
the **intercept** (a) is then given by $a = \bar{y} - b\bar{x}$

Here:

- r is the correlation coefficient
- s_x and s_y are the **standard deviations** of x and y
- $\bar{x}$ and $\bar{y}$ are the **mean** values of x and y.

Warning

If you do not correctly decide which is the explanatory variable (the x-variable) and which is the response variable (the y-variable) before you start calculating the equation of the least squares regression line, you may get the wrong answer.



ISBN 978-1-108-46154-2

Example 1 Determining the equation of the least squares regression line using the formula

The height (x) and weight (y) of 11 people have been recorded, and the values of the following statistics determined:

$\bar{x} = 173.3$ cm $s_x = 7.444$ cm $\bar{y} = 65.45$ cm $s_y = 7.594$ cm and $r = 0.8502$

Use the formula to determine the equation of the least squares regression line that enables weight to be predicted from height. Calculate the value of the slope and intercept correct to two significant figures.

Solution

1 Identify and write down the explanatory variable (EV) and the response variable (RV). Label as x and y, respectively.

Note: In saying that we want to predict weight from height, we are implying that height is the EV.

EV: height (x)
RV: weight (y)

2 Write down the given information.

$\bar{x} = 173.3 \quad s_x = 7.444$
$\bar{y} = 65.45 \quad s_y = 7.594$
$r = 0.8502$

3 Calculate the slope.

Slope:

$$b = \frac{rs_y}{s_x} = \frac{0.8502 \times 7.594}{7.444}$$
$$= 0.87 \text{ (correct to two significant figures)}$$

4 Calculate the intercept.

Intercept:

$$a = \bar{y} - b\bar{x}$$
$$= 65.45 - 0.87 \times 173.3$$
$$= -85 \text{ (correct to two significant figures)}$$

5 Use the values of the intercept and the slope to write down the least squares regression line using the variable names.

$y = -85 + 0.87x$
or
weight $= -85 + 0.87 \times$ height

ISBN 978-1-108-46154-2

Cambridge University Press

Calculator activity 2A Determining the equation of the least squares regression line using a scientific calculator

x	150	157	159	160	161	161	165	170	170	173
y	158	160	162	157	160	162	166	170	167	176

x	177	177	178	184	188	188	188	188	194	200
y	173	176	178	180	189	187	181	192	193	186

Casio fx82

Change the mode to statistics with two random variables.

Press [Mode] > [2] > [2] and you should get a table like this:

	x	y
1		
2		
3		

Insert the first data by pressing 150 then [=]. Continue inserting the data similarly for the x variables.

Use the arrow keys (on the blue circular button near the screen) to navigate to the y column and insert the y data points similarly, ensuring that each observation aligns correctly.

Leave table by pressing [AC]

Press [Shift] > [1] [STAT] > [5] [Reg] > [1] [a] > [=] to get the intercept.

a
32.97194808

Press [Shift] > [1] [STAT] > [5] [Reg] > [2] [b] > [=] to get the slope.

b
0.806640206

ISBN 978-1-108-46154-2

TI-30XB

Press [data] to move to a table.

Insert the first data by pressing [1] then [▼] and continue inserting the rest of the x data. Move to the second column by pressing [▶] and then put in the y data, ensuring that they match up with their respective x data. When finished, press clear to exit the table.

Press 2 nd > [data] [stat] > [2] > [enter] [enter] [enter] and then use the arrow keys on the top right circular button to navigate to the intercept *b*, then press [enter] [enter]. Repeat the steps again but this time navigate to the slope *a*, then press [enter] [enter]. You should get:

b	32.97194808
a	0.806640206

Note that the TI-30XB labels the intercept and slope *differently*.

Sharp

Enter statistics mode for (x, y) observations by pressing [Mode] [1] [1]

Insert the first (x, y) observation by pressing 150 > [STO] [(x, y)] > 158 > [M+] [DATA] and continue following the same process for the other observations.

Press [ALPHA] > [(][a] > [=] to get the intercept *a*:

a=
32.97194808

Press [ALPHA] > [)][b] > [=] to get the slope *b*:

b=
0.806640206

ISBN 978-1-108-46154-2

Cambridge University Press

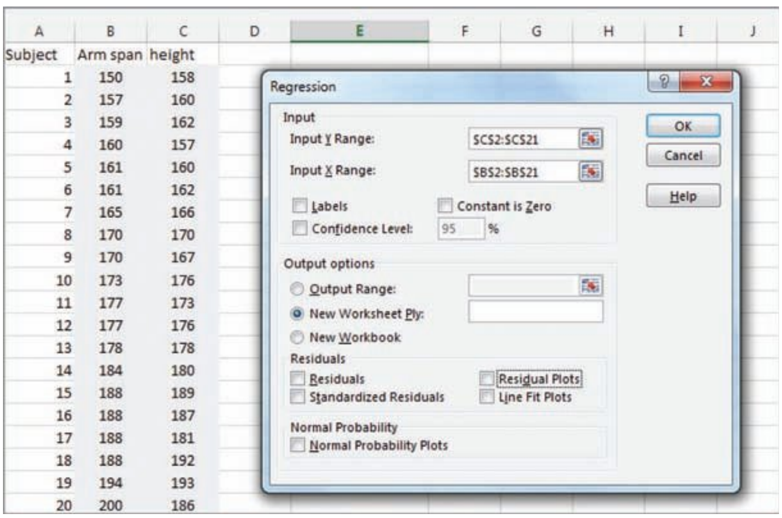
Example 2 Determining the equation of the least squares regression line using Excel

Use Excel to fit a least squares regression line to the data relating height (RV) to arm span (EV) for the group of 20 adults given in Chapter 1, Example 9. Give your answers correct to two decimal places.

Solution

1 First ensure that the Analysis ToolPakSelect is installed, then select the Data Analysis command button on the Data tab. When Excel displays the **Data Analysis** dialog box, select the **Regression** tool from the Analysis Tools list and then click OK.

2 Place the cursor on the **Input X Range** dialogue box, and then select the values of arm span. Next, place the cursor on the **Input Y Range** dialogue box, and then select the values of height.



3 Select OK to generate the following output.

SUMMARY OUTPUT								
Regression Statistics								
Multiple F	0.950668							
R Square	0.903769							
Adjusted I	0.898423							
Standard I	3.788415							
Observati	20							
ANOVA								
	df	*SS*	*MS*	*F*	*gnificance F*			
Regressio	1	2426.212	2426.212	169.0494	1.37E-10			
Residual	18	258.3376	14.35209					
Total	19	2684.55						
	Coefficient	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	32.97195	10.85292	3.038072	0.007074	10.17082	55.77308	10.17082	55.77308
X Variable	0.80664	0.06204	13.0019	1.37E-10	0.676299	0.936982	0.676299	0.936982

ISBN 978-1-108-46154-2

Cambridge University Press

4 The intercept and slope are given in the bottom table of the output. Thus:

- intercept $a = 32.97195 = 32.97$ to two decimal places
- slope $b = 0.80664 = 0.81$ to two decimal places.

Note also that the value of the coefficient of determination R^2 is also given in the output, here $R^2 = 0.903769 = 90.4\%$.

Note also that the least squares regression line can be added to the scatterplot by selecting the scatterplot, choosing Add Chart Element from the menu ribbon and then choosing Trendline.

5 *The equation of the least squares regression line is:*

height = 32.97 + 0.81 × arm span

Exercise 2A

SF

1 What is a residual?

2 The least squares regression line is obtained by:

A minimising the residuals
B minimising the sum of the residuals
C minimising the sum of the squares of the residuals
D minimising the square of the sum of the residuals
E maximising the sum of the squares of the residuals

3 Write down the three assumptions we make about the association we are modelling when we fit a least squares regression line to bivariate data.

Skillsheet

Using a formula to determine the equation of a least squares regression line

Example 1

CF

4 We wish to find the equation of the least squares regression line that enables the *pollution level* beside a freeway to be predicted from *traffic volume*.

a Which is the response variable (RV) and which is the explanatory variable (EV)?

b Use the formula to determine the equation of the least squares regression line that enables the pollution level (y) to be predicted from the traffic volume (x), where:

$$r = 0.940 \quad \bar{x} = 11.4 \quad s_x = 1.87$$
$$\bar{y} = 231 \quad s_y = 97.9$$

Write the equation in terms of *pollution level* and *traffic volume* with the y-intercept and slope written correct to two significant figures.

ISBN 978-1-108-46154-2

Cambridge University Press

5 We wish to find the equation of the least squares regression line that enables *life expectancy* in a country to be predicted from *birth rate*.

a Which is the response variable (RV) and which is the explanatory variable (EV)?

b Use the formula to determine the equation of the least squares regression line that enables life expectancy (y) to be predicted from birth rate (x), where:

$$r = -0.810 \quad \bar{x} = 34.8 \quad s_x = 5.41 \quad \bar{y} = 55.1 \quad s_y = 9.99$$

Write the equation in terms of *life expectancy* and *birth rate* with the y-intercept and slope written correct to two significant figures.

6 We wish to find the equation of the least squares regression line that enables the *distance* travelled by a car (in 1000s of km) to be predicted from its *age* (in years).

a Which is the response variable (RV) and which is the explanatory variable (EV)?

b Use the formula to determine the equation of the least squares regression line that enables distance travelled (y) by a car to be predicted from its age (x), where:

$$r = 0.947 \quad \bar{x} = 5.63 \quad s_x = 3.64 \quad \bar{y} = 78.0 \quad s_y = 42.6$$

Write the equation in terms of *distance travelled* and *age* with the y-intercept and slope written correct to two significant figures.

7 The following questions relate to the formulas used to calculate the slope and intercept of the least squares regression line. CU

a A least squares regression line is calculated and the slope is found to be negative. What does this tell us about the sign of the correlation coefficient?

b The correlation coefficient is zero. What does this tell us about the slope of the least squares regression line?

c The correlation coefficient is zero. What does this tell us about the intercept of the least squares regression line?

8 The table shows the number of sit-ups and push-ups performed by six students. CF

Sit-ups (x)	52	15	22	42	34	37
Push-ups (y)	37	26	23	51	31	45

Let the number of *sit-ups* be the explanatory (x) variable. Use your calculator to show that the equation of the least squares regression line is:

push-ups = 16.5 + 0.566 × *sit-ups* (correct to three significant figures)

ISBN 978-1-108-46154-2

Cambridge University Press

9 The table shows average hours worked and university participation rates (%) in six countries.

Hours	35.0	43.0	38.2	39.8	35.6	34.8
Rate	26	20	36	25	37	55

Use your calculator to show that the equation of the least squares regression line that enables participation *rates* to be predicted from *hours* worked is:

$$rate = 130 - 2.6 \times hours \text{ (correct to two significant figures)}$$

10 The table shows the number of *runs* scored and *balls faced* by batsmen in a cricket match.

Runs (y)	27	8	21	47	3	15	13	2	15	10	2
Balls faced (x)	29	16	19	62	13	40	16	9	28	26	6

a Use your calculator to show that the equation of the least squares regression line enabling *runs* scored to be predicted from *balls faced* is:

$$y = -2.6 + 0.73x$$

b Rewrite the regression equation in terms of the variables involved.

11 The table below shows the number of TVs and cars owned (per 1000 people) in six countries.

Number of TVs (y)	378	404	471	354	381	624
Number of cars (x)	417	286	435	370	357	550

We wish to predict the *number of TVs* from the *number of cars*.

a Which is the response variable?

b Show that, in terms of x and y, the equation of the least squares regression line is:

$$y = 61.2 + 0.930x \text{ (correct to three significant figures)}$$

c Rewrite the regression equation in terms of the variables involved.

ISBN 978-1-108-46154-2

Cambridge University Press

2B Using the least squares regression line to model a relationship between two numerical variables

So far all of our analysis has been based on the assumption that the relationship between the two variables of interest is linear. This is why it has been essential to examine the scatterplot before proceeding with any further analyses. However, sometimes the scatterplot is not sensitive enough to reveal the non-linear structure of a relationship. To gain more information we need to investigate the fit of the regression line to the data, and we do this using a **residual plot**.

▶ The residual plot – assessing the appropriateness of fitting a linear model to data

Residuals are defined as the vertical distances between the regression line and the actual data value.

Residual plot

A residual plot is a graph of the *residuals* (plotted on the vertical axis) against the *explanatory* variable (plotted on the horizontal axis), where:

Residual value = actual data value − predicted data value

Residuals can be positive, negative or zero:

- Data points above the fitted regression line have a positive residual
- Data points below the fitted regression line have a negative residual
- Data points on the fitted regression line have zero residual.

Suppose, for example, that we wish to investigate the nature of the association between the price of a secondhand car and its age. The ultimate aim is to find a mathematical model that will enable the price of a secondhand car to be predicted from its age.

To this end, the age (in years) and price (in dollars) of a selection of secondhand cars of the same brand and model have been collected and are recorded in a table overpage.

ISBN 978-1-108-46154-2

Cambridge University Press

Age (years)	Price (dollars)	Age (years)	Price (dollars)
1	32 500	4	19 200
1	30 500	5	16 000
2	25 600	5	18 400
3	20 000	6	6 500
3	24 300	7	6 400
3	22 000	7	8 500
4	22 000	8	4 200
4	23 000		

The equation of the least squares regression line from these data is:

$$price = 35\,100 - 3940 \times age$$

To determine the appropriateness of fitting the least squares regression line to these data we will construct a residual plot. But first, we need to calculate the residual for each value of the explanatory variable, in this case *age*.

Example 3 Calculating a residual

The actual price of the 6-year-old car is \$6500. Calculate the residual when its price is predicted using the regression equation: $price = 35\,100 - 3940 \times age$

Solution

1. Write down the actual price.

   Actual price: \$6500

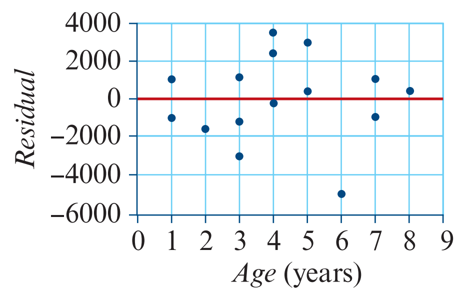
2. Determine the predicted price using the least squares regression equation: $price = 35\,100 - 3940 \times age$

   Predicted price $= 35\,100 - 3940 \times 6$
   $= \$11\,460$

3. Determine the residual.

   Residual = actual − predicted
   $= \$6500 - \$11\,460$
   $= -\$4960$

By completing this calculation for all data points, we can construct a residual plot. Because the mean of the residuals is always zero, we will construct the horizontal axis for the plot at zero (indicated by the red line) as shown.



ISBN 978-1-108-46154-2

Cambridge University Press

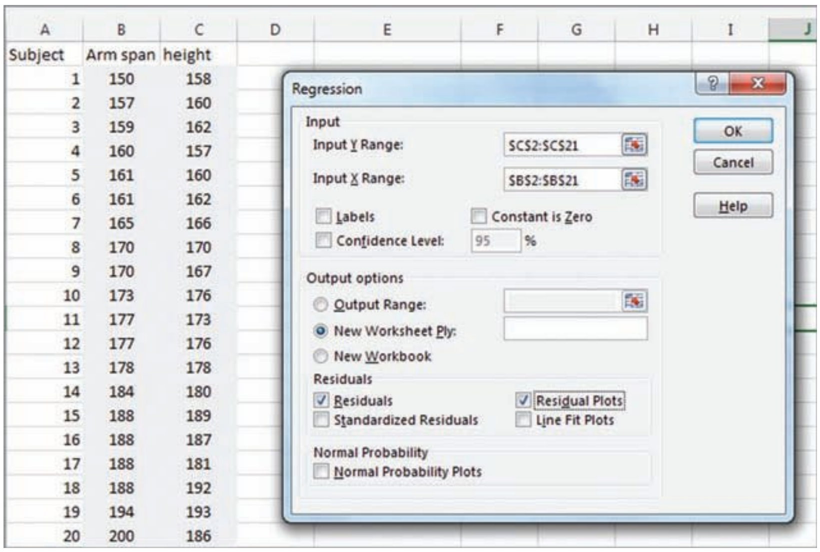
Example 4 Using Excel to construct a residual plot

Construct a residual plot for the least squares regression line fitted to the data in Example 2.

Solution

1 When setting up a spreadsheet to carry out the linear regression, check the boxes Residuals and Residual Plots.

Spreadsheet

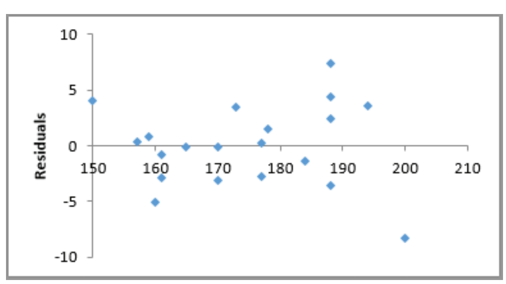


Subject	Arm span	height
1	150	158
2	157	160
3	159	162
4	160	157
5	161	160
6	161	162
7	165	166
8	170	170
9	170	167
10	173	176
11	177	173
12	177	176
13	178	178
14	184	180
15	188	189
16	188	187
17	188	181
18	188	192
19	194	193
20	200	186

2 Select OK. In addition to the regression output columns containing the predicted values of the RV (in this case height), the associated residuals will be produced, together with the associated residual plot.

RESIDUAL OUTPUT

Observatio	Predicted Y	Residuals
1	153.968	4.032021
2	159.6145	0.38554
3	161.2277	0.772259
4	162.0344	-5.03438
5	162.841	-2.84102
6	162.841	-0.84102
7	166.0676	-0.06758
8	170.1008	-0.10078
9	170.1008	-3.10078
10	172.5207	3.479296
11	175.7473	-2.74726
12	175.7473	0.252735
13	176.5539	1.446095
14	181.3937	-1.39375
15	184.6203	4.379693
16	184.6203	2.379693
17	184.6203	-3.62031
18	184.6203	7.379693
19	189.4601	3.539852
20	194.3	-8.29999

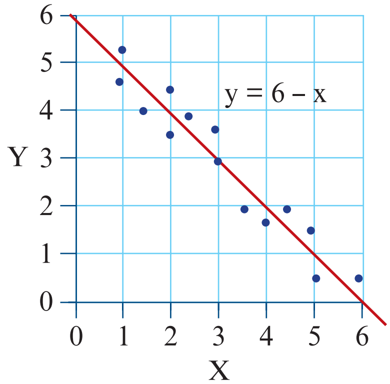


ISBN 978-1-108-46154-2

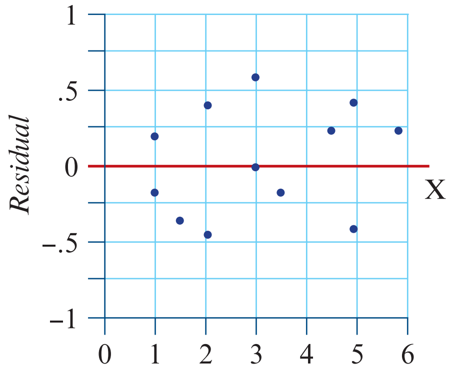
Cambridge University Press

▶ What are we looking for in a residual plot?

The scatterplot below shows a relationship that is clearly linear. When a line is fitted to the data, the resultant residual plot appears to be a random collection of points roughly spread around zero.

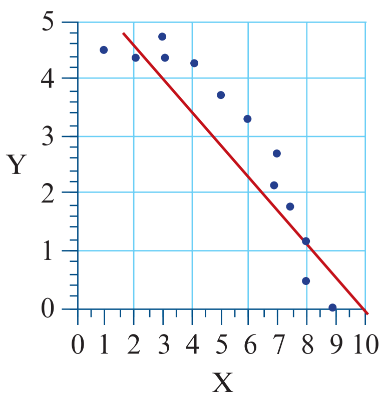


Scatterplot with regression line

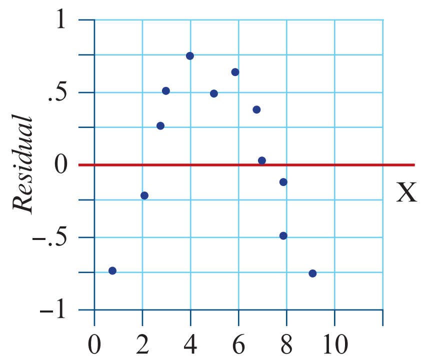


Residual plot

By contrast, the relationship shown in the following scatterplot is clearly non-linear. Fitting a straight line to the data results in the residual plot shown. While there is some random behaviour, there is also a clearly identifiable curve shown in the scatterplot.



Scatterplot displaying a non-linear relationship fitted with a regression line



Residual plot

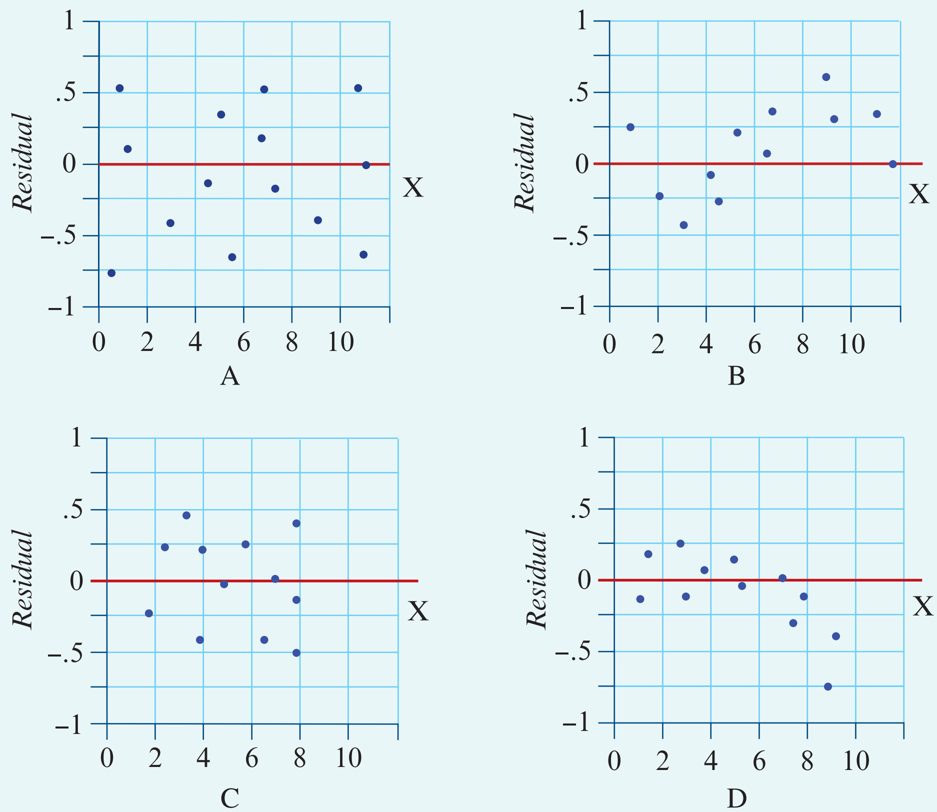
In summary, if a residual plot shows evidence of some sort of systematic behaviour (a *pattern*), then it is likely that the underlying relationship is non-linear. However, if the residual plot appears to be a random collection of points roughly spread around zero, then we can be happy that our original assumption of linearity was reasonable and that we have appropriately modelled the data.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 5 Interpreting a residual plot

Which of the following residual plots would call into question the assumption of linearity in a regression analysis? Give reasons for your answers.



Solution

Examine each plot, looking for a pattern or structure in the residual.

Plot A — residuals look random, so linearity assumption is met.

Plot B — there is a clear curve in the residuals, the linearity assumption is not met.

Plot C — residuals look random, so linearity assumption is met.

Plot B — there is a clear curve in the residuals, the linearity assumption is not met.

ISBN 978-1-108-46154-2

Cambridge University Press

Interpreting the slope and intercept of a regression line

Interpreting the slope and intercept of a regression line

For the regression line $y = a + bx$:

- the slope (b) estimates the average change (increase/decrease) in the *response variable* (y) for each one-unit increase in the *explanatory variable* (x)
- the intercept (a) estimates the average value of the *response variable* (y) when the *explanatory variable* (x) equals 0.

Consider again the least squares regression line relating the *age* of a car to its *price*:

$$price = 35\,100 - 3940 \times age$$

The two key values in this mathematical model are the intercept (35 100) and the slope (−3940). The interpretation of these values is discussed in the following example.

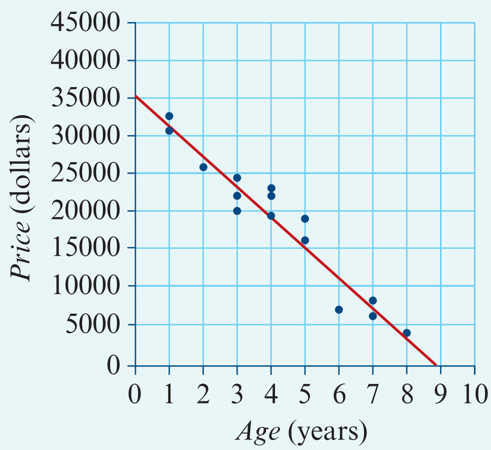
Example 6 Interpreting the slope and intercept of a regression line

The equation of a regression line that enables the *price* of a secondhand car to be predicted from its *age* is:

$$price = 35\,100 - 3940 \times age$$

a Interpret the slope in terms of the variables *price* and *age*.

b Interpret the intercept in terms of the variables *price* and *age*.



Solution

a The *slope* predicts the average change (increase/decrease) in the *price* for each one-year increase in the *age*. Because the slope is negative, it will be a decrease.

On average, the price of these cars decreases by \$3940 each year.

b The *intercept* predicts the value of the *price* of the car when *age* equals 0; that is, when the car was new.

On average, the price of these cars when new was \$35 100.

ISBN 978-1-108-46154-2

Cambridge University Press

▶ Using the coefficient of determination (R^2) to assess the strength of a linear association in terms of the explained variation

In the last chapter, we defined the coefficient of determination R^2 as the percentage of variation in the response variable, which can be explained by the variation in the explanatory variable. It is an assessment of the strength of the linear association, and it can be used to assess the relative importance of associations.

Returning to our example of the association between the price of a secondhand car and its age, here $r = -0.964$ and thus:

$$\text{coefficient of determination, } R^2 = (-0.964)^2 = 0.930 \text{ or } 93\%$$

From this was can say that 93% of the variation in the price of secondhand cars can be explained by the variation in the age of the cars, which makes the age of the car very important in predicting its value. In fact, we often find that there are several explanatory variables that may help explain the value of the response variable, and we can use the value of the coefficient of determination to determine their relative importance.

Example 7 Assessing the strength of the linear association

A recent study in a certain town found the correlation between a father's education level and their child's education level to be 0.4, whereas the correlation between a mother's education level and that of their child was 0.3. Which of these variables has a stronger relationship with a child's education level, father's education level or mother's education level? Explain your answer.

Solution

1. To compare the relative importance of these variables we need to calculate the coefficient of determination for each.

$$R^2_{\text{father, child}} = (0.4)^2 = 0.16 = 16\%$$
$$R^2_{\text{mother, child}} = (0.3)^2 = 0.09 = 9\%$$

2. Interpret and compare in terms of the variables in the study.

The father's education level is the more important explanatory variable, explaining nearly twice the variation in education level as the mother's education level.

ISBN 978-1-108-46154-2

Cambridge University Press

Using the equation of a regression line to make predictions

Once the equation of the regression line is known, we can use it to make predictions, as shown in the following example.

Example 8 **Using the equation of a regression line to make predictions**

The relationship between *eyesight test score* and *age*, based on a sample of adults aged from 20 to 60 years old, was found to be:

$$\textit{eyesight test score} = 4.65 - 0.0387 \times \textit{age}$$

Use the equation to predict (to two decimal places) the eyesight test score for:

a a person aged 40

b a person aged 65

Solution

a Substitute in the formula and evaluate.

Person aged 40
eyesight test score
$= 4.65 - 0.0387 \times 40 = 3.10$

b Substitute in the formula and evaluate.

Person aged 65
eyesight test score
$= 4.65 - 0.0387 \times 65 = 2.13$

Interpolation and extrapolation

When using a regression line to make predictions, we must be aware that, strictly speaking, the equation we have found applies only to the range of data values used to derive the equation.

Predicting *within* the range of data is called **interpolation**.

In general, we can expect a reasonably reliable result when interpolating.

Predicting *outside* the range of data is called **extrapolation**.

With extrapolation we have no way of knowing if the assocation between the two variables continues to be linear outside the data range, and thus we cannot be sure of the reliability of our prediction.

In Example 8, the data used to determine the regression equation included adults aged from 20 to 60 years old. Using this equation to predict the eyesight test score for a person aged 40 is interpolation, as 40 is within the age range used to derive the equation. On the other hand, the prediction for the person aged 65 years is extrapolation, as that person is outside the age range used.

ISBN 978-1-108-46154-2

Consider again the relationship between the price and the age of the secondhand car. This equation was determined based on data obtained from the price of cars aged from 1 to 8 years old.

Using the regression line to predict the price of a 10-year old car, would give:

$$\textit{price} = 35\,100 - 3940 \times 10 = -4300$$

Clearly we do not really expect the car to be worth a negative amount! This is an example of extrapolation, and highlights the dangers of predicting too far outside the data range. Clearly the assumption of linearity no longer holds here.

Reporting the results

After the regression analysis has been carried out, a report can be written about the findings. Here is an example of a report that could be written to summarise the association between the price and age of secondhand cars.

Report

To investigate the association between the price and age of secondhand cars data were collected from a sample of 15 cars. The scatterplot showed a strong, negative, linear relationship between the price and age of secondhand cars ($r = -0.964$), indicating that older cars tend to be lower in price. There were no obvious outliers, and the lack of a clear pattern in the residual plot confirmed the linearity assumption.

The equation of the least squares regression line is:

$$\textit{price} = 35\,100 - 3940 \times \textit{age}$$

The intercept predicts that, on average, the price of the cars when new was \$35 100.

The slope predicts that, on average, the price of the cars decreases by \$3940 each year.

The coefficient of determination indicates that 93% of the variation in the price of these secondhand cars is explained by the variation in their age.

ISBN 978-1-108-46154-2

Cambridge University Press

Performing a regression analysis

To summarise, *performing a regression analysis* involves several processes, which include:

- constructing a *scatterplot* to investigate the nature of an association
- calculating the *correlation coefficient* to indicate the strength of the relationship
- determining the equation of the *regression line*
- *interpreting the coefficients* of the *y-intercept* (a) and the *slope* (b) of the least squares regression line $y = a + bx$
- using the *coefficient of determination* to indicate the *predictive power* of the association
- using the *regression line* to make *predictions*
- calculating residuals and using a *residual plot* to test the *assumption of linearity*
- writing a *report* on your findings.

ISBN 978-1-108-46154-2

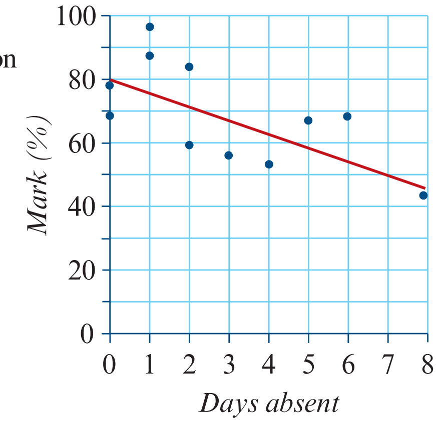
Cambridge University Press

Exercise 2B

Skillsheet

Some basics

1 Use the line on the scatterplot on the right to determine the equation of the least squares regression line in terms of the variables, *mark* and *days absent*. Give the intercept correct to the nearest whole number and the slope correct to one decimal place.



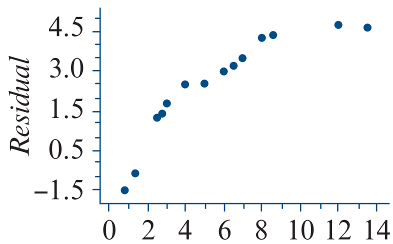
CF

Interpreting residual plots

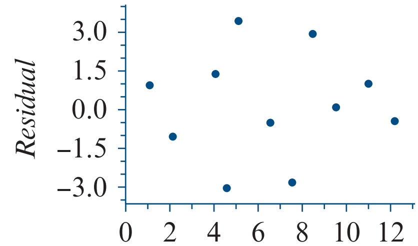
Example 5

2 Each of the following residual plots has been constructed after a least squares regression line has been fitted to a scatterplot. Which of the residual plots suggest that the use of a linear model to fit the data was inappropriate? Why?

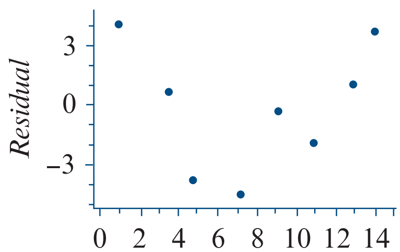
A



B



C



Reading a regression equation, making predictions and calculating residuals

Example 3,6,8

3 The equation of a regression line that enables hand span to be predicted from height is:

$$hand\ span = 2.9 + 0.33 \times height$$

Complete the following sentences, by filling in the boxes:

a The explanatory variable is ☐.

b The slope is ☐ and the intercept is ☐.

c A person is 160 cm tall. The regression line predicts a hand span of ☐ cm.

d This person has an actual hand span of 58.5 cm.

The error of prediction (residual value) is ☐ cm.

ISBN 978-1-108-46154-2

4 For a 100 km trip, the equation of a regression line that enables the fuel consumption of a car (in litres) to be predicted from its weight (kg) is:

$$\textit{fuel consumption} = -0.1 + 0.01 \times \textit{weight}$$

Complete the following sentences:

a The response variable is ______.

b The slope is ______ and the intercept is ______.

c A car weighs 980 kg. The regression line predicts a fuel consumption of ______ litres.

d This car has an actual fuel consumption of 8.9 litres.
The error of prediction (residual value) is ______ litres.

Interpreting a regression equation and its coefficient of determination

Example 7

5 In an investigation of the relationship between the energy content (in calories) and the fat content (in g) in a standard-sized packet of chips, the least squares regression line was found to be:

$$\textit{energy content} = 27.8 + 14.7 \times \textit{fat content} \text{ with } R^2 = 0.7569$$

Use this information to complete the following sentences.

a The slope is ______ and the intercept is ______.

b The regression equation predicts that the energy content in a packet of chips increases by ______ calories for each additional gram of fat it contains.

c $r =$ ______

d ______% of the variation in energy content of a packet of chips can be explained by the variation in their ______.

e The fat content of a standard-sized packet of chips is 8 g.

i The regression equation predicts its energy content to be ______ calories.

ii The *actual* energy content of this packet of chips is 132 calories.
The error of prediction (residual value) is ______ calories.

6 In an investigation of the relationship between the success rate (%) of sinking a putt and the distance of amateur golfers from the hole (in cm), the least squares regression line was found to be:

$$\textit{success rate} = 98.5 - 0.278 \times \textit{distance} \text{ with } R^2 = 0.497$$

a Write down the slope of this regression equation and interpret the meaning of this slope.

b Use the equation to predict the success rate when a golfer is 90 cm from the hole.

c At what distance (in metres) from the hole does the regression equation predict an amateur golfer to have a 0% success rate of sinking the putt?

d Calculate the value of r, correct to three decimal places.

e Write down the value of the coefficient of determination as a percentage and give an interpretation of the result.

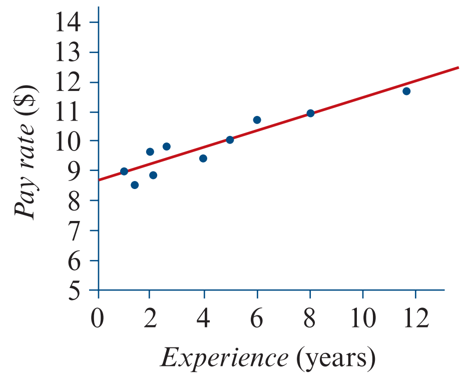
ISBN 978-1-108-46154-2

Cambridge University Press

Performing a regression analysis including the use of residual plots

7 The scatterplot opposite shows the pay rate (dollars per hour) paid by a company to workers with different years of work experience. Using a calculator, show that the least squares regression line has the equation:

$y = 8.56 + 0.289x$ with $r = 0.967$



a Is it appropriate to fit a least squares regression line to the data? Why?

b Work out the coefficient of determination.

c What percentage of the variation in a person's pay rate can be explained by the variation in their work experience?

d Write down the equation of the least squares regression line in terms of the variables *pay rate* and years of *experience*.

e Interpret the y-intercept in terms of the variables *pay rate* and years of *experience*.

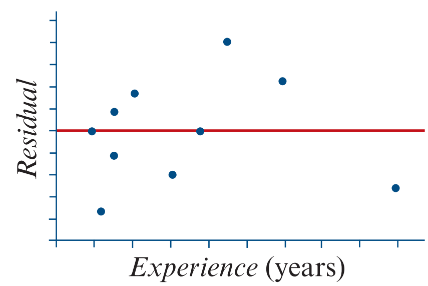
f Interpret the slope in terms of the variables *pay rate* and years of *experience*.

g Use the least squares regression equation to:

i predict the hourly wage of a person with 8 years of experience

ii determine the residual value if the person's actual hourly wage is $11.20 per hour.

h The residual plot for this regression analysis is shown opposite. Does the residual plot support the initial assumption that the relationship between *pay rate* and years of *experience* is linear? Explain your answer.

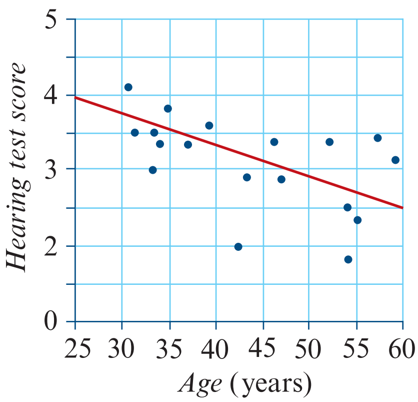


ISBN 978-1-108-46154-2

Cambridge University Press

8 The scatterplot opposite shows scores on a hearing test against age. In analysing the data, a statistician produced the following statistics:

- coefficient of determination: $R^2 = 0.370$
- least squares regression line: $y = 4.9 - 0.043x$



a Determine the value of Pearson's correlation coefficient, r, for the data.

b Interpret the coefficient of determination in terms of the variables *hearing test score* and *age*.

c Write down the equation of the least squares line in terms of the variables *hearing test score* and *age*.

d Write down the slope and interpret.

e Use the least squares regression equation to:

i predict the hearing test score of a person who is 20 years old

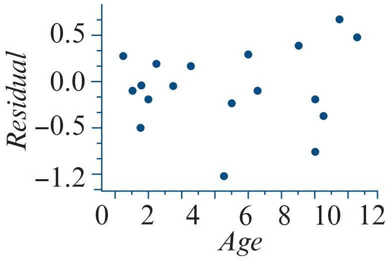
ii determine the residual value if the person's actual hearing test score is 2.0.

f Use the graph to estimate the value of the residual for the person aged:

i 35 years

ii 55 years

g The residual plot for this regression analysis is shown opposite.
Does the residual plot support the initial assumption that the relationship between hearing test score and age is essentially linear? Explain your answer.



ISBN 978-1-108-46154-2

Cambridge University Press

9 The following information was generated using Excel to investigate the relationship between the arm span (cm) and height (cm) for a group of 20 adults.

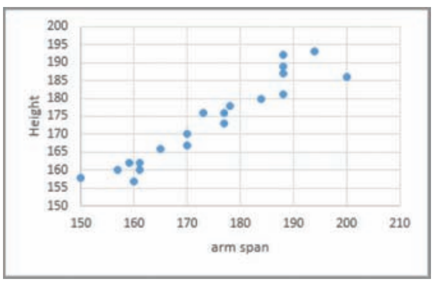
SUMMARY OUTPUT

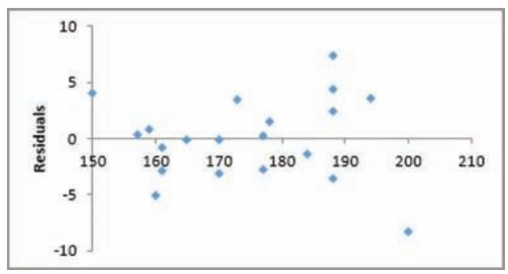
Regression Statistics	
Multiple R	0.950668
R Square	0.903769
Adjusted R	0.898423
Standard E	3.788415
Observati	20

ANOVA

	df	*SS*	*MS*	*F*	*gnificance F*
Regressio	1	2426.212	2426.212	169.0494	1.37E-10
Residual	18	258.3376	14.35209		
Total	19	2684.55			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	32.97195	10.85292	3.038072	0.007074	10.17082	55.77308	10.17082	55.77308
X Variable	0.80664	0.06204	13.0019	1.37E-10	0.676299	0.936982	0.676299	0.936982





a Is it appropriate to fit a least squares regression line to the data? Explain.

b Is the interpretation of the y-intercept meaningful? Explain.

c Use this information to complete the following report.

Report

From the scatterplot, we can see that there is a strong, ______, ______ relationship between height and arm span: r = ______. There are no obvious outliers. The equation of the least squares regression line is:

height = ______ + ______ × arm span

The slope of the regression line predicts an increase/decrease of ______ cm in height for each 1 cm increase in ______.

The coefficient of determination indicates that for this sample ______ % of the variation in ______ is explained by the variation in ______.

ISBN 978-1-108-46154-2

Cambridge University Press

10 In a study of the relationship between height and weight for females, the following data were collected.

Subject	Height (cm)	Weight (kg)
1	169	55
2	155	54
3	175	64
4	168	56
5	170	59
6	168	60
7	160	47
8	153	45
9	166	60
10	165	52
11	160	49
12	183	63
13	170	57
14	173	56
15	154	48

a Use Excel to construct a scatterplot, calculate the value of the correlation coefficient, find the equation of the least squares regression line, construct a residual plot, and find the value of the coefficient of determination (use height as the explanatory variable and weight as the response variable).

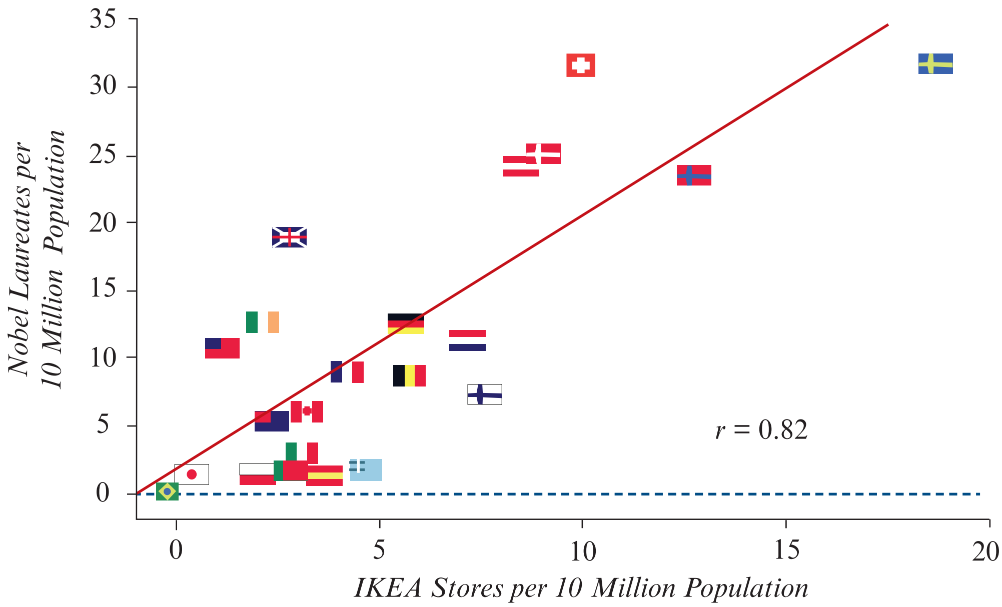
b Use these analyses to construct a report, using the structure of the report from Question 9 as a guide.

ISBN 978-1-108-46154-2

Cambridge University Press

2C Association and causation

Recently there has been interest in the strong association between the number of Nobel prizes a country has won and the number of IKEA stores in that country ($r = 0.82$). This strong association is evident in the scatterplot below. Here country flags are used to represent the data points.



Does this mean that one way to increase the number of Australian Nobel prize winners is to build more IKEA stores?

Almost certainly not, but this association highlights the problem of assuming that a strong *correlation* between two variables indicates the association between them is **causal**.

Correlation does not imply causality

A correlation tells you about the strength of the association between the variables, but no more. It tells you nothing about the source or cause of the association.

Video

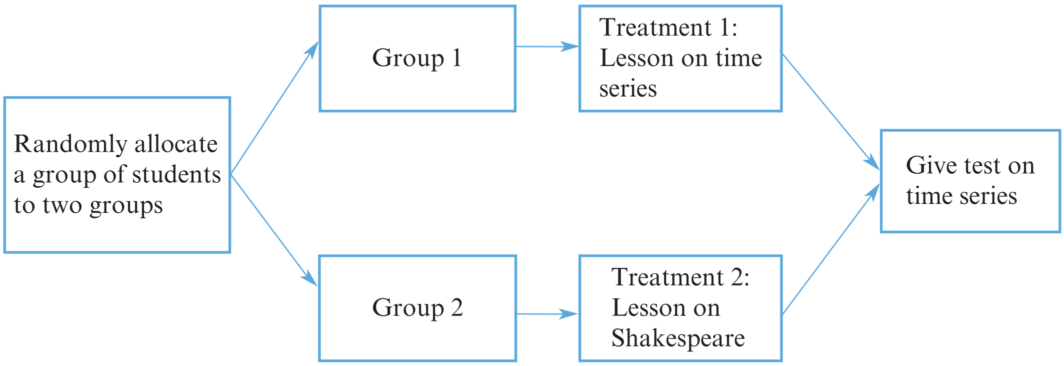
To help you with this concept you should source and watch the video 'The Question of Causation'. It is well worth 15 minutes of your time.

ISBN 978-1-108-46154-2

Cambridge University Press

Establishing causality

To establish causality, you need to conduct an *experiment*. In an experiment, the value of the *explanatory variable* is *deliberately manipulated*, while all other possible explanatory variables are kept constant or controlled. A simplified version of an experiment is displayed below.



In this experiment, a class of students is randomly allocated into two groups. Random allocation ensures that both groups are as similar as possible.

Next, group 1 is given a lesson on time series (treatment 1), while group 2 is given a lesson on Shakespeare (treatment 2). Both lessons are given under the same classroom conditions. When both groups are given a test on time series the next day, group 1 does better than group 2.

We then conclude that this was because the students in group 1 were given a lesson on time series.

Is this conclusion justified?

In this experiment, the students' test score is the response variable and the type of lesson they were given is the explanatory variable. We randomly allocated the students to each group while ensuring that all other possible explanatory variables were controlled by giving the lessons under the same classroom conditions. In these circumstances, the observed difference in the response variable (*test score*) can reasonably be attributed to the explanatory variable (*lesson type*).

Unfortunately, it is extremely difficult to conduct properly controlled experiments, particularly when the people involved are going about their everyday lives.

When data are collected through observation rather than experimentation, we must accept that strong association between two variables is insufficient evidence by itself to conclude that an observed change in the response variable has been caused by an observed change in the explanatory variable. It may be, but unless all of the relevant variables are under our control, there will always be alternative non-causal explanations to which we can appeal. We will now consider the various ways this might occur.

ISBN 978-1-108-46154-2

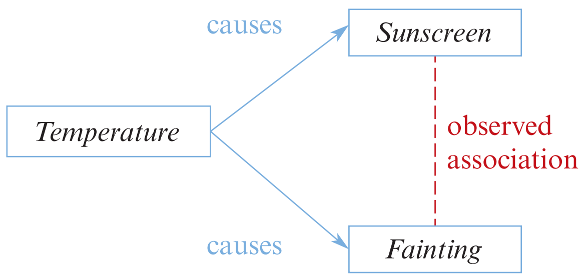
Cambridge University Press

Possible non-causal explanations for an association

Common response

Consider the following. There is a strong positive association between the number of people using sunscreen and the number of people fainting. Does this mean that applying sunscreen causes people to faint?

Almost certainly not. On hot and sunny days, more people *apply sunscreen* and more people *faint* due to heat exhaustion. The two variables are associated because they are both strongly associated with a common third variable, *temperature*. This phenomenon is called a *common response*. See the diagram below.

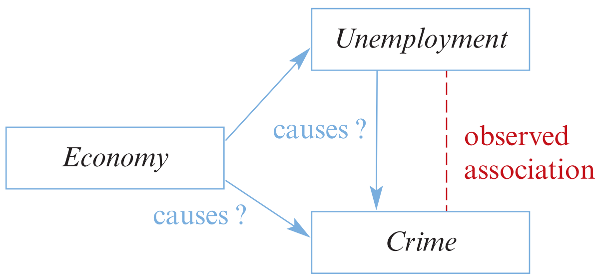


Unfortunately, being able to attribute an association to a single third variable is the exception rather than the rule. More often than not, the situation is more complex.

Confounding variables

Statistics show that *crime* rates and *unemployment* rates in a city are strongly correlated. Can you then conclude that a decrease in unemployment will lead to a decrease in crime rates?

It might, but other possible causal explanations could be found. For example, these data were collected during an economic downturn. Perhaps the state of the economy caused the problem. See the diagram below.



In this situation, we have at least two possible causal explanations for the observed association, but we have no way of disentangling their separate effects. When this happens, the effects of the two possible explanatory variables are said to be confounded, because we have no way of knowing which is the actual cause of the association.

ISBN 978-1-108-46154-2

Cambridge University Press

Coincidence

It turns out that there is a strong correlation ($r = 0.99$) between the consumption of margarine and the divorce rate in the American state of Maine. Can we conclude that eating margarine causes people in Maine to divorce?

A better explanation is that this association is purely coincidental.

Occasionally, it is almost impossible to identify any feasible confounding variables to explain a particular association. In these cases we often conclude that the association is 'spurious' and it has just happened by chance. We call this *coincidence*.

Conclusion

However suggestive a strong association may be, this alone does not provide sufficient evidence for you to conclude that two variables are causally related. Unless the association is totally spurious and devoid of meaning, it will always be possible to find at least one variable 'lurking' in the background that could explain the association.

Association (correlation) and causation

By itself, an observed association between two variables is *never enough* to justify the conclusion that two variables are causally related, no matter how obvious the causal explanation may appear to be.

Exercise 2C

CU

1. A study of primary school children aged 5 to 11 years finds a strong positive correlation between height and score on a test of mathematics ability. Does this mean that taller people are better at mathematics? What common cause might counter this conclusion?

2. There is a clear positive correlation between the number of churches in a town and the amount of alcohol consumed by its inhabitants. Does this mean that religion is encouraging people to drink? What common cause might counter this conclusion?

3. There is a strong positive correlation between the total amount of ice cream consumed and the number of drownings each day. Does this mean that eating ice cream at the beach is dangerous? What common cause might explain this association?

4. The number of days a patient stays in hospital is positively correlated with the number of beds in the hospital. Can it be said that bigger hospitals encourage patients to stay longer than necessary just to keep their beds occupied? What common cause might counter this conclusion?

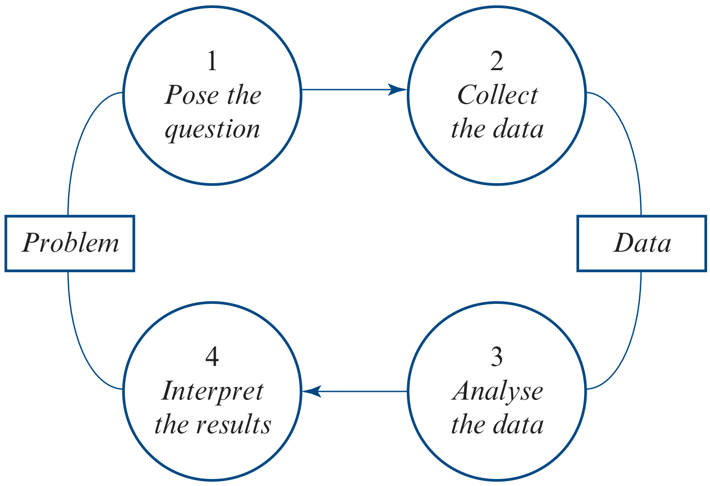
ISBN 978-1-108-46154-2

Cambridge University Press

2D Conducting an investigation – solving practical problems

In Year 11 General Mathematics we introduced the *Statistical Investigation Process*, and the steps involved in undertaking this process:

1. Pose the question – decide what data would allow you to address the problem
2. Data – collect or obtain the data
3. Analyse – summarise and display the data to answer the question posed
4. Conclusion – interpret the results and communicate what has been learned.



In Year 11 we applied these steps to practical questions such as:

- the hours that Year 11 students spent on social media in 2014 compared to 2012
- the age of mothers and fathers when their first child was born in 1970, 1990 and 2010.

The questions involved investigation of the association between a categorical and a numerical variable, and they used the techniques of analysis appropriate for these variables. We now have the tools to extend our investigations to questions concerning two categorical variables and two numerical variables.

ISBN 978-1-108-46154-2

Cambridge University Press

The following table summarises the data analysis tools we now know and can use in an investigation.

One categorical and one numerical variable	Back-to-back stemplots Boxplots **Median**, Interquartile range Mean, Standard deviation
Two categorical variables	Two-way frequency table Column percentages Stacked bar chart
Two numerical variables	Scatterplot Correlation coefficient Least squares regression line (intercept and slope) Residual plot Coefficient of determination

Example 9 **Identifying, analysing and describing the association between two categorical variables**

Does money make us happy? Investigate this relationship. Are your findings the same for males and females?

Solution

One commonly used measure that we could use to measure the amount of money that people have is socioeconomic status (SES). For a measure of happiness, we could ask respondents if they were satisfied with their life overall (yes, no). Thus, we can pose the question: 'Is there a relationship between socioeconomic status and satisfaction with life overall?'

Using a *data set* that recorded each respondent's socioeconomic status (low, mid, high) and answered the question 'Are you satisfied with your life overall?' (yes, no), the following analyses were carried out.

Firstly, the following two-way percentage frequency table summarises the responses observed in the sample of 500 people.

	Socioeconomic status (SES)		
Are you satisified with your life overall?	Low SES	Mid SES	High SES
Yes	84.6%	90.2%	79.5%
No	15.4%	9.8%	20.5%
Total	100.0%	100.0%	100.0%

ISBN 978-1-108-46154-2

The following two-way percentage frequency table summarises the responses for the females ($n = 250$) in the sample.

	Socioeconomic status (SES)		
Are you satisified with your life overall?	Low SES	Mid SES	High SES
Yes	83.9%	85.2%	84.0%
No	16.1%	14.8%	16.0%
Total	100.0%	100.0%	100.0%

The following two-way percentage frequency table summarises the responses for the males ($n = 250$) in the sample.

	Socioeconomic status (SES)		
Are you satisified with your life overall?	Low SES	Mid SES	High SES
Yes	85.2%	95.2%	74.6%
No	14.8%	4.8%	25.4%
Total	100.0%	100.0%	100.0%

Using these analyses, we are now in a position to *answer the question*.

Report

A study was conducted to investigate the relationship between socioeconomic status and satisfaction with life overall. We were particularly interested to know if those of High SES were more likely to be satisfied with their lives. Data were collected from a sample of 500 people, 250 males and 250 females. When the total group was examined, it appeared that there was a relationship between SES and satisfaction with life overall, but contrary to expectations it was the Mid SES group who were more likely to be satisfied (90.2%), followed by the Low SES group (84.6%) and then the High SES group, who were the least likely to be satisfied with their life overall (79.5%).

However, further examination showed that this relationship did not hold for both males and females when each sex was examined separately. For the females, there was no relationship between SES and satisfaction with life, with each group showing similar percentages who were satisfied (Low SES: 83.9%, Mid SES: 85.2%, High SES: 84.0%). However, there was a clear relationship for males, with the Mid SES group more likely to be satisfied (95.2%), followed by the low SES group (85.2%) and then the High SES group, who were the least likely to be satisfied with their life overall (74.6%).

ISBN 978-1-108-46154-2

Example 10 Identifying, analysing and describing the association between two numerical variables

Which of our body measurements is the best predictor of height?

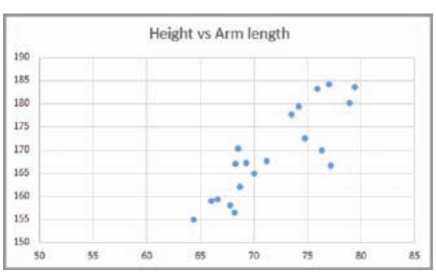
Solution

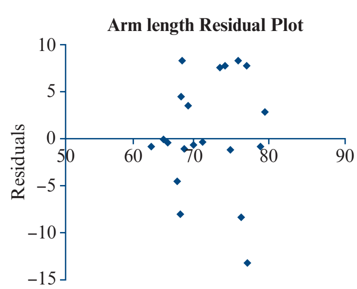
Let's look at two body measurements that are easy to collect: head circumference and arm length. Then we can *pose the question*: 'Which is the better predictor of height, head circumference or arm length?'

The following data were *collected* from a group of 20 students (all measurements are in cm):

Subject	Height	Arm length	Head circumference
1	179.3	74.2	57.1
2	170.4	68.5	57.3
3	159.0	66.0	55.5
4	172.5	74.8	60.4
5	162.0	68.7	56.8
6	167.1	68.3	54.8
7	159.4	66.6	55.9
8	167.3	69.3	57.7
9	167.7	71.2	56.2
10	184.3	77.0	60.0
11	177.7	73.5	57.5
12	169.9	76.3	56.3
13	154.9	64.4	56.0
14	156.5	68.2	57.1
15	158.0	67.8	57.5
16	165.0	70.0	58.5
17	166.6	77.2	60.0
18	183.7	79.4	59.8
19	183.3	75.9	57.3
20	180.2	78.9	57.5

Using the data, the following *analyses* were able to be produced.





$r = 0.8461$

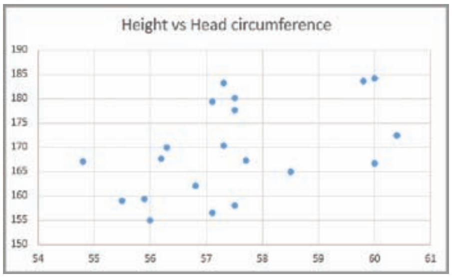
$R^2 = 71.6\%$

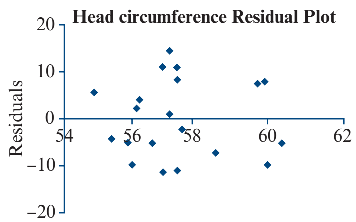
$a = 44.09$

$b = 1.742$

ISBN 978-1-108-46154-2

Cambridge University Press





$r = 0.4796$
$R^2 = 23.0\%$
$a = 2.97$
$b = 2.894$

Based on these analyses the following report could be written to *answer the question*.

Report

A study was conducted to investigate which measure was a better predictor of a person's height, head circumference or arm length. Data were collected from a sample of 20 students.

From the scatterplot of height versus arm length, we can see that there is a strong, positive, linear relationship between height and arm length: r = 0.8461. That is, those students with longer arms also tended to be taller. There are no obvious outliers, and the linearity assumption is confirmed by the residual plot. The equation of the least squares regression line is:

height = 44.09 + 1.742 × arm length

The slope of the regression line predicts an increase of 1.742 cm in height for each 1 cm increase in arm length.

From the scatterplot of height versus head circumference, we can see that there is a moderate, positive, linear relationship between height and arm length: r = 0.480. That is, those students with larger head circumference also tended to be taller. There are no obvious outliers, and the linearity assumption is confirmed by the residual plot. The equation of the least squares regression line is:

height = 2.97 + 2.894 × head circumference

The slope of the regression line predicts an increase of 2.894 cm in height for each 1 cm increase in head circumference.

Comparing the values of the coefficient of determination for each variable we can see that for this sample 71.6% of the variation in height is explained by the variation in arm length, while only 23% of the variation in height is explained by the variation in head circumference. Based on this comparison, we conclude that arm length is a much better predictor of height than head circumference.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 2D

Example 10

1 The following table shows the results of a study of obesity for a sample of 12 women and eight men. The lean body mass, in kilograms, and the resting metabolic rate for each subject in the sample is shown. The researchers hypothesised that lean body mass (a person's weight after allowing for all fat) would have a strong association with metabolic rate. Use the data below to investigate this hypothesis.

Subject	Sex	Mass (kg)	Rate
1	M	53.1	1586
2	M	52.0	1871
3	M	47.3	1363
4	F	40.8	1192
5	F	52.1	1373
6	F	42.1	1421
7	M	63.0	1669
8	F	33.4	921
9	F	34.4	1049
10	M	62.7	1812
11	F	39.8	1174
12	M	51.9	1465
13	F	35.9	989
14	M	47.0	1442
15	F	43.0	1286
16	F	54.4	1420
17	F	42.7	1132
18	M	48.5	1607
19	F	48.5	1405
20	F	49.9	1481

Example 9

2 Researchers were interested in the attitudes to women's role in society. They hypothesised that attitudes might differ based on ethnicity, and that this relationship might also differ for males and females. They also believe that attitudes may have changed in the years between 1990 and 2010. Data were collected and on the basis of these data the following tables were created. Use these analyses to report on the researcher's hypothesis.

ISBN 978-1-108-46154-2

Cambridge University Press

Males 1990	Ethnicity		
A woman should devote her time to her family	Australia	UK	Europe
Agree	139	174	185
Neither	166	140	124
Disagree	299	276	156

Females 1990	Ethnicity		
A woman should devote her time to her family	Australia	UK	Europe
Agree	147	160	145
Neither	126	121	113
Disagree	312	312	270

Males 2010	Ethnicity		
A woman should devote her time to her family	Australia	UK	Europe
Agree	120	151	146
Neither	144	151	126
Disagree	597	604	485

Females 2010	Ethnicity		
A woman should devote her time to her family	Australia	UK	Europe
Agree	58	63	74
Neither	116	121	112
Disagree	406	391	345

3 What is the attitude of Queensland to lowering the driving age? Is that attitude the same for different age groups (under 18, from 18–30, over 30, for example)? Is it the same for males and females? Carry out a statistical investigation to address this question. You will need to collect appropriate data.

4 Is an individual's reaction time faster when they use their dominant hand, compared to their other hand? Carry out a statistical investigation to address this question. You will need to collect appropriate data.

ISBN 978-1-108-46154-2

Key ideas and chapter summary

AS

Linear regression

The process of fitting a line to data is known as **linear regression**.

Least squares method

The **least squares method** is one way of finding the equation of a regression line. It minimises the sum of the squares of the residuals. It works best when there are no outliers.
The equation of the least squares regression line is given by $y = a + bx$, where a represents the *y-intercept* of the line and b the *slope*.

Residuals

The vertical distance from a data point to the straight line is called a **residual**:
residual value = data value − predicted value.

Using the regression line

The regression line $y = a + bx$ enables the value of y to be determined for a given value of x.
For example, the regression line

$$cost = 1.20 + 0.06 \times number\ of\ pages$$

predicts that the cost of a 100-page book is:

$$cost = 1.20 + 0.06 \times 100 = \$7.20$$

Slope and intercept

The **slope** of the regression line above predicts that the cost of a textbook increases by 6 cents ($0.06) for each additional page.The **intercept** of the line predicts that a book with no pages costs $1.20 (this might be the cost of the cover).

Residual plots

Residual plots can be used to test the linearity assumption by plotting the residuals against the EV.
A residual plot that appears to be a random collection of points clustered around zero supports the linearity assumption.
A residual plot that shows a clear pattern indicates that the association is not linear.

Interpolation and extrapolation

Predicting *within* the range of data is called **interpolation**.
Predicting *outside* the range of data is called **extrapolation**.

Correlation and causation

A *correlation* between two variables does not automatically imply that the association is *causal*. Alternative *non-causal explanations* for the association include a *common response* to a common third variable, a *confounded* variable or simply *coincidence*.

ISBN 978-1-108-46154-2

Cambridge University Press

Skills check

Having completed this chapter you should be able to:

- determine the equation of the least squares line using the formulas $b = \frac{rs_y}{s_x}$ and $a = \bar{y} - b\bar{x}$
- for raw data, determine the equation of the least squares line using a scientific calculator
- interpret the slope and intercept of a regression line
- interpret the coefficient of determination as part of a regression analysis
- use the regression line for prediction
- calculate residuals
- construct a residual plot using a spreadsheet
- use a residual plot to determine the appropriateness of using the equation of the least squares line to model the association
- present the results of a regression analysis in report form.

Multiple-choice questions

1 When using a least squares regression line to model a relationship displayed in a scatterplot, one key assumption is that:

A there are two variables
B the variables are related
C the variables are linearly related
D $r^2 > 0.5$
E the correlation coefficient is positive

2 In the least squares regression line $y = -1.2 + 0.52x$:

A the y-intercept $= -0.52$ and slope $= -1.2$
B the y-intercept $= 0$ and slope $= -1.2$
C the y-intercept $= 0.52$ and slope $= -1.2$
D the y-intercept $= -1.2$ and slope $= 0.52$
E the y-intercept $= 1.2$ and slope $= -0.52$

3 If the equation of a least squares regression line is $y = 8 - 9x$ and $R^2 = 0.25$:

A $r = -0.5$ **B** $r = -0.25$ **C** $r = -0.0625$ **D** $r = 0.25$ **E** $r = 0.50$

4 The least squares regression line $y = 8 - 9x$ predicts that, when $x = 5$, the value of y is:

A -45 **B** -37 **C** 37 **D** 45 **E** 53

ISBN 978-1-108-46154-2

5 A least squares regression line of the form $y = a + bx$ is fitted to the data set shown.

x	25	15	10	5
y	10	10	15	25

The equation of the line is:

A $y = -0.69 + 24.4x$ **B** $y = 24.4 - 0.69x$ **C** $y = 24.4 + 0.69x$
D $y = 28.7 - x$ **E** $y = 28.7 + x$

6 A least squares regression line of the form $y = a + bx$ is fitted to the data set shown.

y	30	25	15	10
x	40	20	30	10

The equation of the line is:

A $y = 1 + 0.5x$ **B** $y = 0.5 + x$ **C** $y = 0.5 + 7.5x$
D $y = 7.5 + 0.5x$ **E** $y = 30 - 0.5x$

7 Given that $r = 0.733$, $s_x = 1.871$ and $s_y = 3.391$, the slope of the least squares regression line is closest to:

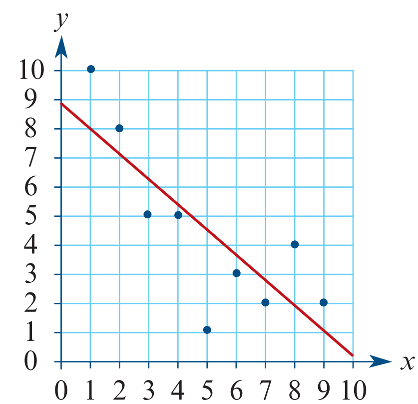
A 0.41 **B** 0.45 **C** 1.33 **D** 1.87 **E** 2.49

8 Using a least squares regression line, the predicted value of a data point is 78.6. The residual value is –5.4. The actual data value is:

A 73.2 **B** 84.0 **C** 88.6 **D** 94.6 **E** 424.4

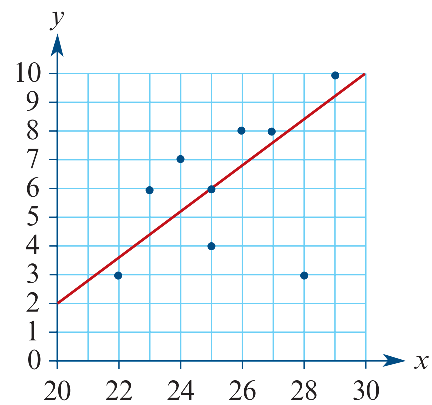
9 The equation of the least squares regression line plotted on the scatterplot opposite is closest to:

A $y = 8.7 - 0.9x$
B $y = 8.7 + 0.9x$
C $y = 0.9 - 8.7x$
D $y = 0.9 + 8.7x$
E $y = 8.7 - 0.1x$



10 The equation of the least squares regression line plotted on the scatterplot opposite is closest to:

A $y = -14 + 0.8x$
B $y = 0.8 + 14x$
C $y = 2.5 + 0.8x$
D $y = 14 - 0.8x$
E $y = 17 + 1.2x$



ISBN 978-1-108-46154-2

Cambridge University Press

Use the following information to answer Questions 11 to 14.

Weight (in kg) can be predicted from height (in cm) using the regression line:

$weight = -96 + 0.95 \times height$, with $r = 0.79$

11 Which of the following statements relating to the regression line is *false*?

A The slope of the regression line is 0.95.
B The explanatory variable in the regression equation is *height*.
C The least squares line does *not* pass through the origin.
D The intercept is 96.
E The equation predicts that a person who is 180 cm tall will weigh 75 kg.

12 This regression line predicts that, on average, weight:

A decreases by 96 kg for each 1 centimetre increase in height
B increases by 96 kg for each 1 centimetre increase in height
C decreases by 0.79 kg for each 1 centimetre increase in height
D decreases by 0.95 kg for each 1 centimetre increase in height
E increases by 0.95 kg for each 1 centimetre increase in height

13 Noting that the value of the correlation coefficient is $r = 0.79$, we can say that:

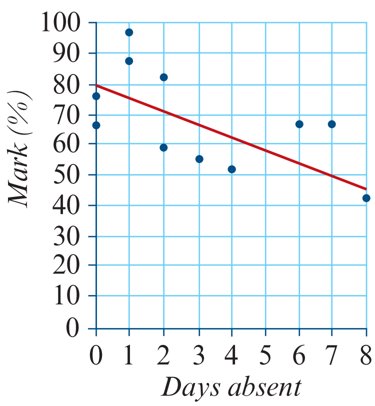
A 62% of the variation in weight can be explained by the variation in height
B 79% of the variation in weight can be explained by the variation in height
C 88% of the variation in weight can be explained by the variation in height
D 79% of the variation in height can be explained by the variation in weight
E 95% of the variation in height can be explained by the variation in weight

14 A person of height 179 cm weighs 82 kg. If the regression equation is used to predict their weight, then the residual will be closest to:

A –8 kg **B** 3 kg **C** –3 kg **D** 9 kg **E** 74 kg

15 The coefficient of determination for the data displayed in the scatterplot opposite is close to $R^2 = 0.5$.
The correlation coefficient is closest to:

A −0.7
B −0.25
C 0.25
D 0.5
E 0.7



ISBN 978-1-108-46154-2

16 There is a strong, linear, positive correlation ($r = 0.85$) between the amount of *garbage recycled* and *salary level*.
From this information, we can conclude that:

A the amount of garbage recycled can be increased by increasing people's salaries
B the amount of garbage recycled can be increased by decreasing people's salaries
C increasing the amount of garbage you recycle will increase your salary
D people on high salaries tend to recycle less garbage
E people on high salaries tend to recycle more garbage

17 There is a strong, linear, positive correlation ($r = 0.95$) between the marriage rate in Kentucky and the number of people who drown falling out of a fishing boat.
From this information, the most likely conclusion we can draw is:

A reducing the number of marriages in Kentucky will decrease the number of people who drown falling out of a fishing boat
B increasing the number of marriages in Kentucky will increase the number of people who drown falling out of a fishing boat
C this correlation is just coincidence, and a change in the marriage rate will not affect the number of people drowning in Kentucky in any way
D only married people in Kentucky drown falling out of a fishing boat
E stopping people from going fishing will reduce the marriage rate in Kentucky

Short-answer questions

1 What are the three assumptions which must be met when we fit a least squares regression line to bivariate data? SF

2 Find the values of a and b in the linear equation:

$$y = a + bx$$

where $r = 0.75$, $\bar{x} = 12.6$, $s_x = 2.4$, $\bar{y} = 124.8$, $s_y = 8.4$

3 If the correlation coefficient is equal to 0: CU

a What does this tell us about the slope of the least squares regression line?
b What does this tell us about the intercept of the least squares regression line?

4 A retailer recorded the number of ice creams sold and the day's maximum temperature over 8 consecutive Saturdays one summer. Use the data in the table to determine the equation of the least squares regression line for these data. Write your equation in terms of the variable in the table. SF

Temperature (°C)	22	25	36	34	21	28	41	31
Number of ice creams sold	145	155	200	198	150	179	230	180

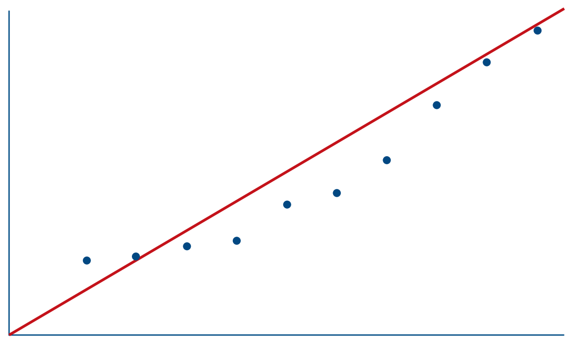
ISBN 978-1-108-46154-2

Cambridge University Press

Review

5 The actual price of a 10-year-old car is \$15 600. Calculate the residual when its price is predicted using the regression equation: price = 57 500 − 4250 × age

6 The relationship between two variables y and x as shown is non-linear as shown in the scatterplot below. CU



Sketch the residual plot which would result when the straight line shown is fitted to this data

7 A regression equation that enables the price of a second-hand caravan to be predicted from its age is: CF

$$\text{price} = 87\,500 - 5675 \times \text{age}$$

a Interpret the slope in terms of the variables price and age

b Interpret the intercept in terms of the variables price and age

8 In a recent study across a number of countries the correlation between educational attainment and the amount spent on education was found to be 0.23, whilst the correlation between educational attainment and the student : teacher ratio was found to be –0.34. CU

a Interpret each of these correlation coefficients in terms of the variables in the study.

b Which of the variables, *amount spent on education* or *student : teacher ratio* is more important in explaining the variation in educational attainment? Explain your answer.

9 The relationship between reaction time (in seconds) in a certain experiment and age (in years), based on a sample of a group of adults aged 40–70 years old, was found to be: CF

$$\text{reaction time} = 1.93 + 0.036 \times \text{age}$$

Use the equation to predict (to two decimal places) the reaction time for:

a A person aged 55 years

b A person aged 35 years

c Comment on the reliability of each of these predictions.

ISBN 978-1-108-46154-2

Cambridge University Press

10 The equation of a regression line that enables scores on a certain test to be predicted from IQ is: SF

$$\text{score} = 32 + 0.34 \times \text{IQ}$$

Complete the following sentences:

a The response variable is ______.

b The slope is ______ and the intercept is ______.

c A person has an IQ of 112. The equation predicts a test score of ______.

d This person has an actual test score of 78. The residual value is ______.

11 Explain the difference between correlation and causation, including an example of each to illustrate your explanation. CU

12 There is a strong correlation between level of maturity and the number of children a person has. Can we assume from this that having children matures a person?

Extended-response questions

1 In an investigation of the relationship between the number of hours of sunshine (per year) and the number of days of rain (per year) for 25 cities, the least squares regression line was found to be: CF

$$\textit{hours of sunshine} = 2850 - 6.88 \times \textit{days of rain}, \text{ with } R^2 = 0.484$$

Use this information to complete the following sentences.

a In this regression equation, the explanatory variable is ______.

b The slope is ______ and the intercept is ______.

c The regression equation predicts that a city that has 120 days of rain per year will have ______ hours of sunshine per year.

d The slope of the regression line predicts that the hours of sunshine per year will ______ by ______ hours for each additional day of rain.

e $r =$ ______, correct to three significant figures.

f ______ % of the variation in sunshine hours can be explained by the variation in ______.

g One of the cities used to determine the regression equation had 142 days of rain and 1390 hours of sunshine.

  i The regression equation predicts that it has ______ hours of sunshine.

  ii The residual value for this city is ______ hours.

h Using a regression line to make predictions within the range of data used to determine the regression equation is called ______.

 ISBN 978-1-108-46154-2 Cambridge University Press

Review

2 The cost of preparing meals in a school canteen is linearly related to the number of meals prepared. To help the caterers predict the costs, data were collected on the cost of preparing meals for different levels of demands. The data are shown below.

Number of meals	30	35	40	45	50	55	60	65	70	75	80
Cost (dollars)	138	154	159	182	198	198	214	208	238	234	244

a Which is the response variable?

b Use your calculator to show that the equation of the least squares regression line that relates the cost of preparing meals to the number of meals produced is:

$$cost = 81.5 + 2.10 \times number\ of\ meals$$

c Use the equation to predict the cost of producing:

i 48 meals. In making this prediction are you interpolating or extrapolating?

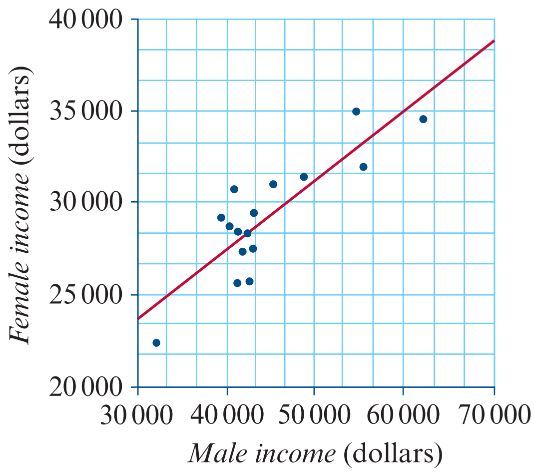
ii 21 meals. In making this prediction are you interpolating or extrapolating?

d **i** Write down and interpret the intercept of the regression line.

ii Write down and interpret the slope of the regression line.

e If $r = 0.978$, write down the coefficient of determination and interpret.

3 In the scatterplot opposite, average annual *female income,* in dollars, is plotted against average annual *male income,* in dollars . This data is provided, for 16 countries. A least squares regression line is fitted to the data.



The equation of the least squares regression line for predicting female income from male income is *female income* = 13 000 + 0.35 × *male income*.

a What is the explanatory variable?

b Complete the following statement by filling in the missing information.

From the least squares regression line equation it can be concluded that, for these countries, on average, female income increases by $ __________ for each $1000 increase in male income.

c **i** Use the equation of the least squares regression line to predict the average annual female income (in dollars) in a country where the average annual male income is $15 000.

ii The prediction made in **part c i** is not likely to be reliable. Explain why.

©VCAA (2000)

ISBN 978-1-108-46154-2

Cambridge University Press

4 We wish to find the equation of the least squares regression line that will enable a person's height (in cm) to be predicted from femur (thigh bone) length (in cm).

a Which is the response variable and which is the explanatory variable?

b Use the following summary statistics to determine the equation of the least squares regression line that will enable height (y) to be predicted from femur length (x).

$$r = 0.9939 \quad \bar{x} = 24.246 \quad s_x = 1.873 \quad \bar{y} = 166.092 \quad s_y = 10.086$$

Write the equation in terms of height and femur length. Give the slope and intercept accurate to three significant figures.

c Interpret the slope of the regression equation in terms of height and femur length.

d Determine the value of the coefficient of determination and interpret in terms of height and femur length.

5 The data below shows the height (in cm) of a group of 10 children aged 2 to 11 years.

Height (cm)	86.5	95.5	103.0	109.8	116.4	122.4	128.2	133.8	139.6	145.0
Age (years)	2	3	4	5	6	7	8	9	10	11

The task is to determine the equation of a least squares regression line that can be used to predict height from age.

a In this analysis, which would be the response variable and which would be the explanatory variable?

b Use your calculator to confirm that the equation of the least squares regression line is $height = 76.64 + 6.366 \times age$ and $r = 0.9973$.

c Use the regression line to predict the height of a 1-year-old child. Give the answer correct to the nearest cm. In making this prediction are you extrapolating or interpolating?

d What is the slope of the regression line and what does it tell you in terms of the variables involved?

e Calculate the value of the coefficient of determination and interpret in terms of the relationship between age and height.

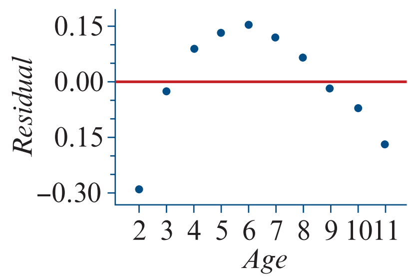
f Use the least squares regression equation to:

i predict the height of the 10-year-old child in this sample

ii determine the residual value for this child

g **i** Confirm that the residual plot for this analysis is shown opposite.

ii Explain why this residual plot suggests that a linear equation is not the most appropriate model for this relationship.



ISBN 978-1-108-46154-2

Cambridge University Press

6 The table below shows the scores obtained by nine students on two tests. We want to be able to predict test B scores from test A scores.

Test A score (x)	18	15	9	12	11	19	11	14	16
Test B score (y)	15	17	11	10	13	17	11	15	19

Use appropriate technology to perform each of the following steps of a regression analysis.

a Construct a scatterplot. Name the variables test A and test B.
b Determine the equation of the least squares line along with the values of r and R^2.
c Display the regression line on the scatterplot.
d Obtain a residual plot.
e Write a report summarising your findings.

7 The table below shows the test scores and number of careless errors made by the same nine students on the two tests from **6**. We want to be able to predict test score from the number of careless errors made.

Test score	18	15	9	12	11	19	11	14	16
Careless errors	0	2	5	6	4	1	8	3	1

Use appropriate technology to perform each of the following steps of a regression analysis.

a Construct a scatterplot. Name the variables *score* and *errors*.
b Determine the equation of the least squares line along with the values of r and R^2. Write answers correct to three significant figures.
c Display the regression line on the scatterplot.
d Obtain a residual plot.
e Write a report summarising your findings.

ISBN 978-1-108-46154-2

Cambridge University Press

8 How well can we predict an adult's weight from their birth weight? The weights of 12 adults were recorded, along with their birth weights. The results are shown.

Birth weight (kg)	1.9	2.4	2.6	2.7	2.9	3.2	3.4	3.4	3.6	3.7	3.8	4.1
Adult weight (kg)	47.6	53.1	52.2	56.2	57.6	59.9	55.3	58.5	56.7	59.9	63.5	61.2

a In this investigation, which would be the response variable and which would be the explanatory variable?

b Construct a scatterplot.

c Use the scatterplot to:

  i comment on the relationship between adult weight and birth weight in terms of direction, outliers, form and strength

  ii *estimate* the value of Pearson's correlation coefficient, r

d Determine the equation of the least squares regression line, the coefficient of determination and the value of Pearson's correlation coefficient, r. Write answers correct to three significant figures.

e Interpret the coefficient of determination in terms of adult weight and birth weight.

f Interpret the slope in terms of adult weight and birth weight.

g Use the regression equation to predict the weight of an adult with a birth weight of:

  i 3.0 kg **ii** 2.5 kg **iii** 3.9 kg

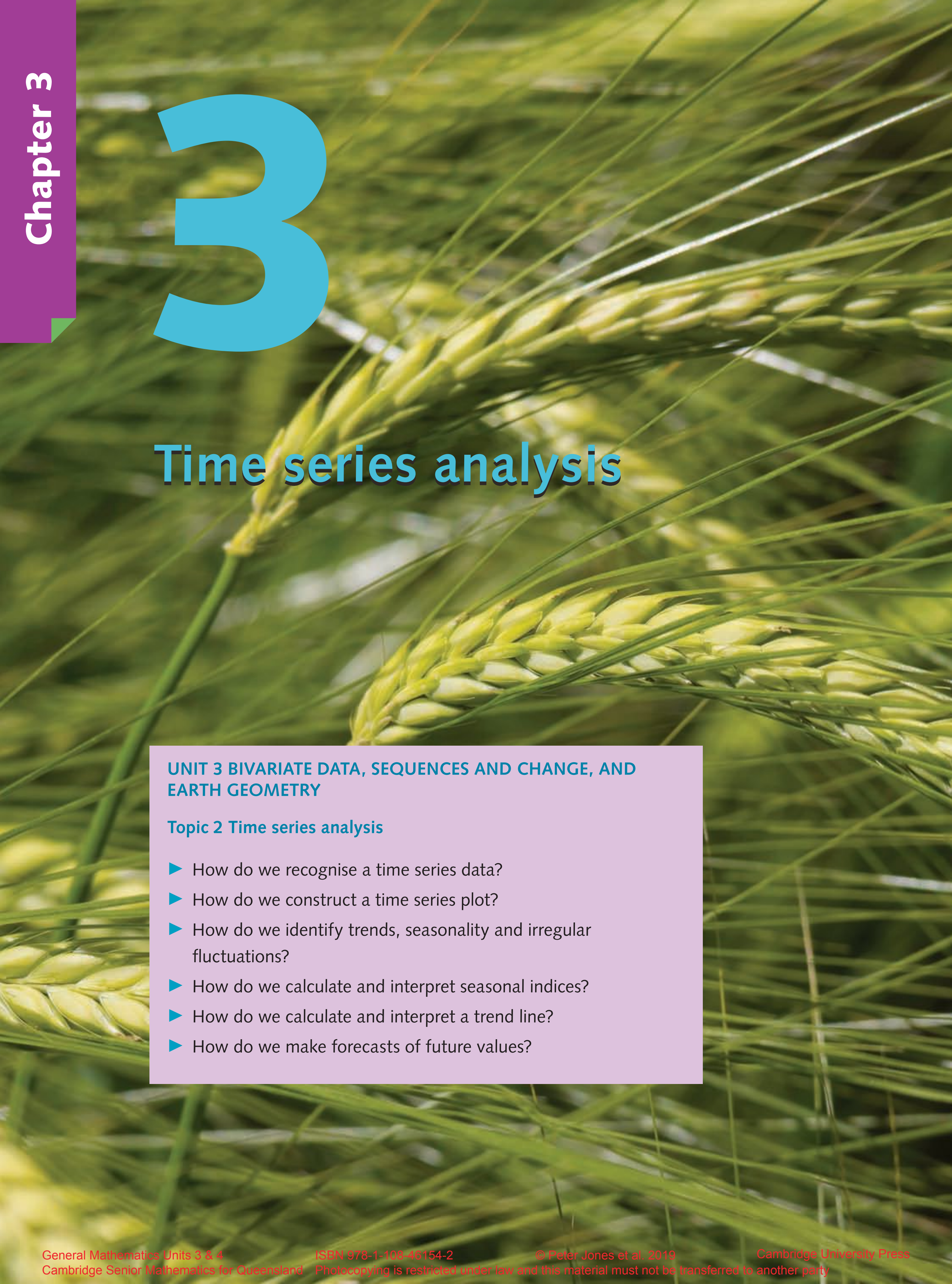
Give answers correct to one decimal place.

h It is generally considered that birth weight is a 'good' predictor of adult weight. Do you think the data support this contention? Explain.

i Construct a residual plot and use it to comment on the appropriateness of assuming that adult weight and birth weight are linearly associated.

ISBN 978-1-108-46154-2

Cambridge University Press



Chapter 3

3

Time series analysis

UNIT 3 BIVARIATE DATA, SEQUENCES AND CHANGE, AND EARTH GEOMETRY

Topic 2 Time series analysis

- How do we recognise a time series data?
- How do we construct a time series plot?
- How do we identify trends, seasonality and irregular fluctuations?
- How do we calculate and interpret seasonal indices?
- How do we calculate and interpret a trend line?
- How do we make forecasts of future values?

General Mathematics Units 3 & 4
Cambridge Senior Mathematics for Queensland
ISBN 978-1-108-46154-2

Introduction

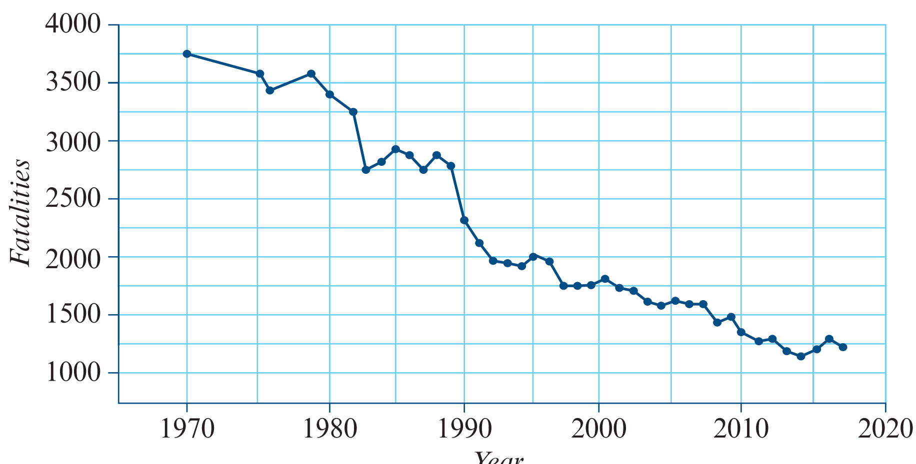
In this chapter we will focus on a special case of numerical bivariate data, called **time series data**. In time series data the *explanatory variable is a measure of time* (for example hour, day, month or year), and we are concerned with understanding how the response variable is changing over time.

3A Time series data

When data concerned with a variable is collected, observed or recorded at successive intervals of time, it is referred to as *time series data*. An example of time series data is Annual road accident fatalities for Australia, 1975–2017.

Year	Fatalities	Year	Fatalities	Year	Fatalities	Year	Fatalities
1975	3634	1988	2887	1998	1755	2008	1437
1976	3456	1989	2801	1999	1764	2009	1491
1979	3587	1990	2331	2000	1817	2010	1353
1980	3403	1991	2113	2001	1737	2011	1277
1982	3252	1992	1974	2002	1715	2012	1300
1983	2755	1993	1953	2003	1621	2013	1187
1984	2822	1994	1928	2004	1583	2014	1150
1985	2941	1995	2017	2005	1627	2015	1209
1986	2888	1996	1970	2006	1598	2016	1293
1987	2772	1997	1767	2007	1603	2017	1225

Since a time series is just a special kind of two numerical variable example, where the explanatory variable is time, we will begin by drawing a scatterplot of the data. In this instance, the scatterplot is called a **time series plot**, with *time* always placed on the horizontal axis. A time series plot differs from a normal scatterplot in that, in general, the points will be joined by line segments in time order. An example of a time series plot, of the road accident fatality data, is given below.



ISBN 978-1-108-46154-2

Cambridge University Press

Looking at the time series plot, we can readily see a clear trend of decreasing road fatalities, which is good news for drivers, as this provides some evidence that the many efforts being made to reduce the road toll across Australia have been effective.

Example 1 Constructing a time series plot

Maximum temperature was recorded each day for a week in a certain town. Construct a time series plot of the data.

Day	Mon	Tues	Wed	Thur	Fri	Sat	Sun
Temperature (°C)	20	21	25	36	34	25	26

Solution

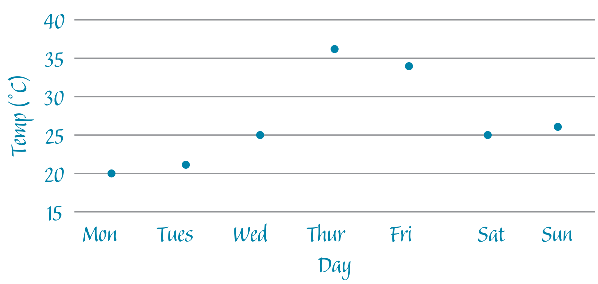
1 In a time series plot, time (day in this case) is always the explanatory variable (EV) and is plotted on the horizontal axis.

Day is the EV – this will label the horizontal axis.
Temperature is the RV – this will label the vertical axis.

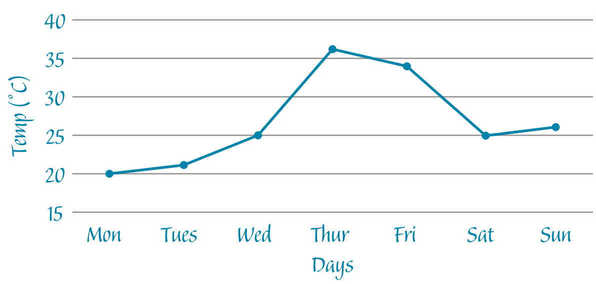
2 Determine the scales for each axis.

A horizontal scale from 0–7 with intervals of 1 for each day would be suitable.
Temperature ranges from 20–36. A vertical scale from 15–40 with intervals of 5 would be suitable.

3 Set up the axes, and then plot all seven data points as for a scatterplot.



4 Complete the graph, by joining consecutive data points with straight lines.



ISBN 978-1-108-46154-2

Cambridge University Press

Example 2 Using Excel to construct a time series plot

Use the data from Example 1 to construct a time series plot using Excel.

Solution

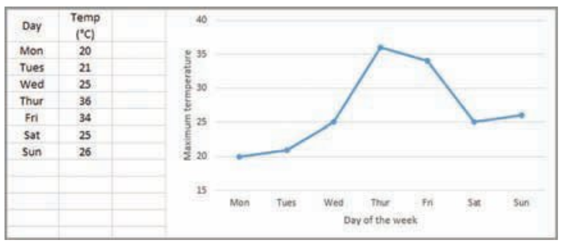
1 Enter the data into two columns as shown below.

2 Select both columns (including headings) and on the **Insert** tab, in the **Charts** group, click on **Scatter** and then the option **Scatter with Straight Lines and Markers**.

3 Double click on each scale separately and edit to include the range of the data.

4 Axis labels can be added using the Add Chart Elements option when editing the scale.

Spreadsheet



ISBN 978-1-108-46154-2

Exercise 3A

SF

Note: Save the time series plots constructed in this exercise for use in Exercise 3B.

Example 1

1 Construct a time series plot to display the following data.

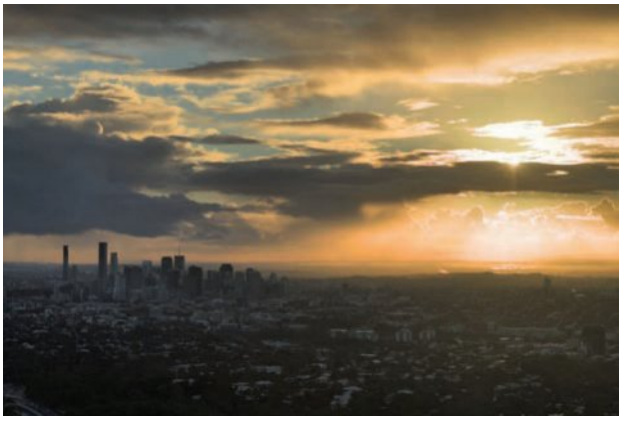
Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
Sales	2	23	35	50	31	45	23	67	70

2 Researchers recorded the number of penguins present on a remote island each month for 12 months. Construct a time series plot of the data.

Month	Number of penguins
January	449
February	214
March	170
April	265
May	434
June	102
July	180
August	241
September	311
October	499
November	598
December	674

3 The following table shows the minimum temperature in Brisbane during one week in January. Construct a time series plot of the data.

Day	Mon	Tues	Wed	Thur	Fri	Sat	Sun
Temperature (°C)	24.0	24.2	17.4	17.7	18.3	19.5	17.4



ISBN 978-1-108-46154-2

4 A time series plot for the population of Australia over the period 2006–2017. Construct a times series plot of the data.

Population of Australia 2006–2017	
Year	Population (1 000 000)
2017	24.5
2016	24.1
2015	23.8
2014	23.5
2013	23.1
2012	22.7
2011	22.3
2010	22.0
2009	21.7
2008	21.3
2007	20.8
2006	20.7

5 The table below shows the motor vehicle theft rate per 100 000 cars in Australia from 2003 to 2014. Construct a time series plot of the data.

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Theft rate	498.1	440.0	396.4	365.8	336.6	319.4	274.0	214.7	220.0	228.4	204.1	190.8

Example 2

6 The table below shows the number of measles cases reported in Australia from 1988 to 2016. Use appropriate technology to construct a time series plot of the data.

Year	Measles	Year	Measles	Year	Measles
1988	248	1998	327	2009	104
1989	169	1999	235	2010	70
1990	880	2000	108	2011	190
1991	1380	2001	141	2012	199
1992	1425	2002	32	2013	158
1993	4536	2003	91	2014	340
1994	4895	2004	70	2015	74
1995	1198	2005	10	2016	99
1996	498	2007	11		
1997	853	2008	65		

ISBN 978-1-108-46154-2

7 The table below shows carbon intensity (a measure of carbon emissions from coal) for Australia for the years from 1961 to 2014. Use appropriate technology to construct a time series plot of the data.

Year	Carbon intensity	Year	Carbon intensity	Year	Carbon intensity
1961	2.77	1979	3.02	1997	3.02
1962	2.78	1980	3.17	1998	3.05
1963	2.81	1981	3.29	1999	3.06
1964	2.91	1982	3.20	2000	3.05
1965	3.07	1983	3.21	2001	3.07
1966	2.91	1984	3.27	2002	3.12
1967	2.97	1985	3.33	2003	3.04
1968	2.98	1986	3.24	2004	3.04
1969	3.06	1987	3.30	2005	3.09
1970	2.90	1988	3.36	2006	3.09
1971	2.96	1989	3.30	2007	3.04
1972	2.96	1990	3.05	2008	3.04
1973	3.00	1991	3.07	2009	3.10
1974	2.92	1992	3.09	2010	3.06
1975	2.91	1993	3.05	2011	3.05
1976	2.82	1994	3.06	2012	3.06
1977	2.83	1995	3.04	2013	2.94
1978	3.04	1996	3.06	2014	2.88

3B Describing time series plots

▶ Looking for patterns in time series plots

The features we look for in a time series are:

- trend
- seasonality
- possible outliers
- cycles
- structural change
- irregular (random) fluctuations.

One or all of these features can be found in a time series plot.

Trend

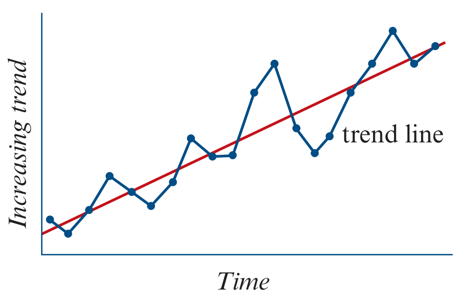
Examining a time series plot we can often see a general upward or downward movement over time. This indicates a long-term change over time that we call a trend.

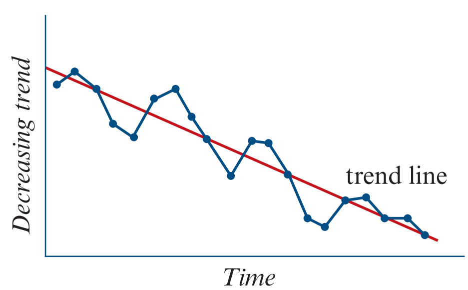
Trend

The tendency for values in a time series to generally increase or decrease over a significant period of time is called a **trend**.

One way of identifying trends on a time series graph is to draw a line that ignores the fluctuations, but which reflects the overall increasing or decreasing nature of the plot. These lines are called *trend lines*.

Trend lines have been drawn on the time series plots below to indicate an *increasing* trend (line slopes upwards) and a *decreasing* trend (line slopes downwards).

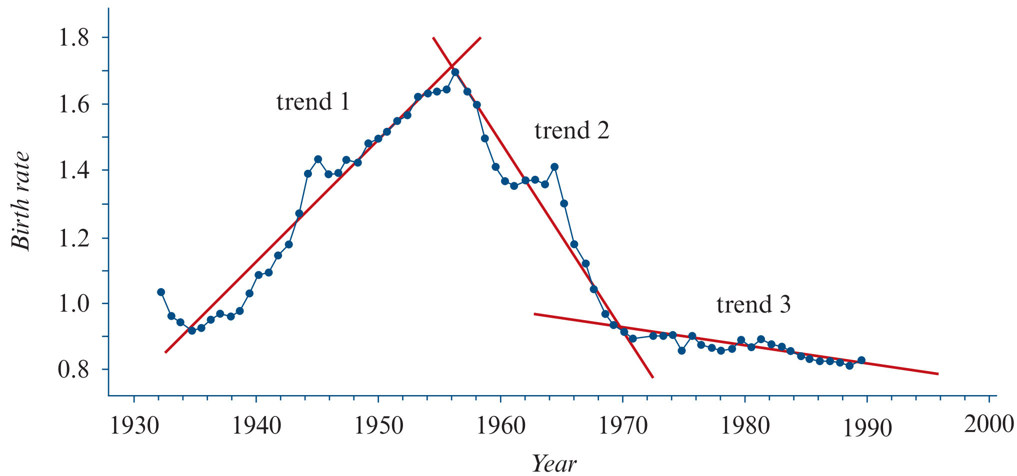




ISBN 978-1-108-46154-2

Cambridge University Press

Sometimes, different trends are apparent in a time series for different time periods. For example, in the time series plot of the birth rate data for Australia, shown below, there are three distinct trends, which can be seen by drawing trend lines on the plot.



Each of these trends can be explained by changing socioeconomic circumstances.

Trend 1: Between 1940 and 1961 the birth rate in Australia grew quite dramatically. Those in the armed services came home from the Second World War, and the economy grew quickly. This rapid increase in the Australian birth rate during this period is known as the 'Baby Boom'.

Trend 2: From about 1962 until 1980 the birth rate declined very rapidly. Birth control methods became more effective, and women started to think more about careers. This period is sometimes referred to as the 'Baby Bust'.

Trend 3: During the 1980s, and up until the early 2000s, the birth rate continues to decline slowly for a complex range of social and economic reasons.

ISBN 978-1-108-46154-2

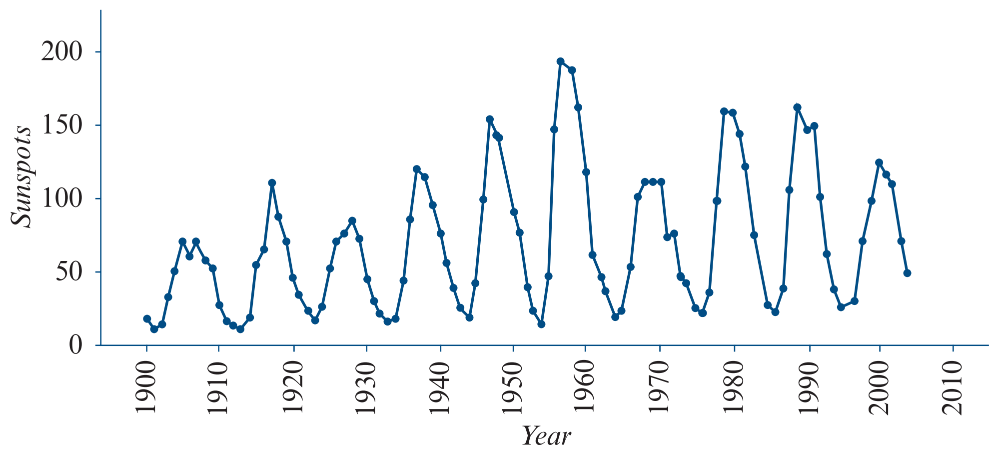
Cambridge University Press

Cycles

Cycles

Cycles are periodic movements in a time series, but over a period greater than 1 year.

Some cycles repeat regularly, and some do not. The following plot shows the sunspot[1] activity for the period 1900 to 2010. The period of this cycle is approximately 11 years.



Many business indicators, such as **interest rates** or unemployment figures, also vary in cycles, but their periods are usually less regular. Cycles with calendar-related periods of 1 year or less are of special interest and give rise to what is called 'seasonality'.

Seasonality

Seasonality

Seasonality is present when there is a periodic movement in a time series that has a calendar-related period, for example, a year, a month or a week.

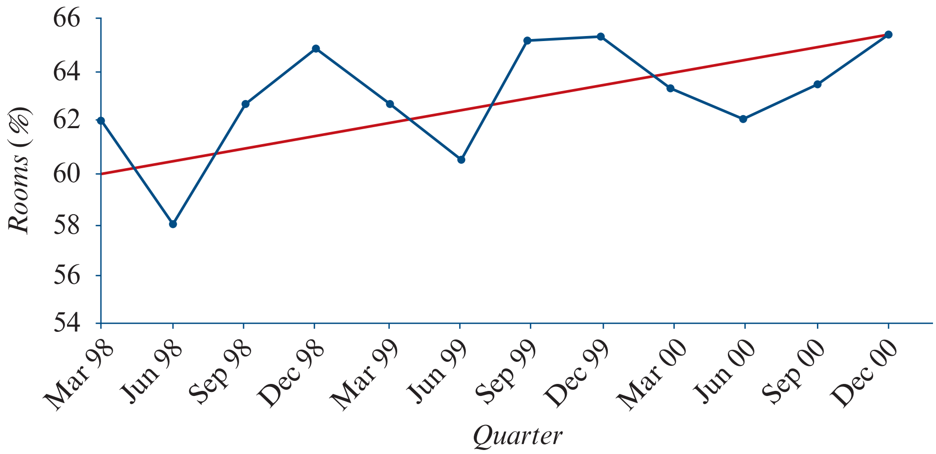
Seasonal movements tend to be more predictable than trends, and occur because of variations in the weather, such as ice-cream sales, or institutional factors, like the increase in the number of unemployed people at the end of the school year.

[1] Sunspots are dark spots visible on the surface of the Sun that come and go over time.

ISBN 978-1-108-46154-2

Cambridge University Press

The plot below shows the total percentage of rooms occupied in hotels, motels and other accommodation in Australia by quarter, over the years 1998–2000.



This time series plot reveals both *seasonality* and *trend* in the demand for accommodation. The *regular peaks and troughs* in the plot that occur at the *same time each year* signal the presence of *seasonality*. In this case, the demand for accommodation is at its lowest in the June quarter and highest in the December quarter.

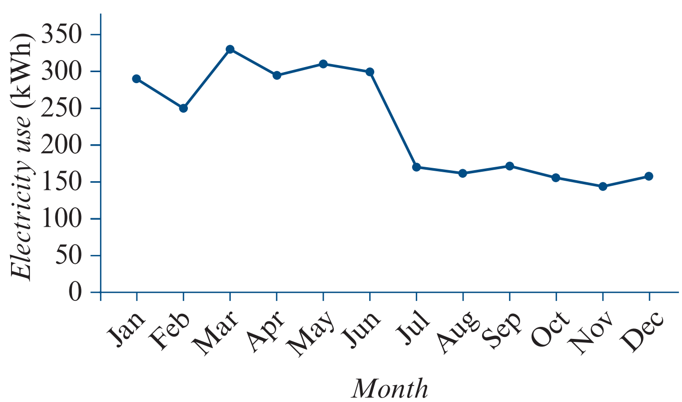
The *upward sloping trend line* signals the presence of a general increasing *trend*. This tells us that, even though demand for accommodation has fluctuated from month to month, demand for hotel and motel accommodation has increased over time.

Structural change

Structural change

Structural change is present when there is a sudden change in the established pattern of a time series plot.

The time series plot below shows the electricity bill for a rental house (in kWh) for the 12 months of a year.



The plot reveals an abrupt change in electricity use in June to July. During this period, monthly electricity use suddenly decreases from around 300 kWh per month from January to June

ISBN 978-1-108-46154-2

Cambridge University Press

to around 175 kWh for the rest of the year. This is an example of structural change that can probably be explained by a change in tenants, from a family with two children to a person living alone.

Structural change is also displayed in the birth rate time series plot we saw earlier. This revealed three quite distinct trends during the period 1900–2010. These reflect significant external events (like a war) or changes in social and economic circumstances.

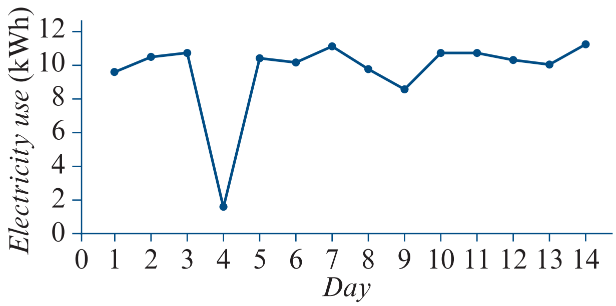
One consequence of structural change is that we can no longer use a single mathematical model to describe the key features of a time series plot.

Outliers

Outliers

Outliers are present when there are individual values that stand out from the general body of data.

The time series plot below shows the daily electricity bill for a house (in kWh) for a fortnight.



For this household, daily electricity use follows a regular pattern that, although fluctuating, averages about 10 kWh per day. In terms of daily electricity use, day 4 is a clear outlier, with less than 2 kWh of electricity used. A follow-up investigation found that, on this day, the house was without power for 18 hours due to a storm, so much less power was used than normal.

Irregular (random) fluctuations

Irregular (random) fluctuations

Irregular (random) fluctuations include all the variations in a time series that we cannot reasonably attribute to systematic changes like trend, cycles, seasonality and structural change or an outlier.

There can be many sources of irregular fluctuations, mostly unknown. A general characteristic of these fluctuations is that they are unpredictable.

One of the aims of time series analysis is to develop techniques to identify regular patterns in time series plots that are often obscured by irregular fluctuations. One of these techniques is smoothing, which you will meet in the next section.

ISBN 978-1-108-46154-2

Cambridge University Press

Identifying patterns in time series plots

The features we look for in a time series are:

- trend
- cycles
- seasonality
- structural change
- possible outliers
- irregular (random) fluctuations.

Trend is present when there is a *long-term* upward or downward movement in a time series.

Cycles are present when there is a periodic movement in a time series. The period is the time it takes for one complete up and down movement in the time series plot. In practice, this term is reserved for periods greater than 1 year.

Seasonality is present when there is a periodic movement in a time series that has a calendar related period, for example, a year, a month or a week.

Structural change is present when there is a sudden change in the established pattern of a time series plot.

Outliers are present when there are individual values that stand out from the general body of data.

Irregular (random) fluctuations are always present in any real-world time series plot. They include all the variations in a time series that we cannot reasonably attribute to systematic changes like trend, cycles, seasonality and structural change or an outlier.

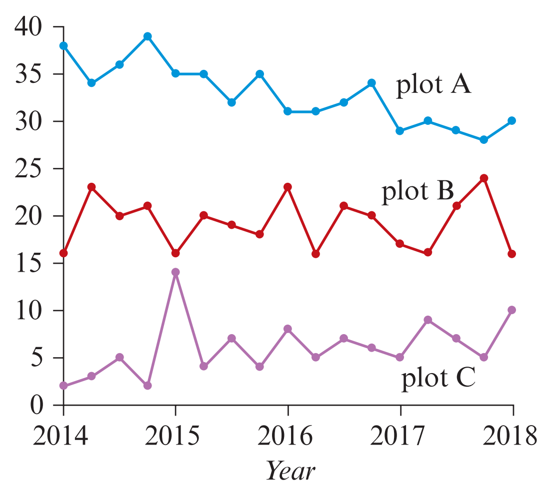
Exercise 3B

Note: You will need the time series plots constructed in Exercise 3A.

Identifying key features in a time series plot

1 Complete the table below by indicating which of the listed features are present in each of the time series plots. CF

Feature	Plot A	Plot B	Plot C
Irregular fluctuations			
Increasing trend			
Decreasing trend			
Cycles			
Outlier			

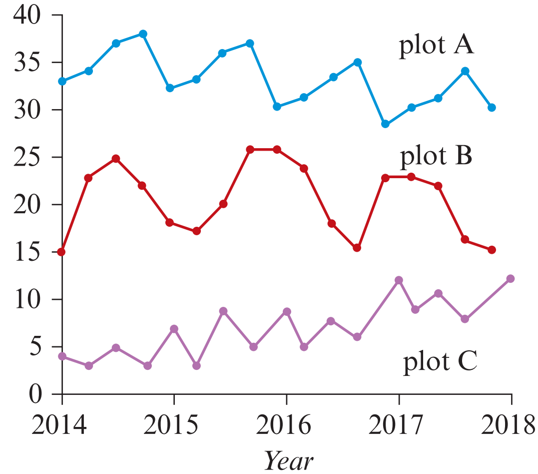


ISBN 978-1-108-46154-2

Cambridge University Press

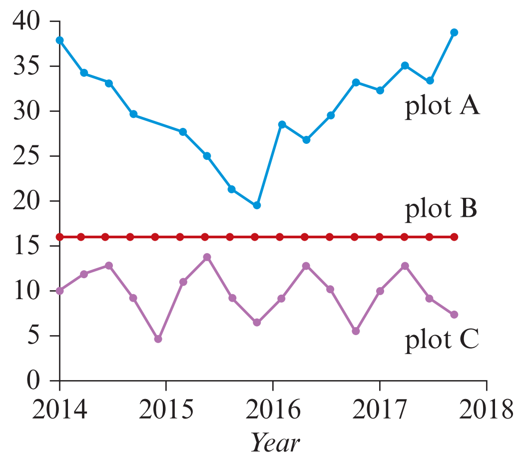
2 Complete the table below by indicating which of the listed features are present in each of the time series plots.

Feature	Plot A	Plot B	Plot C
Irregular fluctuations			
Increasing trend			
Decreasing trend			
Cycles			
Seasonality			



3 Complete the table below by indicating which of the listed features are present in each of the time series plots.

Feature	Plot A	Plot B	Plot C
Irregular fluctuations			
Structural change			
Increasing trend			
Decreasing trend			
Seasonality			

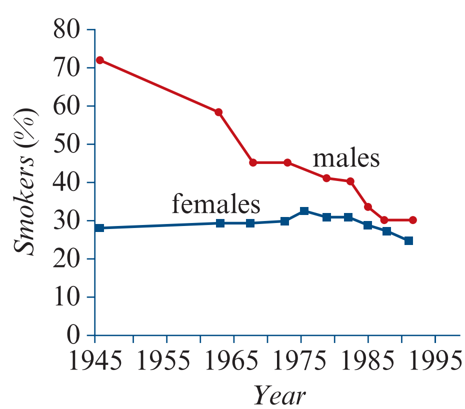


Describing time series plots

4 The time series plot shows the smoking rates (%) of Australian males and females over the period 1945–92.

a Describe the trends in the time series plot.

b Did the *difference* in smoking rates increase or decrease over the period 1945–92?



CU

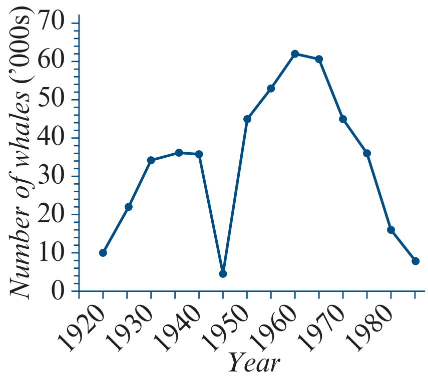
ISBN 978-1-108-46154-2

Cambridge University Press

5 The time series plot opposite shows the number of whales caught between 1920–85. Describe the features of the plot.

Note: This time series exhibits structural change so cannot be described by a single trend. Here is some relevant historical information:

- The 1930s was the time of the Great Depression.
- 1939–45 was the period of the Second World War.
- 1960–85 was a time when countries began to accept that whales were endangered.



6 Describe the features of the time series plot of the Australian population between 2006–2017 constructed in Exercise 3A Question 4.

7 Describe the features of the time series plot of the motor vehicle theft rate per 100 000 cars in Australia from 2003 to 2014 constructed in Exercise 3A Question 5.

8 Describe the features of the time series plot of the number of measles cases reported in Australia from 1988 to 2016 constructed in Exercise 3A Question 6.

9 Describe the features of the time series plot of carbon intensity (a measure of carbon emissions from coal) for Australia for the years from 1961 to 2014 constructed in Exercise 3A Question 7.

ISBN 978-1-108-46154-2

Cambridge University Press

3C Smoothing a time series using moving means

A time series plot can incorporate many of the sources of variation previously mentioned: trend, cycles, seasonality, structural change, outliers and irregular fluctuations. One effect of the irregular fluctuations and seasonality can be to obscure an underlying trend. The technique of **smoothing** can sometimes be used to overcome this problem.

Smoothing a time series plot using moving means

This method of smoothing (**moving mean smoothing**) involves replacing individual data points in the time series with their moving means. The simplest method is to smooth over a small number of odd number points, for example three or five.

The three-moving mean

To use *three-moving mean smoothing*, replace each data value with the mean of that value and the values of its two neighbours, one on each side. That is, if y_1, y_2 and y_3 are sequential data values, then:

$$\text{smoothed}\, y_2 = \frac{y_1 + y_2 + y_3}{3}$$

The first and last points do not have values on each side, so leave them out.

The five-moving mean

To use *five-moving mean smoothing*, replace each data value with the mean of that value and the two values on each side. That is, if y_1, y_2, y_3, y_4, y_5 are sequential data values, then:

$$\text{smoothed}\, y_3 = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5}$$

The first two and last two points do not have two values on each side, so leave them out.

If needed, these definitions can be readily extended for moving means involving 7, 9, 11,… points. The larger the number of points we smooth over, the greater the smoothing effect.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 3 Three- and five-moving mean smoothing

The table below gives the temperature (°C) recorded at a weather station at 9 a.m. each day for a week.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Temperature	18.1	24.8	26.4	13.9	12.7	14.2	24.9

a Calculate the three-mean smoothed temperature for Tuesday.

b Calculate the five-mean smoothed temperature for Thursday.

Solution

a **1** Write down the three temperatures centred on Tuesday.

18.1, 24.8, 26.4

$$\text{Mean} = \frac{(18.1 + 24.8 + 26.4)}{3} = 23.1$$

2 Find the mean and write down your answer.

The three-mean smoothed temperature for Tuesday is 23.1°C.

b **1** Write down the five temperatures centred on Thursday.

24.8, 26.4, 13.9, 12.7, 14.2

2 Find their mean and write down your answer.

$$\text{Mean} = \frac{(24.8 + 26.4 + 13.9 + 12.7 + 14.2)}{5} = 18.4$$

The five-mean smoothed temperature for Thursday is 18.4°C.

The next step is to extend these computations to smooth all terms in the time series.

ISBN 978-1-108-46154-2

Example 4 **Three- and five-moving mean smoothing of a time series**

The following table gives the number of births per month over a calendar year in a country hospital. Use the three-moving mean and the five-moving mean methods, correct to one decimal place, to complete the table.

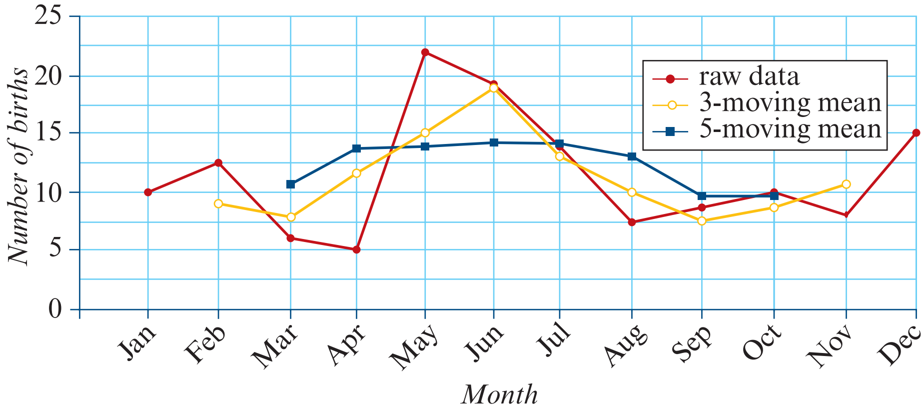
Solution

Complete the calculations as shown below.

Month	Number of births	3-moving mean	5-moving mean
January	10		
February	12	$\frac{10+12+6}{3} = 9.3$	
March	6	$\frac{12+6+5}{3} = 7.7$	$\frac{10+12+6+5+22}{5} = 11.0$
April	5	$\frac{6+5+22}{3} = 11.0$	$\frac{12+6+5+22+18}{5} = 12.6$
May	22	$\frac{5+22+18}{3} = 15.0$	$\frac{6+5+22+18+13}{5} = 12.8$
June	18	$\frac{22+18+13}{3} = 17.7$	$\frac{5+22+18+13+7}{5} = 13.0$
July	13	$\frac{18+13+7}{3} = 12.7$	$\frac{22+18+13+7+9}{5} = 13.8$
August	7	$\frac{13+7+9}{3} = 9.7$	$\frac{18+13+7+9+10}{5} = 11.4$
September	9	$\frac{7+9+10}{3} = 8.7$	$\frac{13+7+9+10+8}{5} = 9.4$
October	10	$\frac{9+10+8}{3} = 9.0$	$\frac{7+9+10+8+15}{5} = 9.8$
November	8	$\frac{10+8+15}{3} = 11.0$	
December	15		

ISBN 978-1-108-46154-2

The result of this smoothing can be seen in the plot below, which shows the raw data, the data smoothed with a three-moving means and the data smoothed with a five-moving means.



Note: In the process of smoothing, data points are lost at the beginning and end of the time series.

Two observations can be made from this plot:

1. five-mean smoothing is more effective in reducing the irregular fluctuations than three-mean smoothing
2. the five-mean smoothed plot shows that there is no clear trend although the raw data suggest that there might be an increasing trend.

There are many ways of smoothing a time series. Moving means of group size other than three and five are common and often very useful.

However, if we smooth over an even number of points, we run into a problem. The centre of the set of points is not at a time point belonging to the original series. Usually, we solve this problem by using a process called **centring**. Centring involves taking a two-moving mean of the already smoothed values so that they line up with the original time values. It is a two-step process.

Two-mean smoothing with centring

We will illustrate the process by finding the two-moving mean, centred on Tuesday, for the daily temperature data opposite.

Day	Temperature
Monday	18.1
Tuesday	24.8
Wednesday	26.4

It is straightforward to calculate a series of two-moving means for this data by calculating the mean for Monday and Tuesday, followed by the mean for Tuesday and Wednesday. However, as we can see in the diagram overpage, these means do not align with a particular day, but lie between days. We solve this problem by finding the average of these two means. This gives a smoothed value that is now centred on Tuesday.

ISBN 978-1-108-46154-2

Cambridge University Press

We call this process two-mean smoothing with centring.

Day	Temperature	Two-moving means	Two-moving mean with centring
Monday	18.1		
		$\frac{(18.1 + 24.8)}{2} = 21.45$	
			$\frac{(21.45 + 25.6)}{2} = 23.525$
Tuesday	24.8		
		$\frac{(24.8 + 26.4)}{2} = 25.60$	
Wednesday	26.4		

In practice, we do not have to draw such a diagram to perform these calculations. The purpose of doing so is to show how the centring process works. In practice, calculating two-moving means is a much briefer and routine process as we illustrate in the following example. However, before proceeding, you might find it useful to view the video for this topic.

Example 5 Two-moving mean smoothing with centring

The temperatures (°C) recorded at a weather station at 9 a.m. each day for a week are displayed in the table.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Temperature	18.1	24.8	26.4	13.9	12.7	14.2	24.9

Calculate the two-mean smoothed temperature for Tuesday with centring.

Solution

1 For two-mean smoothing with centring, write down the *three* data values centred on Tuesday (highlighted in red).

18.1 24.8 26.4

2 Calculate the mean of the first two values (mean 1). Calculate the mean of the second two values (mean 2).

$$\text{Mean 1} = \frac{(18.1 + 24.8)}{2}$$

$$\text{Mean 2} = \frac{(24.8 + 26.4)}{2}$$

3 The centred mean is then the average of mean 1 and mean 2.

$$\text{Centred mean} = \frac{(\text{mean 1} + \text{mean 2})}{2} = \frac{(21.45 + 25.6)}{2} = 23.525$$

4 Write down your answer.

The two-mean smoothed temperature, centred on Tuesday, is 23.5°C (to one decimal place).

ISBN 978-1-108-46154-2

Cambridge University Press

The process of four-mean smoothing with centring is the same as two-mean smoothing except that you smooth values in groups of four.

Example 6 Four-moving mean smoothing with centring

The table below gives the temperature (°C) recorded at a weather station at 9 a.m. each day for a week.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Temperature	18.1	24.8	26.4	13.9	12.7	14.2	24.9

Calculate the four-mean smoothed temperature with centring for Thursday.

Solution

1 For four-mean smoothing with centring, write down the *five* data values centred on Thursday.

24.8 26.4 13.9 12.7 14.2

2 Calculate the mean of the first four values (mean 1) and the mean of the last four values (mean 2).

$$\text{Mean 1} = \frac{(24.8 + 26.4 + 13.9 + 12.7)}{4} = 19.45$$

$$\text{Mean 2} = \frac{(26.4 + 13.9 + 12.7 + 14.2)}{4} = 16.8$$

3 The centred mean is then the average of mean 1 and mean 2.

$$\text{Centred mean} = \frac{(\text{mean 1} + \text{mean 2})}{2} = \frac{(19.45 + 16.8)}{2} = 18.125$$

4 Write down your answer.

The four-mean smoothed temperature centred on Thursday is 18.1°C (to one decimal place).

The next step is to extend these computations to smooth all terms in the time series.

Example 7 Smoothing a time series using Excel

Consider again the data in Example 3:

Day	Mon	Tues	Wed	Thur	Fri	Sat	Sun
Temperature (°C)	18.1	24.8	26.4	13.9	12.7	14.2	24.9

Use Excel to construct a plot showing the data, the three-mean smoothed plot, and the five-mean smoothed plot, all on the same axes.

ISBN 978-1-108-46154-2

Cambridge University Press

Solution

Spreadsheet

1 Enter the data into two columns. Place the cursor in cell D4, next to the temperature on Tuesday and enter the formula =(**C3**+**C4**+**C5**)/**3** as shown below. Press Enter and the mean of these three cells will be calculated.

A	B	C	D
	Day	Temperature	3-moving mean
	Mon	18.1	
	Tues	24.8	=(C3+C4+C5)/3
	Wed	26.4	
	Thur	13.9	
	Fri	12.7	
	Sat	14.2	
	Sun	24.9	

2 **Fill down** the column with the formula, remembering to omit the first and last row, to calculate all three-mean smoothed values.

3 Place the cursor in cell E5, next to the temperature on Wednesday and enter the formula =(**C3**+**C4**+**C5**+**C6**+**C7**)/**5** as shown below. Press Enter and the mean of these five cells will be calculated.

Day	Temperature	3-moving mean	5-moving mean
Mon	18.1		
Tues	24.8	23.1	
Wed	26.4		=(C3+C4+C5+C6+C7)/5
Thur	13.9	17.7	
Fri	12.7	13.6	
Sat	14.2	17.3	
Sun	24.9		

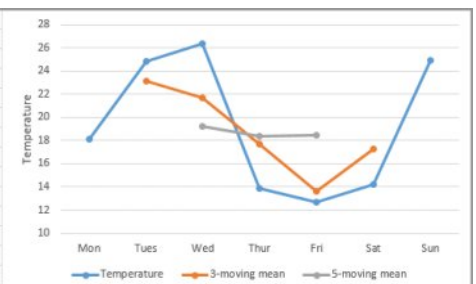
4 **Fill down** the column with the formula, remembering to omit the first and last two rows, to calculate all five-mean smoothed values (see below).

5 Select all four data columns (including headings) and on the **Insert** tab, in the **Charts** group, click on **2D-Line** and then the option **Lines with Markers**.

6 Double click on each scale separately and edit as appropriate.

7 Axis labels can be added using the Add Chart Elements option when editing the scale.

Day	Temperature	3-moving mean	5-moving mean
Mon	18.1		
Tues	24.8	23.1	
Wed	26.4	21.7	19.18
Thur	13.9	17.7	18.4
Fri	12.7	13.6	18.42
Sat	14.2	17.3	
Sun	24.9		



ISBN 978-1-108-46154-2

Exercise 3C

Basic skills

Use the information below to answer Questions 1 to 5.

The table below gives the temperature (°C) recorded at a weather station at 3 p.m. each day for a week.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Temperature (°C)	28.9	33.5	21.6	18.1	16.2	17.9	26.4

Example 3

SF

1 The three-mean smoothed temperature for Wednesday is closest to:
A 20.0 **B** 23.2 **C** 24.4 **D** 29.4 **E** 31.2

2 The five-mean smoothed temperature for Friday is closest to:
A 20.0 **B** 23.2 **C** 24.4 **D** 29.4 **E** 31.2

3 The seven-mean smoothed temperature for Thursday is closest to:
A 20.0 **B** 23.2 **C** 24.4 **D** 28.0 **E** 31.2

4 The two-mean smoothed temperature with centring for Tuesday is closest to:
A 19.1 **B** 20.0 **C** 24.4 **D** 29.4 **E** 31.2

5 The four-mean smoothed temperature for Friday is closest to:
A 19.1 **B** 20.0 **C** 23.2 **D** 28.0 **E** 31.2

Calculating the smoothed value of individual data points

Example 4

6

t	1	2	3	4	5	6	7	8	9
y	5	2	5	3	1	0	2	3	0

For the time series data in the table above, find:

a the three-mean smoothed y-value for $t = 4$
b the three-mean smoothed y-value for $t = 6$
c the three-mean smoothed y-value for $t = 2$
d the five-mean smoothed y-value for $t = 3$
e the five-mean smoothed y-value for $t = 7$
f the five-mean smoothed y-value for $t = 4$
g the two-mean smoothed y-value centred at $t = 3$
h the two-mean smoothed y-value centred at $t = 8$
i the four-mean smoothed y-value centred at $t = 3$
j the four-mean smoothed y-value centred at $t = 6$.

Note: Copies of the tables in Questions 7 to 11 can be accessed via the skillsheet icon in the Interactive Textbook.

ISBN 978-1-108-46154-2

Smoothing a table of values

Spreadsheet

7 Complete the following table.

t	1	2	3	4	5	6	7	8	9
y	10	12	8	4	12	8	10	18	2
Three-mean smoothed y	–								–
Five-mean smoothed y	–	–						–	–

Smoothing and plotting a time series (three- and five-mean smoothing)

8 The maximum temperature of a city over a period of 10 days is given below.

Day	1	2	3	4	5	6	7	8	9	10
Temperature (°C)	24	27	28	40	22	23	22	21	25	26
Three-moving mean										
Five-moving mean										

a Construct a time series plot of the temperature data.
b Use the three-mean and five-mean smoothing method to complete the table.
c Plot the smoothed temperature data and compare and comment on the plots. This is best done if all plots are on the same graph.

9 The value of the Australian dollar in US dollars (exchange rate) over 10 days is given below.

Day	1	2	3	4	5	6	7	8	9	10
Exchange rate	0.743	0.754	0.737	0.751	0.724	0.724	0.712	0.735	0.716	0.711
Three-moving mean										
Five-moving mean										

a Construct a time series plot of the data. Label and scale the axes.
b Use the three-mean and five-mean smoothing method to complete the table.
c Plot the smoothed exchange rate data and compare the plots and comment on the plots. This is best done if all three plots are on the same graph.

Smoothing a time series (two- and four-mean smoothing)

Example 5

10 Construct a table with four columns: 'Month', 'Number of births', 'Two-moving mean' and 'Two-moving mean with centring' using the following data.

Month	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
Number of births	10	12	6	5	22	18	13	7	9	10	8	15

Example 6

11 Construct a table with four columns: 'Month', 'Internet usage', 'Four-moving mean' and 'Four-moving mean with centring' using the following data.

Month	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
Internet usage	21	40	52	42	58	79	81	54	50

CF

ISBN 978-1-108-46154-2

Cambridge University Press

3D Seasonal indices

When the data is seasonal, it is often necessary to **deseasonalise** the data before further analysis. To do this we need to calculate seasonal indices.

The concept of a seasonal index

Consider the (hypothetical) monthly seasonal indices for unemployment given in the table.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Total
1.1	1.2	1.1	1.0	0.95	0.95	0.9	0.9	0.85	0.85	1.1	1.1	12.0

Key fact 1

Seasonal indices are calculated so that their *average* is 1. This means that the *sum* of the seasonal indices equals the *number of seasons*.

Thus, if the seasons are months, the seasonal indices add to 12. If the seasons are quarters, then the seasonal indices would to 4, and so on.

Key fact 2

Seasonal indices tell us how a particular season (generally a day, month or quarter) compares to the *average season*.

For example:

- seasonal index for unemployment for the month of February is 1.2 or 120%.

  This tells us that February unemployment figures tend to be 20% *higher* than the monthly average. Remember, the average seasonal index is 1 or 100%.

- seasonal index for August is 0.90 or 90%.

  This tells us that the August unemployment figures tend to be only 90% of the monthly average. Alternatively, August unemployment figures are 10% *lower* than the monthly average.

We can use seasonal indices to remove the seasonal component (deseasonalise) from a time series, or to put it back in (**reseasonalise**).

ISBN 978-1-108-46154-2

Cambridge University Press

Using seasonal indices to deseasonalise or reseasonalise a time series

To calculate deseasonalised figures, each entry is divided by its seasonal index as follows.

Deseasonalising data

Time series data are deseasonalised using the relationship:

$$\text{deseasonalised figure} = \frac{\text{actual figure}}{\text{seasonal index}}$$

Example 8 Using a seasonal index to deseasonalise data

The seasonal index (SI) for cold drink sales for summer is SI = 1.33.

Last summer a beach kiosk's actual cold drink sales totalled $15 653.

What were the *deseasonalised* sales?

Solution

Use the rule

$$\text{deseasonalised sales} = \frac{\text{actual sales}}{\text{seasonal index}}$$

with actual sales = $15 653 and SI = 1.33.

$$\text{Deseasonalised sales} = \frac{15\,653}{1.33} = 11\,769.17$$

The deseasonalised sales for summer were $11 769.17.

The rule for determining deseasonalised data values can also be used to reseasonalise data; that is, convert a deseasonalised value into an actual data value.

Reseasonalising data

Time series data are reseasonalised using the rule:

$$\text{actual figure} = \text{deseasonalised figure} \times \text{seasonal index}$$

Example 9 Using a seasonal index to reseasonalise data

The seasonal index for cold drink sales in spring is SI = 0.85.

Last spring a beach kiosk's deseasonalised cold drink sales totalled $10 870.

What were the *actual* sales?

Solution

Use the rule

actual sales = deseasonalised sales × seasonal index
with deseasonalised sales = $10 870 and SI = 0.85.

$$\text{Actual sales} = 10\,870 \times 0.85 = 9239.50$$

The actual sales for spring were $9239.50.

ISBN 978-1-108-46154-2

Example 10 Deseasonalising a time series

The quarterly sales figures for Mikki's shop over a 3-year period are given below.

Year	Summer	Autumn	Winter	Spring
1	920	1085	1241	446
2	1035	1180	1356	541
3	1299	1324	1450	659

Use the seasonal indices shown to deseasonalise these sales figures. Write answers correct to the nearest whole number.

Summer	Autumn	Winter	Spring
1.03	1.15	1.30	0.52

Solution

1 To deseasonalise each sales figure in the table, divide by the appropriate seasonal index.
For example, for summer, divide the figures in the 'Summer' column by 1.03.
Round results to the nearest whole number.

$$\frac{920}{1.03} = 893$$

$$\frac{1035}{1.03} = 1005$$

$$\frac{1299}{1.03} = 1261$$

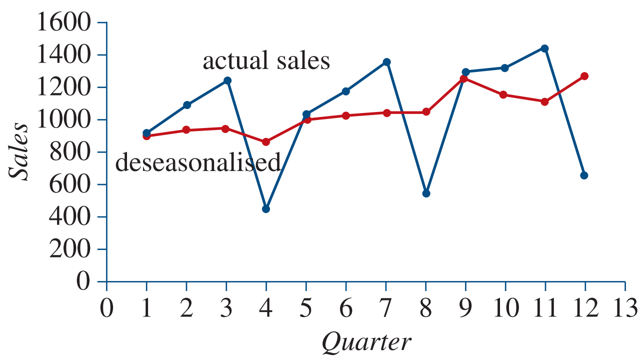
2 Repeat for the other seasons.

Deseasonalised sales figures

Year	Summer	Autumn	Winter	Spring
1	893	943	955	858
2	1005	1026	1043	1040
3	1261	1151	1115	1267

▶ Comparing a plot of the raw data with the deseasonalised data

The plot below shows the time series deseasonalised sales.



ISBN 978-1-108-46154-2

Two things to notice are that deseasonalising has:

- removed the seasonality from the time series plot
- revealed a clear underlying trend in the data.

It is common to deseasonalise time series data before you fit a trend line.

Calculating seasonal indices

To complete this section, you will now learn to calculate a seasonal index. We will start by using only 1 year's data to illustrate the basic ideas and then move onto a more realistic example where several years' data are involved.

Example 11 Calculating seasonal indices (1 year's data)

Mikki runs a shop and wishes to determine quarterly seasonal indices based on last year's sales (shown in table opposite).

Summer	Autumn	Winter	Spring
920	1085	1241	446

Solution

1 The seasons are quarters. Write the formula in terms of quarters.

$$\text{Seasonal index} = \frac{\text{value for season}}{\text{seasonal average}}$$

2 Find the quarterly average for the year.

$$\text{Quarterly average} = \frac{920 + 1085 + 1241 + 446}{4} = 923$$

3 Work out the seasonal index (SI) for each time period.

$$SI_{Summer} = \frac{920}{923} = 0.997$$

$$SI_{Autumn} = \frac{1085}{923} = 1.176$$

$$SI_{Winter} = \frac{1241}{923} = 1.345$$

$$SI_{Spring} = \frac{446}{923} = 0.483$$

4 Check that the seasonal indices sum to 4 (the number of seasons). The slight difference is due to rounding error.

Check: $0.997 + 1.176 + 1.345 + 0.483 = 4.001$

5 Write out your answers as a table of the seasonal indices.

Seasonal indices

Summer	Autumn	Winter	Spring
0.997	1.176	1.345	0.483

ISBN 978-1-108-46154-2

The next example illustrates how seasonal indices are calculated with 3 years' data. While the process looks more complicated, we just repeat what we did in Example 11 three times and average the results for each year at the end.

Example 12 Calculating seasonal indices (several years' data)

Suppose that Mikki has 3 years of data, as shown. Use the data to calculate seasonal indices, correct to two decimal places.

Year	Summer	Autumn	Winter	Spring
1	920	1085	1241	446
2	1035	1180	1356	541
3	1299	1324	1450	659

Solution

The strategy is as follows:

- Calculate the seasonal indices for years 1, 2 and 3 separately, as Example 11 (as we already have the seasonal indices for year 1 from Example 11, we will save ourselves some time by simply quoting the result).
- Average the three sets of seasonal indices to obtain a single set of seasonal indices.

1 Write down the result for year 1.

Year 1 seasonal indices:

Summer	Autumn	Winter	Spring
0.997	1.176	1.345	0.483

2 Now calculate the seasonal indices for year 2.

a The seasons are quarters. Write the formula in terms of quarters.

$$\text{Seasonal index} = \frac{\text{value for quarter}}{\text{quarterly average}}$$

b Find the quarterly average for the year.

$$\text{Quarterly average} = \frac{1035 + 1180 + 1356 + 541}{4}$$
$$= 1028$$

c Work out the seasonal index (SI) for each time period.

$$SI_{Summer} = \frac{1035}{1028} = 1.007$$

$$SI_{Autumn} = \frac{1180}{1028} = 1.148$$

$$SI_{Winter} = \frac{1356}{1028} = 1.319$$

$$SI_{Spring} = \frac{541}{1028} = 0.526$$

ISBN 978-1-108-46154-2

Cambridge University Press

d Check that the seasonal indices sum to 4.

Check: 1.007 + 1.148 + 1.319 + 0.526 = 4.000

e Write out your answers as a table of the seasonal indices.

Year 2 seasonal indices:

Summer	Autumn	Winter	Spring
1.007	1.148	1.319	0.526

3 Now calculate the seasonal indices for year 3.

a Find the quarterly average for the year.

$$\text{Quarterly average} = \frac{1299 + 1324 + 1450 + 659}{4} = 1183$$

b Work out the seasonal index (SI) for each time period.

$$SI_{Summer} = \frac{1299}{1183} = 1.098$$

$$SI_{Autumn} = \frac{1324}{1183} = 1.119$$

$$SI_{Winter} = \frac{1450}{1183} = 1.226$$

$$SI_{Spring} = \frac{659}{1183} = 0.557$$

c Check that the seasonal indices sum to 4.

Check: 1.098 + 1.119 + 1.226 + 0.557 = 4.000

d Write out your answers as a table of the seasonal indices.

Year 3 seasonal indices:

Summer	Autumn	Winter	Spring
1.098	1.119	1.226	0.557

4 Find the 3-year averaged seasonal indices by averaging the seasonal indices for each season.

Final seasonal indices:

$$SI_{Summer} = \frac{0.997 + 1.007 + 1.098}{3} = 1.03$$

$$SI_{Autumn} = \frac{1.176 + 1.148 + 1.119}{3} = 1.15$$

$$SI_{Winter} = \frac{1.345 + 1.319 + 1.226}{3} = 1.30$$

$$SI_{Spring} = \frac{0.483 + 0.526 + 0.557}{3} = 0.52$$

5 Check that the seasonal indices sum to 4.

Check: 1.03 + 1.15 + 1.30 + 0.52 = 4.00

6 Write out your answers as a table of the seasonal indices.

Summer	Autumn	Winter	Spring
1.03	1.15	1.30	0.52

ISBN 978-1-108-46154-2

Example 13 Using Excel to calculate seasonal indices

Using the information in Example 12 repeat the calculation of seasonal indices for Mikki's shop using Excel.

Solution

1 Enter the data table as given in the question into Excel.

2 In cell G3 enter the formula =**(C3+D3+E3+F3)/4** to calculate the quarterly average for the first year. Press enter then **Fill down** to complete all quarterly averages.

Spreadsheet

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	=(C3+D3+E3+F3)/4
4		2	1035	1180	1356	541	
5		3	1299	1324	1450	659	
6							

3 Now determine the seasonal indices for year 1, by entering the formulas as follows:

C7: =**C3/G3**
D7: =**D3/G3**
E7: =**E3/G3**
F7: =**F3/G3**

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			=C3/G3				
8							

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			0.99675	1.175515	1.344529	=F3/G3	
8							

ISBN 978-1-108-46154-2

4 Select cells C7, D7, E7, and F7 and **Fill down** two further rows, to determine the seasonal indices for each of the three years.

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			0.99675	1.175515	1.344529	0.483207	
8			1.006809	1.14786	1.319066	0.526265	
9			1.098056	1.119189	1.225697	0.557058	
10							

5 Finally, determine the average of the three years seasonal indices by entering into C10 the formula =(**C7+C8+C9**)**/3** as shown.

A	B	C	D	E	F	G
	Year	Summer	Autumn	Winter	Spring	Quarterly average
	1	920	1085	1241	446	923
	2	1035	1180	1356	541	1028
	3	1299	1324	1450	659	1183
		0.99675	1.175515	1.344529	0.483207	
		1.006809	1.14786	1.319066	0.526265	
		1.098056	1.119189	1.225697	0.557058	
		=(C7+C8+C9)/3				

6 Select cell C10 and then **Fill right** to complete the seasonal indice calculations. Round to two decimal places.

A	B	C	D	E	F	G
	Year	Summer	Autumn	Winter	Spring	Quarterly average
	1	920	1085	1241	446	923
	2	1035	1180	1356	541	1028
	3	1299	1324	1450	659	1183
		0.99675	1.175515	1.344529	0.483207	
		1.006809	1.14786	1.319066	0.526265	
		1.098056	1.119189	1.225697	0.557058	
		1.03	**1.15**	**1.30**	**0.52**	

ISBN 978-1-108-46154-2

Cambridge University Press

Interpreting the seasonal indices

Having calculated these seasonal indices, what do they tell us?

The seasonal index of:

- 1.03 for summer tells us that summer sales are typically 3% above average
- 1.15 for autumn tells us that autumn sales are typically 15% above average
- 1.30 for winter tells us that winter sales are typically 30% above average
- 0.52 for spring tells us that spring sales are typically 48% below average.

Correcting for seasonality

Also, using the rule

$$\text{deseasonalised figure} = \frac{\text{actual figure}}{\text{seasonal index}}$$

we can work out how much we need to increase or decrease the actual sales figures to correct for seasonality.

For example, we see that for winter:

$$\begin{aligned}\text{deseasonalised figure} &= \frac{\text{actual figure}}{1.30} \\ &= 0.769\ldots \times \text{actual figure} \approx 77\% \text{ of the actual figures}\end{aligned}$$

Thus, to correct the seasonality in winter, we need to decrease the actual sales by about 23%.

Similarly we can show that, to correct for seasonality in spring ($SI_{spring} = 0.52$), we need to increase the actual spring sales figure by around 92% $\left(\frac{1}{0.52} \approx 1.92\right)$.

ISBN 978-1-108-46154-2

Exercise 3D

Skillsheet Basic skills and interpretation

Use the following information to answer Questions 1 to 8.

The table below shows the monthly sales figures (in $’000s) and seasonal indices (for January to November) for a product produced by the U-beaut company.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Seasonal index	1.2	1.3	1.1	1.0	1.0	0.9	0.8	0.7	0.9	1.0	1.1	
Sales ($'000s)	9.6	10.5	8.6		7.1	6.0	5.4		6.4	7.2	8.3	7.4

SF

1 The seasonal index for December is:
A 0.8 **B** 0.9 **C** 1.0 **D** 1.1 **E** 1.2

Example 8

2 The deseasonalised sales (in $’000s) for March is closest to:
A 7.7 **B** 7.8 **C** 8.6 **D** 9.5 **E** 10.3

3 The deseasonalised sales (in $’000s) for June is closest to:
A 5.4 **B** 5.9 **C** 6.0 **D** 6.6 **E** 6.7

Example 9

4 The deseasonalised sales (in $’000s) for August are 5.6. The actual sales are closest to:
A 2.7 **B** 3.9 **C** 5.6 **D** 5.9 **E** 7.3

5 The deseasonalised sales (in $’000s) for April are 6.9. The actual sales are closest to:
A 5.4 **B** 6.3 **C** 6.9 **D** 7.6 **E** 8.3

CF

6 The seasonal index for February tells us that, over time, February sales tend to be greater than the average monthly sales by:
A 0% **B** 10% **C** 20% **D** 30% **E** 70%

7 The seasonal index for September tells us that September sales tend to be less than the average monthly sales by:
A 0% **B** 10% **C** 20% **D** 30% **E** 90%

8 The seasonal index for January is 1.2. To correct the monthly sales figure for seasonality we need to:
A decrease the actual sales figures by around 20%
B increase the actual sales figures by around 20%
C decrease the actual sales figures by around 17%
D increase the actual sales figures by around 17%
E increase the actual sales figures by around 80%

ISBN 978-1-108-46154-2

Cambridge University Press

Use the following information to answer Questions 9 to 12.

The table below shows the quarterly newspaper sales (in $'000s) of a corner store. Also shown are the seasonal indices for newspaper sales in the first, second and third quarters.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Sales		1060	1868	1642
Seasonal index	0.8	0.7	1.3	

9 The seasonal index for quarter 4 is:
A 0.8 **B** 0.7 **C** 1.0 **D** 1.2 **E** 1.3

10 The deseasonalised sales (in $'000s) for quarter 2 are closest to:
A 742 **B** 980 **C** 1060 **D** 1514 **E** 1694

11 The deseasonalised sales (in $'000s) for quarter 3 are closest to:
A 1437 **B** 1678 **C** 1868 **D** 2428 **E** 2567

12 The deseasonalised sales (in $'000s) for quarter 1 are 1256. The actual sales are closest to:
A 986 **B** 1005 **C** 1256 **D** 1570 **E** 1678

SF

Deseasonalising a time series

Example 10

13 The following table shows the number of students enrolled in a 3-month computer systems training course along with seasonal indices that have been calculated from the previous year's enrolment figures. Complete the table.

CF

	Summer	Autumn	Winter	Spring
Number of students	56	125	126	96
Deseasonalised numbers				
Seasonal index	0.5	1.0	1.3	

ISBN 978-1-108-46154-2

Cambridge University Press

SF

14 The number of waiters employed by a restaurant chain in each quarter, along with some seasonal indices, are given in the following table.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Number of waiters	198	145	86	168
Seasonal index	1.30		0.58	1.10

a What is the seasonal index from the second quarter?

b The seasonal index for quarter 1 is 1.30. Explain what this means in terms of the average quarterly number of waiters.

c Deseasonalise the data.

Calculating seasonal indices

15 The table below records quarterly sales (in $'000s) for a shop.

Quarter 1	Quarter 2	Quarter 3	Quarter 4
60	56	75	78

Use the data determine the seasonal indices for the four quarters. Give your results correct to two decimal places. Check that your seasonal indices add to 4.

Example 11

16 The table below records the monthly sales (in $'000s) for a shop.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
12	13	14	17	18	15	9	10	8	11	15	20

Use the data to determine the seasonal indices for the 12 months. Give your results correct to two decimal places. Check that your seasonal indices add to 12.

17 The table below records the monthly sales (in $'000s)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
22	19	25	23	20	18	20	15	14	11	23	30

Use the data to determine the seasonal indices for the 12 months. Give your results correct to two decimal places. Check that your seasonal indices add to 12.

ISBN 978-1-108-46154-2

3E Fitting a trend line and forecasting

Fitting a trend line

If there appears to be a linear trend, we can use the least squares method to fit a line to the data to model the trend.

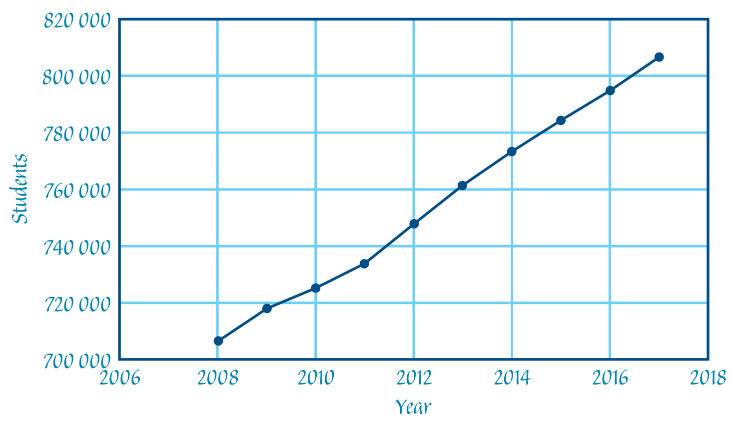
Example 14 Fitting a trend line

Fit a trend line to the data in the following table, which shows the total number of school students in Queensland over the years 2008–2017. Interpret the slope.

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Students	706462	717988	724956	733652	747682	761411	773309	784224	794815	806555

Solution

1 Construct a time series plot of the data to confirm that the trend is linear.



2 Fit a least squares regression trend line to the data with Year as the EV. Write down its equation.

intercept = −22 024 760

slope = 11 319

Number of students = −22 024 760 + 11 319 × year

3 Interpret the slope.

The number of students at school in Queensland increased on average by 11 319 students per year.

ISBN 978-1-108-46154-2

Cambridge University Press

Forecasting

Using a trend line fitted to a time series plot to make predictions about future values is known as **trend line forecasting**.

Example 15 Using a trend line to forecast a future value

Use the data and least squares regression trend line from Example 14. How many students do we predict will be attending school in Queensland in 2025 if the same increasing trend continues? Give your answer correct to the nearest 10 000 students.

Solution

Substitute the appropriate value for year in the equation determined using the least squares regression trend line.

$$\begin{aligned}\text{Number of students} &= -22\,024\,760 + 11\,319 \times 2025 \\ &= 896\,215 \\ &\cong 900\,000\end{aligned}$$

Forecasting taking seasonality into account

When time series data is seasonal, it is usual to deseasonalise the data before fitting the trend line.

Example 16 Fitting a trend line (seasonality)

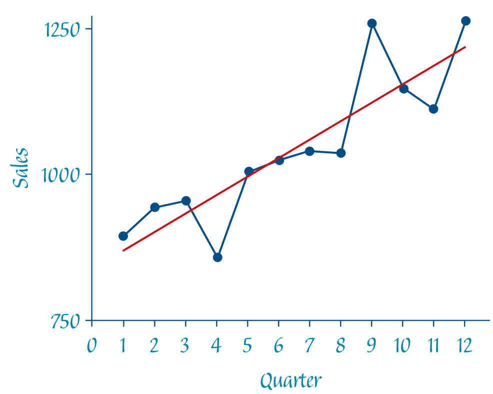
The *deseasonalised* quarterly sales data from Mikki's shop are shown below.

Quarter	1	2	3	4	5	6	7	8	9	10	11	12
Sales	893	943	955	858	1005	1026	1043	1040	1261	1151	1115	1267

Fit a trend line and interpret the slope.

Solution

1. Plot the time series.
2. Using the calculator (with *Quarter* as the EV and *Sales* as the RV), find the equation of the least squares regression trend line. Plot it on the time series.



3. Write down the equation of the least squares regression trend line.

$Sales = 838.0 + 32.1 \times quarter$

4. Interpret the slope in terms of the variables involved.

Over the 3-year period, sales at Mikki's shop increased at an average rate of 32 sales per quarter.

ISBN 978-1-108-46154-2

Making predictions with deseasonalised data

When using deseasonalised data to fit a trend line, you must remember that the result of any prediction is a deseasonalised value. To be meaningful, this result must then be reseasonalised by multiplying by the appropriate seasonal index.

Example 17 Forecasting (seasonality)

What sales do we predict for Mikki's shop in the winter of the 4th year? (Because many items have to be ordered well in advance, retailers often need to make such decisions.)

Solution

1 Substitute the appropriate value for the time period in the equation for the trend line. Since summer year 1 was designated as quarter '1', then winter year 4 is quarter '15'.

Sales = 838.0 + 32.1 × quarter
= 838.0 + 32.1 × 15
= 1319.5

Deseasonalised sales prediction for winter of year 4 = 1319.5

2 The value just calculated is the deseasonalised sales figure for the quarter in question.
To obtain the *actual* predicted sales figure we need to reseasonalise this predicted value. To do this, we multiply this value by the seasonal index for winter, which was found to be 1.30 in Example 12.

Seasonalised sales prediction for winter of year 4 = 1319.5 × 1.30
≈ 1715

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 3E

Example 14

1 Consider the population of Australia over the period 2006–2017.

Population of Australia 2006–2017	
Year	Population (1 000 000)
2017	24.5
2016	24.1
2015	23.8
2014	23.5
2013	23.1
2012	22.7
2011	22.3
2010	22.0
2009	21.7
2008	21.3
2007	20.8
2006	20.7

a Fit a least squares regression trend line to this data.

b Use this equation to predict the population of Australia in 2024.

Example 15

2 The table below shows the percentage of total retail sales that were made in department stores over an 11-year period:

Year	1	2	3	4	5	6	7	8	9	10	11
Sales (%)	12.3	12.0	11.7	11.5	11.0	10.5	10.6	10.7	10.4	10.0	9.4

a Construct a time series plot.

b Comment on the time series plot in terms of trend.

c Fit a least squares regression trend line to the time series plot, find its equation and interpret the slope.(Give answer to three significant figures.)

d Draw the trend line on your time series plot.

e Use the least squares equation to forecast the percentage of retails sales which will be made by department stores in year 15.

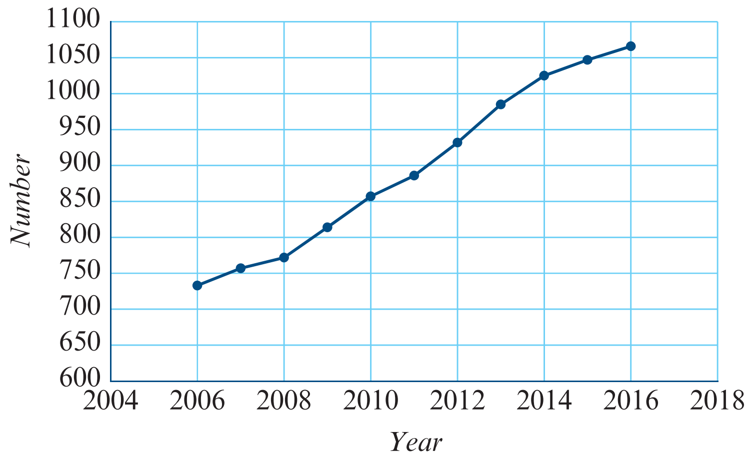
ISBN 978-1-108-46154-2

Cambridge University Press

3 The data shows the number of students enrolled (in thousands) in university in Australia for the period 2006–2016.

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Number	733	757	772	814	857	886	932	985	1025	1047	1066

The time series plot of the data is shown below.



a Comment on the plot.

b Fit a least squares regression trend line to the data. Interpret the slope.

c Use this equation to predict the number of university students in Australia in 2025.

Fitting a least squares regression trend line to a time series with seasonality

Example 16

4 **a** The table below shows the *deseasonalised* quarterly washing-machine sales of a company over 3 years. Use least squares regression to fit a trend line to the data.

	Year 1				Year 2				Year 3			
Quarter number	1	2	3	4	5	6	7	8	9	10	11	12
Deseasonalised	53	51	54	55	64	64	61	63	67	69	68	66

b Use this trend equation for washing-machine sales, with the seasonal indices below, to forecast the sales of washing machines in the fourth quarter of year 4.

Quarter	1	2	3	4
Seasonal index	0.90	0.81	1.11	1.18

 ISBN 978-1-108-46154-2 Cambridge University Press

Example 17

5 The number of international visitors (in thousands) arriving in Australia each year from 1996–2015 is given in the following table.

Year	Number	Year	Number
1996	4165	2006	5532
1997	4318	2007	5644
1998	4167	2008	5586
1999	4459	2009	5490
2000	4931	2010	5790
2001	4856	2011	5771
2002	4841	2012	6032
2003	4746	2013	6382
2004	5215	2014	6868
2005	5499	2015	7444

a Fit a least squares regression trend line to the data.

b Use this trend line to predict the number of international visitors arriving in Australia in 2030. Explain why this prediction may not be reliable.

6 The number of boogie boards sold by a surf shop over a two-year period is given in the table.

CU

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	138	60	73	230
2	283	115	163	417

The quarterly seasonal indices are given below.

Seasonal index	1.13	0.47	0.62	1.77

a Use the seasonal indices to calculate the deseasonalised sales figures for this period to the nearest whole number.

b Plot the actual sales figures and the deseasonalised sales figures for this period and comment on the plot.

c Fit a least squares regression trend line to the deseasonalised sales data. Write the slope and intercept correct to three significant figures.

d Use the relationship calculated in **c**, together with the seasonal indices, to forecast the sales for the first quarter of year 4 (you will need to reseasonalise here).

ISBN 978-1-108-46154-2

Cambridge University Press

3F Conducting a statistical investigation involving time series

We now have the tools to extend our statistical investigations to questions concerning time series data.

Example 18 Conducting a statistical investigation involving time series

To help with forward planning a tourism authority wishes to examine the trends of overseas visitors to Australia, specifically those whose primary reason for visiting Australia is for a holiday. Conduct a statistical investigation and write a report on the findings.

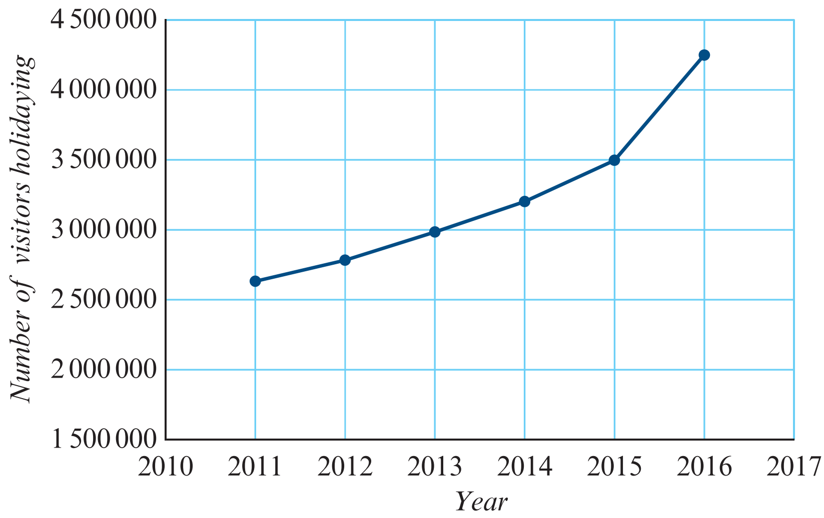
Solution

Data for the number of overseas tourism arrivals can be accessed through the Australian Bureau of Statistics website.

Data for the years 2011–2016 for those visitors who reported the primary purpose of their visit as holidays is given in the table below.

Year	2011	2012	2013	2014	2015	2016
Holiday visitors	2632300	2782900	2984700	3202000	3496900	4249800

A time series plot of the data is given below.



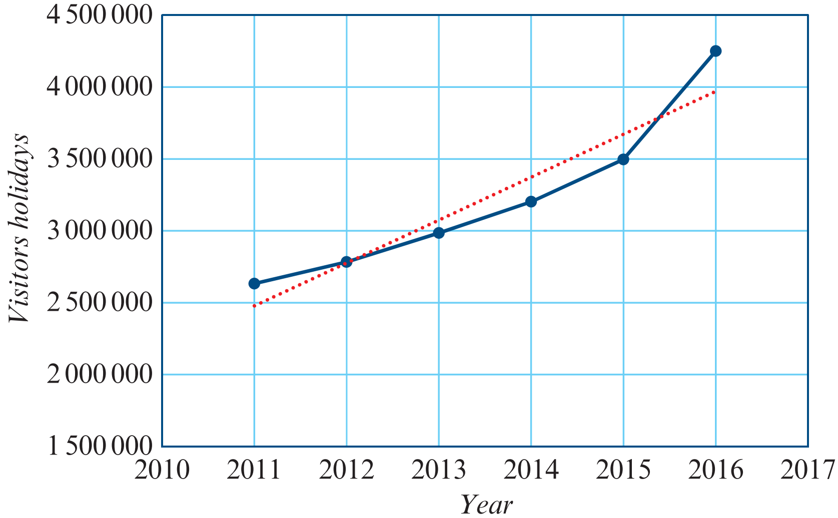
If a least squares regression line is fitted to the data, the following equation is obtained:

$$\textit{holiday visitors} = -597\,764\,713 + 298\,480 \times \textit{year}$$

ISBN 978-1-108-46154-2

Cambridge University Press

This model predicts that the number of overseas holiday visitors to Australia is increasing by approximately 300 000 per year. The line is shown on the following plot:



It also seems likely that the number of overseas holiday visitors to Australia might be seasonal.

Using monthly arrival data for 2015 and 2016, the monthly seasonal indices can be calculated by determining seasonal indices for each year separately then averaging to combine the information for the two years. These calculations are as shown in the following table.

Month	Monthly visitors		Seasonal indices		
	2015	2016	2015	2016	Average
Jan	228 900	318 600	0.785	0.900	0.84
Feb	342 800	411 600	1.176	1.162	1.17
Mar	308 300	397 000	1.058	1.121	1.09
April	247 500	310 100	0.849	0.876	0.86
May	214 900	274 000	0.737	0.774	0.76
Jun	214 000	277 600	0.734	0.784	0.76
Jul	271 400	344 500	0.931	0.973	0.95
Aug	269 100	340 500	0.923	0.961	0.94
Sept	265 500	329 200	0.911	0.930	0.92
Oct	314 900	345 900	1.081	0.977	1.03
Nov	342 800	384 900	1.176	1.087	1.13
Dec	476 800	515 900	1.636	1.457	1.55

ISBN 978-1-108-46154-2

Cambridge University Press

We can use the previous analyses to write the following report.

Report

The number of overseas visitors to Australia in the years 2011–2016 who reported their primary reason for visiting to be for a holiday was investigated. A time series plot showed an increasing trend. A trend forecasting line fitted to the data using least squares regression has the equation:

$$\text{holiday visitors} = -597\,764\,713 + 298\,480 \times \text{year}$$

This tell us that on average the number of visitors is increasing by approximately 300 000 per year. Using this rule, we would forecast the number of overseas holiday visitors in 2017 to be 4 269 447. However, looking carefully at the plot we can see that there was a marked increased in the number of visitors between 2015 and 2016, and we should be cautious with using the least squares line for forecasting too far into the future.

The number of visitors also exhibited monthly seasonality. Based on monthly visitor data for 2014 and 2015, the following seasonal indices were determined:

Month	Jan	Feb	Mar	April	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
SI	0.84	1.17	1.09	0.86	0.76	0.76	0.95	0.94	0.92	1.03	1.13	1.55

Thus, we can say that the peak month for overseas holiday visitors is December, where we can expect 55% more visitors than the average month, while the least popular months for overseas holiday visitors are May and June, both recording 24% less visitors than the average month.

ISBN 978-1-108-46154-2

Exercise 3F

Example 18

1 The table below gives the number of births in Australia ('000s), quarterly from March 2011.

	Q1	Q2	Q3	Q4
2011	76.0	76.3	76.2	74.2
2012	78.1	77.5	78.7	77.9
2013	77.4	77.1	76.9	75.7
2014	77.5	76.9	79.0	77.1
2015	76.2	75.4	78.8	75.9
2016	78.8	78.2	75.7	70.7

a Construct a time series plot and comment on its features.

b Construct a table of the centred four-point moving averages of the number of births in Australia. Superimpose the moving average on the time series plot of the data. Comment on what this graph shows and why a four-point moving average has been recommended.

c Select data for three years and use it to calculate seasonal indices for birth rate.

d Discuss in a few sentences what these seasonal indices tell you about the seasonal pattern of the number of births in Australia.

e Determine the equation of the least squares regression line for the data. Interpret the intercept and slope of the regression line in terms of number of births and time, and use the equation to predict the number of births in Q3 in 2020.

ISBN 978-1-108-46154-2

2 The tables below relate to the number of people in Queensland who are employed and unemployed.

a The table below gives the number of people in Queensland who are employed, by month from Jan 2013 to December 2017.

	Year				
Employed ('000s)	2013	2014	2015	2016	2017
Jan	2276.0	2259.8	2282.5	2356.7	2327.8
Feb	2281.7	2312.5	2340.6	2396.8	2360.0
March	2282.8	2333.9	2313.6	2358.0	2374.6
April	2291.7	2340.9	2326.4	2361.9	2397.9
May	2294.3	2342.6	2351.1	2361.8	2407.5
June	2289.6	2340.6	2345.3	2352.7	2400.1
July	2312.2	2334.7	2339.2	2365.0	2431.9
Aug	2286.4	2339.2	2338.1	2342.7	2439.6
Sept	2338.6	2321.6	2359.6	2341.3	2440.0
Oct	2330.0	2318.6	2385.0	2338.5	2469.1
Nov	2321.2	2308.6	2381.1	2373.3	2468.8
Dec	2326.1	2345.5	2404.7	2377.3	2479.2

i Construct a times series plot of employment and comment on its features.

ii Select data for two years and use it to calculate seasonal indices for employment.

iii Discuss in a few sentences what these seasonal indices tell you about the employment pattern in Queensland.

iv Use the seasonal indices to compute deseasonalised data. Use the deseasonalised data to produce a time series plot of employment against time, and comment on the plot.

v Determine the equation of the least squares regression line for the deseasonalised data. Interpret the intercept and slope of the regression line in terms of employment and time.

ISBN 978-1-108-46154-2

b The table below gives the number of people in Queensland who are unemployed, by month from Jan 2013 to December 2017.

	Year				
Unemployed ('000s)	2013	2014	2015	2016	2017
Jan	145.2	159.0	169.3	173.7	165.6
Feb	165.2	181.2	186.5	158.1	191.3
March	157.5	166.8	178.9	169.2	173.7
April	131.0	151.7	160.0	161.3	159.7
May	140.8	153.1	154.6	157.9	153.9
June	149.8	153.7	146.1	157.2	164.6
July	132.8	159.6	154.8	146.5	153.3
Aug	134.7	159.7	156.1	151.9	144.8
Sept	144.6	154.4	152.6	151.1	146.9
Oct	136.7	167.6	148.2	135.7	144.7
Nov	127.4	153.5	134.9	134.5	135.8
Dec	144.7	148.7	140.4	145.1	147.1

i Construct a times series plot of unemployment and comment on its features.

ii Select data for two years and use it to calculate seasonal indices for unemployment.

iii Discuss in a few sentences what these seasonal indices tell you about the unemployment pattern in Queensland.

iv Use the seasonal indices to compute deseasonalised data. Use the deseasonalised data to produce a time series plot of unemployment against time, and comment on the plot.

v Determine the equation of the least squares regression line for the deseasonalised data. Interpret the intercept and slope of the regression line in terms of unemployment and time.

c Does there seem to be any association between employment and unemployment numbers in Queensland?

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Key ideas and chapter summary

AS

Time series data — **Time series data** are a collection of data values along with the times (in order) at which they were recorded.

Time series plot — **A time series plot** is a line graph where the values of the response variable are plotted in time order.

Features to look for in a time series plot

- Trend
- Cycles
- Seasonality
- Structural change
- Possible outliers
- Irregular (random) fluctuations

Trend — **Trend** is present when there is a long-term upward or downward movement in a time series.

Cycles — **Cycles** are present when there is a periodic movement in a time series. The period is the time it takes for one complete up and down movement in the time series plot. This term is generally reserved for periodic movements with a period greater than one year.

Seasonality — **Seasonality** is present when there is a periodic movement in a time series that has a calendar related period, for example, a year, a month, a week.

Structural change — **Structural change** is present when there is a sudden change in the established pattern of a time series plot.

Outliers — **Outliers** are present when there are individual values that stand out from the general body of data.

Irregular (random) fluctuations — **Irregular (random) fluctuations** are always present in any real-world time series plot. They include all of the variations in a time series that we cannot reasonably attribute to systematic changes like trend, cycles, seasonality, structural change or the presence of outliers.

Smoothing — **Smoothing** is a technique used to eliminate some of the irregular fluctuations in a time series plot so that features such as trend are more easily seen.

Moving mean smoothing — In **moving mean smoothing**, each original data value is replaced by the mean of itself and a number of data values on either side. When smoothing over an even number of data points, centring is required to ensure the smoothed mean is centred on the chosen point of time.

ISBN 978-1-108-46154-2

Seasonal indices — **Seasonal indices** are used to quantify the seasonal variation in a time series.

Deseasonalise — The process of accounting for the effects of seasonality in a time series is called **deseasonalisation**.

Reseasonalise — The process of a converting seasonal data back into its original form is called **reseasonalisation**.

Trend line forecasting — **Trend line forecasting** uses the equation of a trend line to make predictions about the future.

Skills check

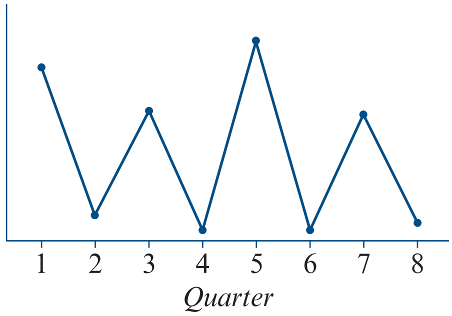
Having completed this chapter you should be able to:

- recognise time series data
- construct a times series plot
- identify the presence of trend, cycles, seasonality, structural change and irregular (random) fluctuations in a time series plot
- smooth a time series to help identify any trend
- calculate and interpret seasonal indices
- calculate and interpret a trend line for linear trends
- use a trend line to make forecasts.

Multiple-choice questions

1 The pattern in the time series in the graph shown is best described as:

A trend
B cyclical but not seasonal
C seasonal
D irregular
E average



Use the following table to answer Questions 2 to 5.

Time period	1	2	3	4	5	6
Data value	2.3	3.4	4.4	2.7	5.1	3.7

ISBN 978-1-108-46154-2

Cambridge University Press

2 The three-moving mean for time period 2 is closest to:
A 3.4 **B** 3.6 **C** 3.9 **D** 4.0 **E** 4.2

3 The five-moving mean for time period 3 is closest to:
A 3.4 **B** 3.6 **C** 3.9 **D** 4.1 **E** 4.2

4 The two-moving mean for time period 5 with centring is closest to:
A 2.7 **B** 3.6 **C** 3.9 **D** 4.0 **E** 4.2

5 The four-moving mean for time period 4 with centring is closest to:
A 2.7 **B** 3.6 **C** 3.9 **D** 4.1 **E** 4.2

6 The seasonal indices for the number of customers at a restaurant are as follows.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.0	p	1.1	0.9	1.0	1.0	1.2	1.1	1.1	1.1	1.0	0.7

The value of p is:
A 0.5 **B** 0.7 **C** 0.8 **D** 1.0 **E** 1.2

Use the following information to answer Questions 7–10.

The seasonal indices for the number of bathing suits sold at a surf shop are given in the table.

Quarter	Summer	Autumn	Winter	Spring
Seasonal index	1.8	0.4	0.3	1.5

7 The number of bathing suits sold one summer is 432. The deseasonalised number is closest to:
A 432 **B** 240 **C** 778 **D** 540 **E** 346

8 The *deseasonalised* number of bathing suits sold one winter was 380. The actual number was closest to:
A 114 **B** 133 **C** 152 **D** 380 **E** 1267

9 The seasonal index for spring tells us that, over time, the number of bathing suits sold in spring tends to be:
A 50% less than the seasonal average
B 15% less than the seasonal average
C the same as the seasonal average
D 15% more than the seasonal average
E 50% more than the seasonal average

10 To correct for seasonality, the actual number of bathing suits sold in Autumn should be:
A reduced by 60% **B** reduced by 40% **C** increased by 40%
D increased by 60% **E** increased by 150%

ISBN 978-1-108-46154-2

11 The number of visitors to an information centre each quarter was recorded for one year. The results are tabulated below.

Quarter	Summer	Autumn	Winter	Spring
Visitors	1048	677	593	998

Using this data, the seasonal index for autumn is estimated to be closest to:

A 0.25 **B** 1.0 **C** 1.23 **D** 0.82 **E** 0.21

Use the following information to answer Questions 12 and 13.

A trend line is fitted to a time series plot displaying the average age at marriage of males (in years) for the period 2005–2016.

The equation of this line is: $age = 31.1 + 0.236 \times year$

Here year 1 is 2005, year 2 is 2006, and so on.

12 Using this trend line, the average age of marriage of males in 2020 is forecasted to be:

A 31.3 **B** 34.6 **C** 34.9 **D** 35.1 **E** 500.0

13 From the slope of the trend line it can be said that:

A on average, the age of marriage for males is increasing by about 3 months per year
B on average, the age of marriage for males is decreasing by about 3 months per year
C older males are more likely to marry than younger males
D no males married at an age younger than 27 years
E on average, the age of marriage for males is increasing by 0.236 months per year

Use the following information to answer Questions 14 and 15.

Suppose that the seasonal indices for the wholesale price of petrol are:

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Index	1.2	1.0	0.9	0.8	0.7	1.2	1.2

The daily deseasonalised prices for a petrol outlet for a week (in cents/litre) are given in the following table.

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Price	88.3	85.4	86.7	88.5	90.1	91.7	94.6

14 The equation of the least squares regression line that could enable us to predict the deseasonalised price is closest to:

A $price = 84.3 + 1.25 \times day$
B $price = -49.7 + 0.601 \times day$
C $price = 1.25 + 84.3 \times day$
D $price = 0.601 - 49.7 \times day$
E $price = 84.3 - 1.25 \times day$

ISBN 978-1-108-46154-2

Cambridge University Press

Review

15 The seven-mean deseasonalised smoothed price of petrol (in cents/litre) for this week was closest to:

A 87.4 **B** 88.3 **C** 88.5 **D** 89.3 **E** 90.0

16 The deseasonalised price (in cents/litre) on Thursday was 90.1. The actual price on Thursday was closest to:

A 63.1 **B** 75.6 **C** 110.8 **D** 128.7 **E** 135.4

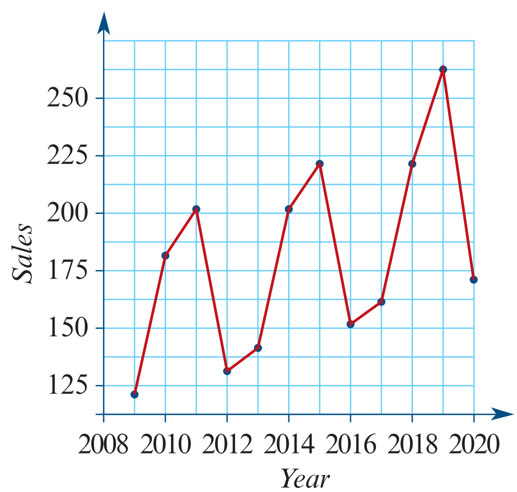
Short-answer questions

1 Construct a time series plot to display the following data: SF

Year	2013	2014	2015	2016	2017	2018	2019	2020
Sales	12	15	35	23	67	56	78	70

2 Briefly explain the concepts of trend, seasonality, and cycles in a time-series plot. CF

3 The time series plot below shows house sales per year in a certain town over a 12-year period.



Describe the features of the plot.

4 What is the purpose of smoothing? SF

5 The following table shows the maximum daily temperature in a certain town over a one-week period:

Day	Mon	Tues	Wed	Thu	Fri	Sat	Sun
Temp (°C)	20	21	26	36	34	42	26

a Determine the three mean smoothed temperature for Tuesday.

b Determine the five mean smoothed temperature for Friday.

ISBN 978-1-108-46154-2

Questions 6 & 7 refer to the following information.

The value of an Australian dollar in US dollars (exchange rate) over a 10-day period is given in the table.

Day	1	2	3	4	5	6	7	8	9	10
Exchange rate	0.7183	0.7146	0.7093	0.7095	0.7076	0.7035	0.7049	0.7009	0.7044	0.7069

6 Determine the centred two mean smoothed value for Day 5. CF

7 Determine the centred four mean smoothed value for Day 6.

8 The seasonal index for sales in December in a retail shop is 1.25. If the actual sales last December were recorded as 2768, what is the deseasonalised sales figure? SF

9 The seasonal indices for the daily sales in an ice-cream shop are as follows: CF

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Seasonal index	0.55	0.50	0.58	0.64	1.0	?	1.87

a What is the seasonal index for Saturday?
b Interpret this index in terms of the average daily sales for this shop.

10 The number of staff employed by a restaurant chain in each quarter of one year is given in the following table. SF

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Number of staff	35	98	65	23

Use these data to calculate quarterly seasonal indices for the number of staff employed by the restaurant chain.

11 The number of international visitors (in thousands) arriving in a certain country each year from 2014–2020 is given in the following table. CF

Year	000's of visitors
2014	6490
2015	6790
2016	6771
2017	7032
2018	7382
2019	7868
2020	8244

a Fit a least squares trend line to the data.
b Use this to predict the number of visitors in 2025. Explain why this prediction may be unreliable.

Review

ISBN 978-1-108-46154-2

Cambridge University Press

12 The number of people arriving at a certain station each day is seasonal, with seasonal indices as shown below:

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Seasonal index	1.32	1.20	0.95	0.95	1.48	0.65	0.45

It is also known that over a specific time period the number of travellers each day has generally been increasing, according to the following equation which was determined from deseasonalised data:

$$\text{number of travellers} = 135 + 22.5 \times \text{day}$$

If a certain Saturday is designated as Day 1, how many travellers would you predict for the following Saturday?

Extended-response questions

1 Gross Domestic Product per capita (GDP per capita) is a measure of the economic performance of a country. The following table shows the Australian GDP per capita for the 12 years from 2005–2016. CU

Year	GDP per capita ($US)
2005	36 179
2006	37 908
2007	45 163
2008	49 224
2009	45 604
2010	56 360
2011	66 773
2012	68 048
2013	64 734
2014	61 232
2015	51 220
2016	51 737

a Construct a time series plot of the data.
b Briefly describe the pattern in the data.
c Find the equation of the least squares regression line that would enable GDP per capital to be predicted from year to year.
d In 2017 the GDP per capital of Australia was $US56 135. Find the error in the prediction if the least squares regression line in **c** is used to predict the GDP per capita in 2017.

ISBN 978-1-108-46154-2

Cambridge University Press

2 The table below shows the rainfall (in mm) for each month in a region near Cairns in 2015 and 2016:

	2015	2016
Jan	251.2	129.4
Feb	605.0	125.2
Mar	280.2	283
Apr	128.1	187.3
May	56.3	178.4
Jun	243.1	108.6
Jul	26.9	99.0
Aug	22.6	47.3
Sept	29.1	55.5
Oct	57.3	56
Nov	131.5	26.5
Dec	420.4	89.1

a Construct a time series plot for the rainfall in 2015.

b Use three-mean smoothing to smooth the time series. Plot the smoothed time series for 2015 on the plot. Describe the general pattern of the rainfall for 2015.

c Construct a time series plot for the rainfall in 2016.

d Use three-mean smoothing to smooth the time series. Plot the smoothed time series for 2016 on the plot. Describe the general pattern in rainfall for 2016.

e Compare the rainfall patterns in 2015 and 2016, identifying similarities and differences.

3 The table below shows the average interest rate for the period 1987–97. Also shown are the three-mean smoothed average rates but with one missing.

Year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Average rate (%)	15.50	13.50	17.00	16.50	13.00	10.50	9.50	8.75	10.50	8.75	7.55
Three-mean (%)		15.33	15.67	15.50	13.33		9.58	9.58	9.33	8.93	

a Complete the table by showing that the three-mean smoothed interest rate for 1992 is 11.0%.

b Construct a time series plot for the average interest rate during the period 1987–97.

c Plot the smoothed interest rate data on the graph and comment on any trend.

ISBN 978-1-108-46154-2

Cambridge University Press

Chapter 4

4

Arithmetic and geometric sequences

UNIT 3 BIVARIATE DATA, SEQUENCES AND CHANGE, AND EARTH GEOMETRY

Topic 3 Growth and decay in sequences

- How do we identify a sequence?
- How do we generate a sequence?
- How do we define a recurrence relation in a sequence?
- How do we calculate a certain term in a sequence?
- How do we use sequences to determine the value of items depreciating?
- How do we use sequences to determine the value of loans and investments?

ISBN 978-1-108-46154-2

Cambridge University Press

4A Sequences and simple recursion

Sequences

A list of numbers, written down in succession, is called a **sequence**. Each of the numbers in a sequence is called a *term*. We write the terms of a sequence as a list, separated by commas. If a sequence continues indefinitely, or if there are too many terms in the sequence to write them all, we use an *ellipsis* '…' at the end of a few terms of the sequence like this:

12, 22, 5, 6, 16, 43, …

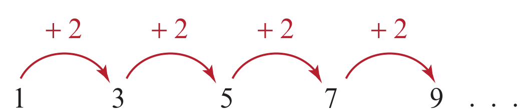
The terms in this sequence of numbers could be the ages of the people boarding a plane.

The ages of these people is random, so this sequence of numbers is called a random sequence. There is no pattern or rule that allows the next number in the sequence to be predicted.

Some sequences of numbers do display a pattern. For example, this sequence:

1, 3, 5, 7, 9, …

has a definite pattern and so this sequence is said to be rule-based. The sequence of numbers has a starting value, which is 1. We add 2 to this number to generate the term 3. Then, add 2 again to generate the term 5, and so on. The rule is 'add 2 to each term'.



Recursion

In the following two examples we generate a sequence of terms by *recursion*. A simple definition of recursion is that from a given starting point the sequence is built up term by term always moving from one term to the next by applying the same rule.

In the sequence 1, 3, 5, 7, 9, …, the starting point was 1 and the rule was 'add 2'.

Example 1 Generating a sequence by recursion (by adding)

Write down the first five terms of the sequence with a starting value of 6 and the rule 'add 4'.

Solution

1 Write down the starting value.	6
2 Apply the rule (add 4) to generate the next term.	$6 + 4 = 10$
3 Calculate three more terms.	$10 + 4 = 14,\ 14 + 4 = 18,\ 18 + 4 = 22$
4 Write your answer.	6, 10, 14, 18, 22

General Mathematics Units 3 & 4 ISBN 978-1-108-46154-2

Cambridge University Press
Cambridge Senior Mathematics for Queensland

Example 2 Generating a sequence by recursion (by multiplying)

Write down the first five terms of the sequence with a starting value of 3 and the rule 'multiply by 4'.

Solution

1 Write down the starting value. — 3

2 Apply the rule (multiply by 4) to generate the next term. — $3 \times 4 = 12$

3 Calculate three more terms. — $12 \times 4 = 48,\ 48 \times 4 = 192,\ 192 \times 4 = 768$

4 Write your answer. — 3, 12, 48, 192, 768

Using a calculator to generate a sequence of numbers by recursion

A calculator can perform recursive calculations very easily, because it automatically stores the answer to the last calculation it performed, as well as the method of calculation. The way it does this depends on the calculator you are using.

Example 3 Generating sequences of numbers from a rule using a calculator

Use a calculator to generate the first five terms of the sequence with a starting value of 5 and the rule 'double and then subtract 3'.

Solution

1 Press **clear** (Casio scientific calculators) or **clear** (TI scientific calculators) to create a blank computation screen.

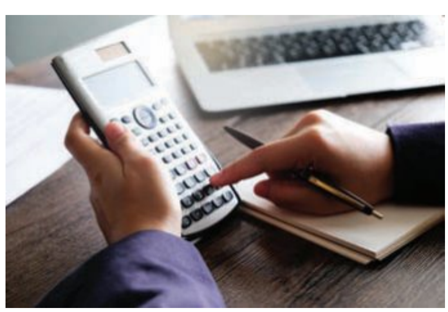
| |
| 0

2 Type 5 and then press = (Casio) or **enter** (TI). This stores the starting value of the sequence in the calculator memory.

5
5

3 Next, type **× 2 − 3** and then press = (Casio) or **enter** (TI). The second term of the sequence will be calculated and displayed on the screen.

Ans × 2 − 3
7



ISBN 978-1-108-46154-2

4 Press = (Casio) or **enter** (TI) repeatedly to apply the sequence rule to the previous calculated term.

Pressing '=' or enter 1 more time

Ans × 2 − 3
11

Note: The screen will show the calculation as 'Ans × 2 − 3' where 'Ans' is the previously calculated term value.

Pressing '=' or enter another time

Ans × 2 − 3
19

5 Write down the sequence terms as they are calculated.

The sequence is 5, 7, 11, 19, 35, ...

Exercise 4A

Generating a sequence by recursion

Example 1, 2

1 Use the following starting values and rules to generate the first five terms of the following sequences recursively by hand.

a Starting value: 2, rule: add 6
b Starting value: 5, rule: subtract 3
c Starting value: 10, rule: divide by 2
d Starting value: 6, rule: multiply by 2

SF

Generating a sequence by recursion using a calculator

Example 3

2 Use the following starting values and rules to generate the first five terms of the following sequences recursively using a calculator.

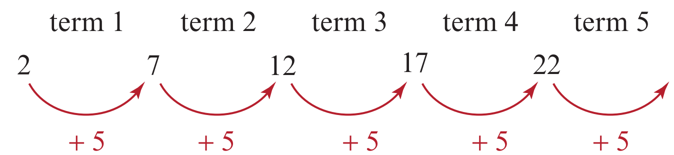
a Starting value: 4, rule: add 2
b Starting value: 50, rule: divide by 5
c Starting value: 24, rule: subtract 4
d Starting value: 5, rule: multiply by 3
e Starting value: 2, rule: multiply by 5

ISBN 978-1-108-46154-2

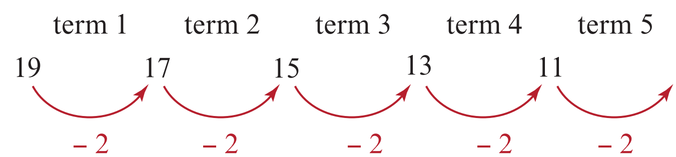
Cambridge University Press

4B Defining an arithmetic sequence by recursion

A sequence in which each successive term can be found by adding the same number to the previous term is called an **arithmetic sequence**. For example, the sequence 2, 7, 12, 17, 22, … is arithmetic because each successive term can be found by adding 5.



The sequence 19, 17, 15, 13, 11, … is also arithmetic because each successive term can be found by adding −2.



The common difference

Because of the way in which an arithmetic sequence is formed, the *difference* between successive terms is constant. We call this difference the *common difference*. In the sequence:

2, 7, 12, 17, 22, …

the common difference is +5 , while in the sequence:

19, 17, 15, 13, 11, …

the common difference is −2.

Method of recursion to generate an arithmetic sequence

Method for using recursion to generate an arithmetic sequence

The method for using recursion to generate an arithmetic sequence has two parts:

1. A starting point, the value of the first term $t_1 = a$ of the sequence
2. Successively adding the common difference d to each term to obtain the next term.

This will generate the sequence.

ISBN 978-1-108-46154-2

Cambridge University Press

For example, in words, a rule for the recursion that can be used to generate the sequence:

10, 15, 20, 25, …

can be written as follows:

1 Start with 10.
2 To obtain the next term, add 5 to the current term and repeat the process.

A more compact way of communicating this information is to translate this rule into symbolic form. We do this in the following way.

- t_1 is the first term of a sequence
- t_2 is the second term of a sequence
- t_3 is the third term of a sequence

and so on. t_n is the nth term of the sequence.

Here we have used t_n, but the t can be replaced by any letter of the alphabet.

For example, for the sequence 10, 15, 20, 25, …

$$t_1 = 10,\ t_2 = 15,\ t_3 = 20,\ t_4 = 25$$

This notation gives us a way to describe the generation of an arithmetic sequence by recursion.

$t_1 = a$ and $t_{n+1} = t_n + d$, where d is the common difference.

We call this a **recurrence relation.**

Note: This also gives us a way to test whether a sequence is arithmetic or not.

We must have:

$$t_{n+1} - t_n = d$$

for all whole numbers $n \geqslant 1$.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 4 Generating an arithmetic sequence with a recurrence relation

An arithmetic sequence is defined by the recurrence relation:

$$t_1 = 9, t_{n+1} = t_n - 4$$

Find the first five terms.

Solution

1. We start with the first term $t_1 = 9$.
2. Subtract 4 each time to generate the sequence.

$t_1 = 9$
$t_2 = t_1 - 4 = 9 - 4 = 5$
$t_3 = t_2 - 4 = 5 - 4 = 1$
$t_4 = t_3 - 4 = 1 - 4 = -3$
$t_5 = t_4 - 4 = -3 - 4 = -7$

Graphing an arithmetic sequence

Consider the arithmetic sequences defined by:

- $P_1 = 22,\ P_{n+1} = P_n + 2$ (Sequence 1)
- $Q_1 = 18,\ Q_{n+1} = Q_n - 2$ (Sequence 2)

The first generates the arithmetic sequence 22, 24, 26, …

It is an *increasing sequence*.

The second generates the arithmetic sequence 18, 16, 14, …

It is a *decreasing sequence*.

Sequence 1

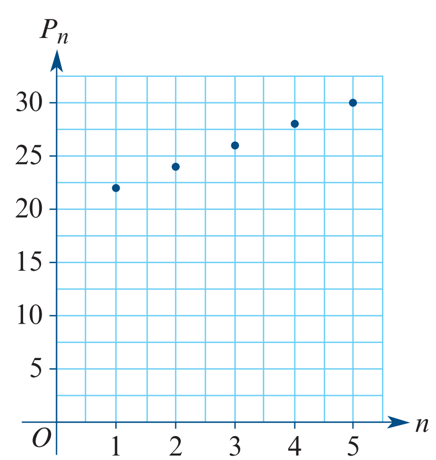
The first few values of the sequence can be shown in a table.

The first five values of Sequence 1 are shown in the table:

n	1	2	3	4	5
P_n	$P_1 = 22$	$P_2 = 24$	$P_3 = 26$	$P_4 = 28$	$P_5 = 30$

and this gives ordered pairs (1, 22), (2, 24), (3, 26), (4, 28), (5, 30).

These can be graphed as shown opposite.



ISBN 978-1-108-46154-2

Cambridge University Press

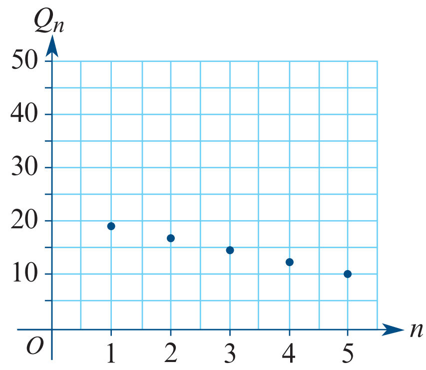
Sequence 2

The first five values of Sequence 2 are shown in the table:

n	1	2	3	4	5
Q_n	$Q_1 = 18$	$Q_2 = 16$	$Q_3 = 14$	$Q_4 = 12$	$Q_5 = 10$

and this gives ordered pairs
(1, 18), (2, 16), (3, 14), (4, 12), (5, 10).

These can be graphed as shown opposite.



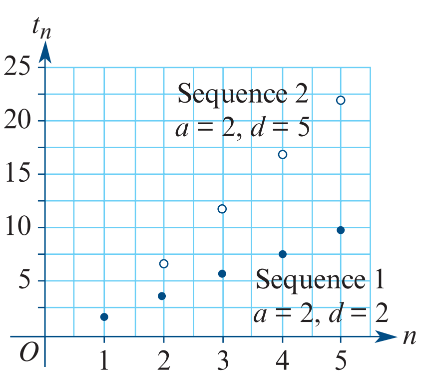
The graphs below display the terms in four different arithmetic sequences:

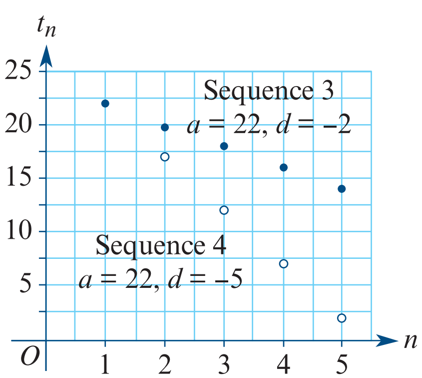
Sequence 1: $t_1 = a = 2,\ d = 2,\ n = 1, 2, 3, \ldots$

Sequence 2: $t_1 = a = 2, d = 5,\ n = 1, 2, 3, \ldots$

Sequence 3: $t_1 = a = 22,\ d = -2,\ n = 1, 2, 3, \ldots$

Sequence 4: $t_1 = a = 22,\ d = -5, n = 1, 2, 3, \ldots$





Linear growth and decay

The key characteristics to note from the graphs above:

- The points in the graphs are *collinear*; that is, the points lie on a *straight line*.
- If the common difference, d, is positive, the terms in the sequence increase. The bigger the value of d, the more rapid the increase. An arithmetic sequence with a positive common difference can be used to model **linear growth**.
- If the common difference, d, is negative, the terms in the sequence decrease. The bigger the absolute value of d, the more rapid the decrease. An arithmetic sequence with a negative common difference can be used to model **linear decay**.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 5 Graphing arithmetic sequences

Prepare a table of values and plot a graph to illustrate the first five terms of the sequence defined by

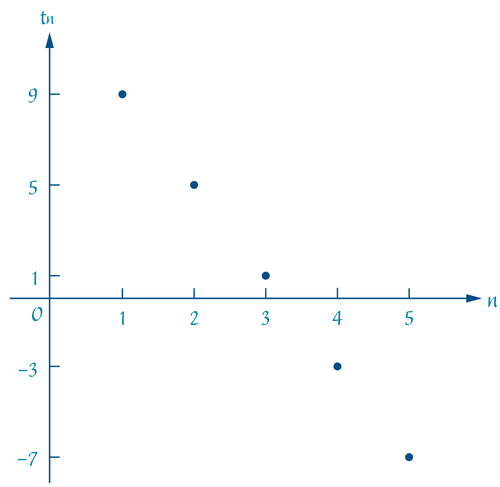
$$t_1 = 9, t_{n+1} = t_n - 4$$

Solution

1 Generate the table from the recurrence relation by subtracting 4 from a term to give the next term.

n	1	2	3	4	5
t_n	9	5	1	−3	−7

2 Plot the ordered pairs (1, 9), (2, 5), (3, 1), (4, −3) and (5, −7).



ISBN 978-1-108-46154-2

Exercise 4B

Generating an arithmetic sequence with a recurrence relation

Example 4

SF

1 An arithmetic sequence is defined by:

$$t_1 = 3,\ t_{n+1} = t_n + 4$$

Find the first five terms.

2 An arithmetic sequence is defined by:

$$t_1 = 15,\ t_{n+1} = t_n - 4$$

Find the first five terms.

3 Consider the arithmetic sequence 20, 24, 28, 32, 36, …

a What is the common difference?

b What is the next term in the sequence?

c Starting with 20, how many times do you have to add the common difference to get to term 8? What is the value of term 8?

d Starting with 20, how many times do you have to add the common difference to get to term 13? What is the value of term 13?

4 Consider the arithmetic sequence 5, 3, 1, −1, −3, …

a What is the common difference?

b What is the next term in the sequence?

c Starting with 5, how many times do you have to add the common difference to get to term 7? What is the value of term 7?

d What is the value of term 10? What is the value of term 50?

Graphing arithmetic sequences

Example 5

5 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 9,\ t_{n+1} = t_n - 4$$

6 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 10,\ t_{n+1} = t_n + 5$$

7 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 12,\ t_{n+1} = t_n - 2$$

8 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 12,\ t_{n+1} = t_n - 4$$

9 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$s_0 = 12,\ s_{n+1} = s_n + 3$$

ISBN 978-1-108-46154-2

Cambridge University Press

4C A general rule for finding the *n*th term of an arithmetic sequence

The first term t_1 of an arithmetic sequence is often denoted by a.

In the sequence 2, 7, 12, 17, 22, …

$$t_1 = a = 2$$

We use the symbol d to represent the common difference.

In the sequence 2, 7, 12, 17, 22, …

$$d = 5$$

The rule for the *n*th term of an arithmetic sequence

Consider an arithmetic sequence with first term a and common difference d.

Then:

$$t_1 = a$$

$$t_2 = t_1 + d = a + d$$

$$t_3 = t_2 + d = a + 2d$$

$$t_4 = t_3 + d = a + 3d$$

$$t_5 = t_4 + d = a + 4d$$

$$t_6 = t_5 + d = a + 5d$$

and so on.

Thus, following the pattern, we can write:

The rule for the *n*th term of an arithmetic sequence

$t_n = a + (n - 1)d$ or $t_n = t_1 + (n - 1)d$

which gives us a rule for finding the nth term of an arithmetic sequence in terms of the first term $t_1 = a$ and the common difference d.

Example 6 Using the rule for an arithmetic sequence (positive difference)

The first term of an arithmetic sequence is $a = 6$ and the common difference is $d = 2$.

Use a rule to determine the value of the 11th term.

Solution

Substitute $a = 6$, $d = 2$ and $n = 11$ in the rule $t_n = a + (n - 1)d$.

$$t_{11} = 6 + (11 - 1)2$$
$$= 6 + 10 \times 2$$
$$= 26$$

ISBN 978-1-108-46154-2

Cambridge University Press

Example 7 Using the rule for an arithmetic sequence (negative difference)

Use a rule to determine the value of the 15th term of the arithmetic sequence 18, 15, 12, 9, …

Solution

Substitute $a = 18$, $d = -3$ and $n = 15$ in the rule $t_n = a + (n - 1)d$.

$$t_{15} = 18 + (15 - 1)(-3)$$
$$= 18 + 14 \times (-3)$$
$$= -24$$

Example 8 Determining an arithmetic sequence given two terms

In an arithmetic sequence, the fifth term is 10 and the ninth term is 18. Write down the first three terms of the sequence.

Solution

1. $t_5 = 10$, substituting in $t_n = a + (n - 1)d$ $\qquad 10 = a + 4d \quad (1)$
   $t_9 = 18$, substituting in $t_n = a + (n - 1)d$ $\qquad 18 = a + 8d \quad (2)$
2. We can now find a and d by solving equations (1) and (2).
   Subtract (1) from (2). $\qquad 8 = 4d$
   so $d = 2$
3. Substitute $d = 2$ in equation (1). $\qquad 10 = a + 4 \times 2$
   $= a + 8$
   $\therefore a = 2$
   The first three terms of the sequence are 2, 4, 6.

Example 9 Determining how many terms of an arithmetic sequence are required to reach a particular number

How many terms would we have to write down in the arithmetic sequence 10, 14, 18, 22, … before we found a term greater than 50?

Solution

We want to find n so that:

$t_n = 10 + (n - 1) \times 4 > 50$

$t_{11} = 50$, so the first term to exceed 50 is t_{12}.

Solving for n we have:

$$10 + (n - 1) \times 4 > 50$$
$$(n - 1) \times 4 > 40$$
$$(n - 1) > 10$$
$$n > 11$$

The first term to exceed 50 is t_{12}.

ISBN 978-1-108-46154-2

Graphing an arithmetic sequence

If we use the rule for finding the *n*th term of an arithmetic sequence, we understand a little more about the graphs of arithmetic sequences. We will work with the rule for which the first term is $t_1 = a$.

$$t_n = a + (n - 1)d$$

When $a = 4$ and $d = -3$

$$\begin{aligned} t_n &= 4 + (n - 1)(-3) \\ &= -3n + 7 \end{aligned}$$

When we plot the graph of t_n against n, the points lie on a straight line. This line has gradient −3. Of course, we do not join these points with this line.

Example 10 Graphing an arithmetic sequence using a rule

Consider the arithmetic sequence with $t_1 = 2$ and $d = 3$.

a Find the rule for this arithmetic sequence.

b Prepare a table of values for the sequence for $n = 1$ to $n = 5$.

c Plot a graph from the table of values.

Solution

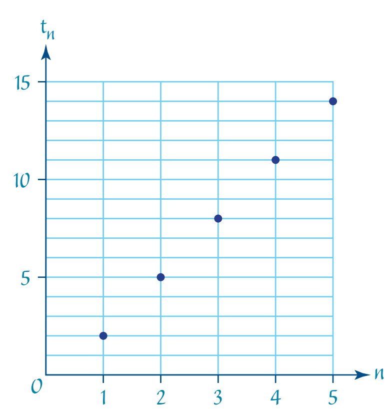
a Substitute in $t_n = a + (n - 1)d$.

$$\begin{aligned} t_n &= 2 + (n - 1) \times 3 \text{ for } n = 1, 2, 3, \ldots \\ &= 3n - 1 \end{aligned}$$

b Use the rule to generate the values of the table. For example, $t_4 = 3 \times 4 - 1 = 11$

n	1	2	3	4	5
t_n	2	5	8	11	14

c Plot the ordered pairs (1, 2), (2, 5), …



ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 4C

Using the rule for an arithmetic sequence

Example 6

SF

1 The first term of an arithmetic sequence is $a = 200$ and the common difference is $d = 10$. Use this rule to determine the value of:

a the third term

b the 10th term

c the 151st term

2 An arithmetic sequence can be described by the recurrence relation $t_1 = 6, t_{n+1} = t_n + 7$. Write down the rule for the *n*th term.

3 An arithmetic sequence can be described through the recurrence relation $t_1 = 48, t_{n+1} = t_n - 11$. Write down the rule for the *n*th term.

4 An arithmetic sequence can be described through the recurrence relation $t_1 = 8, t_{n+1} = t_n + 11$. Write down the rule for the *n*th term.

5 An arithmetic sequence can be described through the recurrence relation $t_1 = 1000, t_{n+1} = t_n - 20$. Write down the rule for the *n*th term.

6 Use the rule for the *n*th term of an arithmetic sequence with $t_1 = a$ to determine the value of:

a the 11th term of the arithmetic sequence 5, 10, 15, 20, 25, …

b the 8th term of the arithmetic sequence 12, 8, 4, 0, …

Example 7

c the 27th term of the arithmetic sequence 0.1, 0.11, 0.12, 0.13, …

d the 13th term of the arithmetic sequence −55, −42, −29, −16, …

e the 10th term of the arithmetic sequence −1.0, −1.5, −2.0, −2.5, …

f the 95th term of the arithmetic sequence 130, 123, 116, 109, …

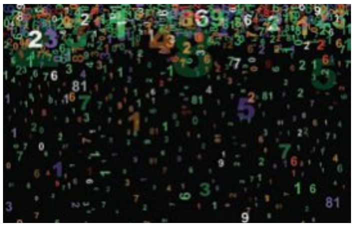
g the 7th term of the arithmetic sequence $\frac{1}{2}, \frac{1}{4}, 0, -\frac{1}{4}, \ldots$

Determining an arithmetic sequence given two terms

Example 8

7 The first term of an arithmetic sequence is 6 and $t_8 = 34$. Determine the common difference.

8 The common difference for an arithmetic sequence is 20 and $t_{10} = 188$. Determine the first term.



ISBN 978-1-108-46154-2

Cambridge University Press

CF

9 Write down the first three terms (starting with t_1) of the arithmetic sequences in which:

a the seventh term is 37 and the ninth term is 47
b the 11th term is 31 and the 15th term is 43
c the sixth term is 0 and the 11th term is −20
d the eighth term is 134 and the 13th term is 159
e the sixth term is 0 and the 10th term is −16
f the seventh term is 60 and the 12th term is 10
g the 10th term is 20 and the 21st term is 75

Determining how many terms of an arithmetic sequence are required to reach a particular number

Example 9

10 How many terms would we have to write down in the arithmetic sequence:

a 5, 7, 9, 11, … before we found a term greater than 25?
b 132, 182, 232, 282, … before we found a term greater than 1000?
c 100, 96, 92, 88, … before we found a term less than 71?
d 10, 16, 22, 28, … before we found a term equal to 52?
e 0.33, 0.66, 0.99, 1.32, … before we found a term greater than 2?
f −17, −15, −13, −11, … before we found a positive term?
g 127, 122, 117, 112, … before we found a term less than zero?

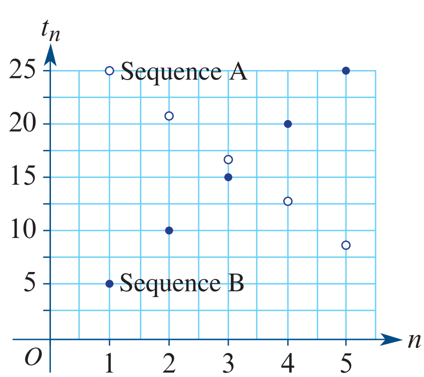
Graphing an arithmetic sequence using a rule

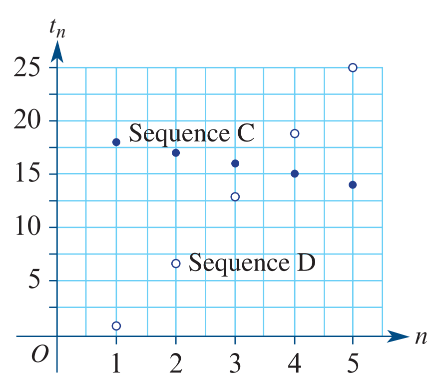
SF

11 Plot the first five terms of each of the following arithmetic sequences:

a Sequence A: $t_1 = a = 3$, $d = 2$, $n = 1, 2, 3, \ldots$
b Sequence B: $t_1 = a = 12$, $d = -2$, $n = 1, 2, 3, \ldots$
c Sequence C: $t_1 = a = 0$, $d = 4$, $n = 1, 2, 3, \ldots$
d Sequence D: $t_1 = a = -6$, $d = 3$, $n = 1, 2, 3, \ldots$
e Sequence E: $t_1 = a = 10$, $d = -5$, $n = 1, 2, 3, \ldots$

12 Four sequences are displayed in the following graphs. For each sequence determine the value of the first term, a, the common difference, d, and an expression for the nth term, t_n, in terms of n.





ISBN 978-1-108-46154-2

Cambridge University Press

CF

13 In an arithmetic sequence, the fifth term is 10 and the ninth term is 18.

a Write down the rule for the arithmetic sequence.

b Write down the first 4 terms of the sequence.

14 In an arithmetic sequence, the second term is 17 and the ninth term is 59.

a Write down the rule for the arithmetic sequence.

b Write down the first 4 terms of the sequence.

15 In an arithmetic sequence, the eighth term is 35 and the twelfth term is 23.

a Write down the rule for the arithmetic sequence.

b Write down the first 4 terms of the sequence.

SF

Example 10 **16** Consider the arithmetic sequence with $t_1 = 5$ and $d = 2$.

a Find the rule for this arithmetic sequence.

b Prepare a table of values for the sequence from $n = 1$ to $n = 5$.

c Plot a graph of the sequence from the table of values.

17 Consider the arithmetic sequence with $t_1 = 20$ and $d = -4$.

a Find the rule for this arithmetic sequence.

b Prepare a table of values for the sequence from $n = 1$ to $n = 5$.

c Plot a graph of the sequence from the table of values.

18 Consider the arithmetic sequence with $t_1 = 10$ and $d = 2$.

a Find the rule for this arithmetic sequence.

b Prepare a table of values for the sequence from $n = 1$ to $n = 5$.

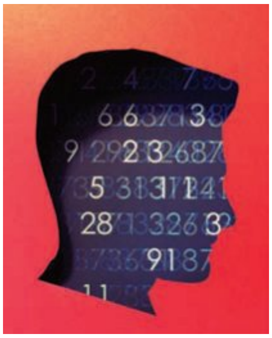
c Plot a graph of the sequence from the table of values.

19 An arithmetic sequence has rule $t_n = 5n + 2$. Find t_1 and the common difference d.

20 An arithmetic sequence has rule $t_n = 4 - 2n$. Find t_1 and the common difference d.

21 An arithmetic sequence has rule $t_n = 7n + 11$. Find t_1 and the common difference d and describe the arithmetic sequence using a recurrence relation.

22 An arithmetic sequence has rule $t_n = 8 - 5n$. Find t_1 and the common difference d and describe the arithmetic sequence using a recurrence relation.



ISBN 978-1-108-46154-2

Cambridge University Press

4D Application of arithmetic sequences

In this section we look at several situations involving linear growth and decay that can be modelled using arithmetic sequences.

Simple interest loans and investments

If you deposit money into a bank account, the bank is effectively borrowing money from you. The bank will pay you a fee for using your money and this fee is called **interest**. If a fixed amount of interest is paid into the account after regular time periods, it is called a simple interest investment. If you borrow money from the bank and are charged a fixed amount of interest after regular time periods, it is called a simple interest loan.

Simple interest is an example of an arithmetic sequence in which the starting value is the amount borrowed or invested. The amount borrowed or invested, P, is called the **principal**. The constant amount added at each step is the interest and is usually a percentage of this principal. We will denote the amount of the investment after n years as A.

If you invest \$2000 at 5% per annum, then the interest each year is \$100.

Let \$$P$ be the initial amount invested.

After one year, the value of investment = \$(2000 + 100) = \$2100

After two years, the value of investment = \$(2000 + 100 + 100) = \$2200

After three years, the value of investment = \$(2000 + 100 + 100 + 100) = \$2300

And so on.

Hence, the value A after n years is given by the rule:

$$A = 2000 + 100n$$

Calculation of simple interest

Let \$$P$ be the amount borrowed or invested (the principal).

Let \$$A$ be the value of the investment after n years.

Let i be the percentage interest rate.

Then $d = i \times P$ is the amount of interest paid per year and is the common difference.

The value A of the simple interest investment after n years is:

$$A = P + nd$$

ISBN 978-1-108-46154-2

Cambridge University Press

Example 11 Calculating the value of an investment using simple interest

David invests \$20 000 into a bank account. He will be paid simple interest at the rate of 5% of the investment per annum.

a Find the expression for A the value of the investment after n years.

b Find the value of the investment after 5 years.

c If David leaves the money in the account, after how many years will the investment be worth \$30 000?

Solution

a **1** Confirm the name of the variable to be used.	*Let A be the value of the investment after n years.*
2 The principal is \$20 000. Start at $n = 0$.	$P = 20\,000$
3 Calculate the yearly interest. This is the common difference of the arithmetic sequence.	5% of $20\,000 = 1000$ $d = 1000$
4 Substitute into the rule for A and d.	$A = 20\,000 + 1000n$
b **1** Substitute $n = 5$ into the rule.	$A = 20\,000 + 1000 \times 5$
2 Evaluate for A.	$= 25\,000$
3 Write your answer in a sentence.	*The investment will be worth \$25 000 after five years.*
c **1** Form a simple equation or see that you need 10 years of \$1000.	$20\,000 + 1000n = 30\,000$
	$1000n = 10\,000$
2 Solve for n.	$n = 10$
3 Write your answer in a sentence.	*The investment will be worth \$30 000 after 10 years.*

This is an example of linear growth.

Note: Our formula for simple interest looks different from our general formula for an arithmetic sequence. We note that if you start with $n = 0$ rather than $n = 1$, the formula is simply $t_n = t_0 + nd$, and this is in fact the form of our simple interest formula. This formula for the nth term of an arithmetic sequence will also be applicable in some of the following examples. It will also be utilised in Chapter 7 of this book.

ISBN 978-1-108-46154-2

Cambridge University Press

Depreciation

Over time, the value of large items will gradually decrease. A car bought new this year will not be worth the same amount of money in a few years time. A new television bought for \$2000 today is unlikely to be worth anywhere near this amount after a few years.

There are a number of techniques for estimating the future value of an asset. Two of them, **flat-rate depreciation** and **unit-cost depreciation**, are discussed below.

Flat-rate depreciation

Flat-rate depreciation is very similar to simple interest, but instead of adding a constant amount of interest, a constant amount is *subtracted* to decay the value of the asset after every time period. This constant amount is called the *depreciation* amount, and, like simple interest, is often given as a percentage of the initial purchase price of the asset.

In the following, A is the value after n years, P is the initial value and i is the interest rate.

If a television is valued at \$2000 and it depreciates at 5% per annum, then the decrease in value each year is \$100.

After one year, the value of the television = \$(2000 − 100) = \$1900

After two years, the value of the television = \$(2000 − 100 − 100) = \$1800

After three years, the value of the television = \$(2000 − 100 − 100 − 100) = \$1700

And so on.

Hence, the value A after n years is given by the rule:

$$A = 2000 - 100n$$

Calculation of flat-rate depreciation

Let P be the initial value of the asset.

Let A be the value of the asset after n years.

Let i be the percentage depreciation rate.

Then $d = i \times P$ is the amount of depreciation per year and is the common difference.

The value A of the asset after n years is:

$$A = P - nd$$

Example 12 Calculating flat-rate depreciation

The value of a machine is flat-rate depreciated in value by 4% of its initial value every year. Initially it was valued at \$100 000.

a Find an expression for A the value of the machine after n years.

b Find the value of the machine after 5 years.

ISBN 978-1-108-46154-2

Solution

a **1** Confirm the name of the variable to be used.
Let A be the value of the machine after n years.

2 The initial value is \$100 000. Start at $n = 0$.
$P = 100\,000$

3 Calculate the yearly depreciation. This is the common difference of the arithmetic sequence.
4% of 100 000 = 4000
$d = 4000$

4 Substitute into the rule for P and d.
$A = 100\,000 - 4000n$

b **1** Substitute in $n = 5$ into the rule.
$A = 100\,000 - 4000 \times 5$

2 Evaluate for A.
$= 80\,000$

3 Write your answer in a sentence.
The machine will be worth \$80 000 after five years.

This is an example of linear decay.

Example 13 Calculating flat-rate depreciation

A new car was bought for \$26 000 in 2014. The value of the car is expected to reduce by 4.5% of its original value every year.

a Write down the rule for the value of the car after n years.

b Find the year when the value of the car is first expected to be less than half its original value.

Solution

a **1** Find the value of the common difference d.
$d = i \times P$

$d = 4.5\% \times 26\,000$

$d = \frac{4.5}{100} \times 26\,000$

$= 1170$

2 Substitute d into rule for A.
$A = P - nd$

$A = 26\,000 - 1170n$

b **1** Half of the original value of the car is \$13 000.

$$\begin{aligned} A &< 13\,000 \\ 26\,000 - 1170n &< 13\,000 \\ 26\,000 - 13\,000 &< 1170n \\ 1170n &> 13\,000 \\ n &> 11.11\ldots \end{aligned}$$

2 We are looking for a term of the sequence that is lower than this.
You can also do this by trial and error using your calculator.

3 Write your answer.
After 12 years, the value of the car will first be less than half its original value.

ISBN 978-1-108-46154-2

Cambridge University Press

Unit-cost depreciation

Some items lose value because of how often they are used, rather than because of their age. A photocopier that is two years old but has never been used could be considered to be in 'brand new' condition and therefore worth the same, or close to, as it was two years ago. But, if the same two-year-old photocopier had printed thousands of copies over those two years, it would be worth much less than its original value.

Cars can also be depreciated according to their use rather than time. When buying a secondhand car, people often consider the number of kilometres that the car has travelled. A secondhand car that has travelled less kilometres could be considered a better buy than a new car that has travelled a large distance.

When the future value of an item is determined based upon use rather than age, we use a unit-cost depreciation method. Unit-cost depreciation can be modelled using a linear decay recurrence relation.

In the following, A is the value after n years, P is the initial value and d is the drop in its value for each use.

Calculation of unit-cost depreciation

Let P be the initial value of the asset.

Let A be the value of the asset after n uses.

Let d be the depreciation per use.

The value of the asset after n uses is:

$$A = P - nd$$

Example 14 Calculating unit-cost depreciation

A lawn mower was purchased for \$270. Every time it is used to mow a lawn, the owner estimates a depreciation in value of 50 cents in the mower's worth.

a Write down the rule for the value of the asset after n mows.

b Find the value of the lawn mower after it has mowed 25 lawns.

Solution

a We use $A = P - nd$ with $P = 270$, $d = 0.50$.

$A = 270 - 0.5n$

b **1** Substitute $n = 25$.

$A = 270 - 0.5 \times 25$
$= 257.50$

2 Write the answer in a sentence.

The value of the lawn mower is \$257.50 after 25 mows.

ISBN 978-1-108-46154-2

Other applications of arithmetic sequences

Example 15 Solving other problems using arithmetic sequences

Before starting on a weight-loss program, a man weighs 124 kg. Using a combination of diet and exercise, he plans to lose weight at a rate of 1.5 kg a week until he reaches his recommended weight of 94 kg. How many weeks will he take to reduce his weight to 94 g?

Solution

1 Write down the man's expected weight at the beginning of each week on the diet by assuming a constant weight loss of 1.5 kg per week.

Beginning of week	Weight (kilograms)
1	124.0
2	124.0 − 1.5 = 122.5
3	122.5 − 1.5 = 121.0
4	121.0 − 1.5 = 119.5
...	...

124.0, 122.5, 121.0, 119.5, ...

2 Check that his weight at the beginning of each week forms an arithmetic sequence.

$a = t_1 = 124.0$ and the common difference $d = -1.5$.

3 Write down a rule that will enable us to predict his weight, w_n, at the beginning of the nth week of his diet.

$w_n = 124.0 - 1.5(n - 1)$ for $n = 1, 2, 3 \ldots$

4 Use the rule to determine for how many weeks he must diet and exercise to reach his recommended weight of 94 kg.

$$94 = 124.0 - 1.5(n - 1)$$
$$-30 = -1.5(n - 1)$$
$$n - 1 = \frac{-30}{-1.5}$$
$$n = 21$$

5 Write the answer in a sentence.

That is, he can expect his weight to be down to 94 kg at the start of the 21st week. Therefore, he must diet and exercise for 20 weeks.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 4D

Calculating simple interest

Example 11

1 Suppose you invest \$5000 a simple interest of 7.5% per annum.

a How much interest is paid to you each year?

b Complete the following table:

End of year	1	2	3	4	5
Value (\$)					

c Write down an expression for the value of the investment at the end of the nth year.

d Determine the value of the investment:

i at the end of the 15th year

ii at the end of 25 years

SF

2 Suppose you borrow \$50 000 with simple interest rate of 9% per annum, for a period of five years.

a How much interest is charged each year?

b Complete the following table:

End of year	1	2	3	4	5
Amount (\$)					

c Write down an expression for the amount you owe at the end of the nth year.

d If you decided to extend the loan, how much would you owe:

i at the end of the 15th year?

ii at the end of the 25th year?

CF

3 Monica invests \$60 000 into a bank account. She will be paid simple interest at the rate of 4.5% per annum.

a Find the expression for \$$A$, the value of the investment after n years.

b Find the value of the investment after 5 years.

c If Monica continues to invest the money, after how many years will the investment be worth more than \$80 000?

4 Two thousand dollars is invested at an interest rate of 3.8% per annum.

a Write down the rule for the value \$$A$ for the simple interest investment.

b Use the rule to find the value of the investment after 6 years.

c Determine how many years it takes for the value of the investment to be more than \$3000.

ISBN 978-1-108-46154-2

Cambridge University Press

5 A simple interest investment of \$7000 has an interest rate of 7.4% per annum.

- **a** Write down the rule for the value of the simple interest investment A_n in terms of n.
- **b** Use this rule to find the value of the investment after 6 years.
- **c** Determine how many weeks it takes for the value of the investment to be more than \$10 000.

Calculating flat-rate depreciation

Example 12, 13

6 The value of a machine is flat-rate depreciated in value by 4% of its initial value every year. Initially it was valued at \$100 000.

- **a** Find an expression for \$$A$, the value of the machine after n years.
- **b** Find the value of the machine after 5 years.

Calculating unit-cost depreciation

Example 14, 13

7 The value of a delivery van with purchase price \$48 000 is depreciated by \$200 for every 1000 kilometres travelled.

- **a** Write a rule for the value of the delivery van after it has travelled n kilometres.
- **b** Determine the value of the van after 15 000 kilometres.
- **c** How many kilometres does it take for the value of the van to reach \$43 000?

Solving other problems using arithmetic sequences

Example 13, 15

8 To make up a set of notes, a printer charges \$2.25 for the cover and binding and an additional 2 cents per page.

- **a** Write down a rule for the cost, C_n, of making up a set of notes with n pages.
- **b** What would it cost to make up a set of notes with:
  - **i** 10 pages?
  - **ii** 35 pages?
- **c** How many pages do the notes contain if the printer charges:
  - **i** \$3.85?
  - **ii** \$6.25?

9 When a garbage truck starts collecting rubbish, it first stops at a corner store where it collects 86 kg of rubbish. It then travels down a long suburban street where it picks up 40 kg of rubbish at each house.

- **a** Write down a rule for the amount of garbage collected by the truck, g_n, after n pick-ups from houses.
- **b** How much garbage would be carried by the truck after:
  - **i** 15 pick-ups from houses?
  - **ii** 27 pick-ups from houses?
- **c** The maximum amount of garbage that can be carried by the truck is 1500 kg. After picking up from the corner store, what is the maximum number of houses it can pick up rubbish from before it is fully loaded?

ISBN 978-1-108-46154-2

Cambridge University Press

10 A coffee urn contains 15 litres of coffee. Coffee is served in 200 mL cups.

a Write down a rule for determining the amount of coffee in the urn, a_n, after n cups of coffee have been served from the urn. Assume that each cup is completely filled.

b How much coffee will be left in the urn after:

i 23 cups have been served?

ii 45 cups have been served?

c How many cups of coffee can be served from the urn if it is necessary to keep 1.5 litres of coffee in the urn for latecomers?

11 You are offered a job with a starting salary of \$20 500 per year and yearly pay rises of \$450.

a Write down a rule for determining your salary, s_n, at the start of each year you work on the job.

b What would your salary be:

i at the start of the fifth year on the job?

ii at the start of the eighth year on the job?

c At this rate, how many years would you have to be on the job to have a salary of \$50 000 per year?

12 You have \$430 to spend while on an overseas holiday. To make the money last as long as possible, you budget for \$25.75 per day.

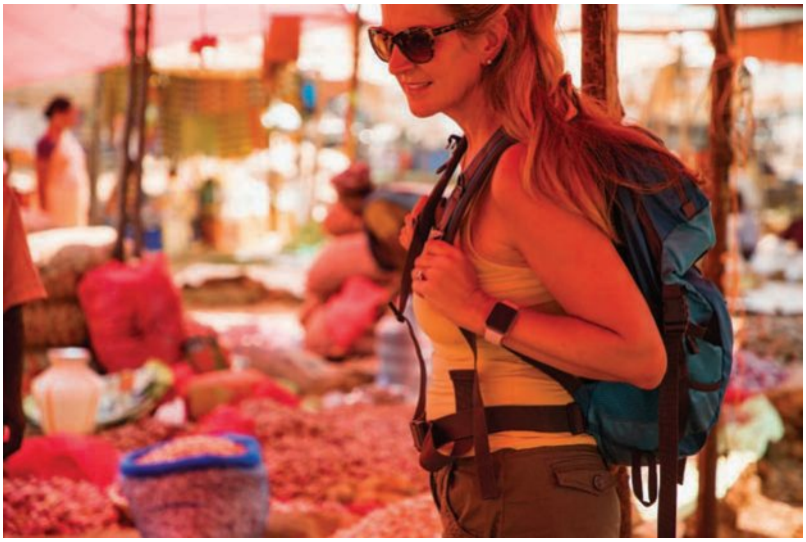
a Write down a rule for determining the amount of spending money, m_n, you will have left at the start of the nth day of your holiday.

b How much spending money would you have left:

i at the start of the seventh day?

ii at the start of the thirteenth day?

c At this spending rate, for how many days can you afford to stay on holidays?



ISBN 978-1-108-46154-2

4E Geometric sequences

A sequence in which each successive term can be found by multiplying the previous term by a constant factor is called a **geometric sequence**.

For example, the sequence 1, 2, 4, 8, 16, … is geometric because each successive term can be found by multiplying the previous term by 2.

term 1	term 2	term 3	term 4	term 5	
1	2	4	8	16	
× 2	× 2	× 2	× 2	× 2	

The sequence 40, 20, 10, 5, 2.5, … is also geometric because each successive term can be found by multiplying the previous term by 0.5.

term 1	term 2	term 3	term 4	term 5	
40	20	10	5	2.5	
× 0.5	× 0.5	× 0.5	× 0.5	× 0.5	

The common ratio

Because of the way in which a geometric sequence is formed, the ratio between successive terms is constant. In the language of geometric sequences, we call it the *common ratio*. In the sequence:

1, 2, 4, 8, 16, …

the common ratio is 2, while in the sequence:

40, 20, 10, 5, 2.5, …

the common ratio is 0.5.

Common ratiows can also be negative; for example, the common ratio for the geometric sequence 1, −2, 4, −8, 16, … is −2.

Method of recursion to generate a geometric sequence

The method for using recursion to generate a geometric sequence has two parts.

1. A starting point, the value of the first term $t_1 = a$ of the sequence.
2. Successively multiply each term by the common ratio r to obtain the next term.

This will generate the sequence.

For example, in words, a rule for the recursion that can be used to generate the sequence 10, 20, 40, 80, ... can be written as follows:

1. Start with 10.
2. To obtain the next term, multiply the current term by 2 and repeat the process.

ISBN 978-1-108-46154-2

A more compact way of communicating this information is to translate this rule into symbolic form.

This notation gives us a way to describe the generation of a geometric sequence by recursion.

$t_1 = a$ and $t_{n+1} = r \times t_n$, where r is the constant ratio.

This is the general recurrence relation for a geometric sequence.

Note: This also gives us a way to test whether a sequence is geometric or not.

We must have:

$$\frac{t_{n+1}}{t_n} = r$$

for all whole numbers $n > 1$.

Example 16 **Generating a geometric sequence with a recurrence relation**

A geometric sequence is defined by $t_1 = 5$, $t_{n+1} = 2t_n$.

Find the first five terms.

Solution

1 Write down the starting term.	$t_1 = 5$
2 Apply the rule (multiply by 2) to generate the next term.	$t_2 = 2 \times t_1 = 2 \times 5 = 10$
3 Calculate three more terms.	$t_3 = 2 \times t_2 = 2 \times 10 = 20$ $t_4 = 2 \times t_3 = 2 \times 20 = 40$ $t_5 = 2 \times t_4 = 2 \times 40 = 80$
4 Write your answer.	5, 10, 20, 40, 80

Graphing a geometric sequence

Consider the arithmetic sequences defined by:

- $V_1 = 15$, $V_{n+1} = 3V_n$
- $V_1 = 4$, $V_{n+1} = 0.5V_n$

The first generates the geometric sequence 15, 45, 135, …

It is an *increasing sequence*, which can be used to model **geometric growth**.

The second generates the geometric sequence 4, 2, 1, …

It is a *decreasing sequence*, which can be used to model **geometric decay**.

ISBN 978-1-108-46154-2

Cambridge University Press

Sequence 1

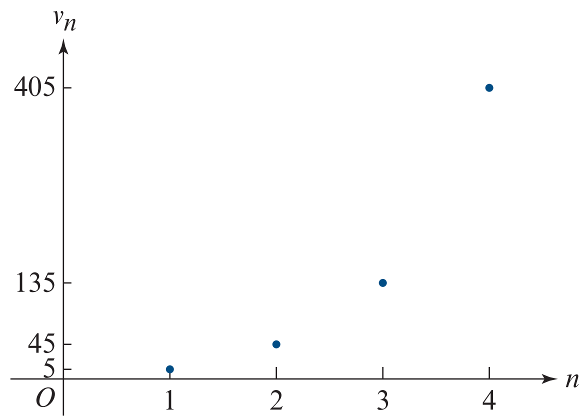
The first few values of the sequence can be shown in a table.

The first four values of Sequence 1 are shown in the table:

n	1	2	3	4
V_n	$V_1 = 15$	$V_2 = 45$	$V_3 = 135$	$V_4 = 405$

and this gives ordered pairs (1, 15), (2, 45), (3, 135), (4, 405).

These can be graphed as shown opposite.



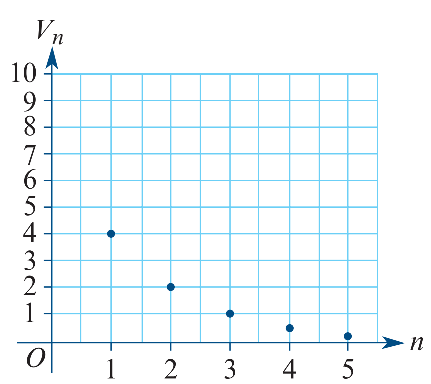
Sequence 2

The first five values of Sequence 2 are shown in the table:

n	1	2	3	4	5
V_n	$V_1 = 4$	$V_2 = 2$	$V_3 = 1$	$V_4 = \frac{1}{2}$	$V_5 = \frac{1}{4}$

and this gives ordered pairs (1, 4), (2, 2), (3, 1), $\left(4, \frac{1}{2}\right)$, $\left(5, \frac{1}{4}\right)$.

These can be graphed as shown opposite.



ISBN 978-1-108-46154-2

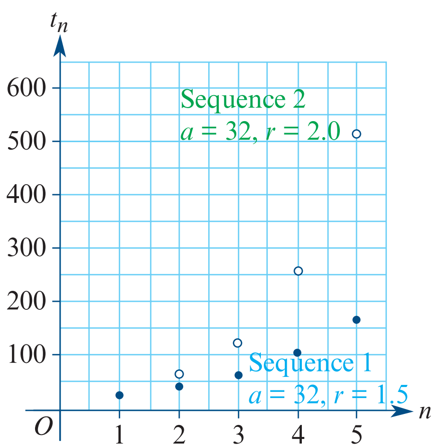
In general:

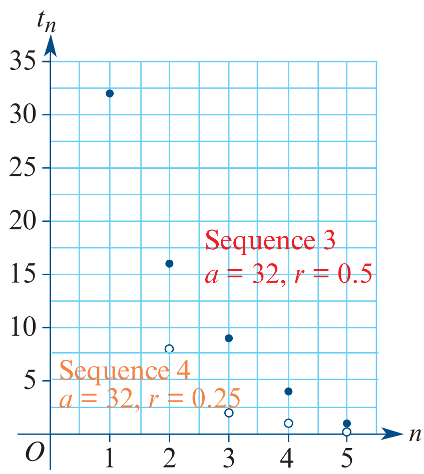
- if $r > 1$, the recurrence relation $t_{n+1} = rt_n$ can be used to model geometric growth
- if $r < 1$, the recurrence relation $t_{n+1} = rt_n$ can be used to model geometric decay.

Using a graph to investigate the pattern of growth in a geometric sequence

The graphs below display the terms in four different geometric sequences:

- Sequence 1: $t_1 = a = 32$, $r = 1.5$, $n = 1, 2, 3, \ldots$
- Sequence 2: $t_1 = a = 32$, $r = 2.0$, $n = 1, 2, 3, \ldots$
- Sequence 3: $t_1 = a = 32$, $r = 0.5$, $n = 1, 2, 3, \ldots$
- Sequence 4: $t_1 = a = 32$, $r = 0.25$, $n = 1, 2, 3, \ldots$





The key characteristics to note are:

- that the points in the graphs are not *collinear* but lie on what is called an *exponential curve*.
- if the common ratio, r, is *greater than* 1, the terms in the sequence *increase* in value. The bigger the value of r, the more rapid the increase.
- if the common ratio, r, is *less than* 1, the terms in the sequence *decrease* in value. The closer the value to 0 the more rapid the decrease.

Note: When r is negative (not shown), terms oscillate between positive and negative values.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 17 Graphing geometric sequences

Prepare a table of values and plot the first five terms to illustrate the sequence defined by:

$t_1 = 400 \quad t_{n+1} = 0.75t_n$

Solution

1 Use the recursion relation $t_{n+1} = 0.75t_n$ to complete the table.

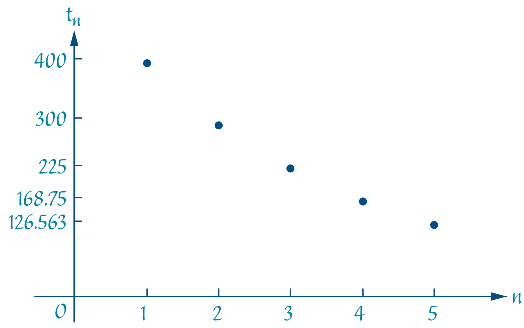
n	1	2	3	4	5
t_n	400	300	225	168.75	126.5625

Start with $t_1 = 400$.

$t_2 = 0.75t_1 = 0.75 \times 400 = 300$

$t_3 = 0.75t_2 = 0.75 \times 300 = 225$

2 Use the values in the table to plot the ordered pairs (1, 400), (2, 300), (3, 225), …



Exercise 4E

Generating a geometric sequence with a recurrence relation

Example 16

SF

1 A geometric sequence is defined by:

$t_1 = 3, \; t_{n+1} = 4t_n$

Find the first five terms.

2 A geometric sequence is defined by:

$t_1 = 15, \; t_{n+1} = 2t_n$

Find the first five terms.

3 Consider the geometric sequence 2, 20, 200, 2000, …

- **a** What is the common ratio?
- **b** What is the next term in the sequence?
- **c** Starting with 2, how many times do you have to multiply by the common ratio to get to term 5?
- **d** Starting with 2, how many times do you have to multiply by the common ratio to get to term 15? What is the value of term 15? Give your answer using scientific notation.

ISBN 978-1-108-46154-2

Cambridge University Press

4 Consider the geometric sequence 1024, 256, 64, 16, …

a What is the common ratio?

b What is the next term in the sequence?

c Starting with 1024, how many times do you have to multiply by the common ratio to get to term 7? What is the value of term 7?

d What is the value of term 10? What is the value of term 50?

5 Consider the sequence −1, 5, −25, 125, …

a What is the common ratio?

b What is the next term in the sequence?

c Starting with −1, how many times do you have to multiply by the common ratio to get to term 6? What is the value of term 6?

d Starting with −1, how many times do you have to multiply by the common ratio to get to term 15? What is the value of term 15?

Graphing a geometric sequence

Example 17

6 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 9, \; t_{n+1} = 2t_n$$

7 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 12, \; t_{n+1} = \frac{1}{2}t_n$$

8 Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 4, \; t_{n+1} = 4t_n$$

ISBN 978-1-108-46154-2

4F A general rule for finding the *n*th term of a geometric sequence

As for arithmetic sequences, we use t_n to indicate the *n*th term in the sequence and *a* to represent the *first term* in the sequence.

Finally, we use the symbol *r* to represent the common ratio. For the sequence 2, 4, 8, 16, … the common ratio *r* equals 2.

The rule for the *n*th term of a geometric sequence

Consider a geometric sequence with first term *a* and common ratio *r*.

Then:

$$t_1 = a$$
$$t_2 = t_1 \times r = ar$$
$$t_3 = t_2 \times r = ar^2$$
$$t_4 = t_3 \times r = ar^3$$
$$t_5 = t_4 \times r = ar^4$$
$$t_6 = t_5 \times r = ar^5$$

and so on.

The rule for the *n*th term of a geometric sequence

Thus, following the pattern, we can write:

$$t_n = ar^{n-1}$$

which gives us a rule for finding the *n*th term of a geometric sequence in terms of the first term, $t_1 = a$, and the common ratio, *r*.

Example 18 Using the general rule for a geometric sequence

The first term of a geometric sequence is $t_1 = a = 6$ and the common ratio is $r = 2$.

Use the rule to determine the value of the seventh term.

Solution

Substitute $t_1 = a = 6$ and $r = 2$ in the rule $t_n = ar^{n-1}$.

$$t_7 = 6 \times 2^{7-1}$$
$$= 6 \times 2^6$$
$$= 384$$

ISBN 978-1-108-46154-2

Example 19 Using the general rule for a geometric sequence

Use the rule to determine the value of the eighth term of the geometric sequence 100, 50, 25, 12.5, …

Solution

For this sequence $t_1 = a = 100$ and $r = 0.5$.
Substitute into the rule $t_n = ar^{n-1}$.

$$t_8 = 100 \times 0.5^{8-1}$$
$$= 100 \times 0.5^7$$
$$= 0.78125$$

Example 20 Determining a geometric sequence given two terms

In a geometric sequence, the fourth term is 24 and the ninth term is 768. Write down the first three terms of the sequence.

Solution

1 Substitute $t_4 = 24$ in $t_n = ar^{n-1}$ with $n = 4$.
$$24 = ar^{4-1} = ar^3 \quad (1)$$

2 Substitute $t_9 = 768$ in $t_n = ar^{n-1}$ with $n = 9$.
$$768 = ar^{9-1} = ar^8 \quad (2)$$

3 Find a and r by solving the simultaneous equations (1) and (2).
Divide (2) by (1).
$$\frac{768}{24} = \frac{ar^8}{ar^3}$$
$$32 = r^5$$
$$r = 2$$

4 Substitute $r = 2$ in equation (1).
$$24 = a \times 2^3$$
or $24 = 8a$ so $a = 3$

5 Write the first three terms of the sequence.
The first three terms of the sequence are 3, 6, 12.

Example 21 Determining how many terms of a geometric sequence are required to reach a particular number

How many terms would we have to write down in the geometric sequence 0.5, 5, 50, 500, … before we found a term greater than 1 000 000?

Solution

Substitute $a = 0.5$ and $r = 10$ in the general rule for a geometric sequence.

$$0.5 \times 10^{n-1} > 1\,000\,000$$
$$10^{n-1} > 2\,000\,000$$
$$n - 1 > 6$$
$$n > 7$$

We want to find n so that:
$t_n = 0.5 \times 10^{n-1} > 1\,000\,000$

Solve by trial and error with your calculator.

$10^2 = 100,\ 10^3 = 1000,\ \ldots, 10^6 = 1\,000\,000,$
$10^7 = 10\,000\,000$

We would have to write down eight terms to find the first term that exceeds 1 000 000.

ISBN 978-1-108-46154-2

Graphing a geometric sequence

In investigating the properties of geometric sequences, particularly the rate at which the terms in the series either grow or decay, it is useful to have a graphical representation of the sequence.

Example 22 Graphing a geometric sequence

Consider the geometric sequence with $t_1 = a = 3$ and the common ratio $r = 2$.

a Find the rule for this geometric sequence.

b Prepare a table of values for the sequence for $n = 1$ to $n = 5$.

c Plot a graph from the table of values.

Solution

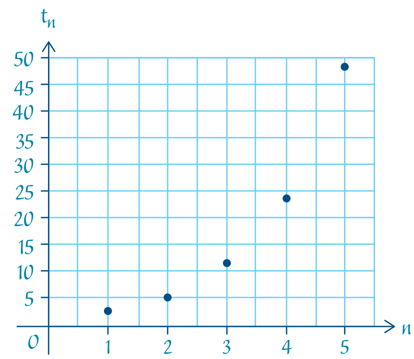
a Find an expression for the nth term, t_n, in terms of n.

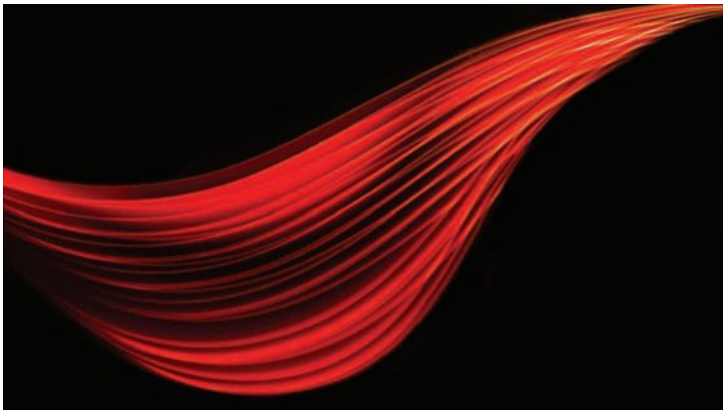
$t_n = 3 \times 2^{n-1}$ for $n = 1, 2, 3, \ldots$

b Generate a table of values using the rule. For example, $t_3 = 3 \times 2^{3-1} = 12$.

n	1	2	3	4	5
t_n	3	6	12	24	48

c Plot the points on a graph with n on the horizontal axis and t_n on the vertical axis. Do not join up the points as the terms in the sequence are only defined for $n = 1, 2, 3, \ldots$





ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 4F

Using the rule to generate a geometric sequence

Example 18

1 The first term of a geometric sequence is $a = 1$ and the common ratio is $r = 4$. Use the rule to determine the value of:

- **a** the third term
- **b** the seventh term
- **c** the 15th term

SF

2 A geometric sequence is described by the recurrence relation $t_1 = 8$ and $t_{n+1} = 5t_n$. Write the rule for the nth term of this sequence.

3 A geometric sequence is described by the recurrence relation $t_1 = 5$ and $t_{n+1} = 0.6t_n$. Write the rule for the nth term of this sequence.

Example 19

4 Use the rule to determine the value of:

- **a** the seventh term of the geometric sequence 1, 5, 25, 125, 625, …
- **b** the eighth term of the geometric sequence 10 000, 2000, 400, …
- **c** the 10th term of the geometric sequence −1, 2, −4, 8, …
- **d** the ninth term of the geometric sequence −20, −60, −180, …
- **e** the eighth term of the geometric sequence $2, 1, \frac{1}{2}, \dots$
- **f** the fifth term of the geometric sequence 110, 121, 133.1, …
- **g** the seventh term of the geometric sequence $1, -\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \dots$

Determining a geometric sequence given two terms

Example 20

CF

5 Write down the first three terms of the geometric sequence in which:

- **a** the fifth term is 81 and the eighth term is 2187
- **b** the second term is 10 000 and the fifth term is 1250
- **c** the third term is 40 and the sixth term is −320
- **d** the second term is 160 and the fourth term is 250 ($r > 0$)

Determining how many terms of a geometric sequence are required to reach a particular number

Example 21

6 How many terms would we have to write down in the geometric sequence:

- **a** 2, 4, 8, 16, … before we found a term greater than 250?
- **b** 1, 1.1, 1.21, … before we found a term greater than 2?
- **c** 100, 80, 64, … before we found a term less than 10?
- **d** −8, −16, −32, … before we found a term equal to −4096?
- **e** 0.9, 0.81, 0.729, … before we found a term less than 0.1?
- **f** 2000, 2100, 2205, … before we found a term greater than 4000?
- **g** 6000, 5700, 5415, … before we found a term less than 3000?

ISBN 978-1-108-46154-2

Graphing geometric sequences

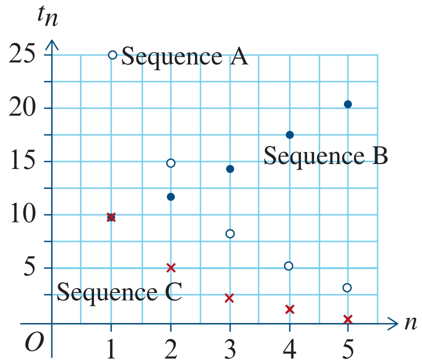
Example 22

7 Plot the first five terms of each of the following geometric sequences.

- **a** Sequence A: $t_1 = a = 1,\ r = 2,\ n = 1, 2, 3, \ldots$
- **b** Sequence B: $t_1 = a = 100,\ r = 0.5,\ n = 1, 2, 3, \ldots$
- **c** Sequence C: $t_1 = a = 1024,\ r = 0.75,\ n = 1, 2, 3, \ldots$
- **d** Sequence D: $t_1 = a = 32,\ r = 1.5,\ n = 1, 2, 3, \ldots$
- **e** Sequence E: $t_1 = a = 1024,\ r = -0.25,\ n = 1, 2, 3, \ldots$

8 Three sequences are displayed in the graphs below. For each sequence:

- **a** determine the value of the first term, a
- **b** determine from the trend of the points whether the value of r is greater than 1 or between 0 and 1
- **c** use the first two points to estimate the value of r



9 The sequence defined by the rule for the nth term, $t_n = 2^n$ is a geometric sequence.

- **a** List the first four terms of the sequence.
- **b** State the value of the common ratio.

SF

ISBN 978-1-108-46154-2

Cambridge University Press

4G Applications of geometric sequences

Populations grow in size and genuine antiques gain value in time while most cars depreciate in value with time. These are examples of growth and decay of quantities that can be modelled with a geometric sequence.

Many situations that we experience in our everyday lives can be mathematically represented, or **modelled**, with geometric sequences.

Example 23 Applying of geometric sequence

A dish in a laboratory contains 150 000 bacteria. The population of bacteria is expected to double in size every day.

a Using recursion, find the number of bacteria after 1, 2, 3 and 4 days.

b Write a rule for the number, N, of bacteria after n days.

c Use the rule to find the number of bacteria after 7 days.

Solution

a The number of bacteria is doubled each day. Therefore, we multiply by 2 each time to find the number of bacteria the next day.

Number of bacteria at end of day one
$= 2 \times 150\,000 = 300\,000$
Number of bacteria at end of day two
$= 2 \times 300\,000 = 600\,000$
Number of bacteria at end of day three
$= 2 \times 600\,000 = 1\,200\,000$
Number of bacteria at end of day four
$= 2 \times 1\,200\,000 = 2\,400\,000$

b Multiplying by 2 n times results in a factor of 2^n.

$N = 150\,000 \times 2^n$

c Substitute $n = 7$ in rule.

$N = 150\,000 \times 2^7$
$= 19\,200\,000$

Percentage increase and decrease

The number of koalas in a reserve might increase by 5% each year. If there are 400 koalas this year, we can expect there to be an extra 5% next year, in addition to the 400 koalas that are already there.

This year: number of koalas = 400

$$\begin{aligned}\text{Next year: number of koalas} &= 400 + 5\% \text{ of } 400\\ &= 400 + \frac{5}{100} \times 400\\ &= 400 + 20\\ &= 420 \text{ koalas}\end{aligned}$$

ISBN 978-1-108-46154-2

There is an easier way to calculate the number of koalas next year.

If this year we have 400 koalas, this represents ALL of the koalas currently in the reserve, or 100% of them. If the number of koalas grows by 5% in one year, then next year there will be 100% of the koalas plus an extra 5% in the reserve. Next year there will be 105% of this year's number.

This year: number of koalas = 400

Next year:
$$\begin{aligned}\text{number of koalas} &= 105\% \text{ of } 400\\ &= \frac{105}{100} \times 400\\ &= 1.05 \times 400\\ &= 420 \text{ koalas}\end{aligned}$$

If, on the other hand, the koala population is decaying by 5% each year, next year there would be 5% less koalas than this year. This would be 100% − 5% or 95% of this year's number of koalas.

The *rate* of geometric growth and decay is often given in the form of a percentage of the starting term.

As usual we use r for the common ratio in these problems and $i\%$ for the percentage growth.

Percentage increase and decrease

If the percentage *increase* for geometric growth is $i\%$, then $r = 1 + i\%$.

If the percentage *decrease* for geometric decay is $i\%$, then $r = 1 - i\%$.

Compound interest

Most interest calculations are not as straightforward as simple interest. The more usual form of interest is **compound interest** where any interest that is earned after one time period is added to the principal and then contributes to the earning of interest in the next time period. This means that the value of the investment grows in increasing amounts, or grows geometrically, instead of by the same amount as in simple interest.

Consider the investment of $5000 into an account that pays interest of 8% per annum. The interest will be paid into the account after each year and this interest will effectively be re-invested and will earn interest in the next year. When the interest is itself earning interest, we say the interest is *compounding*.

ISBN 978-1-108-46154-2

Cambridge University Press

We can model this compound interest investment with a recurrence relation. The sequence of numbers that represent the value of the investment from year to year will have a starting term 5000. The value of the investment will *grow* at a rate of 8% each year, so the multiplying factor will be:

$$r = 1 + i\%$$
$$r = 1.08$$

\$5000 is the initial amount invested

After one year the value of investment $= \$(5000 \times 1.08) = \$5400$

After two years the value of investment $= \$(5000 \times 1.08 \times 1.08) = \$5832$

After three years the value of investment $= \$(5000 \times 1.08 \times 1.08 \times 1.08) = \$6298.56$

And so on.

Hence, the value \$$A$ after n years is given by the rule:

$$A = 5000 \times 1.08^n$$

Calculation of compound interest loans and investments

Let A be the value of a compound interest loan or investment after n years.

Let P be the amount borrowed or invested (principal).

Let i be the annual interest rate of the loan or investment.

$$A = P \times r^n$$

where $r = 1 + i$

Example 24 Calculating compound interest

An amount of \$2000 is invested with compound interest at an interest rate of 7.5% per annum.

a Find the rule for the value of the investment after n years.

b Find the value of the investment after 4 years.

c Determine when the value of the investment will first exceed \$3000.

Solution

a Substitute the values for P and r into the formula $A = P \times r^n$ with $r = 1 + i$

$P = 2000$ and $r = 1 + \frac{7.5}{100} = 1.075$

$A = 2000 \times 1.075^n$

b Substitute $n = 4$.

$A = 2000 \times 1.075^4$
$= 2670.938\ldots$

The value of the investment is \$2670.94 after 4 years.

ISBN 978-1-108-46154-2

Solution

c Work with trial and error to determine when the investment will exceed \$3000.

After 5 years, $A = 2000 \times 1.075^5 = 2871.258\ldots$

After 6 years, $A = 2000 \times 1.075^6 = 3086.60\ldots$

After 6 years the investment will exceed \$3000.

Reducing-balance depreciation

Earlier in the chapter, we studied two different methods for depreciating the value of an asset, both of which were examples of linear decay. **Reducing-balance depreciation** is another method of depreciation, one where the value of an asset decays geometrically. Each year, the value will be reduced by a percentage, $i\%$, of the previous year's value. The calculations are very similar to compounding interest, but with decay in value, rather than growth.

Calculation of reducing-balance depreciation

Let A be the value of the asset after n years.

Let P be the initial cost of the asset.

Let i be the annual interest rate of the loan or investment.

$$A = P \times r^n$$

where $r = 1 - i\%$

Example 25 Calculating reducing-balance depreciation

An item of office furniture has a purchase price of \$6900. It can be considered to be depreciating at a reducing-balance rate of 8.4% per annum.

a Find the rule for the value of the office furniture after n years.

b Find the value of the office furniture after 4 years.

c Determine when the value of the office furniture will be less than \$3000.

Solution

a **1** Write down the purchase price of the furniture, A.

$A = 6900$

2 Calculate the value or r.

The depreciation rate is 8.4% per annum.

$$r = 1 - \frac{8.4}{100}$$
$$= 0.916$$

3 Use the rule $A = P \times r^n$

$A = 6900 \times 0.916^n$

b Substitute $n = 4$ in rule.

$$A = 6900 \times 0.916^4$$
$$= 4857.7033\ldots$$

Answer to the nearest cent

The value of the office furniture after 4 years is \$4857.70.

ISBN 978-1-108-46154-2

Cambridge University Press

c Use trial and error to determine when the value of the furniture will be less than \$3000 by substituting values of n in the rule $A = 6900 \times 0.916^n$.

When $n = 5$, $A = 6900 \times 0.916^5$
$= 4449.656\ldots$

When $n = 6$, $A = 6900 \times 0.916^6$
$= 4075.885\ldots$

When $n = 7$, $A = 6900 \times 0.916^7$
$= 3733.510\ldots$

When $n = 8$, $A = 6900 \times 0.916^8$
$= 3419.895\ldots$

When $n = 9$, $A = 6900 \times 0.916^9$
$= 3132.624\ldots$

When $n = 10$, $A = 6900 \times 0.916^{10}$
$= 2869.484\ldots$

After 10 years the value is less that \$3000.

Note: Be careful if the question asks for the 'value' at the beginning of the year.

The same idea can be applied to other situations where an amount is decreasing by a contant percentage for each period of time.

Example 26 Applying geometric sequences

A swimming pool is filled with 200 000 litres of water.

Under certain conditions, 2% of the water in the pool will evaporate each day.

a Determine the rule for the amount of water in the pool at the end of the nth day.

b Find the amount of water in the pool at the end of the eighth day.

Solution

a **1** Determine the value of r.

Let V_n be the volume of water in the pool after n days.
The water volume decays by 2% every day.

$$r = 1 - \frac{2}{100}$$
$$= 0.98$$

2 Substitute the value r into the rule $V_n = 200000 \times r^n$

$V_n = 200\,000 \times 0.98^n$

b $V_8 = 200000 \times r^8$

$V_8 = 200\,000 \times 0.98^8$
$= 170\,153$ litres to the nearest litre

ISBN 978-1-108-46154-2

Interest rates over different periods

Compound interest rates are usually quoted as annual rates, or interest rate per annum. The time period after which compound interest is calculated and paid is called the compounding period. For example, the compounding period could be months or days.

An interest rate given per annum can, by convention, be converted to a compounding interest rate for a shorter period by *dividing* this interest rate per annum by the appropriate number.

For example, an interest rate of 3.6% per annum gives a monthly interest rate of $(3.6 \div 12) = 0.3\%$ per month.

Note: Problems involving interest rates for compounding periods other than a year appear in the review section of this chapter. This is discussed in more detail in Chapter 7.

Exercise 4G

Applying geometric sequences

Example 23

SF

1 The following rule can be used to model the number of shares an investor owns after n months, if the investor sells 4% of the shares he owns after every month and the investor initially owns P shares.

$$A = P \times 0.96^n$$

where A is the number of shares owned by the investor after n months.

- **a** How many shares does the investor have after m years?
- **b** Write down the rule for the number of shares owned after n months if the investor owned 30 000 shares and sold 3.5% of the shares owned after every month.

2 The population of kangaroos in a national park is increasing by 5% every year. There are currently 2700 kangaroos in the national park.

- **a** Write down the rule for the number of kangaroos after n years.
- **b** How many kangaroos are there after 5 years?
- **c** How many years does it take for the kangaroo population to double?

There are also 830 wombats in the park. Their numbers are increasing by 4% every year.

- **d** How many wombats are there in the park after 4 years?
- **e** How long does it take for the number of wombats in the park to double?

ISBN 978-1-108-46154-2

Cambridge University Press

Calculating compound interest

Example 24

3 An investment of \$6000 earns compounding interest at a rate of 4.2% per annum. Write down the value of the investment after n years.

SF

4 A loan of \$20 000 is charged compounding interest at the rate of 6.3% per annum.

CF

a Write down the value of the loan after n years.

b Determine how many years it takes for the value of the loan to first exceed \$30 000.

c Write down the rule for the value of a loan of \$18 000 at a compounding interest rate of 9.4% per annum after n years.

Calculating reducing-balance depreciation

Example 25

5 A motorcycle was purchased new for \$9800 and is depreciated using a reducing-balance depreciation method with an annual depreciation rate of 3.5%.

SF

a Determine the rule for the value of the motorcycle after n years.

b What is the value of the motorcycle after 5 years?

c What is the depreciation of the motorcycle in the third year?

6 An investment of \$8000 earns 12.5% compound interest each year.

CF

a Write down a rule for the value of the investment after n years.

b Use the rule to find the value of the investment after three years.

c How much interest has been earned over three years?

d How much interest was earned in the third year of the investment?

7 A loan of \$3300 is charged 7.5% compound interest each year.

a Write down a rule for the amount owed after n years.

b Use the rule to find the value of the loan after 10 years.

c How much interest has been charged over 10 years?

d How much interest was charged in the 10th year of the investment?

8 A stereo system, initially valued at \$1200, is depreciated using reducing-balance depreciation of 12%.

a Write down a rule for the value of the stereo system after n years.

b Use the rule to find the value of the stereo system after seven years.

ISBN 978-1-108-46154-2

Applying geometric sequences

Example 26

9 A ball is dropped vertically from a tower 3.6 metres high and the height of its rebound is recorded for four successive bounces. The results are shown in the table below:

Bounce number	0	1	2	3
Height (centimetres)	360.00	270.00	202.50	151.875

a Do the heights of the bouncing ball given in the table form a geometric sequence? Explain.

b Assuming that the height of the bouncing ball follows a geometric sequence:

i predict the height of the fourth bounce

ii write down an expression for the height of the nth bounce

iii predict the height of the 15th bounce

10 Suppose a newly discovered virulent bacteria replicates itself every five minutes.

a If we start off with 10 bacteria, how many bacteria will there be after five minutes?

b Label the first five minutes Time period 1, the second five minutes Time period 2, etc. Complete the following table.

Time period	1	2	3
Number of bacteria	10		

c Write down an expression for the number of bacteria at the start of the nth time period.

d Determine the number of bacteria:

i at the start of the fifth time period

ii after 30 minutes

iii after 1 hour

11 A fish population increases its size by 40% every six months, provided the conditions are ideal.

a Starting with a population of 1000 fish, how many fish would there be after six months?

b Label the first six months Time period 1, the second six months Time period 2, etc. Complete the following table.

Time period	1	2	3	4
Number of fish	1000			

c Write down an expression for the number of fish in the population at the start of the nth time period.

d Determine the number of fish:

i at the start of the seventh time period

ii after 5 years

iii after 10 years

ISBN 978-1-108-46154-2

Cambridge University Press

12 Suppose a car costs \$20 000 when new. Assume that it loses 7.5% of its value each year.

a What is the value of the car at the start of the second year?

b Label the first year Time period 1, the second year Time period 2, etc. Complete the following table.

Time period	1	2	3	4
Value of car (\$)	20 000			

c Write down an expression for the value of the car at the start of the nth year.

d Determine the value of the car:

i at the start of the seventh year of its life

ii after 8 years

iii after 15 years

13 When a man purchases an antique table for \$12 000, he is told its value will increase by 125% every 25 years.

a Making this assumption, how much do you expect the table to be worth after 25 years?

b Label the first 25 years Time period 1, the second 25 years Time period 2, etc. Complete the following table.

Time period	1	2	3
Value of antique table (\$)	12 000		

c Write down an expression for the value of the table at the start of the nth time period.

d Determine the value of the table:

i at the start of the fifth time period

ii after 150 years

iii after 250 years

ISBN 978-1-108-46154-2

Key ideas and chapter summary

AS

Sequence

A **sequence** is a list of numbers or symbols written in succession. For example, 5, 15, 25, …
Sequences often have patterns that mean we can write rules and predict the **terms** that make up the sequence.

Term

Each number or symbol that makes up a sequence is called a **term**.

Recursion

Recursion involves repeating the same calculation over and over, using the previous result to calculate the next result.

Recurrence relation

Recurrence relations define the terms of a sequence using recursive calculations. The rule of the recurrence relation relies on one term in the sequence, t_n, to generate the next term in the sequence, t_{n+1}. The recurrence relation must show the starting value, t_1, and the rule.

Modelling

Modelling is the use of a mathematical rule or formula to represent or model real-life situations. Recurrence relations can be used to model situations involving the **growth** (increase) or **decay** (decrease) in values of a quantity.

Arithmetic sequences

A sequence is **arithmetic** if it satisfies the recurrence relation:
$t_{n+1} = t_n + d$ and a starting point usually t_1.
Arithmetic sequences are used to model linear growth and linear decay situations.
The rule for the nth term of an arithmetic sequence is:

$$\begin{aligned} t_n &= t_1 + (n - 1)d \\ &= a + (n - 1)d, \text{ where } a = t_1 \text{ is the starting value.} \end{aligned}$$

Linear growth

When a recurrence relation rule involves adding a constant amount, d, to each term, the terms of the sequence will increase uniformly through the sequence. The terms will grow linearly.
Linear growth can be modelled by the recurrence relation or
$a = t_1 =$ starting value, $t_{n+1} = t_n + d$
where d is a positive real number.

 ISBN 978-1-108-46154-2 Cambridge University Press

Linear decay	When a recurrence relation rule involves subtracting a constant amount from each term, the terms of the sequence will decrease uniformly through the sequence. The terms will decay linearly. **Linear decay** can be modelled by the recurrence relation t_1 = starting value, $t_{n+1} = t_n - d$ where d is a positive real number.
Linear growth and decay rule	The value of a quantity that grows or decays linearly can be found using the general rule $t_n = t_1 + (n - 1)d$, where t_1 is the starting value, t_n is the value of the quantity after n steps and d is a real number.
Geometric sequences	A sequence is **geometric** if it satisfies the recurrence relation: $t_{n+1} = r \times t_n$ and a starting point usually t_1. Geometric sequences are used to model geometric growth and linear decay situations. The rule for the nth term of a geometric sequence is: $t_n = r^{n-1}t_1 = r^{n-1}a$, where $a = t_1$ is the starting value.
Geometric growth	When a recurrence relation rule involves multiplying by a factor, r, that is larger than 1, the terms of the sequence increase through the sequence. The terms will grow geometrically. **Geometric growth** can be modelled by the recurrence relation $t_1 = a$ = starting value, $t_{n+1} = r \times t_n$, where $r > 1$.
Geometric decay	When a recurrence relation rule involves multiplying by a factor that is smaller than 1, the terms of the sequence decrease through the sequence. The terms will decay geometrically. **Geometric decay** can be modelled by the recurrence relation $t_1 = a$ = starting value, $t_{n+1} = r \times t_n$, where $r < 1$.
Geometric growth and decay rule	The value of a quantity that grows or decays geometrically can be found using the general rule $t_n = r^{n-1} \times t_1 = ar^{n-1}$ where t_n is the value of the quantity after n steps.

ISBN 978-1-108-46154-2

Principal	The **principal** is the initial amount that is invested or borrowed.
Balance	The value of a loan or investment at any time during the loan or investment period is the **balance**.
Interest	**Interest** is the fee that is added to a loan or the payment for investing money.
Simple interest	**Simple interest** is a fixed amount of interest that is paid at regular time intervals. Simple interest is an example of linear growth.
Depreciation	**Depreciation** is the amount by which the value of an item decreases after a period of time.
Flat-rate depreciation	A constant amount that is subtracted from the value of an item at regular time intervals. **Flat-rate depreciation** is an example of linear decay.
Unit-cost depreciation	Depreciation that is calculated based on units of use rather than time. **Unit-cost depreciation** is an example of linear decay.
Compounding period	Interest rates are often quoted as annual rates (per annum). Interest is sometimes calculated more regularly than each year, for example, each quarter, month, fortnight, week or day. The time period for the calculation of interest is called the **compounding period**.
Compound interest	When interest is added to a loan or investment and then contributes to earning more interest, the interest is said to compound. **Compound interest** is an example of geometric growth.
Reducing-balance depreciation	When the value of an item decreases as a percentage of its value after each time period, it is said to be depreciating using a reducing-balance method. **Reducing-balance depreciation** is an example of geometric decay.

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Skills check

Having completed this chapter, you should be able to:

- generate a sequence of terms from a recurrence relation by hand and using calculator recursion
- identify arithmetic sequences
- define a recurrence relation for an arithmetic sequence
- work with the formula $t_n = a + (n - 1)d$ for the nth term of an arithmetic sequence
- determine the value of simple interest loans and investments
- determine the value of items depreciating using flat-rate depreciation
- determine the value of items depreciating using unit-cost depreciation
- identify geometric sequences
- define a recurrence relation for a geometric sequence
- work with the formula $t_n = ar^{n-1}$ for the nth term of an geometric sequence
- determine the value of compound interest loans and investments
- determine the value of items using reducing-balance depreciation.

Multiple-choice questions

1 Which of the following could be the first five terms of an arithmetic sequence?

A 2, 4, 2, 4, 2
B 1, 10, 100, 1000, 10000
C −189, −89, 11, 111, 211
D 1, 4, 9, 16, 25
E 4, 4, 6, 6, 8

2 Which of the following is not an arithmetic sequence?

A 11, 2, –8, –19, ...
B 4, 7, 10, 13, ...
C 57, 51, 45, 39, ...
D –3, –5, –7, –9, ...
E 1, 2, 3, 4, ...

3 The first term of a sequence is 3. Each subsequent term is 0.6 times the previous term. What is the sixth term correct to two decimal places?

A 0.14 **B** 0.23 **C** 6.92 **D** 7.15 **E** 7.5

ISBN 978-1-108-46154-2

Cambridge University Press

4 The nth term of the sequence defined by the recurrence relation $t_1 = 50$ and $t_{n+1} = \frac{1}{2}t_n$ is:

A $t_n = 50 \times 2^{-n}$
B $t_n = 50 \times 2^{-n+1}$
C $t_n = 50 \times 2^{-2n+1}$
D $t_n = 50 \times 2^{-2n}$
E $t_n = 50 \times 2^{-2n-1}$

5 The ninth term of the arithmetic sequence 44, 41, 38, … is:

A 8 **B** 17 **C** 20 **D** 23 **E** 26

6 The following is a geometric sequence 15, −45, 135, −405, … The common ratio r is equal to:

A −3 **B** −2 **C** 1 **D** 2 **E** 3

7 A recurrence relation is defined by $T_{n+1} = 4T_n + 5$, where $T_1 = 3$. T_4 is equal to:

A 73 **B** 75 **C** 85 **D** 297 **E** 1193

8 The rungs of a ladder diminish uniformly in length from 30 cm at the bottom of the ladder to 22.5 cm at the top of the ladder. There are 16 rungs altogether. The length, in centimetres, of the 10th rung up the ladder is:

A 24.5 **B** 25.0 **C** 25.3 **D** 25.5 **E** 25.8

9 The sequence generated by the recurrence relation $V_0 = 5$, $V_{n+1} = V_n - 3$ is:

A 5, 15, 45, 135, 405, …
B 5, 8, 11, 14, 17, …
C 5, 2, −1, −4, −7, …
D 5, 15, 45, 135, 405, …
E 5, −15, 45, −135, 405, …

10 Brian has two trees in his backyard. Every month, he will plant three more trees. A recurrence relation model, T_n, for the number of trees in Brian's backyard at the start of the month n, is:

A $T_1 = 2, T_{n+1} = 3T_n$
B $T_1 = 2, T_{n+1} = 3T_n + 3$
C $T_1 = 2, T_{n+1} = T_n + 3$
D $T_1 = 2, T_{n+1} = T_n - 3$
E $T_1 = 2, T_{n+1} = 3T_n - 3$

ISBN 978-1-108-46154-2

Cambridge University Press

Review

11 Jennifer invests \$2000 with a bank. She will be paid simple interest at the rate of 5.1% per annum. If V_n is the value of Jennifer's investment after $n - 1$ years, the recurrence relation model for Jennifer's investment is:

A $V_1 = 2000,\ V_{n+1} = V_n + 5.1$
B $V_1 = 2000,\ V_{n+1} = 5.1\,V_n$
C $V_1 = 2000,\ V_{n+1} = 0.051\,V_n + 102$
D $V_1 = 2000,\ V_{n+1} = V_n + 102$
E $V_1 = 2000,\ V_{n+1} = 5.1V_n + 2000$

12 A sequence is generated from the recurrence relation $V_1 = 40,\ V_{n+1} = V_n - 16$. The rule for the value of the term V_n is:

A $V_n = 40n - 16$
B $V_n = 56 - 16n$
C $V_n = 40n$
D $V_n = 40 + 16n$
E $V_n = 40n - 16$

13 A computer is depreciated using a flat-rate depreciation method. It was purchased for \$2800 and depreciates at the rate of 8% per annum. The amount of depreciation after 4 years is:

A \$224 **B** \$448 **C** \$672 **D** \$896 **E** \$1904

14 A car purchased for \$18 990 is depreciated using a unit-cost depreciation method. After travelling a total of 20 000 kilometres, it has an estimated value of \$15 990. The depreciation amount, per kilometre, is:

A \$0.15 **B** \$0.80 **C** \$0.95 **D** \$6.67 **E** \$3000

15 A population of penguins is decreasing by 8% every year. There are currently 2700 penguins in the population. The number of penguins in the population after n years, A, is:

A 2700×1.8^n
B 2700×1.08^n
C 2700×0.92^n
D $2700 + 1.08^n$
E $2700 + 0.08n$

16 Sandra invests \$6000 in an account that pays compounding interest at the rate of 4.57% per annum. The number of years it takes for the investment to exceed \$8000 is:

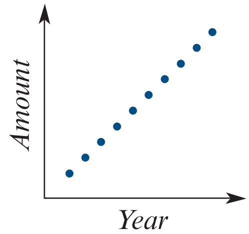
A 5 **B** 6 **C** 7 **D** 8 **E** 9

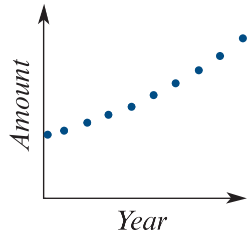
ISBN 978-1-108-46154-2

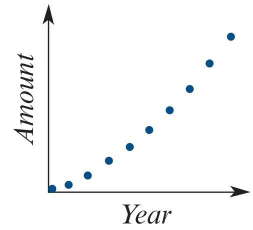
Cambridge University Press

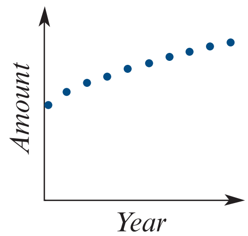
Review

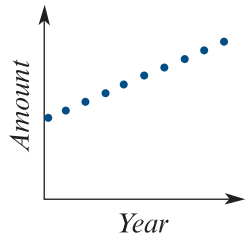
17 An investment of $50 000 is made at a fixed rate of interest compounding annually over a number of years. Which graph best represents the value of the investment at the end of each year?

A


B


C


D


E


18 After 10 years, a compound interest investment of $8000 earned a total of $4000 in interest. The annual interest rate of this investment was closest to:

A 2.5%
B 4.14%
C 5.03%
D 7.2%
E 50%

19 The second and fifth terms of a geometric sequence are –24 and 1536, respectively. The rule or the *n*th term is:

A $t_n = 6 \times (-4)^{n-1}$
B $t_n = 6 \times 4^n$
C $t_n = 6 \times 4^{n-1}$
D $t_n = 4 \times 3^{n-1}$
E $t_n = 4 \times (-3)^{n-1}$

20 The seventeenth and nineteenth terms of an arithmetic sequence are –28 and −102, respectively. The rule or the *n*th term is:

A $t_n = 564 - 37(n - 2)$
B $t_n = 564 - 37(n + 1)$
C $t_n = 564 + 37(n + 1)$
D $t_n = 564 - 37(n - 1)$
E $t_n = 564 - 37(n + 2)$

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Short-answer questions

SF

1 For an arithmetic sequence with $a = t_1 = 6$ and $d = 5$. Determine t_9.

2 For an arithmetic sequence with $a = t_1 = 20$ and $d = 4$. Determine t_{10}.

3 For a geometric sequence with $t_1 = 20$ and $r = 2$. Determine t_5.

4 For a geometric sequence with $t_1 = 2000$ and $r = 0.25$. Determine t_5.

5 In an arithmetic sequence with $t_5 = 22$ and $t_{10} = 47$. Determine t_{15}.

6 Find an expression for the nth term of the arithmetic sequence 13, 9, 5, …

7 For an arithmetic sequence with rule $t_n = 5n - 4$, write down the values of t_1, t_2 and t_3.

8 For a geometric sequence with rule $t_n = 5 \times 2^n$ write down the values of t_1, t_2 and t_3.

9 An arithmetic sequence has $a = t_1 = 9$ and $d = 3$. If $t_n = 36$, find the value of n.

10 A geometric sequence has $a = t_1 = 2$ and $r = 3$. If $t_n = 13\,122$, find the value of n.

11 A sequence is defined by the recurrence relation $t_n = t_{n-1} + 6$, with $t_1 = 6$. Write down the first three terms.

12 A sequence is defined by the recurrence relation $t_n = 0.5t_{n-1}$, with $t_1 = 1000$. Write down the first three terms.

13 A sequence is defined by the recurrence relation $t_n = t_{n-1} - 5$, with $t_1 = 8$. Write down a rule for the nth term of the sequence.

14 A sequence is defined by the recurrence relation $t_n = 3t_{n-1}$, with $t_1 = 1$. Write down a rule for the nth term of the sequence.

CF

15 In a geometric sequence in which all the terms are positive, $t_3 = 18$ and $t_5 = 162$. Determine t_8.

16 A sequence is generated from the recurrence relation $t_1 = 2$, $t_{n+1} = 6t_n$.

- **a** Write down a rule for the value of the nth term of this sequence.
- **b** Use the rule to find t_5.
- **c** Use the rule to find the value of t_{18}.

17 A sequence is generated from the recurrence relation $t_1 = 120$, $t_{n+1} = \frac{1}{4}t_n$

- **a** Write down a rule for the value of the nth term of this sequence.
- **b** Use the rule to find t_3.
- **c** Use the rule to find the value of t_{14}.

ISBN 978-1-108-46154-2

Review

18 A car was purchased for \$38 500. It depreciates in value at a rate of 9.5% per year, using a reducing-balance depreciation method.

a Write down a rule for the value of the car after n years.

b Use the rule to find the value of the car after five years.

c What is the total depreciation of the car over five years?

19 Jack borrows \$20 000 from a bank and is charged simple interest at a rate of 9.4% per annum. Let t_n be the value of the loan after n years.

a Write down a rule for the value of the loan after n years.

b How much will Jack need to pay the bank after 5 years?

c How many years does it take the value of Jack's loan to reach \$40 680?

20 A commercial cleaner bought a new vacuum cleaner for \$1650. The value of the vacuum cleaner decreases by \$10 for every 50 offices that it cleans.

a By how much does the cleaning of one office depreciate the value of the vacuum cleaner?

b Write down rule for the value of the vacuum cleaner after n offices are cleaned.

c The cleaner has a contract to clean 10 offices per night, 5 nights a week for 40 weeks in a year. What is the value of the vacuum cleaner after one year?

Extended-response questions

CU

1 Kelly bought her current car five years ago for \$22 500.

a Kelly assumes a flat-rate depreciation of 12% per annum. Let A be the value of Kelly's car after n years.

i Write down the value A of Kelly's car after n years.

ii Find the current value of Kelly's car.

b Kelly assumes reducing-balance depreciation at 16% per annum. Let B be the value of Kelly's car after n years.

i Write down the value of Kelly's car after n years.

ii Find the current value of Kelly's car.

c On the same axes, sketch a graph of the value of Kelly's car against the number of years for both flat-rate and reducing-balance depreciation.

2 Meghan has \$5000 to invest. Company A offers her an account paying 6.3% per annum simple interest.

a How much will she have in this account at the end of 5 years?

Company B offers her an account paying 6.1% per annum compound interest, compounding monthly.

b How much will she have in this account at the end of 5 years?

c Find, correct to one decimal place, the simple interest rate that Company A should offer if the two investments are to have equal value after 5 years.

ISBN 978-1-108-46154-2

Review

3 An iron ore smelting works has a tall chimney stack from which a pollutant gas is emitted at a rate of 1500 kilograms per day. New technology has been developed that enables the emissions to be reduced in stages to a minimum of 200 kilograms per day. There are two methods of installing new equipment to reduce the emissions.

a Using the *first* installation method, the emissions will be reduced by a *constant amount* each day until the minimum emission of 200 kilograms per day is reached. Consider the case where the emissions are reduced by 130 kilograms each day. The installation will be completed by the end of the 10th day, and from the 11th day the emissions will be 200 kilograms per day. Use this information to complete the table below.

Day	1	2	3	4	5	6	7	8	9	10
Emission each day (kilograms)	1500	1370	1240	…	…	…	…	…	…	330

b Now suppose that the installation is to be completed by the end of the eighth day so that from the ninth day the emission will be 200 kilograms per day. By what *constant amount* must the emission be reduced each day during the installation period?

c Using the *second* installation method, the emissions will be reduced by a *constant percentage* each day until the minimum emission of 200 kilograms per day is reached. Consider the case where the constant percentage is 25%. While the daily emissions are being reduced, the emissions each day will form a geometric sequence.

i Write down the common ratio of this geometric sequence.

ii Complete the table below, giving the entries correct to two decimal places.

Day	1	2	3	4	5	6	7	8	9	10
Emission each day (kilograms)	1500	1125	843.75	…	…					

d In the case described in part **c**, on which day will the daily emission first reach the minimum of 200 kilograms, within one kilogram?

e If the case described in part **d** is used rather than the case described in part **a**, how much less, to the nearest kilogram, is the total emission during days 1 to 10 inclusive?

f Now suppose that the second installation method is used, but the minimum daily emission of 200 kilograms is not reached until the 10th day. By what *constant percentage* must the emissions be reduced each day in this case? Give your answer correct to one decimal place.

ISBN 978-1-108-46154-2

Cambridge University Press

5

Earth geometry and time zones

UNIT 3 BIVARIATE DATA, SEQUENCES AND CHANGE, AND EARTH GEOMETRY

Topic 4 Earth geometry and time zones

- How do we define a great circle?
- How do we use latitude and longitude to describe a location on Earth?
- How do we find the latitude and longitude of a location on Earth?
- How do we calculate the distance between two places on Earth?
- How do we find time differences between two places on Earth using time zones?
- How do we solve problems in time planning associated with time differences between two places on Earth?

ISBN 978-1-108-46154-2

Cambridge University Press

5A Angle measurement and arc length

We start with a brief section to remind you how to convert angle measurements given in decimal form to degrees and minutes and vice versa, and how to calculate the length of an arc. This section will help you to understand and make calculations relating to Earth geometry.

▶ Conversion of angle measurements

There is a provision to further increase accuracy of angle measurements with a third measurement of angle, which is seconds, but we will not do this here. Answers will be given to the nearest minute.

There are 60 minutes in a degree. We write 56 minutes as $56'$. To get a feeling for the conversion consider the following.

- $34.50° = 34°30'$
- $34.25° = 34°15'$
- $34.75° = 34°45'$

Changing from decimal form to degrees and minutes

Multiply the decimal part of the number by 60.

- For $34.7°$, multiply $0.7°$ by 60. The result is $42'$ and we have $34.7° = 34°42'$.
- For $34.321°$, multiply $0.321°$ by 60. The result is $19.26'$ and we have $34.321° = 34°19'$ to the closest minute.

Changing from degrees and minutes to decimal form

Divide the minutes by 60.

- For $34°56'$, divide 56 by 60. The result is 0.9333… and we have $34°56' = 34.93°$, correct to two decimal places.
- For $54°19'$ divide 19 by 60. The result is 0.31666… and we have $54°19' = 54.317°$ correct to three decimal places.

Example 1 Converting angle measurements

a Change $32.45°$ to degrees and minutes.

b Change $44°32'$ to decimal form. Give your answer to 2 decimal places.

Solution

a Multiply the decimal part of the number by 60.	$0.45° = (0.45 \times 60) = 27'$ Therefore, $32.45° = 32°27'$.
b Divide the minutes part of the angle by 60.	$32' = (32 \div 60) = 0.533\ldots$ Therefore, $44°32' = 44.53°$ correct to two decimal places.

ISBN 978-1-108-46154-2

Calculator activity 5A Converting angle measurements on a calculator

Solve the following using a calculator.

a. Change 32.45° to degrees and minutes.

b. Change 44°32′ to decimal form.

Casio fx82

a. Type 32.45 press [° ′ ″] > = to get the degrees and minutes form
   To get the result 32.45° = 32°27′

b. Type 44 > [° ′ ″] > 32 > [° ′ ″] > = > [° ′ ″] to get the decimal form
   To get the result 44°32′ = 44.533…

TI-30XB

a. Type 32.45 press [2nd] > [Angle]. Scroll down the list to find [▸DMS] > enter
   To get the result 32°27′.

b. Type 44 then press [2nd] > [Angle] > 1 > type 32 > [Angle] > 2 > enter
   To get the result 44.533…

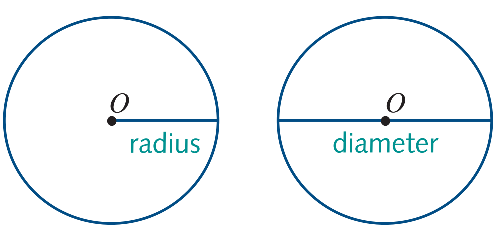
Sharp

a. Type 32.45 press [2ndF] > [D°M′S] [↔ **DEG**]
   To get the result 32°27′.

b. Type 44, then press [D°M′S] > type 32 > [D°M′S] > [2ndF] > [D°M′S] [↔ DEG]
   To get the result 44.533…

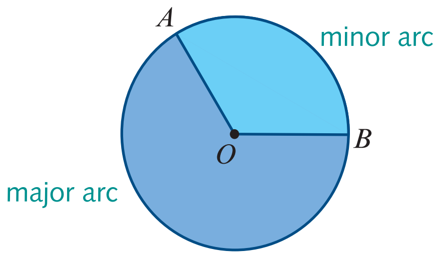
Arc length

Any line segment drawn from the centre of a given circle to any point on the circle is called a **radius** (plural radii).



Any line segment joining two points on the circle and passing through the centre is called the *diameter* of the circle.

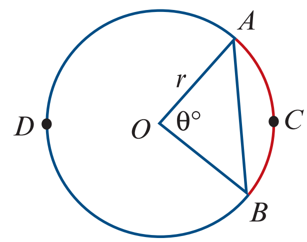
Any two points on a circle divide the circle into **arcs**. The shorter arc is called the *minor arc*, the longer is the *major arc*.



ISBN 978-1-108-46154-2

Cambridge University Press

The arc ACB is said to subtend the angle $\angle AOB$ at the centre of the circle. If $\angle AOB = \theta$ and radius length is r units, then the length of arc ACB will be a fraction of the circumference.



The fraction of the circumference will be $\frac{\theta}{360}$.

Recall that the circumference C of a circle of radius r is given by

$$C = 2\pi r.$$

Therefore, the length s of an arc that subtends an angle of θ at the centre is:

$$s = \frac{\theta}{360} \times 2\pi r$$

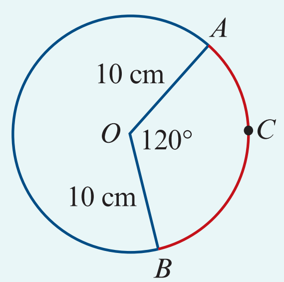
Length of an arc

The length, s, of an arc of a circle of radius, r, that subtends an angle of θ at the centre is given by:

$$s = \frac{\pi r \theta}{180}$$

Example 2 Calculating the length of an arc

In this circle, centre O, and radius length 10 cm, the angle subtended at O by arc ACB has magnitude 120°. Find the length of the arc ACB correct to one decimal place.



Solution

1 Write down the formula for the length of an arc.

$$s = \frac{\pi r \theta}{180}$$

2 Substitute $\theta = 120°$ and $r = 10$ into the equation.

$$s = \frac{\pi \times 10 \times 120}{180}$$

$$= \frac{20\pi}{3}$$

≈ 20.9 cm (correct to one decimal place)

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 5A

Converting angle measurements

Example 1

1 Convert each of the following angle measurements from decimal form to degrees and minutes.

a 32.45° **b** 43.20° **c** 122.46° **d** 91.12° **e** 0.75°

SF

2 Convert each of the following angle measurements from degrees and minutes to decimal form.

a 32°45′ **b** 15°35′ **c** 7°22′ **d** 142°44′ **e** 67°15′

Calculating the length of an arc

Example 2

3 What is the circumference of each circle? Answer correct to two decimal places.

a Radius of 8 cm

b Radius of 14 m

c Radius of 45 mm

d Diameter of 12 mm

e Diameter of 14 m

4 What fraction of a circle is each sector?

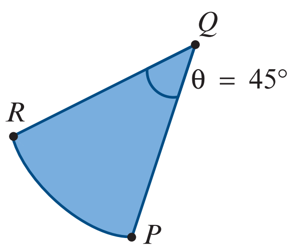
a Angle at the centre is 90° **b** Angle at the centre is 270°

c Angle at the centre is 30° **d** Angle at the centre is 120°

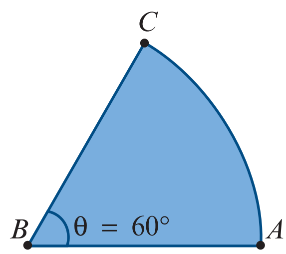
e Angle at the centre is 60° **f** Angle at the centre is 150°

5 Find the arc length of each sector. The radius is 10 cm. Answer correct to two decimal places.

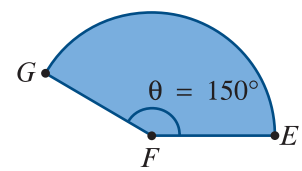
a



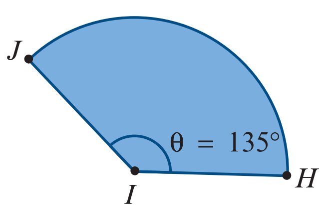
b



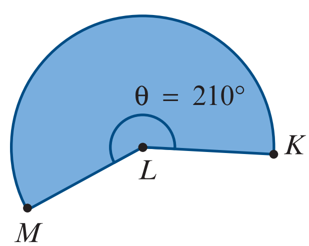
c



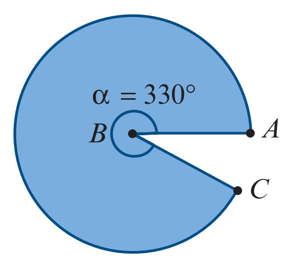
d



e



f



ISBN 978-1-108-46154-2

Cambridge University Press

6 Find the arc length where the radius of the circle (given in cm) and the angle subtended at the centre are as given. Give your answer correct to two decimal places.

a $r = 15,\ \theta = 50°$

b $r = 20,\ \theta = 15°$

c $r = 30,\ \theta = 150°$

d $r = 16,\ \theta = 135°$

e $r = 40,\ \theta = 175°$

f $r = 30,\ \theta = 210°$

7 Find the arc length that subtends an angle of magnitude 105° at the centre of a circle of radius 25 cm.

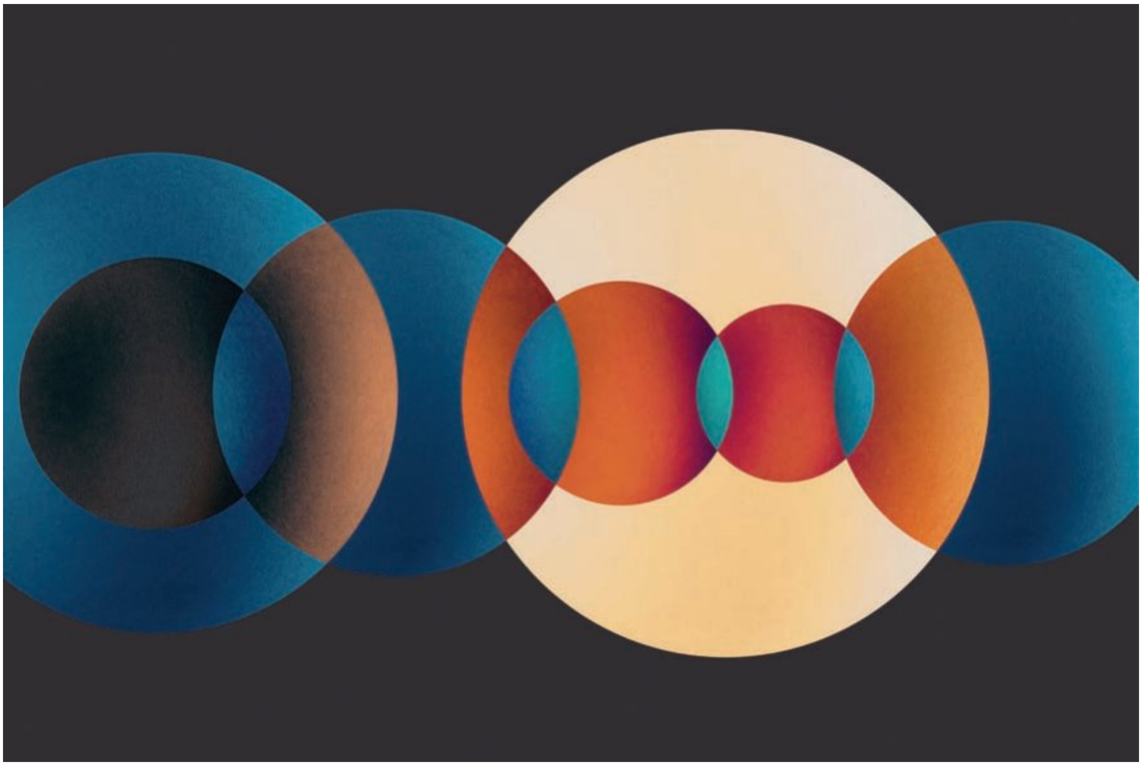
8 Find the size of the angle subtended at the centre of a circle of radius length 30 cm by an arc length of:

a 50 cm

b 25 cm

CF

9 A chord of length 6 cm is drawn in a circle of radius 7 cm. Find the length of the minor arc cut off by the chord.



ISBN 978-1-108-46154-2

5B Latitude and longitude

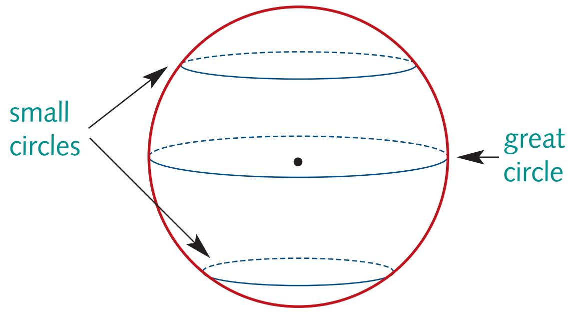
The Earth can be modelled by a sphere of radius 6371 km.
The radius at the equator is 6378.14 km, but the radius at the poles is only 6356.75 km.
Here we will use the average radius for our calculations of distances on the Earth.

Elements of Earth geometry

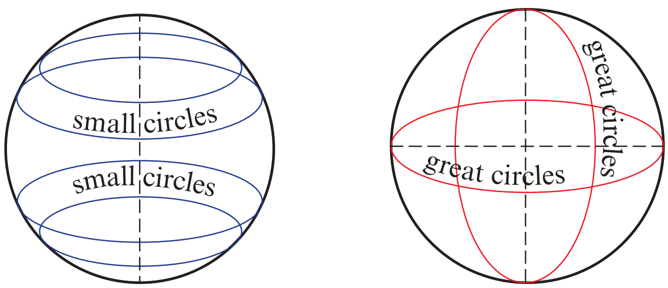
Great circles and small circles

A **great circle** is the circular boundary of a cross-section of a sphere that contains a diameter of the sphere. The cross-section contains the centre of the sphere.

The circular boundary of cross-sections of the sphere that do not contain a diameter of the sphere are called **small circles**. The cross-section for a small circle does not contain the centre of the sphere.



The great circle shown in the diagram above is in the *plane of the equator*.



The shortest distance between two points on the surface of the Earth is the distance along the great circle that passes through those two points.

The great circle path from New Delhi to New York is shown in the figure to the right.

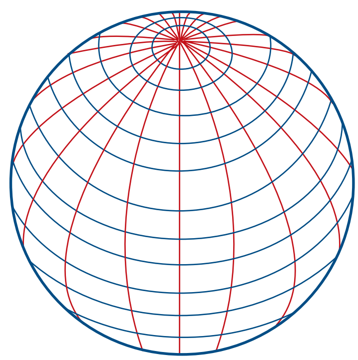
ISBN 978-1-108-46154-2

Cambridge University Press

Below is a representation of the great circle route from Brisbane to London.

In this section, we are interested in describing the location of points on the surface of the Earth. We do this in a manner similar to how we described points in the plane with Cartesian coordinates.

This is done using a grid of lines as shown here. They are used to give *coordinates* called **latitude** and **longitude**. The red lines will be used for longitude and the blue lines for latitude.



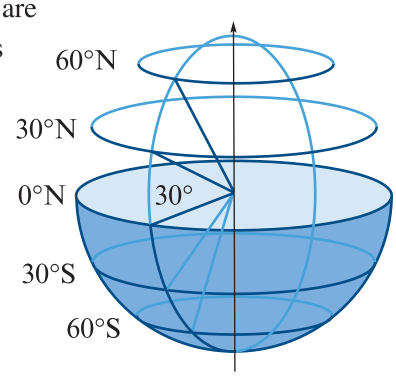
Meridians and parallels

Meridians of longitude are *semi-great circles* (an arc that goes from pole to pole) that pass through the north and south poles. The red lines on the sphere are meridians of longitude.

Parallels of latitude are small circles whose planes are parallel to that of the equator. The blue lines in the sphere are parallels of latitude. The equator is the only latitude that is a great circle.

Latitude

The blue lines in the diagram opposite are parallels of latitude. The latitude of a point on a sphere is the elevation of the point from the plane of the equator.



- The equator has latitude 0°N.
- The north pole has latitude 90°N.
- The south pole has latitude 90°S.

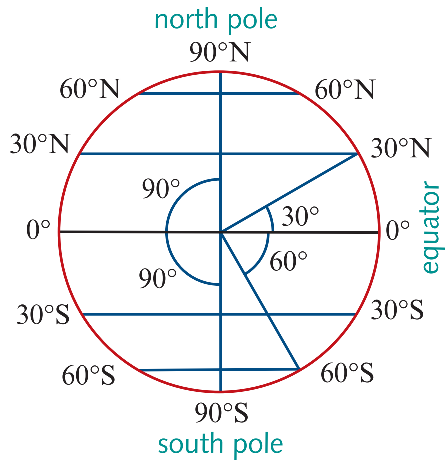
In the diagram the latitudes 60°N, 30°N, 60°S and 30°S are shown.

ISBN 978-1-108-46154-2

In the diagram opposite, the Earth has been sliced in half along a great circle. The vertical line through the poles is perpendicular, or at 90°, to the plane of the equator.

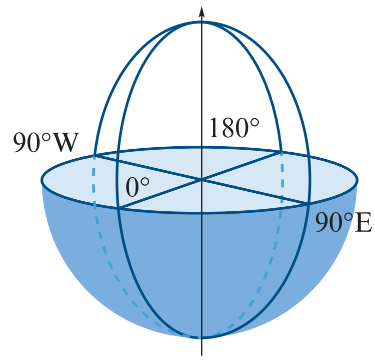
At the surface of the Earth, at a given latitude, draw a line from that location to the centre of the Earth. The angle between this line and the equator is the latitude measurement.

The diagram shows two examples: one for 30°N and one for 60°S.



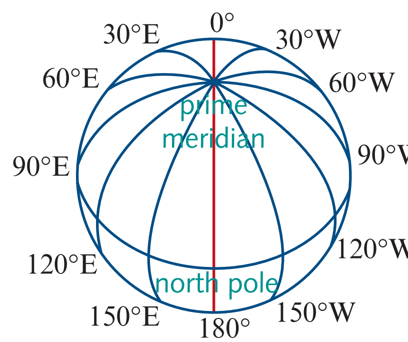
Longitude and the prime meridian

Lines of longitude are measured in degrees east or west of the **prime meridian** (0°). The lines of longitude shown in the diagram opposite are 0°, 90°E, 180° and 90°W. Note that you don't need to add E or W to the 0° or the 180°.



The prime meridian passes through Greenwich in England.

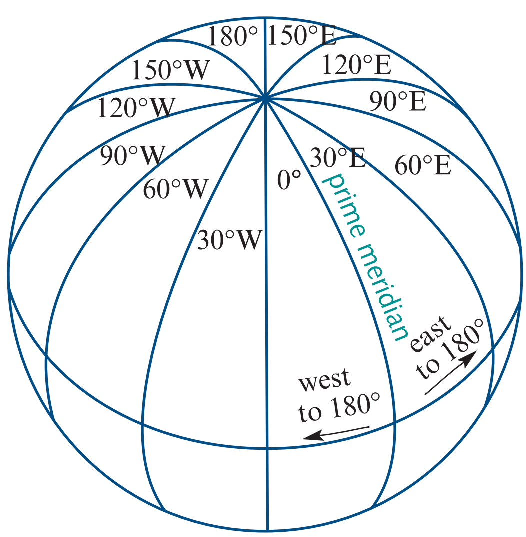
The diagram opposite is a diagram of the Earth looking down from the north pole. The evident plane in the diagram above is the plane of the equator. The angle formed between the prime meridian and the line from the centre of the Earth to the point where the meridian of longititude meets that plane is the longitude.



The meridian 120°W is on the same great circle as the meridian 60°E.

The meridian 30°W is on the same great circle as the meridian 150°E.

This is shown in a different way in the diagram opposite.



ISBN 978-1-108-46154-2

Cambridge University Press

Latitude and longitude

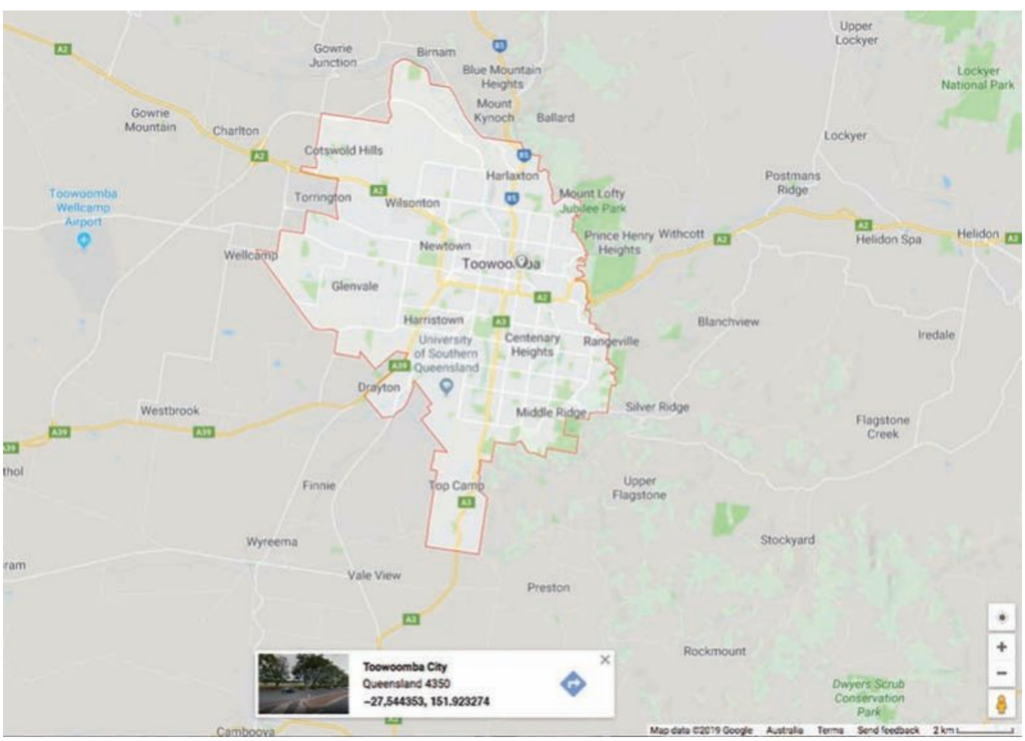
Any point on the Earth's surface can be described by giving its latitude and longitude. For example, Brisbane has latitude 27.4698°S and longitude 153.0251°E. Townsville has latitude 19.2590°S and longitude 146.8169°E. These are called the coordinates of the location.

Hemispheres

- Northern hemisphere, the half of the Earth that lies north of the equator. Locations in the northern hemisphere, such as London, have their latitude described using °N. For London this is 51.5°N.
- Southern hemisphere, the half of the Earth that lies south of the equator. Locations in the southern hemisphere, such as Brisbane, have their latitude described using °S. For Brisbane this is 27.5°S.
- Western and eastern hemispheres are defined through the prime meridian. Brisbane is in the eastern hemisphere and New York in the western hemisphere.

Finding the longitude and latitude of a location

Here is a map of a section of South East Queensland and northern NSW obtained from Google Maps. By clicking on a location, information is obtained including the latitude and longitude. (The precise method will depend on your device and browser or app, check Google Maps Help if needed.) In the diagram of the map below, the details for Toowoomba are shown.



ISBN 978-1-108-46154-2

The coordinates of Toowoomba are given as (−27.544353, 151.932374). Changing this form to the hemisphere notations is 27.5598°S, 151.9507°E and changing to degrees and minutes is 27°34′S, 151°57′E. This may be done with any location. Other methods of finding a location include using a Global Positioning System (GPS) or using an atlas.

GPS

GPS is a system used for worldwide navigation and surveying. It is commonly used for determining an exact location anywhere on Earth by obtaining the current time at a specific location. This is made possible by the network of 24 man-made satellites, called GPS satellites.

The system was used for military purposes, the GPS system became available for use by all people approximately 30 years ago. GPS provides latitudes and longitudes given to a high degree of accuracy. For example, through Google Maps you could enter an address and the GPS coordinates will be given for the place, or vice-versa.

Distance along a meridian

On a flat surface, the shortest distance between two points is a straight line. Since the Earth's surface is curved, the shortest distance between A and B is the arc length AB of the great circle (the meridian) that passes through A and B. This is called the *great circle distance*.

We can calculate the distance between two points on Earth using the difference in their latitudes. Great circles of Earth have a radius of about 6371 km, so their circumference is $2 \times \pi \times 6371 \approx 40030$ km.

If two points subtend an angle of 1° at the centre of a great circle, the distance between them is:

$$\frac{1}{360} \times 40030 \approx 111.2 \text{ km}$$

Note: This result will be used throughout the chapter.

Angular distance with respect to a meridian

Beijing (China), and Perth (Australia), have coordinates (40°N, 116°E) and (32°S, 116°E), respectively. These two cities have the same longitude to the nearest degree. As the cities are in different hemispheres, north and south, we need to add the latitudes to determine the angular distance. The angular distance = (40 + 32) = 72°.

For Brisbane and Coffs Harbour, the latitudes are 27°S and 30°S, respectively (to the nearest degree). As the cities are in the same hemisphere, south, we need to subtract the latitudes to determine the angular distance. The angular distance = (30 − 27) = 3°. The cities are in the same hemisphere.

ISBN 978-1-108-46154-2

Cambridge University Press

For locations on the same meridian in different hemispheres, we add the latitudes. For locations on the same meridian in the same hemisphere, we find the difference between the latitudes (always subtract smaller from larger).

The distance, D km, between two points on the same meridian is given by:

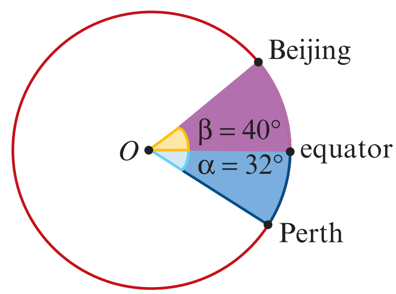
$$D = 111.2 \times \textit{angular distance}$$

Example 3 Finding a distance along a meridian

Beijing (China), and Perth (Australia), have coordinates (40°N, 116°E) and (32°S, 116°E), respectively. Calculate the shortest distance between Beijing and Perth, to the nearest kilometre.

Solution

1 The two cities have the same longitude correct to the nearest degree. Therefore, they are on the great circle that is the meridian of longitude 116°E.



2 Add the latitudes of each city to find the angle subtended at the centre of the arc.

Angle = (40 + 32) = 72°

3 As an angle of 1° at the centre of the great circle is subtended by 111.2 km on Earth, calculate the length of the arc by multiplying the angle measurement by 111.2 km.

Therefore, the distance along the meridian
≈ 111.2 × 72
= 8006 km

Note: The calculated distance given on the internet is 7985 km, which is quite close to our approximate result.

ISBN 978-1-108-46154-2

Cambridge University Press

Finding the distance along a meridian using a search engine

In a search engine, for example, Google, type 'Distance from Perth to Beijing'. The shortest arc along the great circle is shown.

Distance between two points on the equator

The equator is a great circle and therefore the distance between points on the equator are possible to find using our knowledge of length of arcs.

Example 4 **Finding the distance between two points on the equator**

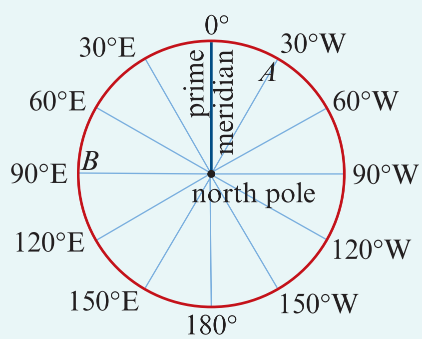
Point A has longitude 30°W and latitude 0°.

Point B has longitude 90°E and latitude 0°.

Find the distance between the two points:

a if you fly east from A to B

b if you fly west from A to B



Solution

a **1** Flying east the angle between the locations is 120°.

$\theta = (30 + 90) = 120°$

2 We can use the formula obtained for locations on the same meridian because we are on a great circle.
$D = 111.2 \times$ angular distance

$D = 111.2 \times 120$
$= 13\,344$ km

3 Write the answer in a sentence.

Flying east the distance between A and B is 13 344 km.

ISBN 978-1-108-46154-2

Cambridge University Press

b **1** Flying west the angle is 240° found by subtracting 120° from 360°.	$\theta = (360 - 120) = 240°$
2 Use the formula to calculate the distance between the points. $D = 111.2 \times$ angular distance	$D = 111.2 \times 240$ $= 26\,688$ km
3 Write the answer in a sentence.	Flying west the distance between A and B is 26 688 km.

Example 5 Finding a distance between two places on the equator

The cities of Pontianak (Indonesia), and Quito (Ecuador), are on the equator to the nearest degree. The longitude of Pontianak is 109°E and the longitude of Quito is 78°W. Find the distance between the two cities flying east from Pontianak to Quito.

Solution

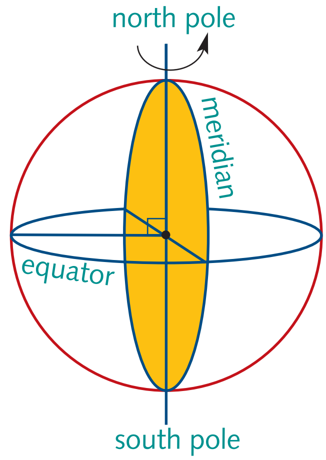
1 Add to find the total angle between the longitudes – we are considering the distance covered by flying west.	$\theta = 78 + 109 = 187°$
2 Determine the angle between the two longitudes if flying east.	Hence, the required angle is $\theta = 360 - 187 = 173°$
3 Use the formula to calculate the distance between Pontianak and Quito if flying east. $D = 111.2 \times$ angular distance	$D = 111.2 \times 173$ $= 19\,238$ km
4 Write the answer in a sentence.	The approximate distance when flying east from Pontianak to Quito is 19 238 km.

ISBN 978-1-108-46154-2

Cambridge University Press

Distance from a pole or from the equator

All meridians pass through the poles and a point on the equator. Therefore, you can find the distance from any point on the surface of the Earth to a pole or the equator if we know its latitude.



Example 6 Finding the distance to the equator or a pole

Brisbane has latitude 27°S and longitude 153°E. Find the distance of Brisbane to:

a the equator **b** the south pole **c** the north pole

Solution

a **1** Along the plane of the meridian 153°E, the difference between the equator and Brisbane is $(27 - 0 = 27°)$. Use the formula $D = 111.2 \times$ angular distance.

$D = 111.2 \times 27$
$= 3002.4$ km

2 Write the answer in a sentence.

The approximate distance between Brisbane and the equator is 3002.4 km.

b **1** Along the plane of the meridian 153°E, the difference between the south pole and Brisbane is $(90 - 27 = 63°)$. Use the formula $D = 111.2 \times$ angular distance.

$D = 111.2 \times 63$
$= 7005.6$ km

2 Write the answer in a sentence.

The approximate distance between Brisbane and the south pole is 7005.6 km.

c **1** Along the plane of the meridian 153°E, the difference between the north pole and Brisbane is $(90 + 27 = 117°)$. Use the formula $D = 111.2 \times$ angular distance.

$D = 111.2 \times 117$
$= 13\,010.4$ km

2 Write the answer in a sentence.

The approximate distance between Brisbane and the north pole is 13 010.4 km.

Note: The distance from the north pole to the south pole along a great circle is approximately

$$= \pi \times 6371 \approx 20015 \text{ km.}$$

ISBN 978-1-108-46154-2

Cambridge University Press

Distance along a parallel of latitude

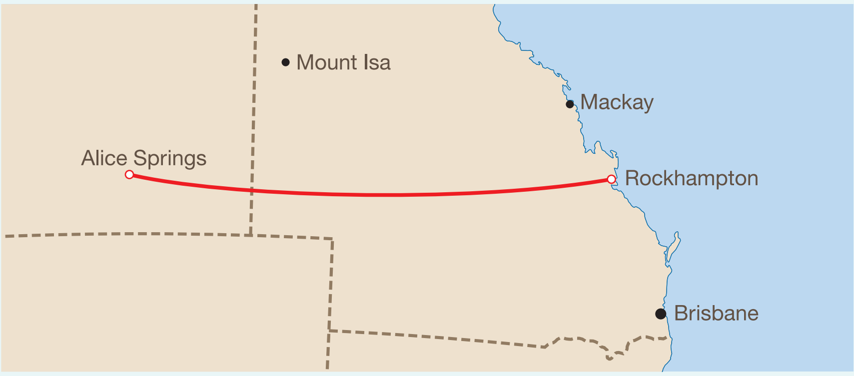
In the following example, we find out how to find the distance between two points with the same latitude. We will illustrate this by finding the distance between Rockhampton and Alice Springs.

Example 7 Calculating distance along a parallel of latitude

Rockhampton, Queensland, has latitude 23°S and longitude 150°E.

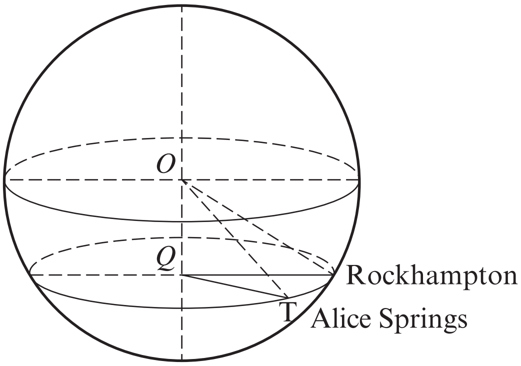
Alice Springs, Northern Territory, has latitude 23°S and longitude 134°E.

Find the distance along the parallel of latitude 23°S from Rockhampton to Alice Springs.



Solution

1 In the diagram of the Earth, Alice Springs and Rockhampton are shown. The circle passing through these two cities that is parallel to the equator is the parallel of latitude at 23°S.



2 Find the radius of the small circle of latitude 23°S.

Using right-angled triangle OTQ, the radius (QT) of the small circle of the latitude 23°S is

$$\cos 23^\circ = \frac{QT}{6371}$$

$$QT = 6371 \times \cos 23^\circ$$

$$\approx 5864.54 \text{ km}$$

ISBN 978-1-108-46154-2

Cambridge University Press

3 Find the length of the arc connecting Alice Springs and Rockhampton.

The required angle $= 150^\circ - 134^\circ$
$= 16^\circ$

$$\text{Distance} = \frac{16}{360} \times 2 \times \pi \times 5864.54$$
$$= 1637.68$$
$$\approx 1638 \text{ km}$$

These calculations are always the same for calculating distances around small circles so we use $111.2 \cos 23 \times 16 \approx 1638$.

The distance D km between two points that have the same parallel of latitude is given by

$$D = 111.2 \cos\theta \times \textit{angular distance}$$

where the parallel of latitude is $\theta°$N or $\theta°$S.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 8 Calculating distance along a parallel of latitude

The latitude of both Rockhampton in Queensland and Sao Paolo in Brazil is 23°S. Their longitudes are 151°E and 47°W, respectively. Find the distance between the two cities:

a by flying west from Rockhampton to Sao Paolo

b by flying east from Rockhampton to Sao Paolo

Solution

a

1 For flying west, add the two angles as they are both measured from either side of the prime meridian.

Angular distance = (151 + 47) = 198°
Latitude = 23°S

2 Calculate the distance between the two locations using the formula.

$D = 111.2 \cos 23^\circ \times 198$
$= 20\,267.31$ km

3 Write the answer in a sentence.

Flying west the distance is 20 267.31 km.

b

1 For flying east, subtract the angle in part **a** from 360°.

Angular distance = (360 − 198) = 162°

2 Calculate the distance using the formula.

$D = 111.2 \cos 23^\circ \times 162$
$= 16\,582.34$ km

3 Write the answer in a sentence.

Flying east the distance is 16 582.34 km.

Example 9 Calculating distance using degrees and minutes

Yarraden in Queensland and Wyndham in Western Australian both have latitude 15°S. Yarraden has longitude 143°18′E and Wyndham has longitude 128°07′E. Find the distance along the small circle between the locations at the latitude 15°S.

Solution

1 Convert the longitudes to decimal notation first.

$143^\circ 18'E = 143.3^\circ E$
$128^\circ 07'E = 128.12^\circ E$

2 The two locations are both east of the prime meridian. Find the difference between the two longitudes.

Angular difference = 15.18°

3 Calculate the distance using the formula.

$D = 111.2 \times \cos 15^\circ \times 15.18$
$= 1630.50$ km

4 Write the answer in a sentence.

The distance between Yarraden and Wyndham is 1630.50 km.

ISBN 978-1-108-46154-2

Exercise 5B

Making latitude and longitude calculations

Skillsheet

SF

1 Use an atlas or Google Earth to name the place situated with the following coordinates.

a 44°26′N, 26°06′E (44.43°N, 26.10°E)

b 59°54′N, 10°45′E (59.91°N, 10.75°E)

c 29°55′N, 95°22′W (29.76°N, 95.36°W)

d 1°21′N, 103°49′E (1.35°N, 103.82°E)

e 33°55′S, 18°25′E (33.92°S, 18.42°E)

2 Use an atlas, Google Earth or another method to state the coordinates of:

a Cairo

b Mexico City

c Buenos Aires

d Mumbai

e Lagos

f Harare

3 Use Google maps or an atlas or another method to find the coordinates of each of the following Queensland locations.

a St George

b Caloundra

c Mackay

d Weipa

e Bundaberg

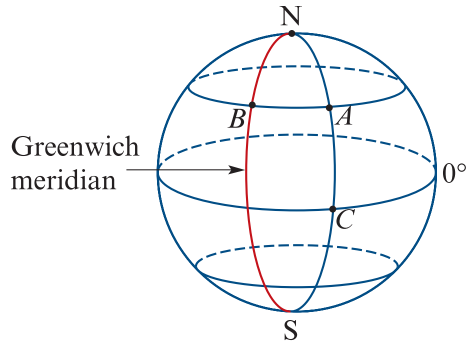
f Charters Towers

g The Queensland and Northern Territory border (Longitude only)

4 On the diagram to the right, the latitude and longitude of point A are $(65°\text{N},\ 75°\text{E})$.

a What are the coordinates of point B?

b What are the coordinates of point C?



5 Athens has the coordinates $(38°\text{N},\ 24°\text{E})$ and Sofia has the coordinates $(43°\text{N},\ 23°\text{E})$.

a What is the latitude and longitude of a point 20° due south of Athens?

b What is the latitude and longitude of a point 20° due east of Athens?

c What is the latitude and longitude of a point 60° due south of Sofia?

d What is the latitude and longitude of a point 60° due west of Sofia?

ISBN 978-1-108-46154-2

Cambridge University Press

6 The following table shows the latitude and longitude of cities around the world given to the nearest degree.

City	Latitude	Longitude
Brisbane	27°S	153°E
Melbourne	38°S	145°E
Cooktown	15°S	145°E
Townsville	19°S	147°E
Marseilles	43°N	5°E
Mexico City	19°N	99°W
Wellington	41°S	175°E
Zurich	47°N	8°E
London	52°N	0°(actually west of Greenwich)
Lima	12°S	77°W
Plymouth	50°N	4°W
Yangon	19°N	96°E

a Which city (or cities) is closest to the following latitudes?
i 15°N **ii** 28°S
b Which city (or cities) is closest to the following longitudes?
i 151°E **ii** 20°W
iii Greenwich meridian **iv** Longitude of Sofia (from Question 5)
c Which cities are in the northern hemisphere?
d Which cities are in the western hemisphere?
e Which cities have the same latitude?
f Which cities have the same longitude?
g Which city is closest to the north pole?
h Which city is closest to the south pole?

Calculating the distance between two points on the same meridian

Example 3, 4

7 Two places on the same meridian have latitudes 22°N and 35°S. Find the distance between the two places. Give your answer to the nearest kilometre.

8 Cairns 17°S and Melbourne 37.8°S are nearly on the same meridian. Assuming they are, find the distance between them. Give your answer to the nearest kilometre.

9 How far apart are Esperance 34°S, 122°E and Broome 18°S, 122°E. Give your answer to the nearest kilometre.

10 Cairns and Griffith are 1920 km apart and both are on the same meridian. If the latitude of Cairns is 17°S, find the latitude of Griffith. Griffith is south of Cairns.

ISBN 978-1-108-46154-2

Cambridge University Press

11 Calculate the shortest distance along a meridian in each of the following cases. Give your answer to the nearest kilometre.

a Point X: latitude 10°N, longitude 18°W
Point Y: latitude 45°N, longitude 18°W

b Point X: latitude 14°N, longitude 35°W
Point Y: latitude 13°S, longitude 35°W

c Point X: latitude 23°S, longitude 140°W
Point Y: latitude 67°S, longitude 140°W

d Point X: latitude 15°N, longitude 60°W
Point Y: latitude 25°S, longitude 60°W

e Point X: latitude 15°N, longitude 70°W
Point Y: latitude 15°S, longitude 70°W

Calculating the distance between two points on the equator

Example 4

12 The difference of longitude between two points on the equator is 32°. Find the distance between them in kilometres.

Example 5

13 There are places in Ecuador (South America), and Somalia (Africa), which are on the equator. In Ecuador, the longitude of a place X on the equator is 78°W and in Somalia the longitude of a place Y is 42°W. Find the distance between them in kilometres.

14 Calculate the shortest distance along the equator between places A and B having the following longitudes:

a A 137°E B 87°E
b A 57°E B 13°W
c A 57°W B 27°W
d A 140°E B 160°W
e A 95°W B 113°E

15 The distance between two points on the equator is 600 km. What is the difference in their longitudes?

Finding the distance to the equator or a pole

Example 6

16 New Orleans has latitude 30°N and longitude 90°W. Find the distance of New Orleans to:

a the equator **b** the north pole **c** the south pole

17 Izmir has latitude 38°N and longitude 27°E. Find the distance of Izmir to:

a the equator **b** the north pole **c** the south pole

18 Ballarat has latitude 37.5500°S and longitude 143.8500°E. Find the distance of Ballarat to:

a the equator **b** the north pole **c** the south pole

19 Find the distance from the given point to the equator and to each of the poles.

a Latitude 42°N, longitude 134°E
b Latitude 55°N, longitude 45°W
c Latitude 15°S, longitude 35°E
d Latitude 14°S, longitude 75°W

ISBN 978-1-108-46154-2

Cambridge University Press

Calculating the distance around a parallel of latitude

20 Find the radius of the small circle that is parallel to the latitude:

a 15°S **b** 30°S **c** 45°S **d** 60°S

21 For which parallels of latitude is the radius of the small circle half the radius of the equator. CF

Example 7

22 Find the distance around the parallel of latitude for the following locations. SF

a X: latitude 22°N, longitude 134°E; Y: latitude 22°N, longitude 145°E

b X: latitude 32°S, longitude 50°E; Y: latitude 32°S, longitude 80°E

c X: latitude 12°S, longitude 30°E; Y: latitude 12°S, longitude 80°E

23 The position of Salzburg is 48°N, 13°E and the position of Seattle is 48°N, 122°W. Find the distance around the 48°N parallel of latitude between Seattle and Salzburg.

Calculations using degrees and minutes

Example 9

24 In the following, each pair of locations lie on the same meridian. Find the approximate distance between them.

a X(24°15′N), Y(36°35′N) **b** X(24°15′N), Y(36°45′S)

c X(15°15′S), Y(26°55′N) **d** *X*(58°15′S), Y(36°45′S)

25 In the following, each pair of locations lie on the same latitude. Find the approximate shortest distance between them travelling along the small circle.

a X(124°15′E, 20°15′S), Y(36°45′E, 20°15′S)

b X (104°15′E, 32°25′S), Y(28°45′W, 32°25′S)

c X (120°15′E, 0°25′N), Y(120°45′W, 0°25′N)

d X (158°15′E, 40°45′N), Y(26°45′E, 40°45′N)

Mixed exercises

Example 7

26 An aircraft flies from Bairnsdale (38°S, 148°E) due north to Grenfell NSW (34°S, 148°E). CF

a How far is this (great circle) distance?

It then flies due west to Renmark SA (34°S, 141°E).

b How far is it from Grenfell to Renmark (distance around the parallel of latitude)?

c What is the total distance flown?

27 An aircraft flies from Warragul (38°S, 146°E) due north to Cairns, Qld (17°S, 146°E).

a How far is this (great circle) distance?

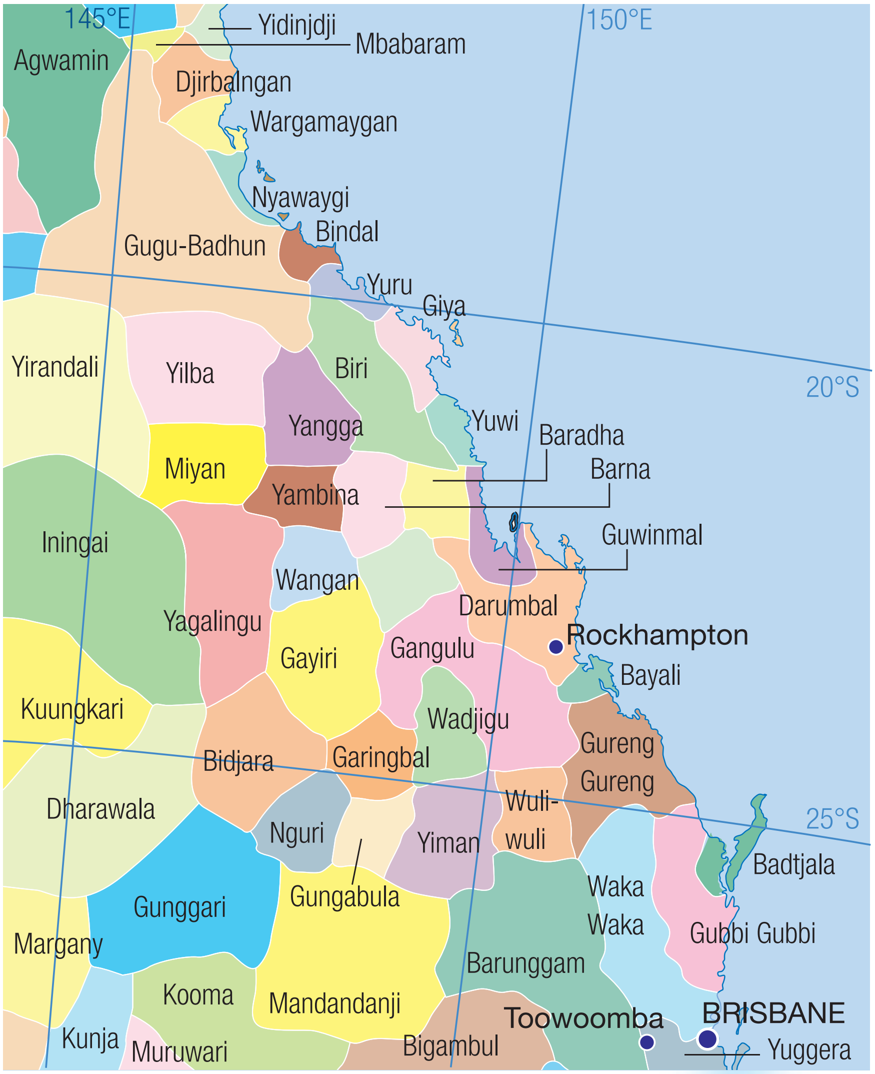
It then flies due west to Derby WA (17°S, 124°E).

b How far is it from Cairns to Derby (distance around the parallel of latitude)?

c What is the total distance flown?

ISBN 978-1-108-46154-2

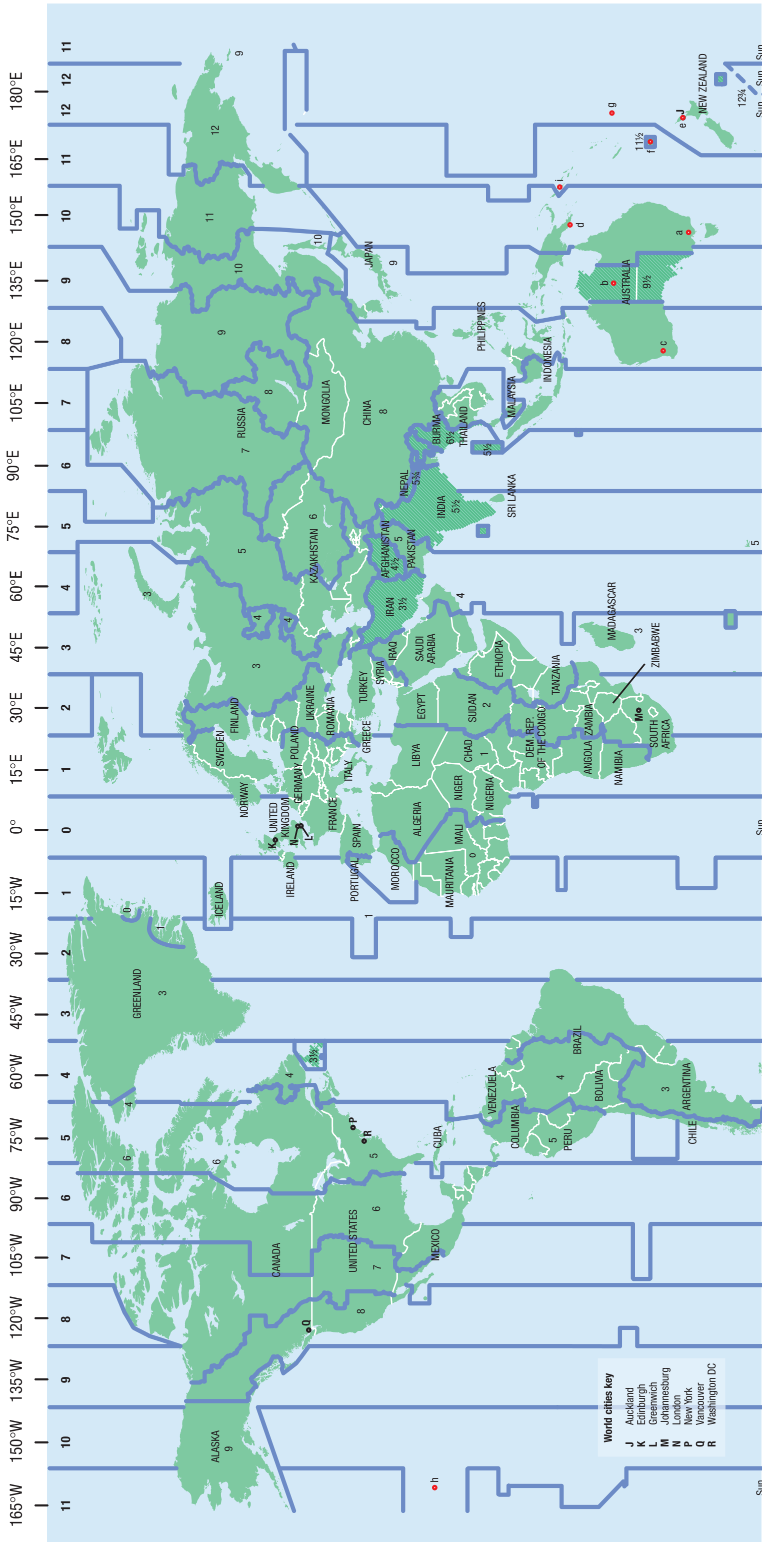
28 The map shows various regions of Queensland based on language, social or nation groups of the Indigenous Australians. The lines of latitude show the parallels of latitude 20°S and 25°S and meridians 145°E and 150°E. Use this information, other resources and the map to help describe boundary positions of three chosen regions.



ISBN 978-1-108-46154-2

Cambridge University Press

5C Time zones and time differences



ISBN 978-1-108-46154-2

Cambridge University Press

The **time zones** are largely determined by the meridians of longitude. You can see from the map on page 240 that there are exceptions to this because of local requirements.

If it is 12 noon along a meridian, then on the other side of the world along the meridian that makes up the other half of the great circle it is midnight. For example, when it is noon in Victoria on the 145°E meridian, it is midnight in the far east of Brazil on the 35°W meridian.

Since the Earth turns 360° in 24 hours, it turns 15° in 1 hour. For every 15° of longitude, the time difference is 1 hour, and so for every 1° of longitude the time difference is 4 minutes.

> $15°$ longitude $=$ 1 hour time difference
>
> $1°$ longitude $=$ 4 minutes time difference

Local times around the world are given relative to the time along the prime meridian. The time at the prime meridian is taken as **Coordinated Universal Time (UTC)**. This is not a time zone but is the primary time standard by which the world regulates clocks and time. For most purposes, UTC is considered interchangeable with **Greenwich Mean Time (GMT)** , but GMT is no longer precisely defined by the scientific community. In the map on page 240, the time zones show the adjustments to UTC, which are taken in various locations in the world.

Places east of the prime meridian are ahead of GMT (UTC), while places west are behind GMT (UTC).

Example 10 Using time difference without time zones or summer time

Singapore is located at 1°N 104°E and Sydney is located at 34°S 151°E. What is the time difference between Singapore and Sydney?

Solution

1 Calculate the difference in longitude.	*Difference in longitude = 151 − 104 = 47*
2 Use 1 hour for each 15° of longitude.	$47 \div 15 = 3\frac{2}{15}$ *Therefore, the time difference is 3 hours. Sydney is three hours ahead.*

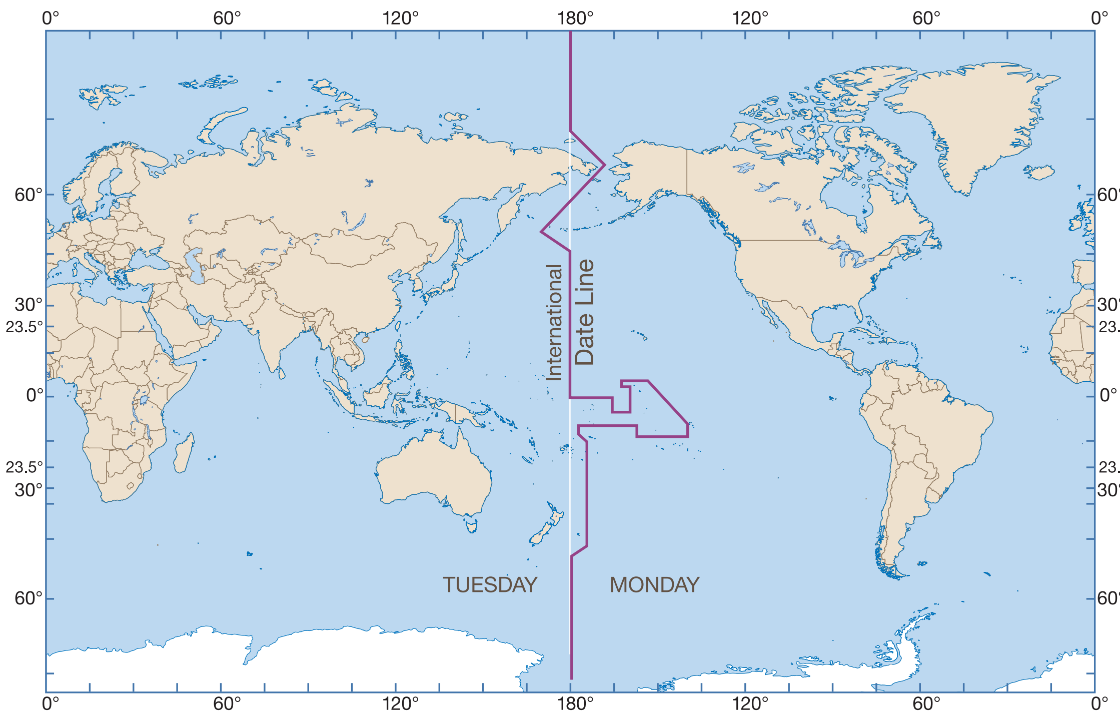
ISBN 978-1-108-46154-2

Cambridge University Press

International Date Line

The International Date Line is an imaginary line on Earth's surface. **The International Date Line** is located halfway around the world from the prime meridian (0° longitude) at about the 180° meridian. The dateline runs from the north pole to the south pole. It is not straight but zigzags to avoid political and country borders and to not cut some countries in half. The location of the international dateline is shown in the map below. When you cross the International Date Line:

- from west to east, you subtract a day
- from east to west, you add a day.

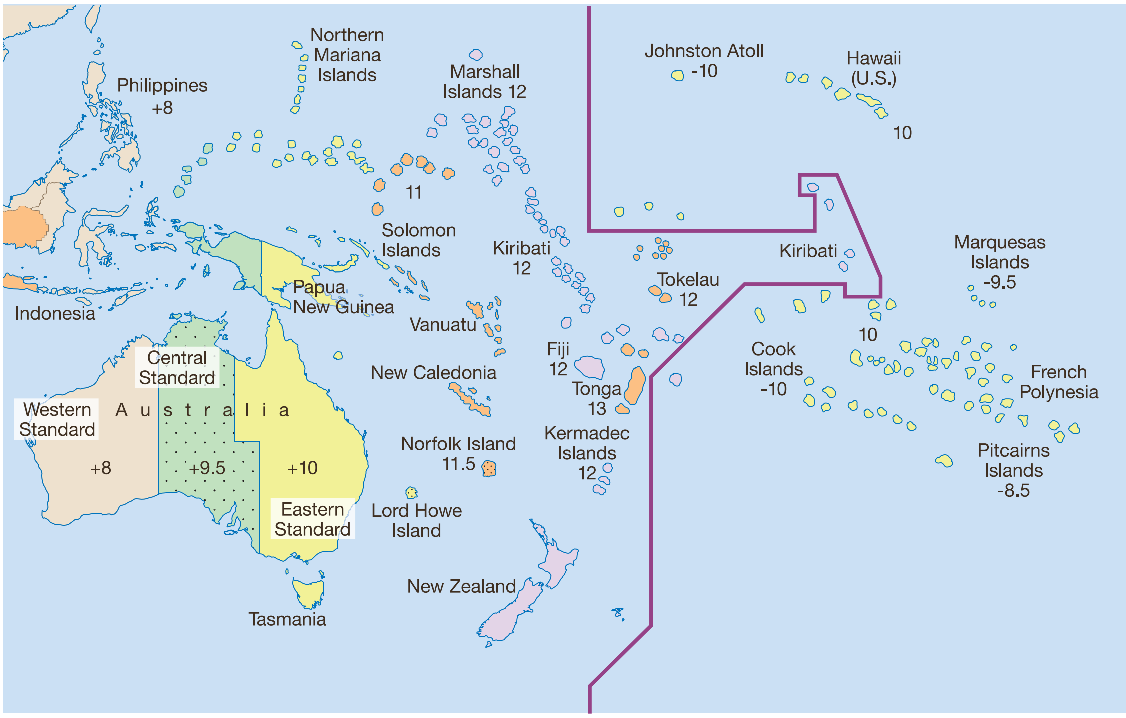


ISBN 978-1-108-46154-2

Cambridge University Press

Time zones for Australia and its neighbours

Australia is divided into three time zones: Eastern Standard, Central Standard and Western Standard. Some states change the clocks in summer to include daylight saving. Queensland time is 10 hours ahead of GMT. New Zealand time is ahead of eastern standard time by 2 hours and is 12 hours ahead of GMT.



ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 5C

Using time difference without time zones or summer time

Example 10

SF

1 Give the time differences between the following places.

a X longitude 150°E, Y longitude 120°E
b X longitude 0°, Y longitude 75°W
c X longitude 0°, Y longitude 75°E
d X longitude 123°E, Y longitude 150°E

2 The following table shows the latitude and longitude of cities around the world.

City	Latitude	Longitude
Brisbane	27°S	153°E
Melbourne	38°S	145°E
Marseilles	43°N	5°E
Mexico City	19°N	99°W
Wellington	41°S	175°E
Zurich	47°N	8°E
London	52°N	0° (actually west of Greenwich)
Lima	12°S	77°W
Plymouth	50°N	4°W
Yangon	19°N	96°E

a Find the time difference between Brisbane and:
  i Wellington **ii** Marseilles
  iii Yangon **iv** London
  v Mexico City

b If it is 6 a.m. in Brisbane what time is it in:
  i Wellington? **ii** Marseilles?
  iii Yangon? **iv** London?
  v Mexico City?

3 The time in a Pacific island, is 10 hours behind GMT.

a What is the longitude of the island?
b What is the time in London when it is 4 a.m. on the island?
c What is the time on the island when it is 5:30 a.m. in London?

4 The longitude of Hanoi is 105°E, while the longitude of Cape Howe is 150°E. When the time in Hanoi is 2:30 p.m., what is the time at:

a Cape Howe? **b** London?

ISBN 978-1-108-46154-2

5 Kalgoorlie has longitude 121°E while the Pacific island of Nauru has longitude 166°E.
 a Calculate the difference in longitude between these two places.
 b Calculate the time difference between the two places.
 c What is the time in Nauru when it is 11:45 a.m. in Kalgoorlie?

6 You live in Broadbeach (28.0308°S, 153.4319°E) and want to telephone a friend in one of the places listed below at 9 a.m. on a Saturday (their time). For each city, at what time (your time) should you call?
 a Vancouver (123°W)
 b Helsinki (25°E)
 c Bologna (11°E)
 d San Francisco (122°W)
 e Vladisvostock (132°E)
 f Fiji (178°E)

Using time zones to solve problems

CF

7 Dimitri is in Athens, which is two hours ahead of Greenwich Mean Time. Allan is in New York, which is five hours behind Greenwich Mean Time.
 a Dimitri is going to ring Allan at 10 a.m. on Wednesday, Athens time. What day and time will it be in New York when he rings?
 b Allan is going to fly from New York to Athens. His flight will leave on Wednesday at 10 p.m., New York time, and will take 10 hours. What day and time will it be in Athens when he arrives?

8 Louise is in Dubai, which is three hours ahead of Greenwich Mean Time. David is in Sydney, which is ten hours ahead of Greenwich Mean Time.
 a Louise is going to ring David at 10 a.m. on Wednesday, Sydney time. What day and time will it be in Dubai when she rings?
 b David is going to fly from Sydney to Dubai. His flight will leave on Wednesday at 10 p.m., Sydney time, and will take 14 hours. What day and time will it be in Dubai when he arrives?

9 Los Angeles is 8 hours behind GMT and Brisbane is 10 hours ahead of GMT.
 a If it is 10 a.m. in Brisbane on Tuesday, what time is it in Los Angeles?
 b A plane leaves Brisbane at 10 a.m. on Tuesday and arrives in Los Angeles at 6 a.m. on Tuesday. What was the length of the flight? (Note that the plane crosses the International Date Line)
 c A plane leaves Los Angeles at 23:20 on Tuesday and arrives in Brisbane at 7:15 a.m. on Thursday. What was the length of the flight? (Note that the plane crosses the International Date Line)

Using time zones of Australia and its neighbours

SF

10 Use the map on page 240 to answer these questions using the time zones indicated on the map. Ignore summer time. You may need to find the location in an atlas or using the internet. If it 12 midday in Brisbane. what time and day is it in:
 a Melbourne?
 b Alice Springs?
 c Perth?
 d Port Moresby?
 e Auckland?
 f Norfolk Island?
 g Suva?
 h Honolulu?
 i Solomon Islands?

ISBN 978-1-108-46154-2

Cambridge University Press

Key ideas and chapter summary

Length of an arc — AS

The length s of an arc of a circle of radius r that subtends an angle of θ at the centre is given by

$$s = \frac{\pi r \theta}{180}$$

Great circle

A **great circle** is a section of a sphere that contains a diameter of the sphere. The section contains the centre of the sphere.

Small circles

Sections of the sphere that do not contain a diameter are called **small circles**. A small circle does not contain the centre of the sphere.

Meridians of longitude

Meridians of longitude are semi-great circles that pass through the north and south poles.

Parallels of latitude

Parallels of latitude are small circles whose planes are parallel to that of the equator.

Australian time zones

Australia is divided into three time zones: Eastern Standard, Central Standard and Western Standard. Some states change the clocks in summer to include daylight saving.

Skills check

Having completed this chapter, you should be able to:

- find the length of an arc of a circle
- convert degrees expressed in decimal form to degrees and minutes and vice versa
- use the meridians of longitude and parallels of latitude to describe locations on Earth
- determine a distance around a parallel of latitude
- find the distance around a great circle of two points with the same longitude
- relate time zones with longitude.

Multiple-choice questions

1 A great circle on a newly found planet has a circumference of 11 000 km. The diameter of the planet is closest to:

A 320 km **B** 300 km **C** 3500 km
D 250 km **E** 1100 km

ISBN 978-1-108-46154-2

2 The coordinates of two points M and N on the Earth's surface are (40°N, 40°E) and (25°S, 55°E). Which statement is most likely to be correct about the time difference?

A M is 5 hours behind N.
B M is 1 hour behind N.
C N is 5 hours behind M.
D N is 1 hour behind M.
E M is 15 hours ahead of N.

3 Point X on the Earth's surface has coordinates (29°S, 32°E), while point Y is at (8°S, 32°E). The distance between X and Y is closest to:

A 2335 km
B 750 km
C 111 km
D 1350 km
E 2010 km

4 X and Y are two towns on the equator. The longitude of X is 18°E and the longitude of Y is 48°W. Approximately how far apart are these two towns?

A 4000 km
B 7400 km
C 8500 km
D 10 100 km
E 7340 km

5 Trevor lives in Albany, which has a longitude of 118°E. He wants to watch a basketball game being played in Ottawa, which has a longitude of 76°W. The game starts at 10 p.m. on Wednesday Ottawa time. What is the time in Albany when the game starts? (Ignore time zones and daylight saving.)

A 9 a.m. on Wednesday
B 11 a.m. on Wednesday
C 1 a.m. on Thursday
D 9 a.m. on Thursday
E 11 a.m. on Thursday

6 Perth in Western Australia is 8 hours ahead of GMT. Pretoria in South Africa is 2 hours ahead of GMT. What is the time in Pretoria when it is 1 p.m. in Perth?

A 3 a.m.
B 7 a.m.
C 11 a.m.
D 9 p.m.
E 11 p.m.

7 Stockholm has coordinates 59°N, 18°E and Darwin has coordinates 13°S, 131°E. What is the time difference between Stockholm and Darwin? (Ignore time zones and daylight saving.)

A 184 minutes
B 288 minutes
C 452 minutes
D 596 minutes
E 620 minutes

Short-answer questions

SF

1 Two locations lie on the same meridian of longitude. One is 30° north of the other. What is the distance between the two locations, correct to the nearest kilometre?

2 Two locations lie on the same meridian of longitude. One is 35° south of the other. What is the distance between the two locations, correct to the nearest kilometre?

ISBN 978-1-108-46154-2

Cambridge University Press

3 Dunedin has longitude 170°E, while Albany has longitude 118°E.
  a Calculate the difference in longitude between these two places.
  b Calculate the time difference between the two places (ignore time zones and daylight saving).
  c What is the time in Dunedin when it is 2:45 a.m. in Albany? (Ignore time zones and daylight saving.)

4 The position of Rabaul is (4°S, 152°E). An island is 4° to the north of Rabaul and 48° east of Rabaul. Give the latitude and longitude of the island.

5 Pontianak has a longitude of 109°E, and Amazonas, a town in Brazil, has a longitude of 70°W. Both places lie on the equator. Find the shortest distance between these two places. Give you answer to the nearest kilometre.

6 Singapore has longitude 104°E and Sydney has longitude 151°E. What is the time difference between Singapore and Sydney? (Ignore time zones and daylight saving.)

7 Anthony lives in Rockhampton and wants to phone his grandfather in London. It is 6 p.m. on Saturday in Queenland. What time is it in London? (London GMT +0, Rockhampton GMT +10)

8 Arlene lives in Townsville and has a baby at 2 a.m. on Saturday. She wants to phone her mother who is on holiday in Samoa with the good news. What time is it in Samoa when she calls? (Townsville GMT +10, Samoa GMT +13)

9 Arisa is from Japan and is studying in Brisbane. She plans to phone home on Sunday night at 8 p.m. What time is it in Tokyo? (Brisbane GMT +10, Tokyo GMT +9)

10 Terry is playing hockey at a tournament in Buenos Aires. After her team wins the semi-final at 5 p.m. on Friday, she phones her father in Toowoomba to tell him the news. What time is it in Toowoomba? (Toowoomba GMT +10, Buenos Aires GMT –3)

11 Elizabeth is in Rome, which is one hour ahead of Greenwich Mean Time. Leslie is in Boston, which is five hours behind Greenwich Mean Time. CF
  a Elizabeth is going to ring Leslie at 10 p.m. on Tuesday, Rome time. What day and time will it be in Boston when she rings?
  b Elizabeth is going to fly from Boston to Rome. Her flight will leave on Wednesday at 10 a.m., Boston time, and will take 10 hours. What day and time will it be in Rome when she arrives?

 ISBN 978-1-108-46154-2 Cambridge University Press

12 Osaka is at 34°N, 135°E, and Dallas is at 33°N, 97°W.

- **a** Find the time difference between the two cities. (Ignore time zones.)
- **b** Rex lives in Dallas and wants to ring a friend in Osaka. In Dallas it is 8 p.m. Monday. What time and day is it in Osaka?
- **c** Rex's friend in Osaka sent him a text message, which happened to take 14 hours to reach him. It was sent at 10 a.m. Thursday, Osaka time. What was the time and day in Dallas when Rex received the text?

Extended-response questions

CU

1 Two locations X and Y have the same latitude 30°S. The longitude of X is 145°E and the longitude of Y is 130°E. (In this question take the radius of the Earth to be 6400 km.)

- **a** Find the radius of the small circle of the 30°S parallel of latitude.
- **b** Find the distance between X and Y around the parallel of latitude 30°S.
- **c** Find the radius of the great circle through X and Y.
- **d** Find the difference between the distance around the great circle and the distance around the parallel of latitude. Give answer correct to the nearest km.

2 The Tropic of Cancer is at latitude 23.5°N, while the Tropic of Capricorn is at latitude 23.5°S.

- **a** Calculate the distance between these two tropics along the same great circle (correct to the nearest km).
- **b** Calculate the radius of the small circles of the Tropic of Capricorn and the Tropic of Cancer.
- **c** Rockhampton in Queensland and Sao Paulo in Brazil are both on the Tropic of Capricorn. Rockhampton has longitude 150.5°E and Sao Paolo has longitude 46.5°W. Find the distance around the Tropic of Capricorn from Sao Paolo to Rockhampton.

ISBN 978-1-108-46154-2

Cambridge University Press

6 Revision of Unit 3 Chapters 1–5

UNIT 3 BIVARIATE DATA, SEQUENCES AND CHANGE, AND EARTH GEOMETRY

Topic 1 Bivariate data analysis

Topic 2 Time series analysis

Topic 3 Growth and decay in sequences

Topic 4 Earth geometry and time zones

The revision exercises are arranged by chapter with these categories of questions:

- Simple familiar question types
- Complex familiar question types
- Complex unfamiliar question types
- Problem-solving and modelling questions

ISBN 978-1-108-46154-2

Cambridge University Press

Revision

6A Topic 1 Bivariate data analysis

Multiple-choice questions

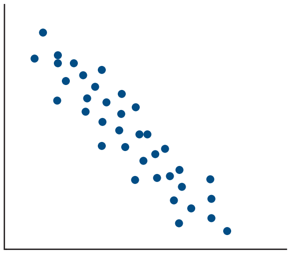
1 The table below shows the percentage of respondents in three age groups (18–25 years, 26–44 years, 45–60 years) and their response to the question 'Are you satisfied with your career choice?'

	Age group			
Are you satisfied with your career choice?	18–25	26–44	45–60	Total
Yes	65%	45%	60%	57%
No	35%	55%	40%	43%
Total	100%	100%	100%	100%

For the people surveyed, which of the following statements, by itself, supports the contention that there is an association between satisfaction with career choice and age group?

A 65% of 18–25-year-olds are satisfied with their career choice, and 35% are not.
B 65% of 18–25-year-olds are satisfied with choice of work and 35% of 18–25-year-olds are not.
C Only 45% of 26–44-year-olds are satisfied with their career choice, much less than the 60% of 45–60-year-olds, and 65% of 18–25-year-olds who are satisfied.
D 65% of 18–55-year-olds are satisfied with their career choice, more than the 55% of 26–44-year-olds, and 40% of 45–60-year-olds who are not.
E Overall 57% of respondents are satisfied with their career choice, more than the 43% who are not.

2 The value of r for the scatterplot is closest to:



A 0.8
B 0.5
C 0
D −0.5
E −0.9

3 The association pictured in the scatterplot in the previous question is best described as:

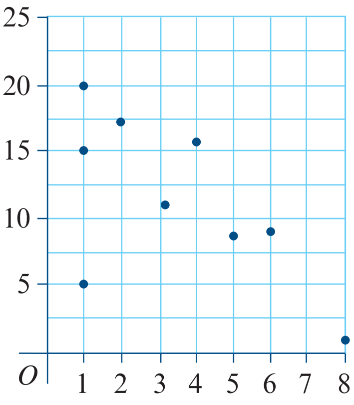
A strong, positive, linear
B strong, negative, linear
C weak, negative, linear
D strong, negative, non-linear with an outlier
E strong, negative, non-linear

ISBN 978-1-108-46154-2

Cambridge University Press

Revision

4 When the correlation coefficient, r, was calculated for the data displayed in the scatterplot, it was found to be $r = -0.64$.
If the point (1, 5) was replaced with the point (6, 5) and the correlation coefficient, r, recalculated, then the value of r would be:



A unchanged
B positive but closer to 1
C negative but closer to 0
D positive but closer to 0
E negative but closer to –1

5 The correlation between computer ownership (number of computers/1000 people) and car ownership (number of computers/1000 people) in six countries is $r = 0.92$, correct to two decimal places.
Based on this information, which of the following statements is *not* true?
A Computer ownership and car ownership are both numerical variables.
B Around 85% of the variation in computer ownership is explained by car ownership.
C Either *computer ownership* or *car ownership* could be the explanatory variable.
D For these countries, computer ownership increases as car ownership increases.
E For these countries, computer ownership decreases as car ownership increases.

6 The correlation between the score on a maths test and height for a group of primary school students is found to be 0.7. From this information, it is reasonable to conclude that:
A learning maths makes children grow taller.
B there is no association between height and maths test scores.
C a child's maths ability depends only on their height.
D the children who obtained high maths test scores tended to be taller.
E all tall children are better at maths than shorter children.

7 Given that $r = 0.675$, $s_x = 2.567$ and $s_y = 4.983$, the slope of the least squares regression line $y = a + bx$ is closest to:
A 0.35 **B** 0.68 **C** 1.3 **D** 1.7 **E** 3.36

The following data relate to Questions 8 and 9.

Number of hot dogs sold	190	168	146	155	150	170	185
Temperature (°C)	10	15	20	15	17	12	10

We wish to determine the equation of the least squares regression line for the data that will enable the number of hot dogs sold to be predicted from temperature.

ISBN 978-1-108-46154-2

Cambridge University Press

8 The slope of the regression line will be closest to:

A −4.3 **B** −0.2 **C** 0.2 **D** 4.3 **E** 227

9 The coefficient of determination will be closest to:

A −0.94 **B** −0.89 **C** 0.21 **D** 0.89 **E** 0.94

The following information relates to Questions 10 to 15.

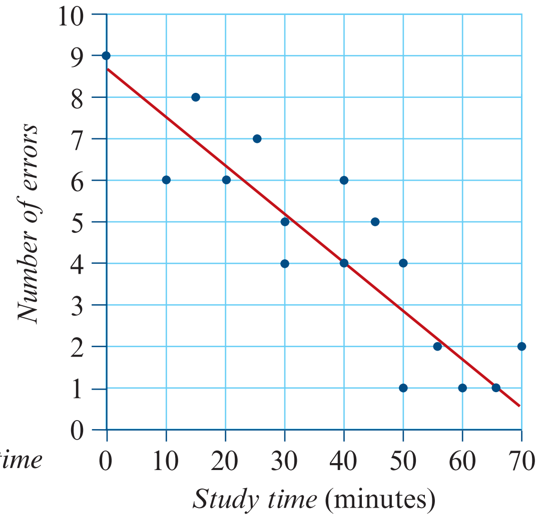
Eighteen students sat for a 15-question multiple-choice test. In the scatterplot opposite, the number of errors made by each student on the test is plotted against the time they reported studying for the test.

A least squares regression line has been determined for the data and is also displayed on the scatterplot.

The equation for the least squares regression line is:

number of errors = 8.8 − 0.12 × *study time*

and the coefficient of determination is 0.8198.



10 The least squares regression line predicts that a student reporting a study time of 35 minutes would make:

A 4.3 errors **B** 4.6 errors **C** 4.8 errors

D 5.0 errors **E** 13.0 errors

11 The value of Pearson's correlation coefficient, r, is closest to:

A −0.91 **B** −0.82 **C** 0.67 **D** 0.82 **E** 0.91

12 The student who reported a study time of 10 minutes made six errors. The predicted score for this student would have a residual of:

A −7.6 **B** −1.6 **C** 0 **D** 1.6 **E** 7.6

13 Which of the following statements that relate to the regression line are *not* true?

A The slope of the regression line is −0.12.

B The equation predicts that a student who spends 40 minutes studying will make approximately four errors.

C The least squares regression line does *not* pass through the origin.

D On average, a student who does not study for the test will make around 8.8 errors.

E The explanatory variable in the regression equation is *number of errors*.

ISBN 978-1-108-46154-2

Cambridge University Press

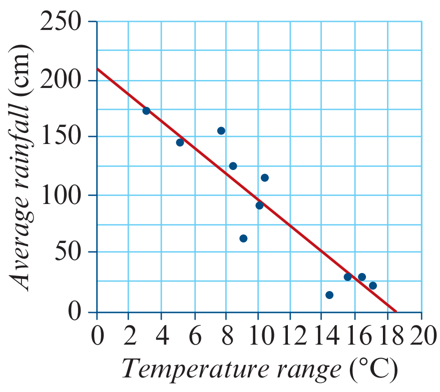
14 This regression line predicts that, on average, the number of errors made:

A decreases by 0.82 for each extra minute spent studying.
B decreases by 0.12 for each extra minute spent studying.
C increases by 0.12 for each extra minute spent studying.
D increases by 8.8 for each extra minute spent studying.
E decreases by 8.8 for each extra minute spent studying.

15 Given that the coefficient of determination is 0.8198, we can say that close to:

A 18% of the variation in the number of errors made can be explained by the variation in the time spent studying.
B 33% of the variation in the number of errors made can be explained by the variation in the time spent studying.
C 67% of the variation in the number of errors made can be explained by the variation in the time spent studying.
D 82% of the variation in the number of errors made can be explained by the variation in the time spent studying.
E 95% of the variation in the number of errors made can be explained by the variation in the time spent studying.

16 The average rainfall and temperature range at several locations in the South Pacific region are displayed in the scatterplot opposite.



A least squares regression line has been fitted to the data, as shown. The equation of this line is closest to:

A *average rainfall* = 210 − 11 × *temperature range*
B *average rainfall* = 210 + 11 × *temperature range*
C *average rainfall* = 18 − 0.08 × *temperature range*
D *average rainfall* = 18 + 0.08 × *temperature range*
E *average rainfall* = 250 − 13 × *temperature range*

ISBN 978-1-108-46154-2

Cambridge University Press

Short-answer questions

Simple familiar questions

1 For the following pairs of variables, classify each variable as either categorical or numerical, and choose which of the following analysis techniques you would use to investigate the association between them:

- two-way frequency table
- scatterplot
- parallel boxplots

a age (years) and reaction time (seconds)
b sex (males, female) and reaction time (seconds)
c sex (male, female) and reaction time (fast, average, slow)

2 The data in the table below is based on a study of dolphin behaviour. In this study, the main activities of dolphins observed in the wild were classified as 'travelling', 'feeding', and 'socialising'. The time of day was also noted.

	Time of observation		
Activity	Morning	Afternoon	Evening
Travelling	11.4%	53.3%	16.5%
Feeding	38.0%	6.7%	70.9%
Socialising	50.6%	40.0%	12.6%
Total	100.0%	100.0%	100.0%

a Which is the explanatory variable and which is the response variable?
b Of the dolphins who were observed in the morning, what percentage were feeding?
c Does the information in the table support the contention that the behaviours of the dolphins are associated with the time of day? Justify your answer by quoting appropriate percentages.

3 The number of hours spent studying for an examination by each member of a class, and the marks they received are given in the table:

Student	1	2	3	4	5	6	7	8	9	10
Hours	4	36	23	28	25	11	18	13	4	8
Mark	27	87	67	84	66	52	61	43	38	52

a Which of these variables is the explanatory variable and which is the response variable?
b On graph paper, construct a scatterplot to display the data.
c Describe the scatterplot.

ISBN 978-1-108-46154-2

Cambridge University Press

4 The following table gives the life expectancies, in years, for males and females across a group of countries.

	Male life expectancy (years)	Female life expectancy (years)
Australia	80.40	84.50
Canada	79.60	83.80
China (People's Republic of)	74.50	77.50
France	79.20	85.50
India	66.90	69.90
Indonesia	67.00	71.20
Italy	80.30	84.90
Japan	80.80	87.10
Mexico	72.30	77.70
Russia	65.90	76.70
South Africa	55.50	59.50
United Kingdom	79.20	82.80
United States	76.30	81.20

a For these data, calculate the value of the correlation coefficient r, and interpret.

b What assumptions have you made about the relationship between male and female life expectancies in these countries for the calculation of the correlation coefficient to be valid?

5 The following table gives the adult heights (in cm) of ten pairs of mothers and daughters:

Mother	170	163	157	165	175	160	164	168	152	173
Daughter	178	175	165	173	168	152	163	168	160	178

a Identify which variable is the explanatory variable and which is the response variable.

b Construct a scatterplot of the data, and then describe the association between mother's height and daughter's height.

c Calculate the value of the correlation coefficient r, and classify its strength.

d Find the equation of the least squares regression line, and interpret the intercept and slope.

e Find the value of the coefficient of determination, R^2, and interpret in terms of the variables in this question.

f Estimate the adult height of a girl whose mother is 170 cm tall.

g Is your prediction in part **f** interpolation or extrapolation? Explain.

 ISBN 978-1-108-46154-2 Cambridge University Press

6 The following table shows the daily maximum temperature and the number of ice-creams sold at a kiosk on the beach over a nine-day period:

Temperature	Sales
18	280
21	298
22	333
24	359
25	360
26	355
27	378
32	427
36	465

The equation of the least squares regression line that allows the number of ice-creams sold to be predicted from the temperature is:

sales = 97.2 + 10.3 × *temperature*

a Complete the following table of the residuals.

Temperature	18	21	22	24	25	26	27	32	36
Residual	−2.6		9.2	14.6	5.3	−10.0		0.2	−3.0

b Construct a residual plot, and comment on the linearity assumption.

7 A marketing firm wanted to investigate the relationship between the number of times a song was played on the radio (*played*) and the number of downloads sold the following week (*weekly sales*).
The following data was collected for a random sample of ten songs.

Played	47	34	40	34	33	50	28	53	25	46
Weekly sales	3950	2500	3700	2800	2900	3750	2300	4400	2200	3400

a Which is the explanatory variable and which is the response variable?
b Construct a scatterplot of this data.
c Determine the value of the Pearson correlation coefficient, r, for this data.
d Describe the relationship between *weekly sales* and *played* in terms of direction, strength and form and outliers (if any).
e Determine the equation for the least squares regression line and write it down in terms of the variables *weekly sales* and *played*.
f Interpret the slope and intercept of the least squares regression line in the context of the problem.
g Use your equation to predict the number of downloads of a song when it was played on the radio 100 times in the previous week.
h In making this prediction, are you interpolating or extrapolating?

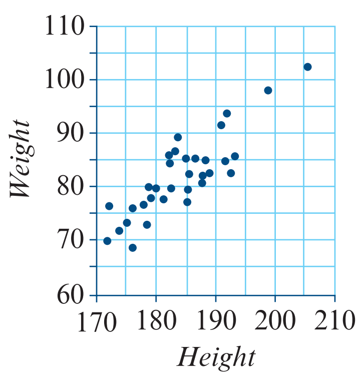
 ISBN 978-1-108-46154-2 Cambridge University Press

8 The data below shows heights and weights of 35 players listed to play for an AFL team.

Height (centimetres)	Weight (kilograms)	Height (centimetres)	Weight (kilograms)
174	72	180	80
179	78	178	77
183	85	198	98
193	86	176	69
173	76	181	78
186	85	206	103
188	85	188	81
191	91	184	86
183	80	175	73
187	85	192	92
178	73	185	80
172	70	193	83
188	82	185	77
183	84	185	89
185	83	185	89
189	83	176	76
179	80	192	85
185	83		

a In this analysis, which would be the response variable and which would be the explanatory variable?

b A scatterplot for this data is shown below. Does this support the use of a least squares regression line to model the relationship between the weight and height of these footballers? Explain your answer.



ISBN 978-1-108-46154-2

Cambridge University Press

c Calculate the least squares regression line for this data and write it in terms of the variables under investigation.

d What is the slope of the regression line, and what does it tell you in terms of the problem at hand?

e Why does it not make sense to try to interpret the y-intercept in this problem?

f On average, how much do you expect players who are 195 cm tall to weigh?

g Calculate the value of the coefficient of determination and interpret in terms of these variables.

Complex familiar questions

9 A caterer collected the following data on the cost to the company of the preparation of a differing number of meals.

Number of meals (x)	30	70	90	25	50	60	75	100
Cost in dollars (y)	345	595	720	300	485	530	585	750

a Using the method of least squares regression, find the equation of a straight line that relates the two variables.

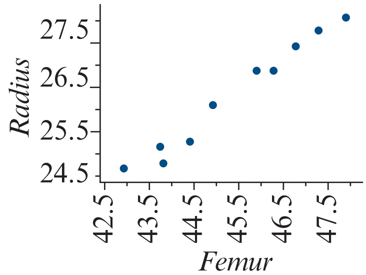
b What is the caterer's base cost for the standard menu?

c What is the cost of each meal, over and above this base cost?

ISBN 978-1-108-46154-2

Cambridge University Press

10 A regression analysis was conducted to investigate the nature of the relationship between femur (thigh bone) length and radius (the short thicker bone in the forearm) length in 18-year-old males. The bone lengths are measured in centimetres. The results of this analysis are reported below. In this investigation, femur length was treated as the explanatory variable.



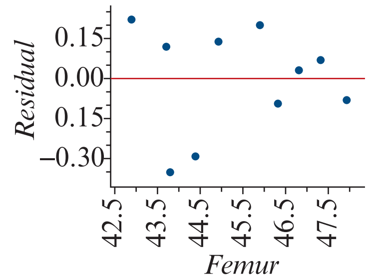
Regression equation $y = a + bx$

$a = -7.24946$

$b = 0.739556$

$R^2 = 0.975291$

$r = 0.987568$



Based on these analyses, write a report describing the association between femur length and radius.

11 The following table gives the correlation coefficients between infant mortality rate and a range of possible explanatory variables for countries across the world.

	Mortality rate, infant (per 1000 live births)
Birth rate per 1000 people	$r = 0.876$
Births attended by skilled health staff (% total)	$r = -0.714$
Exclusive breastfeeding (% of children under 6 months)	$r = 0.170$
Health expenditure per capita (current US$)	$r = -0.495$
Literacy rate, adult female (% of females ages 15 and above)	$r = -0.800$
Literacy rate, adult male (% of males ages 15 and above)	$r = -0.849$
People using safely managed sanitation services (% of population)	$r = -0.636$
People using safely managed drinking water services (% of population)	$r = -0.817$

a Determine the values of the coefficient of determination for each of these variables.

b Discuss the relative importance of each of the explanatory variables in understanding infant mortality.

ISBN 978-1-108-46154-2

Cambridge University Press

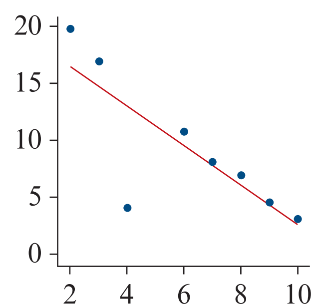
12 In a large university, students were offered the choice of attending traditional lectures and tutorials on campus or working independently using an online method of instruction. The study intentions were recorded for a random sample of full-time students from three different faculties (Arts, Business and Science). Researchers suggested that students enrolled in the Arts faculty would be more likely to select the traditional on campus study than students from other faculties. Write a report addressing the researcher's hypothesis.

		Faculties			
		Arts	Science	Business	Total
Type of instruction	On campus	102 63.8%	128 64.0%	58 41.4%	288
	Online	58 36.3%	72 36.0%	82 58.6%	212
		160 100%	200 100%	140 100%	500

Extended-response questions

Complex unfamiliar questions

1 When the Pearson correlation coefficient was calculated for the data displayed in the following scatterplot, it was found to be −0.433, and the slope of the least squares regression line is −1.8.



a If the outlier is removed, what would be the effect on the value the correlation coefficient?

b If the outlier is removed, what would be the effect on the slope (b) of the least squares regression line?

2 Suppose that the manager of a company determines that the cost of manufacturing a pair of jeans is:

$$cost = 12\,500 + 22.90 \times number\ of\ pairs\ of\ jeans\ produced$$

a What is the marginal cost of manufacturing each pair of jeans?

b What is the cost per pair of jeans if 100 pairs of jeans are produced?

c For the manufacturer to make 75% profit on each pair of jeans, how much should he sell them each for:

i if 100 pairs of jeans are produced?

ii if 500 pairs of jeans are produced?

ISBN 978-1-108-46154-2

Cambridge University Press

3 **a** A sample of 608 males were asked whether Australia should retain the Queen or become a republic. Their answers, together with the political affiliation of the respondent, are summarised in the following two-way frequency table.

	Political affiliation		
Queen or republic	Liberal	Labor	Total
Definitely keep Queen	74	28	102
Probably keep Queen	60	39	99
Probably become republic	81	78	159
Definitely become republic	47	201	248
	262	346	608

Is there a relationship between political affiliation and attitude to the monarchy in this sample of males? Discuss quoting appropriate percentages.

b A sample of 560 females were asked whether Australia should retain the Queen or become a republic. Their answers, together with the political affiliation of the respondent, are summarised as follows.

	Political affiliation		
Queen or republic	Liberal	Labor	Total
Definitely keep Queen	93	51	144
Probably keep Queen	67	43	110
Probably become republic	49	72	121
Definitely become republic	38	147	185
	247	313	560

Is there a relationship between political affiliation and attitude to the monarchy in this sample of females? Discuss quoting appropriate percentages.

c Is the relationship between political affiliation and attitude to the monarchy the same for the females as for the males? Discuss.

ISBN 978-1-108-46154-2

Cambridge University Press

4 The following table displays the distance fallen by an object across one-second intervals.

Time (seconds)	Distance (metres)
0	0
1	5.2
2	18
3	42
4	79
5	128
6	168

a **i** Construct a scatterplot of the data, with time as the explanatory variable and distance as the response variable.

ii Determine the equation of the least squares regression line for the data.

iii Construct a residual plot, and comment on the linearity assumption.

b To find a better rule relating time and distance, the science teacher suggested the student consider $(\text{time})^2$ as the explanatory variable.

i Complete the following table:

Time (seconds)	$(\text{Time})^2$	Distance (metres)
0		0
1		5.2
2		18
3		42
4		79
5		128
6		168

ii Construct a scatterplot of the data, with $(\text{time})^2$ as the explanatory variable and distance as the response variable.

iii Determine the equation of the least squares regression line for the data.

iv Construct a residual plot, and comment on the linearity assumption.

ISBN 978-1-108-46154-2

Cambridge University Press

Problem-solving and modelling

5 The table below gives the data pertaining to Australian One Day International cricket captains.

No.	Name	Period of captaincy	Played	Won	Lost	Tied	No result
1	Bill Lawry	1971	1	1	0	0	0
2	Ian Chappell[6]	1972–1975	11	6	5	0	0
3	Greg Chappell	1975–1983	49	21	25	0	3
4	Bob Simpson	1978	2	1	1	0	0
5	Graham Yallop	1979	4	2	1	0	1
6	Kim Hughes	1979–1984	49	21	23	1	4
7	David Hookes	1983	1	0	1	0	0
8	Allan Border	1985–1994	178	107	67	1	3
9	Ray Bright	1986	1	0	1	0	0
10	Geoff Marsh	1987–1991	4	3	1	0	0
11	Mark Taylor	1992–1997	67	36	30	1	0
12	Ian Healy	1996–1997	8	5	3	0	0
13	Steve Waugh	1997–2002	106	67	35	3	1
14	Shane Warne	1998–1999	11	10	1	0	0
15	Adam Gilchrist	2001–2007	17	12	4	0	1
16	Ricky Ponting[7]	2002–2012	229	164	51	2	12
17	Michael Hussey	2006–2007	4	0	4	0	0
18	Michael Clarke	2008–2015	74	50	21	0	3
19	Cameron White	2011	1	1	0	0	0
20	Shane Watson	2012–2013	9	5	3	1	0
21	George Bailey	2013–2015	29	16	10	0	3
22	Steve Smith	2015–present	51	25	23	0	3
23	David Warner	2016	3	3	0	0	0
24	Aaron Finch	2017	2	0	2	0	0
Grand total			**911**	**556**	**312**	**9**	**34**

Use the data to investigate the question ‘Who is Australia’s most successful one-day international cricket captain?’

6 Investigate the association between pulse rate before and after exercise, collecting data as required. Determine if this association is the same for a range of other variables, such as sex, age group, or whether or not the subject regularly participates in exercise.

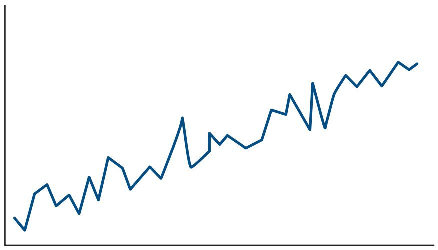
ISBN 978-1-108-46154-2

Cambridge University Press

6B Topic 2 Time series analysis

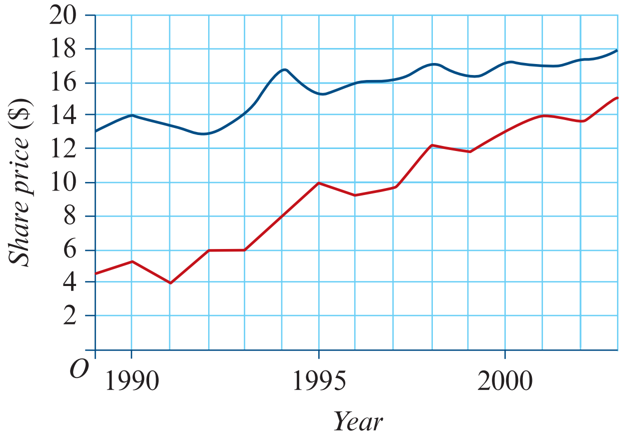
Multiple-choice questions

1 The pattern in the time series graph shown is best described as:



A trend
B cyclical, but not seasonal
C seasonal
D random
E average

2 The time series plot shows the share price of two companies over a period of time.
From the plot, it can be concluded that over the period 1990–2000, the *difference* in share price between the two companies has shown:



A a decreasing trend
B an increasing trend
C seasonal variation
D a 5-year cycle
E no trend

Use the information in the table below to answer Questions 3 to 6.

t	1	2	3	4	5	6	7	8	9	10
y	4	5	4	4	8	6	9	10	9	12

3 The three-smoothed mean for $t = 2$ is closest to:
A 4.3 **B** 6.2 **C** 6.4 **D** 6.5 **E** 7.25

4 The five-smoothed mean for $t = 5$ is closest to:
A 4.3 **B** 6.2 **C** 6.4 **D** 6.5 **E** 7.25

5 The centred two-smoothed mean for $t = 6$ is closest to:
A 4.3 **B** 4.75 **C** 6.25 **D** 6.5 **E** 7.25

6 The centred four-smoothed mean for $t = 3$ is closest to:
A 4.3 **B** 4.75 **C** 6.25 **D** 7.25 **E** 9.75

ISBN 978-1-108-46154-2

Cambridge University Press

Use the following information to answer Questions 7 and 8.

The *long-term* quarterly sales figures of a car dealer are shown in the table below. Also shown are the seasonal indices for the first and second quarters.

Quarter	1	2	3	4
Sales (number)	21	36	49	28
Seasonal index	0.6	1.0		

7 The car dealer sells 18 cars in the first quarter of this year. The deseasonalised number sold is:

A 11 **B** 13 **C** 18 **D** 20 **E** 30

8 The seasonal index for the fourth quarter is closest to:

A 0.6 **B** 0.8 **C** 1.1 **D** 1.5 **E** 4.0

Use the information below to answer Questions 9 to 10.

The quarterly sales figures for a soft drink company and the seasonal indices are as shown.

Quarter	1	2	3	4
Sales ($'000s)	1200	1000	800	1200
Seasonal index	1.1	0.90	0.8	

9 The deseasonalised figure (in $'000s) for quarter 3 is:

A 640 **B** 667 **C** 800 **D** 1000 **E** 1500

10 The seasonal index for quarter 4 is:

A 0.6 **B** 0.8 **C** 1.00 **D** 1.1 **E** 1.2

11 The deseasonalised sales (in dollars) for June were $91 564. The seasonal index for June is 1.45.
The *actual* sales for June were closest to:

A $41 204 **B** $61 043 **C** $63 148 **D** $91 564 **E** $132 768

12 Sales for a major department store are reported quarterly. The seasonal index for the third quarter is 0.85. This means that sales for the third quarter are typically:

A 85% below the quarterly average for the year
B 15% below the quarterly average for the year
C 15% above the quarterly average for the year
D 18% above the quarterly average for the year
E 18%below the quarterly average for the year

 ISBN 978-1-108-46154-2 Cambridge University Press

Revision

13 The seasonal index for headache tablet sales in summer is 0.80.
To correct for seasonality, the headache tablet sales figures for summer should be:

A reduced by 80% **B** reduced by 25% **C** reduced by 20%
D increased by 20% **E** increased by 25%

14 The table below shows the number of broadband users in Australia for the years 2004 to 2008.

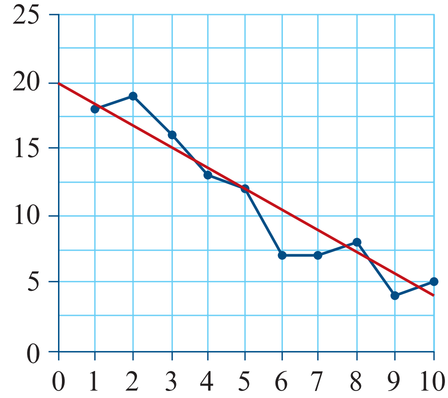
Year	2004	2005	2006	2007	2008
Number	1012000	2016000	3900000	4830000	5140000

A two-point moving mean with centring is used to smooth the time series.
The smoothed value for the number of broadband users in Australia in 2006 is:

A 2958000 **B** 3379600 **C** 3455500
D 3661500 **E** 3900000

15 A time series for y is shown in the graph, where t represents time. If a linear trend line is fitted to this data, as shown, then the equation of the line is closest to:

A $y = 20 - 1.6t$
B $y = -1.6t$
C $y = 20 + 1.6t$
D $y = 20 - 0.6t$
E $y = 20 + 0.6t$



ISBN 978-1-108-46154-2

Cambridge University Press

Extended-response questions

Simple familiar questions

1 Construct a time series plot of the number of new vehicles purchased in Queensland in 2016 given in the following table and describe the plot.

Month	New vehicle sales
Jan	17 193
Feb	18 711
Mar	21 470
Apr	17 753
May	19 565
Jun	27 270
Jul	18 445
Aug	18 062
Sep	19 702
Oct	16 870
Nov	18 733
Dec	19 252

2 The number of purchases made at a shop over the period of three years from 2016 to 2018 is recorded in the table below.

Number of purchases	Summer	Autumn	Winter	Spring
2016	1380	1627	1840	720
2017	1552	1770	2056	725
2018	1949	1986	2150	990

a Construct a time series graph for this data and describe the plot.

b The seasonal indices for autumn and winter are 1.15 and 1.30.

  i Calculate the seasonal index for summer.

  ii Hence, find the seasonal index for spring.

c Construct a table to show the number of purchases after deseasonalisation.

d Construct a time series plot the deseasonalised data.

e Determine the equation of the least squares regression line for this time series (deseasonalised data vs quarter number).

ISBN 978-1-108-46154-2

Cambridge University Press

f Draw the least squares regression line on your scatterplot.

g Interpret the value of the gradient of the least squares regression line in this case.

h Determine the predicted number of purchases for summer 2020.

i Calculate the coefficient of determination for the deseasonalised data for number of purchases vs time (quarter number). Interpret this result in terms of the variables involved.

3 The following table gives the annual percentage of young people (aged 15–24) who were unemployed in Australia for the years 2007–2016.

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
% youth unemployed	9.37	8.82	11.46	11.55	11.38	11.73	12.22	13.32	13.14	12.50

a Construct a time series plot of these data.

b Calculate a three-point moving average of the percentage of young people who are unemployed in Australia.

c Superimpose the moving average on the time series plot of the data.

d Comment on the smoothed graph.

4 The earnings for one particular person over a period of 12 months are as shown.

Month	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Income $000's	5.0	2.6	5.2	6.0	2.4	3.2	0.2	8.4	6.2	3.2	3.6	4.0

a Determine the three-point moving mean for March.

b Determine the five-point moving mean for October.

c The seasonal indices for 12 months are shown with the exception of December.

Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1.200	0.672	1.248	1.44	0.576	0.768	0.100	2.016	1.488	0.768	0.864	

i Find the seasonal index for December.

ii In the following year the person earns $4200 in January. What is this equivalent to in deseasonalised terms?

ISBN 978-1-108-46154-2

Cambridge University Press

Complex familiar question

5 Average weekly earnings in Australia are reported twice yearly, in May and November. The following tables give this data for the years 2007–2016.

	Average weekly earnings (adult) ($)
May 2007	1124.10
Nov 2007	1151.00
May 2008	1171.50
Nov 2008	1210.80
May 2009	1234.90
Nov 2009	1276.70
May 2010	1300.30
Nov 2010	1328.50
May 2011	1357.90
Nov 2011	1390.50
May 2012	1413.10
Nov 2012	1458.00
May 2013	1482.50
Nov 2013	1500.10
May 2014	1515.80
Nov 2014	1542.40
May 2015	1541.50
Nov 2015	1560.50
May 2016	1573.30
Nov 2016	1595.50

a Construct a time series plot of these data.

b Calculate a centred two-point moving average of the average weekly adult earnings.

c Use the actual data values to fit a least squares regression line to the data.

d The actual average weekly earnings in May 2017 was reported as $1605.60. Comment on the accuracy of the prediction obtained for May 2017 using the least squares regression equation from **c**.

e Write a report describing how average weekly earnings have changed over the years 2007–2016.

ISBN 978-1-108-46154-2

Cambridge University Press

Complex unfamiliar questions

6 The following table shows the GDP of Australia since 1980.

Year	GDP	Year	GDP
1980	163.08	1999	411.82
1981	188.59	2000	399.28
1982	187.23	2001	376.93
1983	179.65	2002	424.52
1984	197.33	2003	540.42
1985	174.55	2004	657.43
1986	181.65	2005	734.85
1987	213.31	2006	781.95
1988	271.39	2007	949.15
1989	308.28	2008	1057.11
1990	323.93	2009	997.16
1991	324.59	2010	1249.65
1992	318.14	2011	1504.23
1993	309.64	2012	1561.18
1994	353.22	2013	1509.71
1995	379.78	2014	1449.52
1996	425.59	2015	1229.94
1997	426.66	2016	1261.65
1998	381.10	2017	1390.15

a Construct a time series plot of the data and describe.

b Consider the data for the years 1980–1989.

i Fit a least squares regression model to the data.

ii Use the model to predict for the years 1990, 2000, 2010, 2015, 2017, then compare them to the actual figures. Using percentage error defined as:

$$\frac{actual - predicted}{predicted} \times \frac{100}{1}$$

iii As a measure of accuracy, comment on the differences between estimated and actual values and the reliability of your model.

ISBN 978-1-108-46154-2

Cambridge University Press

c Consider the data for the years 2000–2009.

- **i** Fit a least squares regression model to the data.
- **ii** Use the model to predict for the years 2010, 2015 and 2017, then compare them to the actual figures. Using percentage error defined as:

$$\frac{actual - predicted}{predicted} \times \frac{100}{1}$$

- **iii** As a measure of accuracy, comment on the differences between estimated and actual values and the reliability of your model.

d Consider the data for the years 1980–2009.

- **i** Fit a least squares regression model to the data.
- **ii** Use the model to predict for the years 2010, 2015 and 2017, then compare them to the actual figures. Using percentage error defined as:

$$\frac{actual - predicted}{predicted} \times \frac{100}{1}$$

- **iii** As a measure of accuracy, comment on the differences between estimated and actual values and the reliability of your model.

e Compare your three models and discuss the limitations of each.

Problem-solving and modelling

7 Investigate the fluctuation in exchange rates between the Australian dollar and other currencies. Do these relationships have some part in explaining the variation we see in overseas tourism (both outbound and inbound) from particular countries?

8 Investigate a range of indicators of climate change and how they have changed over time. Possible indicators include temperature, rainfall, greenhouse gas emissions, and global snow and ice coverage, but there are many others. Investigate at least two. Do they tell the same story about climate change?

ISBN 978-1-108-46154-2

Cambridge University Press

6C Topic 3 Growth and decay in sequences

Multiple-choice questions

1 For the recurrence relation $t_{n+1} - t_n = 6$ with $t_1 = 3$, the nth term is:

A $t_n = 6 - 3n$ **B** $t_n = 3 + 3n$ **C** $t_n = 6n - 3$
D $t_n = -3 + 3n$ **E** $t_n = -4 + 7n$

2 A car purchased on 1 June 2017 loses value at a reducing-balance depreciation rate of 20% per year. The original purchase price was \$65 000. The value of the car on 1 June 2022 will be closest to:

A \$32 000 **B** \$22 000 **C** \$26 600
D \$21 300 **E** \$17 000

3 The tenth term of the arithmetic sequence 45, 39, 33, 27, … is:

A 18 **B** 6 **C** 0 **D** –6 **E** –9

4 A population of small marsupials in a particular area is found to be decreasing by 2% each year. When the population was first counted, there were 8000 small marsupials. If t_n is the marsupials at the beginning of the nth year, then:

A $t_n = -0.02 \times t_{n-1}$, where $t_1 = 8000$ **B** $t_n = 0.98 \times t_{n-1}$, where $t_1 = 8000$
C $t_n = 1.02 \times t_{n-1}$, where $t_1 = 8000$ **D** $t_n = t_{n-1} + 0.02$, where $t_1 = 8000$
E $t_n = t_{n-1} - 0.02$, where $t_1 = 8000$

5 The following is a geometric sequence 100 000, 90 000, 81 000, 72 900, ...
The common ratio r is equal to:

A –10 000 **B** 0.1 **C** –9000 **D** 0.9 **E** 1.1

Short-answer questions

Simple familiar questions

1 For an arithmetic sequence with $a = t_1 = 6$ and $d = 5$, determine t_{10}.

2 For an arithmetic sequence with $a = t_1 = 400$ and $d = -10$, determine t_{10}.

3 For a geometric sequence with $a = t_1 = 10000$ and $r = 0.8$, determine t_5.

4 For a geometric sequence with $a = t_1 = 2$ and $r = 3$, determine t_5.

5 In an arithmetic sequence with $t_3 = 99758$ and $t_{11} = 98790$, determine t_{10}.

6 Find an expression for the nth term of the arithmetic sequence 500, 475, 450, …

ISBN 978-1-108-46154-2

Cambridge University Press

7 For an arithmetic sequence with rule $t_n = 8n - 3$, write down the values of t_1, t_2 and t_3.

8 For a geometric sequence with rule $t_n = 2 \times 5^{n-1}$, write down the values of t_1, t_2 and t_3.

9 An arithmetic sequence has $a = t_1 = 45$ and $d = -3$. If $t_n = -39$, find the value of n.

10 A sequence is defined by the recurrence relation $t_n = t_{n-1} + 11.25$, with $t_1 = 8$. Write down the first three terms.

11 A sequence is defined by the recurrence relation $t_n = t_{n-1} - 15$, with $t_1 = 800$. Write down a rule for the nth term of the sequence.

12 A sequence is defined by the recurrence relation $t_n = 6t_{n-1}$, with $t_1 = 2$. Write down a rule for the nth term of the sequence.

13 You have \$200 000 to invest and a bank offers you an interest rate of 2.8% per annum compounded annually for the duration of the investment. How much would your investment be worth in eight years time? Give your answer correct to the nearest dollar.

Complex familiar questions

14 For an arithmetic sequence $t_6 = 13$ and $t_{10} = 5$, find t_{20}.

15 For a geometric sequence $t_6 = 80$ and $t_{10} = 1280$, find t_4.

16 A car was purchased for \$48 000. It depreciates in value at the rate of 8% per year, using a reducing-balance depreciation method.

- **a** Write down a rule for the value of the car after n years.
- **b** Use this rule to find the value of the car after six years.
- **c** What is the total depreciation of the car over five years? Give you answer to the nearest dollar.

17 Philip borrows \$25 500 from a bank and is charged simple interest at the rate of 16% per annum. Let t_n be the value of the loan after n years.

- **a** Write down a rule for the value of the loan after n years.
- **b** How much will Philip need to pay the bank after 3 years?
- **c** How many years does it take the value of the loan to reach \$50 000?

18 The cost of hiring a photocopier for a year involves a flat rate of \$5000 and then a cost of 2 cents per copy.

- **a** Write down the cost of hiring the machine for a year where n copies are made.
- **b** How much will it cost if 120 000 copies are made for the year?
- **c** How many copies can be made in a year if no more than \$10 000 can be spent on the hiring the photocopier?

ISBN 978-1-108-46154-2

Extended-response questions

Complex unfamiliar questions

1 A gardener moves a load of sand used as a foundation for a brick path by emptying barrow loads of sand in line at points 5 metres apart. The first drop-off point is 20 metres from where the sand was originally dumped.

a Write down a mathematical model of the form:

$$L_{n+1} = L_n + d \qquad \text{where} \qquad L_1 = a$$

that can be used to describe this situation. (L_n represents the distance of the nth drop-off point from the load of sand.)

b Write down an expression for the distance of the nth drop-off point from the load of sand.

c Calculate the distance of the 10th drop-off point from the load of sand.

2 Wild deer are causing a problem in a nature reserve. Under normal conditions, the deer population grows at a rate of 22% per year. When counted at the start of the year, there were 1356 deer in the nature reserve.

a Write down a mathematical model of the form:

$$N_{n+1} = rN_n \qquad \text{where} \qquad N_1 = a$$

that can be used to describe the growth of the deer population in the nature reserve under normal conditions. (N_n represents the number of deer in the nature reserve at the end of the nth year.)

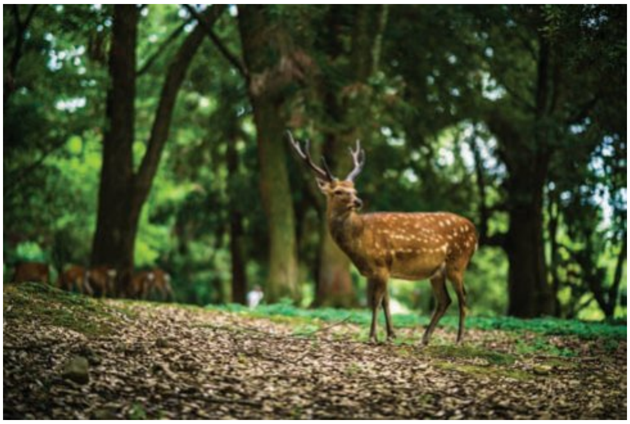
b Use the mathematical model to complete the following table:

End of year	1	2	3	4	5
Number of deer (N_n)					

c Plot a graph of deer numbers against year.

d At the end of which year will the population be in excess of 5000?

e Discuss the reality of the model as the numbers increase.



ISBN 978-1-108-46154-2

Cambridge University Press

Revision

3 When purchased new, a machine for manufacturing car components costs \$2 500 000.

a Let V_n be the value of the machine at the end of the nth year of its working life assuming that it depreciates in value by a constant amount of \$200 000 per year.

i Write down an expression for the value of the machine after n years.

ii Calculate the value of the machine at the end of each of the first five years of its working life.

iii Plot these values on a graph.

b Let U_n be the value of the machine at the end of the nth year of its working life assuming that it depreciates in value by 10% of its value each year.

i Write down an expression for the value of the machine after n years.

ii Calculate the value of the machine at the end of each of the first five years of its working life.

iii Plot these values on a graph.

c From some points of view, the best depreciation method is the one that gives you the lowest value of the machine at the time of disposal. Which depreciation method should you use if you plan to keep the machine for:

i two years?

ii 10 years?

ISBN 978-1-108-46154-2

Cambridge University Press

6D Topic 4 Earth geometry and time zones

Multiple-choice questions

1 City M has latitude 5°N and longitude 5°E. City N has latitude 35°S and longitude 5°E. The shortest distance along the meridian between M and N, in kilometres, is closest to:

A 4431 **B** 4448 **C** 6200 **D** 3336 **E** 3580

2 Location A has latitude 25°N and longitude 5°E. Location B has latitude 25°N and longitude 50°W. The distance along the small circle of the 25°N latitude between A and B, in kilometres, is closest to:

A 1112 **B** 5543 **C** 3833 **D** 6062 **E** 11 120

3 The coordinates of two points X and Y on the Earth's surface are (25°N, 15°E) and (25°S, 45°W). Which statement is most likely to be correct about the time difference?

A Y is 2 hours behind X.
B X is 2 hours behind Y.
C Y is 4 hours ahead of X.
D Y is 5 hours behind X.
E X is 4 hours ahead of Y.

4 Catherine lives in Cooktown 145°E. She wants to watch a basketball game in Los Angeles 118°W starting at 8 p.m. on Friday. What is the closest time in Cooktown to when the game starts in Los Angeles? (Ignore time zones and daylight saving.)

A 2 a.m. on Friday
B 11 p.m. on Friday
C 2 a.m. on Saturday
D 9 a.m. on Saturday
E 2 p.m. on Saturday

Short-answer questions

▶ Simple familiar questions

1 Two locations lie on the same meridian of longitude. One is 37° south of the other. What is the distance between the two locations, correct to the nearest kilometre?

2 Two locations lie on the same parallel of latitude 15°S. One is 140° west of the other. What is the distance around the small circle between the two locations, correct to the nearest kilometre?

3 Vancouver, Canada, has longitude 118°W, while Cooktown has longitude 146°E.

a Calculate the difference in longitude between these two places.
b Calculate the time difference between the two places (Ignore time zones and daylight saving.).
c What is the time in Cooktown when it is 6 p.m. in Vancouver? (Ignore time zones and daylight saving.)

4 How far is London (51.5°N, 0.1°W) from the:

a the equator? **b** the north pole? **c** the south pole?

ISBN 978-1-108-46154-2

Complex familiar questions

5 Two locations A and B are situated on the 140°E meridian. Location A is on the 5°S parallel of latitude and B is north of A. If the distance between the two locations is 3000 km, find the approximation position of location B.

6 Two places are situated on latitude 25°S. If their difference of longitude is 48°, find the distance between the two places measured along the parallel of latitude.

7 Find the length of the parallel of latitude.
a 27°S **b** 51°N

8 A plane flies from M(40°N, 40°E) over the north pole to N(40°N, 140°W).
a How far does the plane fly?
b If the plane flies from M to N around the 40°N parallel of latitude, how far is this?
c What is the difference between the distances?

9 A plane flies from a point with longitude 5°E on the equator and flies around the equator in a westerly direction for 3500 km. Give the approximate location of the plane.

Extended-response questions

Complex unfamiliar questions

1 Complete the following questions relating to the distance between meridians.
a What is the distance between the 145°E and 150°E meridians at the equator?
b What is the distance between the 145°E and 150°E meridians at the 30°S parallel?
c At what latitude is the distance between 145°E and 150°E meridians 500 km?

2 A plane leaves A(5°N, 5°E).
a If the plane flies south along the 5°E meridian until it reaches the 10°S parallel of latitude and then flies west along the 10°S parallel of latitude until it reaches the 50ºW meridian, what is the total distance flown? (Final location B(10°S, 50°W)).
b If the plane flies west along the 5°N parallel of latitude until it reaches the 50°W meridian and then flies south along the 50°W meridian until it reaches the 10°S meridian parallel of latitude, what is the total distance flown? (Final location B (10°S, 50°W)).
c Find the difference of the two total distances.
d Find the time difference between points A and B. (Ignore time zones and summertime.)

ISBN 978-1-108-46154-2

Cambridge University Press

6E List of Unit 3 assessment and examination practice online items

These assessment practice items can be found in the interactive textbook and in the teacher resources of the online teaching suite.

Interactive Textbook

For student and teacher access:

1 IA1: A practice PSMT from Unit 3 Topics 1–3.

2 IA2: A practice internal exam on Unit 3.

Online Teaching Suite

For teacher access:

1 IA1: A PSMT from Unit 3 Topics 1–3.

2 IA2: An internal exam on Unit 3.

Assessment items for Unit 4, and for Units 3 and 4 together, are listed at the end of Chapter 13.

General Mathematics Units 3 & 4
Cambridge Senior Mathematics for Queensland
ISBN 978-1-108-46154-2

Cambridge University Press

Chapter 7

7

Compound interest loans and investments

UNIT 4 INVESTING AND NETWORKING

Topic 1 Loans, investments and annuities

- How do we construct a recurrence relation model for simple and compound interest?
- How do we find the compounding interest rate from a nominal interest rate?
- How do we see the difference in value of loans or investments having different compounding periods?
- How do we find the effective annual rate of a loan or investment?
- How do we find the present and future values of a loan or investment?

ISBN 978-1-108-46154-2

Cambridge University Press

7A Sequences and recurrence relations

Recurrence relations were introduced in Chapter 4 and will be used extensively to model compound interest loans and investments in this chapter. They will also be used to model reducing-balance loans and annuities in later chapters, so it is important to review the definition of recurrence relations in Section 4A before studying this chapter.

Recurrence relations

A **recurrence relation** is a mathematical rule that we can use to generate a sequence. It has two parts:

1. a *starting point*: the value of one of the terms in the sequence
2. a *rule* that can be used to generate successive terms in the sequence.

For example, in words, a recursion rule that can be used to generate the sequence:

10, 15, 20, …

can be written as follows:

1. Start with 10.
2. To obtain the next term, add 5 to the current term and repeat the process.

A more compact way of communicating this information is to translate this rule into symbolic form. We do this by defining a subscripted variable. Here we will use the variable A_n, but the A can be replaced by any letter of the alphabet.

Let A_n be the term in the sequence *after n* applications of the rule.

Using this definition, we now proceed to translate our rule written in words into a mathematical rule.

Starting value ($n = 0$)	Rule for generating the next term	Recurrence relation (two parts: starting value plus a rule)
$A_0 = 10$	$A_{n+1} = A_n + 5$ next term = current term + 5	$A_0 = 10$, starting value; $A_{n+1} = A_n + 5$ rule

Note: Because of the way we defined A_n, the starting value of n is 0. At the start there have been no applications of the rule. This is the most appropriate starting point for financial modelling.

The key step in using a recurrence relation to generate the terms of a sequence is to be able to translate the mathematical recursion rule into words.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 1 Writing a sequence of numbers from a recurrence relation

Write down the first five terms of the sequence defined by the recurrence relation:

$A_0 = 9, A_{n+1} = A_n - 4$

Solution

1 Write down the starting value A_0. $A_0 = 9$

2 Use the rule to find the next term, A_1.
$A_1 = A_0 - 4$
$= 9 - 4$
$= 5$

3 Use the rule find three more terms.
$A_2 = A_1 - 4 = 5 - 4 = 1$
$A_3 = A_2 - 4 = 1 - 4 = -3$
$A_4 = A_3 - 4 = -3 - 4 = -7$

4 Write your answer. The sequence is 9, 5, 1, −3, −7, . . .

Example 2 Constructing a recurrence relation

A sequence has starting value 3 and rule 'triple and then add six'.

Construct a recurrence relation that defines this sequence.

Solution

1 Write down the starting value A_0. $A_0 = 3$

2 The rule must be translated into symbolic form: $A_{n+1} = 3 \times A_n + 6$
- triple means the term A_n is multiplied by three
- then add six.

3 Write your answer. The recurrence relation is $A_0 = 3, A_{n+1} = 3 \times A_n + 6$

ISBN 978-1-108-46154-2

Example 3 **Using a calculator to generate sequences from a recurrence relation**

A sequence is generated by the recurrence relation $A_0 = 300$, $A_{n+1} = 0.5\,A_n - 9$.

Use a calculator to generate this sequence and determine how many terms of the sequence are positive.

Solution

1 Press **clear** (Casio) or **clear** (TI) to create a blank computation screen.

2 Type **300** and then press = (Casio) or enter (TI).

300
300

3 Next, type $\times$ **0.5** – **9** and then press = (Casio) or **enter** (TI) to generate the terms of the sequence. Stop when the first negative term is generated.

Pressing ' = ' or enter once

Ans × 0.5 – 9
141

Pressing ' = ' or enter again

Ans × 0.5 – 9
61.5

4 Count how many positive terms there are in the sequence.

The sequence is
300, 141, 61.5, 21.75, 1.875, −8.0625, ...

5 Write your answer.

The first five terms of this sequence are positive.

Exercise 7A

Generating sequences from a rule

SF

1 Use the following starting values and rules to generate the first five terms of the following sequences recursively by hand.

a Starting value: 2 rule: add 6

b Starting value: 5 rule: subtract 3

c Starting value: 1 rule: multiply by 4

d Starting value: 10 rule: divide by 2

e Starting value: 6 rule: multiply by 2 then add 2

f Starting value: 12 rule: multiply by 0.5 then add 3

ISBN 978-1-108-46154-2

Cambridge University Press

2 Use the following starting values and rules to generate the first five terms of the following sequences recursively using a calculator.

a Starting value: 4
rule: add 2

b Starting value: 24
rule: subtract 4

c Starting value: 2
rule: multiply by 3

d Starting value: 50
rule: divide by 5

e Starting value: 5
rule: multiply by 2 then add 3

f Starting value: 18
rule: multiply by 0.8 then add 2

Generating sequences from a recurrence relation

Example 1

3 Without using your calculator, write down the first five terms of the sequences generated by each of the recurrence relations below.

a $W_0 = 2, W_{n+1} = W_n + 3$

b $D_0 = 50, D_{n+1} = D_n - 5$

c $M_0 = 1, M_{n+1} = 3M_n$

d $L_0 = 3, L_{n+1} = -2L_n$

e $K_0 = 5, K_{n+1} = 2K_n - 1$

f $F_0 = 2, F_{n+1} = 2F_n + 3$

g $S_0 = -2, S_{n+1} = 3S_n + 5$

h $V_0 = -10, V_{n+1} = -3V_n + 5$

Constructing recurrence relations

4 Using your calculator, write down the first five terms of the sequence generated by each of the recurrence relations below.

a $A_0 = 12, A_{n+1} = 6A_n - 15$

b $Y_0 = 20, Y_{n+1} = 3Y_n + 25$

c $V_0 = 2, V_{n+1} = 4V_n + 3$

d $H_0 = 64, H_{n+1} = 0.25H_n - 1$

e $G_0 = 48\,000, G_{n+1} = G_n - 3000$

f $C_0 = 25\,000, C_{n+1} = 0.9C_n - 550$

Example 2

5 Write down a recurrence relation that defines each sequence with the starting value and rule given in Question 2 above.

Exploring sequences and recurrence relations using a calculator

Example 3

6 How many terms of the sequence formed from the recurrence relation $F_0 = 150, F_{n+1} = 0.6F_n - 12$ are positive?

7 How many terms of the sequence formed from the recurrence relation $Y_0 = -45, Y_{n+1} = 0.8Y_n + 2$ are negative?

CF

ISBN 978-1-108-46154-2

7B Modelling simple and compound interest situations with recurrence relations

Mathematical modelling

Mathematical **modelling** is the process of describing or explaining a real-life situation using mathematical terms and symbols. The mathematical *model* that is created can be used to help us understand the situation clearly, to explore and make predictions about that situation.

Recurrence relations can be used as a mathematical model for many different situations and they will be particularly useful to explain or understand financial situations.

Modelling simple interest situations with recurrence relations

Simple interest was introduced in Chapter 4 and will be extended here. Review section 4D, if necessary, to remind yourself of its basic concepts and skills.

A *recurrence relation* model can be used to represent simple interest loans and investments.

Recurrence relation model for simple interest loans and investments

Let A_n be the value of the loan or investment after n years.

Let x be the annual percentage interest rate for the loan or investment.

Let I be the amount of interest that is charged or earned each year.

The recurrence relation for the value of the loan or investment after n years is:

$$A_0 = \text{principal}, \qquad A_{n+1} = A_n + I$$

where $I = \dfrac{x}{100} \times A_0$.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 4 Constructing a recurrence relation model for simple interest

Cheryl deposits \$5000 in a bank account that pays simple interest at a rate of 4.8% each year.

Construct a recurrence relation model for this financial situation.

Solution

1 Define the symbol A_n in the model.

Let A_n be the amount of money in Cheryl's account after n years.

2 Write down the principal amount of the investment, A_0.

$A_0 = 5000$

3 Use the annual percentage rate of interest to calculate the amount of interest added each year, I.

$x = 4.8\%$ per year

$$I = \frac{x}{100} \times A_0$$

$$I = \frac{4.8}{100} \times 5000$$

$$= 240$$

4 Write the recurrence relation.

$A_0 = 5000, \ A_{n+1} = A_n + 240$

The recurrence relation allows us to explore more about this simple interest investment.

Example 5 Using a recurrence relation model to analyse a simple interest investment

Cheryl's investment from example 4 is modelled by the recurrence relation:

$$A_0 = 5000, \ A_{n+1} = A_n + 240$$

a Use this model to determine the value of Cheryl's investment after three years.

b After how many years will Cheryl's investment first exceed \$6000?

Solution

a 1 Write down the recurrence relation.

$A_0 = 5000, \ A_{n+1} = A_n + 240$

2 On a clear calculator screen, type **5000** and press = (Casio) or enter (TI).

5000	
	5000

3 Type +**240** and press = (Casio), or **enter** (TI), three times to obtain the value of Cheryl's investment after three years.

Ans + 240	
	5240
Ans + 240	
	5480
Ans + 240	
	5720

ISBN 978-1-108-46154-2

Cambridge University Press

4 Write your answer.

After three years, Cheryl's investment has a value of $5720.

b 1 Continue pressing = (Casio) or **enter** (TI), counting how many times you press before the value is greater than 6000.

Ans + 240
6200

2 Write your answer.

It takes five years for Cheryl's investment to exceed $6000. It will have value $6200.

Decimal interest rates

The interest rate for a compound interest investment or loan is usually given as a percentage annual rate of interest. When we perform calculations using this percentage rate of interest, it must be converted to a decimal rate of interest by dividing by 100.

Converting percentage interest rates to decimal interest rates

Let i be the decimal interest rate for the percentage interest rate $x\%$ per annum.

$$i = \frac{x}{100}$$

Modelling compound interest situations with recurrence relations

Compound interest was introduced in section 4G as the more usual form of interest, where any interest that is earned after one time period is added to the principal and then contributes to the earning of interest in the next time period.

Consider an investment of $5000 that will be paid compound interest at the rate of 8% every year, or *per annum*.

Every year the balance of the investment will increase by 8% of the previous year's balance.

So: balance next year = balance this year + interest earned
= balance this year + 8% of the balance this year
= 100% of the balance this year + 8% of the balance this year
= 108% of the balance this year
= 1.08 × balance this year

Let A_n be the balance of the investment after n years.

The starting value of the recurrence relation is the principal value of the investment, $A_0 = 5000$.

In recurrence relation symbols, the rule is:

$$A_{n+1} = 1.08 \times A_n$$

We now have a recurrence relation that can be used to model the balance of a compound interest loan or investment.

ISBN 978-1-108-46154-2

Cambridge University Press

A recurrence relation model for a compound interest loan or investment

Let A_n be the balance of a compound interest loan or investment after n years.

Let i be the annual decimal rate of interest.

A recurrence relation model for the balance of a compound interest investment or loan is:

$$A_0 = \text{principal of loan or investment}, \quad A_{n+1} = r \times A_n$$

where $r = 1 + i$

The total interest earned or charged after n years $= A_n - A_0$

The interest earned or charged after the nth year $= A_n - A_{n-1}$

Example 6 Constructing a recurrence relation model for compound interest

Darren borrows \$4000 from a bank. The bank will charge him interest at the annual rate of 9.8%.

Construct a recurrence relation model for the value of Darren's loan after n years.

Solution

1 Define the variable of the recurrence relation.

Let A_n be the value of the loan after n years.

2 The amount borrowed initially, the principal, is the starting value, A_0.

$A_0 = 4000$

3 Calculate the value of r, using the decimal rate of interest, i.

$x = 9.8$

$i = \frac{9.8}{100} = 0.098$

$r = 1 + 0.098$

$r = 1.098$

4 Write your answer.

The recurrence relation is

$A_0 = 4000, \quad A_{n+1} = 1.098 \times A_n$

ISBN 978-1-108-46154-2

Cambridge University Press

Example 7 Using recurrence relation models for compound interest loans and investments

The following recurrence relation can be used to model a compound interest loan of \$2000 charged interest at the percentage rate of 7.5% per annum.

$$A_0 = 2000, \qquad A_{n+1} = 1.075 \times A_n$$

In the recurrence relation, A_n is the balance of the loan after n years.

a Use the recurrence relation to find the balance of the loan after one, two and three years.

b How much interest in total has been charged after three years?

c How much interest has been charged after the third year?

d Determine when the balance of the loan will first exceed \$2500.

Solution

a **1** Write down the principal of the loan, A_0.

$A_0 = 2000$

2 Use calculator recursion to apply the recurrence relation rule and calculate A_1, A_2 and A_3.

Note: The value after three years must be rounded to the nearest cent.

2000	2000
ans × 1.075	2150
ans × 1.075	2311.25
ans × 1.075	2484.59375

3 Write your answer.

After one year, the balance is \$2150.00
After two years, the balance is \$2311.25
After three years, the balance is \$2484.59

b **1** Subtract the principal from the balance to calculate the total interest charged.

Interest after three years
$= \$2484.59 - \$2000$
$= \$484.59$

2 Write your answer.

After three years, a total of \$484.59 in interest has been charged.

ISBN 978-1-108-46154-2

Cambridge University Press

c **1** Subtract the balance after two years from the balance after three years.

Interest after third year

$= A_3 - A_2$

$= \$2484.59 - \$2311.25$

$= \$173.34$

2 Write your answer.

After the third year, $173.34 in interest has been charged.

d Use calculator recursion to count how many years are required to reach the balance of $2500.

1 Press = (Casio), or **enter** (TI), repeatedly, counting the number of times before the balance first exceeds $2500.

2000	2000
ans × 1.075	2150
ans × 1.075	2311.25
ans × 1.075	2484.59375
ans × 1.075	2670.938281

2 Write your answer.

After 4 years, the balance of the loan will first exceed $2500.

▶ Nominal and compounding interest rates

Compound interest rates are usually quoted as an annual rate, or an interest rate per annum. This rate is called the **nominal interest rate** for the investment or loan. Sometimes an annual rate might be quoted, but the interest can be calculated and paid according to a different time period, such as monthly. The time period after which compound interest is calculated and paid is called the **compounding period**.

For example, a compound interest investment may earn interest at the rate of 4.2% per annum, but if the interest is calculated and added to the balance of the investment after every month, then the interest is said to compound monthly.

ISBN 978-1-108-46154-2

Annual (nominal) interest rates can be converted to interest rates for other compounding periods using simple arithmetic.

It must be assumed that there are:

- 12 equal months in every year (even though some months have different numbers of days)
- 4 quarters in every year (a quarter is equal to 3 months)
- 26 fortnights in a year (even though there are slightly more than this)
- 52 weeks in a year (even though there are slightly more than this)
- 365 days in a year (ignore the existence of leap years).

Converting between annual percentage rates of interest and decimal rates of interest per compounding period

Let x be the nominal annual percentage rate of interest.
Let k be the number of compounds per year.
Let i be the decimal interest rate per compounding period for the loan or investment.

$$i = \frac{x}{k \times 100}$$

$$x = k \times 100 \times i$$

Example 8 Converting annual interest rates to decimal interest rates per compounding period

An investment account will pay interest at the rate of 6.24% per annum.
Convert this to a decimal interest rate, i, if the interest compounds are:

a monthly **b** fortnightly **c** quarterly

Solution

a Divide the interest rate by 12 × 100.

$$i = \frac{6.24}{12 \times 100} = 0.0052$$

b Divide the interest rate by 26 × 100.

$$i = \frac{6.24}{26 \times 100} = 0.0024$$

c Divide the interest rate by 4 × 100.

$$i = \frac{6.24}{4 \times 100} = 0.0156$$

ISBN 978-1-108-46154-2

Cambridge University Press

Example 9 **Converting decimal interest rates to annual interest rates per compounding period**

Convert the following decimal interest rates to annual rates.

a 0.0024 (monthly) **b** 0.015 (quarterly) **c** 0.00312 (weekly)

Solution

a Multiply the interest rate by 12×100.

$$\text{Annual interest rate} = 0.0024 \times 12 \times 100 = 2.88\% \text{ per annum}$$

b Multiply the interest rate by 4×100.

$$\text{Annual interest rate} = 0.015 \times 4 \times 100 = 6\% \text{ per annum}$$

c Multiply the interest rate by 52×100.

$$\text{Annual interest rate} = 0.00312 \times 52 \times 100 = 16.224\% \text{ per annum}$$

The recurrence relation model for compound interest with compounding periods other than one year can now be written.

A recurrence relation model for a compound interest loan or investment

Let A_n be the balance of a compound interest loan or investment after n compounding periods.

Let i be the decimal rate of interest, per compounding period.

A recurrence relation model for the balance of a compound interest investment or loan is:

$$A_0 = \text{principal of loan or investment}, \quad A_{n+1} = r \times A_n$$

where $r = 1 + i$

The total interest earned or charged after n compounding periods $= A_n - A_0$

The interest earned or charged after the nth compounding period $= A_n - A_{n-1}$

ISBN 978-1-108-46154-2

Cambridge University Press

Example 10 **Constructing recurrence relation models for compound interest loans and investments**

Diego will invest \$7500 and will earn compound interest at the nominal rate of 9.6% per annum.

Let A_n be the balance of the investment after n compounding periods.

Construct a recurrence relation to model the balance of Diego's investment if interest is compounded:

a yearly **b** quarterly **c** monthly

Solution

a **1** Calculate the decimal rate of interest per year.

$i = \frac{9.6}{1 \times 100} = 0.096$ (one compound per year)

2 Calculate the value of r.

$r = 1 + 0.096$
$= 1.096$

3 Write the recurrence relation.

$A_0 = 7500,\ A_{n+1} = 1.096 \times A_n$

b **1** Calculate the decimal rate of interest per quarter.

$i = \frac{9.6}{4 \times 100} = 0.024$ (four compounds per year)

2 Calculate the value of r.

$r = 1 + 0.024$
$= 1.024$

3 Write the recurrence relation.

$A_0 = 7500,\ A_{n+1} = 1.024 \times A_n$

c **1** Calculate the decimal rate of interest per month.

$i = \frac{9.6}{12 \times 100} = 0.008$ (twelve compounds per year)

2 Calculate the value of r.

$r = 1 + 0.008$
$= 1.008$

3 Write the recurrence relation.

$A_0 = 7500,\ A_{n+1} = 1.008 \times A_n$

ISBN 978-1-108-46154-2

Exercise 7B

Constructing recurrence relation models for simple interest

Example 4

1 Construct a recurrence relation model for the following simple interest investments, where A_n is the value of the investment after n years.

a $2000 earning interest at the annual rate of 2.5%
b $6000 earning interest at the annual rate of 4.2%
c $25 000 earning interest at the annual rate of 6.4%

Converting between annual percentage rates of interest and decimal rates of interest

Example 8

2 Convert the following annual percentage rates of interest to decimal rates of interest for the given compounding periods.

a 7.2% per annum, compounding monthly
b 11.16% per annum, compounding monthly
c 8.06% per annum, compounding fortnightly
d 13.52% per annum, compounding fortnightly
e 7.6% per annum, compounding quarterly
f 10.44% per annum, compounding quarterly

Example 9

3 Convert the following decimal rates of interest to annual percentage rates of interest for the given compounding periods.

a 0.0375 (compounding yearly)
b 0.064 (compounding yearly)
c 0.008 (compounding monthly)
d 0.0094 (compounding monthly)
e 0.0153 (compounding quarterly)
f 0.0042 (compounding fortnightly)

Constructing recurrence relation models for compound interest

Example 10

4 Construct a recurrence relation model for the following compound interest loans or investments, where A_n is the value of the loan or investment after n years.

a $5000 borrowed and charged interest at the annual rate of 8.4%
b $8500 invested and earning interest at the annual rate of 4.2%
c $26 000 borrowed and charged interest at the annual rate of 12.6%

SF

ISBN 978-1-108-46154-2

Cambridge University Press

Example 10

CF

5 Construct a recurrence relation model for the following compound interest loans or investments, where A_n is the value of the loan or investment after n compounding periods.

a $2000 borrowed and charged interest at the rate of 8.4% per annum, compounding monthly.

b $24 000 invested and earning interest at the rate of 7.68% per annum, compounding monthly.

c $16 000 borrowed and charged interest at the rate of 6.72% per annum, compounding quarterly.

d $2800 invested and earning interest at the rate of 4.96% per annum, compounding quarterly.

e $34 000 invested and earning interest at the rate of 9.36% per annum, compounding weekly.

f $18 000 borrowed and charged interest at the rate of 12.48% per annum, compounding weekly.

Problem-solving and modelling

Example 5

SF

6 The following recurrence relation can be used to model a simple interest investment of $2000 earning interest at the rate of 3.8% per annum.

$$A_0 = 2000, A_{n+1} = A_n + 76$$

In the recurrence relation, A_n is the value of the investment after n years.

a Apply the recurrence relation to find the value of the investment after 1, 2 and 3 years.

b Use your calculator to determine how many years it takes for the value of the investment to first be more than $3000.

7 The following recurrence relation can be used to model a simple interest loan of $7000 being charged interest at the rate of 7.4% per annum.

$$A_0 = 7000, A_{n+1} = A_n + 518$$

In the recurrence relation, A_n is the value of the loan after n years.

a Apply the recurrence relation to find the value of the loan after 1, 2 and 3 years.

b Use your calculator to determine how many years it takes for the value of the loan to first be more than $10 000.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 7,10

CF

8 An investment of \$6000 earns compounding interest at the rate of 5.76% per annum, compounding monthly.

Let the balance of the investment after n months be A_n.

A recurrence relation that can be used to model the balance of the investment is shown below.

$$A_0 = 6000, A_{n+1} = 1.0048 \times A_n$$

a Apply the recurrence relation to find the balance of the investment after 1, 2 and 3 months.

b How much interest is earned in total after three months?

c How much interest is earned in the second month?

d How many months will it take for the total interest earned to exceed \$200?

9 A loan of \$8400 is charged compounding interest at the rate of 12.6% per annum, compounding monthly.

Let the balance of the loan after n months be A_n.

A recurrence relation that can be used to model the balance of the loan is:

$$A_0 = 8400, A_{n+1} = 1.0105 \times A_n$$

a Apply the recurrence relation to find the balance of the loan after 1, 2 and 3 months.

b How much interest is charged after the first month?

c How much interest is charged in the second month?

d How much interest is charged in total after three months?

e How many months will it take for the total interest charged to exceed \$1000?

10 Wayne has invested \$7600 and will earn compound interest at the rate of 6% per annum, compounding monthly. Let the balance of Wayne's investment be A_n after n months.

a What is the monthly interest rate for Wayne's investment?

b Construct a recurrence relation model for the balance of Wayne's investment.

c What is the balance of Wayne's investment after six months?

CU

11 Jessica has borrowed \$3500 and will be charged compound interest at the rate of 8% per annum, compounding quarterly. Let the balance of Jessica's loan be A_n after n months.

a What is the quarterly interest rate for Jessica's loan?

b Construct a recurrence relation model for the balance of Jessica's loan.

c If Jessica fully repays her loan after one year, how much money will she need to repay?

d How much extra interest will Jessica be charged if she waits another year to repay the loan?

ISBN 978-1-108-46154-2

Cambridge University Press

7C Investigating compound interest loans and investments

Graphs of compound interest

Consider an investment of \$5000 earning interest at the rate of 10% per annum.

If this was a simple interest investment, the recurrence relation model would be:

$$A_0 = 5000, A_{n+1} = A_n + 500.$$

If this was a compound interest investment, with interest compounding annually, the recurrence relation model would be:

$$A_0 = 5000, A_{n+1} = 1.1 \times A_n$$

The recurrence relations can be used to calculate the balance of each type of investment after each year.

The table on the right shows the rounded balances (where necessary) for both simple interest and compound interest, after each year for a period of ten years.

Both investments grow in value over time. After one year, the balance is the same, but after each subsequent year, the compound interest investment balance is higher than that of the simple interest investment.

Year (n)	Balance (simple interest)	Balance (compound interest)
0	5000.00	5000.00
1	5500.00	5500.00
2	6000.00	6050.00
3	6500.00	6655.00
4	7000.00	7320.50
5	7500.00	8052.55
6	8000.00	8857.81
7	8500.00	9743.59
8	9000.00	10717.94
9	9500.00	11789.74
10	10000.00	12968.71

A graph can make the difference in the rate of growth of simple and compound interest investments much clearer to see.

The graph on the following page shows the year number, n, on the horizontal axis and the balance of the investment on the vertical axis.

Crosses are used to represent the balances of the simple interest investment and dots are used to represent the balances of the compound interest investment.

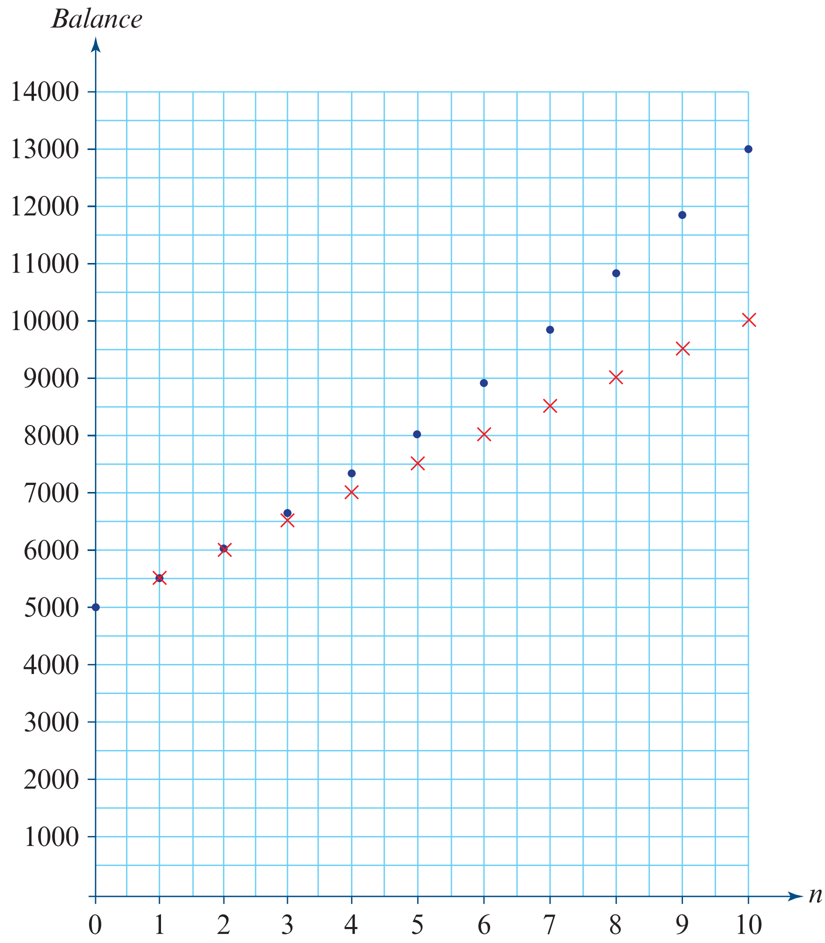
Notice that there is very little difference in the balance of both types of investments after one or two years. As the number of years for the investments increases, the balance of the compound interest investment grows much larger than the simple interest balance. After ten years, the compound interest investment has grown to have almost \$3000 more than the simple interest investment.

ISBN 978-1-108-46154-2

Cambridge University Press

The graph of the simple interest investment is a straight line because it grows by a constant amount each year.

The graph of the compound interest investment curves upwards because it grows by an amount that increases each year.



Desmos activity 7C: To investigate and compare the compound interest growth rates for different loans and investments

ISBN 978-1-108-46154-2

Cambridge University Press

The effect of the compounding period on total interest

Changing the compounding period of a compound interest loan or investment has an effect on the total interest that is earned or charged.

Investigating the effect of compounding period using recurrence relations

You can use recurrence relations to investigate this change using the investigation below.

Investigation 7C-1: Investigating the effect of compounding period using a recurrence relation

You can use a spreadsheet to investigate this change using the investigation below. There are instructions for setting up this spreadsheet yourself and a link to the completed spreadsheet below.

Spreadsheet

Investigation 7C-2: Investigating the effect of compounding period using a spreadsheet

The effect of changing the compounding period on compound interest

Increasing the number of compounds per year of a reducing balance loan or investment increases the balance of that loan or investment. For investments, from the viewpoint of the investor interest should compound as often as possible so that the interest earned is as large as possible. For loans, From the viewpoint of the borrower interest should compound as infrequently as possible so that the interest charged is as small as possible.

The effect of changing the compounding period on compound interest

- Increasing the number of compounding periods per year will mean more interest is earned or charged over the same period of time.
- For investments, from the viewpoint of the investor interest should compound as often as possible to maximise total interest.
- For loans, From the viewpoint of the borrower interest should compound as infrequently as possible to minimise total interest.

ISBN 978-1-108-46154-2

Example 11 Using a spreadsheet or recurrence relation to analyse compounding periods for loans and investments

a Use the compound interest spreadsheet or a recurrence relation to complete the table below.

Principal: \$4000 Annual interest rate: 4.2%	
Compounds per year	Balance after 1 year
1	
2	
4	
12	
52	

b If this was an investment that was closed after one year, how much extra interest is earned by choosing monthly compounds instead of yearly compounds?

c If this was a loan that was paid out after one year, how much interest is saved by choosing monthly compounds instead of weekly compounds?

Solution

a Use the spreadsheet to complete the table, or use the recurrence relations:

$A_0 = 4000, A_{n+1} = 1.042 \times A_n$
(1 compound)

$A_0 = 4000, A_{n+1} = 1.021 \times A_n$
(2 compounds)

$A_0 = 4000, A_{n+1} = 1.0105 \times A_n$
(4 compounds)

$A_0 = 4000, A_{n+1} = 1.0035 \times A_n$
(12 compounds)

$A_0 = 4000, A_{n+1} = 1.000807692 \times A_n$
(52 compounds)

Principal: \$4000 Annual interest rate: 4.2%	
Compounds per year	Balance after 1 year
1	\$4168.00
2	\$4169.76
4	\$4170.66
12	\$4171.27
52	\$4171.51

b **1** Extra interest = monthly value − annual value

Extra interest
= \$4171.27 − \$4168.00
= \$3.27

2 Write your answer.

Compared to yearly compounds, monthly compounds would earn an extra \$3.27 interest.

c **1** Interest saved = weekly value − monthly value

Interest saved = \$4171.51 − \$4171.27
= \$0.24

2 Write your answer.

Compared to weekly compounds, monthly compounds would save \$0.24 interest.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 7C

Using a spreadsheet or recurrence relation to analyse compounding periods for compound interest loans and investments.

Example 11

1 Use a spreadsheet or recurrence relation to complete the tables below. If you use a recurrence relation, do not fill the last row of the table (52 compounds per year).

a

Principal: \$12 000 Annual interest rate: 5.2%	
Compounds per year	Balance after 1 year
1	
2	
4	
12	
52	

b

Principal: \$25 000 Annual interest rate: 9.4%	
Compounds per year	Balance after 1 year
1	
2	
4	
12	
52	

c

Principal: \$67 500 Annual interest rate: 5.7%	
Compounds per year	Balance after 1 year
1	
2	
4	
12	
52	

d

Principal: \$180 000 Annual interest rate: 3.47%	
Compounds per year	Balance after 1 year
1	
2	
4	
12	
52	

Problem solving and modelling

2 A sum of \$12 800 is invested into an account earning compound interest at the rate of 5.8% per annum.

a If there is a choice, should the investor choose weekly compounds or monthly compounds?

b Use a spreadsheet or recurrence relation to calculate the difference in interest earned during the first year of investment by monthly and weekly compounds.

3 A sum of \$3500 is borrowed from a money lender that charges compound interest at the rate of 14.8% per annum.

a If there is a choice, should the borrower choose quarterly compounds or fortnightly compounds?

b Use a spreadsheet or recurrence relation to calculate the difference in interest charged during the first year of the loan by quarterly and fortnightly compounds.

CF

ISBN 978-1-108-46154-2

Cambridge University Press

7D Effective annual rate of interest

▶ Defining effective annual rate of interest

As a general principle, the more frequently interest is calculated and added to a compound interest investment or loan, the more rapidly the value of the investment or loan increases.

The table below compares the value of a $5000 investment earning interest at the nominal rate of 4.8% per annum with the value of the investment with interest calculated on a quarterly and monthly basis.

Principal of investment: $5000			
Nominal annual interest rate: 4.8%			
	Value of investment for interest earned at the rate of:		
Month	4.8% per annum	1.2% per quarter	0.4% per month
0	5000.00	5000.00	5000.00
1			5020.00
2			5040.08
3		5060.00	5060.24
4			5080.48
5			5100.80
6		5120.72	5121.21
7			5141.69
8			5162.26
9		5182.17	5182.91
10			5203.64
11			5224.45
12	5240.00	5244.35	5245.35
Total interest earned*	240.00	244.35	245.35
Effective annual interest rate	4.80%	4.89%	4.91%

*Note: The total interest earned is the value of the investment at the end of the year less the principal.

ISBN 978-1-108-46154-2

Cambridge University Press

The table shows that the more frequently interest is calculated and added, the greater the value of the investment at the end of the year.

When interest is added monthly, the investment earned $245.35 in interest, which is greater than the $240.00 earned with interest added monthly.

The **effective annual rate of interest** of a loan or investment is the annual interest rate that would generate the same amount of interest with one single compound per year as that generated by the original loan.

For example, the effective annual rate of interest for the investment above with 12 compounds per year (monthly) is the interest rate that would earn $245.25 interest in one annual compound. This can be calculated by writing the interest amount as a percentage of the principal value.

$$\text{effective interest rate for 1 compound per year} = \frac{240}{5000} \times 100\% = 4.8\%$$

$$\text{effective interest rate for 4 compounds per year} = \frac{244.35}{5000} \times 100\% = 4.887\%$$

$$\text{effective interest rate for 12 compounds per year} = \frac{245.35}{5000} \times 100\% = 4.907\%$$

An investment with a nominal interest rate of 4.907% per annum with one yearly compound will earn the same interest ($245.35) as an investment with a nominal interest rate of 4.8% but with monthly compounds.

A rule for effective annual rate of interest

The calculations above required the amount of interest earned or charged to be known before the effective annual rate of interest was calculated. It is possible to use a rule to calculate the effective annual rate of interest for a loan or investment given the nominal interest rate and the number of compounding periods per year.

Effective annual rate of interest

A loan with a particular nominal interest rate and number of compounding periods will generate a certain amount of interest over the course of one year.

The effective annual rate of interest of this loan is the nominal interest rate with one annual compound that generates the same amount of interest over the course of one year.

Let n be the number of compounding periods in one year.

Let i be the annual decimal rate of interest.

Let $i_{effective}$ be the effective annual decimal rate of interest for the loan.

$$i_{effective} = \left(1 + \frac{i}{n}\right)^n - 1$$

Note: The annual decimal rate of interest can be converted to an annual percentage rate of interest by multiplying by 100%.

ISBN 978-1-108-46154-2

Example 12 Comparing loans and investments with effective annual rates of interest

Brooke would like to borrow \$20 000. She is deciding between two loan options:

- option A: 5.95% per annum compounding weekly
- option B: 6% per annum compounding quarterly.

a Calculate the effective annual rate of interest for each investment.

b Which investment option is the best and why?

Solution	Option A	Option B
a 1 Write down the values of n for each loan.	$n = 52$	$n = 4$
2 Calculate the annual decimal rate of interest for each loan.	$i = \frac{5.95}{100} = 0.0595$	$i = \frac{6}{100} = 0.06$
3 Apply the effective interest rate formula.	$i_{effective} = \left(1 + \frac{0.0595}{52}\right)^{52} - 1$ $= 0.06126\ldots$	$i_{effective} = \left(1 + \frac{0.06}{4}\right)^{4} - 1$ $= 0.06136\ldots$
4 Convert the effective rates to annual percentage rates.	Annual percentage effective rate $= 0.06126\ldots \times 100\%$ $= 6.127\%$ $= 6.13\%$ (to two decimal places)	Annual percentage effective rate $= 0.06136\ldots \times 100\%$ $= 6.136\%$ $= 6.14\%$ (to two decimal places)
b Compare the effective interest rates.	Brooke is borrowing money, so the best option is the one with the lowest effective interest rate. She will pay less interest with option A.	

Note: Either the effective decimal interest rate or effective percentage interest rate can be compared.

The amount of interest charged or earned over a particular time period depends on the number of compounds within that time period. In a short period of time, the number of compounds per year has little effect on the total interest charged or earned. Over a long period of time, however, the number of compounds per year can have a significant effect on the total interest charged or earned.

Spreadsheet

Spreadsheet activity 7D: A spreadsheet calculator for effective annual rate of interest

ISBN 978-1-108-46154-2

Cambridge University Press

Example 13 Using a spreadsheet calculator to compare loans or investments using effective interest

Ronnie has $30 000 to invest. She has the choice of two investment accounts:

- Account 1 pays compound interest at 6.2% per annum, compounding monthly.
- Account 2 pays compound interest at the rate of 6.05% per annum, compounding weekly.

Which investment should Ronnie choose?

Spreadsheet

Solution

1 Use the effective annual rate of interest spreadsheet calculator (or formula) to find the effective annual rate of interest for Account 1.

	A	B	C	D	E
1	Effective annual rate of interest calculator				
2	*Enter the nominal percentage interest rate and compounds per year into the boxes. Click 'Calculate' to find the effective annual rate of interest. Click 'Clear' to remove all numbers and start again.*				Clear
3					
4			nominal percentage interest rate	6.2	
5			compounds per year	12	
6			effective annual rate of interest	6.379253169	Calculate

2 Use the effective annual rate of interest spreadsheet calculator (or formula) to find the effective annual rate of interest for Account 2.

	A	B	C	D	E
1	Effective annual rate of interest calculator				
2	*Enter the nominal percentage interest rate and compounds per year into the boxes. Click 'Calculate' to find the effective annual rate of interest. Click 'Clear' to remove all numbers and start again.*				Clear
3					
4			nominal percentage interest rate	6.05	
5			compounds per year	52	
6			effective annual rate of interest	6.233023748	Calculate

3 Ronnie is investing and so should choose the account that has the highest effective annual rate of interest.

Account 1 has the higher effective annual rate of interest (6.379%) compared to Account 2 (6.233%) and so Ronnie will earn more interest with Account 1.

ISBN 978-1-108-46154-2

Exercise 7D

Calculating effective annual rates of interest

SF

1 Calculate the effective annual rate of interest for the following nominal annual interest rates and compounding periods. Round your answer to two decimal places.

- **a** 6.2% per annum compounding monthly
- **b** 8.4% per annum compounding daily
- **c** 4.8% per annum compounding weekly
- **d** 12.5% per annum compounding quarterly
- **e** 7.5% per annum compounding every six months

Problem-solving and modelling

CF

2 Brenda invests \$15 000 in an account earning nominal compound interest of 4.60% per annum, compounding quarterly.

- **a** Explain why Brenda would be better off with *more frequent* compounds per year.
- **b** Calculate the effective annual rate of interest for the current loan with quarterly compounds, correct to two decimal places.
- **c** Calculate the effective annual rate of interest for this investment with monthly compounds, correct to two decimal places.
- **d** Explain how these effective annual rates of interest support your answer to part **a**.

3 Stella borrows \$25 000 from a bank and pays nominal compound interest of 7.94% per annum, compounding fortnightly.

- **a** Explain why Stella would be better off with *less frequent* compounds per year.
- **b** Calculate the effective annual rate of interest for the current loan with fortnightly compounds, correct to two decimal places.
- **c** Calculate the effective annual rate of interest for this loan with monthly compounds, correct to two decimal places.
- **d** Explain how these effective annual rates of interest support your answer to part **a**.

Example 12

4 Luke is considering a loan of \$35 000. His bank has two compound interest rate options:

A: 8.3% per annum, compounding monthly

B: 7.8% per annum, compounding weekly.

- **a** Calculate the effective annual rate of interest for each of the loan options.
- **b** Calculate the amount of interest Luke would pay in the first year for each of the loan options.
- **c** Which loan should Luke choose and why?

ISBN 978-1-108-46154-2

Cambridge University Press

5 Sharon is considering investing $140 000. Her bank has two compound interest investment options:

A: 5.3% per annum, compounding monthly.

B: 5.5% per annum, compounding quarterly.

a Calculate the effective annual rate of interest for each of the loan options.

b Calculate the amount of interest Sharon would earn in the first year for each of the loan options.

c Which investment option should Sharon choose and why?

ISBN 978-1-108-46154-2

7E Solving problems involving compound interest

▶ A rule for the balance of compound interest loans and investments

Consider an investment of \$2000 that earns compound interest at the rate of 5% per annum, compounding yearly. If we let A_n be the balance of this investment after n years, the following recurrence relation can be used to model this investment:

$$A_0 = 2000, \; A_{n+1} = 1.05 \times A_n$$

Using this recurrence relation, we can write out the sequence of terms it generates as follows:

$$\begin{aligned}
A_0 &= 2000 \\
A_1 &= 1.05 \times A_0 \\
A_2 &= 1.05 \times A_1 = 1.05\,(1.05 \times A_0) = 1.05^2 \times A_0 \\
A_3 &= 1.05 \times A_2 = 1.05\,(1.05^2 \times A_0) = 1.05^3 \times A_0 \\
A_4 &= 1.05 \times A_3 = 1.05\,(1.05^3 \times A_0) = 1.05^4 \times A_0
\end{aligned}$$

and so on.

Following this pattern, after n years, the balance of the investment will be $A_n = 1.05^n \times A_0$.

Instead of using recurrence relation symbols in this rule, we can use P to represent the principal amount of the loan (A_0) or investment and A to represent the *future value* of the loan or investment after n compounding periods (A_n).

This rule allows the balance of a compound interest loan or investment after any number of compounding periods to be calculated.

A rule for the future value of compound interest loans and investments

Let A be the future value of a compound interest loan or investment.

Let n the total number of compounding periods.

Let i be the decimal rate of interest per compounding period.

Let P be the principal of the loan or investment.

The future value of the compound interest loan or investment after n compounding periods is

$$A = P \times (1 + i)^n$$

ISBN 978-1-108-46154-2

Cambridge University Press

Example 14 Using the rule for the future value of compound interest loans and investments

Bongile would like to invest $25 000 into an account that will pay her compound interest at the rate of 4.2% per annum, compounding monthly.

What is the balance of Bongile's investment after 10 years?

Solution

1 Write down the value of P and i.

$P = 25\,000$

$i = \dfrac{4.2}{12 \times 100}$

$= 0.0035$

2 Calculate the number of compounding periods in the time of the loan, n.

$n = 10 \times 12$ months

$= 120$ months

3 Apply the rule to find the future value, A.

$A = P \times (1 + i)^n$

$A = 25\,000 \times (1 + 0.0035)^{120}$

$= 38\,021.14816$

4 Write your answer, rounded to the nearest cent.

The balance of Bongile's investment is $38 021.15 after 10 years of investment.

The compound interest rule above can be algebraically rearranged into alternate forms. These forms allow the calculation of the principal amount of the loan or investment and the annual percentage interest rate of the loan or investment.

Rule for the principal and interest rates of compounding interest loans and investments

Let A be the future value of a compound interest loan or investment.

Let n the total number of compounding periods.

Let i be the decimal rate of interest per compounding period.

Let P be the principal of the loan or investment.

The principal of the compound interest loan or investment is:

$$P = \frac{A}{(1 + i)^n}$$

The decimal interest rate per compounding period for the loan or investment is:

$$i = \left(\frac{A}{P}\right)^{\frac{1}{n}} - 1$$

Note: The decimal interest rate per compounding period, i, can be converted to a percentage rate per compounding period by multiplying by 100 and then to an annual percentage rate by multiplying by the number of compounds per year.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 15 Using the rule for the principal of compound interest loans and investments

Lowanna has been offered the opportunity to invest some money. She will earn interest at the annual percentage interest rate of 9% per annum, compounding quarterly. Lowanna would like to have at least $15 000 in her investment after 4 years.

How much should Lowanna invest in order to achieve this goal? Round your answer to the nearest dollar.

Solution

1 Write down the values of n and i

$n = 4 \times 4$ quarters
$= 16$ quarters

$i = \frac{9}{4 \times 100} = 0.0225$

2 The future value is the amount of Lowanna's savings goal.

$A = 15\,000$

3 Apply the rule to find the value of P.

$P = \frac{15\,000}{(1 + 0.0225)^{16}}$

$P = 10\,506.98693$

4 Write your answer, rounding to the nearest dollar.

In order to have a balance of $15 000 after 4 years, Lowanna should invest $10 507 now.

Example 16 Using the rule for the annual percentage interest rate of compound interest loans and investments

Daaruk has $35 000 to invest now and would like this investment to grow to at least $45 000 over a period of six years.

If Daaruk's investment earns interest that compounds monthly, what is the minimum annual percentage interest rate that he would require in order to achieve his savings goal?

Round your answer to one decimal place.

Solution

1 Write down the value of P, n and A.

$P = \$35\,000$
$n = 6 \times 12$ months $= 72$ months
$A = \$45\,000$

2 Apply the rule to find the annual percentage rate of interest.

$$i = \left(\frac{A}{P}\right)^{\frac{1}{n}} - 1$$

$$i = \left(\frac{45\,000}{35\,000}\right)^{\frac{1}{72}} - 1$$

$i = 0.003496\ldots$

ISBN 978-1-108-46154-2

3 Convert i to an annual percentage interest rate.	*Annual percentage interest rate* $= 0.003496 \times 12 \times 100$ $= 4.19589\ldots$
4 Write your answer rounded to one decimal place.	*Daaruk would need an annual percentage interest rate of 4.2% to achieve his savings goal.*

Solving compound interest problems using technology

Compound interest calculations involve five different values:

- Principal
- Annual interest rate
- Number of compounds per year
- Future value after n compounds
- n, the total number of compounding periods.

If any four of these values are known, the fifth can be calculated using technology such as a spreadsheet.

Spreadsheet **Spreadsheet activity 7E:** A spreadsheet compound interest calculator

Using the compound interest calculator spreadsheet

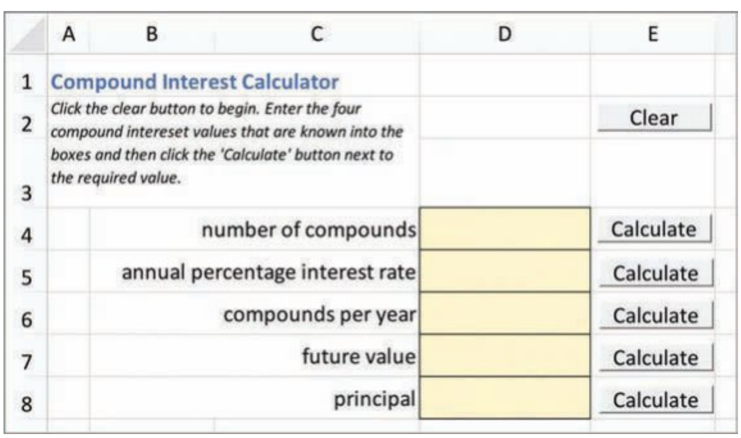
The Compound Interest Calculator spreadsheet is shown here.

Click the 'Clear' button before every calculation.

Enter the four known quantities into the orange box.

Click the 'calculate' button next to the quantity that you want to find. It will appear in the box.

Spreadsheet



ISBN 978-1-108-46154-2

Cambridge University Press

Example 17 Solving compound interest problems using a spreadsheet

The balance of Ahmet's investment account is $15 480.03 after a period of 2 years. His initial investment was $12 000. If compound interest is calculated and added to the account monthly, what is the annual percentage interest rate for Ahmet's investment? Round your answer to one decimal place.

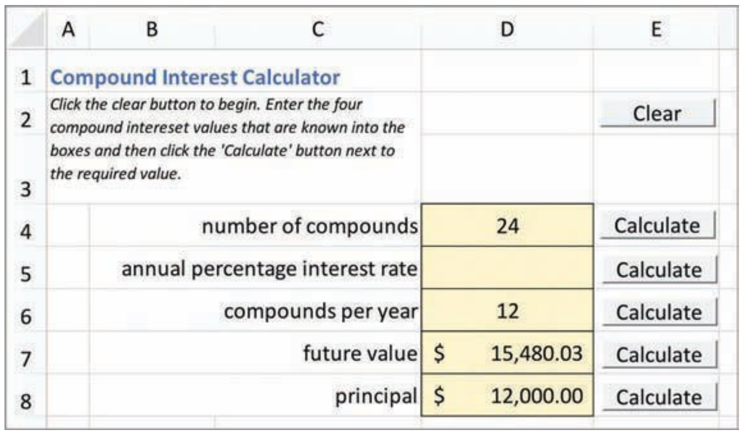
Solution

1 Identify the known quantities.

Principal = $12 000
Balance = $15 480.03
Number of compounds per year = 12 (monthly)
Number of compounds = 2 × 12 months (2 years)
= 24

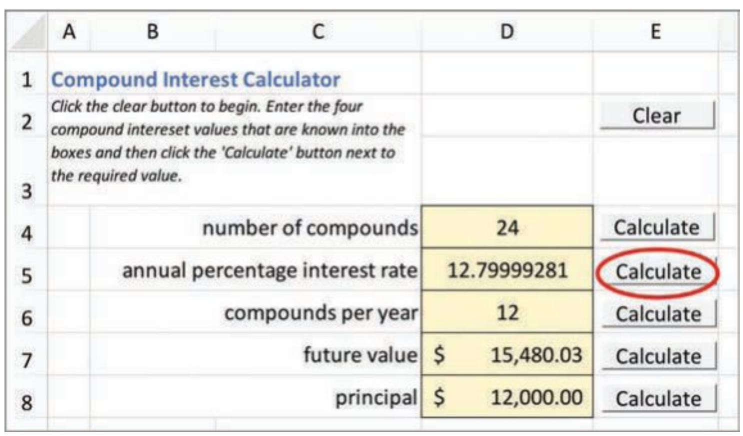
2 Click 'Clear' on the compound interest spreadsheet and enter the known values.

Notes: You do not need to type the dollar sign or thousands comma. Cells for dollar values have currency formatting applied which includes them automatically.



	A	B	C	D	E
1	Compound Interest Calculator				
2	Click the clear button to begin. Enter the four compound intereset values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			number of compounds	24	Calculate
5			annual percentage interest rate		Calculate
6			compounds per year	12	Calculate
7			future value	$ 15,480.03	Calculate
8			principal	$ 12,000.00	Calculate

3 Click the 'Calculate' button next to annual interest rate. The annual percentage interest rate will be calculated and entered into the box for you.



	A	B	C	D	E
1	Compound Interest Calculator				
2	Click the clear button to begin. Enter the four compound intereset values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			number of compounds	24	Calculate
5			annual percentage interest rate	12.79999281	Calculate
6			compounds per year	12	Calculate
7			future value	$ 15,480.03	Calculate
8			principal	$ 12,000.00	Calculate

4 Write your answer, rounding to one decimal place.

The annual percentage interest rate for Ahmet's investment is 12.8%.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 7E

Using the rule to determine the future value of compound interest loans and investments

Example 14

SF

1 Use the compound interest rule to determine the balance of the following loans and investments after the given period of time. Round your answers to the nearest cent.

- **a** \$2500 borrowed for 3 years with compound interest of 3.5% per annum, compounding yearly.
- **b** \$15 000 invested for 4 years with compound interest of 2.8% per annum, compounding yearly.
- **c** \$5800 borrowed for 4 quarters with compound interest of 9.6% per annum, compounding quarterly.
- **d** \$26 000 invested for 60 months with compound interest of 4.6% per annum, compounding monthly.
- **e** \$6400 borrowed for 5 years with compound interest of 8.5% per annum, compounding weekly.
- **f** \$12 500 invested for 6 years with compound interest of 4.7% per annum, compounding monthly.

Using the rule to determine the principal and interest rate of compound interest loans and investments

Example 15

2 Use the rule for the principal of compound interest loans and investments to determine the principal value of the following loans and investments after the given period of time. Round your answers to the nearest cent.

- **a** An investment earning compound interest at the rate of 6.9% per annum, compounding quarterly, with a future value of \$14 692.82 after 8 years.
- **b** A loan charged compound interest at the rate of 12.6% per annum, compounding quarterly, with a future value of \$34 821.06 after 3 years.
- **c** An investment earning compound interest at the rate of 4.2% per annum, compounding monthly, with a future value of \$43 162.90 after 5 years.
- **d** A loan charged compound interest at the rate of 14.5% per annum, compounding monthly, with a future value of \$7944.62 after 18 months.
- **e** An investment earning compound interest at the rate of 3.8% per annum, compounding weekly, with a future value of \$33 446.91 after 2 years.

ISBN 978-1-108-46154-2

Example 16

3 Use the rule for the annual percentage interest rate of compound interest loans and investments to determine the interest rate for the following loans and investments after the given period of time. Round your answers to two decimal places.

a An investment of \$2000.00 earning compound interest that compounds quarterly and with a future value of \$2560.67 after 6 years.

b An investment of \$8500.00 earning compound interest that compounds quarterly and with a future value of \$10 198.86 after 3 years.

c An investment of \$50 000.00 earning compound interest that compounds monthly and with a future value of \$63 828.57 after 4 years.

d An investment of \$15 000 earning compound interest that compounds monthly and with a future value of \$33 059.63 after 15 years.

e An investment of \$45 000 earning compound interest that compounds weekly and with a future value of \$52 153.57 after 3 years.

Problem-solving and modelling with the aid of technology

Example 17

CF

4 Sarah invested \$3500 at 6.75% per annum, compounding annually. After how many years will the value of Sarah's investment first exceed \$5000?

5 Tenile invested \$20 000 that has grown to a balance of \$21 522.15 after 18 months with interest compounding monthly. Determine the percentage annual interest rate for Tenile's investment. Round your answer to two decimal places.

6 How long will it take \$2000 to exceed \$20 000 if it was invested at a compound interest rate of 4.75% per annum, compounding annually?

7 If \$45 000 was invested in a compound interest account earning interest at the rate of 6.8% per annum, compounding quarterly, how many quarters would it take for the balance of the investment to exceed \$100 000?

8 Suppose that an investment of \$1000 has grown to \$1601.03 after 12 years. If this investment earned compound interest at the rate of i% per annum compounding yearly, what is the value of i? Round your answer to two decimal places.

9 Jannie invested \$25 000 in an account earning compound interest at the rate of i% per annum, compounding monthly. Jannie's investment had a balance of \$29 216.11 after 30 months. Calculate the value of i. Round your answer to two decimal places.

ISBN 978-1-108-46154-2

Key ideas and chapter summary

AS

Sequence
A **sequence** is a list of numbers or symbols written in succession. For example: 5, 15, 25, …

Term
Each number or symbol that makes up a sequence is called a **term**.

Recurrence relation
A **relation** that enables the value of the next term in a sequence to be obtained by one or more current terms. Examples include 'to find the next term, add two to the current term' and 'to find the next term, multiply the current term by three and subtract five'.

Mathematical modelling
Mathematical modelling is the use of mathematical terms and symbols to describe or explain real-life situations. **Recurrence relations** can be used as models for many different real-life situations, including financial situations.

Interest
The fee that is added to a loan or the payment received for investing money is called the **interest**.

Principal
The **principal** is the initial amount that has been invested or borrowed.

Simple interest
When a fixed amount of interest is added to a loan or investment at regular time intervals, the interest is called **simple interest**.

Compound interest
When interest is added to a loan or investment and then contributes to earning more interest, the interest is called **compound interest**.

Compounding period
The time period for the calculation of interest is called the compounding period. **Compounding periods** are usually daily, fortnightly, monthly, quarterly, six-monthly or annually.

Decimal rate of interest per compounding period
If x is the annual percentage rate of interest, and if k is the number of compounding periods per year, then the decimal rate of interest per compounding period is

$$i = \frac{x}{k \times 100}$$

 ISBN 978-1-108-46154-2 Cambridge University Press

Review

Recursive model for compound interest

A recurrence relation that can be used to determine the balance of a compound interest loan or investment after n compounding periods.
If the number of compounding periods per year is k, and if the annual percentage interest rate is i, then the **recursive model** for the loan or investment is
A_0 = principal of loan or investment, $A_{n+1} = r \times A_n$
where $r = 1 + i$

Nominal interest rate

The annual percentage interest rate for a loan or investment is called the **nominal interest rate**.

Effective annual rate of interest

The **effective annual rate of interest $i_{effective}$** is used to compare the interest paid on loans or investments with different nominal rates of interest and compounding periods.
If the number of compounding periods per year is k, then the annual decimal rate of interest is

$$i_{effective} = (1 + i)^k - 1$$

The annual percentage rate of interest is

$$i_{effective} = \left((1 + i)^k - 1\right) \times 100\%$$

Future value

The **future value** (A) of a compound interest loan or investment is the balance of that loan or investment after some number of compounding periods.

Rule for the future value of a compound interest loan or investment

If P is the principal of the loan or investment, i is the decimal rate of interest per compounding period, then the future value of the loan after n compounding periods is:

$$A = P \times (1 + i)^n$$

Rule for the principal of a compound interest loan or investment

If A is the future value of a loan or investment and i is the decimal rate of interest per compounding period, then the principal of the loan or investment after n compounding periods is

$$P = \frac{A}{(1 + i)^n}$$

Rule for the decimal interest rate per compounding period of a compound interest loan or investment

If P is the principal of the loan or investment, A is the future value after n compounding periods then the decimal interest rate per compounding period is

$$i = \left(\frac{A}{P}\right)^{\frac{1}{n}} - 1$$

ISBN 978-1-108-46154-2

Review

Skills check

Having completed this chapter, you should be able to:

- construct a recurrence relation model for simple interest
- analyse simple interest using a recurrence relation model
- construct a recurrence relation model for compound interest
- calculate a compounding interest rate from a nominal annual interest rate
- calculate a nominal annual interest rate from a compounding interest rate
- analyse compound interest using a recurrence relation model
- understand and compare the graph of the balance of simple interest and compound interest loans and investments
- understand the effect that the compounding period has on compound interest
- calculate the effective annual rate of interest for a given compound interest loan or investment
- compare compound interest loans or investments using the effective interest rate
- use a rule to calculate the future value, principal or interest rate of a compound interest loan or investment
- solve compound interest problems using technology (spreadsheet calculation tool).

Multiple-choice questions

1 Which one of the following recurrence relations defines the sequence with starting value 4 and the rule 'multiply by 3 and subtract 1'?

A $A_0 = 1, A_{n+1} = 3A_n - 4$
B $A_0 = 4, A_{n+1} = 3A_n - 1$
C $A_0 = 3, A_{n+1} = 4A_n - 1$
D $A_0 = 4, A_{n+1} = A_n - 3$
E $A_0 = 1, A_{n+1} = 4A_n - 3$

2 How many terms of the sequence defined by the recurrence relation $A_0 = 25, A_{n+1} = 2A_n - 30$ are positive?

A 1 **B** 2 **C** 3 **D** 4 **E** 5

3 The annual percentage rate of interest for a compound interest loan is 12.6% per annum, compounding monthly.

The balance of this loan after n months, A_n, can be modelled by the recurrence relation:

$$A_0 = 4000, A_{n+1} = 1.0105 \times A_n.$$

If the loan and interest is fully repaid after 5 months, how much in total will be paid?

A \$4127.33 **B** \$4170.66 **C** \$4214.46 **D** \$4258.71 **E** \$4303.42

ISBN 978-1-108-46154-2

4 A compound interest investment of principal $12 000 will earn interest at the rate of 10.8% per annum, compounding every six months. Which one of the following is a recurrence relation model for A_n, the balance of the investment after n six-month periods?

A $A_0 = 12000,\ A_{n+1} = 1.054 \times A_n$

B $A_0 = 12000,\ A_{n+1} = 1.108 \times A_n$.

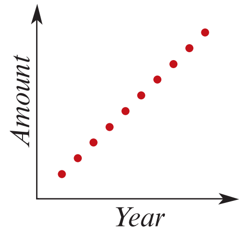
C $A_0 = 12000,\ A_{n+1} = 1.15 \times A_n$.

D $A_0 = 12000,\ A_{n+1} = 1.018 \times A_n$.

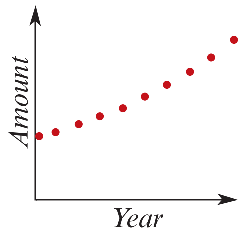
E $A_0 = 12000,\ A_{n+1} = 10.8 \times A_n$.

5 An investment of $50 000 is made at a fixed rate of interest compounding annually over a number of years. Which graph best represents the value of the investment at the end of each year?

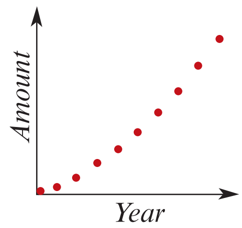
A



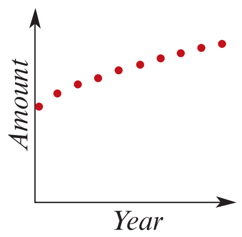
B



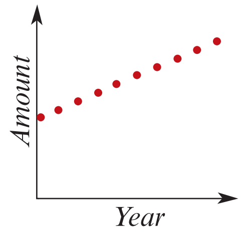
C



D



E



6 Issa has some money to invest and would like to earn as much interest as possible in the first year of the investment. Which one of the following interest rates should he choose?

A 3.1% per annum, compounding weekly

B 3.1% per annum, compounding monthly

C 3.2% per annum, compounding quarterly

D 3.2% per annum, compounding monthly

E 3.2% per annum, compounding annually

Review

7 A compound interest investment earns interest that compounds monthly. The balance of this investment after n months, A_n, can be found using the recurrence relation:

$$A_0 = 15\,000,\ A_{n+1} = 1.0024 \times A_n.$$

The balance of the investment can also be found using the rule:

A $1.0024 \times (15\,000)^n$

B $15\,000 \times (2.88)^n$

C $1.0024 \times (2.88)^n$

D $15\,000 \times (1.24)^n$

E $15\,000 \times (1.0024)^n$

8 A principal of \$2000 is invested and will earn compound interest at the rate of 5.4% per annum, compounding quarterly. The effective annual rate of interest for this investment is closest to:

A 5.2% **B** 5.3% **C** 5.4% **D** 5.5% **E** 5.6%

9 A compound interest investment earns interest at the rate of 3.6% per annum, compounding quarterly. If the balance after 4 years is \$10 964.33, the principal investment amount is closest to:

A \$6200 **B** \$9500 **C** \$10 400 **D** \$10 600 **E** \$10 800

10 A compound interest investment of \$10 000 earns \$590.69 interest over a period of 15 months. If interest compounds monthly, the annual percentage rate of interest is closest to:

A 1.5% **B** 3.8% **C** 4.6% **D** 7.3% **E** 9.8%

Short-answer questions

SF

1 Eli invested \$12 500 into an account that pays compound interest at the rate of 7.8% per annum, compounding monthly.

a Construct a recurrence relation model for the balance of Eli's investment after n months.

b Apply the recurrence relation to determine the balance of Eli's investment after 4 months.

2 The following recurrence relation can be used to model a compound interest investment of \$5800 earning interest at the rate of 6.72% per annum, compounding monthly.

$$A_0 = 5800,\ A_{n+1} = 1.0056 \times A_n.$$

In this recurrence relation, A_n is the balance of the investment after n months.

a Apply the recurrence relation to find the balance of the investment after one, two and three months.

b After how many months will the balance of the investment first exceed \$6000?

ISBN 978-1-108-46154-2

Cambridge University Press

3 The following recurrence relation can be used to model a compound interest loan. The interest is calculated and added to the loan weekly.

$$A_0 = 1600, A_{n+1} = 1.0035 \times A_n.$$

In this recurrence relation, A_n is the balance of the investment after n weeks.

- **a** What is the principal of this loan?
- **b** Calculate the annual percentage rate of interest.
- **c** Apply the recurrence relation to find the balance of the loan after one, two and three weeks.
- **d** After how many weeks will the balance of the loan first exceed $1650?

4 Hansie has borrowed $2200 and will be charged compound interest at the rate of 15.2% per annum, compounding quarterly. Let the balance of Hansie's loan be A_n after n quarters.

- **a** What is the quarterly interest rate for Hansie's loan?
- **b** Construct a recurrence relation that models Hansie's loan.
- **c** If Hansie fully repays her loan with one payment after the first year, how much money will she need to repay?

5 Eva borrows $15 000 and will be charged compound interest at the rate of 4.6% per annum.

- **a** Calculate the balance of the loan after two years.
- **b** What is the total interest that will be charged after two years?

6 Rodney will borrow $1500. He will be charged compound interest at the rate of 11.28% per annum compounding monthly.

- **a** Use the compound interest rule to determine the balance of Rodney's loan after two months.
- **b** What is the total interest that has been charged after two months

7 Ilana uses a credit card to buy a dress that costs $300. Interest on this loan will be compounded monthly. If Ilana does not make any repayments on her credit card, she will need to repay $323.98 after five months.
Use the compound interest rule to determine the annual percentage interest rate for Ilana's credit card. Round your answer to one decimal place.

 ISBN 978-1-108-46154-2 Cambridge University Press

Extended-response questions

CF

1 Meghan has \$5000 to invest.

- Bank A offers investment accounts that pay interest at the rate of 6.3% per annum compounding quarterly.
- Bank B offers investment accounts that pay interest at the rate of 6.1% per annum, compounding monthly.

a Calculate the effective annual rate of interest for each of these investment accounts. Round your answers to two decimal places.

b Which investment account should Meghan choose? Justify your answer by explaining how you compared the two investment options.

c Calculate the extra interest that Meghan will earn in one year by choosing the investment account in **b**, compared to the other account. Write your answer to the nearest dollar.

2 Darius has \$25 000 to invest. He has two investment options:

- Bank A offers to pay 8.2% per annum compounding six-monthly
- Bank B offers to pay 8.1% per annum compounding quarterly

Darius would like his money to remain invested for a period of 18 months.

a Which of the two investment options would earn Darius the most interest? Justify your answer by explaining how you compared the two investment options.

b Calculate the difference between the total interest earned by both investment options. Round your answer to the nearest dollar.

3 Eva invests \$15 000 and will earn compound interest at the annual percentage interest rate of 4.6% per annum.

a Calculate the amount of money Eva has invested after one year if the interest earned compounds:

i quarterly **ii** monthly **iii** weekly

b What financial principle is used to compare the investment conditions in **1a** above?

c Write a paragraph to explain to Eva why weekly compounding interest on her investment will be of most benefit to her.

4 Chi borrowed some money and will be charged compound interest that compounds weekly. The balance of Chi's loan after n weeks can be found from the rule $A_n = 2300 \times 1.0034^n$.

In the following questions, round your answers to the nearest cent.

a What is the principal of Chi's loan?

b Calculate the annual interest rate for Chi's loan.

c Find the:

i balance of Chi's loan after 16 weeks

ii total amount of interest that has been charged after 16 weeks

iii total amount of interest that has been charged after 1 year

ISBN 978-1-108-46154-2

Cambridge University Press

5 A 'payday loan' company offers short term loans. Interest on these loans is charged at the annual percentage interest rate of 12.48%. CU

Lucille needs $2500 to pay for urgent repairs of her car.

a Complete the following table that shows how much Lucille will have to pay back after different periods of time and for quarterly and monthly compounding periods.

Month	Quarterly compounds	Monthly compounds
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

b **i** Construct a graph that shows the values in the table from question **5a** above.

ii Quarterly compounding interest is better for Lucille than monthly compounding interest.

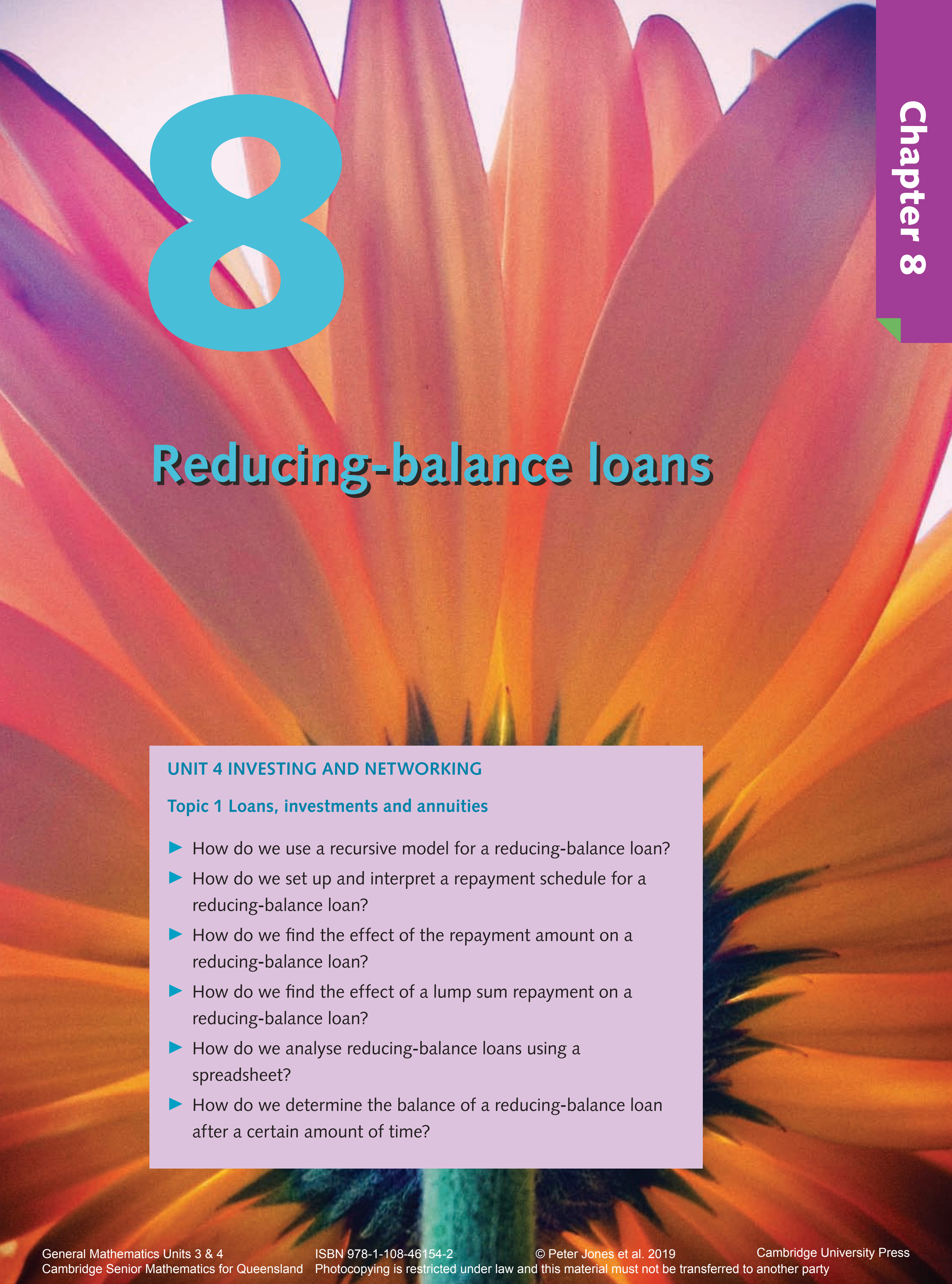
Describe the features of the table and graph you have drawn in question **bi** above that support this statement.

c Lucille will repay the principal and all interest on her loan after one year.

How much extra interest will Lucille pay if she is charged interest that compounds monthly instead of quarterly?

ISBN 978-1-108-46154-2

Cambridge University Press



8

Reducing-balance loans

UNIT 4 INVESTING AND NETWORKING

Topic 1 Loans, investments and annuities

- How do we use a recursive model for a reducing-balance loan?
- How do we set up and interpret a repayment schedule for a reducing-balance loan?
- How do we find the effect of the repayment amount on a reducing-balance loan?
- How do we find the effect of a lump sum repayment on a reducing-balance loan?
- How do we analyse reducing-balance loans using a spreadsheet?
- How do we determine the balance of a reducing-balance loan after a certain amount of time?

ISBN 978-1-108-46154-2

Cambridge University Press

8A A recursive model for reducing-balance loans

Reducing-balance loans

In the compound interest loans from Chapter 7, a principal amount of money was borrowed, interest was calculated and charged at regular compounding time periods and then the money was paid back, along with any interest charged, at the end of the loan.

In practice, it is very unusual for a borrower to wait until the end of the loan to repay the principal and interest to the bank. Instead, loans are usually repaid by making regular repayments that coincide with the compounding time periods. This has the effect of gradually reducing the **balance** of the loan over time, until it is fully repaid.

This kind of loan is called a **reducing-balance loan**. Home loans and other personal loans are examples of reducing-balance loans.

A recursive model for a reducing-balance loan

Consider a reducing-balance loan for \$5000 that is charged interest at the rate of 8% per annum, compounding yearly. Repayments of \$1500 will be made each year.

Let A_n be the balance of the loan after n years.

The starting value of the recurrence relation is the principal value of the loan, $A_0 = 5000$.

Each year, the loan balance increases by the amount of interest charged, that is 8% of the previous balance, and then reduces by the amount of the repayment

So: balance next year = balance this year + interest charged − repayment
= balance this year + 8% of the balance this year − repayment
= 100% of the balance this year + 8% of the balance this year − repayment
= 108% of the balance this year − repayment

In recurrence relation symbols: $A_{n+1} = 1.08 \times A_n - 1500$

We now have a recurrence relation that can be used to model the balance of a reducing-balance loan for differing compounding periods.

A recurrence relation model for a reducing-balance loan

Let A_n be the balance of a reducing-balance loan after n compounding periods.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period.

Let R be the repayment amount per compounding period.

A recurrence relation model for the balance of a reducing-balance loan is:

$$A_0 = \text{principal of loan}, \qquad A_{n+1} = r \times A_n - R$$

where $r = 1 + i$

ISBN 978-1-108-46154-2

Example 1 Modelling a reducing-balance loan with a recurrence relation

Alyssa will borrow \$4800 and will be charged compound interest at the rate of 15% per annum, compounding monthly. She will make monthly repayments of \$300 to repay this loan.

a Construct a recurrence relation model for Alyssa's loan.

b Apply the recurrence relation to determine how much Alyssa will still owe on the loan after two repayments.

Solution

a 1 Write down the values of A_0 and R.

$A_0 = 4800$ (principal of loan)
$R = 300$ (monthly repayment)

2 Calculate the value of the decimal interest rate, i.

15% per annum compounding monthly

$$i = \frac{15}{12 \times 100} = 0.0125$$

3 Calculate the value of r.

$$r = 1 + i = 1 + 0.0125 = 1.0125$$

4 Write your answer.

$A_0 = 4800$, $A_{n+1} = 1.0125 \times A_n - 300$

b 1 Apply the recurrence relation two times to find A_2.

$A_0 = 4800$
$A_1 = 1.0125 \times 4800 - 300 = 4560$
$A_2 = 1.0125 \times 4560 - 300 = 4317$

2 Calculator recursion can also be used to find A_2.

Press **AC** (Casio) or **clear** (TI) to create a blank calculation screen.

| 0

Type **4800** and then press = (Casio) or **enter** (TI)

4800
4800

Next, type × **1.0125 − 300** and then press = (Casio), or enter (TI), twice to find A_2

Pressing ' = ' or enter once for A_1

Ans × 1.0125 − 300
4560

Pressing ' = ' or enter once again for A_2

Ans × 1.0125 − 300
4317

3 Write your answer.

After two repayments, Alyssa will still owe \$4317.00.

ISBN 978-1-108-46154-2

Cambridge University Press

Calculating the total interest charged for reducing-balance loans

When a repayment on a reducing-balance loan is made, the first priority is to pay the interest that was charged after that compounding period. This interest amount is usually smaller than the repayment amount and so any remaining amount of the repayment will pay back some of the principal of the loan. In this way, the principal of the loan; that is, the amount owed after each compounding period, will gradually be reduced in value.

After n compounding periods, the reduction in the principal can be calculated as:

$$\text{Reduction in principal} = A_0 - A_n$$

The total repayment amount after n compounding periods can be calculated as:

$$\text{Total repayment} = n \times \text{repayment per compounding period}$$

The total repayment amount pays both the reduction in principal and the interest charged and so the total interest charged after n repayments can be calculated as:

$$\text{Total interest} = \text{total repayment} - \text{reduction in principal}$$

Total interest charged for reducing-balance loans

Let A_0 be the principal amount of a reducing-balance loan.

Let A_n be the balance of a reducing-balance loan after n compounding periods.

Let R be the repayment amount per compounding period.

Let I be the total interest charged after n compounding periods.

The reduction in principal after n compounding periods $= A_0 - A_n$.

The total repayments made after n compounding periods $= n \times R$.

$$I = \text{total repayments made} - \text{reduction in principal}$$
$$= n \times R - (A_0 - A_n)$$

Another way of writing this rule is

$$I = A_n + n \times R - A_0$$

Example 2 Analysing reducing-balance loans with a recurrence relation

Henry will borrow \$20 000 and will be charged compound interest at the rate of 8.4% per annum, compounding quarterly. He will make quarterly repayments of \$1500 to repay this loan.

a Construct a recurrence relation model for Henry's loan.

b How much interest will Henry pay in the first year of his loan?

ISBN 978-1-108-46154-2

Cambridge University Press

Solution

a **1** Write down the values of A_0, i, and R.

$A_0 = 20\,000$ (loan principal)

$i = \frac{8.4}{4 \times 100}$ (decimal interest rate per quarter)

$= 0.021$

$R = 1500$ (quarterly repayment)

2 Calculate the value of r.

$r = 1 + 0.021$

$= 1.021$

3 Write your answer.

$A_0 = 20\,000$, $A_{n+1} = 1.021 \times A_n - 1500$

b **1** One year has four quarters.

Apply the recurrence relation four times to find A_4. Using calculator recursion:

Press **AC** (Casio) or **clear** (TI) to create a blank calculation screen.

|	0

Type **20000** and then press = (Casio) or **enter** (TI)

20000	20000

Next, type × **1.021 − 1500** and then press = (Casio), or **enter** (TI), four times to find A_4.

Pressing ' = ' or enter once for A_1

Ans × 1.021 − 1500	18920

Round your answer to the nearest cent if necessary.

Pressing ' = ' or enter three more times for A_4

Ans × 1.021 − 1500	15542.00488

2 Write down the value of n, R, A_0 and A_n.

$n = 4$ (quarterly repayments)

$R = 1500$ (repayment amount)

$A_0 = 20\,000$ (principal amount)

$A_n = 15\,542$ (balance after 4 repayments)

3 Calculate the interest paid after one year (four repayments).

Interest paid $= n \times R - (A_0 - A_n)$

$= 4 \times 1500 - (20\,000 - 15\,542)$

$= 6000 - 4458$

$= 1542$

4 Write your answer.

In the first year of his loan, Henry will pay $1542.00 in interest.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 8A

Constructing a recurrence relation model for a reducing-balance loan

Example 1

1 The table below shows the principal amount, annual percentage interest rate, compounding period and repayment amount for six reducing-balance loans.

	Principal	Annual percentage interest rate	Compounding period	Repayment per compounding period
a	\$6500	14%	Yearly	\$1800
b	\$14 000	11.2%	Quarterly	\$2000
c	\$22 000	7.2%	Quarterly	\$1000
d	\$85 000	8.04%	Monthly	\$1800
e	\$150 000	6.48%	Monthly	\$1700
f	\$245 000	4.16%	Fortnightly	\$1200

For each of these loans:

i construct a recurrence relation model

ii apply the recurrence relation model to find how much is still owed on the loan after three compounding periods

SF

Using a recurrence relation model to analyse a reducing-balance loan

Example 2

2 A reducing-balance loan can be modelled by the recurrence relation

$$A_0 = 2500,\ A_{n+1} = 1.08 \times A_n - 600$$

where A_n is the balance of the loan after n repayments have been made.

a What is the value of the repayment each compounding period?

b Use calculator recursion to determine how much is still owed on the loan after three repayments have been made.

c How much interest has been paid after three repayments?

CF

3 A reducing-balance loan can be modelled by the recurrence relation

$$A_0 = 5000,\ A_{n+1} = 1.01 \times A_n - 860$$

where A_n is the balance of the loan after n repayments have been made.

a What is the value of the repayment each compounding period?

b Use calculator recursion to determine how much is still owed on the loan after five repayments have been made.

c How much interest has been paid after five repayments?

ISBN 978-1-108-46154-2

4 A reducing-balance loan with interest compounding monthly and with monthly repayments can be modelled by the recurrence relation

$$A_0 = 14500, A_{n+1} = 1.0072 \times A_n - 1500$$

where A_n is the balance of the loan after n repayments have been made.

a Calculate the annual percentage rate of interest.

b Use calculator recursion to determine how much is owed on the loan after three months.

c How much interest has been paid after three months?

5 A reducing-balance loan with interest compounding monthly and with monthly repayments can be modelled by the recurrence relation

$$A_0 = 6300, A_{n+1} = 1.0095 \times A_n - 450$$

where A_n is the balance of the loan after n repayments have been made.

a Calculate the annual percentage rate of interest.

b Use calculator recursion to determine how much is owed on the loan after four months.

c How much interest has been paid after four months?

Problem-solving and modelling

CU

6 Andrea needs to borrow \$20 000. Her bank will charge interest at the annual percentage interest rate of 7.08%, compounding monthly. Andrea will be required to make monthly repayments of \$600, but Andrea thinks she can afford to pay \$800 instead.

a Construct a recurrence relation model for the loan with monthly repayments of \$600 and apply it to evaluate:

i how much Andrea would owe after five months

ii the total interest that Andrea would pay after five months

b Construct a recurrence relation model for the loan with monthly repayments of \$800 and apply it to evaluate:

i how much Andrea would owe after five months

ii the total interest that Andrea would pay after five months

c If Andrea makes monthly repayments of \$800 instead of \$600, how much interest will she save over the first five months of her loan?

ISBN 978-1-108-46154-2

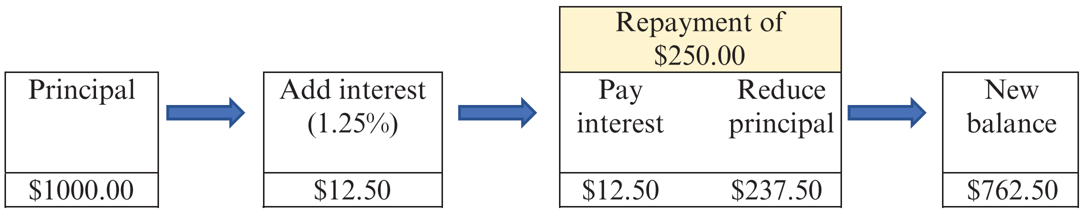
Cambridge University Press

8B Investigating reducing-balance loans

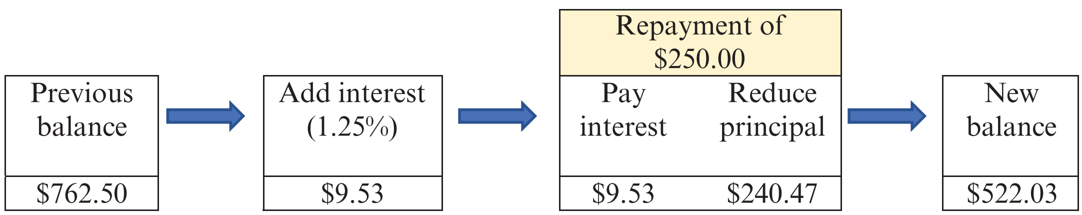
▶ Repayment schedules for reducing-balance loans

Consider a loan with principal \$1000. Interest will be charged at the rate of 1.25% per month and a repayment of \$250 will be made every month.

The calculation of the amount still owed after the first repayment has been made is shown here.



The calculation of the amount still owed after the second repayment has been made is shown here.



It is convenient to record all of the results of these calculations in a table called a *repayment schedule*.

A repayment schedule for the first three months of this reducing-balance loan is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	1000.00
1	250.00	12.50	237.50	762.50
2	250.00	9.53	240.47	522.03
3	250.00	6.53	243.47	278.56

Note: Some of the money values in the repayment schedule have been rounded to the nearest cent and may differ slightly to the values calculated using a recurrence relation model. The repayment schedule values are rounded after every calculation (if necessary) while the recurrence relation calculations are not.

ISBN 978-1-108-46154-2

Constructing a repayment schedule for a reducing-balance loan

At each step of the loan:

1 interest paid = interest rate per compounding period × unpaid balance
For example, when repayment 2 is made:
interest paid = 1.25% of \$762.50 = \$9.53

2 principal reduction = repayment − interest
For example, when repayment 2 is made:
principal reduction = \$250.00 − \$9.53 = \$240.47

3 balance of loan = previous balance − principal reduction
For example, when repayment 2 is made:
balance = \$762.50 − \$240.47 = \$522.03

4 total interest paid = total repayments made − (principal − balance)

Example 3 Constructing a repayment schedule for a reducing-balance loan

A repayment schedule for the first six repayments of a reducing-balance loan are shown in the table below. Interest is charged at the annual percentage interest rate of 7.68%, compounding monthly, with monthly repayments of \$600.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	8500.00
1	600.00	54.40	545.60	7954.40
2	600.00	50.91	549.09	7405.31
3	600.00	47.39	A	6852.70
4	600.00	43.86	556.14	6296.56
5	600.00	B	559.70	5736.86
6	600.00	36.72	563.28	5173.58

a What is the principal value of this loan?

b Calculate the value of A, the principal reduction from repayment number 3.

c Calculate the value of B, the interest paid with repayment number 5.

d Calculate the total interest paid after six repayments.

Solution

a The principal of the loan is the balance after repayment number 0.

The principal value of the loan is \$8500.

b Principal reduction = repayment amount − interest charged

A = \$600.00 − \$47.39
= \$552.61

ISBN 978-1-108-46154-2

c The interest paid with repayment number 5 is the interest rate percentage of the balance after repayment 4, rounded to the nearest cent.

$$B = \frac{7.68/12}{100} \times 6296.56$$
$$= 40.2979$$
$$= 40.30$$

d **1** Calculate the total of the repayments made.

Total repayments = 6 × 600
= 3600

2 Calculate the total interest paid.

Total interest = total repayments − (principal − balance)
= 3600 − (8500 − 5173.58)
= 273.58

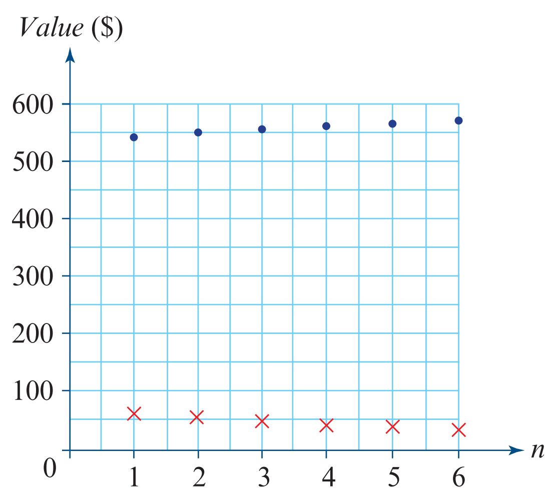
Note: This answer can be verified by adding all of the values in the interest column of the repayment schedule.

3 Write your answer.

The total interest paid on this loan after six repayments is $273.58.

Graphs of reducing-balance loans

The repayment schedule for a reducing-balance loan shows that after each successive repayment on the loan, the amount of interest that is charged decreases. It also shows that the principal reduction increases after each successive repayment.



The graph on the right shows the interest paid with each repayment as a red cross. This interest value decreases with each repayment.

The graph shows the reduction in the principal of the loan with each repayment as a blue dot. This value increases with each repayment.

The effect of the repayment amount on reducing-balance loans

A reducing-balance loan repayment schedule spreadsheet

The repayment schedule for a reducing-balance loan is very easy to construct using a spreadsheet. Such a spreadsheet could be used to further explore and investigate reducing-balance loans. There are instructions for setting up a reducing-balance loan spreadsheet and a copy of the completed spreadsheet on the next page.

ISBN 978-1-108-46154-2

Spreadsheet

Spreadsheet activity 8B: Repayment schedule for a reducing-balance loan
Use the spreadsheet to investigate the effect the repayment amount has on a reducing-balance loan.

The effect of the repayment amount on reducing-balance loans

The larger a regular repayment amount for a particular reducing-balance loan, the quicker that loan will be repaid; that is, the shorter the term of the loan will be. The term of a reducing-balance loan can be reduced by increasing the regular repayment amount.

A lump sum repayment is a repayment that is larger than the usual amount required by a reducing-balance loan. A lump sum repayment can reduce the amount of interest that is paid in a reducing-balance loan, but it can also reduce the term of the loan, if it is large enough. The larger the lump sum repayment, the more interest is saved and the shorter the term of the loan will be.

The earlier that a lump sum repayment is paid, the more interest is saved. This is because the amount of interest charged towards the end of the loan is much smaller than the amount of interest charged towards the start of a loan; that is, the interest charged decreases after each repayment of a reducing-balance loan.

Home loans are an example of how this feature of reducing-balance loans can be used to great advantage by a borrower. Many people choose to repay more than is required by their home loan agreements because this will mean they will pay less interest overall than if they paid only the required amount. They may also choose to pay larger amounts, perhaps the funds from the sale of another property, into their home loan as lump sum repayments, therefore significantly reducing the balance and the interest charged overall.

The effect of the repayment amount on reducing-balance loans

In general, for reducing-balance loans:

- increasing the repayment amount
  - can mean the loan is repaid in a shorter time
  - will mean less interest is paid overall
- lump sum repayments
  - can mean the loan is repaid in a shorter time
  - will mean less interest paid overall
- the earlier a lump sum repayment is made, the less interest is paid overall (or the more interest is saved overall).

ISBN 978-1-108-46154-2

Exercise 8B

Constructing a repayment schedule for a reducing-balance loan

Example 3 SF

1 Misaki borrowed $8400 and will be charged compounding interest at the rate of 11.4% per annum, compounding monthly. Misaki will make monthly repayments of $500.

- **a** Construct a repayment schedule to determine the balance of Misaki's loan after five repayments.
- **b** How much interest in total has Misaki paid after five repayments?

2 Vadik borrowed $2000 and will be charged compounding interest at the rate of 18.2% per annum, compounding weekly. Vadik will make weekly repayments of $100.

- **a** Construct a repayment schedule to determine the balance of Vadik's loan after five repayments.
- **b** How much interest in total has Vadik paid after:
  - **i** three weeks?
  - **ii** four weeks?

3 Sofia borrowed $4600 and will be charged compounding interest at the rate of 4.8% per annum, compounding monthly. The first three repayments that Sofia made were for $300. The next two repayments Sofia made were for $450.

Construct a repayment schedule to determine the balance of Sofia's loan after five repayments.

4 Jabulani borrowed $7500 and will be charged compounding interest at the rate of 3.8% per annum, compounding quarterly. He made two repayments of $300 and then doubled this amount for the next three repayments.

Construct a repayment schedule to determine the balance of Jabulani's loan after five repayments.

Using a spreadsheet to analyse reducing-balance loans

Spreadsheet 8BQ5 CF

5 Consider a reducing-balance loan of $25 000. This loan is charged interest at the rate of 3.8% per annum, compounding monthly. Monthly repayments of $2130 will be used to repay this loan, except for the final repayment.

- **a** Enter these loan details into the repayment schedule for a reducing-balance loan spreadsheet.
  - **i** What is the balance of the loan after five repayments?
  - **ii** The balance of the loan is negative for the first time after twelve repayments. What does this tell you?
  - **iii** How much interest in total has been charged on this loan?
- **b** The borrower made a lump sum repayment of $3000 as repayment number four.
  - **i** What is the value of the final repayment required now?
  - **ii** How much interest in total has been charged on this loan with the lump sum repayment?
  - **iii** How much interest has been saved by making this lump sum repayment?
- **c** If the lump sum repayment was made as a different repayment number, which one(s) should it be and why?

ISBN 978-1-108-46154-2

8C Solving problems involving reducing-balance loans

▶ The annuities formula

If the total number of repayments made to repay a reducing-balance loan is small, it is convenient to use a recurrence relation to model that loan. But for many reducing-balance loans, such as home loans, the principal is very large, and the repayments are made over a long period of time, usually decades.

For example, a home loan for $250 000 might be repaid over 25 years using monthly repayments. Analysing this reducing-balance loan using a recurrence relation would take $25 \times 12 = 300$ interest calculations. A repayment schedule for this loan would require 300 rows.

There is a formula that can be used to calculate the balance of any reducing-balance loan after any number of repayments and it is called the *annuities formula*. Instead of using recurrence relation symbols in this formula, we can use P to represent the principal amount of the loan and A to represent the future value of the loan after n compounding periods.

The annuities formula

Let P be the principal amount of the loan.

Let n be the total number of compounding periods.

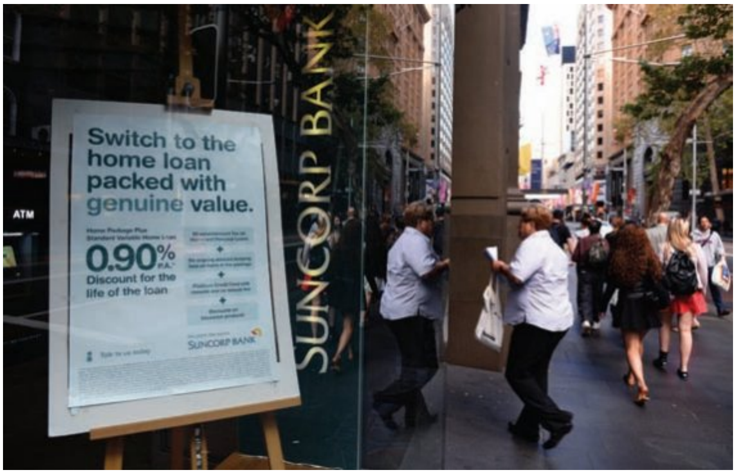
Let A be the balance of the loan after n repayments.

Let i be the decimal interest rate per compounding period.

Let R be the repayment made after each compounding period.

The balance of the reducing-balance loan after n compounding periods is

$$A = P(1+i)^n - R\frac{((1+i)^n - 1)}{i}$$



ISBN 978-1-108-46154-2

Example 4 Using the annuities formula

Nomsa has borrowed $8000 at an interest rate of 11.28% per annum, compounding monthly.

If she makes monthly repayments of $500, what will be the balance of the loan after one year?

Round your answer to the nearest cent.

Solution

1 Write down the values of P, i, R and n.	$P = 8000$ $i = \frac{11.28}{12 \times 100} = 0.0094$ $R = 500$ $n = 12$ (one year of monthly repayments)
2 Apply the annuities formula to calculate A.	$A = 8000(1 + 0.0094)^{12} - 500 \times \frac{((1 + 0.0094)^{12} - 1)}{0.0094}$ $A = 2630.419$
3 Write your answer, rounding to the nearest cent.	After 1 year (12 months), the balance of this loan is $2630.42.

The annuities formula can be rearranged algebraically in order to calculate other loan values, such as the repayment amount, annual interest rate or the number of compounds. This involves some complex calculations that are beyond the scope of this course. Technology, however, can be used to solve more complex problems involving reducing-balance loans.

ISBN 978-1-108-46154-2

Cambridge University Press

Solving reducing-balance loan problems using technology

Reducing-balance loan calculations involve six different values:

- Principal
- Annual percentage rate of interest
- Number of compounds per year
- Future balance after n compounds
- n, the total number of compounds
- Repayment amount

If the number of compounds per year is known, along with any four other values, the sixth can be calculated using technology such as an online calculation tool or a spreadsheet.

Spreadsheet **Spreadsheet activity 8C:** Reducing-balance Loan Calculator

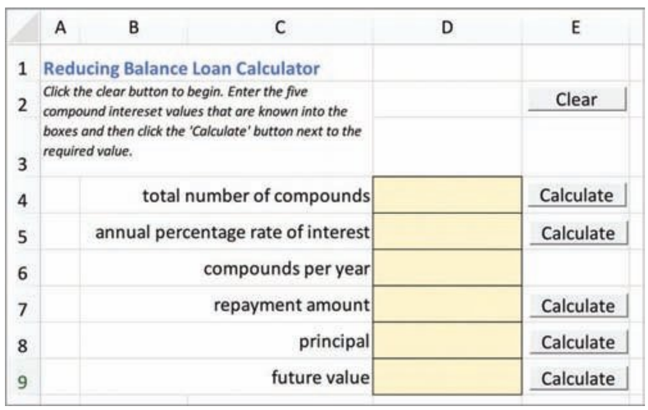
Using the reducing-balance loan calculator spreadsheet

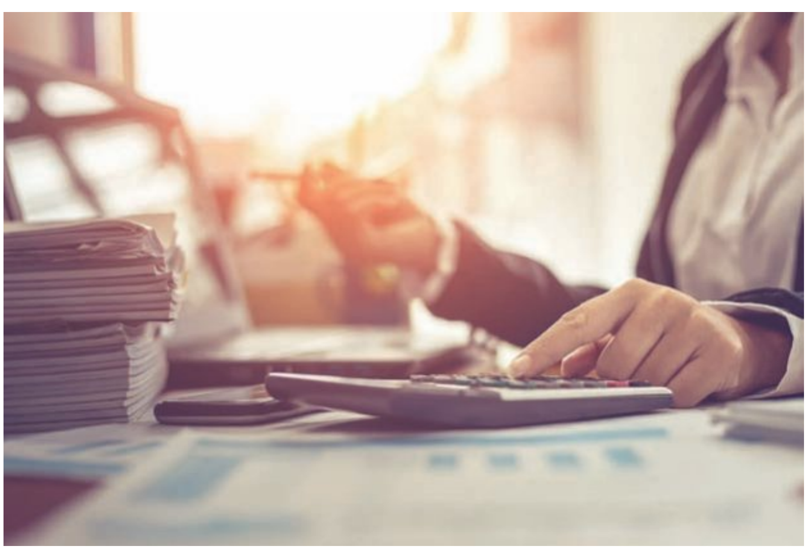
The Reducing-balance Loan Calculator spreadsheet is shown here.

Click the 'Clear' button before every calculation.

Enter the five known quantities into the corresponding orange box.

Click the 'calculate' button next to the quantity that you want to find. It will appear in the box.





ISBN 978-1-108-46154-2

Cambridge University Press

Example 5 Solving reducing-balance loan problems using a spreadsheet

Julian has a home loan with a principal value of \$250 000. He will be charged interest at the rate of 3.75% per annum, compounding fortnightly. If he makes regular fortnightly repayments of \$950, how long will it take for the balance to first be below \$200 000?

Solution

1 Identify the known quantities.

Principal = \$250 000
Future Balance = \$200 000
Compound per year = 26 (fortnightly)
Annual percentage rate of interest = 3.75%
Repayment = \$950

Spreadsheet
8CQ2

2 Click 'Clear' on the reducing balance loan spreadsheet and enter the known values.

Note: Type the numbers into the balance and principal boxes without the dollar sign or thousands comma. When you press Enter, the value will automatically be shown in currency format.

	A	B	C	D	E
1	Reducing Balance Loan Calculator				
2	*Click the clear button to begin. Enter the five compound intereset values that are known into the boxes and then click the 'Calculate' button next to the required value.*				Clear
3					
4			total number of compounds		Calculate
5			annual percentage rate of interest	3.7500000%	Calculate
6			compounds per year	26	
7			repayment amount	\$950.00	Calculate
8			principal	\$250,000.00	Calculate
9			future value	\$200,000.00	Calculate

3 Click the 'Calculate' button next to total number of compounds. This will be calculated and entered into the box for you.

	A	B	C	D	E
1	Reducing Balance Loan Calculator				
2	*Click the clear button to begin. Enter the five compound intereset values that are known into the boxes and then click the 'Calculate' button next to the required value.*				Clear
3					
4			total number of compounds	80.08493681	Calculate
5			annual percentage rate of interest	3.7500000%	Calculate
6			compounds per year	26	
7			repayment amount	\$950.00	Calculate
8			principal	\$250,000.00	Calculate
9			future value	\$200,000.00	Calculate

4 Round the number of compounds up to the nearest whole number and write your answer.

It will take 81 fortnights for Julian's loan balance to first be below \$200 000.

Note: If the number of compounding periods was rounded down to 80, the future value of the loan will be slightly more than \$200 000.

ISBN 978-1-108-46154-2

Example 6 Solving reducing-balance loan problems using a spreadsheet

Andrew has a loan of principal \$20 000 and is charged interest at the rate of 7.25% per annum, compounding monthly. Andrew would like to repay this loan fully after three years of equal monthly repayments, except for the final repayment.

a What is the usual monthly repayment for this loan?

b Which repayment number is Andrew's final repayment?

c How much will this final repayment be?

d How much interest in total has Andrew paid on this loan?

Solution

a **1** Identify the known quantities.

Principal = \$20 000
Future value = \$0
Compound per year = 12 (monthly)
Annual interest rate = 7.25%
Number of compounding periods = 3 years × 12
= 36 months

Spreadsheet
8CQ2

2 Click 'Clear' on the reducing balance loan spreadsheet and enter the known values.

Note: Type the numbers into the balance and principal boxes without the dollar sign or thousands comma. When you press Enter, the value will automatically be shown in currency format.

	A	B	C	D	E
1	Reducing Balance Loan Calculator				
2	Click the clear button to begin. Enter the five compound intereset values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			total number of compounds	36	Calculate
5			annual percentage rate of interest	7.2500000%	Calculate
6			compounds per year	12	
7			repayment amount		Calculate
8			principal	\$20,000.00	Calculate
9			future value	\$0.00	Calculate

3 Click the "Calculate" button next to repayment. The required fortnightly repayment will be calculated and entered into the box for you.

Note: this repayment amount is a rounded value. Repayments calculated this way will not usually repay the loan fully in the exact number of compounds used in the calculations. Almost all reducing-balance loans require the final repayment to be adjusted to fully repay the loan.

	A	B	C	D	E
1	Reducing Balance Loan Calculator				
2	Click the clear button to begin. Enter the five compound intereset values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			total number of compounds	36	Calculate
5			annual percentage rate of interest	7.2500000%	Calculate
6			compounds per year	12	
7			repayment amount	\$619.83	Calculate
8			principal	\$20,000.00	Calculate
9			future value	\$0.00	Calculate

4 Write your answer.

The usual monthly repayment is \$619.83 for this loan.

ISBN 978-1-108-46154-2

Cambridge University Press

b If Andrew wants to repay this loan after 3 years, then the last repayment of the third year is the final repayment of the loan.

The final repayment of Andrew's loan will be repayment number 36.

c **1** Click 'Clear' and enter all the known values, except for the future value.

	A	B	C	D	E
1	Reducing Balance Loan Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			total number of compounds	36	Calculate
5			annual percentage rate of interest	7.2500000%	Calculate
6			compounds per year	12	
7			repayment amount	$619.83	Calculate
8			principal	$20,000.00	Calculate
9			future value		Calculate

2 Click 'calculate' next to the future value.

Note: A future value that is negative indicates that the repayments have overpaid the loan. A future value that is positive indicates that the repayments have underpaid the loan. In both cases, the value shown in the future value must be added to the usual repayment amount.

	A	B	C	D	E
1	Reducing Balance Loan Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			total number of compounds	36	Calculate
5			annual percentage rate of interest	7.2500000%	Calculate
6			compounds per year	12	
7			repayment amount	$619.83	Calculate
8			principal	$20,000.00	Calculate
9			future value	$0.02	Calculate

3 Calculate the final repayment.

Final repayment $= \$619.83 + \$0.02$
$= \$619.85$

4 Write your answer.

Andrew's final repayment will be $619.85 to fully repay his loan.

d **1** Calculate the total repayments made.

Total repayments $= 35 \times \$619.83 + \$619.85$
$= \$22\,313.90$

2 Total interest = total repayments − principal

Total interest $= \$22\,313.90 - \$20\,000$
$= \$2313.90$

3 Write your answer.

Andrew will pay a total of $2313.90 interest on this loan.

ISBN 978-1-108-46154-2

Exercise 8C

All money amounts in this exercise should be rounded to the nearest cent.

Using the annuities formula

Example 4

1 Use the annuities formula to find the future value of each of the following reducing-balance loans after the given number of compounding periods. Round your answers to the nearest cent.

SF

	Principal	Annual percentage rate of interest	Compounding period	Repayment per compounding period	Balance after ...
a	\$8000	4.5%	Monthly	\$350	6 months
b	\$25 000	7.8%	Monthly	\$1200	1 year
c	\$240 000	8.3%	Quarterly	\$7900	5 years
d	\$75 000	6.9%	Quarterly	\$4800	2 years
e	\$50 000	4.6%	Weekly	\$350	1 year

Solving reducing-balance loan problems using technology

Example 5

2 Use technology to find the future value of each of the following reducing-balance loans after the given number of compounding periods. Round your answers to the nearest cent.

CF

	Principal	Annual percentage rate of interest	Compounding period	Repayment per compounding period	Balance after ...
a	\$2000	15.4%	Weekly	\$200	10 weeks
b	\$16 000	10.4%	Monthly	\$800	12 months
c	\$48 000	6.8%	Monthly	\$1400	2 years
d	\$135 000	4.72%	Quarterly	\$5000	5 years
e	\$350 000	3.81%	Monthly	\$4000	8 years

Problem-solving and modelling

3 A reducing-balance loan of \$90 000 is to be repaid with monthly repayments. Interest on this loan will be charged at the rate of 11% per annum, compounding monthly.

CF

The loan will be repaid over 30 years, that is 360 monthly repayments.

a Determine the repayment amount. Round your answer to the nearest cent.

b Use the rounded repayment amount from **3a** to determine the balance of the loan after 4 years.

ISBN 978-1-108-46154-2

CF

4 A building society offers \$240 000 home loans at an interest rate of 10.25% per annum, compounding monthly. Chelsea would like to pay off this loan over 15 years with monthly repayments.

a Determine the repayment amount. Round your answer to the nearest dollar.

b Use the rounded repayment amount from **4a** to determine how many months it takes for Chelsea's loan balance to first fall below \$200 000.

5 Dan arranges to make repayments of \$450 per month to repay a loan of \$20 000. If interest is being charged at the rate of 9.5% per annum, compounding monthly, find:

a the number of months required to repay this loan, rounded to the nearest month.

b the total interest paid after 2 years of repayments.

CU

6 Isla borrowed \$35 000 and will be charged interest at the rate of 10.5% per annum, compounding monthly. Isla would like to repay this loan over a period of 20 years.

a What monthly repayments would be required for this loan? Round your answer to the nearest cent.

b What is the balance of the loan after 4 years of repayments? Use the rounded answer from **6a** in your calculation.

After 4 years, the interest rate for Isla's loan was increased to 13.75% per annum. Isla would still like to repay this loan over a 20 year period.

c What is the new monthly repayment required for this loan? Round your answer to the nearest cent.

ISBN 978-1-108-46154-2

Review

Key ideas and chapter summary

AS

Reducing-balance loan

A **reducing-balance loan** is a type of compound interest loan that is repaid in regular repayments. The interest of a reducing-balance loan is calculated on the amount still owing after each repayment is made.

Recursive model for a reducing-balance loan

A recurrence relation that can be used to determine the balance of a reducing-balance loan after n compounding periods.
If the decimal rate of interest per compounding period is i and the regular repayment amount is R, then the recursive model for a reducing-balance loan is

$$A_0 = \text{principal of loan}, \quad A_{n+1} = r \times A_n - R$$

where $r = 1 + i$

Repayment schedule for a reducing-balance loan

A table that summarises the interest calculations at every compounding stage of a reducing-balance loan is called **repayment schedule**. A repayment schedule shows the repayment number, repayment amount, interest paid, principal reduction and loan balance after each repayment for some, or all, of the repayments of a loan.

Annuities formula

A formula that can be used to calculate the balance of a reducing-balance loan given the principal (P), decimal rate of interest per compounding period (i), number of compounding periods (n) and repayment amount (R). The annuities formula is

$$A = P(1 + i)^n - R\frac{((1 + i)^n - 1)}{i}$$

Total interest

The **total interest** paid on a reducing-balance loan after n repayments:
= total repayments made − (principal − balance after n repayments)

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Skills check

Having completed this chapter, you should be able to:

- use a recursive model to determine the balance of a reducing-balance loan after n compounding periods
- write a recursive model for a reducing-balance loan
- construct and interpret a repayment schedule for a reducing-balance loan
- understand the effect of the repayment amount on reducing-balance loans
- understand the effect of a lump sum repayment on a reducing-balance loan
- analyse reducing-balance loans with a spreadsheet
- use the annuities formula to determine the balance of a reducing-balance loan after n compounding periods
- solve reducing-balance loan problems using technology (spreadsheet calculation tool).

Multiple-choice questions

1 A reducing-balance loan is modelled by the recurrence relation shown below.

$$A_0 = 25\,000, \; A_{n+1} = 1.007 \times A_n - 400$$

where A_n is the balance of the loan after n months.
The balance of the loan after five months is:

A \$23 626.15
B \$23 859.14
C \$24 090.51
D \$25 707.38
E \$25 887.34

2 Hermione borrows \$18 000 and will be charged compound interest at the rate of 6.96% per annum compounding monthly. Hermione will repay the loan with monthly repayments of \$850.
If A_n is the balance of the loan after n months, a recurrence relation model for this reducing balance loan is:

A $A_0 = 18\,000, \; A_{n+1} = 1.0058 \times A_n - 850$
B $A_0 = 18\,000, \; A_{n+1} = 1.0174 \times A_n - 850$
C $A_0 = 18\,000, \; A_{n+1} = 1.0696 \times A_n - 850$
D $A_0 = 18\,000, \; A_{n+1} = 1.0174 \times A_n - 1242$
E $A_0 = 18\,000, \; A_{n+1} = 1.0696 \times A_n - 1242$

ISBN 978-1-108-46154-2

Use the following information to answer questions 3, 4 and 5.

A repayment schedule for the first five repayments of a reducing-balance loan is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	15 000.00
1	500.00	97.50	402.50	14 597.50
2	500.00	94.88	405.12	14 192.38
3	500.00	92.25	A	13 784.63
4	500.00	89.60	410.40	13 374.23
5	500.00	86.93	413.07	12 961.16

3 What is the principal of this loan?

A \$97.50 **B** \$402.50 **C** \$500.00 **D** \$12 961.16 **E** \$15 000

4 What is the value of A, the reduction in principal by repayment number 3?

A \$312.87 **B** \$407.75 **C** \$497.37 **D** \$500.00 **E** \$592.25

5 The interest charged on this loan compounds monthly and monthly repayments are made. The annual percentage rate of interest for this loan is closest to:

A 6.5% **B** 6.67% **C** 7.8% **D** 8.01% **E** 8.9%

6 A reducing-balance loan has principal \$180 000 and will be repaid with monthly repayments of \$1500. The interest rate for this loan is 6.5% per annum, compounding monthly.
Using the annuities formula, the balance of this loan after two years will be:

A \$144 000.00 **B** \$166 583.04 **C** \$167 667.37
D \$173 508.88 **E** \$176 807.03

7 Mei Hui has borrowed \$28 000 and will be charged compound interest at the rate of 6.4% per annum compounding monthly. She will repay this loan with exactly 24 repayments. The monthly repayment amount is closest to:

A \$850 **B** \$1046 **C** \$1246 **D** \$2415 **E** \$2854

8 A reducing-balance loan with principal \$40 000 is repaid with monthly repayments of \$550. The annual interest rate is 5.2%. How many repayments does it take to reduce the balance of this loan below \$35 000?

A 9 **B** 10 **C** 11 **D** 12 **E** 13

 ISBN 978-1-108-46154-2 Cambridge University Press

9 If the balance of a reducing-balance loan with principal \$175 000 has been reduced to \$89 573.97 after 35 quarterly repayments of \$4218, the annual percentage rate of interest for the loan is closest to:

A 5.2% **B** 5.4% **C** 5.6% **D** 5.8% **E** 6.0%

10 A loan of \$6000, plus interest, is to be repaid with exactly 12 quarterly repayments. Interest is charged at the rate of 10% per annum, compounding quarterly. If the first 11 repayments are of value \$580, the 12th and final repayment must be:

A \$228.97 **B** \$351.03 **C** \$580.00 **D** \$647.91 **E** \$931.03

Short-answer questions

SF

1 A reducing-balance loan is modelled using the recurrence relation shown below.

$$A_0 = 9500,\ A_{n+1} = 1.0035 \times A_n - 250$$

In the recurrence relation, A_n is the balance of the loan after n fortnightly repayments.

a What is the principal of this loan?

b What is the value of the fortnightly repayments?

c What is the balance of this loan after six repayments?

2 Barry is considering borrowing \$250 000 to buy a house. His bank will charge interest at the rate of 5.88% per annum, compounding monthly. Barry can afford to make repayments of \$2400 per month.

Let A_n be the balance of Barry's loan after n months.

a What is the monthly percentage rate of interest for Barry's loan?

b Construct a recurrence relation model for Barry's loan.

c What is the balance of Barry's loan after 6 months?

d How many months will it take for Barry's loan to have a balance below \$240 000 for the first time?

ISBN 978-1-108-46154-2

Cambridge University Press

3 A repayment schedule for a reducing-balance loan repaid with monthly repayments is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	3500.00
1	600.00	32.20	567.80	2932.20
2	600.00	26.98	573.02	2359.18
3	600.00	21.70	578.30	1780.88
4	600.00	16.38	583.62	1197.26
5	600.00	11.01	588.99	608.27
6	600.00	5.60	594.40	13.87

a What is the principal of this loan?

b What is the value of the monthly repayment?

c Calculate the:

i monthly interest rate, rounded to two decimal places.

ii annual interest rate, rounded to two decimal places.

d If this loan was fully repaid with repayment number six, how much should that repayment be?

4 A reducing-balance loan of $125 000 is to be repaid with monthly repayments of $1000. The annual percentage rate of interest for this loan is 3.72%.

a Construct a repayment schedule that shows the first four repayments of this loan.

b Use the annuities formula to determine the balance of the loan after two years.

5 Chelsea has borrowed $75 000 and will be charged interest at the rate of 7.44% per annum, compounding monthly. Chelsea will repay this loan with monthly repayments of $1500 per month.

a What is the balance of Chelsea's loan after two repayments?

b How much interest in total has been paid after two repayments?

6 Anton borrowed $149 000 to buy an apartment and has been paying $1000 per fortnight to repay this loan. The balance of Anton's loan is $84 987.19 after 3 years of payments. What is the annual percentage rate of interest for this loan? Round your answer to two decimal places.

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Extended-response questions

CF

1 Merlin borrowed \$84 000 and has been charged compound interest at the annual percentage interest rate of 7.08% per annum, compounding monthly. After three years of repayments, Merlin will still owe \$43 845.98 on his loan.

a Determine the monthly repayment amount that Merlin has been paying.

b How much will Merlin owe after a further two years of repayments?

After a total of four years of repayments, Merlin increases his repayment amount so that his loan is fully repaid after one further year.

c Determine the new monthly repayment for this loan. Round your answer to the nearest cent.

2 Karrie has borrowed \$62 000 and will repay this loan with monthly repayments of \$1200 per month. Interest is charged at the rate of 7.32% per annum, compounding monthly.

a Calculate the balance of the loan after one year.

b How much interest has been paid after one year?

After the 12th repayment, the annual percentage interest rate for Karrie's loan increased to 7.44% per annum, compounding monthly.

c If Karrie continues to pay repayments of \$1200, what is the balance of her loan after a further two years?

3 Angelique has borrowed \$160 000 and will be charged interest at the annual percentage interest rate of 5.64% per annum, compounding monthly.

Angelique is considering paying monthly repayments of \$840 per month.

a How many repayments of \$840 will Angelique need to make to reduce the balance of the loan to below \$100 000 for the first time?

b If Angelique pays monthly repayments of \$1000 per month, how many fewer repayments will she need to make to reduce the balance of the loan to below \$100 000 for the first time, when compared to repayments of \$840 per month?

4 A reducing-balance loan of \$180 000 is to be repaid with monthly repayments. Interest on this loan will be charged at the rate of 3.72% per annum, compounding monthly.

a How many monthly repayments of \$1000 will it take to halve the balance of the loan? Round your answer to the nearest whole number.

b What monthly repayment is required in order to halve the balance in 100 repayments? Round your answer to the nearest dollar.

ISBN 978-1-108-46154-2

5 Nicholas needs to borrow \$165 000 to purchase a flat. His bank has offered him two loans, each with different conditions as shown below: CU

Fixed Interest Home Loan:

- monthly repayments
- annual percentage rate of interest 4.6% per annum, compounding monthly
- fixed interest for the first five years of the loan
- monthly repayments of \$1071.00
- lump sum repayments allowed
- no extra repayments

Variable Interest Home Loan:

- monthly repayments
- annual percentage rate of interest 4.5% per annum, compounding monthly
- annual percentage rate of interest may vary at any time
- no lump sum repayments
- extra repayments allowed

Suppose that Nicholas chooses the Fixed Interest Home Loan.

a **i** What will be the balance of Nicholas' loan after five years?

ii Calculate the amount of interest that Nicholas will have paid after five years.

b Nicholas is expecting to receive funds from a family trust. This will become available after two years; that is, after 24 repayments of \$1071 have been made. He thinks he could make a lump sum repayment of \$20 000 as repayment number 25.

i What is the balance of Nicholas' loan after five years if this lump sum repayment is made?

ii How much interest will be saved by making this lump-sum repayment?

iii Explain why Nicholas should make this lump sum repayment as soon as the money is available to him rather than leaving it until later in the loan.

Suppose that Nicholas chooses the Variable Interest Home Loan.

Assume that the interest rate does not change in the first five years of the loan.

c What monthly repayments would be required to ensure the Fixed Interest Home Loan (with no lump sum repayment) and Variable Interest Home Loan would have the same balance after five years?

After repayment number 24, the annual percentage rate of interest for this loan increased to 4.6%.

d What new monthly repayment amount would be required to ensure the Fixed Interest Home Loan (with no lump sum repayment) and Variable Interest Home Loan (with increased interest rate) would have the same balance after five years?

Nicholas has decided not to make lump sum repayments on his loan.

The annual percentage rate of interest for the Variable Interest Home Loan is expected to increase by 0.05% each year over the next five years.

Nicholas can afford repayment of no more than \$1300 per month.

e Provide some advice to Nicholas about the best loan for him to choose. Justify your advice by explaining your mathematical reasoning.

ISBN 978-1-108-46154-2

Cambridge University Press

Chapter 9

9

Annuities and perpetuities

UNIT 4 INVESTING AND NETWORKING

Topic 1 Loans, investments and annuities

- How do we use a recursive model for an annuity?
- How do we determine the balance of an annuity after a certain amount of time?
- How do we construct and interpret a payment schedule for an annuity?
- How do we find the effect of the repayment amount on an annuity?
- How do we analyse annuities using a spreadsheet?
- How do we find the payment withdrawn, required principal and interest rate for a perpetuity?

ISBN 978-1-108-46154-2

9A Recursive model for annuities

Annuities

An **annuity** is a type of investment that can be used to provide a regular income to the investor. The *principal* of the annuity is the amount of money initially invested and this money will earn *interest* that is calculated and added to the investment at regular time periods. In addition to this interest, an investor may deposit a payment at regular time periods. This deposit adds to the principal and will continue to earn interest over the life, or term, of the annuity. The balance of the annuity will increase during the *deposit phase* of the investment.

After some time, the regular deposits an investor makes may stop and the investor may choose to withdraw a payment at regular time periods instead of adding to the investment. If the payment withdrawn is larger than the interest earned, then the balance of the annuity will decrease during this *withdrawal phase* of the investment.

Superannuation is an example of annuity investment. When a person begins working, their employer must make a regular payment into an investment fund or superannuation account. Over their working life, the balance of the superannuation account grows with regular deposits and interest. After retirement, the balance of the superannuation account will decrease because of the regular payments that are withdrawn to provide the retirement income.

A recursive model for an annuity

The deposit phase of an annuity and the withdrawal phase of an annuity have different recursive models.

A recursive model for the deposit phase of an annuity

Consider an annuity with a principal investment of \$100 000. Interest will be earned at the annual percentage interest rate of 6%, compounding yearly. A deposit of \$15 000 will be added to the annuity every year.

Let A_n be the balance of the annuity after n years.

The starting value of the recurrence relation is the principal value of the annuity, $A_0 = 100\,000$.

Each year, the annuity balance increases by the amount of interest that is earned, that is 6% of the previous balance, and then increases by the amount of the deposit.

So: balance next year = balance this year + interest earned + deposit
= balance this year + 6% of the balance this year + deposit
= 100% of the balance this year
+ 6% of the balance this year + deposit
= 106% of the balance this year + deposit
= 1.06 × balance this year + deposit

In recurrence relation symbols:

$$A_{n+1} = 1.06 \times A_n + 15\,000$$

ISBN 978-1-108-46154-2

A recursive model for the withdrawal phase of an annuity

Consider an annuity with a current balance of \$200 000. Interest will be earned at the annual percentage interest rate of 6%, compounding yearly. Each year, \$24 000 will be withdrawn from this annuity.

Let A_n be the balance of the annuity after n years.

The starting value of the recurrence relation is the principal value of the annuity, $A_0 = 200\,000$.

Each year, the annuity balance increases by the amount of interest that is earned, that is 6% of the previous balance, and then reduces by the amount withdrawn.

So: balance next year = balance this year + interest earned − withdrawal
= balance this year + 6% of the balance this year − withdrawal
= 100% of the balance this year + 6% of the balance this year − withdrawal
= 106% of the balance this year − withdrawal
= 1.06 × balance this year − withdrawal

In recurrence relation symbols:

$$A_{n+1} = 1.06 \times A_n - 24\,000$$

We now have a recurrence relation that can be used to model the balance of an annuity during the deposit and withdrawal phases of that investment.

A recurrence relation model for an annuity

Let A_n be the balance of an annuity after n compounding periods.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period for the annuity.

Let d be the payment deposited or withdrawn each compounding period.

A recurrence relation model for the balance of an annuity investment is:

Deposit phase: $A_0 =$ balance of annuity investment, $A_{n+1} = r \times A_n + d$

Withdrawal phase: $A_0 =$ balance of annuity investment, $A_{n+1} = r \times A_n - d$

where $r = 1 + i$

ISBN 978-1-108-46154-2

Example 1 Modelling an annuity during the deposit phase with a recurrence relation

Julian has just started a new job and his employer will deposit \$1500 every month into a superannuation account. Interest is earned at the rate of 4.5% per annum, compounding monthly.

a Construct a recurrence relation model for this annuity.

b Apply the recurrence relation to determine the balance of Julian's superannuation account after 5 months.

Solution

a **1** Write down the initial balance of the annuity, A_0.

$A_0 = 0$

2 Calculate the value of the decimal interest rate, i.

4.5% per annum compounding monthly

$$i = \frac{4.5}{12 \times 100}$$

$i = 0.00375$

3 Calculate the value of r.

$$r = 1 + i$$
$$= 1 + 0.00375$$
$$= 1.00375$$

4 Write your answer.

$A_0 = 0, \quad A_{n+1} = 1.00375 \times A_n + 1500$

b **1** Use calculator recursion to determine the value of A_5 (balance after five months).

0	
	0

Pressing '=' or enter once for A_1

ans × 1.00375 + 1500	
	1500

Pressing '=' or enter four more times for A_5

ans × 1.00375 + 1500	
	7556.461333

2 Write your answer (rounding to the nearest cent if necessary).

After five months, Julian's superannuation account will have a balance of \$7556.46.

ISBN 978-1-108-46154-2

Example 2 Modelling an annuity investment with a recurrence relation

Reza plans to travel overseas. He will invest \$12 000 in an annuity investment that earns interest at the rate of 6% per annum, compounding monthly. Reza will withdraw a payment of \$1500 per month while he travels.

a Construct a recurrence relation model for this annuity investment.

b Apply the recurrence relation to determine the balance of Reza's investment after three months.

c How many payments of \$1500 can Reza receive from this investment?

d After all payments of \$1500 have been received, Reza can receive one final payment to close the annuity. How much will this payment be?

Solution

a **1** Write down the principal of the investment, A_0.

$A_0 = 12\,000$

2 Calculate the value of the decimal interest rate, i.

6% per annum compounding monthly

$$i = \frac{6}{12 \times 100}$$

$i = 0.005$

3 Calculate the value of r.

$r = 1 + i$
$= 1 + 0.005$
$= 1.005$

4 Write your answer.

$A_0 = 12\,000, \; A_{n+1} = 1.005 \times A_n - 1500$

b **1** Use calculator recursion to determine the value of A_3 (balance after three months)

2 Press **AC** (Casio) or **clear** (TI) to create a blank calculation screen.

3 Type **12000** and then press = (Casio) or **enter** (TI)

12000	
	12000

Pressing '=' or enter once for A_1

4 Next, type **× 1.005 − 1500** and then press = (Casio), or **enter** (TI), four times to find A_3.

Ans × 1.005 − 1500	
	10560

Pressing '=' or enter three times for A_3

5 Write your answer (rounding to the nearest cent if necessary).

Ans × 1.005 − 1500	
	7658.364

After three months, Reza's annuity will have a balance of \$7658.36

ISBN 978-1-108-46154-2

Cambridge University Press

c	**1** Use calculator recursion to count the number of payments before the balance of the annuity is less than \$1500.	*Pressing '=' or enter four times for A_4* Ans × 1.005 − 1500 6196.65582 *Pressing '=' or enter eight times for A_8* Ans × 1.005 − 1500 276.3713495
	2 Write your answer.	*Reza can withdraw 8 payments of \$1500.*
d	**1** Use calculator recursion to apply the recurrence relation one last time.	*The balance of the annuity after 8 payments of \$1500 is \$276.37.*
	2 The balance is now negative. Add this negative amount (rounded to the nearest cent) to the usual payment amount to calculate the final payment.	*Final payment = \$1500 − \$1222.25* *= \$277.75*
	3 Write your answer.	*Reza's final payment will be \$277.75, after which there will be nothing left in his annuity.*

Calculating the total interest earned for annuities

When a payment is withdrawn from an annuity, part of this payment is made up of the interest that has been earned after that compounding period. The interest amount is usually smaller than the payment amount and so the balance of the annuity is used to make up the remaining part of the payment. This will cause the balance of the annuity to reduce after each compounding period.

The rule for calculating interest earned for an annuity are very similar to those for reducing-balance loans.

ISBN 978-1-108-46154-2

Cambridge University Press

Total interest earned by annuities

Let A_0 be the principal amount of an annuity.
Let n be the total number of compounding periods.
Let A_n be the balance of annuity after n compounding periods.
Let d be the payment deposited or withdrawn each compounding period.
Let I be the total interest earned after n compounding periods.

Deposit phase:

The increase in principal after n compounding periods $= A_n - A_0$.
The total amount deposited after n compounding periods $= n \times d$.

$$I = \text{increase in principal} - \text{total amount deposited}$$
$$= (A_n - A_0) - n \times d$$

Withdrawal phase:

The reduction in principal after n compounding periods $= A_0 - A_n$.
The total payments withdrawn after n compounding periods $= n \times d$.

$$I = \text{total payments withdrawn} - \text{reduction in principal}$$
$$= n \times d - (A_0 - A_n)$$

Another way of writing these rules is

$I = A_n - n \times d - A_0$ during the deposit phase
$I = A_n + n \times d - A_0$ during the withdrawal phase

Example 3 Analysing annuities with a recurrence relation

Diego has an annuity with a current balance of \$120 000.

His investment is earning interest at the rate of 7.68% per annum, compounding monthly.

For the next three months, Diego will deposit \$500 per month into his annuity, after which he will withdraw \$2500 per month for a further five months.

a Construct a recurrence relation model for Diego's annuity during the deposit phase.

b Apply this recurrence relation to determine the balance of Diego's account after three months.

c Construct a recurrence relation model for Diego's annuity during the withdrawal phase.

d How much interest has Diego's investment earned during these eight months?

ISBN 978-1-108-46154-2

Solution

a 1 Write down the values of A_0.
$A_0 = 120\,000$ (current balance of annuity)

2 Calculate the values of i, r and d.
7.68% per annum compounding monthly

$$i = \frac{7.68}{12 \times 100}$$
$$= 0.0064$$
$$r = 1 + i$$
$$= 1 + 0.0064$$
$$= 1.0064$$

$d = 500$ (deposit of \$500)

3 Write your answer.
$A_0 = 120\,000,\ A_{n+1} = 1.0064 \times A_n + 500$

b 1 Use calculator recursion to apply the recurrence relation three times to find A_3.

120000	
	120000

Pressing '=' or enter once for A_1

ans × 1.0064 + 500	
	121268

Pressing '=' or enter two more times for A_3

ans × 1.0064 + 500	
	123828.3975

2 Write your answer, rounding to the nearest cent if necessary.
The balance of Diego's investment is \$123 828.40 after three months.

c 1 The value of A_0 for this recurrence relation is the balance of the investment after the deposit phase.
$A_0 = 123\,828.40$

2 Write down the values of i, r and d.
$i = 0.0064$
$r = 1.0064$
$d = 2500$ (withdrawal of \$2500)

3 Write your answer.
$A_0 = 123\,828.40,\ A_{n+1} = 1.0064 \times A_n - 2500$

ISBN 978-1-108-46154-2

d **1** Calculate the amount of interest earned during the deposit phase.

$I = A_n - n \times d - A_0$

$I = 123\,828.40 - 3 \times 500 - 120\,000$

$I = \$2328.40$

2 Calculate the balance of the annuity after five months of withdrawals, A_5, by applying the recurrence relation five times.

123828.40
123828.40

Pressing '=' or enter once for A_1

ans × 1.0064 – 2500
122120.90

Pressing '=' or enter four more times for A_5

ans × 1.0064 – 2500
115180.93

3 Calculate the amount of interest earned during the withdrawal phase.

$I = A_n + n \times d - A_0$

$I = 115\,180.93 + 5 \times 2500 - 123\,828.40$

$I = \$3852.53$

4 Calculate the total interest earned.

Total interest $= \$2328.40 + \$3852.53$

$= \$6180.93$

Exercise 9A

Constructing a recurrence relation model for an annuity during deposit phase

Example 1

SF

1 The table below shows the principal amount, annual percentage interest rate, compounding period and payment amount for 6 annuities.

	Principal	Annual percentage interest rate	Compounding period	Deposit per compounding period
a	$0	2.5%	Yearly	$5000
b	$0	6.4%	Quarterly	$6500
c	$320 000	3.6%	Quarterly	$8000
d	$460 000	6.96%	Monthly	$4200
e	$845 000	4.92%	Monthly	$7500
f	$1 250 000	6.24%	Weekly	$2700

For each of these investments:

i construct a recurrence relation model

ii apply the recurrence relation model to find how much is left in the annuity after three compounding periods

ISBN 978-1-108-46154-2

Constructing a recurrence relation model for an annuity during withdrawal phase

2 The table below shows the principal amount, annual percentage interest rate, compounding period and withdrawal amount per compounding period for six annuities.

	Principal	Annual percentage interest rate	Compounding period	Withdrawal per compounding period
a	\$120 500	2.8%	Yearly	\$8000
b	\$276 000	5.04%	Quarterly	\$4600
c	\$358 000	5.72%	Quarterly	\$25 000
d	\$440 000	4.32%	Monthly	\$5000
e	\$845 000	8.04%	Monthly	\$9600
f	\$1 360 000	7.8%	Weekly	\$2900

For each of these investments:

i construct a recurrence relation model

ii apply the recurrence relation model to find the balance of the annuity after three compounding periods

Applying a recurrence relation model to analyse an annuity

Example 2

3 An annuity can be modelled by the recurrence relation

$$A_0 = 5000, \quad A_{n+1} = 1.01 \times A_n - 1030$$

where A_n is the balance of the annuity after n payments have been received.

a Explain how we can tell this annuity is in withdrawal phase.

b What is the payment that is withdrawn from this annuity each compounding period?

c Use calculator recursion to apply the recurrence relation and determine the amount left in the annuity after three payments have been received.

d How much interest has been earned after three payments have been received?

4 An annuity can be modelled by the recurrence relation

$$A_0 = 6000, \quad A_{n+1} = 1.005 \times A_n + 300$$

where A_n is the balance of the investment after n payments have been withdrawn.

a Explain how we can tell this annuity is in deposit phase.

b What is the payment that is deposited into this annuity each compounding period?

c Use calculator recursion to apply the recurrence relation and determine the balance of the annuity after five deposits.

d How much interest has been earned after five deposits have been made?

ISBN 978-1-108-46154-2

Example 3

5 An annuity can be modelled by the recurrence relations below.

Deposit phase: $A_0 = 40\,000,\quad A_{n+1} = 1.0018 \times A_n + 4500$
Withdrawal phase: $A_0 = P,\quad A_{n+1} = 1.0018 \times A_n - 7500$

where A_n is the balance of the investment after n monthly payments have been withdrawn or deposited.

a For the deposit phase, calculate:
  i the annual percentage rate of interest for this investment
  ii the balance of the annuity after five months

b After five months, the annuity will enter the withdrawal phase.
  i What is the monthly withdrawal amount?
  ii What is the value of P?
  iii What is the balance of the annuity after five withdrawals?

c How much interest has been earned:
  i during the deposit phase?
  ii during the withdrawal phase for five withdrawals?
  iii in total over this period of ten months?

6 An annuity can be modelled by the recurrence relations below.

Deposit phase: $A_0 = 265\,000,\quad A_{n+1} = 1.0031 \times A_n + 750$
Withdrawal phase: $A_0 = P,\quad A_{n+1} = 1.0031 \times A_n - 1800$

where A_n is the balance of the investment after n monthly payments have been withdrawn or deposited.

a For the deposit phase, calculate:
  i the annual percentage rate of interest for this investment
  ii the balance of the annuity after three months

b After three months, the annuity will enter the withdrawal phase.
  i What is the monthly withdrawal amount?
  ii What is the value of P?
  iii What is the balance of the annuity after three withdrawals?

c How much interest has been earned:
  i during the deposit phase?
  ii during the withdrawal phase for three withdrawals?
  iii in total over this period of six months?

7 A superannuation annuity has a current balance of \$125 000 and will earn interest at the annual percentage interest rate of 3.38% per annum, compounding fortnightly. Each fortnight, an employer deposits \$695 into this account.
Let A_n be the balance of the annuity after n deposits.

a Construct a recurrence relation model for this annuity.

b Use calculator recursion to determine the number of deposits that are required to raise the balance of the annuity above \$130 000.

c How much interest is earned by this annuity after nine fortnights?

CF

8 A sum of \$32 000 is invested in an annuity and will earn interest at the annual percentage interest rate of 4.32% per annum, compounding monthly. Monthly payments of \$3500 will be withdrawn.

Let A_n be the balance of the investment after n payments.

a Construct a recurrence relation model for this annuity.

b Use calculator recursion to determine the number of payments until the balance of the annuity first falls below \$20 000.

c How many payments of \$3500 can be withdrawn before the annuity is exhausted?

d After all payments of \$3500 have been received, how much money is left in the annuity?

e What is the final payment that can be received to fully exhaust the annuity?

9 The amount of \$54 000 is invested in an annuity. This investment will increase with monthly deposits of \$1500 for a period of six months. After these six months, monthly payments of \$1800 will be withdrawn.

When deposits are being made, interest is earned at the annual percentage interest rate of 7.68%. When withdrawals are being made, interest is earned at the annual percentage interest rate of 7.56%.

Let A_n be the balance of the investment after n payments.

a Construct a recurrence relation model for the deposit phase of this annuity.

b Find the balance of the annuity at the end of the deposit phase.

c Construct a recurrence relation model for the withdrawal phase of this annuity.

d How much interest is earned in total after these twelve months of investment?

ISBN 978-1-108-46154-2

9B Investigating annuities

▶ Payment schedules for annuities

Deposit phase

Consider an annuity with principal \$200 000. Interest will be charged at the rate of 1.05% per month and a deposit of \$3000 will be made every month.

The calculation of the balance of the annuity after the first deposit has been made is shown here.

Principal	→	Add interest (1.05%)	→	Add deposit	→	New balance
\$200 000		\$2100		\$3000		\$205 100

The calculation of the balance of the annuity after the second deposit has been made is shown here.

Previous balance	→	Add interest (1.05%)	→	Add deposit	→	New Balance
\$205 100		\$2153.55		\$3000		\$210 253.55

It is convenient to record all of the results of these calculations in a payment schedule.

A payment schedule for the first three months of this annuity is shown below.

Payment number	Deposit amount	Interest	Principal increase	Balance of annuity
0	0	0	0	200 000.00
1	3000.00	2100.00	5100.00	205 100.00
2	3000.00	2153.55	5153.55	210 253.55
3	3000.00	2207.66	5207.66	215 461.21

Note: Some of the money values in the payment schedule have been rounded to the nearest cent and may differ slightly to the values calculated using a recurrence relation model. The payment schedule values are rounded after every calculation (if necessary) while the recurrence relation calculations are not.

ISBN 978-1-108-46154-2

Cambridge University Press

Constructing a payment schedule for an annuity (deposit phase)

At each step of the investment:

1 interest earned = interest rate per compounding period × current balance
For example, before deposit 2 is made:
interest earned = 1.05% of \$205 100.00 = \$2153.55

2 principal increase = deposit + interest
For example, after deposit 2 is made:
principal increase = \$3000 + \$2153.55 = \$5153.553

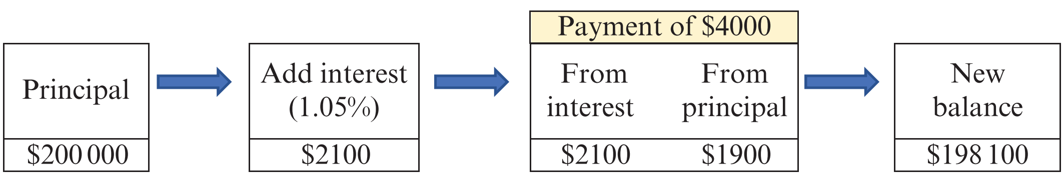
3 balance of annuity = previous balance + principal increase
For example, after deposit 2 is made:
balance = \$205 100.00 + \$5153.55 = \$210 253.55

4 total interest earned = (balance − principal) − total deposits made

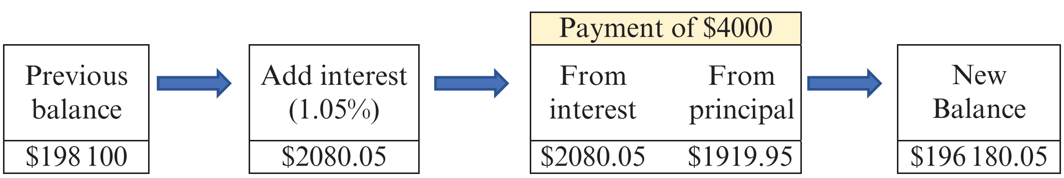
Withdrawal phase

Consider an annuity with principal \$200 000. Interest will be charged at the rate of 1.05% per month and a payment of \$4000 will be withdrawn every month.

To calculate the amount still invested in the annuity after the first payment:



To calculate the amount still invested in the annuity after the second payment:



It is convenient to record all of the results of these calculations in a payment schedule. This is very similar to the repayment schedules constructed for reducing-balance loans. The payment schedule and repayment schedules differ only in the headings and interpretations of the calculation results. The calculations are all performed in exactly the same way.

A payment schedule for the first three months of this annuity is shown below.

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	200 000.00
1	4000.00	2100.00	1900.00	198 100.00
2	4000.00	2080.05	1919.95	196 180.05
3	4000.00	2059.89	1940.11	194 239.94

Note: Some of the money values in the payment schedule have been rounded to the nearest cent and may differ slightly to the values calculated using a recurrence relation model. The payment schedule values are rounded after every calculation (if necessary) while the recurrence relation calculations are not.

ISBN 978-1-108-46154-2

Constructing a payment schedule for an annuity

At each step of the investment:

1 interest earned = interest rate per compounding period × current balance
For example, before withdrawal 2 is made:
interest earned = 1.05% of \$198 100 = \$2080.05

2 principal reduction = payment − interest
For example, after withdrawal 2 is made:
principal reduction = \$4000 – \$2080.05 = \$1919.95

3 balance of annuity = previous balance − principal reduction
For example, after withdrawal 2 is made:
balance = \$198 100 – \$1919.95 = \$196 180.05

4 total interest earned = total payments received − (principal − balance)

Example 4 Interpreting a payment schedule for an annuity in withdrawal phase

The payment schedule for the first six payments from an annuity are shown in the table below. The interest compounds monthly and payments are also withdrawn monthly.

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	65 000.00
1	1500.00	455.00	1045.00	63 955.00
2	1500.00	447.69	A	62 902.69
3	1500.00	440.32	1059.68	61 843.01
4	1500.00	432.90	1067.10	60 775.91
5	1500.00	B	1074.57	59 701.34
6	1500.00	417.91	1082.09	58 619.25

a What is the principal value of this annuity?

b Calculate the annual percentage rate of interest for this annuity.

c Calculate the value of A, the principal reduction from payment number 2.

d Calculate the value of B, the interest earned before payment number 5.

e What is the total interest earned after 6 payments have been made?

ISBN 978-1-108-46154-2

Solution

a The principal of the annuity is the balance after payment number 0.

The principal value of the annuity is \$65 000.

b **1** Choose any quarter.

Choose quarter 1.

Note: It is best to choose a year that does not contain values calculated previously in the question.

2 Write the interest amount as a percentage of the previous balance.

$$\text{Interest rate} = \frac{\text{interest amount}}{\text{previous balance}} \times 100\%$$
$$= \frac{455.00}{65000} \times 100\%$$
$$= 0.007$$

Note: Because the values in the table have been rounded to the nearest cent, the interest rates calculated using each of the years may be slightly different.

3 Convert this quarterly interest rate to an annual interest rate by multiplying by 12 (12 months per year).

$$\text{Interest rate} = 0.007 \times 12$$
$$= 8.4\% \text{ per annum}$$

4 Write your answer.

The annual percentage rate of interest for this annuity is 8.4%.

c Principal reduction = payment withdrawn − interest charged

$$A = \$1500 - \$447.69$$
$$= \$1052.31$$

d The interest earned before payment number 5 is the interest rate percentage of the balance after payment number 4, rounded to the nearest cent.

$$B = \frac{8.4}{12 \times 100} \times \$60\,775.91$$
$$= \$425.43$$

1 Alternative solution

2 Interest = payment − principal reduction

$$B = \$1500 - \$1074.57$$
$$= \$425.43$$

e **1** Calculate the total payments received.

Total payments received
$$= 6 \times \$1500$$
$$= \$9000$$

2 Calculate the total interest earned.

Note: this answer can be verified by adding the interest amounts from the interest column of the payment schedule.

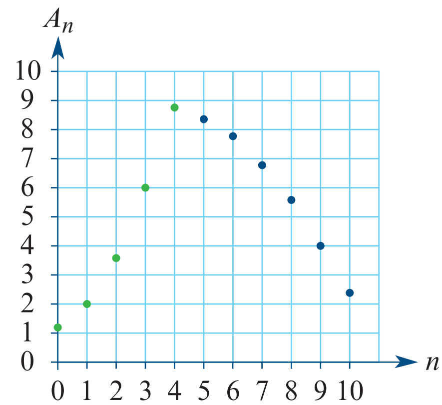
Total interest
$$= \text{total payments} - (\text{principal} - \text{balance})$$
$$= \$9000 - (\$65\,000 - \$58\,619.25)$$
$$= \$2619.25$$

The payment schedule for an annuity in deposit phase shows that after each successive deposit, the amount of interest that is earned increases. This is because the balance of the investment is increasing over time meaning more and more interest is earned.

ISBN 978-1-108-46154-2

The payment schedule for an annuity in withdrawal phase shows that after each successive withdrawal, the amount of interest that is earned decreases. This is because the balance is decreasing over time meaning less and less interest is earned.

If the annuity continues in deposit phase, it will continue to grow in value indefinitely. At some point in the investment, however, it is usual for it to enter withdrawal phase, typically to provide an income for the investor. The longer the annuity exists in withdrawal phase, the more of the balance is required to make the regular payment and so the balance of the annuity decreases more rapidly later in the life of the annuity than it did in the early stages, as shown in the example graph on the right.



The green dots represent the balance of the investment during the deposit phase. After each additional deposit, the amount of interest earned increases.

The blue dots represent the balance of the investment during the withdrawal phase. After each withdrawal is made, the amount of interest earned decreases.

Investigating the effect of the payment amount on the duration of an annuity

As you have seen, the calculations used to create repayment schedules for reducing-balance loans and payment schedules for annuities in the withdrawal phase are very similar. The only difference is the interpretation of the quantities involved. A payment withdrawn from an annuity has the same effect on the balance of that annuity as a repayment does on a reducing-balance loan. The only difference is that the annuity balance is money that you are owed, and a reducing-balance loan balance is money the bank is owed.

Because of the similarities in calculation, the reducing-balance loan spreadsheet can be used to perform the interest calculations of an annuity in the withdrawal phase, with only small alterations to the column headings.

Spreadsheet **Spreadsheet activity 9B:** Payment schedule for an annuity

Investigation 9B: How does the withdrawal amount affect the length of an annuity here?

In general, increasing the payment withdrawn results in a shorter term of the investment. The funds will not last as long as if the payment withdrawn was smaller.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 9B

Constructing a payment schedule for an annuity in growth phase

SF

1 Create a payment schedule for the following annuities in growth phase, showing five deposits.

- **a** Initial balance of \$50 000 earning interest at the rate of 3.2% per annum, compounding quarterly, with deposits of \$4000 per quarter.
- **b** Initial balance of \$135 000 earning interest at the rate of 4.32% per annum, compounding monthly, with deposits of \$1200 per month.

Constructing a payment schedule for an annuity in withdrawal phase

2 Create a payment schedule for the following annuities in withdrawal phase, showing five withdrawals.

- **a** \$25 000 earning interest at the rate of 6.9% per annum, compounding monthly, with payments of \$1000 per month.
- **b** \$380 000 earning interest at the rate of 4.8% per annum, compounding quarterly, with payments of \$12 000 per quarter.

Example 4

CF

3 A payment schedule for the first five withdrawals from an annuity is shown below. The interest compounds monthly and payments will be withdrawn after each month.

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of loan
0	0	0	0	164 000.00
1	3500.00	787.20	2712.80	161 287.20
2	3500.00	774.18	2725.82	158 561.38
3	3500.00	761.09	2738.91	155 822.47
4	3500.00	747.95	2752.05	153 070.42
5	3500.00	734.74	2765.26	150 305.16

- **a** What is the principal of the investment?
- **b** What is the annual percentage rate of interest for the investment? Round your answer to one decimal place.
- **c** What is the balance of the investment after three payments have been withdrawn?
- **d** How much interest was received before the fourth payment?
- **e** By how much did the fifth payment reduce the balance of the loan?
- **f** Complete the following.
  - **i** Calculate the total interest earned after five payments.
  - **ii** Verify your answer to **f i** above by adding values from the interest column of the table.
- **g** Construct the next two rows of this payment schedule.

ISBN 978-1-108-46154-2

Cambridge University Press

9C Solving problems involving annuities with technology

*An online update is available in the Interactive Textbook

Annuities are often used for long-term investments that involve many payments. Modelling these investments with recurrence relations is cumbersome because of the number of calculations that must be performed. Similarly, a payment schedule for an annuity that has a life of many years would be large and difficult to manage.

Problems involving annuities with a large number of payments can be solved using a variation of the annuities formula (studied in Chapter 8) or with the use of technology.

The annuities formula

The annuities formula can be used to find the balance of an annuity investment after n payments have been received.

The annuities formula

Let P be the principal amount of the annuity.

Let n be the total number of compounding periods.

Let A be the balance of the annuity after n deposits.

Let i be the decimal interest rate per compounding period.

Let M be the amount of each deposit or withdrawal.

The balance of the annuity after n compounding periods is

Deposit phase:

$$A = M\frac{((1 + i)^n - 1)}{i} \quad \text{(for principal = \$0)}$$

$$A = P(1 + i)^n + M\frac{((1 + i)^n - 1)}{i}$$

Withdrawal phase:

$$A = P(1 + i)^n - M\frac{((1 + i)^n - 1)}{i}$$

Example 5 Using the annuities formula

Edward is starting a new job. His salary each week will be \$495 and his employer will pay 9% of this into a superannuation account for him. This superannuation fund will earn interest at the rate of 4.16% per annum, compounding weekly.

a How much money will Edward's employer deposit into the account each week?

b Assume that the interest rate for this account does not change. Calculate the balance of Edward's superannuation account after 15 weeks of work.

ISBN 978-1-108-46154-2

Solution

a **1** Each deposit, M, is 9% of Edward's salary.

$M = 9\% \text{ of } \$495.00$

$= \frac{9}{100} \times \$495.00$

$= \$44.55$

2 Write down your answer.

Each week, Edward's employer will deposit \$44.55 into Edward's superannuation account.

b **1** Write down the value of i and n.

4.16% per annum compounding weekly

$i = \frac{4.16}{52 \times 100}$

$= 0.0008$

$n = 15$

2 Apply the annuities formula to calculate A.

$A = M\frac{((1 + i)^n - 1)}{i}$

$= 44.55 \times \frac{((1 + 0.0008)^{15} - 1)}{0.0008}$

$= 672.0052\ldots$

3 Write your answer, rounding to the nearest cent.

After 15 weeks, the balance of Edward's superannuation account is \$672.01.

Example 6 Using the annuities formula

Wendy invested her superannuation funds of \$375 000 in an annuity that will pay her interest at the rate of 6.12% per annum, compounding monthly. Wendy will withdraw a monthly payment of \$4000 from this investment. Use the annuities formula to calculate the balance of Wendy's investment after 10 years.

Solution

1 Write down the values of P, i, M and n.

$P = 375\,000.00$

$i = \frac{6.12}{12 \times 100} = 0.0051$

$M = 4000.00$

$n = 10 \times 12$ (10 years of monthly repayments)

$= 120$

2 Apply the annuities formula to calculate A.

$A = P(1 + i)^n - M\frac{((1 + i)^n - 1)}{i}$

$= 375\,000 \times (1 + 0.0051)^{120}$

$- 4000\frac{((1 + 0.0051)^{120} - 1)}{0.0051}$

$= 30\,664.83067$

3 Write your answer, rounding to the nearest cent.

After 10 years, the balance of Wendy's investment is \$30 664.83.

ISBN 978-1-108-46154-2

The annuities formula is very easy to use to calculate the balance of an annuity investment given all of the other values for the investment. It isn't as easy to use to calculate other values such as the number of compounding periods, withdrawal amount or interest rate. These sorts of problems can easily be solved using technology.

Solving annuity problems using technology

Annuity calculations involve six different values:

- Principal
- Annual interest rate
- Number of compounds per year
- Future balance after n compounds
- n, the number of compounds that will be considered
- Withdrawal amount

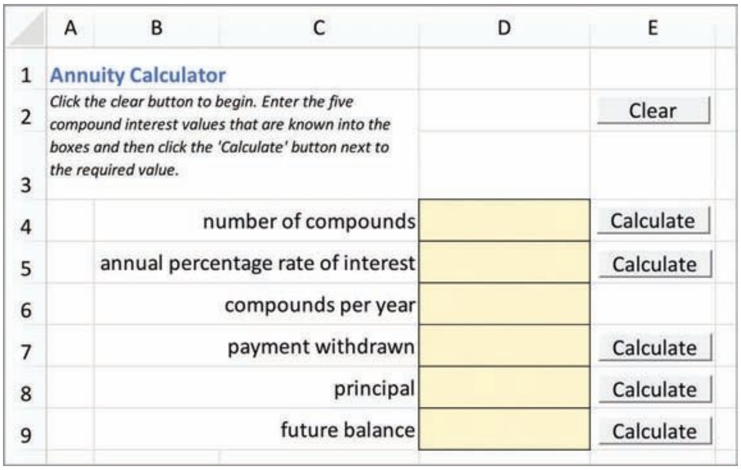
If the number of compounds per year is known, along with any four other values, the sixth can be calculated using technology such as a CAS calculator, online calculation tool or a spreadsheet.

Spreadsheet

Spreadsheet activity 9C: Annuity Calculator

Using the Annuity Calculator spreadsheet

The Annuity Calculator is shown here. It works in the same way as the reducing-balance loan calculator from Chapter 8.



Click the 'Clear' button before every calculation.

Enter the five known quantities into the spreadsheet and click the 'Calculate' button next to the quantity that you want to find. It will appear in the box.

ISBN 978-1-108-46154-2

Example 7 Solving annuity problems using a spreadsheet

Nino has \$475 000 to invest in an annuity. His investment will earn interest at the rate of 5.2% per annum, compounding quarterly and Nino plans to withdraw \$12 000 from his investment every quarter.

a Verify that Nino's investment will last for a period of at least ten years.

b How many payments of \$12 000 will Nino be able to withdraw?

c What is the final withdrawal that Nino can make?

d If Nino withdraws \$10 000 each month instead of \$12 000, how much longer will his investment last?

Solution

Spreadsheet

a **1** Identify the known quantities.

Principal = \$475 000
Compounds per year = 4 (quarterly)
Percentage annual rate of interest = 5.2%
Number of compounding periods = 40 (10 × 4 quarters)
Quarterly withdrawal = \$12 000

2 Calculate the balance after ten years.

3 Click 'Clear' on the annuity spreadsheet and enter the known quantities.

Note: Type the numbers into the balance and principal boxes without the dollar sign or thousands comma. When you press Enter, the value will automatically be shown in currency format.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			number of compounds	40	Calculate
5			percentage rate of interest	5.2000000%	Calculate
6			compounds per year	4	
7			payment withdrawn	\$12,000.00	Calculate
8			principal	\$475,000.00	Calculate
9			future balance		Calculate

4 Click 'Calculate' next to future balance. The value of the balance after ten years will be automatically entered.

Note: this balance amount is a rounded value.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			number of compounds	40	Calculate
5			annual percentage rate of interest	5.2000000%	Calculate
6			compounds per year	4	
7			payment withdrawn	\$12,000.00	Calculate
8			principal	\$475,000.00	Calculate
9			future balance	\$171,920.49	Calculate

ISBN 978-1-108-46154-2

5 If the balance of the annuity is positive, the annuity will last for at least ten years.	*The balance after ten years is positive.*
6 Write your answer.	*The balance of Nino's investment after ten years is positive and so his investment will last for at least ten years.*

b **1** Click 'Clear' and add all known values to calculate the number of compounding periods (number of withdrawals). Make the future balance equal to zero to calculate the number of possible withdrawals.

	A	B	C	D	E
1	Annuity Calculator				
2	*Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.*				Clear
3					
4			number of compounds		Calculate
5			annual percentage rate of interest	5.2000000%	Calculate
6			compounds per year	4	
7			payment withdrawn	$12,000.00	Calculate
8			principal	$475,000.00	Calculate
9			future balance	$0.00	Calculate

2 Click 'Calculate' next to number of compounds.

	A	B	C	D	E
1	Annuity Calculator				
2	*Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.*				Clear
3					
4			number of compounds	55.956569	Calculate
5			annual percentage rate of interest	5.2000000%	Calculate
6			compounds per year	4	
7			payment withdrawn	$12,000.00	Calculate
8			principal	$475,000.00	Calculate
9			future balance	$0.00	Calculate

3 Write your answer.

Nino's investment will allow 55 withdrawals of $12 000.

Note: the number of compounds must be rounded down to 55. The 56th withdrawal will need to be smaller than usual.

c **1** Use the spreadsheet to calculate the future balance after 56 withdrawals.

	A	B	C	D	E
1	Annuity Calculator				
2	*Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.*				Clear
3					
4			number of compounds	56	Calculate
5			annual percentage rate of interest	5.2000000%	Calculate
6			compounds per year	4	
7			payment withdrawn	$12,000.00	Calculate
8			principal	$475,000.00	Calculate
9			future balance	-$517.96	Calculate

2 Calculate the 56th and final withdrawal amount.

Final withdrawal = $12 000 − $517.96
= $11 482.04

d **1** Use the spreadsheet to calculate the number of $10 000 withdrawals that are possible.

	A	B	C	D	E
1	Annuity Calculator				
2	*Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.*				Clear
3					
4			number of compounds	74.40460397	Calculate
5			annual percentage rate of interest	5.2000000%	Calculate
6			compounds per year	4	
7			payment withdrawn	$10,000.00	Calculate
8			principal	$475,000.00	Calculate
9			future balance	$0.00	Calculate

2 Compare the two options.

Note: The same result is obtained by subtracting the total number of repayments. The number of extra withdrawals = 75 − 56 = 19.

$12 000 withdrawals will last for 55 quarters + 1 smaller withdrawal.

$10 000 withdrawals will last for 74 quarters + 1 smaller withdrawal.

Extra withdrawals = 74 − 55
= 19

3 Write your answer.

Nino's investment will last an extra 19 quarters if he withdraws $10 000 per quarter instead of $12 000.

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 9C

Using the annuities formula

Example 5

1 Use the annuities formula to find the balance of each of the following annuities (initial principal = $0) after the given number of compounding periods. Round your answers to the nearest cent.

SF

	Annual percentage rate of interest	Compounding period	Payment per compounding period	Balance after …
a	4.14%	Monthly	$800.00	8 months
b	5.25%	Monthly	$480.00	2 years
c	6.03%	Quarterly	$2500.00	8 years
d	4.28%	Quarterly	$1800.00	5 years
e	6.14%	Weekly	$100.00	2 years

Example 6

2 Use the annuities formula to find the balance of each of the following annuities in the withdrawal phase after the given number of compounding periods. Round your answers to the nearest cent.

	Principal	Annual percentage rate of interest	Compounding period	Payment per compounding period	Balance after …
a	$160 000.00	4.14%	Monthly	$2500.00	8 months
b	$375 000.00	5.25%	Monthly	$4800.00	2 years
c	$415 000.00	6.03%	Quarterly	$8400.00	8 years
d	$520 000.00	4.28%	Quarterly	$9700.00	5 years
e	$256 000.00	6.14%	Weekly	$500.00	2 years

Solving annuity problems using technology

Example 7

3 Leigh invests $64 000 in an annuity and will be charged interest at the rate of 6.25% per annum, compounding monthly. Leigh will withdraw $1275 per month.

CF

a How many withdrawals of $1275 will Leigh be able to make?
b His final withdrawal will be smaller than $1275. What is the value of this final payment?
c What is the total interest that Leigh has earned from this investment?

4 Raj invests $85 500 in an annuity that will pay interest at the rate of 7.25% per annum, compounding quarterly. If Raj receives a regular quarterly payment of $5000, how many payments in total will he receive?

ISBN 978-1-108-46154-2

5 Stephanie invests \$40 000 in an annuity and would like to receive a monthly payment for exactly 10 years. Interest on her investment is earned at the rate of 7.5% per annum, compounding monthly.

a How many payments will Stephanie receive?

b The last of these payments will be smaller than all of the others.

i What is the value of the usual payment? Round your answer to the nearest cent.

ii What is the value of the final payment? Round your answer to the nearest cent.

c What is the total interest that Stephanie will earn from her investment?

6 Kaspar has an annuity investment that earns interest at the rate of 8.16% per annum, compounding monthly, from which he receives a monthly payment of \$3600. The balance of Kaspar's investment was \$391 262.50 after the first two years.

a What was the principal of Kaspar's investment?

b How much interest has Kaspar earned after two years?

7 Simon has an annuity investment of principal \$480 000, from which he receives a payment of \$1250 per week. The balance of Simon's investment is \$439 252.37 after the first year of payments.

a What is the annual interest rate for Simon's investment?

b Calculate the balance of Simon's investment after two years.

c How many payments in total can Simon expect before his investment is fully exhausted?

d If the interest rate on Simon's investment decreases by 0.2%, and if Simon continues to withdraw \$1250 per week, how many fewer payments will he receive compared to the number of payments calculated in **7c** above?

ISBN 978-1-108-46154-2

9D Perpetuities

Investigation

This section can begin with an investigation.

Spreadsheet

Spreadsheet activity 9D-1: Payment schedule for an annuity

Investigation 9D: Investigating payments withdrawn from an annuity.

For any annuity in withdrawal phase, there exists a withdrawal amount for which the balance of the annuity does not change over time. A withdrawal higher than this will result in a balance that decreases over time, eventually to zero. A withdrawal lower than this will result in a balance that increases over time.

Perpetuities

An annuity that has regular withdrawals of a value equal to the interest earned after one compounding period is called a **perpetuity**. As observed in the investigation above, the balance of an annuity will remain constant forever, or *in perpetuity*, if the payment withdrawn is the same as the interest earned. This means that, while the investment remains in place, the regular payment can be withdrawn for as long as required.

Example 8 Calculating the payment withdrawn from a perpetuity

A university has invested \$80 000 into a perpetuity, the interest from which will provide an annual prize for one of their students. The investment earns interest at the rate of 8.4% per annum, compounding yearly.

What is the value of the student prize?

Solution

1 The interest each year is 8.4% of the principal.

$$\text{Interest} = 8.4\% \text{ of } \$80\,000$$
$$= \frac{8.4}{100} \times \$80\,000$$
$$= \$6720$$

2 Write your answer.

The annual student prize has value \$6720.

The payments withdrawn from a perpetuity are equal in value to the interest that is earned after each compounding time period. This means that the balance of the perpetuity does not change.

ISBN 978-1-108-46154-2

Perpetuities

Let A be the future value of the perpetuity after n payments have been withdrawn.

Let n be the total number of compounds.

Let P be the principal amount of the investment.

Let i be the decimal rate of interest per compounding period.

Let d be the payment withdrawn after each compounding period.

The balance of a perpetuity investment is constant and equal to the principal, $A = P$

$d = i \times P \qquad P = \dfrac{d}{i} \qquad i = \dfrac{d}{P}$

Example 9 Calculating the payment withdrawn from a perpetuity using a recurrence relation

The balance of an annuity investment after n months, A_n, is modelled by the recurrence relation below.

$$A_0 = 500\,000, A_{n+1} = 1.005 \times A_n - d$$

What payment should be withdrawn from this investment every month if the payments are to be withdrawn in perpetuity?

Solution

There are two ways to solve this problem. One involves an interest calculation and the other uses the rule for the recurrence relation.

Method 1

1 Calculate the interest rate i using the multiplying factor from the recurrence relation.

$1.005 = 1 + i$
$i = 1.005 - 1$
$i = 0.005$

2 Calculate the payment withdrawn.

$d = i \times P$
$= 0.005 \times 500\,000$
$= 2500$

3 Write your answer.

The payment withdrawn from this perpetuity investment should be \$2500.

ISBN 978-1-108-46154-2

Cambridge University Press

Method 2

1 The balance of the perpetuity is always equal to P. — $A_{n+1} = 1.005 \times A_n - d$

2 Write the recurrence relation with A_n and A_{n+1} both equal to P. — $P = 1.005 \times P - d$

3 Use the value of $P(A_0)$ from the recurrence relation. — $500\,000 = 1.005 \times 500\,000 - d$

4 Solve for d — $d = 1.005 \times 500\,000 - 500\,000$

$d = 2500$

5 Write your answer. — The payment withdrawn from this perpetuity investment should be \$2500.

Example 10 Calculating the investment required to establish a perpetuity

How much money will need to be invested in a perpetuity account, earning interest at the rate of 4.2% per annum compounding monthly, if \$200 will be withdrawn every month? Round your answer to the nearest cent.

Solution

1 Write down the values of i and d.

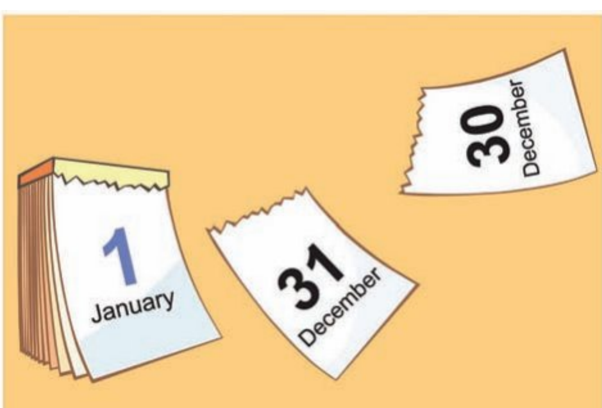
$$i = \frac{4.2}{12 \times 100} = 0.0035$$

$$d = 200$$

2 Use the rule to calculate the principal.

$$P = \frac{d}{i} = \frac{200}{0.0035} = 57142.85714$$

3 Write your answer, rounding to the nearest cent. — The principal invested in this perpetuity should be \$57142.86



ISBN 978-1-108-46154-2

Example 11 Calculating the interest rate required for a perpetuity

A university mathematics faculty has $30 000 to invest. It intends to award an annual mathematics prize of $1500 to a student using the interest from this investment. If the award is to be made in perpetuity, what is the annual interest rate required for this investment?

Solution

1 Write down the values of P and d.

$P = 30\,000$
$d = 1500$

2 Use the rule to calculate the decimal rate of interest per compounding period.

$i = \dfrac{d}{P}$
$= \dfrac{1500}{30000}$
$= 0.05$

3 Convert the value of i to an annual percentage rate.

Annual percentage interest rate $= 0.05 \times 1 \times 100\%$ (one compound per year)
$= 5\%$

4 Write your answer.

The annual interest rate required for this perpetuity investment is 5% per annum.

Perpetuities are just special cases of annuities in withdrawal phase. The payment that is withdrawn just happens to be the same as the interest that is earned after each compounding period. The spreadsheet ‘Annuity Calculator’ can be used to solve problems involving perpetuities.

Spreadsheet **Spreadsheet activity 9D-2:** Annuity Calculator

ISBN 978-1-108-46154-2

Cambridge University Press

Example 12 Using technology to solve a problem involving a perpetuity investment

A university mathematics faculty has \$30 000 to invest. It intends to award an annual mathematics prize of \$1500 to a student using the interest from this investment. If the award is to be made in perpetuity, use technology to find the annual interest rate required for this investment.

Solution

Spreadsheet 9D-2

1 Write down the values of P and d.

$P = 30\,000$

$d = 1500$

2 Enter values into the spreadsheet.

Note: The number of compounds is entered as '1'. The balance of the perpetuity will be the same no matter how many compounding periods are considered.

The future balance is the same as the principal for a perpetuity.

The compounds per year is 1 because the payment is withdrawn after one compounding period.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			number of compounds	1	Calculate
5			annual percentage rate of interest		Calculate
6			compounds per year	1	
7			payment withdrawn	\$1,500.00	Calculate
8			principal	\$30,000.00	Calculate
9			future balance	\$30,000.00	Calculate

3 Click 'Calculate' to find the annual interest rate required.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4			number of compounds	1	Calculate
5			annual percentage rate of interest	5.0000000%	Calculate
6			compounds per year	1	
7			payment withdrawn	\$1,500.00	Calculate
8			principal	\$30,000.00	Calculate
9			future balance	\$30,000.00	Calculate

4 Write your answer.

An interest rate of 5% per annum is required for this perpetuity investment.

ISBN 978-1-108-46154-2

Exercise 9D

SF

Example 8

1 Craig has won $1 000 000 in a lottery and has decided to invest this money in a perpetuity that pays interest at the annual percentage rate of interest of 5.75%, compounding monthly.

a What is the monthly payment that Craig can withdraw from his investment?

b Use technology to verify your answer to **1a** above.

2 Suzie has invested her inheritance of $642 000 in a perpetuity that pays interest at the rate of 6.1% per annum, compounding quarterly.

a What quarterly payment does Suzie receive?

b After five quarterly payments, how much money remains invested in the perpetuity?

3 Geoff would like to establish a perpetuity, the interest from which will be donated to the RSPCA. He would like the annual payment to be $2500. The perpetuity account will pay interest at the rate of 2.5% compounding annually.

a Use a rule to verify that Geoff will need $100 000 for this investment.

b Geoff only has $80 000 to invest. Use a rule to determine the interest rate that Geoff would need to provide the annual payment to the RSPCA.

c Verify your answer to **3b** above using technology.

4 Barbara would like to establish a scholarship that will reward the hardest working mathematics student in Year 12 each year with a $500 prize.

a If the interest on her investment is 2.7% per annum, how much should Barbara invest?

b Barbara has $12 000 to invest in the perpetuity. What annual interest rate does Barbara require in order to pay the prize in perpetuity? Round your answer to two decimal places.

c Use technology to verify your answer to **4b** above.

CF

5 Kathy is a fan of the Brisbane Broncos. Her annual club membership costs $350. Kathy invests some money and intends to use the interest to pay her membership every year in perpetuity. Assume that the membership fee remains the same from year to year.

a If Kathy can invest her money to earn interest at the rate of 3.5% per annum, compounding annually, what principal amount will she need? Round your answer to the nearest cent.

b If Kathy paid her membership fee with quarterly instalments:

i what quarterly instalment would she pay to the club?

ii what quarterly interest rate (compounding quarterly) would be required from her perpetuity investment?

iii calculate the equivalent annual percentage rate of interest for her investment.

c From your answers to **5b** above, what impact does the number of compounds per year have on the interest earned from a perpetuity?

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Key ideas and chapter summary

AS

Annuity

An **annuity** is a type of compound interest investment from which either a regular payment is withdrawn, or into which a regular payment is deposited. The interest of an annuity is calculated on the balance of the investment after each payment is withdrawn or deposited.

Recursive model for an annuity

A recurrence relation can be used to determine the balance of an annuity after n compounding periods.

If A_n is the balance of the annuity after n compounding periods, the decimal rate of interest per compounding period is i and the regular amount that is deposited or withdrawn after each compounding period is d, then the recursive model for an annuity can have one of two forms:

Deposit: $A_0 =$ principal of investment, $A_{n+1} = r \times A_n + d$

Withdrawal: $A_0 =$ principal of investment, $A_{n+1} = r \times A_n - d$

where $r = 1 + i$

Payment schedule for an annuity

A table that summarises the interest calculations for an annuity is called a **payment schedule**. A payment schedule shows the payment number, withdrawal amount, interest paid, principal reduction and investment balance after each payment for some, or all, of the payments of an annuity.

Annuities formula

A formula that can be used to calculate the balance of an annuity after n compounding periods, given the principal (P), decimal rate of interest per compounding period (i), total number of compounding periods (n), and payment amount (M). The annuities formula is:

Deposit: $A = M\dfrac{((1 + i)^n - 1)}{i}$ (for principal $= \$0$)

$$A = P(1 + i)^n + M\frac{((1 + i)^n - 1)}{i}$$

Withdrawal: $A = P(1 + i)^n - M\dfrac{((1 + i)^n - 1)}{i}$

Total interest

An annuity will earn interest in both deposit and withdrawal phases. If I is the total interest earned by the annuity, A_n is the balance of the annuity after n compounding periods and d is the amount of each regular deposit or withdrawal, then:

$I = A_n - n \times d - A_0$ during the deposit phase

$I = A_n + n \times d - A_0$ during the withdrawal phase

ISBN 978-1-108-46154-2

Perpetuity

A perpetuity is a special case of an annuity. The regular payment withdrawn from a perpetuity is equal to the interest earned by the principal. The value of a perpetuity will remain constant.
The future value of a perpetuity is the same as the principal, $A = P$.
Given the decimal rate of interest per compounding period (i), the regular payment amount (d), or the principal (P) the rules that allow the calculation of d, P and i are:

$$d = i \times P \qquad P = \frac{d}{i} \qquad i = \frac{d}{P}$$

Skills check

Having completed this chapter, you should be able to

- use a recursive model to determine the balance of an annuity after n compounding periods
- write a recursive model for an annuity
- construct and interpret a payment schedule for an annuity
- understand the effect of the repayment amount on the duration of an annuity
- analyse annuities with a spreadsheet
- use the annuities formula to determine the balance of an annuity after n compounding periods
- solve annuity problems using technology (spreadsheet calculation tool)
- calculate the payment withdrawn from a perpetuity using a rule
- calculate the principal required by a perpetuity using a rule
- calculate the interest rate required by a perpetuity using a rule
- calculate the interest rate and principal for a perpetuity using technology (spreadsheet calculation tool).

Multiple-choice questions

1 Julie has started a new job. She has a new superannuation account and her employer will deposit $850 each month into this account. Assume that the money in this account will earn interest at the rate of 8.28% per annum, compounding monthly. A recurrence relation model for the balance of this investment after n months, A_n, is

A $A_0 = 0,\ A_{n+1} = 1.0069 \times A_n + 850$
B $A_0 = 0,\ A_{n+1} = 1.069 \times A_n + 850$
C $A_0 = 0,\ A_{n+1} = 1.0828 \times A_n + 850$
D $A_0 = 0,\ A_{n+1} = 1.69 \times A_n + 850$
E $A_0 = 0,\ A_{n+1} = 1.828 \times A_n + 850$

ISBN 978-1-108-46154-2

Use the following information to answer questions 2 and 3.

An annuity is modelled by the recurrence relation shown below.

$$A_0 = 386000, \quad A_{n+1} = 1.0065 \times A_n - 5600$$

where A_n is the balance of the investment after n months.

2 The balance of the investment after four months is:

A \$370 342.77 **B** \$373 514.93 **C** \$376 666.59
D \$379 797.91 **E** \$382 909.00

3 The annual percentage rate of interest for this investment is:

A 1.5% **B** 5.4% **C** 6.5% **D** 7.8% **E** 17.4%

Use the following information to answer questions 4, 5 and 6.

A payment schedule for the first five payments from an annuity is shown below.

Payment number	Payment withdrawn	Interest earned	Principal reduction	Balance of investment
0	0	0	0	285 000.00
1	5500.00	1653.00	3847.00	281 153.00
2	5500.00	1630.69	3869.31	277 283.69
3	5500.00	1608.25	3891.75	273 391.94
4	5500.00	1585.67	3914.33	A
5	5500.00	1562.97	3937.03	265 540.58

4 What is the principal amount for this annuity?

A \$1653 **B** \$3847 **C** \$5500 **D** \$281 153 **E** \$285 000

5 What is the value of A, the balance of the annuity after payment 4?

A \$261 626.25 **B** \$267 891.94 **C** \$269 477.61
D \$269 500.19 **E** \$271 806.27

6 The interest charged on this annuity compounds monthly and monthly payments are received.

The annual percentage rate of interest for this annuity is closest to:

A 1.35% **B** 5.80% **C** 6.96% **D** 16.12% **E** 23.16%

7 Michelle will spend one year travelling the world. She invests \$25 000 into an annuity that earns interest at the rate of 7.08% per annum, compounding monthly. Michelle expects to receive exactly 12 payments from this investment, the first 11 of which are equal in value, before it is fully exhausted.

The monthly payment that Michelle receives is closest to:

A \$1160 **B** \$2160 **C** \$3160 **D** \$4160 **E** \$5160

ISBN 978-1-108-46154-2

8 An annuity of principal \$285 000 earns interest at the rate of 6.36% per annum, compounding monthly. Monthly payments of \$3200 are received from this investment. After how many payments will the balance of this investment first be below \$270 000?

A 6 **B** 7 **C** 8 **D** 9 **E** 10

9 A perpetuity will be set up to provide an annual prize of \$400 to the best mathematics student in a school. Interest will be earned on the principal of the investment at a rate of 3.4% per annum and will be used to pay for the prize every year. The amount that must be invested is closest to:

A \$400 **B** \$800 **C** \$1176 **D** \$11 765 **E** \$136 000

10 A perpetuity has a balance of \$120 000 after six years. Interest is earned at the percentage annual interest rate of 5% per annum. The perpetuity is used to provide an annual prize of value \$6000.

After a further six years, what is the balance of the perpetuity?

A \$84 000.00 **B** \$86 654.33 **C** \$120 000.00
D \$123 563.76 **E** \$168 243.44

Short-answer questions

SF

1 An annuity in withdrawal phase is modelled using the recurrence relation shown below.

$$A_0 = 624000, \ A_{n+1} = 1.0013 \times A_n - 2500$$

In the recurrence relation, A_n is the balance of the investment after n weekly repayments.

a What is the principal of this investment?

b What is the value of the weekly payments withdrawn?

c What is the balance of this loan after six payments have been withdrawn?

2 Carys has \$345 000 to invest in an annuity. Interest will be paid at the annual percentage interest rate of 4.6%, compounding quarterly. Carys will withdraw a payment of \$12 000 per quarter from the investment.

Let A_n be the balance of Carys' annuity after n quarters.

a Construct a recurrence relation model for the balance of Carys' investment after n quarters.

b Apply the recurrence relation to calculate the balance of Carys' investment after six payments have been withdrawn.

c How much interest will have been earned in total after six payments have been received?

ISBN 978-1-108-46154-2

Review

3 The payment schedule for an annuity with monthly payments is shown below.

Payment number	Payment withdrawn	Interest earned	Principal reduction	Balance of investment
0	0	0	0	84 000.00
1	14 500.00	394.80	14 105.20	69 894.80
2	14 500.00	328.51	14 171.49	55 723.31
3	14 500.00	261.90	14 238.10	41 485.21
4	14 500.00	194.98	14 305.02	27 180.19
5	14 500.00	127.75	14 352.25	12 827.94
6	14 500.00	60.29	14 439.71	–1611.77

a What is the principal of this investment?

b What is the value of the monthly payment?

c Calculate the:

i monthly interest rate **ii** annual interest rate

d If this annuity was fully exhausted by the sixth payment, how much will this payment be?

4 Michael's employer has started a superannuation fund for him. Each month, $250 will be placed in this account. If the money in the superannuation fund earns interest at the annual percentage interest rate of 3.89%, compounding monthly, how much money is in the account after five years? Round your answer to the nearest cent.

5 A university would like to establish a perpetuity, the interest from which will fund a scholarship for a talented student each year. The value of this prize should be $3500 per year.

a If the interest on this investment is 3.6% per annum, how much will the university need to invest?

The university only has $60 000 to invest.

b What annual interest rate is required to pay the scholarship in perpetuity? Round your answer to two decimal places.

c Use technology to verify you answer to **5b** above.

ISBN 978-1-108-46154-2

Extended-response questions

1 Mary has just started a new job. She will be paid a salary of \$63 000 per year. Mary receives her salary as 12 equal payments on the first of every month. Mary's employer will pay 8.4% of her monthly salary into a superannuation account each month. The money in this account earns interest at the rate of 1.8% per annum, compounding monthly.

a What is Mary's monthly salary

b How much will Mary's employer pay into the superannuation account each month?

c Use the annuity formula (principal = \$0) to determine the balance of Mary's superannuation account after:

i one full year of work

ii ten years of work

2 Henry has just started a new job. He will be paid a salary of \$63 960 per year. Henry receives his salary as 26 equal fortnightly payments. Every fortnight, Henry's employer will pay 7.8% of his fortnightly salary into a superannuation account. The money in this account earns interest at the rate of 3.9% per annum, compounding fortnightly.

a How much will Henry receive each fortnight?

b How much will Henry's employer pay into the superannuation account each month?

c Use the annuity formula (principal = \$0) to determine the balance of Henry's superannuation account after:

i one year of work

ii five years of work.

Henry will add \$100 of his own money into the superannuation account each fortnight.

d Use the annuity formula (principal = \$0) to determine:

i the balance of the superannuation account after one year with these extra payments

ii how much higher the balance of the superannuation account will be after one year with extra payments compared to the balance with no extra payments?

3 Ludwig inherited \$150 000 from his aunt. He decided to invest this money into an account that pays interest at the rate of 5.76% per annum, compounding monthly.

a If Ludwig's account was a perpetuity, what monthly payment would he receive?

b If Ludwig's account was an annuity and he withdrew \$2000 per month:

i write a recurrence relation to model this annuity

ii how much money would be left in the account after 6 months?

iii how many months would it take for the balance of the investment to first fall below \$130 000?

ISBN 978-1-108-46154-2

Cambridge University Press

Review

c If Ludwig's account was an annuity and he withdrew $4500 per month:

i how many payments of $4500 could he receive?

ii what would be the value of his final repayment if it was smaller than all the others?

4 Jagathi receives monthly payments of $5250 from an annuity that is earning interest at the rate of 5.28% per annum, compounding monthly. The balance of Jagathi's investment is $376 623.14 after three years of investment.

a What is the principal amount of Jagathi's investment?

b How much interest has Jagathi earned after three years of investment?

c How many more payments of $5250 can Jagathi withdraw?

d What final amount can Jagathi withdraw to fully exhaust his annuity?

5 Ebrahim has recently retired and will invest his superannuation payment into an annuity that will earn interest at the rate of 5.04% per annum, compounding monthly. The principal amount of this investment will be $534 000.

a If Ebrahim withdraws a payment of $6000 per month, how many payments in total can he expect to withdraw?

b What is the balance of Ebrahim's investment after four years of withdrawals?

Ebrahim decided to decrease his monthly payment after four years of investment. He will withdraw monthly payments of $4500 until his investment is fully exhausted.

c How many payments of $4500 can he expect?

d What is the final amount that Ebrahim can withdraw to fully exhaust his annuity?

6 Byron has begun a new job and a superannuation account was opened by his employer. Byron will be paid $1400 per fortnight. His employer will deposit 8% of his salary into the superannuation account every fortnight. CU

Assume that the superannuation account earns interest at the annual percentage interest rate of 2.1%.

a Use the annuities formula to determine the balance of Byron's superannuation account after one year of work.

After one year of work, Byron's employer offers him a change of salary conditions. Byron has a choice:

Option 1 – a salary of $1800 per fortnight with no change to superannuation conditions.

Option 2 – a salary of $1600 per fortnight with 8.2% of salary deposited into the superannuation account each fortnight.

ISBN 978-1-108-46154-2

b Which option would result in the greater amount of money deposited in the superannuation account?

Assume that Byron chooses Option 1 and that the annual percentage interest rate of the superannuation account remains constant at 2.1% per annum.

After ten years, Byron changes jobs. He transfers his superannuation balance to a new account that earns interest at the rate of 2.4% per annum, compounding monthly.

Byron's new job pays a monthly salary of $3500 and his new employer will contribute 9.4% of this salary to the superannuation account each month.

c **i** Calculate the balance of Byron's superannuation account after ten more years.

ii What total interest will Byron earn over the twenty-one years he will hold the superannuation account?

7 Jarrod's superannuation account has a balance of $250 000. His employer adds $1000 to this account every month.

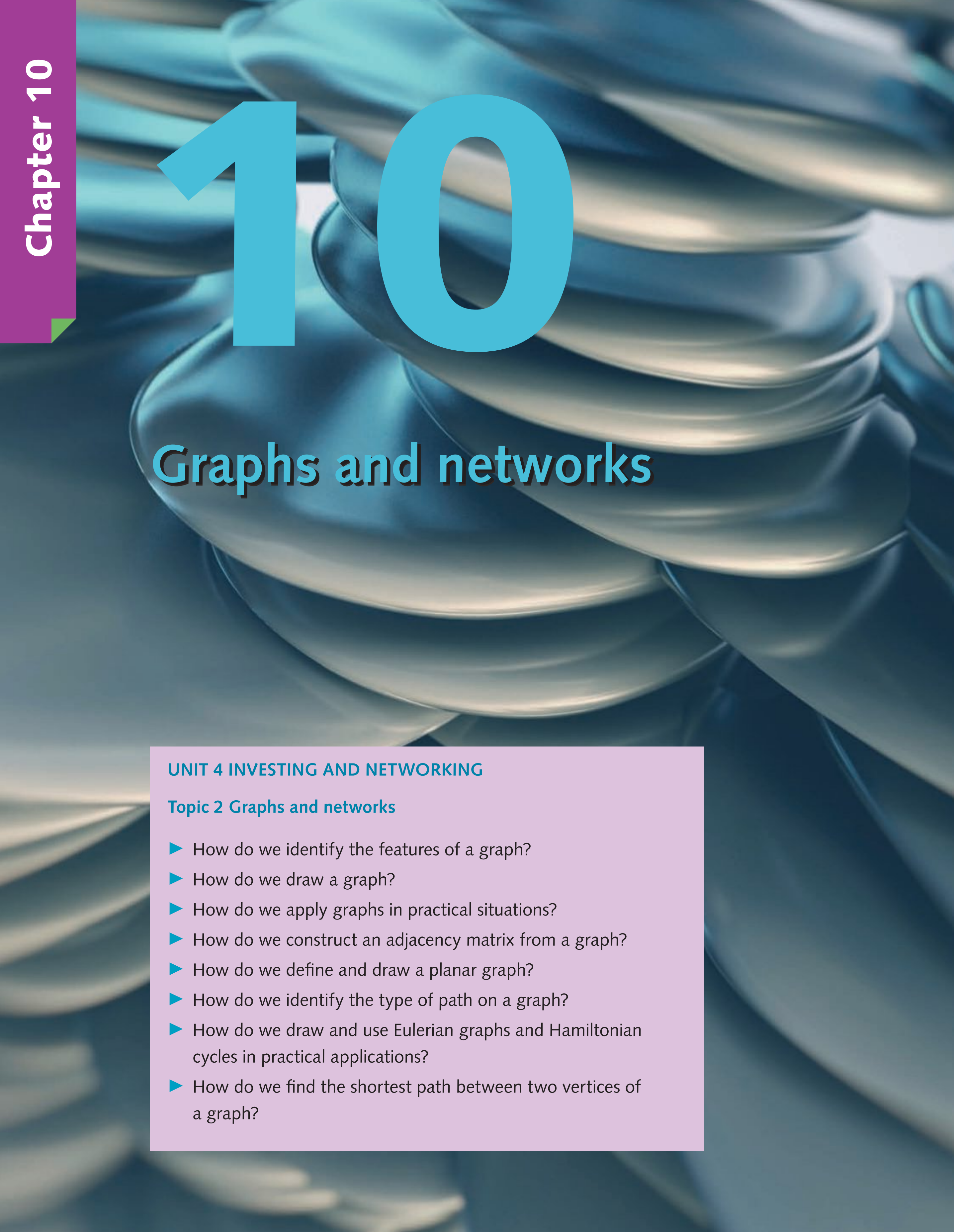
Jarrod's superannuation fund pays him interest at the annual percentage interest rate of 3.36% with monthly compounds.

a If Jarrod intends to retire after a further 3 years, how much money will he receive from his superannuation account?

b Jarrod will invest this money in an annuity that will pay interest at the annual percentage interest rate of 4.08% with quarterly compounds. If he would like this money to last for at least 15 years, what quarterly payment can he withdraw?

 ISBN 978-1-108-46154-2 Cambridge University Press

Chapter 10



10

Graphs and networks

UNIT 4 INVESTING AND NETWORKING

Topic 2 Graphs and networks

- How do we identify the features of a graph?
- How do we draw a graph?
- How do we apply graphs in practical situations?
- How do we construct an adjacency matrix from a graph?
- How do we define and draw a planar graph?
- How do we identify the type of path on a graph?
- How do we draw and use Eulerian graphs and Hamiltonian cycles in practical applications?
- How do we find the shortest path between two vertices of a graph?

ISBN 978-1-108-46154-2

Cambridge University Press

10A Graphs and associated terminology

Representing connections with graphs

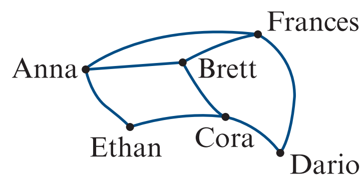
There are many situations in everyday life that involve connections between people or objects. Towns are connected by roads, computers are connected to the internet and people connect to each other through being friends on social media. A diagram that shows these connections is called a **graph**.

Networks: Basic concepts Watch the video in the Interactive Textbook for an illustration of the terms and concepts in action.

Vertices and edges

The graph below represents the connections between friends on a social media website.

There are six people in this graph and each person is represented by a dot called a **vertex**.



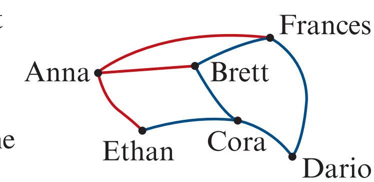
Each vertex in the graph is joined to some of the other *vertices* (plural of vertex) by a line called an **edge**. These lines represent the connection between the people represented by the vertices.

For example, the vertex for Anna is connected by an edge to the vertex for Brett, which means Anna and Brett are connected as friends on the website.

There is no edge between the vertices for Frances and Cora, which means they are not connected as friends on the website.

The degree of a vertex

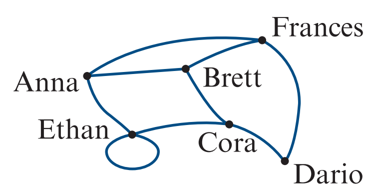
The graph of the social media connections above shows that Anna has three friends, Frances, Brett and Ethan. There are three edges that connect Anna to other people. This number is called the **degree of the vertex** representing Anna. It is the number of times an edge connects to that vertex.



In symbolic form, the degree of the vertex representing Anna can be written as deg(Anna) = 3.

Loops

Imagine that Ethan is able to add himself as a friend on the social media website.



The edge representing this connection would connect the vertex representing Ethan back to itself. This type of edge is called a **loop**.

A loop is attached twice to a vertex and so it will add two to the degree for that vertex. In this graph, deg(Ethan) = 4.

ISBN 978-1-108-46154-2

Cambridge University Press

Representing connections with graphs

- A *graph* consists of vertices joined together by *edges*.
- The number of times an edge connects to a vertex is called the *degree* of that vertex.
- In symbolic form, the degree of vertex V is written as deg (V).
- A *loop* connects a vertex to itself.
- A loop connects twice to a vertex and so it will add two to the degree of that vertex.

Example 1 Drawing a graph to represent connections

Five people, Anthony, Ronnie, Robyn, George and Evan have accounts on a social media website.

- Anthony is a friend of Robyn, Ronnie and Evan.
- Ronnie is a friend of everyone.
- Robyn is a friend of Anthony, Ronnie and Evan.
- George is a friend of Ronnie only.
- Evan is a friend of Anthony, Ronnie and Robyn.

a Draw a graph to represent the connections between the five people above.

b Write down the value of deg(Anthony).

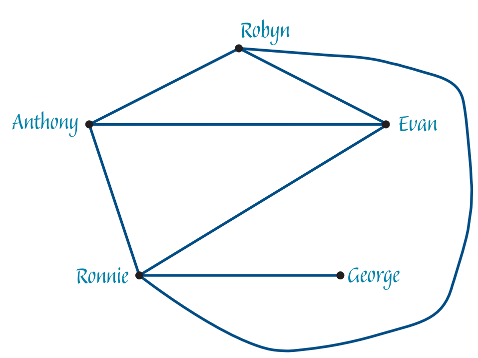
c Which person has the vertex with the:

i smallest degree? **ii** largest degree?

Solution

a The graph must have an edge between:

- Anthony and Robyn
- Anthony and Ronnie
- Anthony and Evan
- Ronnie and Robyn
- Ronnie and George
- Ronnie and Evan
- Robyn and Evan

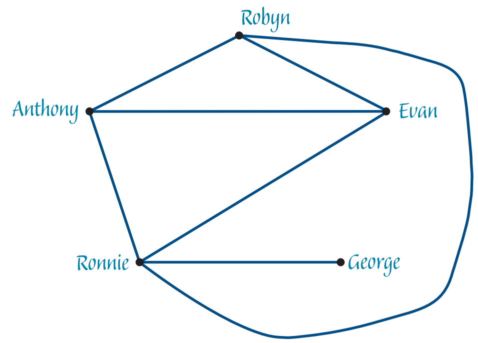


There will be one edge for every pair of people that are friends on the social media website.

Note: The position of the vertices representing the people do not have to be in the same position as they are shown in the diagram. As long as the edges connecting the people are the same, the graph can be drawn with vertices in many different positions.

ISBN 978-1-108-46154-2

Cambridge University Press



b The vertex representing Anthony has three edge connections to it.

deg (Anthony) = 3

c **i** The vertex representing George has only one edge connection to it, less than all the other vertices, so it has the smallest degree.

George has the vertex with the smallest degree.
deg (George) = 1

ii The vertex representing Ronnie has four edge connections to it, more than all the other vertices, so it has the largest degree.

Ronnie has the vertex with the largest degree.
deg (Ronnie) = 4

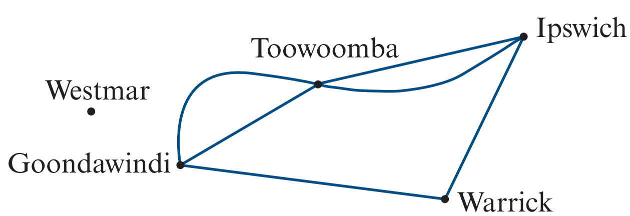
Describing graphs

Graphs that represent connections between objects can take different forms and have different features. This means that there is a variety of ways to describe these graphs.

Multiple edges and isolated vertices

The graph below shows six cities in Queensland represented as vertices and the major highways connecting these towns represented as edges.

There are two different major highways that can be travelled to drive from Ipswich to Toowoomba and the graph shows this using **multiple edges**. Multiple edges connect the same two vertices in a graph.

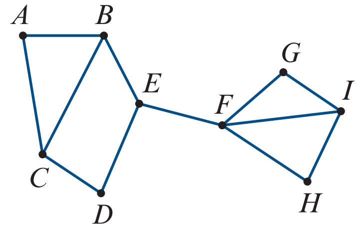


Westmar is not connected to any of the cities by a major highway. In this graph, the vertex representing Westmar is called an **isolated vertex** because it is not connected to any other vertex in the graph. The degree of the vertex representing Westmar is zero.

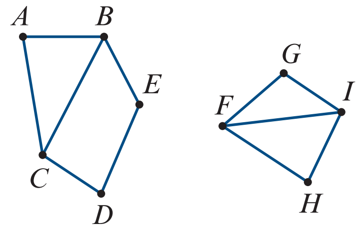
ISBN 978-1-108-46154-2

Cambridge University Press

The graph on the right is called a **connected graph** because all of the vertices are connected in some way into the graph. There are no isolated vertices and no separate parts.



The edge between vertex E and vertex F is called a **bridge**. If this edge was removed, the graph would be left with two separate parts and would no longer be connected.



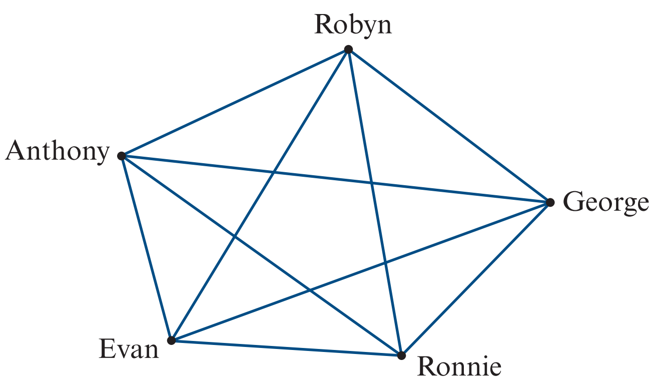
A bridge in a network is much like a bridge over a river. If the bridge collapses, the two sides of the river are disconnected. If we remove an edge that is a bridge in a graph, we are left with two separated parts of the graph.

Simple and complete graphs

Simple graphs do not have any loops and they do not have any multiple edges.

Complete graphs have an edge between every pair of vertices. Every vertex in a complete graph is connected directly by an edge to every other vertex in the graph.

If all five people from Example 1 were friends with each other on the social media website, the graph representing this would be both simple and complete, as shown below.



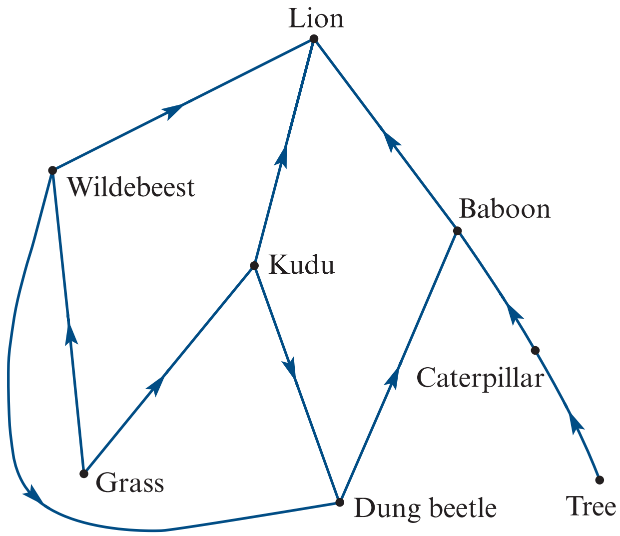
ISBN 978-1-108-46154-2

Cambridge University Press

Directed graphs (digraphs)

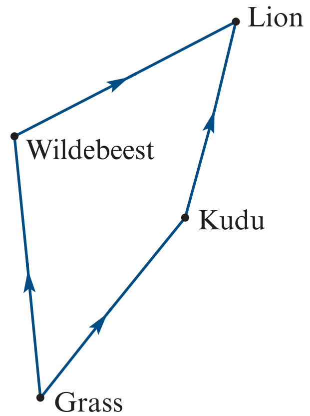
A **directed graph**, or **digraph**, is a graph where there is a direction associated with the edge. An edge of a directed graph is sometimes called an arc or a directed edge.

The directed graph below shows the food connections between some African animals and plants. The arrow on the directed edge (arc) between the vertex representing wildebeest and the vertex representing lion points towards the lion. This means that the lion eats the wildebeest.

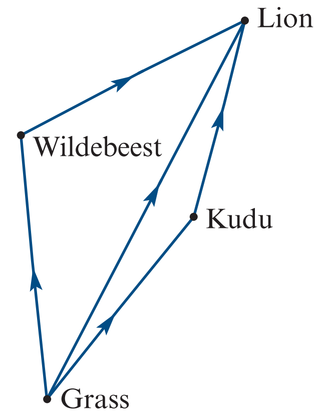


Subgraphs

A **subgraph** is part of a larger graph. All of the edges and vertices in the subgraph must exist in the original graph.



The directed graph above shows only the Wildebeest, Lion, Kudu and Grass. It is a small part of the graph from above. All of the vertices and directed edges (arcs) exist in the original graph and so this is a subgraph of the original food connection graph.



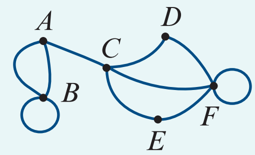
The directed graph on the right contains the same vertices, Wildebeest, Lion, Kudu and Grass, but there is an extra directed edge (arc) that shows the lion also eats grass. This edge (arc) was not in the original graph and so this is not a subgraph of the original food connection graph.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 2 Describing and interpreting graphs

A connected graph is shown on the right.



a Which two vertices have multiple edges between them?

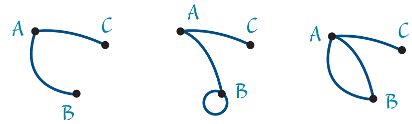
b Draw a subgraph of this graph that only involves vertices A, B and C.

Solution

a Look for two or more edges that connect the same vertices.

Vertex A and B are connected by two edges and so have multiple edges between them.

b There are a few possible answers for this question. Some are shown on the right.



Describing graphs

- A **connected graph** is a graph that has no isolated vertices and no separate parts.
- A **bridge** is an edge that, if removed, would cause the graph to no longer be connected.
- An *isolated vertex* is a vertex in a graph that is not connected to any other vertex by an edge.
- The degree of an isolated vertex is zero.
- *Multiple edges* connect the same two vertices of a graph.
- *Simple graphs* do not have loops and do not have multiple edges.
- *Complete graphs* have an edge between every pair of vertices.
- *Directed graphs* (digraphs) have a directional meaning associated with the edges.
- The edges of a directed graph are called *arcs* or *directed edges*.
- *Subgraphs* are a small section of an existing graph, with no extra vertices and no extra edges.

Investigation 10A: The shaking hands problem

The investigation above showed that each edge of a graph contributes two to the sum of the degrees of the vertices. This can easily be verified by the fact that in order for an edge to exist, it must involve two vertices.

In any graph, the sum of the degrees of the vertices is always twice the number of edges in that graph.

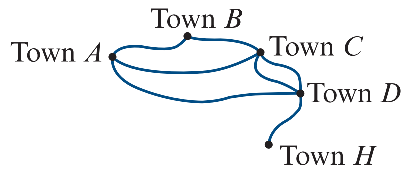
ISBN 978-1-108-46154-2

Exercise 10A

Describing graphs

Example 2

1 The graph below shows five towns, A, B, C, D and H, represented as vertices and the roads between the towns are represented by edges.



a Write down the degree of the vertex representing:

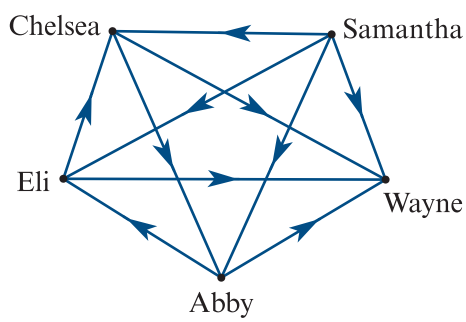
  i Town A
  ii Town B
  iii Town H

b **i** How many edges are in this graph?
  ii Using the result of the shaking hands problem, what will be the sum of the degrees of the vertices in this graph?
  iii Verify your answer to **bii** by adding all the degrees of the vertices.

c Draw a subgraph of this graph that contains only towns H, D and C.

d Which two towns are connected by the bridge in this graph?

2 The directed graph on the right shows the results of a squash tournament. The vertices represent the people involved in the tournament and the arcs represent the games between the two people whose vertices are joined by it. The arrow on the arc points to the winner of the game.



a Who was the winner in the game between Chelsea and Samantha?

b Which player won all of their games?

c Which player lost all their games?

d Write down:

  i deg(Samantha)
  ii deg(Eli)

SF

Drawing graphs

3 Draw a graph that has:

a three vertices, two of which have an odd degree

b four vertices and five edges, one of which is a loop

c six vertices, one of which is isolated, and eight edges

d six vertices, two of which have an odd degree, and which contain at least one subgraph that is a triangle

CF

ISBN 978-1-108-46154-2

Cambridge University Press

Example 1

4 A national park in Africa contains a number of animal species. In the park:

- lions eat impala
- leopards eat impala and warthogs
- warthogs eat lizards
- lizards eat flies
- eagles eat lizards and small birds
- small birds eat lizards and flies

a Draw a directed graph to represent the information above. Use a vertex to represent each animal and an arc to represent the connection between them. The arrow of the arc should point to the animal that eats the other.

b Write down deg(Warthogs).

c Draw a subgraph that contains only lizards, flies, small birds and eagles.

SF

5 Complete the following for a graph.

a Draw a complete graph that has 6 vertices.

b How many edges does this graph have?

ISBN 978-1-108-46154-2

10B The adjacency matrix

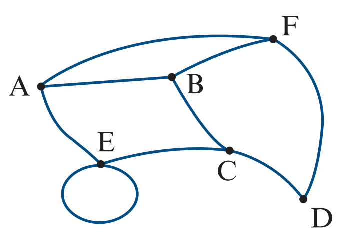
Vertices in a graph that are connected by an edge are said to be *adjacent* to, or next to, each other. A **matrix** can be used to record which vertices are adjacent to each other and also the number of connections between adjacent vertices. This matrix is called an **adjacency matrix**.

Adjacency matrices for simple graphs

Consider the simple graph that shows the social media connections between people from earlier in the chapter. For the purposes of creating an adjacency matrix for this graph, we will represent the vertices with the first letter of the names of the people. For example, Anna will be represented by A, Brett will be represented by B and so on.

The graph and the adjacency matrix for that graph are shown below.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array}\begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$



The adjacency matrix has:

- six rows and six columns, one for each vertex of the graph
- row and column labels that match the vertices of the graph
- a ‘0’ in the intersection of row A and column C because there is no edge connecting A to C
- a ‘1’ in the intersection of row B and column F because there is one edge connecting B to F
- a ‘0’ in the intersection of row D and column D because there is no edge connecting D to D (that is there is no loop at vertex D)
- a ‘1’ in the intersection of row E and column E because there is one edge connecting E to E (that is there is a loop at vertex E).

The number of edges between every other pair of vertices in the graph is recorded in the adjacency matrix in the same way. Notice that the adjacency matrix is symmetric about the main diagonal. This means that the number in row m, column n is the same as the number in row n, column m.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 3 Drawing a graph from an adjacency matrix

Draw the graph that has this adjacency matrix.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

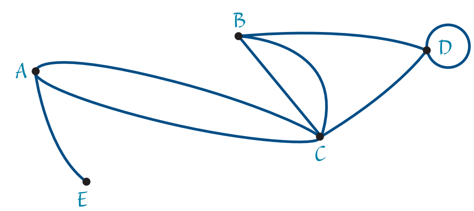
Solution

1 Draw a dot for each vertex and label them A to E.

2 There is a '2' in the intersection of row A and column C. This means that there are two edges connecting vertex A and vertex C. These will be multiple edges.

3 There is a '1' in the intersection of row D and column D. This means that there is a loop at vertex D (vertex D is connected to itself by a loop).

4 Look at every intersection of row and column in the matrix and add edges to the graph, if they do not already exist.



Note: The graph has been drawn so that the edges do not cross over each other. This is not strictly necessary.

Adjacency matrices for directed graphs

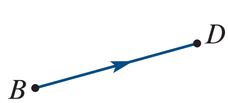
An adjacency matrix can also be drawn to show the directed connections between vertices in a directed graph. The adjacency matrix for a directed graph may be, but is not necessarily, symmetric about the main diagonal.

The matrix labels for the rows of the matrix are the *origin* vertices for the arcs. The column labels in the matrix are the destination vertices. The arrow on the arc will point from the origin vertex (row) to the destination vertex (column).

$$\text{origin}\ \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} - & - & - & - & - \\ - & - & - & 1 & - \\ - & - & - & - & - \\ - & 0 & - & - & - \\ - & - & - & - & - \end{bmatrix} \end{array}$$

In the graph for this adjacency matrix:

- there is one arc from vertex B to vertex D, shown by a '1' in row B, column D
- there is no arc from vertex D back to vertex B shown by a '0' in row D, column B.



This means the arc between vertex B and vertex D has a directional arrow from B to D.

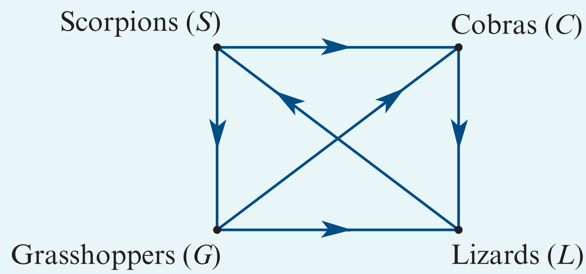
ISBN 978-1-108-46154-2

Cambridge University Press

Example 4 Writing an adjacency matrix from a directed graph

A directed graph is used to represent the winners in a round robin sporting competition. Each of the teams in the competition are represented by a vertex and the arc represents the game between the two teams that it connects. The arrow on the arc points to the winner of the game.

Use the graph on the right to answer the following questions.



a Which team was the winner in the game between the Grasshoppers and the Cobras?

b Which is the only team to beat the Scorpions?

c Write down the adjacency matrix for this graph.

Solution

a Look at the arc that joins the vertices for Grasshoppers and Cobras. It is pointing to the Cobras.

The Cobras won the game against the Grasshoppers.

b Look at all the arcs connected to the vertex for Scorpions. The only one pointing to Scorpions is from Lizards.

The Lizards are the only team to beat the Scorpions.

c In the matrix, use a '1' for an arc that starts at the row team and points towards the column team. Otherwise, use a '0'.

Note: the vertices can be written in any order within the row headings and column headings of the matrix, but it is usual to keep the order consistent between the rows and columns.

$$\begin{array}{c} \\ S \\ G \\ C \\ L \end{array} \begin{array}{c} \begin{array}{cccc} S & G & C & L \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

ISBN 978-1-108-46154-2

Cambridge University Press

Adjacency matrices

An *adjacency matrix* is a square matrix that summarises the connections between vertices of a graph.

For undirected graphs:

- the entry in row m and column n shows the number of edges that join the vertices from this row and column
- the adjacency matrix is symmetric about the main diagonal.

For directed graphs:

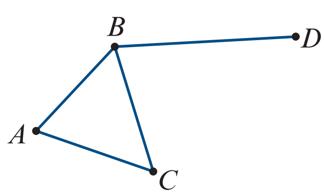
- the entry in row m and column n shows the number of arcs from the vertex in this row to the vertex in this column
- the adjacency matrix may be, but is not necessarily, symmetric about the main diagonal.

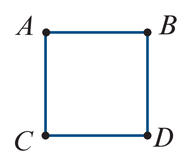
Loops are counted as one edge.

Exercise 10B

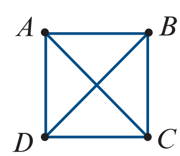
Writing adjacency matrices for graphs

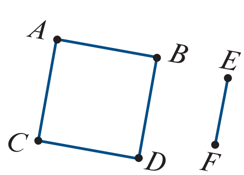
1 For each of the following graphs, write down the adjacency matrix.

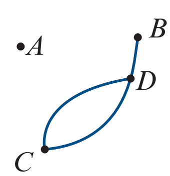
a


b


c


d


e


f


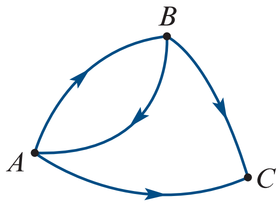
ISBN 978-1-108-46154-2

Cambridge University Press

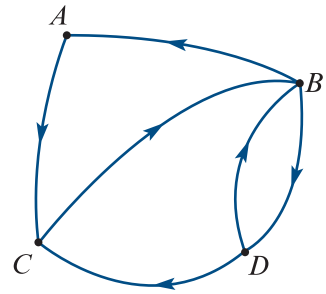
Example 4

2 For each of the following directed graphs, write down the adjacency matrix.

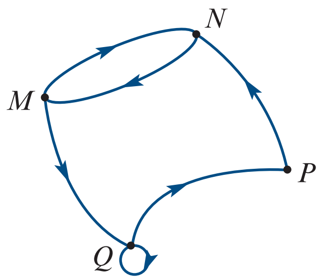
a



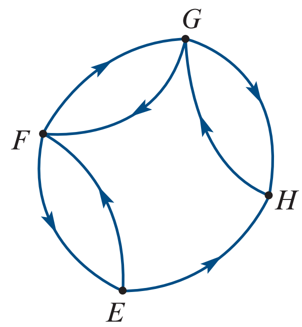
b



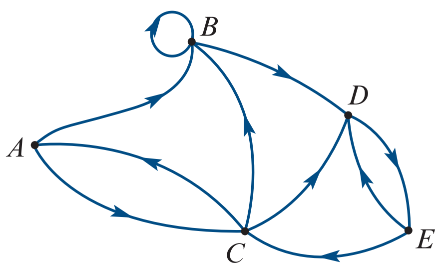
c



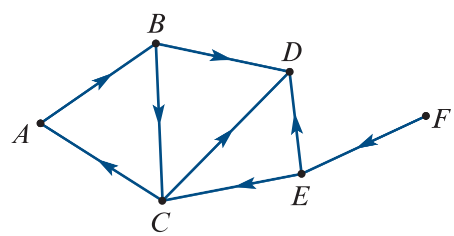
d



e



f



Drawing graphs from adjacency matrices

Example 3

3 Draw a graph for the following adjacency matrices.

a

	A	B	C
A	0	1	1
B	1	0	1
C	1	1	0

b

	A	B	C	D
A	1	0	0	0
B	0	0	1	1
C	0	1	1	0
D	0	1	0	0

c

	A	B	C	D
A	0	1	2	1
B	1	0	1	1
C	2	1	0	0
D	1	1	0	0

 ISBN 978-1-108-46154-2

4 Draw a directed graph for the following adjacency matrices.

a
$$\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{array}{c} \begin{matrix} A & B & C \end{matrix} \\ \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

b
$$\begin{array}{c} \\ S \\ T \\ U \end{array}\begin{array}{c} \begin{matrix} S & T & U \end{matrix} \\ \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

c
$$\begin{array}{c} \\ P \\ Q \\ R \\ S \end{array}\begin{array}{c} \begin{matrix} P & Q & R & S \end{matrix} \\ \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

d
$$\begin{array}{c} \\ P \\ Q \\ R \\ S \end{array}\begin{array}{c} \begin{matrix} P & Q & R & S \end{matrix} \\ \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

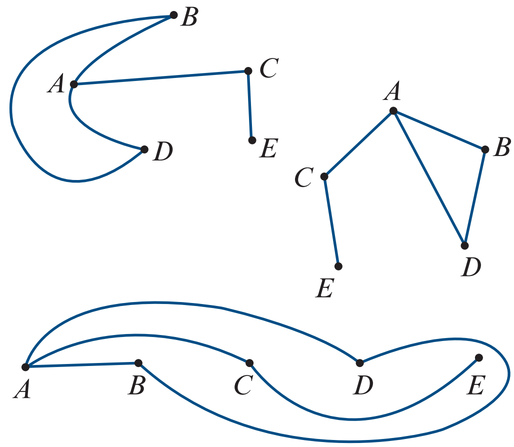
ISBN 978-1-108-46154-2

10C Planar graphs and Euler's formula

▶ Planar graphs

Equivalent graphs

All of the graphs shown in the diagram below contain the same information. For example, the edge between vertex E and C exists in all three of the graphs.



The physical location of the vertices and edges in the diagram is unimportant. As long as the connection information is represented accurately, the graph can be drawn with the vertices in any location.

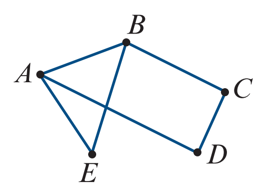
The first of the graphs has some curved edges and the second has all straight edges. The third has the vertices arranged in a straight line.

All of them, regardless of how they are drawn, contain exactly the same connections between vertices and so these graphs are considered to be *equivalent* to each other.

Equivalent graphs contain identical information and are sometimes called **isomorphic graphs**.

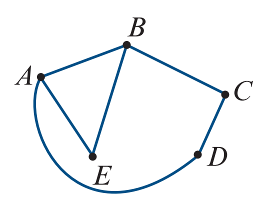
Planar graphs

The graph opposite has two edges that cross over each other (EB and AD). It is helpful to think of edges that cross like this as insulated electrical wires. It is quite safe to cross two insulated wires because the wires themselves never touch and never interfere with each other. We can think of crossing edges in a graph in a similar way.



The edges that cross over in this diagram are similar in that they do not intersect. It is important to note that there is no vertex at the point where these edges cross over.

Graphs with edges that cross in this way *may* be able to be redrawn so that the edges no longer cross. In this diagram the edge between vertices A and D has been moved, but none of the information in the graph has changed.



A **planar graph** is a graph that can be drawn in such a way that the edges it contains do not cross over each other. If a graph is drawn so that no edges cross over, then it is said to be drawn in *planar form*.

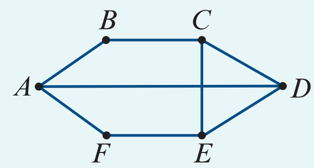
ISBN 978-1-108-46154-2

Cambridge University Press

If it is impossible to draw an equivalent graph without crossing edges, then that graph is called a non-planar graph. It is impossible to draw a non-planar graph in planar form.

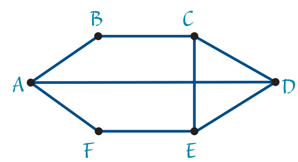
Example 5 Redrawing a graph in planar form

Show that this graph is planar by redrawing it so that no edges cross.

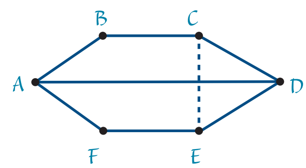


Solution

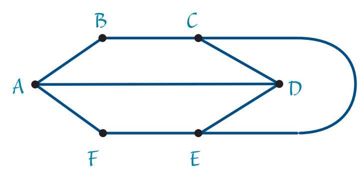
1 Choose one of the edges that crosses over another edge. For example, choose the edge between vertex C and vertex E. Alternatively, the edge between vertex A and vertex D could be chosen.



2 Remove this edge temporarily from the graph.



3 Redraw the edge between the same two vertices (C and E), but without the edge crossing any other edge.

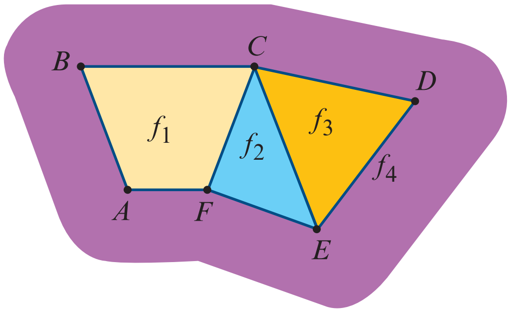


Euler's formula

Leonard Euler (pronounced 'oiler') was one of the most prolific mathematicians of all time. He contributed to many areas of mathematics and his proof of the rule named after him is considered to be the beginning of the branch of mathematics called topology.

Faces

A planar graph defines separate regions of the paper it is drawn on. These regions, called **faces**, could be coloured in as you can see in the diagram shown here. There are three faces inside the graph, one coloured cream, one blue and one gold, but there is also a fourth face coloured purple that totally surrounds the graph.



ISBN 978-1-108-46154-2

Cambridge University Press

In the diagram, the faces are labelled f_1, f_2, f_3 and f_4.

A face is an area in a graph that can only be reached by crossing an edge.

Euler's formula

For any connected planar graph, we can count the number of vertices (v), the number of faces (f) and the number of edges (e). There is a relationship between these numbers, called **Euler's formula**.

In words: 'the number of vertices + the number of faces – the number of edges = 2'

In symbols: $v + f - e = 2$

Euler's formula

For any connected planar graph:

$$v + f - e = 2$$

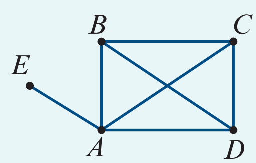
where v is the number of vertices, f is the number of faces and e is the number of edges in the graph.

Example 6 Verifying Euler's formula

For the graph shown on the right:

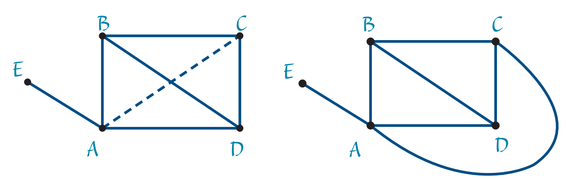
a redraw the graph in planar form

b verify Euler's formula



Solution

a Temporarily remove an edge that crosses another edge and redraw it so that it does not cross another edge.



b **1** Count the number of vertices, faces and edges in the graph.

In this graph, there are: five vertices, four faces and seven edges.

2 Substitute into Euler's formula to verify.

$v + f - e = 5 + 4 - 7 = 2$

Euler's formula is verified.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 7 Using Euler's formula

A connected planar graph has six vertices and nine edges. How many faces does this graph have?

Solution

1. Write down the known values. $v = 6$ and $e = 9$
2. Substitute into Euler's formula and solve for the unknown value.

$$v + f - e = 2$$
$$6 + f - 9 = 2$$
$$f - 3 = 2$$
$$f = 2 + 3$$
$$f = 5$$

3. Write your answer. This graph has five faces.

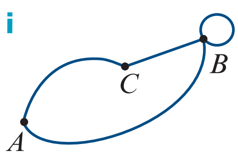
Exercise 10C

Equivalent (isomorphic) graphs

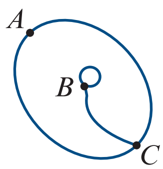
SF

1 In each question below, three of the graphs are equivalent (isomorphic) and the fourth is not. Identify the graph which is not equivalent (isomorphic) to the others.

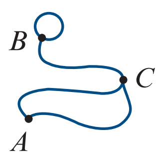
a i



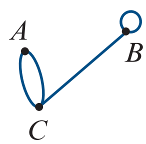
ii



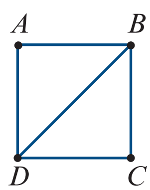
iii



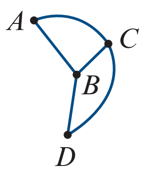
iv



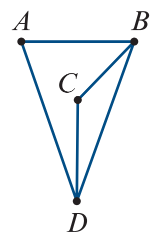
b i



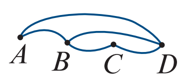
ii



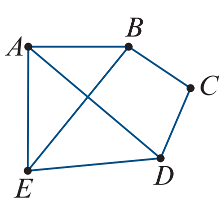
iii



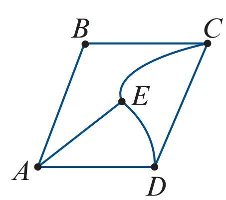
iv



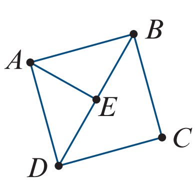
c i



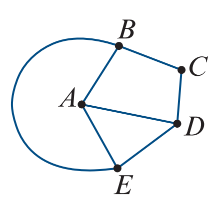
ii



iii



iv

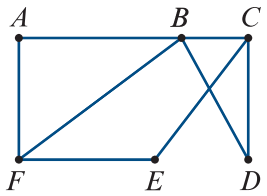


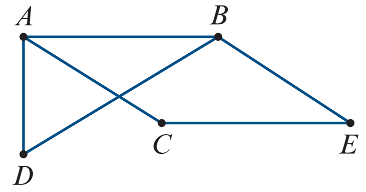
ISBN 978-1-108-46154-2

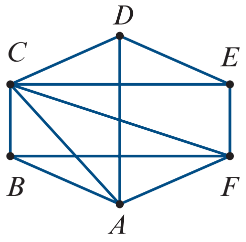
Drawing graphs in planar form

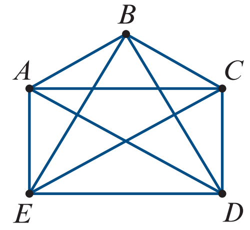
Example 5

2 Where possible, show that the following graphs are planar by redrawing them in a suitable planar form.

a 

b 

c 

d 

CF

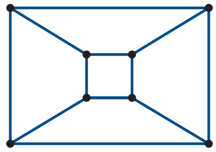
Euler's formula

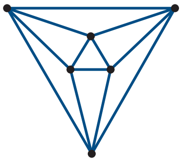
Example 6

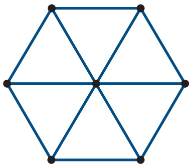
3 For each of the following graphs:

i state the values of v, e and f

ii verify Euler's formula

a 

b 

c 

SF

Example 7

4 For a planar connected graph, find:

a f, if $v = 8$ and $e = 10$

b v, if $e = 14$ and $f = 4$

c f, if $v = 5$ and $e = 14$

d e, if $v = 10$ and $f = 11$

ISBN 978-1-108-46154-2

Cambridge University Press

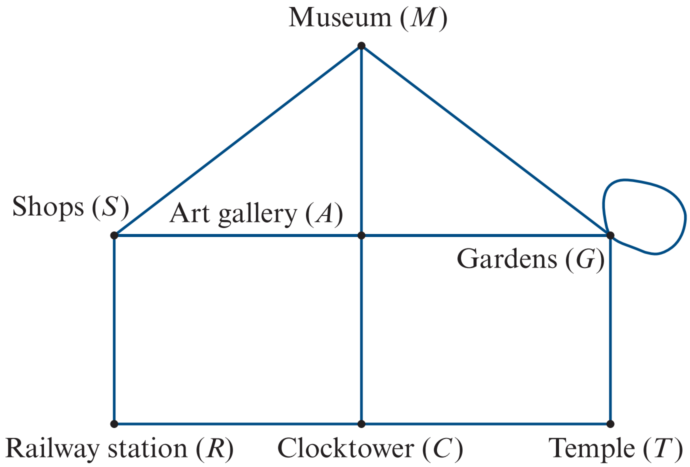
10D Exploring a graph

▶ The language of exploring a graph

Graphs can be used to model and analyse problems involving exploring or travelling. These problems can include minimising the distance travelled or minimising the time taken to travel between different locations using different routes. For example, a courier driver would like to know the shortest route to use for deliveries to minimise the total distance that is driven.

To solve these types of problems, you will need to learn the language used to describe the different ways of navigating through a graph, or *exploring the graph*, from one vertex to another.

The graph on the right shows six tourist locations in a city represented by vertices. The edges of the graph represent the roads between each of the locations.



This graph will be used to help explain the language and define the terms used to describe *exploring a graph*.

Travelling through a network: Watch the video in the Interactive Textbook to see the five types of routes that can be travelled through networks.

ISBN 978-1-108-46154-2

Cambridge University Press

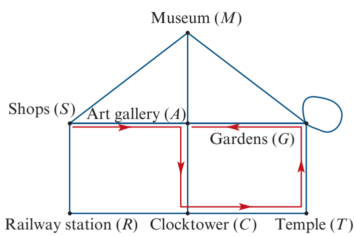
Walks

On one tour, a tour guide meets a group of tourists at the Museum. The group will visit the Art gallery, the Gardens and will finish at the Temple. The route that they take can be written as a list of locations visited in order; that is, $M - A - G - T$.

A route through a graph, from one vertex to another along edges of the graph, is called a **walk**.

The following routes are all walks:

$S - A - C - T - G - A$

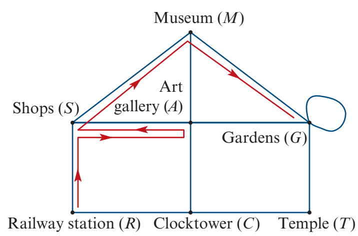


This walk starts at the Shops (S) and ends at the Art gallery (A).

It starts and ends at different vertices and is called an *open walk*.

Notice that the tour group will visit the Art gallery twice.

$R - S - A - S - M - G$

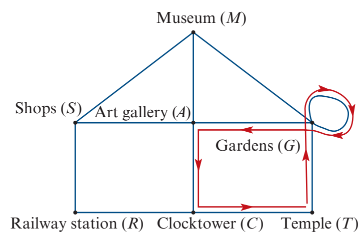


This walk starts at the Railway station (R) and ends at the Gardens (G).

This route is also an *open walk* because it starts and ends at different vertices.

Notice that the tour group will traverse the road between the Shops (S) and Art gallery (G) twice.

$T - G - G - A - C - T$



This walk starts and ends at the Temple (T).

It starts and ends at the same vertex and is called a *closed walk*.

The route traverses the loop at the Gardens and so the tour group will visit the Gardens vertex (G) twice.

ISBN 978-1-108-46154-2

Cambridge University Press

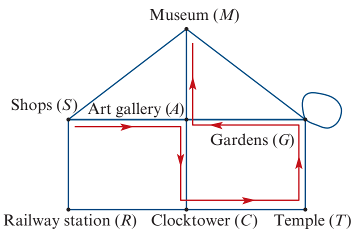
Trails and paths

Trails and paths are special types of walks:

- A **trail** is a walk that has no repeated edges but may contain repeated vertices.
- A **path** is a walk that has no repeated edges and no repeated vertices.

The following routes are all walks but may be trails, paths or neither.

$S - A - C - T - G - A - M$



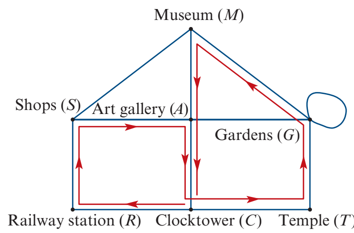
This *walk* starts at the Shops (S) and ends at the Museum.

It starts and ends at different vertices and so it is an *open walk*.

There are no repeated edges and so this is a *trail*.

There are repeated vertices (A) and so this is not a path.

$C - R - S - A - C - T - G - M - A - C$

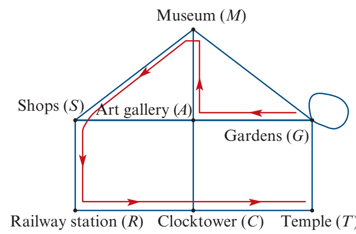


This *walk* starts and ends at the Clocktower (C).

It starts and ends at the same vertex and so it is a *closed walk*.

There are some repeated edges ($A - C$) and so this is neither a trail nor path.

$G - A - M - S - R - C - T$



This *walk* starts at the Gardens (G) and ends at the Temple (T).

It starts and ends at different vertices and so it is an *open walk*.

There are no repeated edges and so this is a *trail*.

There are no repeated vertices and so this is also a *path*.

Walks, trails and paths

A *walk* is a route through a graph, from one vertex to another, along the edges of the graph:

- an open walk starts and ends at different vertices
- a closed walk starts and ends at the same vertex.

A *trail* is a walk that has no repeated edges. It may have repeated vertices.

A *path* is a walk that has no repeated edges and no repeated vertices.

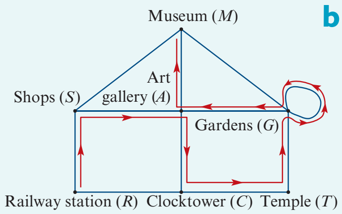
ISBN 978-1-108-46154-2

Cambridge University Press

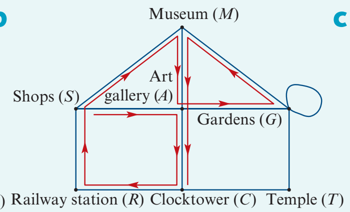
Example 8 Describing walks

Identify the walk shown in each of the graphs below as a trail, a path or a walk only.

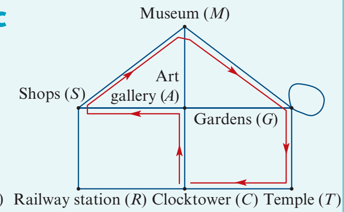
a



b



c



Solution

a This walk has repeated vertices (A and G) but no repeated edge. — *This walk is a trail.*

b This walk has repeated vertices (S, M, A, C) and some repeated edges ($M - A$, $A - C$). — *This walk is a walk only.*

c This walk has no repeated vertices and no repeated edges. — *This walk is a path.*

Closed trails and cycles

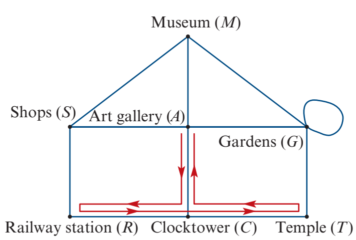
If a walk starts and ends at the same vertex, then it is called a closed walk.

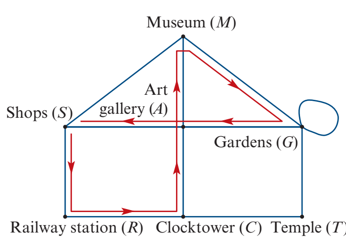
A closed walk will also be a *closed trail* if it has no repeated edges.

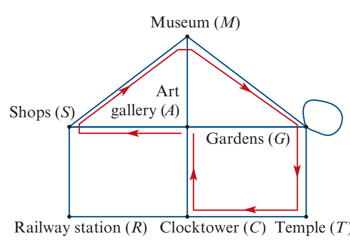
A closed walk will also be a **cycle** if it has no repeated edges and no repeated vertices, except for the starting and ending vertex. Cycles are sometimes called *closed paths*.

The following routes are all *closed walks* but may also be *closed trails* or *cycles*.

$A - C - R - C - T - C - A$ $\quad$ $S - R - C - A - M - G - A - S$ $\quad$ $A - S - M - G - T - C - A$







This walk is a *closed walk* because it starts and ends at the same vertex (A).

There are repeated edges so it is not a closed trail.

There are repeated vertices and repeated edges so it is not a cycle.

This is a *closed walk* because it starts and ends at the same vertex (S).

There is no repeated edge so it is also a *closed trail*.

There are repeated vertices so it is not a cycle.

This is a *closed walk* because it starts and ends at the same vertex (A).

There is no repeated edge so it is also a *closed trail*.

There are no repeated vertices (except for the start and end vertex) and so it is also a *cycle*.

ISBN 978-1-108-46154-2

Cambridge University Press

Closed walks, closed trails and cycles

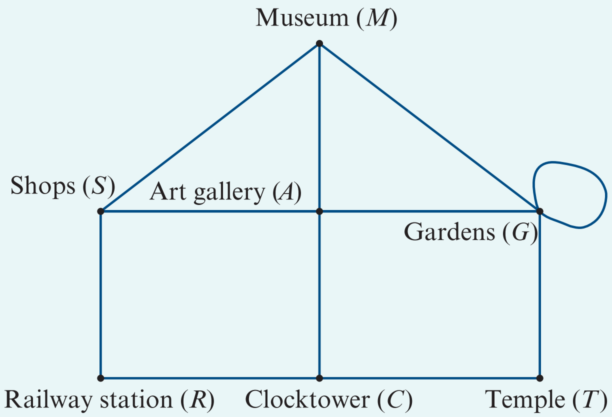
A *closed walk* is a walk that begins and ends at the same vertex.

A *closed trail* is a closed walk that has no repeated edges.

A *closed path* is a closed walk that has no repeated edges and no repeated vertex, except for the starting and ending vertex. A *closed path* can also be called a *cycle*.

Example 9 Describing closed walks

Use the graph of six tourist locations in a city shown on the right to describe the following walks as either a trail, path, closed trail, cycle, open walk only or closed walk only.



a $M - G - G - A - S - R$

b $S - R - C - A - G - T - C - A$

c $C - M - S - R - C$

d $G - G - T - C - A - M - G$

Solution

a This walk has repeated vertices (G). It starts and ends at different vertices. — *This walk is a trail.*

b This walk has repeated vertices (C) and repeated edges (C–A). It starts and ends at different vertices. — *This walk is an open walk only.*

c This walk has no repeated vertices (except for the start and end vertex) and no repeated edge. — *This walk is a closed path or cycle.*

d This walk has repeated vertices (G). It starts and ends at the same vertex. — *This walk is a closed trail.*

ISBN 978-1-108-46154-2

Exercise 10D

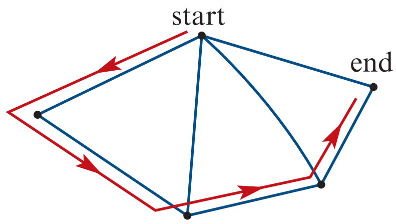
Describing walks through a graph

Example 8

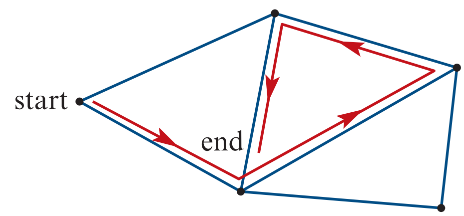
1 Describe the walk shown in each of the following graphs as a trail, path, closed trail, cycle, open walk only or closed walk only.

SF

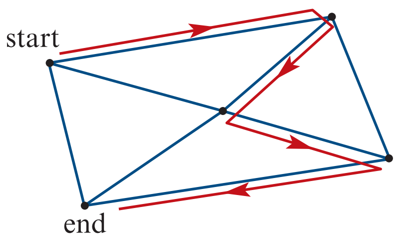
a



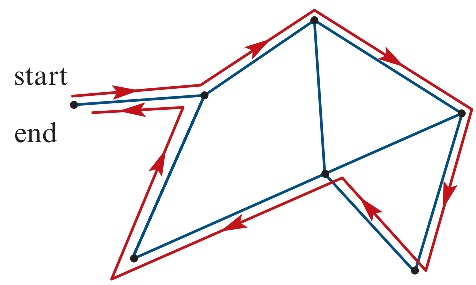
b



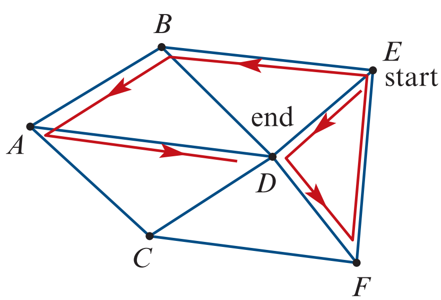
c



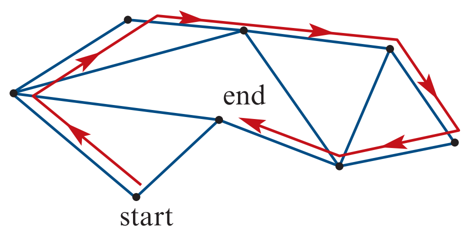
d



e



f



Example 9

2 Use the graph on the right to describe the walks below as a trail, path, closed path, closed trail, cycle, open walk only or closed walk only.

a $B-E-D-C-A-B$

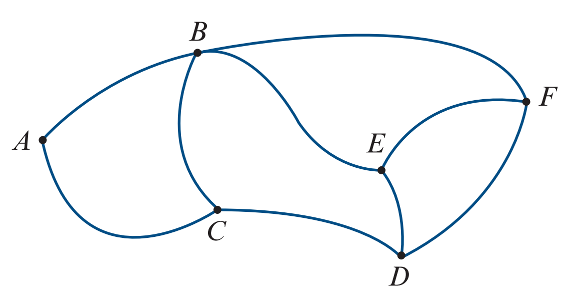
b $A-B-E-B-C-D-E$

c $F-E-B-F-E-D$

d $A-B-C-D-E-B-F$

e $C-A-B-C-D-E-B-C$

f $D-E-B-A-C-D$



ISBN 978-1-108-46154-2

Cambridge University Press

10E Eulerian graphs and applications

The explorer problem

Consider a national park that contains a number of campsites and some roads that lead between them. A park ranger might need to travel along each of these roads in order to check their condition. It is in the best interests of the park ranger to minimise the number of times each road is travelled and to avoid backtracking wherever possible. Ideally, each road should be travelled only once.

This situation is an example of the *explorer problem*. The explorer problem asks the question 'Is it possible to travel every edge in a network only once?'

Investigation 10E: The explorer problem

Whether or not the explorer problem has a solution for a particular graph depends upon the degrees of the vertices in that graph.

Eulerian graphs and Eulerian trails

In Investigation 10E above you may have noticed that if the degree of every vertex in the graph was even, then the ranger could begin at any campsite, travel along every road only once, and would end up back at the starting campsite. This is a *closed trail* that involves every edge of the graph and can be called an **Eulerian trail**.

A graph that contains an *Eulerian trail* is called an *Eulerian graph*. A graph is Eulerian if all of the vertices in that graph have even degrees.

Semi-Eulerian graphs

If there were two odd vertices in the graphs from Investigation 10E above, the ranger could start at either one of these campsites, travel along every road only once, and would end up at the other odd-degree campsite. This is an *open trail* that involves every edge of the graph.

A graph where such an open trail exists is called a *semi-Eulerian graph*. A graph is semi-Eulerian if it contains exactly two vertices that have odd degrees and all other vertices with even degrees.

Eulerian trails and circuits: Watch the video in the Interactive Textbook to see them in action.

ISBN 978-1-108-46154-2

Eulerian and semi-Eulerian graphs

Eulerian graphs:

- contain an *Eulerian trail* (a closed trail that involves every edge of the graph)
- have all vertices with even degree
- allow explorers to begin at any vertex, travel every edge only once and return to the starting vertex.

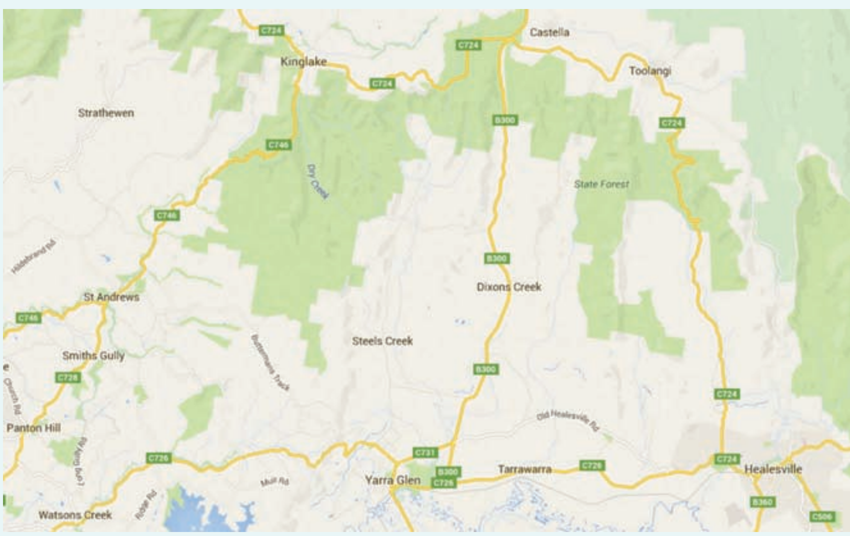
Semi-Eulerian graphs:

- contain an open trail that involves every edge of the graph
- have two vertices with odd degree
- allow explorers to begin at one of the odd-degree vertices, travel every edge only once and end at the other odd-degree vertex.

Eulerian and semi-Eulerian graphs must be connected.

Example 10 The explorer problem – Eulerian graphs

A map showing the towns of St Andrews, Kinglake, Yarra Glen, Toolangi and Healesville is shown below.



a Draw a graph with a vertex representing each of these towns and with edges representing the direct road connections between towns. Ignore any towns on the map not listed in the question.

b Explain why this graph is semi-Eulerian and not Eulerian.

c **i** Write down an open trail that begins at Toolangi and follows every edge only once.

ii Explain how you could tell this trail would end at Kinglake.

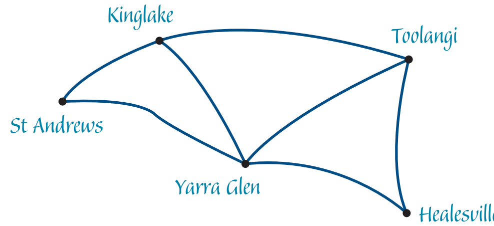
ISBN 978-1-108-46154-2

Cambridge University Press

Solution

a A road connection exists between:

- St Andrews and Kinglake
- St Andrews and Yarra Glen
- Kinglake and Yarra Glen
- Kinglake and Toolangi
- Yarra Glen and Toolangi
- Yarra Glen and Healesville
- Healesville and Toolangi



b The graph has exactly two vertices with odd degrees (Kinglake and Toolangi).

Graphs that have exactly two vertices with odd degrees are semi-Eulerian.

c **i** This question has many answers, one of which is shown.

Toolangi – Healesville – Yarra Glen – Toolangi – Kinglake – Yarra Glen – St Andrews – Kinglake

ii Toolangi was one of the two vertices with an odd degree.

An open trail through a semi-Eulerian graph will start at one of the vertices with odd degree and end at the other. Toolangi was one of the vertices with odd degree and Kinglake was the other, so the open trail must end at Kinglake.

Exercise 10E

Eulerian and semi-Eulerian graphs

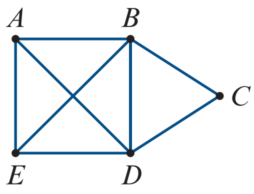
Example 10

1 For each of the graphs shown below:

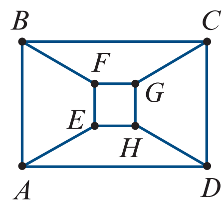
i identify whether the graph is Eulerian, semi-Eulerian or neither

ii name any open or closed Eulerian trails found

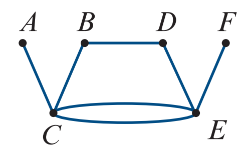
a



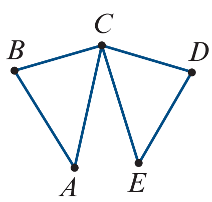
b



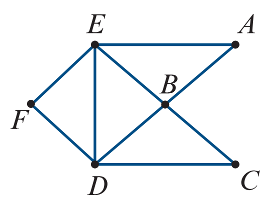
c



d



e



SF

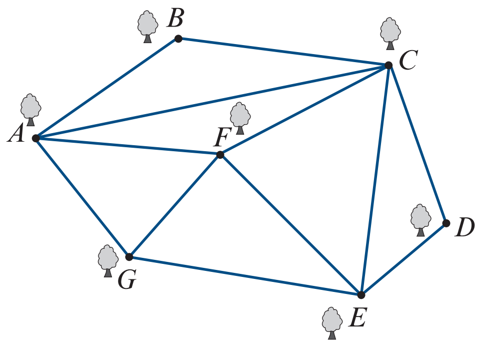
ISBN 978-1-108-46154-2

Cambridge University Press

CF

2 A housing estate has large open parklands that contain seven large trees.

The trees are denoted as vertices A to G on the graph below.



Walking tracks link the trees and are shown as edges on the graph.

a Determine the degree of each of the vertices in the graph.

b One day, Jamie decided to go for a walk that will take him along each path only once at most.

  i At which vertices could Jamie start?

  ii At which vertex will Jamie end?

c A new track is to be made between two trees. This track will mean that Jamie could start at any vertex, walk along each of the tracks only once and return to his starting point.

  i Between which two vertices should the new track be made?

  ii What is the name of the walk that Jamie will follow once the new track is made?

ISBN 978-1-108-46154-2

Cambridge University Press

10F Hamiltonian graphs and applications

▶ The traveller problem

You have seen an application of graphs that involved tourists visiting a number of sites of interest. A traveller in such a situation would be much more interested in seeing the sites at the vertices of the graph, than the roads that connect them. *The traveller problem* involves a situation where every vertex in a graph is visited once.

Unlike the explorer problem, the traveller problem does not have a set of rules to define whether it has a solution. Rather, finding a route that allows a traveller to visit all vertices in a graph relies purely on inspection.

Investigation 10F: The traveller problem

Hamiltonian graphs

In 1857, Hamilton's puzzle in Investigation 10F above was commercially produced as a board game called 'The Icosian Game'. The object of the game was to find a route through the graph that:

- starts and ends at the same vertex
- visits all vertices exactly once (except for the start and end vertex).

Copyright (c) 2019 The Puzzle Museum – J. Dalgety

The game consisted of a wooden board with holes at vertices. Numbered ivory plugs were put into the board to mark out the route. There are only four of these games known to still exist.

The solution to Hamilton's puzzle is a path because it has no multiple edges and no multiple vertices. The solution is also a cycle, because it starts and ends at the same vertex. Because of this puzzle, any path through any graph that visits every vertex of the graph only once is called a **Hamiltonian path**.

A **Hamiltonian graph** has a cycle (closed path) that visits every vertex only once. This cycle is called a **Hamiltonian cycle**.

A *semi-Hamiltonian graph* has an open path that visits every vertex only once. This path is called a *Hamiltonian path*.

Hamiltonian and semi-Hamiltonian graphs

Hamiltonian graphs:

- contain a cycle (closed path) called a *Hamiltonian cycle* that involves every vertex of the graph
- allow travellers to begin at any vertex, visit every vertex only once and return to the starting vertex.

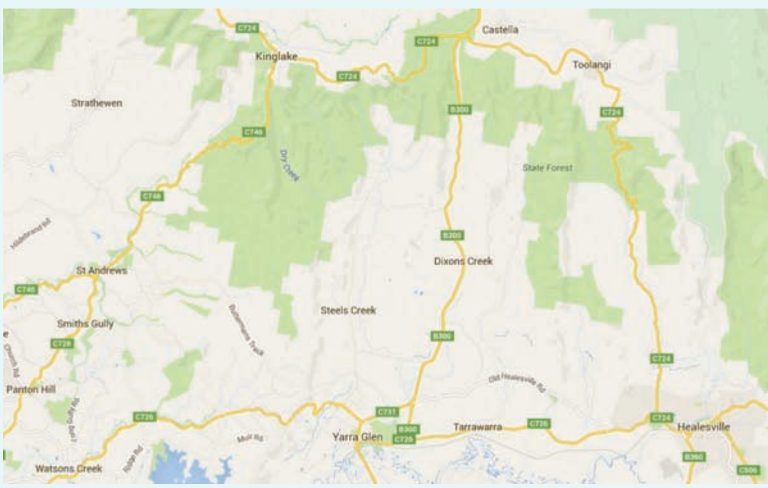
Semi-Hamiltonian graphs:

- contain an open path called a *Hamiltonian path* that involves every vertex of the graph
- allow travellers to begin at one vertex and to visit every other vertex only once but not return to the starting vertex.

Hamiltonian and semi-Hamiltonian graphs must be connected.

Example 11 Solving a traveller problem

A map showing the towns of St Andrews, Kinglake, Yarra Glen, Toolangi and Healesville is shown below.



a Draw a graph with a vertex representing each of the towns and edges representing the direct road connections between the towns. Ignore any towns on the map not listed in the question.

b Explain why this graph is Hamiltonian.

c Write down one:

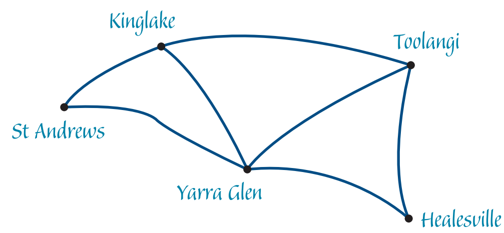
- **i** Hamiltonian path
- **ii** Hamiltonian cycle that begins at Healesville

ISBN 978-1-108-46154-2

Solution

a A road connection exists between:

- St Andrews and Kinglake
- St Andrews and Yarra Glen
- Kinglake and Yarra Glen
- Kinglake and Toolangi
- Yarra Glen and Toolangi
- Yarra Glen and Healesville
- Healesville and Toolangi



b Can a closed walk be found that visits every town?

It is possible to visit every town exactly once and return to the starting town, so the graph is Hamiltonian.

c **i** There are many solutions to this question, one of which is shown.

One Hamiltonian path is:
St Andrews – Kinglake – Yarra Glen – Toolangi – Healesville.

ii There are two solutions to this question, one of which is shown.

A Hamiltonian cycle that begins at Healesville is:
Healesville – Yarra Glen – St Andrews – Kinglake – Toolangi – Healesville.

Exercise 10F

Hamiltonian and semi-Hamiltonian graphs

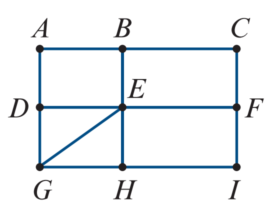
Example 11

1 For each of the graphs below:

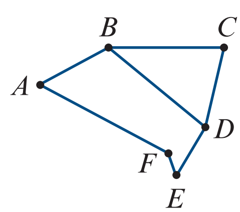
i List one Hamiltonian path starting at vertex A.

ii List one Hamiltonian cycle starting at vertex E.

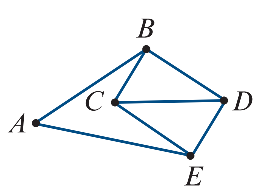
a



b



c

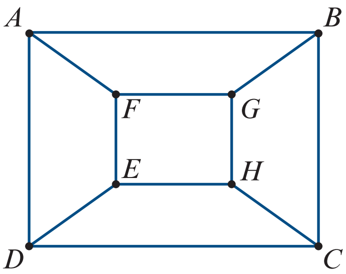


SF

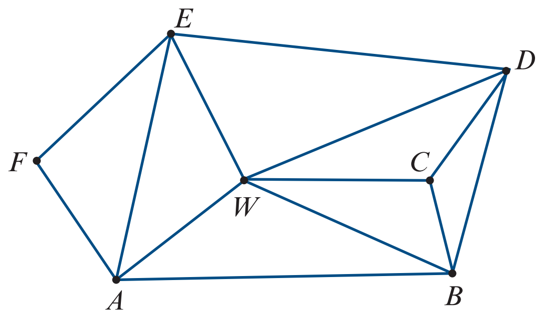
ISBN 978-1-108-46154-2

Cambridge University Press

2 List a Hamiltonian path for this graph, starting at vertex F and finishing at vertex G.



3 The network diagram below shows the location of a warehouse at vertex W.



This warehouse supplies equipment to six factories A, B, C, D, E and F.

a What is the degree of vertex W?

b A salesperson plans to leave factory E, first visit the warehouse, W, and then visit every other factory. They will visit each location only once and will not return to factory E.

i Write down the mathematical term used to describe the planned route.

ii Write down an order in which the salesperson can visit the factories.

based on VCAA (2005)

CF

ISBN 978-1-108-46154-2

Cambridge University Press

10G Weighted graphs, networks and shortest path problems

Weighted graphs

The edges of a graph represent the connections between the vertices of that graph. Sometimes there is more information known about that connection. For example, if the edge of a graph represents a road between two towns, the length of the road, or perhaps the time it takes to travel that road, might be known.

Extra numerical information about the edge can be written next to the edge in a graph. Graphs that have numerical information on each edge are called **weighted graphs** and the numbers themselves are called the *weights* of the edges. Weighted graphs in which the weights are physical quantities, such as distance, time or cost are called **networks**.

The *total weight* of a walk through a network is the sum of all the weights for the edges that are travelled in that walk.

The travelling salesperson problem

The travelling salesperson problem is solved using a weighted graph.

Investigation 10G: The travelling salesperson problem

The solution to the travelling salesperson problem for a particular network is the Hamiltonian cycle in the network that had the smallest total weight. While there are some algorithms (mathematical procedures) that can be used to find this cycle, you will solve travelling salesperson problems, and shortest path problems, using observation and trial-and-error only.

Shortest path problems

Shortest path problems involve finding the shortest path from one vertex to another. This path does not have to be Hamiltonian; that is, it does not need to visit every vertex in the network. The length of the shortest path will be the sum of the weights for every edge that the path covers.

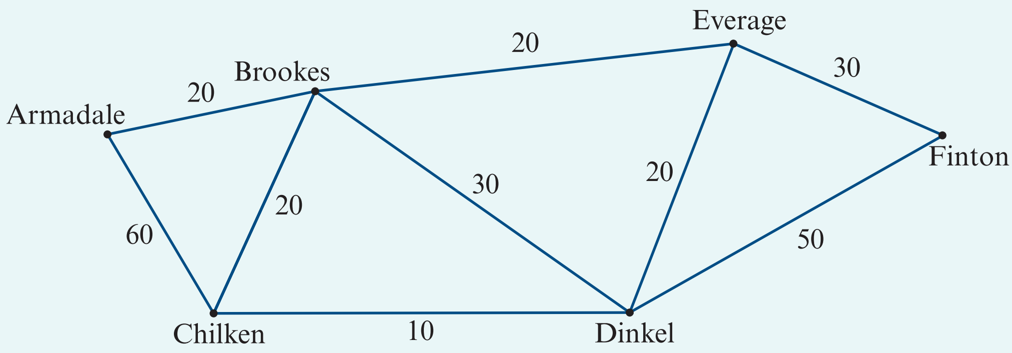
Shortest path problems are easy to solve by inspection if the network has a small number of vertices. Some edges may have very large weights compared to others and are probably best avoided, while other edges may have very small weights compared to others. It may be useful to try and incorporate these small weights into any shortest path.

Careful inspection of the network can help rule out any routes that are unlikely to contribute to the shortest path overall.

ISBN 978-1-108-46154-2

Example 12 Finding a shortest path

The graph below shows six towns represented by vertices and the roads between those towns represented by edges.



The weights on each of the edges show the travel times in minutes between each town.

Find the total time it takes to travel the shortest path from Armadale to Finton.

Solution

1 Consider smaller sections of the network, gradually moving through the network until the ending vertex is reached.	The shortest path from A to B is 20 (direct). The shortest path from A to C is 40 (via B). The shortest path from A to D is 50 (via B). The shortest path from A to E is 40 (via B). The shortest path from A to F is 70 (via B and E).
2 Write down the shortest path.	The shortest path from A to F is A – B – E – F.
3 Add the weights to find the total time.	Total time = 20 + 20 + 30 = 70 minutes
4 Write your answer.	The shortest time to travel from Armadale to Finton is 70 minutes, via Brookes and Everage.

ISBN 978-1-108-46154-2

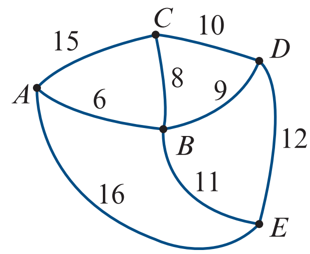
Cambridge University Press

Exercise 10G

Weighted graphs and networks

Example 12

1 The graph on the right shows towns A, B, C, D and E represented by vertices. The edges represent road connections between the towns and the weights on these edges are the average time, in minutes, it takes to travel along each road.



a Which two towns are 12 minutes apart by road?

b How long will it take to drive from C to D via B?

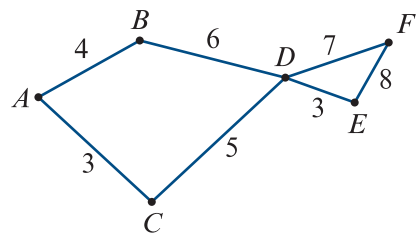
c A motorist intends to drive from D to E via B. How much time will they save if they travel directly from D to E?

d Find the shortest time it would take to start at A, finish at E and visit every town exactly once.

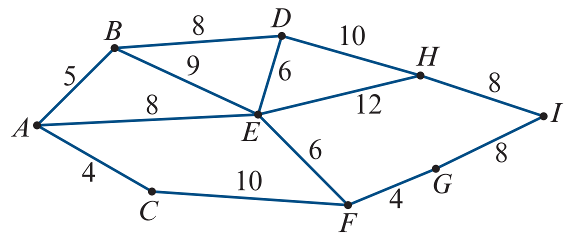
SF

Shortest path problems

2 Find the length of the shortest path from A to E in the graph on the right.



3 The network on the right shows the distance, in kilometres, along walkways that connect the landmarks A, B, C, D, E, F, G, H and I in a national park.

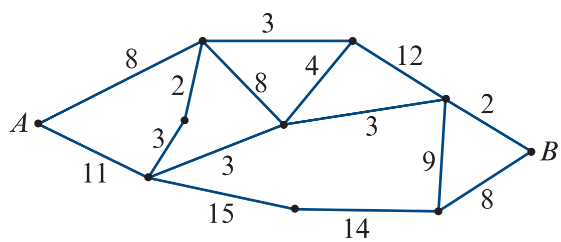


a What distance is travelled on the path $A - B - E - H - I$?

b What distance is travelled on the closed trail $F - E - D - H - E - A - C - F$?

c Find the shortest path from A to I.

4 In the network on the right, the vertices represent small towns and the edges represent roads. The weights on the edges indicate the distance, in kilometres, between towns.



Determine the length of the shortest path between towns A and B.

ISBN 978-1-108-46154-2

Cambridge University Press

Key ideas and chapter summary

AS

Graph	A **graph** is a diagram that shows the connections between objects, represented by vertices, as edges between those vertices.
Vertex	A **vertex** is a point in a graph that represents an object.
Edge	An **edge** is a line in a graph that connects two vertices and represents the connection between them.
Degree of a vertex	The **degree of a vertex** is the number of edges that are attached to that vertex. The degree of vertex A is written deg (A).
Loop	A **loop** is an edge that connects a vertex of a graph to itself. A loop contributes 2 to the degree of that vertex.
Multiple edge	Sometimes a graph has two or more edges that connect the same vertices. These are called **multiple edges**.
Isolated vertex	An **isolated vertex** in one that is not connected to any other vertex. Isolated vertices have degree zero.
Simple graph	A **simple graph** does not have any loops and does not have multiple edges.
Connected graph	A **connected graph** is a graph that has no isolated vertices and no separate parts.
Bridge	A **bridge** is an edge in a connected graph that, if removed, would leave the graph no longer connected.
Directed graph (digraph)	A **directed graph** (digraph) is a graph with a direction associated with the edges.
Subgraph	A **subgraph** is a graph that is part of a larger graph and has some of the same vertices and edges as that larger graph. A subgraph does not have any extra vertices or edges that do not appear in the larger graph.
Adjacency matrix	An **adjacency matrix** is a square matrix that uses a zero or an integer to record the number of edges connecting each pair of vertices in the graph.

ISBN 978-1-108-46154-2

Equivalent (isomorphic) graphs Graphs that contain identical information (connections between vertices) to each other are called **equivalent graphs** or **isomorphic graphs**.

Planar graph A **planar graph** is one that can be drawn so that no two edges cross over.

Face A **face** is an area in a graph enclosed between edges. It can only be reached by crossing an edge. One face on a graph is always the area surrounding the graph.

Euler's formula **Euler's formula** applies to planar graphs. It states that: 'the number of vertices plus the number of faces minus the number of edges always equals 2'.
If $v =$ the number of vertices, $f =$ the number of faces and $e =$ the number of edges in a planar graph, then
$v + f - e = 2$.

Exploring a graph Movement through a graph from one vertex to another along the edges of the graph is called exploring the graph.

Walk A **walk** is a route through a graph, from one vertex to another, along the edges of the graph. An open walk will start and end at different vertices. A closed walk will start and end at the same vertex.

Trail A **trail** is a walk that has no repeated edges. It may contain repeated vertices.

Path A **path** is a walk that has no repeated edges and no repeated vertices.

Closed trail A closed trail is a trail (no repeated edges) that starts and ends at the same vertex.

Cycle A **cycle** is also known as a closed path. It is a path (no repeated edges and no repeated vertices) that starts and ends at the same vertex.

Eulerian trail A closed trail (no repeated edges) that involves every edge of the graph is called an **Eulerian trail**.

ISBN 978-1-108-46154-2

Cambridge University Press

Eulerian graph	An Eulerian graph is a graph that contains an Eulerian trail.
Semi-Eulerian graph	A semi-Eulerian graph is a graph that contains an open trail that involves every edge of the graph.
Hamiltonian cycle	A **Hamiltonian cycle** is a closed path (no repeat edges, no repeat vertices) that visits every vertex of the graph. It will start and end at the same vertex.
Hamiltonian graph	A Hamiltonian graph is a graph that contains a Hamiltonian cycle.
Hamiltonian path	A **Hamiltonian path** is an open path (no repeat edges, no repeat vertices) that visits every vertex of the graph.
Semi-Hamiltonian graph	A semi-Hamiltonian graph is a graph that contains a Hamiltonian path.
Weighted graph	A **weighted graph** has numbers, called weights, associated withs the edges of a graph. The weights often represent physical quantities as additional information to the edge, such as time, distance or cost.
Network	A **network** is a weighted graph where the weights represent physical quantities such as time, distance or cost.
The travelling salesperson problem	The travelling salesperson problem involves finding the Hamiltonian cycle in a graph (visits all vertices in a graph once) that has the smallest total weight. It often refers to minimising the total distance travelled through a graph to return to the starting vertex.
The shortest path problem	The **shortest path problem** involves finding the shortest path from one vertex to another. The shortest path does not have to visit every vertex.

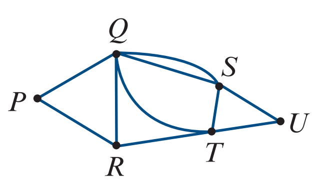
Skills check

Having completed this chapter, you should be able to:

- identify edges, vertices and loops in a graph
- determine the degree of a vertex in a graph
- define multiple edges, isolated vertices, connected graphs, bridges, simple graphs and completed graphs
- define directed graphs
- identify and draw subgraphs
- write adjacency matrices for graphs
- draw a graph from an adjacency matrix
- define planar graphs and redraw graphs in planar form
- identify equivalent graphs
- use Euler's formula
- define and identify walks, open walks, closed walks, trails and paths
- define and identify closed trails and cycles (closed paths)
- identify Eulerian trails
- identify Eulerian and semi-Eulerian graphs
- identify Hamiltonian cycles
- identify Hamiltonian and semi-Hamiltonian graphs
- define a weighted graph
- solve the travelling salesperson problem for a graph by finding the Hamiltonian cycle with the smallest total weight
- solve the shortest path problem by finding the shortest path between two vertices.

Multiple-choice questions

1 For the graph shown below, which vertex has degree 5?

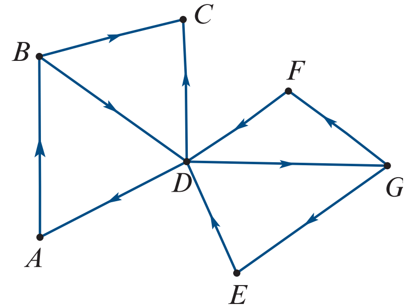


A Q **B** T **C** S **D** R **E** U

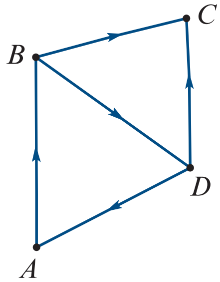
ISBN 978-1-108-46154-2

Cambridge University Press

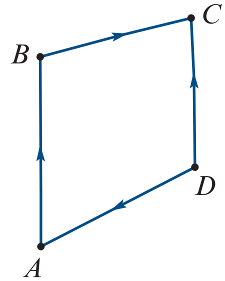
2 Which one of the following is **not** a subgraph of the graph shown on the right?



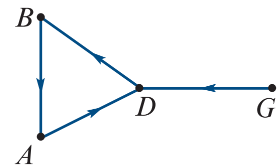
A



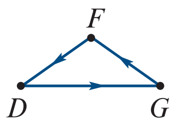
B



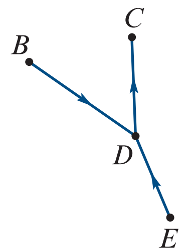
C



D



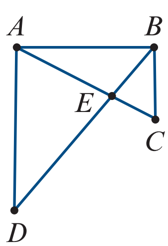
E



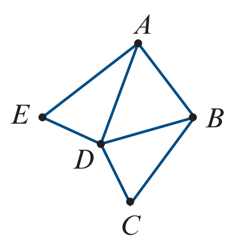
3 The graph that has been drawn from the adjacency matrix shown on the right is:

	A	B	C	D
A	0	1	0	1
B	1	0	1	0
C	0	1	0	1
D	1	0	1	0

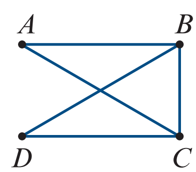
A



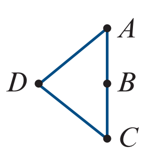
B



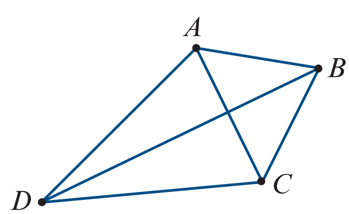
C



D



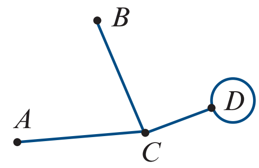
E



ISBN 978-1-108-46154-2

Cambridge University Press

4 The adjacency matrix that corresponds to the graph on the right is:



A
$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

B
$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

C
$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix} \end{array}$$

D
$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

E
$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

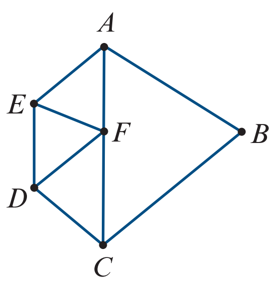
5 A connected graph with 15 vertices divides the plane into 12 faces. How many edges does this graph have?

A 15 **B** 23 **C** 24 **D** 25 **E** 27

6 A connected planar graph divides a plane into a number of faces. If the graph has eight vertices and these are linked by 13 edges, then how many faces does it have?

A 5 **B** 6 **C** 7 **D** 8 **E** 10

Use the graph below to answer questions 7 and 8.



7 The walk $A - E - D - C - B - A$ is best described as a:

A cycle
B Hamiltonian path
C Eulerian trail
D closed trail
E open walk

8 The walk $E - A - F - C - D - E$ is best described as a:

A Eulerian trail
B closed path
C Hamiltonian cycle
D closed trail
E open walk

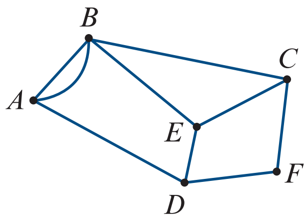
ISBN 978-1-108-46154-2

Cambridge University Press

Review

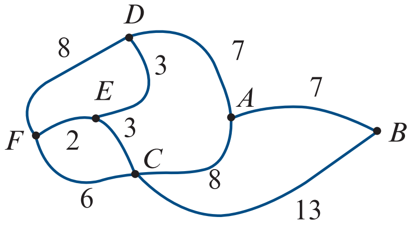
9 For the graph shown on the right, which edge could be removed to result in a semi-Eulerian graph?

A $A-B$
B $A-D$
C $B-E$
D $C-F$
E $D-F$



10 The length of the shortest path from F to B in the network shown on the right is:

A 17
B 18
C 19
D 20
E 21

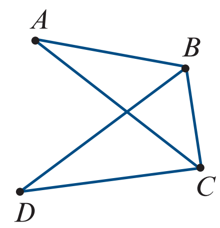


Short-answer questions

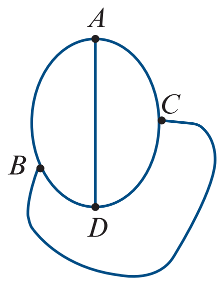
SF

1 Consider the graph shown on the right.

a What is the degree of vertex C?
b How many odd-degree vertices does this graph have?
c Write down the vertices that have an even degree.
d Redraw the graph in planar form.



2 Construct an adjacency matrix for the graph below.



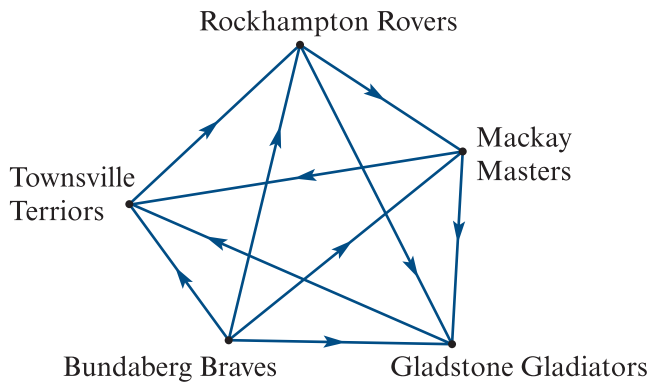
3 Draw a directed graph using the adjacency matrix below.

	A	B	C	D
A	0	1	0	1
B	1	0	1	0
C	1	0	0	0
D	0	0	2	0

ISBN 978-1-108-46154-2

Review

4 The directed graph on the right shows the results of a round-robin darts tournament. Each team is represented by a vertex and the arrows point to the winner of that match.

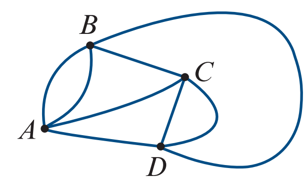


a Which team(s) beat the Rockhampton Rovers?

b How many teams did the Mackay Masters beat?

c Which team lost all of its games?

5 Consider the graph on the right.



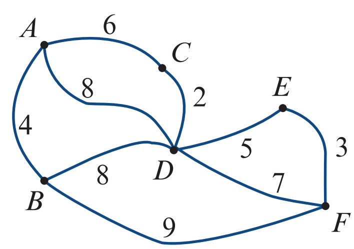
a How many vertices does this graph have?

b How many faces does this graph have?

c How many edges does this graph have?

d Verify Euler's formula for this graph.

6 Consider the graph below. The vertices represent cities in a particular state. The numbers on the arcs shows the time take, in hours, to drive between each city.



a In hours, which two cities are two hours driving time apart?

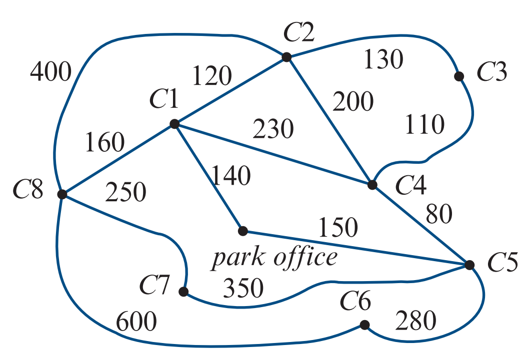
b In hours, what is the shortest driving time between E and B?

c In hours, what is the shortest driving time between F and A?

Extended-response questions

CF

1 The network opposite shows the walking tracks in a small national park. The tracks, represented by arcs, connect campsites to each other and to the park office. The weights on the arcs show the distance, in metres, between each location.



a The graph shown is planar. Explain what this means.

b Verify Euler's formula for this network.

ISBN 978-1-108-46154-2

c A ranger at campsite $C8$ plans to visit campsites $C1$, $C2$, $C3$, $C4$ and $C5$ on her way back to the park office. What is the shortest distance she will have to travel?

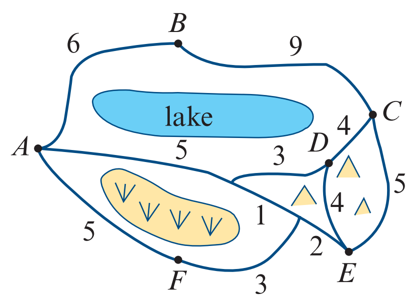
d Each day, the ranger on duty has to inspect each of the tracks to make sure that they are all safe to walk on.

- **i** Is it possible for her to do this starting and finishing at the park office, while only walking along a path once? Explain your answer.
- **ii** Identify one route that she could take.

e Another ranger wants to inspect each of the campsites but not pass through any campsite more than once on the inspection route. He wants to start and finish his inspection route at the park office.

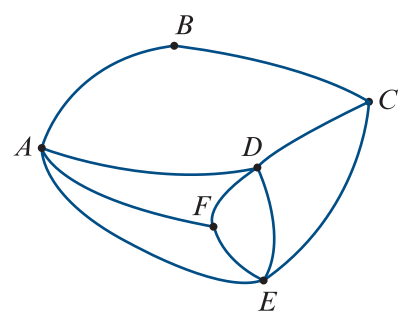
- **i** What is the mathematical name of the route he wants to take?
- **ii** With the present layout of tracks, he cannot inspect all the tracks without passing through at least one campsite twice. Suggest where an additional track could be added to solve this problem.
- **iii** With this new track, write down a route that he could follow.

2 The map on the right shows six campsites, A, B, C, D, E and F, which are joined by tracks. The numbers on the edges give the lengths, in kilometres, of those sections of track.



a Complete the following questions related to the map of the campsites.

- **i** Complete the graph opposite, which shows the shortest direct distances between campsites. The campsites are represented by vertices and the tracks are represented by edges.
- **ii** Fill in the missing entries for the adjacency matrix of the graph on the right.



	A	B	C	D	E	F
A	0	1	0	1	1	1
B	1	0	1	0	0	0
C	0	1	0	1	1	0
D	1	0	–	–	–	–
E	1	0	–	–	–	–
F	1	0	–	–	–	–

 ISBN 978-1-108-46154-2 Cambridge University Press

Review

b A walker follows the route $A - B - A - F - E - D - C - E - F - A$.

i How far does this person walk if they take the shortest track between each point?

ii Why is the route not a Hamiltonian cycle?

iii Write down a route that a walker could follow that is a Hamiltonian cycle.

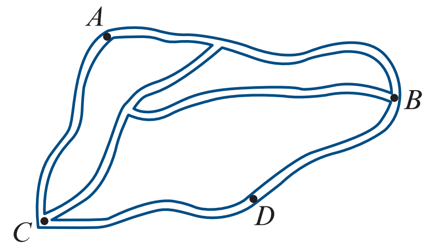
iv Find the distance walked by following this Hamiltonian cycle.

c It is impossible to start at A and return to A by going along each track exactly once. An extra track joining two campsites can be constructed so that this is possible.

Which two campsites need to be joined by a track to make this possible?

CU

3 Four children each live in a different town. The diagram below is a map of the roads that link the four towns, A, B, C and D.



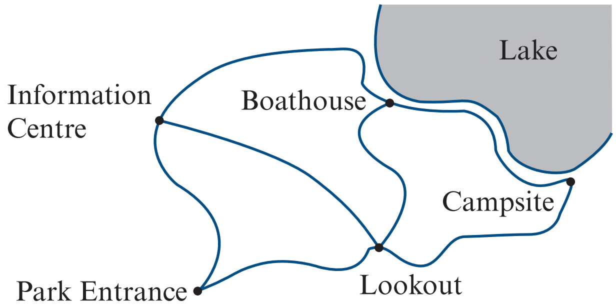
a How many different trails are there from town A to town D?

b How many different ways can a vehicle travel between town A and town B without visiting any other town?

c Draw this map as a graph by representing towns as vertices and each different route between two towns as an edge.

d Explain why a vehicle at A could not follow a trail through this graph.

4 A national park contains five locations connected by bushwalking tracks. The graph below shows the park entrance, information centre, lookout, boathouse and camping site represented by vertices and the bushwalking tracks represented by edges.



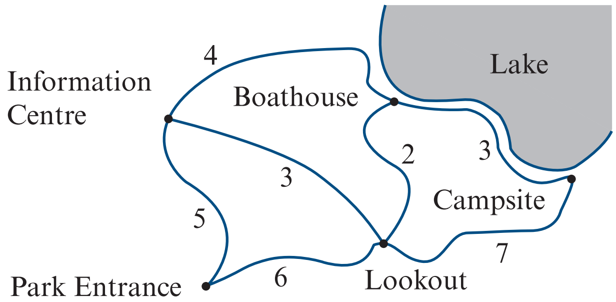
a If a bushwalker starts at the park entrance, write down two Hamiltonian cycles that he can follow.

b Is an Eulerian trail possible for this graph? Explain your answer with mathematical reasoning.

ISBN 978-1-108-46154-2

Cambridge University Press

c The network below has the distances between locations (in kilometres) added to each edge.



The bushwalker can walk at a speed of 4 km/hour.

i How long would it take the bushwalker to walk directly from the lookout to the boathouse?

ii Determine the shortest time it would take the bushwalker to walk between the park entrance and the camping site.

 ISBN 978-1-108-46154-2 Cambridge University Press

Chapter 11

11

Connector, assignment and flow problems

UNIT 4 INVESTING AND NETWORKING

Topic 3 Networks and decision mathematics

- How do we define a tree or a spanning tree?
- How do we draw a tree from any graph?
- How do we find the minimum distance between two vertices?
- How do we draw and use a bipartite graph to solve allocation problems?
- How do we find the optimal allocation of multiple groups of objects?
- How do we define a flow?
- How do we calculate the maximum flow through a network?

ISBN 978-1-108-46154-2

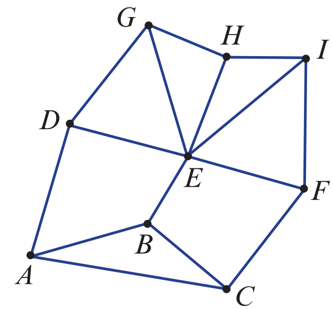
Cambridge University Press

11A Trees and connector problems

Trees

A **tree** is a connected graph that has no loops, no multiple edges, and no cycles. Recall that a cycle starts and ends at the same vertex and has no repeated vertices, nor repeated edges. Because a tree has no cycle, it will be impossible to find a walk in a tree that starts and ends at the same vertex without repeating an edge.

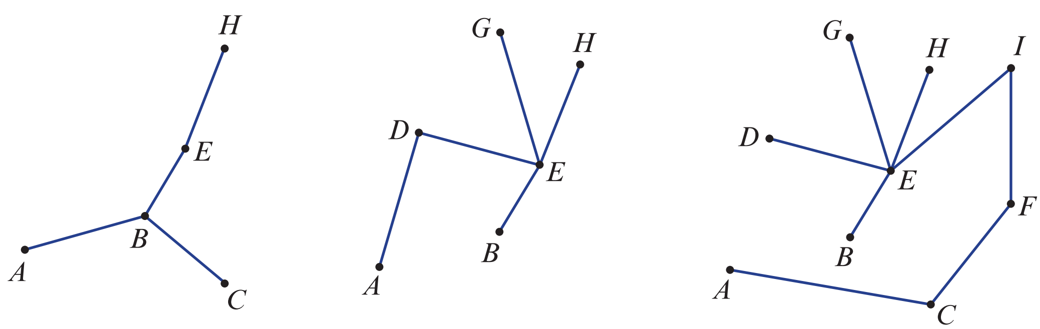
The number of edges in a tree will always be one less than the number of vertices.



Every connected graph will contain at least one subgraph that is a tree.

Some of the trees in the graph on the right are shown below.

In all three of the graphs below, a cycle does not exist. The trees are all subgraphs of the original graph. The first two trees involve some of the vertices of the original graph, but the third tree involves every vertex from the original graph.



Spanning trees

A **spanning tree** of a graph is a tree that includes every vertex of that graph. A connected graph may have more than one spanning tree.

A spanning tree can also be found for a weighted graph. The total weight of the spanning tree is the sum of all the weights in that spanning tree.

A spanning tree can be found by removing edges of the graph until there is one less edge than the number of vertices in the graph, making sure that all vertices remain connected.

A guide to trees: Watch the video in the Interactive Textbook to see trees, spanning trees and minimum spanning trees in action.

ISBN 978-1-108-46154-2

Cambridge University Press

Trees

A tree has no loops, multiple edges or cycles.

If a tree has n vertices, it will have $n - 1$ edges.

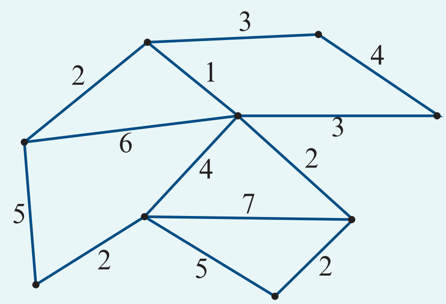
A *spanning tree* is a tree that connects all of the vertices of a graph.

The weight of a spanning tree for a weighted graph is the sum of the weights of the edges in that tree.

Example 1 Finding the weight of a spanning tree

a Draw one spanning tree for the graph shown on the right.

b Calculate the weight of the spanning tree.



Solution

a **1** Count the number of vertices and edges in the graph.

There are 9 vertices and 13 edges.

2 Calculate the number of edges in the spanning tree.

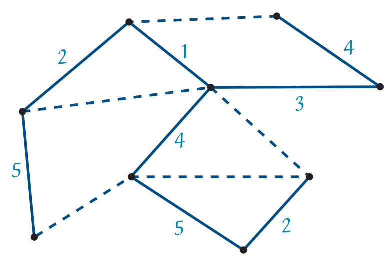
The spanning tree will have 9 − 1 = 8 edges.

3 Calculate how many edges must be removed.

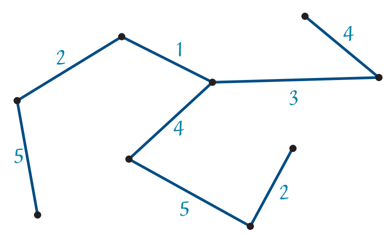
Remove 13 − 8 = 5 edges.

4 Choose edges to remove. Make sure that no vertex is left isolated.

Note: there are many other spanning trees that are possible. This is just one example.



b Add the weights of the remaining edges.



Weight = 5 + 2 + 1 + 4 + 5 + 2 + 3 + 4

= 26

ISBN 978-1-108-46154-2

Cambridge University Press

Minimum spanning trees

For any graph, there may be more than one spanning tree possible. If the weight of every spanning tree is found, there will always be a tree, or trees, that has a smaller total weight than all the others. This tree is called the **minimum spanning tree** for that graph.

For small graphs, the minimum spanning tree can be found by inspection, but for larger trees, an **algorithm** must be used. An algorithm is a series of instructions that can be used to solve a particular problem. The algorithm used to find the minimum spanning tree for a network is called **Prim's algorithm**.

Prim's algorithm

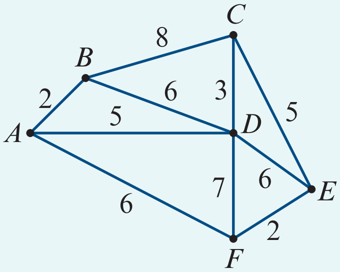
Prim's algorithm is used to find the minimum spanning tree for a network.

- Choose a starting vertex:
  - the spanning tree will contain every vertex and so any vertex can be chosen as the starting vertex.
- Inspect the edges that are connected to the starting vertex:
  - choose the edge that has the lowest weight
  - if there is more than one edge that has the lowest weight, it does not matter which one you choose
  - the starting vertex, the edge chosen, and the vertex connected by this edge form the beginning of the minimum spanning tree.
- Inspect the edges that are connected to the vertices chosen so far:
  - choose the one that has the lowest weight, but ignore any that would connect the tree back to itself
  - add the chosen edge and the vertex connected by this edge to the minimum spanning tree.

Repeat the process until all vertices have been added to the minimum spanning tree.

Example 2 Finding the minimum spanning tree

Apply Prim's algorithm to find the minimum spanning tree for the graph shown on the right. Write down the total weight of the minimum spanning tree.



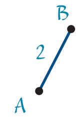
ISBN 978-1-108-46154-2

Cambridge University Press

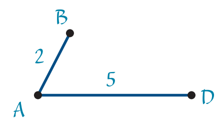
Solution

1 Choose any vertex to begin. We will start with vertex A.

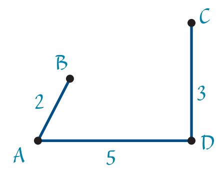
2 The smallest weighted edge from vertex A is to B with weight 2.



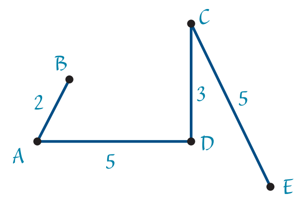
3 Look at vertices A and B. The smallest weighted edge from either vertex A or vertex B is from A to D with weight 5.



4 Look at vertices A, B and D. The smallest weighted edge from vertex A, B or D is from D to C with weight 3.

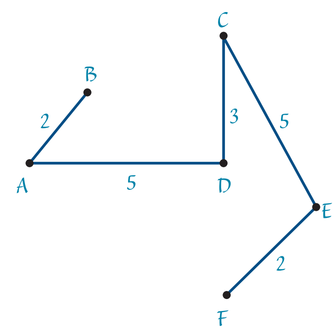


5 Look at vertices A, B, D and C. The smallest weighted edge from vertex A, B, D or C is from C to E with weight 5.



6 Look at vertices A, B, D, C and E. The smallest weighted edge from vertex A, B, D, C or E is from E to F with weight 2.

All vertices have been included in the graph. This is the minimum spanning tree.



7 Add the weights to find the total weight of the minimum spanning tree.

The total weight of the minimum spanning tree is $2 + 5 + 3 + 5 + 2 = 17$.

Kruskal's algorithm This alternative method is covered in the Interactive Textbook.

Connector problems

A connector problem involves a situation where it is important to minimise the connections between vertices for an overall minimum total weight. For example, if towns are to be connected to a gas pipeline, then it would be important to minimise the cost of the connections used.

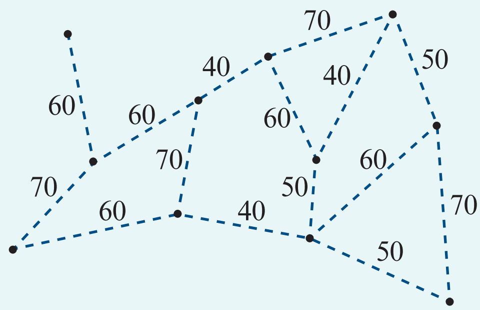
A minimum spanning tree is used to solve connector problems. The minimum spanning tree gives the edges required to keep all vertices connected in the graph for the smallest possible total weight.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 3 Solving a connector problem

At a showgrounds, 11 locations require access to water. These locations are represented by vertices on the network diagram shown. The dashed lines on the network diagram represent possible water pipe connections between adjacent locations. The numbers on the dashed lines show the minimum length of pipe, in metres, required to connect these locations.



a On the diagram, show where these water pipes should be placed to minimise the total amount of piping used.

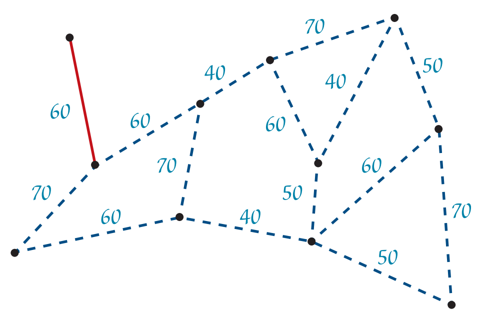
b Calculate the total length, in metres, of pipe that is required.

©VCAA (Further Maths, 2, 2011)

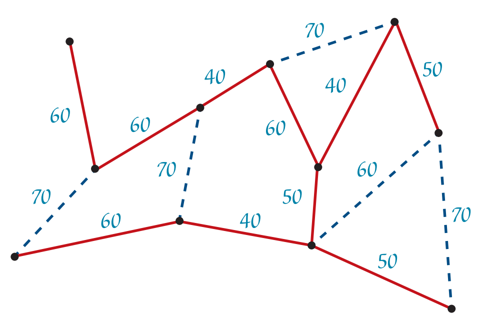
Solution

a **1** The water pipes will be a minimum length if they are placed on the edges of the minimum spanning tree for the network.

A good starting point for Prim's algorithm is the vertex that is connected by just one edge. This vertex must be connected to the minimum spanning tree by this edge.



2 Follow Prim's algorithm to find the minimum spanning tree.



b Add the weights of the minimum spanning tree. Write your answer in a sentence.

The length of water pipe required is
$60 + 60 + 40 + 60 + 40 + 50$
$+ 50 + 60 + 40 + 50 = 510$ *meters*

ISBN 978-1-108-46154-2

Exercise 11A

Trees and spanning trees

SF

1 Complete the following for the different trees.

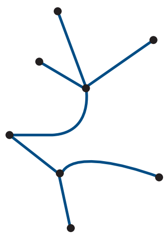
a How many edges are there in a tree with 12 vertices?

b How many vertices are there in a tree with 8 edges?

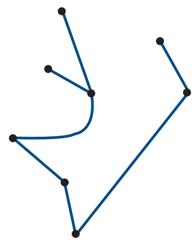
c Draw two different trees that have 5 vertices.

2 Which of the following graphs are trees?

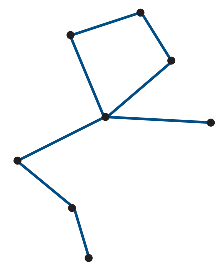
A



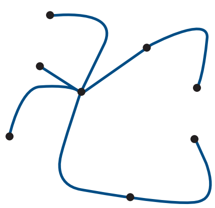
B



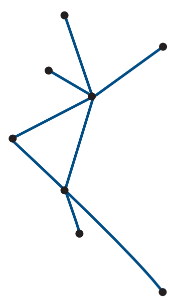
C



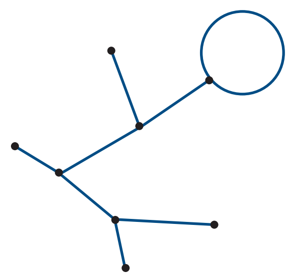
D



E



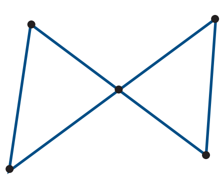
F



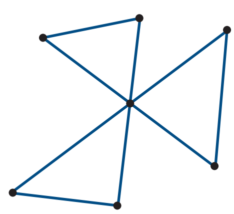
Example 1

3 For each of the following graphs, draw two different spanning trees.

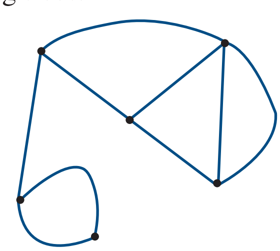
a



b



c



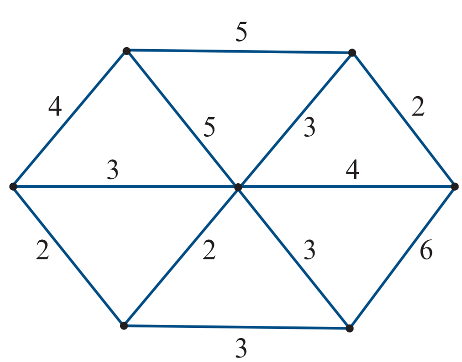
CF

4 A network is shown on the right.

a How many edges must be removed in order to leave a spanning tree?

b Remove some edges to form two different trees.

c For each tree in part **b**, find the total weight.



ISBN 978-1-108-46154-2

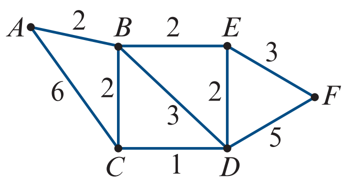
Cambridge University Press

Prim's algorithm

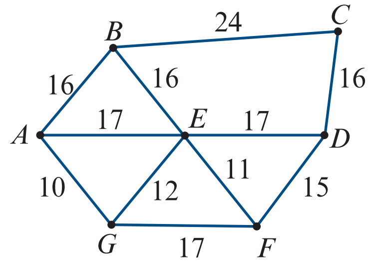
Example 2

5 Apply Prim's algorithm to determine a minimum spanning tree for each of the following graphs and then write down the total weight.

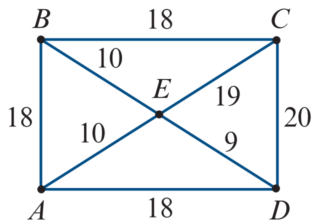
a



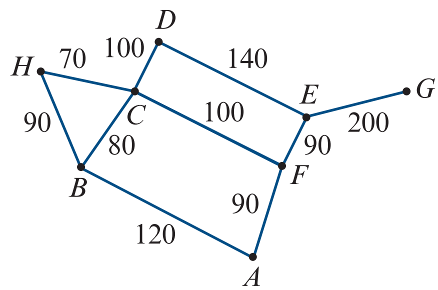
b



c



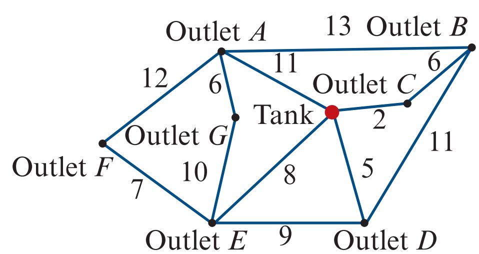
d



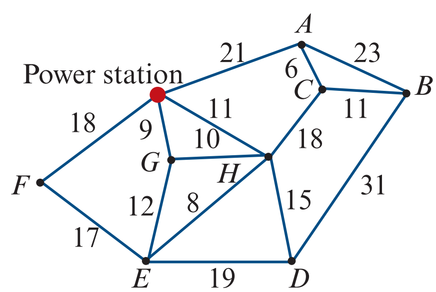
Connector problems

Example 3

6 Water is to be piped from a water tank to seven outlets on a property. The distances (in metres) of the outlets from the tank and from each other are shown in the network to the right. Starting at the tank, determine the minimum length of pipe needed to have water piped to all outlets in the property.



7 Power is to be connected by cable from a power station to eight substations (A to H). The distances (in kilometres) of the substations from the power station and from each other are shown in the network to the right. Determine the minimum length of cable needed to provide all substations with power.



ISBN 978-1-108-46154-2

Cambridge University Press

11B Assignment problems and the Hungarian algorithm

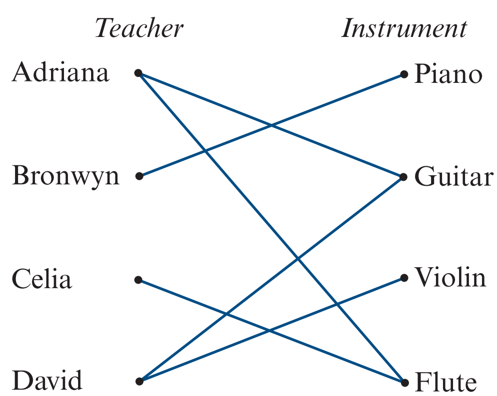
▶ Assignment problems

Assignment problems involve matching two groups of objects to each other based on particular needs or circumstances. For example, a school has particular subjects that need teachers and the school also has teachers that can teach particular subjects. The school would need to assign teachers to subjects to ensure that every class had a teacher. Another example of an assignment problem is a factory that has a number of machines and machine operators. The factory may want to assign an operator to a machine so that the total time it takes a process to occur is minimised.

Bipartite graphs

Assignment problems can be represented graphically using a **bipartite graph**. In a bipartite graph, there are two groups of vertices. The vertices from one group are connected to one or more vertices in the other group by edges.

The bipartite graph below has two groups of vertices, one for the music teachers in a music school and one for the instruments that are taught. Each teacher and instrument are represented by a vertex in the relevant group.



The edge in the bipartite graph connects the teachers to the instruments that they can teach. For example, Adriana can teach both Guitar and Flute because there is an edge connecting Adrianna to each of these instruments.

The bipartite graph can help the school assign each teacher to an instrument. For example, Bronwyn is the only teacher that can teach piano and so this assignment is necessary. Celia can only teach Flute and so, even though Adriana can teach both Flute and Guitar, if she teaches Flute then Celia will not be able to teach anything. So, Celia must teach Flute, which in turn means Adriana must teach Guitar.

General Mathematics Units 3 & 4 ISBN 978-1-108-46154-2
Cambridge Senior Mathematics for Queensland

Example 4 Solving an assignment problem with a bipartite graph

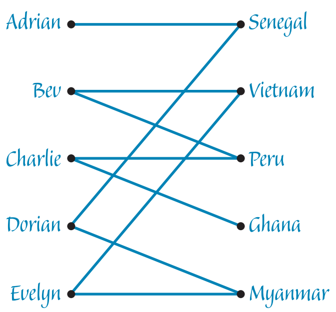
Adrian, Bev, Charlie, Dorian and Evelyn are presenters on a TV travel show. Each presenter will be assigned a story to film about one country that they have visited before.

- Adrian has visited Senegal.
- Bev has visited Vietnam and Peru.
- Charlie has visited Ghana and Peru.
- Dorian has visited Senegal and Myanmar.
- Evelyn has visited Vietnam and Myanmar.

Construct a bipartite graph of the information above and use it to decide on the assignment of each presenter to one country.

Solution

1 The two groups of items are: Presenters and Countries. Draw a vertex for each presenter in one column and each country in another.

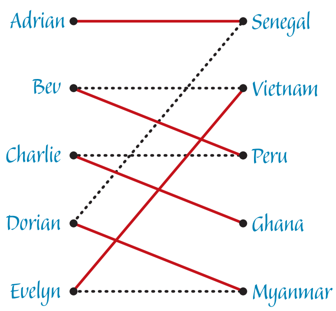


2 Adrian has visited Senegal so she could be sent there to film her story. Join the vertices for Angie and Senegal with an edge.

3 Bev has visited Vietnam and Peru so join the vertex for Bev to the vertices for Vietnam and Peru.

4 Join the other presenter vertices to country vertices in a similar way.

5 Definite assignments are shown in red, impossible assignments are shown with dotted lines in the bipartite graph.



- Adrian is only connected to Senegal and so must visit this country. If Angie visits Senegal, Dorian cannot.
- If Dorian cannot visit Senegal, he must visit Myanmar.
- If Dorian visits Myanmar, then Evelyn cannot.
- If Evelyn cannot visit Myanmar, she must visit Vietnam.
- If Evelyn must visit Vietnam, Bev cannot and so she must visit Peru.
- If Bev must visit Peru, then Charlie cannot. Charlie must visit Ghana.

6 Write the assignments.

Adrian will visit Senegal.
Bev will visit Peru.
Charlie will visit Ghana.
Dorian will visit Myanmar.
Evelyn will visit Vietnam.

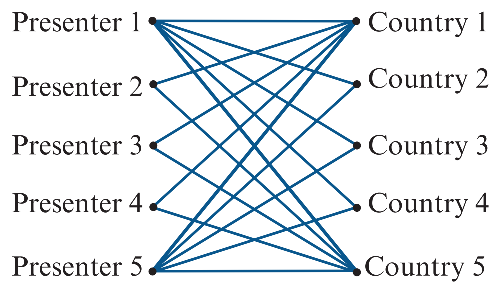
ISBN 978-1-108-46154-2

Cambridge University Press

Complete bipartite graphs

The assignment problem of presenter to country in Example 4 would be greatly simplified if every presenter could visit all of the countries. Every presenter vertex would be connected to every country vertex and there would be many different assignments that were possible. The bipartite graph for this situation would be a complete graph.

Rather than just assign presenters to countries randomly, the producers could use information about the preferences of the presenters, or perhaps the number of times that they have been to each of the countries, to make the assignment with priority. This information would be weights on the edges of an already very complex graph. Rather than write these weights on the graph, they can be recorded in a table instead.



Cost matrix

The table of weights for a bipartite graph is called a **cost matrix**. Even though it is called a cost matrix, the 'cost' does not have to be in terms of money. It could be the time taken to complete a task or the distance that needs to be travelled.

As an example, a factory might need to assign each of four employees to one of four machines.

The cost matrix on the right shows the time each employee takes to complete the task on each machine. This assumes that each of the four employees can perform all of the tasks equally as well as each other in terms of the quality of work.

Employee	A	B	C	D
Wendy	30	40	50	60
Xenofon	70	30	40	70
Yolanda	60	50	60	30
Zelda	20	80	50	70

The cost matrix only shows that the time each person takes to do the task is different. How best to assign the employees to a machine is determined by an algorithm.

▶ The Hungarian algorithm

The **Hungarian algorithm** is used to determine the best assignment of employee to machine so that the overall time taken to complete the tasks is minimised.

ISBN 978-1-108-46154-2

Cambridge University Press

Performing the Hungarian algorithm

Step 1: Subtract the lowest value in each row, from every value in that row.

- 30 has been subtracted from every value in the row for Wendy.
- 30 has been subtracted from every value in the row for Xenofon.
- 30 has been subtracted from every value in the row for Yolanda.
- 20 has been subtracted from every value in the row for Zelda.

Employee	A	B	C	D
Wendy	0	10	20	30
Xenofon	40	0	10	40
Yolanda	30	20	30	0
Zelda	0	60	30	50

Step 2: If the minimum number of lines required to cover all the zeroes in the table is equal to the number of allocations to be made, jump to step 6. Otherwise, continue to step 3.

- The zeroes can be covered with three lines. This is less than the number of allocations to be made (4).
- Continue to step 3.

Employee	A	B	C	D
Wendy	0	10	20	30
Xenofon	40	0	10	40
Yolanda	30	20	30	0
Zelda	0	60	30	50

Step 3: If a column does not contain a zero, subtract the lowest value in that column from every value in that column.

- Column C does not have a zero.
- 10 has been subtracted from every value in column C.

Employee	A	B	C	D
Wendy	0	10	10	30
Xenofon	40	0	0	40
Yolanda	30	20	20	0
Zelda	0	60	20	50

Step 4: If the minimum number of lines required to cover all the zeroes in the table is equal to the number of allocations to be made, jump to step 6. Otherwise, continue to step 5a.

- The zeros can be covered with three lines. This is less than the number of allocations to be made (4).
- Continue to step 5a.

Employee	A	B	C	D
Wendy	0	10	10	30
Xenofon	40	0	0	40
Yolanda	30	20	20	0
Zelda	0	60	20	50

ISBN 978-1-108-46154-2

Cambridge University Press

Step 5a: Add the smallest uncovered value to any value that is covered by two lines. Subtract the smallest uncovered value from all the uncovered values.

- The smallest uncovered **element** is 10.
- 10 has been *added* to Xenofon–A and Xenofon– D because these values are covered by two lines.
- 10 has been *subtracted* from all the uncovered values.

Employee	A	B	C	D
Wendy	0	0	0	30
Xenofon	50	0	0	50
Yolanda	30	10	10	0
Zelda	0	50	10	50

Step 5b: Repeat from step 4.

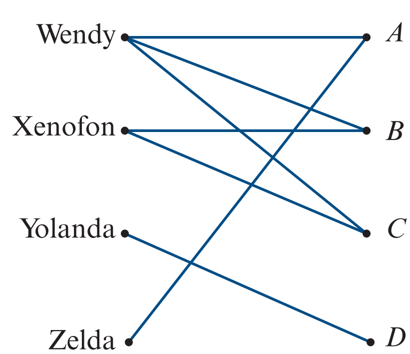
- The zeroes can be covered with a minimum of four lines. This is the same as the number or allocations to make.
- Continue to step 6.

Employee	A	B	C	D
Wendy	0	0	0	30
Xenofon	50	0	0	50
Yolanda	30	10	10	0
Zelda	0	50	10	50

Step 6: Draw a directed bipartite graph with an edge for every zero value in the table.

In the bipartite graph:

- Wendy will be connected to A, B and C
- Xenofon will be connected to B and C
- Yolanda will be connected to D
- Zelda will be connected to A.



Step 7 : Make the allocation and calculate minimum cost.

- Zelda must operate machine A (20 minutes).
- Yolanda must operate machine D (30 minutes).
- Wendy can operate either machine B (40 minutes) or C (50 minutes).
- Xenofon can operate either machine C (40 minutes) or B (30 minutes).

Note: Because Wendy and Xenofon can operate either B or C, there are two possible allocations. Both allocations will have the same minimum cost.

The minimum time taken to finish the work $= 20 + 30 + 40 + 40 = 130$ minutes, or $20 + 30 + 50 + 30 = 130$ minutes.

General Mathematics Units 3 & 4 ISBN 978-1-108-46154-2

Cambridge University Press
Cambridge Senior Mathematics for Queensland

The Hungarian algorithm

- A cost matrix is a table of weights for a bipartite graph.
- The Hungarian algorithm is used to determine an **allocation** for minimum overall cost.
- The Hungarian algorithm assumes that the cost in the cost matrix is the only decision factor for allocation. For example, a cost matrix that shows the time that four people take to complete four jobs assumes that the time taken is the only difference and that the quality of the work for all people is the same.

Exercise 11B

Bipartite graphs

Example 4

SF

1 Gloria, Minh, Carlos and Trevor are buying ice-cream. They have a choice of five flavours: chocolate, vanilla, peppermint, butterscotch and strawberry.

- Gloria likes vanilla and butterscotch, but not the others.
- Minh only likes strawberry.
- Carlos likes chocolate, peppermint and butterscotch.
- Trevor likes all flavours.

a Explain why a bipartite graph can be used to display this information.

b Draw a bipartite graph with the people on the left and flavours on the right.

c What is the degree of the vertex representing Trevor?

The ice-cream shop has no butterscotch ice-cream available. Gloria, Minh, Carlos and Trevor will have only one ice-cream each and will all have a different flavour.

d Who must have the vanilla ice-cream?

e If Carlos chooses peppermint, write down the allocation of ice-cream flavour to these four people.

2 Joni, Ian, Dylan and Joshua are teachers in a school. The school has a Maths class, an English class, a Geography class and a Science class, each of which a teacher. Each teacher can be allocated one class only.

Joni can teach English or Geography.
Ian can teach Maths or Science.
Dylan can teach English or Geography.
Joshua can teach Geography or Science.

a Draw a bipartite graph to show the teachers and the subject that they can teach.

b Explain why Joshua must take the science class.

c Write two different allocations of teachers to subjects.

ISBN 978-1-108-46154-2

Cambridge University Press

3 The table below shows five people in the rows and five sports in the columns. A tick (✓) in a table cell indicates that the person in that row can coach the sport in that column. A cross (×) indicates that they cannot coach that sport.

Each sport in the table must be coached by only one of the people in the table.

	Hockey	Cricket	Soccer	Rugby	Squash
Rob	✓	×	✓	×	×
Janet	×	✓	✓	✓	✓
Tara	×	✓	×	✓	×
Diana	✓	×	×	✓	×
Jason	×	×	×	✓	×

a Draw a bipartite graph to represent the information contained in the table above.
b Explain why Diana must coach hockey.
c Write the allocation of people to sports.

The Hungarian algorithm

4 Three volunteer workers, Joe, Meg and Ali, are available to help with three jobs. The time (in minutes) in which each worker is able to complete each task is given in the table opposite. Which allocation of workers to jobs will enable the jobs to be completed in the minimum time?

	Job		
Student	A	B	C
Joe	20	20	36
Meg	16	20	44
Ali	26	26	44

ISBN 978-1-108-46154-2

5 A football association is scheduling football games to be played by three teams (the Champs, the Stars and the Wests) on one day. On this day, one team must play at their Home ground, one will play Away and one will play at a Neutral ground. The costs (in $'000s) for each team to play at each of the grounds are given in the table to the right.

Team	Home	Away	Neutral
Champs	10	9	8
Stars	7	4	5
Wests	8	7	6

Determine a schedule that will minimise the total cost of playing the three games and determine this cost.

Note: There are two different ways of scheduling the games to achieve the same minimum cost. Identify both of these.

CF

6 A company has four machine operators and four different machines that they can operate. The table shows the hourly cost in dollars of running each machine with each operator. How should the operators be allocated to the machines in order to minimise the totally hourly cost of operating all machines?

	Machine			
Operator	W	X	Y	Z
A	38	35	26	54
B	32	29	32	26
C	44	26	23	35
D	20	26	32	29

7 **a** A cost matrix is shown. Find the allocation(s) by the Hungarian algorithm that will give the minimum cost.

	A	B	C	D
W	110	95	140	80
X	105	82	145	80
Y	125	78	140	75
Z	115	90	135	85

b Find the minimum cost for the given cost matrix and give a possible allocation.

	A	B	C	D
W	2	4	3	5
X	3	5	3	4
Y	2	3	4	2
Z	2	4	2	3

ISBN 978-1-108-46154-2

Cambridge University Press

8 A roadside vehicle assistance organisation has four service vehicles located in four different places. The table below shows the distance (in kilometres) of each of these service vehicles from four motorists in need of roadside assistance.

	Motorist			
Service vehicle	Jess	Mark	Raj	Karla
A	18	15	15	16
B	7	17	11	13
C	25	19	18	21
D	9	22	19	23

Determine a service vehicle assignment that will ensure that the total distance travelled by the service vehicles is minimised. Determine this distance.

Note: There are two ways that the service vehicles can be assigned to minimise the total distance travelled. Identify both of these.

9 A relay race comprises four individual track distances: 100 m, 400 m, 800 m and 1500 m. Four students will each run one of the track distances. The time (in seconds) each student can run each distance is given in the table below.

Student	100 m	400 m	800 m	1500 m
Dimitri	11	62	144	379
John	13	60	146	359
Carol	12	61	149	369
Elizabeth	13	63	142	349

Use the Hungarian algorithm to assign each student to one distance so that the total time taken to complete the race is minimised.

ISBN 978-1-108-46154-2

Cambridge University Press

11C Flow networks

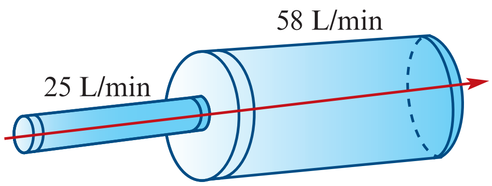
Flow problems

Flow problems involve the transfer or **flow** of material from one point, called the **source**, through a **flow network** to another point called the **sink**. Water flowing through a network of pipes or traffic flowing along a network of roads are examples of flow situations. Flow problems involve determining how much water or traffic in total can pass through the network, based on restrictions such as the **capacity** of the individual pipes.

Maximum flow

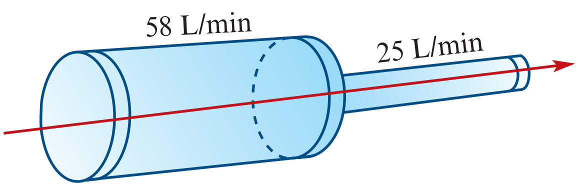
Consider the flow situation of water flowing through pipes. The water starts at the source, flows through the network and then ends up at the sink. The water will flow in one direction only. If the pipes are all different sizes, then water will flow at different rates through different parts of the network. Overall, this will affect the rate at which the water arrives at the sink.

The diagram on the right shows two pipes that are joined together. The small pipe has a capacity of 25 litres per minute and this is joined to a larger pipe with capacity 58 litres per minute. Water flows through from the source, into the small pipe, through the large pipe and out to the sink.



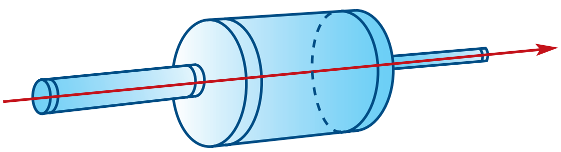
Even though the large pipe has a capacity of 58 litres per minute, the small pipe restricts the flow of water into the large pipe to 25 litres per minute. The flow through the large pipe will never be more than 25 litres per minute.

If the connection is reversed, water will be able to enter the large pipe at the rate of 58 litres per minute, but there will be a 'bottleneck' of flow at the junction between the large and small pipe. The large pipe can deliver 58 litres of water every minute to the small pipe, but the small pipe can only allow 25 litres per minute to pass.



In both of the situations above, the flow through the entire pipe system (both pipes from source to sink) is restricted to a maximum flow of 25 litres per minute. This is the capacity of the smallest pipe in the connection.

If we connect more pipes together, one after the other, we can calculate the overall capacity of **maximum flow** of the pipe system by looking for the smallest capacity pipe in that system.



ISBN 978-1-108-46154-2

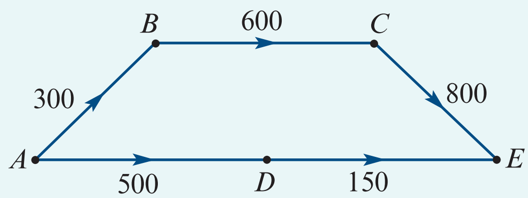
Cambridge University Press

Maximum flow

If pipes of different capacities are connected one after the other, the *maximum flow* through the pipes is equal to the *minimum* capacity of the individual pipes.

Example 5 Calculating the maximum flow

In the flow network shown on the right, the vertices A, B, C, D and E represent towns. The edges of the graph represent roads and the weights of those edges are the maximum number of cars that can travel on the road each hour. The roads allow only one-way travel, as indicated by the arrow.



a Find the maximum traffic flow from A to E through town C.

b Find the maximum traffic flow from A to E overall.

c A new road is being built to allow traffic from town D to town C. This road can carry 500 cars per hour.

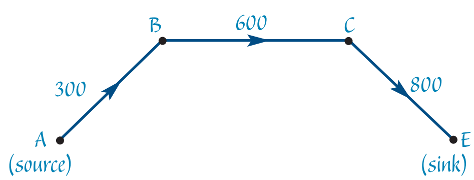
  i Add this road to the flow network.

  ii Find the maximum traffic flow from A to E overall after this road is built.

Solution

a Look at the subgraph that includes town C.

The smallest capacity of the individual roads is 300 cars per hour. This will be the maximum flow through town C.



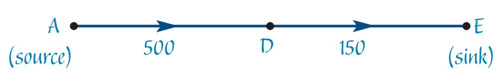
The maximum flow from A to E through town C is equal to the smallest capacity road along that route. The maximum flow is 300 cars per hour.

b Look at the two subgraphs from A to E.

The maximum flow through C will be 300 cars per hour.

The maximum flow through D will be 150 cars per hour (minimum capacity).

Add the maximum flow through C to the maximum flow through D.



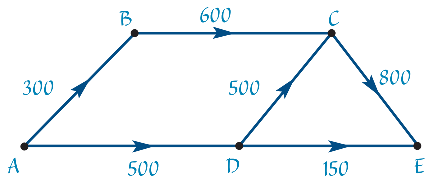
The maximum flow from A to E overall is:

300 + 150 = 450 cars per hour

ISBN 978-1-108-46154-2

Cambridge University Press

c **i** Add the edge from D to C representing the new road to the diagram.



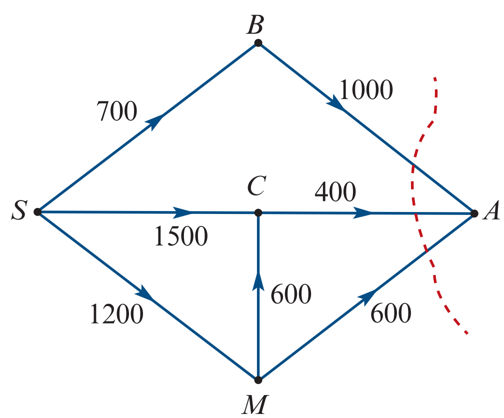
ii Determine the maximum flow.

The maximum flow through A – B – C – E is 300. But C – E has capacity 800. If another 500 cars per hour come through D – C, they will be able to travel from C – D.
The new maximum flow is now 800 cars per hour.

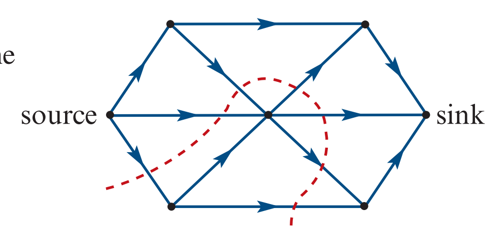
Cuts

For flow networks that contain many vertices and edges, it can be difficult to determine the maximum flow by inspection. We can simplify the search for maximum flow by searching for **cuts** with the network.

A cut divides the flow network into two parts, completely separating the source from the sink. It is helpful to think of cuts as imaginary blocks in the flow that do not allow any flow across them. In the diagram on the right, the dotted line is a cut. It completely blocks the flow from the source (S) to the sink (A).



The second graph shown on the right contains a dotted line that is *not* a cut. It blocks some of the flow through the network but there is still a flow pathway from the source to the sink across the top of the network.



A cut must *completely* block the flow from the source to the sink.

Cut capacity

The **cut capacity** for any cut is the sum of all the weights of the edges that the cut passes through. Only flow from the source side to the sink side is considered in the calculation of cut capacity. Any flow from the sink side across the cut to the source side is ignored.

ISBN 978-1-108-46154-2

Cambridge University Press

Cuts and cut capacity

A *cut* is an imaginary line across a flow network that completely blocks all flow from the source to the sink.

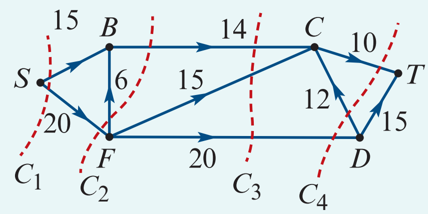
The *cut capacity* of a cut is the sum of the capacities for the edges of the flow network that are blocked by that cut. Only edges that flow from the source side of the cut to the sink side of the cut are considered.

Example 6 Calculating cut capacity

Calculate the capacity of each of the four cuts shown in the flow network on the right.

The cuts are labelled C_1, C_2, C_3 and C_4.

The source is vertex S and the sink is vertex T.



Solution

1 All edges in C_1 are counted as they all flow from S to T across the cut. — The capacity of $C_1 = 15 + 20 = 35$

2 The edge from F to B in C_2 is not counted. F is on the sink side of the cut and the flow crosses the cut back to the source side. — The capacity of $C_2 = 14 + 20 = 34$

3 All edges in C_3 are counted as they all flow from S to T across the cut. — The capacity of $C_3 = 14 + 15 + 20 = 49$

4 The edge from D to C in C_4 is not counted. D is on the sink side of the cut and the flow crosses the cut back to the source side. — The capacity of $C_4 = 20 + 10 = 30$

Maximum flow and cut capacity

The capacity of a cut is important to determine the *maximum flow* through any flow network. Look for the smallest, or **minimum, cut capacity** that exists in the graph. This will be the same as the maximum flow that is possible through that graph.

Maximum flow

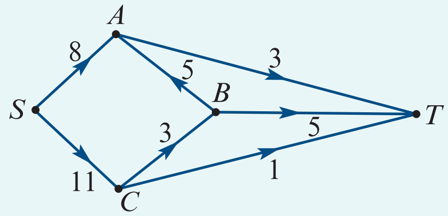
The *maximum flow* that is possible through a flow network is the same as the minimum cut capacity possible for that network.

ISBN 978-1-108-46154-2

Cambridge University Press

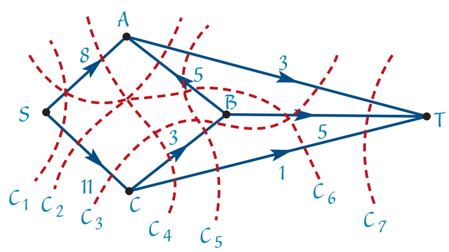
Example 7 Calculating maximum flow

Calculate the maximum flow from S to T for the flow network shown on the right.



Solution

1 Mark in all possible cuts on the network.



2 Calculate the capacity of all the cuts.

The capacity of $C_1 = 8 + 11 = 19$
The capacity of $C_2 = 3 + 11 = 14$
The capacity of $C_3 = 3 + 5 + 11 = 19$
The capacity of $C_4 = 8 + 3 + 1 = 12$
The capacity of $C_5 = 3 + 3 + 1 = 7$
The capacity of $C_6 = 8 + 5 + 1 = 14$
The capacity of $C_7 = 3 + 5 + 1 = 9$

3 Identify the minimum cut capacity and write your answer.

The minimum cut capacity is 7 so the maximum flow from S to T is 7.

ISBN 978-1-108-46154-2

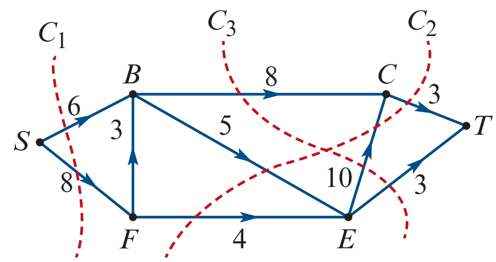
Cambridge University Press

Exercise 11C

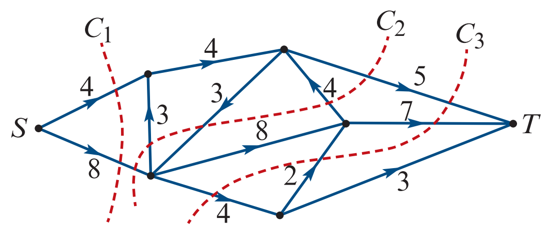
Cuts and cut capacity

Example 6

1 Find the capacity of each of the cuts in the flow network on the right. The source is vertex S and the sink is vertex T.



2 Find the capacity of each of the cuts in the flow network on the right. The source is vertex S and the sink is vertex T.

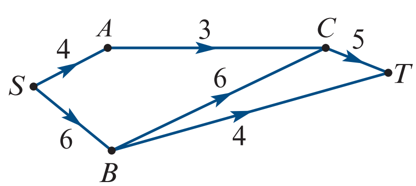


Calculating maximum flow

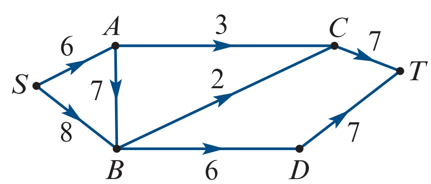
Example 7

3 Find the maximum flow for each of the following flow networks. The source is vertex S and the sink is vertex T.

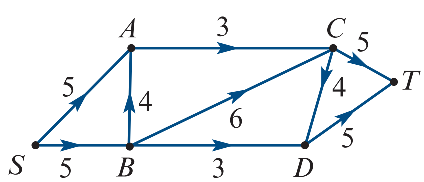
a



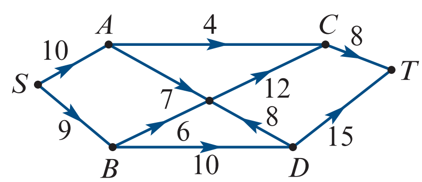
b



c



d



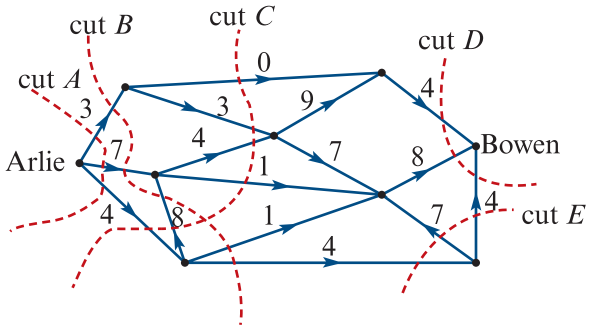
ISBN 978-1-108-46154-2

Cambridge University Press

Applications of maximum flow

Example 5

4 A train has the stages of its journey represented by the edges on the following directed network. The number of available seats for each stage is indicated beside the corresponding edge, a shown on the diagram on below.



The five cuts, A, B, C, D and E, shown on the network are attempts to find the maximum number of available seats that can be booked for a journey from Arlie to Bowen.

a Write down the capacity of cut A, cut B, cut C, cut D and cut E.

b Explain why cut E is not a valid cut when trying to find the minimum cut between Arlie and Bowen.

c Find the maximum number of available seats for a train journey from Arlie to Bowen.

ISBN 978-1-108-46154-2

Cambridge University Press

Review

Key ideas and chapter summary

AS

Tree	A tree is a graph that has no loops, multiple edges nor cycles. If a tree has n vertices, then it will have $n–1$ edges.
Spanning tree	A **spanning tree** for any graph is a tree that connects all vertices of that graph.
Minimum spanning tree	The **minimum spanning tree** for a graph is the spanning tree that has the smallest possible total weight for that graph.
Prim's algorithm	**Prim's algorithm** is an algorithm that is used to determine the minimum spanning tree for a graph.
Connector problems	A connector problem is a problem where it is important to minimise the total weight of connections between objects or locations. The weights in connector problems can be length, time, cost or other physical quantity. Connector problems are solved by finding the minimum spanning tree for the graph that represents the problem.
Assignment problems	Assignment problems involve matching the objects in one group to objects in another so that the overall cost in terms of time, money or other quantity is minimised. Assignment problems are solved with bipartite graphs and/or the Hungarian algorithm.
Bipartite graph	A **bipartite graph** is a graph where the vertices exist in separate groups. The edges of a bipartite graph connect vertices in one group with vertices in the other.
Complete bipartite graph	In a complete bipartite graph, every vertex in one group of the bipartite graph is connected by an edge to every vertex in the other group.
Cost matrix	A **cost matrix** is a table of weights for a complete bipartite graph. It contains the cost, in terms of time, money or other quantity, of assigning objects from one group to objects in another. An example of a cost matrix is a table of the time it will take people (one group) to complete tasks (another group).
Hungarian algorithm	The **Hungarian algorithm** is an algorithm that is used to determine the best allocation to minimise the total cost.

ISBN 978-1-108-46154-2

Flow — **Flow** is the transfer of material, such as water, gas or traffic, through a directed network.

Flow network — A **flow network** occurs where the directed edges of the graph represent the flow of material from one vertex to another. The weight of an edge of a flow network is called the **capacity** of that edge.

Source — The **source** is the origin of material that flows through a network.

Sink — The **sink** is the final destination of material that flows through a network.

Flow problems — A flow problem is a problem that involves maximising the amount of material that flows through a network.
Flow problems can be solved by finding the minimum cut for the network.

Cut — A **cut** is an imaginary line dividing a directed graph into two parts, one containing the source and the other containing the sink. It can be imagined that the cut blocks the flow through any edge it crosses.

Cut capacity — The **cut capacity** of a cut is the sum of all the weights on the edges it cuts. Only edges that flow from the source side of the cut to the sink side of the cut are considered.

Maximum flow — The **maximum flow** possible through a directed graph is the same as the smallest cut capacity of the cuts that are possible in that graph.

Skills check

Having completed this chapter, you should be able to:

- define and describe a tree
- draw a tree from any graph
- identify a spanning tree for any graph
- calculate the weight of a spanning tree
- identify the minimum spanning tree using Prim's Algorithm

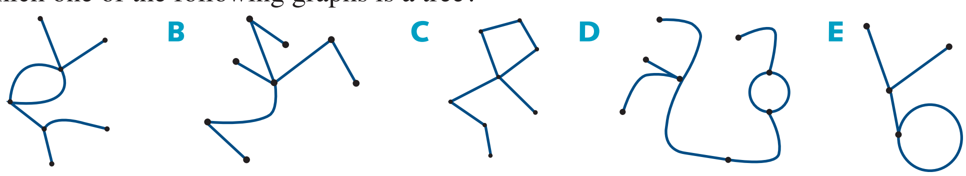
Review

- solve connector problems using Prim's Algorithm
- define and describe bipartite graphs
- solve simple assignment problems using bipartite graphs
- solve assignment problems using the Hungarian Algorithm
- define and describe flow
- calculate the maximum flow through a flow network by observation
- identify cuts and calculate cut capacities
- determine the maximum flow through a flow network by finding the capacity of the minimum cut
- solve flow problems by finding minimum cut capacities.

Multiple-choice questions

1 Which one of the following graphs is a tree?

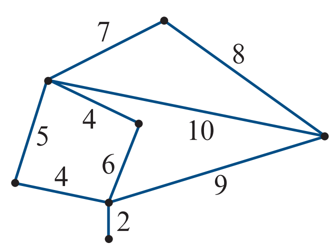
A **B** **C** **D** **E**



2 A graph has 6 vertices and 12 edges. A spanning tree for this graph will have:

A 5 vertices and 12 edges
B 5 vertices and 5 edges
C 6 vertices and 5 edges
D 6 vertices and 6 edges
E 6 vertices and 12 edges

3 For the graph shown here, the minimum spanning tree has length:



A 30 **B** 31 **C** 33
D 34 **E** 36

4 There are four different human blood types: O, A, B and AB. Blood can be donated from one human to another using the following rules:

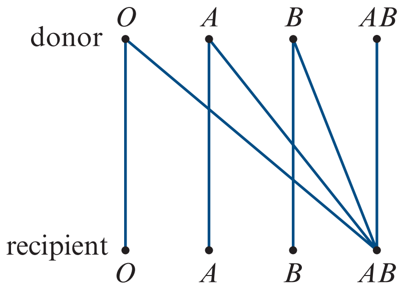
- Type O can donate blood to any type.
- Type AB can receive blood from any type
- Each type can donate blood to its own type.
- Each type can receive blood from its own type.

ISBN 978-1-108-46154-2

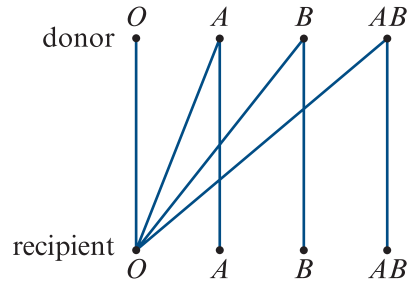
Cambridge University Press

Which one of the following bipartite graphs correctly represents this information?

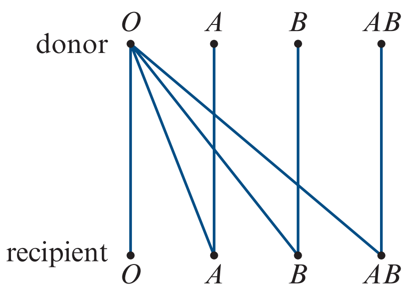
A



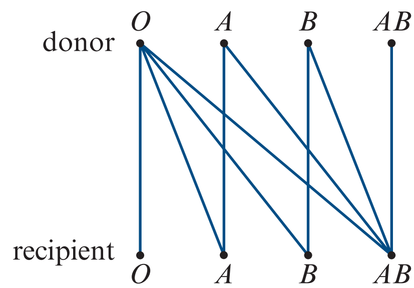
B



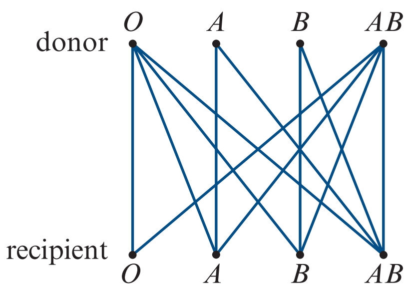
C



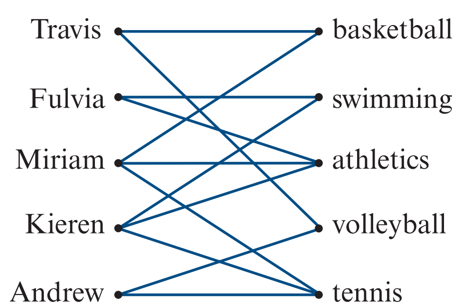
D



E



5 A group of five students represent their school in five different sports. The information is displayed in a bipartite graph. From this graph we can conclude that:



- **A** Travis and Miriam played all the sports between them.
- **B** In total, Miriam and Fulvia played fewer sports than Andrew and Travis.
- **C** Kieren and Miriam each played the same number of sports.
- **D** In total, Kieren and Travis played fewer different sports than Miriam and Fulvia.
- **E** Andrew played fewer sports than any of the others.

ISBN 978-1-108-46154-2

Cambridge University Press

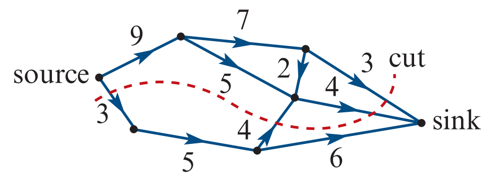
6 Five people are to be each allocated one of five tasks (*A*, *B*, *C*, *D*, *E*). The table shows the time, in hours, that each person takes to complete the tasks. The total time to complete all the tasks is to be minimised. If no person can help another, Francis should be allocated task:

Name	A	B	C	D	E
Francis	12	15	99	10	14
David	10	9	10	7	12
Herman	99	10	11	6	12
Indira	8	8	12	9	99
Natalie	8	99	9	8	11

A *A* **B** *B* **C** *C* **D** *D* **E** *E*

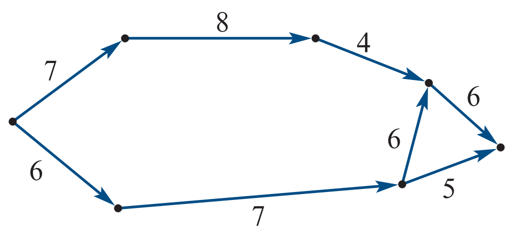
7 For the flow network shown on the right, the capacity of the cut is:

A 3 **B** 6 **C** 9
D 10 **E** 14



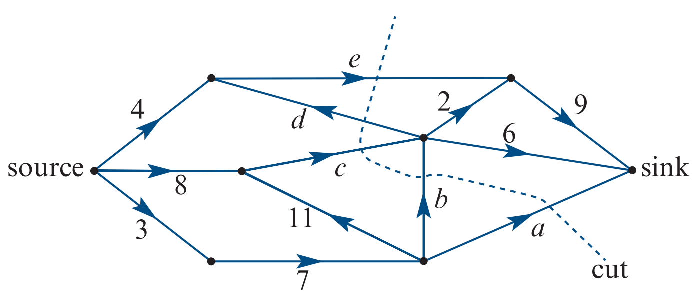
8 The maximum flow, from source to sink, in the flow network shown to the right is:

A 10 **B** 11 **C** 12
D 13 **E** 14



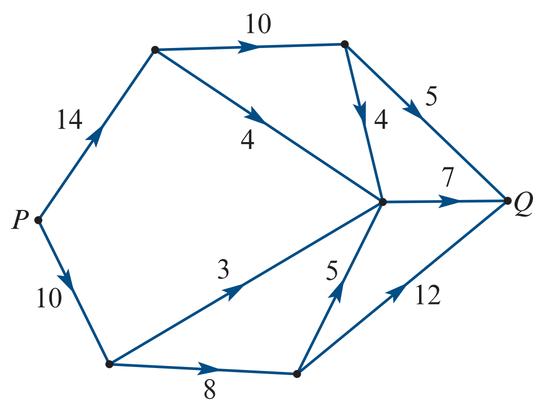
9 In the flow network to the right, the weight of each edge is non-zero. The capacity of the cut shown is:

A $a + b + c + d + e$
B $a + c + d + e$
C $a + b + c + e$
D $a + b + c - d + e$
E $a - b + c - d + e$



10 The flow network to the right shows the capacity of data flow along cables in Megabits per second. What is the maximum flow of data, in Megabits per second, from server *P* to server *Q*?

A 20 **B** 22 **C** 23
D 24 **E** 30



ISBN 978-1-108-46154-2

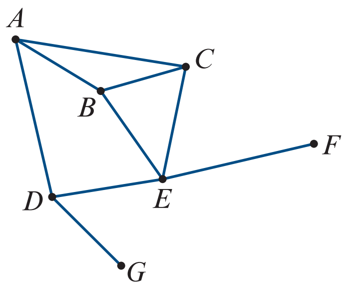
Cambridge University Press

Review

SF

Short-answer questions

1 Consider the graph shown on the right.

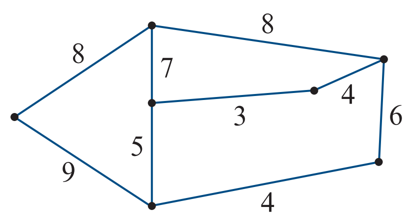


a How many edges must be removed in order to leave a spanning tree?

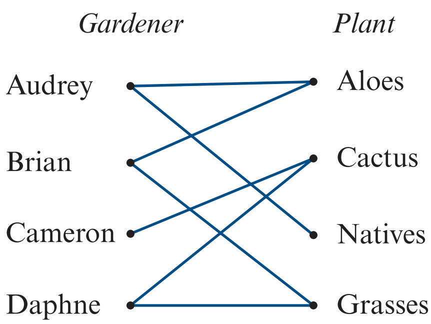
b Two of the edges in this graph must be in every spanning tree. Between which vertices are these edges?

c Remove some edges to form two different spanning trees for this graph.

2 Determine the minimum spanning tree for the network shown on the right.



3 The bipartite graph on the right shows the gardeners in a botanical garden and the plant types that they have experience caring for.



a How many gardeners have experience caring for Aloes?

b Which gardener is the only gardener who has experience caring for natives?

c The head gardener will assign Audrey, Brian, Cameron and Daphne to one plant type each. How must the plant types be allocated to gardeners?

4 Isla has four employees, David, Robyn, Linda and Anthony. She needs to assign each of these people to one of four tasks: Word processing, Editing, Printing, Mailing. Isla has applied the Hungarian algorithm to determine how these tasks should be allocated so that the least overall time is spent completing all the tasks. The result of this is shown in the table above.

	Word processing	Editing	Printing	Mailing
David	5	1	0	2
Robyn	4	0	2	0
Linda	2	0	3	2
Anthony	0	3	5	1

a Construct a bipartite graph from this table.

b Determine the allocation of tasks that Isla should make.

ISBN 978-1-108-46154-2

Cambridge University Press

5 Steve is a supervisor in a furniture factory. An order for chairs needs to be filled. The table below shows the time, in hours, it would take each of four employees (Julia, Mario, Sylvana, George) to perform each of four tasks (cutting the pieces, assembling the chairs, preparing the chairs for painting, painting the chairs) required to complete the chairs for this order.

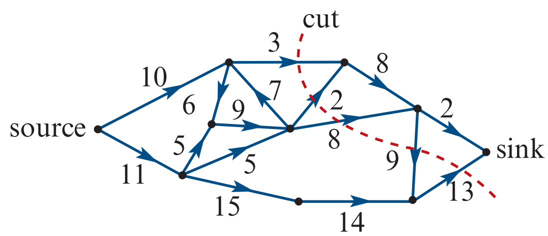
	Cutting	Assembling	Preparing	Painting
Julia	8	4	3	4
Mario	6	6	7	5
Sylvana	8	6	4	6
George	5	8	5	6

a Which employee would take the shortest time to paint the chairs?

b Which employee would take the longest time to prepare the chairs for painting?

c Apply the Hungarian algorithm to determine the allocation of employee to task so that the overall minimum time is taken to fill the order.

6 In the flow network opposite, the values on the edges give the maximum flow possible between each pair of vertices. The arrows show the direction of flow in the network. Also shown is a cut that separates the source from the sink.

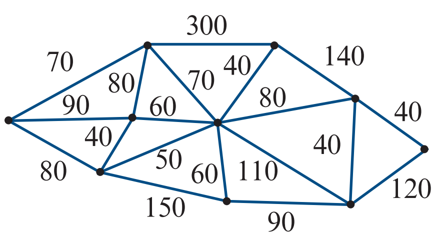


a Determine the capacity of the cut shown.

b Determine the maximum flow through this network.

Extended-response questions

1 In the network opposite, the vertices represent water tanks on a large property and the edges represent pipes used to move water between the tanks. The numbers on each edge indicate the lengths of pipes (in m) connecting different tanks.
Determine the shortest length of pipe needed to connect all water storages.



ISBN 978-1-108-46154-2

Cambridge University Press

2 Bernard, Georgia, Chris and Arthur are student pilots. Their flying instructor, Terry, has four lesson appointments available on a particular Saturday (9 a.m., 10 a.m., 1 p.m. and 3 p.m.).

- Bernard can only fly at 10 a.m.
- Georgia can fly at any time before midday.
- Chris can fly at 9 a.m. and then any time after 11 a.m.
- Arthur can fly any time after 2 p.m.

a Draw a bipartite graph with the student pilots on the left and the times on the right.

b Which student pilot will Terry be teaching at 1 p.m.?

c Write down the appointment times for each of the four student pilots.

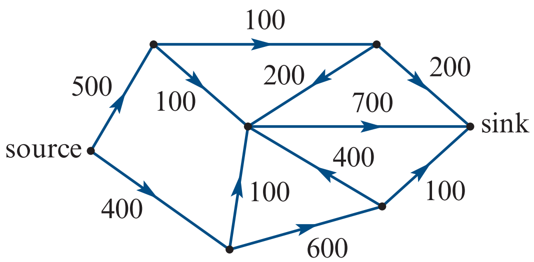
3 Ann, Bianca, Con and David are four examination supervisors. There are four examinations venues: *B*, *C*, *D* and *E*. Each examination venue requires one examination supervisor.

The table shows the times (in minutes) that the examination supervisors would take to travel from their home to each examination venue.

Determine the allocation of examination supervisor to examination venue that will minimise the overall travel time for the supervisors.

Supervisor	B	C	D	E
Ann	25	30	15	35
Bianca	22	34	20	45
Con	32	20	33	35
David	40	30	28	26

4 The flow network on the right shows the maximum rate of water flow (in litres per minute) through a system of water pipes. The water flows from the source to the sink.



a Determine the maximum amount of water that can flow from the source to the sink through this network of pipes.

b How many litres of water will flow into the sink over the 2 hours?

c A tank with capacity 2700 L is placed at the sink. How long will it take this tank to fill?

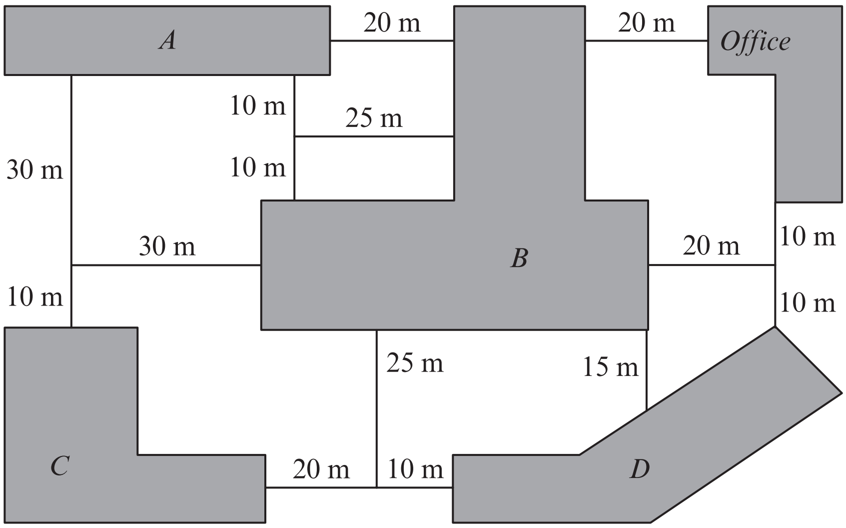
ISBN 978-1-108-46154-2

Cambridge University Press

Review

CU

5 The diagram below shows the buildings of a new university. The lines on the diagram show the location of the pathways between the buildings.



a **i** How many different ways can a student walk directly from building A to building B?

ii Represent this diagram as a weighted graph in planar form.

iii Which buildings are immediately adjacent to building C?

Some of the pathways will be covered to protect students from the rain as they move between buildings. The covering structure will cost $240 per metre to make and install.

b The shortest direct pathway between each building and its adjacent buildings will be covered.

i Modify your planar weighted graph from question **aii** above to show only the shortest direct pathway between adjacent buildings.

ii How much will the covered walkways on these pathways cost to build?

c It has been decided that covering all of these walkways is too expensive. Only the minimum number of pathways that are necessary to allow students to walk from one building to any other while remaining under cover will be built.

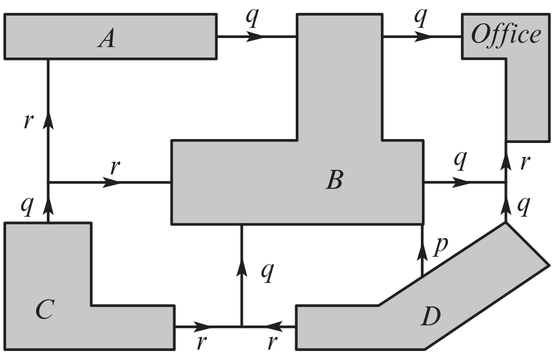
i Draw the graph that shows the pathways that should be covered so that the overall cost of making and installing the covering structure is a minimum.

ii Calculate the cost of the covering structure in part **i**.

ISBN 978-1-108-46154-2

Cambridge University Press

d In emergency situations, some of the doors in building *B* are locked and students are directed to evacuate the university via other pathways. The diagram below shows the locations of these evacuation pathways.



The pathways and doors allow different rates of students to flow along per minute. On the diagram:

- p = flow rate of 80 students per minute
- q = flow rate of 120 students per minute
- r = flow rate of 150 students per minute

  i If there are 675 students in building *D* when the alarm bell rings, what is the minimum time it could take all students to leave this building? Assume there are no students in the other buildings.

  ii Evening school is held in building *C*. On a particular evening there were 840 students in building *C*. Assume that there were no other students in any of the other buildings. How long would it take to evacuate all of these students through the office building?

ISBN 978-1-108-46154-2

Cambridge University Press

Chapter 12

12

Project planning and scheduling

UNIT 4 INVESTING AND NETWORKING

Topic 3 Networks and decision mathematics

- How do we plan for a project?
- How do we identify predecessors of an activity?
- How do we draw an activity network and use it to plan for a project?
- How do we account for float times in our project?
- How do we find the earliest starting time and latest finishing time for an activity in a project?

ISBN 978-1-108-46154-2

Cambridge University Press

12A Project planning – precedence tables and activity networks

Project planning

Building a house, manufacturing a product, and organising a wedding are all examples of a *project*, that is a task that involves a number of individual steps, or *activities*, that must be completed. The individual activities often rely upon each other and some cannot be performed until others are completed.

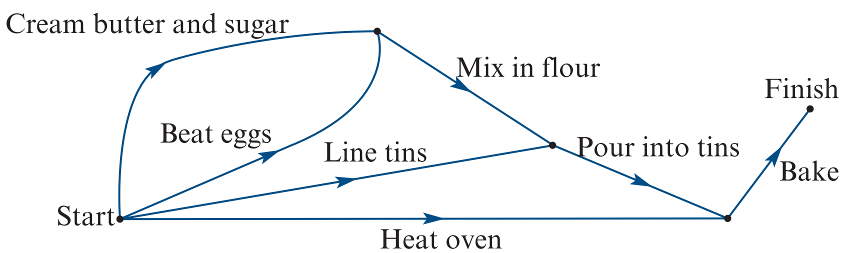
For example, in the organisation of a wedding, invitations would be sent out to guests, but a plan for seating people at the tables during the reception cannot be completed until the invitations have been accepted. When building a house, the plastering of the walls cannot begin until the house has been sealed from the weather.

Project planning involves the analysis of the requirements of each of the activities of a project to determine the order in which they must be completed.

Activity networks

Projects are represented using a directed graph called an **activity network**. Activity networks have a vertex labelled *start* and another labelled *finish*. Each **activity** within the project is represented by an edge and so in activity networks it is the edges that must be labelled, not the vertices. The edges are arranged to display the order in which activities must be completed. Activity networks do not have multiple edges.

This activity network shown below represents the project of making a sponge cake. There is an edge labelled for each of the steps of the recipe. The vertices have not been labelled, except for *start* and *finish*.



The edges for activities 'Cream butter and sugar' and 'Beat eggs' both end at the vertex where the activity 'Mix in flour' begins. This shows that the butter and sugar must be creamed and the eggs beaten before the flour can be mixed in. The activity 'Mix in flour' cannot begin until the other two activities are completed. Similarly, the activity 'Pour into tins' cannot begin until the activity 'Line tins' is also completed. Finally, the activity 'Bake' cannot begin until all the other activities, including 'Heat oven', are completed.

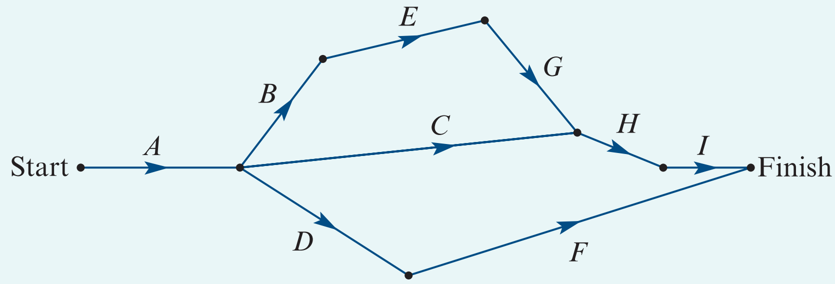
Activity networks show the *precedence* that activities have over each other. The activity 'Mix in flour' must be completed before the activity 'Pour into tins' can begin and so 'Mix in flour' is called an **immediate predecessor** of activity 'Pour into tins'.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 1 Interpreting activity networks

The activity network for a project is shown below.



a How many activities are involved in this project?

b Which activity is an immediate predecessor of activity F?

c Activity B is an immediate predecessor of which activity?

d How many immediate predecessors does activity H have?

Solution

a Count the number of edges in the network. *This project has 9 activities.*

b An immediate predecessor of activity F ends at the vertex at which activity F begins. *Activity D is an immediate predecessor of activity F.*

c Any activity that begins at the same vertex that activity B ends on has activity B as an immediate predecessor. *Activity B is an immediate predecessor of activity E.*

d Count the number of activities that end at the vertex at which activity H begins. *Activity H has two immediate predecessors, G and C.*

Precedence tables

The immediate predecessors for all activities in a project can be recorded in a **precedence table**. The precedence table on the right is for a project that has seven activities, A, B, C, D, E, F and G.

Activity C has only one immediate predecessor, activity A.

Activity B is an immediate predecessor of both activity D and activity E.

Activity F has two immediate predecessors, activities C and D.

Activity	Immediate predecessors
A	–
B	–
C	A
D	B
E	B
F	C, D
G	E, F

Activities A and B have no immediate predecessors, as indicated by '–'. This means that they can begin at the 'start' vertex; that is, at the very start of the whole project.

ISBN 978-1-108-46154-2

Cambridge University Press

Activity networks and precedence tables

A project is made up of individual activities.

An activity network has edges that represent the activities of a project. The vertices of an activity network are not labelled, except for the *start* and *finish* vertices.

Activity networks do not have multiple edges.

When activity M must be completed before activity N begins, activity M is called an immediate predecessor of activity N.

A precedence table records all of the immediate predecessors for each activity of a project.

Example 2 Drawing an activity network from a precedence table

Draw an activity network from the precedence table shown below.

Activity	Immediate predecessors
A	–
B	A
C	A
D	A
E	B
F	C
G	D
H	E, F, G

ISBN 978-1-108-46154-2

Cambridge University Press

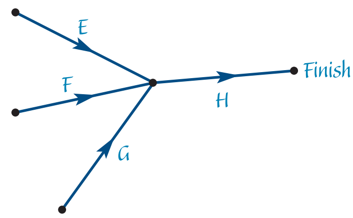
Solution

The activity network can be drawn in any order, beginning with any activity.

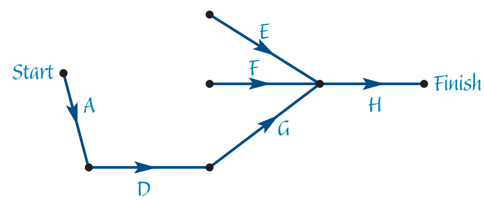
In this solution, the activity network will be drawn from the finishing vertex back to the starting vertex.

H is not an immediate predecessor for any other activity so it will lead to the finish of the project.

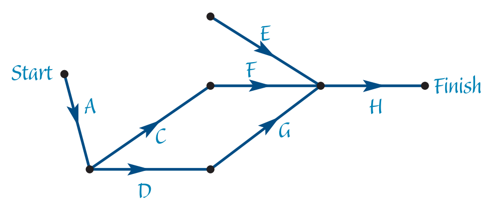
H has immediate predecessors E, F and G and so these three activities will lead into activity H.



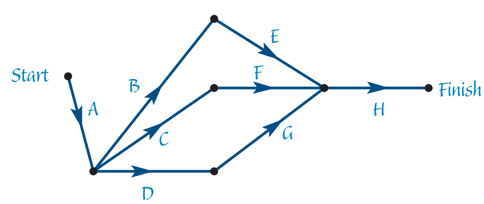
Activity D is an immediate predecessor of activity G *and* has immediate predecessor activity A. There will be a path through activity A, activity D and then activity G.



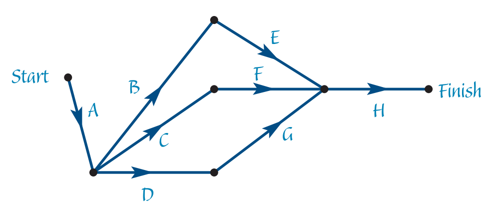
Activity C is an immediate predecessor of activity F *and* has immediate predecessor activity A. There will be a path through activity A, activity C and then activity F.



Activity B is an immediate predecessor of activity E *and* has immediate predecessor activity A. There will be a path through activity A, activity B and then activity E.



Activity A has no immediate predecessors, so it is the start of the project.



ISBN 978-1-108-46154-2

Cambridge University Press

Dummy activities

Although activity networks do not have multiple edges between vertices, there are instances where the conditions of the project imply that multiple edges are required. A **dummy activity** must be used in these circumstances. This is usually indicated by activities that share some immediate predecessors, but not all of them.

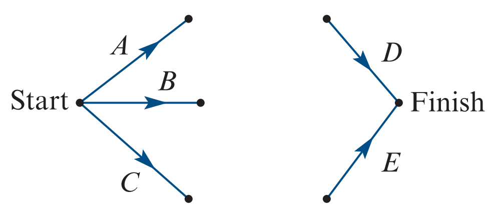
Activity	Immediate predecessors
A	–
B	–
C	–
D	*A*, *B*
E	*B*, *C*

In this very simple precedence table, activity *D* and activity *E* share the immediate predecessor activity *B*, but they both have an immediate predecessor activity that the other does not.

This overlap of predecessors presents some difficulty when constructing the activity network, but this difficulty is easily overcome.

Activity *D* and activity *E* are not immediate predecessors for any other activity, so they will lead directly to the finish vertex of the project.

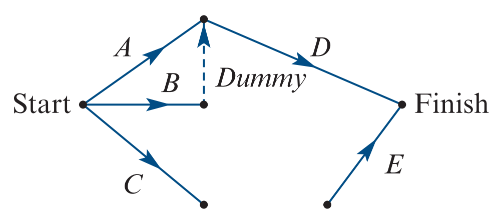
Activites *A*, *B* and *C* have no immediate predecessors, so they will follow directly from the start vertex of the project.



The start and finish of the activity network are shown in the diagram above. We need to use the precedence information for activity *D* and activity *E* to join these two parts together.

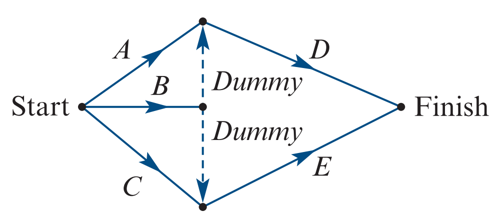
Activity *D* needs to follow directly from activity *A* and activity *B*, but we can only draw one edge for activity *D*. Activity *E* needs to follow directly from both activity *B* and activity *C*, but again we only have one edge for activity *E*, not two.

The solution is to draw the diagram with activity *D* starting after one of its immediate predecessors, and using a dummy activity for the other. The dummy activities are represented by dotted edges and are, in effect, imaginary. They are not real activities, but they allow all of the predecessors from the table to be correctly represented.



The dummy activity for *D* allows activity *D* to directly follow both activity *A* and *B*.

A dummy activity is also needed for activity *E* because it, too, has to start after two different activities, activity *B* and *C*.



ISBN 978-1-108-46154-2

Cambridge University Press

Dummy activities

A *dummy activity* is required if two activities share some, but not all, of their immediate predecessors. This can avoid the activity network having multiple edges between two vertices.

A dummy activity will be required from the end of each shared immediate predecessor to the start of the activity that has additional immediate predecessors.

Dummy activities are represented in the activity network using dotted lines.

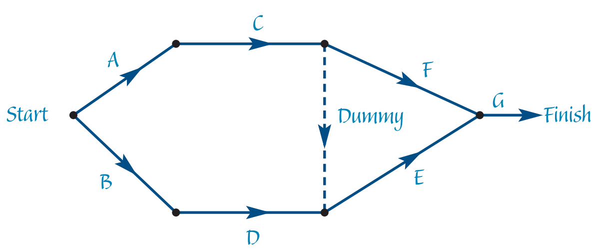
Example 3 Constructing an activity network that requires a dummy activity

Draw an activity network from the precedence table shown on the right.

Activity	Immediate predecessors
A	–
B	–
C	A
D	B
E	C, D
F	C
G	E, F

Solution

- A and B will lead from the start vertex.
- G will lead to the end vertex.
- A dummy will be required from the end of activity C (shared immediate predecessor) to the start of activity E (the activity with an additional immediate predecessor.

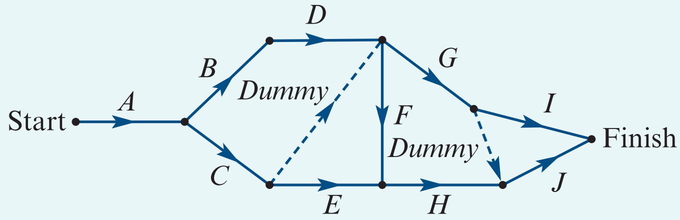


ISBN 978-1-108-46154-2

Cambridge University Press

Example 4 Creating a precedence table from an activity network involving dummy activities

Write a precedence table for the activity network shown below.

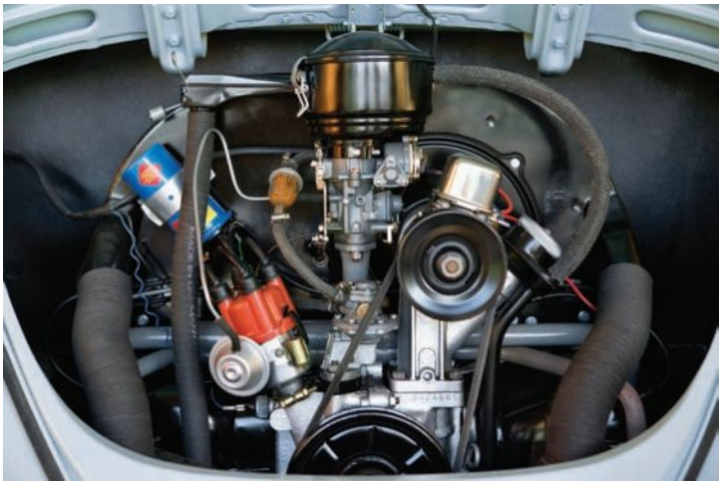


Solution

1 Create a table with a row for each activity.
2 Look at the start of an activity. In the immediate predecessor column, write down all of the activities that lead directly to this activity.
3 Activity *C* is a predecessor of activity *E* and the dummy shows that it is also a predecessor of activities *F* and *G*.
4 Activity *G* is a predecessor of activity *I* and the dummy shows that it is also a predecessor of activity *J*.

Note: The dummy activity is not included in the precedence table. The dummy is there only to ensure all precedence requirements are met and to avoid the need for multiple edges. It is not an activity in the project.

Activity	Immediate predecessors
A	–
B	A
C	A
D	B
E	C
F	D, C
G	D, C
H	E, F
I	G
J	G, H



ISBN 978-1-108-46154-2

Cambridge University Press

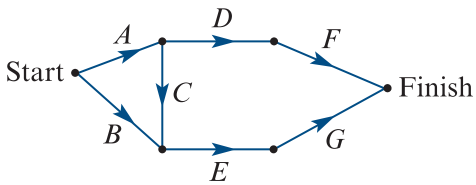
Exercise 12A

Writing precedence tables from activity networks

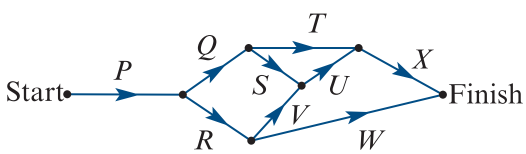
Example 1

1 Write a precedence table for the activity networks shown below.

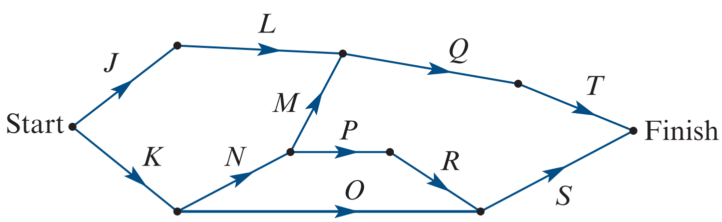
a



b

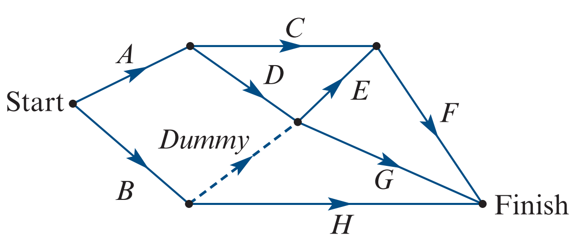


c

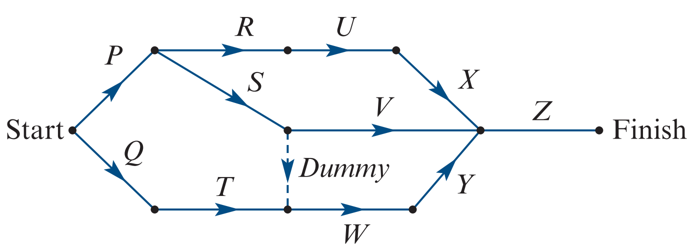


Example 4

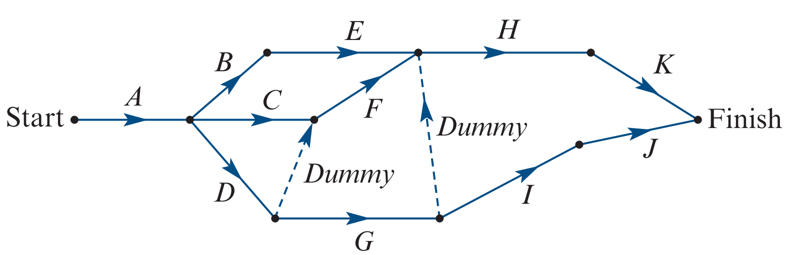
d



e



f



SF

ISBN 978-1-108-46154-2

Cambridge University Press

Drawing activity networks from precedence tables

Example 2

2 Draw an activity network from the precedence tables below.

SF

a

Activity	Immediate predecessors
A	–
B	A
C	A
D	B
E	C

b

Activity	Immediate predecessors
P	–
Q	–
R	P
S	Q
T	R, S

c

Activity	Immediate predecessors
T	–
U	–
V	T
W	U
X	V, W
Y	X
Z	Y

d

Activity	Immediate predecessors
F	–
G	–
H	–
I	F
J	G, I
K	H, J
L	K

Example 3

3 Draw an activity network from the precedence tables below. Dummy activities will be required.

CF

a

Activity	Immediate predecessors
F	–
G	–
H	F
I	H, G
J	G

b

Activity	Immediate predecessors
A	–
B	A
C	A
D	B
E	B, C

c

Activity	Immediate predecessors
P	–
Q	–
R	P
S	Q
T	Q
U	R, S
V	R, S, T

d

Activity	Immediate predecessors
A	–
B	A
C	A
D	B, C
E	C
F	E
G	D
H	F, G

ISBN 978-1-108-46154-2

12B Scheduling problems

Projects that involve multiple activities are usually completed against a time schedule. Knowing how long individual activities within a project are likely to take allows managers of such projects to hire staff, book equipment and also to estimate overall costs of the project. Allocating time to the completion of activities in a project is called scheduling. *Scheduling problems* involve analysis to determine the minimum overall time it would take to complete a project.

▶ Weighted precedence tables

Weighted precedence tables show the completion times, or durations, of each of the activities that make up a particular project.

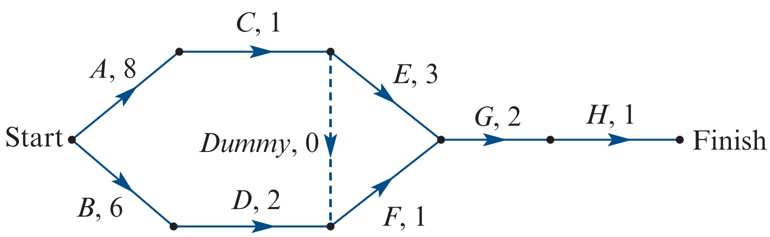
The precedence table on the right shows the activities of a project, the duration of each activity and the immediate predecessors of each of the activities.

The durations are recorded on the activity network that is drawn from the precedence table. It is usual to record the name of the activity followed by a comma and then the duration of that activity on the edge that represents it.

Activity	Duration (days)	Immediate predecessors
A	8	–
B	6	–
C	1	A
D	2	B
E	3	C
F	1	C, D
G	2	E, F
H	1	G

For example, Activity D would be labelled D, 2.

The activity network for this project is shown below. This project requires the use of a dummy activity. Dummy activities are always considered to have a duration of zero.



The weight (duration) of a dummy activity is always zero.

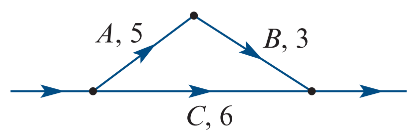
ISBN 978-1-108-46154-2

Cambridge University Press

Float times

The diagram below shows a small section of an activity network. Three activities are shown, A, B, C along with their individual durations, in hours.

Activity A and B form a small sequence of activities. Activity B cannot begin until activity A has finished. The minimum time it would take to complete activity A and B would be $5 + 3 = 8$ hours.



Activity C can begin at the same time as activity A, but must be completed no later than activity B. The activity network shows that activity C can be completed at the same time as the sequence of activites $A – B$. Activity C has a duration of 6 hours, which is two hours less than the time for the sequence $A – B$ and so there is some flexibility around when activity C could start. This value is called the **float time** for activity C. The float time is sometimes called the *slack time*.

The flexibility around the timing of activity C is shown in the diagram below.

	A	A	A	A	A	B	B	B
Start at same time	C	C	C	C	C	C	Slack	Slack
Delay C by 1 hour	Slack	C	C	C	C	C	C	Slack
Delay C by 2 hour	Slack	Slack	C	C	C	C	C	C

The five red squares represent the 5 hours it takes to complete activity A. The three green squares represent the 3 hours it takes to complete activity B.

This six yellow squares represent the 6 hours it takes to complete activity C. Activity C does not have to start at same time as activity A because it has some slack time available (2 hours).

Activity C should not be delayed by more than 2 hours because this would cause delays to the project. The next activity requires B and C to be complete before it can begin.

Critical path analysis

Scheduling problems are concerned with minimising the total time it takes to complete a project and so it is essential that all the activities in a project begin at the earliest possible time.

Critical path analysis is the process of analysing the timing of activities of a project to find the overall minimum time to complete the project. It also involves identifying those activities that have float time and those that do not. Activities that have no float time are said to be **critical** activities because if they are delayed, the whole project will not finish in the minimum possible time.

ISBN 978-1-108-46154-2

Cambridge University Press

Critical path analysis begins with determining the **earliest starting time (EST)** and the latest finishing time (LFT) for each activity.

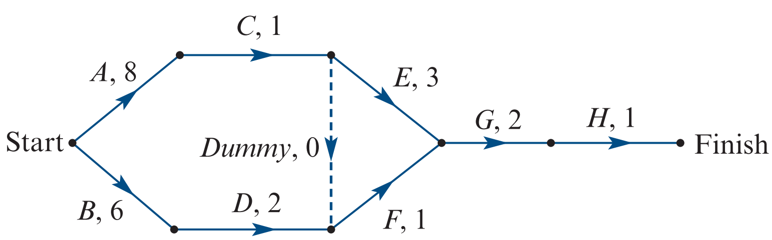
EST values indicate the earliest possible time after the start of the project that a particular activity can begin and still allow the project to be completed in minimum time. For example, an EST of 8 hours means that the activity can begin, at the earliest, 8 hours after the start of the project.

LFT values indicate the latest possible time after the start of the project that a particular activity can finish and still allow the project to be completed in minimum time. For example, an LFT of 14 hours means that at the very latest, the activity can finish 14 hours after the start of the project.

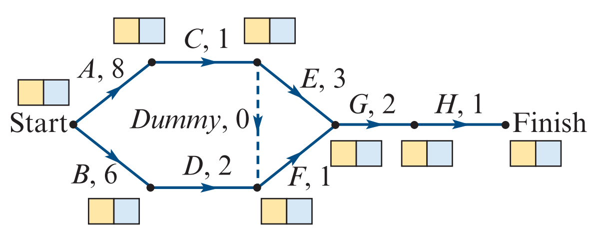
EST values are determined using a process called **forward scanning**.

Forward scanning

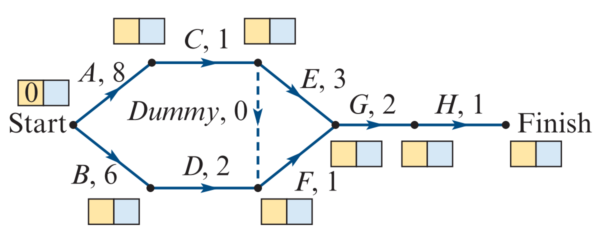
Forward scanning will be demonstrated using the activity network shown below.



1 Draw a box, split into two cells, next to each vertex of the activity network, as shown. The cells in the boxes in the diagram are coloured yellow and blue to help identify which cell we are using. The left cell (yellow) at any vertex will contain the EST for any activity that begins at that vertex.



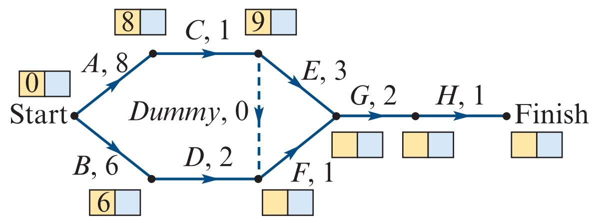
2 Put a zero (0) in the left cell (yellow) of the box at the start vertex. This represents the start of the entire project. It also represents the EST for activities *A* and *B* because they start at this vertex.



ISBN 978-1-108-46154-2

Cambridge University Press

3 Each activity will have a box at the vertices at either end of the edge that represents that activity. Take the left cell (yellow) value of the box at the start of the activity, add it to the duration of the activity and write the answer in the left cell (yellow) of the box at the end of the activity. This is the EST for the activity or activities that follow.



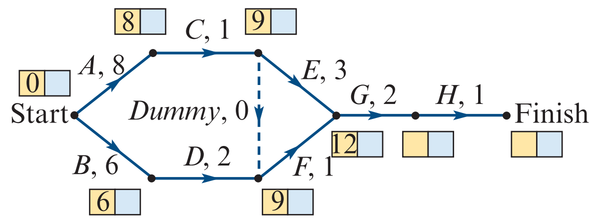
Notes:

1 The cell at the end of activity *A* has value 0 + 8 = 8. This is the EST for activity *C*.

2 The cell at the end of activity *B* has value 0 + 6 = 6. This is the EST for activity *D*.

3 The cell at the end of activity *C* has value 8 + 1 = 9. This is the EST for activity *E* and dummy.

4 If the edges representing more than one activity end at the same vertex, the left cell (yellow) of the box at this vertex must contain the *largest* of the possible values because this activity must wait for all predecessor activities to be completed before it can begin.



Notes:

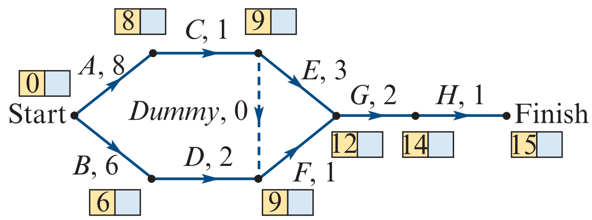
1 The cell at the end of activity *D* and dummy could be:
- from activity *D*: 6 + 2 = 8
- from dummy activity: 9 + 0 = 9

The largest of these options is 9. This is the EST for activity *F*.

2 The cell at the end of activity *E* and *F* could be:
- from activity *E*: 9 + 3 = 12
- from activity *F*: 9 + 1 = 10

The largest of these options is 12. This is the EST for activity *G*.

5 Continue adding the previous EST value to the duration to calculate the following EST values until the final cell is reached.



Identifying minimum project completion time

The box for the finish vertex above contains the EST for the next activity, but there is no activity left to begin. So, this value represents the overall minimum completion time for the entire project. This project can be completed in a minimum of 15 days.

To complete the analysis, latest finishing time (LFT) values need to be calculated for each activity.

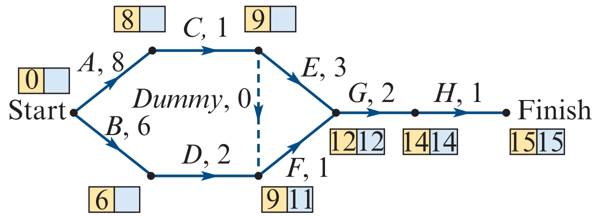
LFT values are determined using a process called **backward scanning**.

ISBN 978-1-108-46154-2

Cambridge University Press

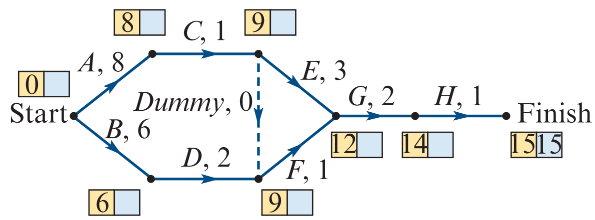
Backward scanning

1 Look at the box at the finish vertex. Copy the value in the left cell (yellow) to the right cell (blue). The right cell (blue) contains the latest finishing time (LFT) for any activity that ends at this vertex.



The LFT for activity H is 15.

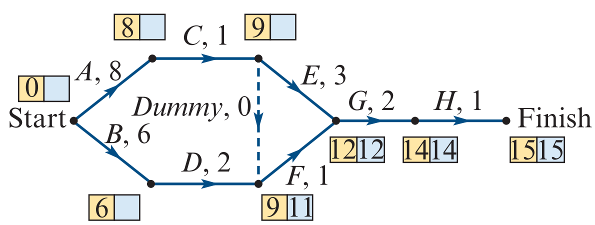
2 For each activity, take the LFT from the right cell (blue) value of the box at the end of the activity and subtract the duration of the activity. Write the answer in the right cell (blue) of the box at the start of the activity. This will give the LFT for any activity that ends at this vertex.



Notes:

1. The cell at the start of activity H has value $15 - 1 = 14$. This is the LFT for activity G.
2. The cell at the start of activity G has value $14 - 2 = 12$. This is the LFT for activity E and F.
3. The cell at the start of activity F has value $12 - 1 = 11$. This is the LFT for activity D and dummy.

3 If the edges representing more than one activity start at the same vertex, the right cell (blue) of the box at this vertex must contain the smallest of the possible values. This ensures that the longest of the activities that follow this vertex will have time to be completed.



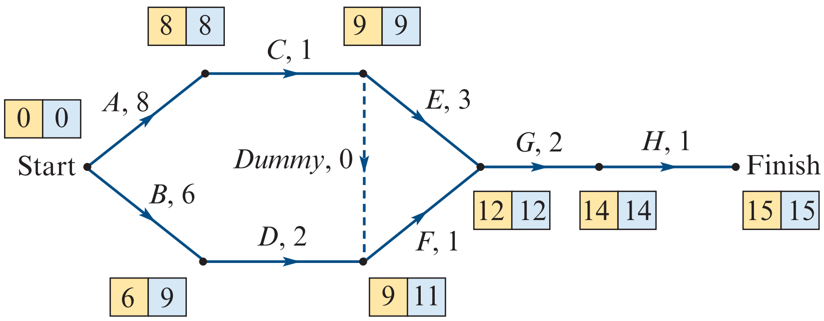
Note:

1. The cell at the start of activity E and dummy could be:
   - from activity E: $12 - 3 = 9$
   - from dummy activity: $11 - 0 = 11$

   The smallest of these options is 9. This is the LFT for activity C.

ISBN 978-1-108-46154-2

4 Complete the backward scanning for all remaining activities.



Notes:

1 The cell at the start of activity C has value $9 - 1 = 8$. This is the LFT for activity A.

2 The cell at the start of activity D has value $11 - 2 = 9$. This is the LFT for activity B.

3 The cell at the start vertex has value, either:

- from activity A: $8 - 8 = 0$
- from activity B: $9 - 6 = 2$

The smallest of these options is 0.

4 The box at the start vertex will always contain a zero in both the left (yellow) and right (blue) cells.

Determining latest starting time (LST)

The forward scanning process identified the EST, earliest starting time, for each activity.

The backward scanning process identified the LFT, latest finishing time, for each activity.

The **latest starting time (LST)** for an activity is the latest possible time it could start after the beginning of the project and still allow the project to be completed in minimum time.

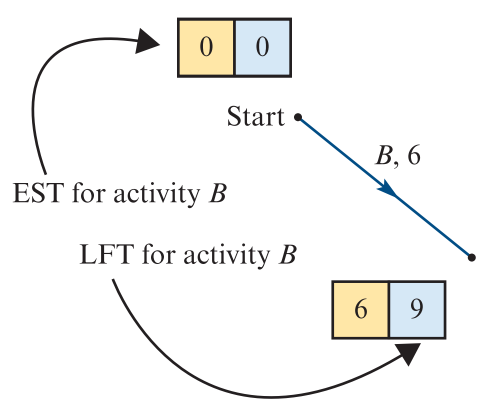
The LST for an activity is determined by a simple calculation:

$$\text{LST} = \text{LFT} - \text{duration}$$

Activity B from the completed forward and backward scanning process above is shown in the diagram on the right.

The EST for activity B is in the left cell (yellow) in the box at the vertex where activity B begins.

The LFT for activity B is in the right cell (blue) in the box at the vertex where activity B ends.



The LST for activity B = LFT – duration
$= 9 - 6$
$= 3$ days

ISBN 978-1-108-46154-2

Cambridge University Press

Activity B must be completed, at the latest, after 9 days. Since it has a duration of 6 days, it can begin 3 days after the start of the project and still finish in time. The latest time it can start (LST) is 3 days.

The LST for all activities in the project can be found using similar calculations, the results of which are shown in the table below.

LST $A = 8 - 8 = 0$
LST $B = 9 - 6 = 3$
LST $C = 9 - 1 = 8$
LST $D = 9 - 2 = 7$
LST $E = 12 - 3 = 9$
LST $F = 12 - 1 = 11$
LST $G = 14 - 2 = 12$
LST $H = 15 - 1 = 14$

Activity	Duration	EST	LFT	LST
A	8	0	8	0
B	6	0	9	3
C	1	8	9	8
D	2	6	9	7
E	3	9	12	9
F	1	9	12	11
G	2	12	14	12
H	1	14	15	14

Determining float time

Some of the activities from the project above have the same EST and LST values in the table. The earliest start time and the latest start time are exactly the same, which means that there really is only one start time that is possible in order to make sure the project is completed in minimum time.

These activities are the critical activities described earlier. They have no flexibility in their starting time and so have a float time of zero. The critical activities in this project are A, C, E, G and H.

The non-critical activities B, D, F have float time that is not zero. Float time can easily be calculated using this rule.

Float time $= LST - EST$

Float time for activity $B = 3 - 0 = 3$

Float time for activity $D = 7 - 6 = 1$

Float time for activity $F = 11 - 9 = 3$

Identifying the critical path

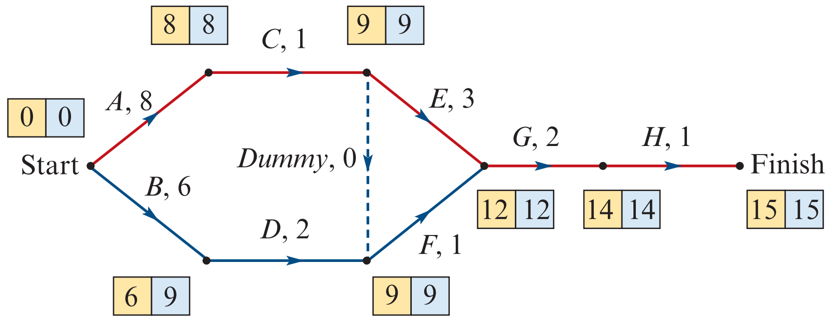
The critical activities have already been identified as those activities that have zero float time; that is, those activities that have equal EST and LST.

The critical path through the network is the sequence of these critical activities, from the start of the project through to the finish.

ISBN 978-1-108-46154-2

Cambridge University Press

The critical path has been highlighted in red on the diagram opposite.



In most projects, there will be a single critical path from start to finish, but it is possible for a project to have a critical path that branches. For example, if the time for activity F was increased to 3 days in the project above, then the LST for F would become $12 - 3 = 9$ and so F and the dummy activity would also be on the critical path.

Critical path analysis

Critical path analysis is the process of analysing the timing of activities in a project to determine the critical path of a project.

The *critical path* of a project is the sequence of activities that cannot be delayed without affecting the overall completion time of the project.

Perform critical path analysis by:

- drawing a box with two cells next to each vertex of the activity network
- *forward scanning* to identify *EST* for each activity
  - add the left cell value at the start of the activity to the duration
  - write the result in the left cell at the end of the activity
  - use the largest of the possibilities if there is more than one activity ending at the same vertex
- identifying the minimum project completion time as the left cell value at the finish vertex
- *backward scanning* to identify *LFT* for each activity
  - subtract the duration from the right cell value at the end of the activity
  - write the result in the right cell at the start of the activity
  - use the smallest of the possibilities if there is more than one activity beginning at the same vertex.

EST values are in the left cell at the start of each activity.

LFT values are in the right cell at the end of each activity.

LST = LFT − duration

Float = LST − EST

Critical activities will have zero float; that is, LST = EST

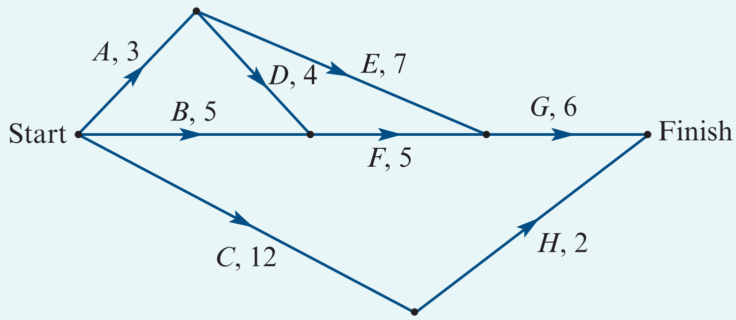
The critical path is the sequence of critical activities through the activity network.

ISBN 978-1-108-46154-2

Cambridge University Press

Example 5 Critical path analysis

The activity network for a project consisting of eight activities, A, B, C, D, E, F, G and H is shown below. The number next to the activity name is the time it takes, in weeks, to complete that activity.



a Complete the forward scanning process to identify the minimum time it will take to complete this project.

b Complete the backward scanning process.

c What is the earliest starting time for activity E?

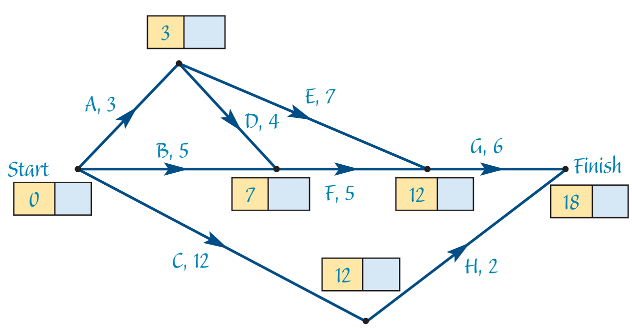
d What is the latest starting time for activity E?

e Identify the critical path for this project.

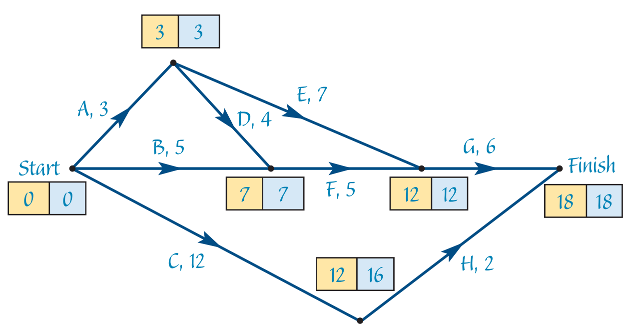
f The person responsible for completing activity E falls sick three weeks into the project. If he will be away from work for two weeks, will this cause the entire project to be delayed?

Solution

a The forward scanning process results are shown in the diagram.

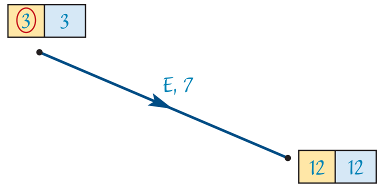


b The backward scanning process results are shown in the diagram.



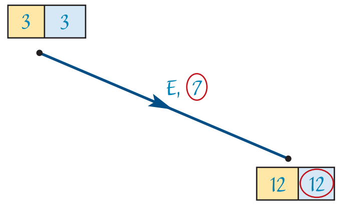
ISBN 978-1-108-46154-2

c Earliest starting time is the left cell value of the box at the beginning of the activity.



The EST for activity E = 3

d Latest starting time is found by subtracting the duration of the activity from the right cell box at the end of the activity.



The LST for activity E = 12 − 7 = 5

e The critical path joins all of the activities that have the same EST and LST, and therefore which have zero float time.

The critical path for this project is A − D − F − G.

f **1** Calculate the float for activity *E*. This tells us how long the start of activity *E* can be delayed, without delaying the entire project.

Float E = LST − EST
= 5 − 3
= 2 weeks

2 If the float time is more or equal to the delay in the start of activity *E*, the project will not be affected.

The person will be away for two weeks, starting 3 weeks into the project. This is equal to the float time for activity E, and so delaying the start of activity E until the person comes back to work will not affect the overall completion time of the project.

ISBN 978-1-108-46154-2

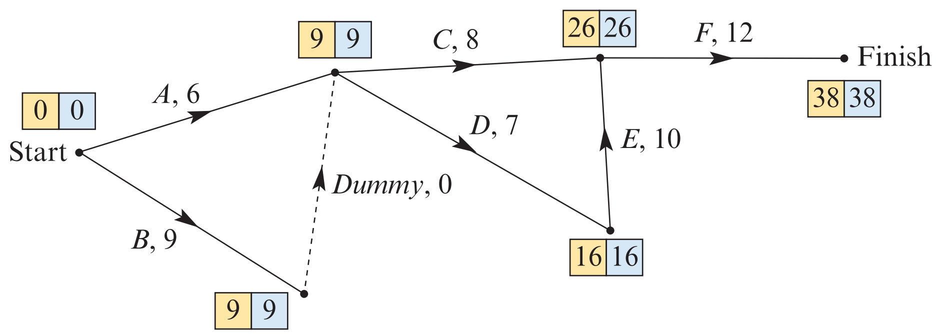
Cambridge University Press

Exercise 12B

Skillsheet Interpreting completed forward and backward scanning

Example 5

1 The activity network below shows the results of forward and backward scanning.



a What is the minimum completion time for this project?

b What is the duration of activity *C*?

c What is the EST of activity *C*?

d What is the LFT of activity *C*?

e **i** Calculate the LST of activity *C*.

ii Explain how you know that activity *C* is not a critical activity.

f There is one other non-critical activity in the project. Which one is it?

2 Consider the section of an activity network shown in the diagram below.

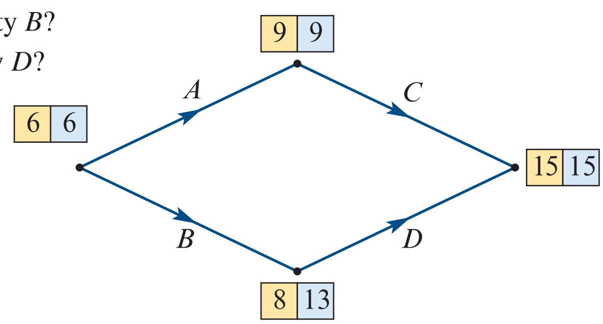
a What is the duration of activity *A*?

b What is the float time of activity *B*?

c What is the duration of activity *D*?

d What is the latest time that activity *D* can start?

e Write down the critical path through this section of the activity network.



Critical path analysis

3 Consider the activity network in the diagram shown below.

a Complete forward and backward scanning for this activity network.

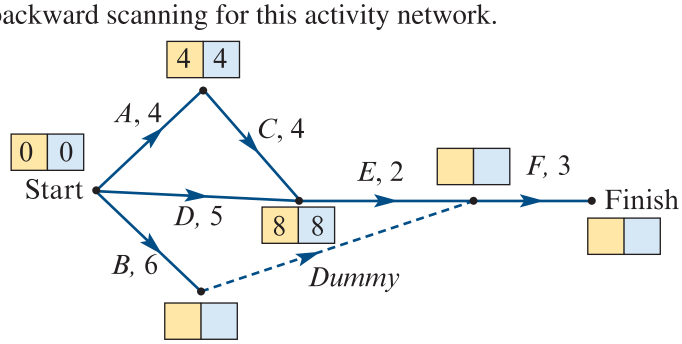
b What is the minimum completion time for this project?

c Write down the critical path for this project.

d For each non-critical activity, calculate the:

i LST

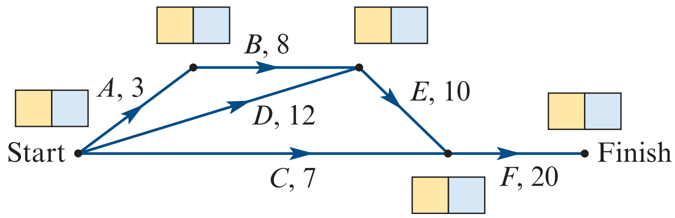
ii float time



ISBN 978-1-108-46154-2

4 Consider the activity network in the diagram shown below.

a Complete forward and backward scanning for this activity network.



b What is the minimum completion time for this project?

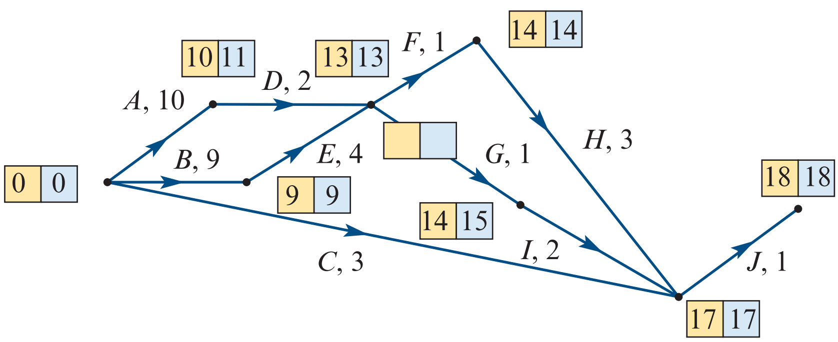
c Write down the critical path for this project.

d For each non-critical activity, calculate the:

i LST

ii float time

5 Consider the activity network shown in the diagram below.



a Complete the table of durations, EST, LFT, LST and float below.

Activity	Duration	EST	LFT	LST	Float
A	10		11	1	1
B		0	9		
C	3	0		14	14
D	2		13	11	
E	4	9		10	0
F		13	14		0
G	1		15	14	1
H		14		14	
I	2	14	17		1
J	1		18	17	0

b Write down the critical path for the project.

ISBN 978-1-108-46154-2

Cambridge University Press

6 Consider the precedence table for the activities in a project shown to the right.

Activity	Duration (weeks)	Immediate predecessors
P	4	–
Q	5	–
R	12	–
S	3	P
T	6	Q
U	3	S
V	4	R
W	8	R, T, U
X	13	V
Y	6	W, X

a Draw an activity network for this project.

b Complete the forward and backward scanning for this project.

c What is the shortest time, in weeks, in which this project could be completed?

d Use the activity network and results of scanning to write down the critical path for this project.

e Complete the table below to find the float times for every activity in the project.

Activity	Duration	EST	LFT	LST	Float
P	4				
Q	5				
R	12				
S	3				
T	6				
U	3				
V	4				
W	8				
X	13				
Y	6				

Note: you could use a spreadsheet to complete the calculations in this table.

f Use the table to verify that you identified the critical path correctly.

ISBN 978-1-108-46154-2

7 Consider the precedence table for the activities in a project shown on the right.

a Draw an activity network for this project.

b Complete the forward and backward scanning for this project.

c What is the shortest time, in weeks, in which this project could be completed?

d Use the activity network and results of scanning to write down the critical path for this project.

e Complete the table below to find the float times for every activity in the project.

Activity	Duration (weeks)	Immediate predecessors
I	2	–
J	3	–
K	5	–
L	4	I
M	8	J, N
N	1	K
O	6	L, M
P	6	J, N
Q	7	J, N
R	5	K
S	1	O
T	9	Q, R

Activity	Duration	EST	LFT	LST	Float
I	2				
J	3				
K	5				
L	4				
M	8				
N	1				
O	6				
P	6				
Q	7				
R	5				
S	1				
T	9				

Note: you could use a spreadsheet to complete the calculations in this table.

f Use the table to verify that you identified the critical path correctly.

ISBN 978-1-108-46154-2

Cambridge University Press

12C Applications of critical path analysis

To complete your study of problem planning and scheduling problems, you will see some realistic applications of critical path analysis.

Example 6 Application of critical path analysis

Linda is building a new house and has decided to manage the project herself.

She has decided on the major activities involved in this build and has been given some advice regarding the length of time each activity is expected to take.

These activities and their durations are shown in the table below.

Activity	Description	Duration (days)	Immediate predecessors
A	Preparing site and laying slab	5	-
B	Constructing frame and roof	25	*A*
C	Preparing floor	3	*A*
D	Landscaping gardens	10	*B*
E	Installing plants and lawn	3	*D*
F	Installing electrical	4	*B*
G	Installing plumbing	6	*B*, *C*
H	Finishing off ready to move in	5	*F*, *G*

a Construct an activity network for this project.

b Apply the forward scanning technique to determine the shortest time in which Linda can expect completion of her house.

c Apply the backward scanning technique and then complete the following.

- **i** Construct a table that shows the EST, LFT, LST and float for each activity.
- **ii** Which of the activities will cause a delay in the entire project if they take longer than expected?
- **iii** Write down the critical path of this project.

d **i** If the electrical work takes 2 days longer than expected, what effect will this have on the overall project?

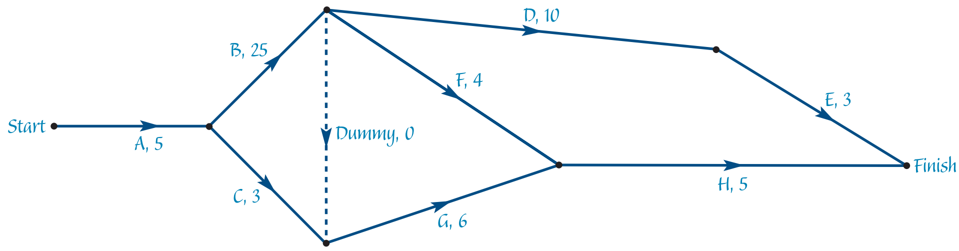
- **ii** Assuming that all other activities have no delay, what is the maximum delay possible for the floor preparation so that the project as a whole is not delayed?

ISBN 978-1-108-46154-2

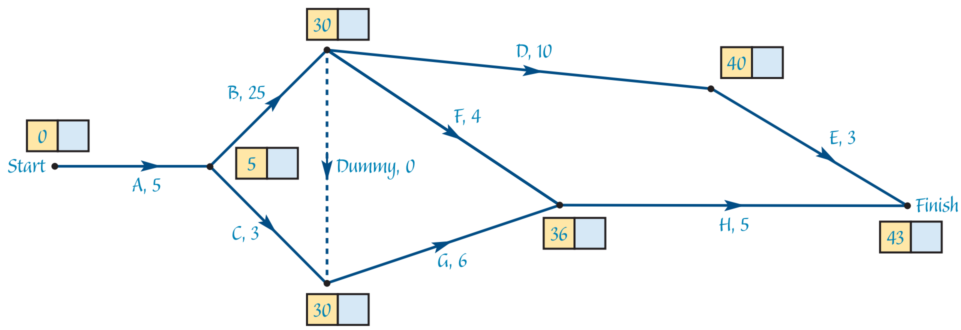
Cambridge University Press

Solution

a

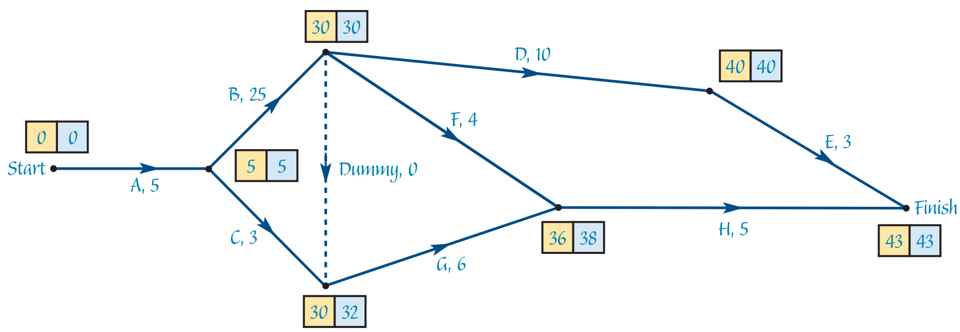


b



The house can be expected to be completed in 43 weeks.

c



i

Activity	Duration (days)	EST	LFT	LST	Float
A	5	0	5	0	0
B	25	5	30	5	0
C	3	5	32	29	24
D	10	30	40	30	0
E	3	40	43	40	0
F	4	30	38	34	4
G	6	30	38	32	2
H	5	36	43	38	2

ISBN 978-1-108-46154-2

ii The activities with zero float time will cause a delay to the entire project if they take longer than expected. These activities are preparing site and laying of slab (A), constructing frame and roof (B), landscaping the gardens (D) and installing the plants (E).

iii The critical path of this project is A – B – D – E.

d **i** The activity Installing electrical is activity F. Since activity F is not on the critical path, a delay will not necessarily affect the completion time of the project. Two days is less than the float time of activity F and so there will be no effect.

ii The activity 'Preparing floor' is activity C. Since C is not on the critical path, it may be delayed without affecting the completion time of the project. The float time for activity C is 24 days and so it may be delayed by a maximum of 24 days without delaying the project as a whole.

Exercise 12C

Example 6

1 Sharon's car washing business offers a premium service that involves five activities. These activities, their durations and their immediate predecessors are shown in the table below.

CF

Activity	Description	Duration (minutes)	Immediate predecessors
A	Wash the car	15	-
B	Dry the car	8	A
C	Wax the car	20	B
D	Clean the interior	35	B
E	Polish and shine the car	30	C

a Construct an activity network for this project.

b Apply the forward scanning technique to determine the shortest time in which the premium service can be expected to be completed.

c Apply the backward scanning technique and then answer the following questions.

  i Construct a table that shows the EST, LFT, LST and float for each activity.

  ii Which of the activities will not cause the premium service to take longer than expected if they were delayed?

  iii Write down the critical path of this project.

d The person responsible for cleaning the interior of the car has been delayed while working on another car. What is the latest time, in minutes after the project begins, that she can start the interior cleaning and still finish cleaning the car on time?

ISBN 978-1-108-46154-2

2 Robyn is planning a reunion of her primary school classmates. The activities involved in planning the event are shown in the table below, along with the duration and immediate predecessors of each activity.

Activity	Description	Duration (days)	Immediate predecessors
A	Locate and contact classmates	5	-
B	Choose a venue for the reunion	3	-
C	Organise decorations	2	B
D	Invite classmates	1	A
E	Wait for responses from classmates	7	D
F	Book tables	1	B, E
G	Send directions to venue	1	B, E
H	Hold the reunion	1	C, F, G

a Construct an activity network for this project. Two dummy activities will be required.

b Apply the forward scanning technique to determine the shortest time in which the reunion is expected to be completed.

c Apply the backward scanning technique and then answer the following questions.

- **i** Construct a table that shows the EST, LFT, LST and float for each activity.
- **ii** This project has two critical paths. Write down both of them.

d Use the information in the table from **ci** to explain what would happen if the classmates took 9 days to send their responses to Robyn instead of 7.

ISBN 978-1-108-46154-2

3 Anthony is creating a robot for a university project. The activities required to design and build the robot are shown in the table below, along with their duration in days and the immediate predecessor for each activity.

Activity	Description	Duration (weeks)	Immediate predecessors
A	Research robot design and control	5	-
B	Design the internal electronics	8	*A*
C	Design the remote control	3	*A*
D	Construct and assemble the robot	15	*B*
E	Write the code to control the robot	10	*B*
F	Construct and program the remote control	6	*C*
G	Debug the code to control the robot	4	*E*
H	Install the software	1	*D*, *F*
I	Test the robot	3	*G*, *H*

a Construct an activity network for this project.

b Apply the forward scanning technique to determine the shortest time in which Anthony can expect to create his robot.

c Apply the backward scanning technique and then answer the following questions.

i Construct a table that shows the EST, LFT, LST and float for each activity.

ii Write down the critical path of this project.

d Use the information in the table from **ci** to describe and explain what would happen if Anthony took:

i 3 weeks to research robot design and control instead of 5

ii 10 weeks to construct and program the remote control instead of 6

iii 20 weeks to construct and assemble the robot instead of 15

ISBN 978-1-108-46154-2

Cambridge University Press

Key ideas and chapter summary

AS

Project	A project is a task that involves a number of individual steps or activities.
Activity	An activity is an individual step in the completion of a project.
Project planning	Project planning involves the analysis and organisation of the activities in a project, taking into account the order in which they must be completed and the time it takes to complete each activity.
Activity network	A directed graph that represents the activities in a project is called an **activity network**. Edges are used to represent the activities, and these are labelled with the name of the activity and the duration. Vertices of an activity network are not labelled, except for the 'Start' and 'Finish' vertices.
Immediate predecessor	If activity M must be completed before activity N, then activity M is an **immediate predecessor** of activity N.
Precedence table	A table that shows all of the activities in a project and their immediate predecessors. **Precedence tables** may also show the duration of each activity.
Dummy activity	A **dummy activity** is used to preserve all precedence information in situations where activities share some, but not all, of the same immediate predecessors. Dummy activities have no duration.
Scheduling problem	Scheduling problems involve analysis of the precedence relationship between activities of a project and their durations in order to determine the minimum overall time it would take to complete the project.
Weighted precedence table	A weighted precedence table is a precedence table that also contains the durations of the activities.
Float time	**Float time** is also called slack time. It is the largest amount of time by which the activity can be delayed without affecting the overall completion time of the project. The float time for an activity is the difference between the latest starting time and the earliest starting time of that activity. Float = LST − EST

ISBN 978-1-108-46154-2

Cambridge University Press

Critical path	The critical path for a project is the sequence of activities that cannot be delayed without affecting the overall completion time of the project. Activities on the critical path have float times of zero, that is the EST and LST are the same. A project may have more than one critical path.
Critical path analysis	**Critical path analysis** is the process of using knowledge of the precedence and duration for each activity to determine the critical path of a project.
Earliest starting time (EST)	The **earliest starting time** for an activity is the earliest time after the start of the project that an activity can begin. Earliest starting time is referred to as EST.
Latest finishing time (LFT)	The **latest starting time** for an activity is the latest time after the start of the project that an activity can finish without affecting the overall completion time of the project. Latest finishing time is referred to as LFT.
Forward scanning	**Forward scanning** is the process of determining the EST for each activity in a project. The EST of an activity is added to the duration of the activity to determine the EST of the activities that immediately follow. The EST of any activity is equal to the largest forward scanning value determined from all immediate predecessors.
Backward scanning	**Backward scanning** is a process of determining the LFT for each activity in a project. The duration of an activity is subtracted from the LFT of activities that immediately follow. The LFT of any activity is equal to the smallest backward scanning value determined from all activities that immediately follow that activity.
Minimum completion time	The overall shortest amount of time in which the project can be completed.
Latest starting time (LST)	The latest starting time of an activity is the latest time that activity can start without affecting the overall completion time of the project. Latest starting time is referred to as LST. For any activity, LST = LFT − duration.

ISBN 978-1-108-46154-2

Cambridge University Press

Skills check

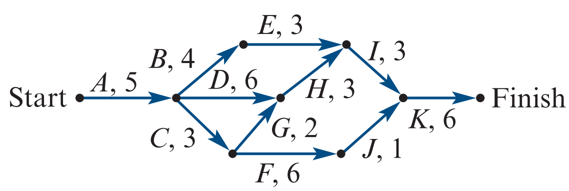
Having completed this chapter, you should be able to:

- identify activities in a project
- understand the precedence that some activities over others in a project
- identify immediate predecessors of activities from an activity network
- draw an activity network from a precedence table
- understand and explain the need for dummy activities in projects
- include dummy activities in activity networks as required
- create activity networks that include weights (durations) of each activity
- understand the existence of float times for some activities
- determine the EST for activities using forward scanning
- determine the LFT for activities using backward scanning
- calculate the LST for activities
- calculate the float time for activities
- complete critical path analysis to identify the critical path for a project.

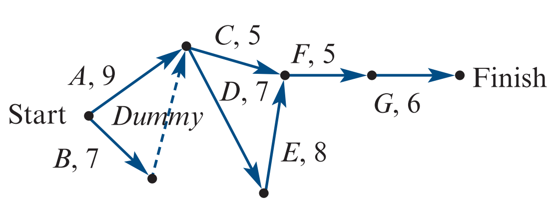
Multiple-choice questions

1 This activity network is for a project where the component times, in days, are as shown. The critical path for the network of this project is given by:

A $A-B-E-I-K$
B $A-D-H-I-K$
C $A-C-G-H-I-K$
D $A-C-F-J-K$
E $A-D-G-F-J-K$



2 The activity network shown represents a project development with activities and their durations (in days) listed on the edges of the graph. Note that the dummy activity takes zero time.



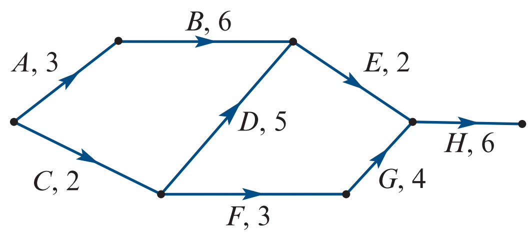
The earliest time (in days) that activity F can begin is:

A 0 **B** 12 **C** 14 **D** 22 **E** 24

ISBN 978-1-108-46154-2

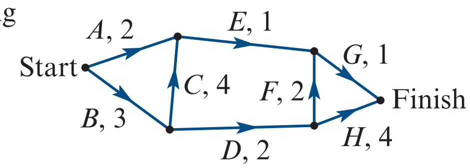
Cambridge University Press

3 The edges in this activity network correspond to the tasks involved in the preparation of an examination. The numbers indicate the time, in weeks, needed for each task. The total number of weeks needed for the preparation of the examination is:



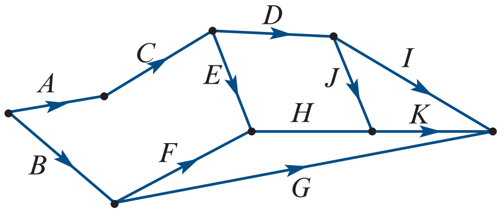
A 14 **B** 15 **C** 16 **D** 17 **E** 27

4 This activity network represents a manufacturing process with activities and their duration (in hours) listed on the edges of the graph. The earliest time (in hours) after the start of manufacturing that activity G can begin is:



A 3 **B** 5 **C** 6 **D** 7 **E** 8

5 This network represents a project development with activities listed on the edges of the graph.
Which of the following statements must be true?

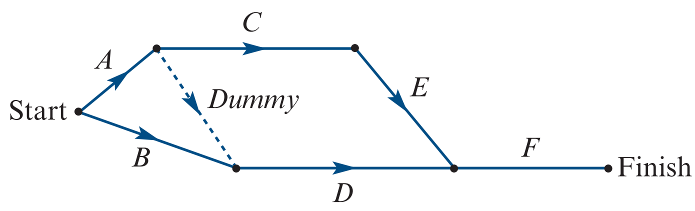


A A must be completed before B can start.
B A must be completed before F can start.
C E and F must start at the same time.
D E and F must finish at the same time.
E E cannot commence while A is still taking place.

ISBN 978-1-108-46154-2

Cambridge University Press

6 The table that shows the immediate predecessors for the activity network shown below is:



A

Activity	Immediate predecessors
A	–
B	–
C	A, B
D	B
E	C
F	E

B

Activity	Immediate predecessors
A	–
B	A
C	A, B
D	B
E	C
F	D, E

C

Activity	Immediate predecessors
A	–
B	A
C	A, B
D	B
E	C
F	D, E

D

Activity	Immediate predecessors
A	–
B	–
C	A
D	A, B
E	C
F	D, E

E

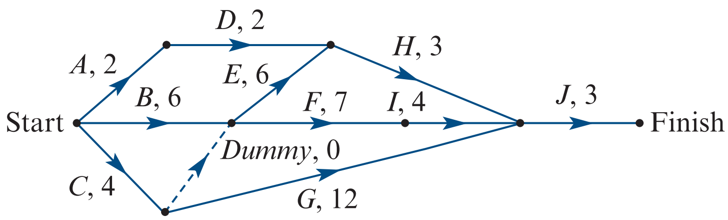
Activity	Immediate predecessors
A	–
B	–
C	A
D	A, B
E	C
F	D

 ISBN 978-1-108-46154-2 Cambridge University Press

Review

Use this information to answer questions 7 to 10

The activity network below shows the activities required to complete a particular project and the durations, in hours, of those activities.



7 The earliest time that activity I can start is:

A 11 **B** 12 **C** 13 **D** 14 **E** 15

8 The latest starting time for activity E is:

A 7 **B** 8 **C** 9 **D** 10 **E** 11

9 The float time of activity H is:

A 2 **B** 3 **C** 4 **D** 5 **E** 6

10 The critical path for the project is:

A $A - D - H - J$

B $C - G - J$

C $C - Dummy - E - H - J$

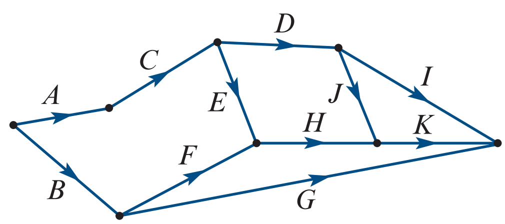
D $C - Dummy - F - I - J$

E $B - F - I - J$

Short-answer questions

SF

1 An activity network is shown below.



a Which activities are immediate predecessors for activity D?

b For which activities is activity B an immediate predecessor?

c How many immediate predecessors does activity H have?

ISBN 978-1-108-46154-2

Cambridge University Press

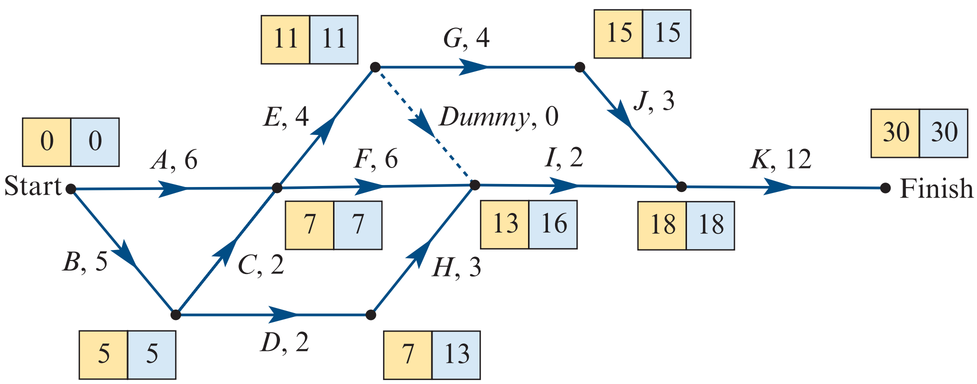
2 A precedence table for a project is shown below.

Activity	Immediate predecessors	Duration (weeks)
A	-	5
B	-	6
C	-	3
D	*C*	8
E	*B*	2
F	*D*, *E*	5
G	*A*, *F*	4

a Draw an activity network for this project.

b Perform the forward scanning process and determine the shortest time in which this project could be completed.

3 The activity network for a project is shown below. All durations in this network are in days.



The forward scanning and backward scanning processes have already been completed.

a How many days will it take to complete this project?

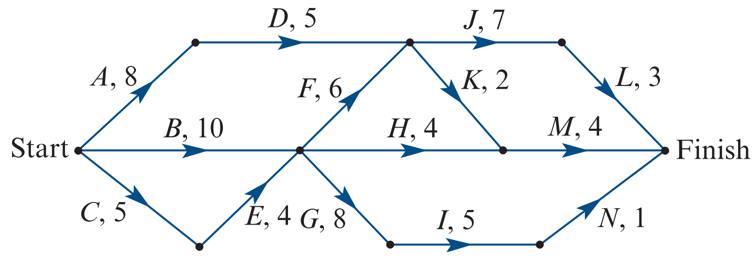
b What is the earliest starting time for activity *I*?

c What is the float time for activity *A*?

d Write down the critical path for this project.

ISBN 978-1-108-46154-2

4 The activity network for a project is shown below. All durations in this network are in hours.



a Construct a precedence table for this project.

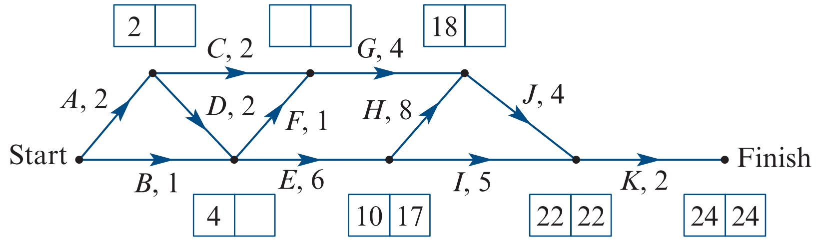
b Complete the forward scanning process to determine the shortest time in which this project can be completed.

c Complete the backward scanning process and write down the critical path for this project.

Extended-response questions

1 The assembly of machined parts in a manufacturing process can be represented by the following network. The activities are represented by letters on the arcs and the numbers represent the time taken (in hours) for the activities scheduled.

CF



Activities	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
EST	0	0	2	2	4	4		10	10	18	22

a The earliest start times (EST) for each activity except *G* are given in the table. Complete the table by finding the EST for *G*.

b What is the shortest time required to assemble the product?

c What the float (Slack time) for activity *I*?

ISBN 978-1-108-46154-2

Cambridge University Press

2 A precedence table for a project is shown below.

Activity	Immediate
A	–
B	–
C	*B*
D	*A*, *C*
E	*A*, *C*
F	*E*
G	*D*
H	*D*
I	*H*, *F*
J	*G*, *I*

a Draw the activity network for this project.

b Complete the forward and backward scanning processes to find the minimum number of days in which this project could be completed.

c Complete the table of EST, LFT, LST and float values.

Activity	Duration (days)	EST	LFT	LST	Float
A	10				
B	5				
C	3				
D	5				
E	4				
F	6				
G	6				
H	7				
I	5				
J	4				

d Explain what it means for an activity to be on the critical path for a project.

e Identify the feature of the table above that allows you to write down the critical path for this project.

f Write down the critical path for this project.

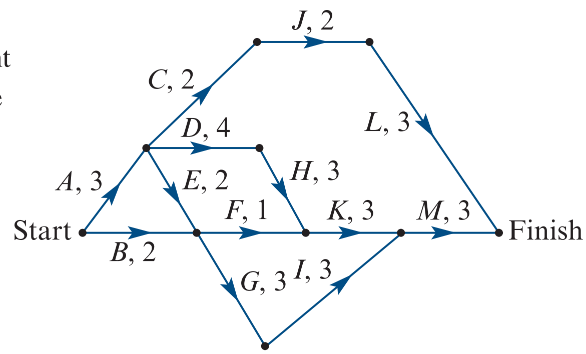
g Explain what would happen to the completion time of the project if activity C started 1 day later than expected.

 ISBN 978-1-108-46154-2

Review

CU

3 The Bowen Yard Buster team specialises in backyard improvement projects. The team has identified the activities required for a backyard improvement. The network diagram to the right shows the activities identified and the actual times, in hours, needed to complete each activity, that is, the duration of each activity.



The table below lists the activities, their immediate predecessor(s) and the earliest starting time (EST), in hours, of each of the activities. Activity X is not yet drawn on the network diagram.

a Use the information in the network diagram to complete the table.

b Draw and label activity X on the network diagram above, including its direction and duration.

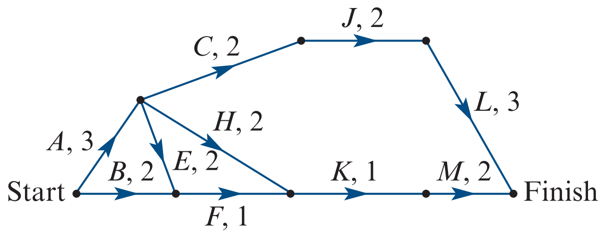
c The path $A - D - H - K - M$ is the only critical path in this project.

i Write down the duration of path $A - D - H - K - M$.

ii Explain the importance of the critical path in completing the project.

Activity	Immediate predecessor(s)	EST
A	–	0
B	–	0
C	A	3
D	A	3
E		3
F	B, E	5
G	B, E	5
H	D	7
I	G	
J	C, X	8
K	F, H	10
L	J	10
M	I, K	
X	D	7

d To save money, Bowen Yard Busters decide to revise the project and leave out activities D, G, I and X. This results in a reduction in the time needed to complete activities H, K and M as shown.

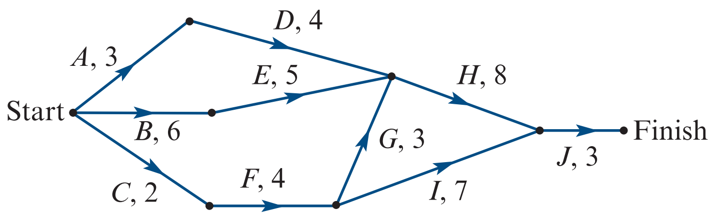


i For this revised project network, what is the earliest starting time for activity K?

ii Write down the critical path for this revised project network.

iii Without affecting the earliest completion time for this entire revised project, what is the latest starting time for activity M?

4 The activity network for a project is shown in the diagram below. The duration for each activity is in hours.



a Identify the critical path for this project.

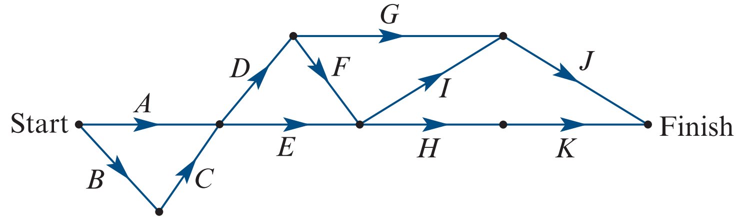
b What is the maximum number of hours by which the completion time for activity E can be reduced without affecting the critical path of the project?

c What is the maximum number of hours by which the completion time for activity H can be reduced without affecting the critical path of the project?

d Every activity can be reduced in duration by a maximum of 2 hours. If every activity was reduced by the maximum amount possible, what is the new minimum completion time for the project?

 ISBN 978-1-108-46154-2

5 In laying a pipeline, the various jobs involved have been grouped into a set of specific tasks $A - K$, which are performed in the precedence described in the network below.



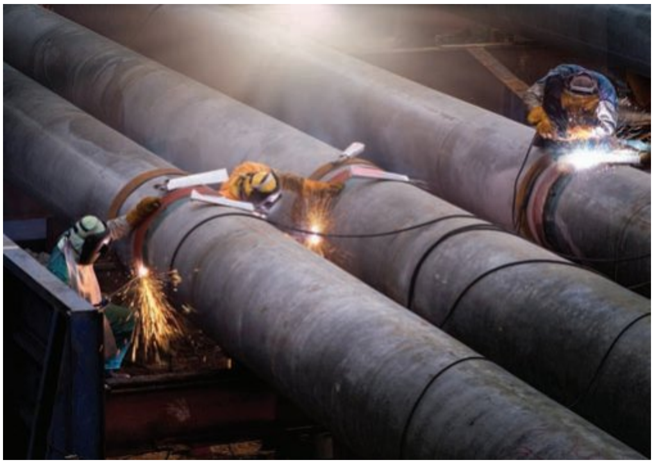
a List all the task(s) that must be completed before task E is started.

The durations of the tasks are given in Table 16.7.

b Use the information in Table 16.7 to complete Table 16.8.

Table 16.7 Task durations

Task	Normal completion time (months)
A	10
B	6
C	3
D	4
E	7
F	4
G	5
H	4
I	5
J	4
K	3



ISBN 978-1-108-46154-2

Table 16.8 Starting times for tasks

Task	EST	LST
A	0	0
B	0	
C	6	7
D	10	10
E		11
F	14	14
G	14	18
H	18	20
I	18	
J	23	23
K	22	24

c For this project:

  i write down the critical path

  ii determine the length of the critical path (that is, the earliest time the project can be completed)

d If the project managers are prepared to pay more for additional labour and machinery, the time taken to complete task *A* can be reduced to 8 months, task *E* can be reduced to 5 months and task *I* can be reduced to 4 months.

Under these circumstances:

  i what would be the critical path(s)?

  ii how long would it take to complete the project?

ISBN 978-1-108-46154-2

13

Revision of Unit 4 Chapters 7–12

UNIT 4 INVESTING AND NETWORKING

Topic 1 Loans, investments and annuities

Topic 2 Graphs and networks

Topic 3 Networks and decision mathematics

The revision exercises are arranged by chapter with these categories of questions:

- Simple familiar question types
- Complex familiar question types
- Complex unfamiliar question types

ISBN 978-1-108-46154-2

Cambridge University Press

Revision

13A Topic 1 Loans, investments and annuities

Multiple-choice questions

Use the following information to answer question 1 and 2.

The balance of a compound interest investment after n quarters, A_n, can be modelled by the recurrence relation $A_0 = 50000$, $A_{n+1} = 1.021 \times A_n$

1 The balance of the investment after one year is:

A \$50000
B \$51050
C \$52122.05
D \$53216.61
E \$54334.16

2 The annual percentage rate of interest for this investment is:

A 0.21% **B** 2.1% **C** 21% **D** 0.84% **E** 8.4%

3 Ali would like to borrow some money with a compound interest loan. The principal and all interest will be paid back at the end of the loan period. He would like to pay as little interest as possible. Which one of the following interest rates should he choose?

A 6.1% per annum, compounding weekly
B 6.1% per annum, compounding monthly
C 6.2% per annum, compounding quarterly
D 6.2% per annum, compounding monthly
E 6.2% per annum, compounding annually

4 A principal of \$14000 is invested and will earn compound interest at the rate of 2.8% per annum, compounding weekly. The effective annual rate of interest for this investment is closest to:

A 2.80% **B** 2.81% **C** 2.82%
D 2.83% **E** 2.84%

5 Eli borrowed some money. He will be charged compound interest at the rate of 7.08% per annum, compounding monthly. After one year, Eli repaid \$6674.95 as principal and interest. The amount borrowed was closest to:

A \$6000 **B** \$6100 **C** \$6200
D \$6300 **E** \$6400

6 The balance of a reducing-balance loan after n months, A_n, can be modelled by the recurrence relation $A_0 = 250000$, $A_{n+1} = 1.003125 \times A_n - 2300$.
The total interest that has been paid after one repayment is closest to:

A \$781 **B** \$1519 **C** \$1558
D \$2300 **E** \$3042

ISBN 978-1-108-46154-2

Cambridge University Press

7 A loan of $180 000 is charged compound interest at the annual percentage interest rate of 3.24% per annum, compounding monthly. The loan is repaid with monthly repayments of $1200. If A_n is the balance of the loan after n months, which one of the following is a recurrence relation model for this loan?

A $A_0 = 180000,\ A_{n+1} = 1.0324 \times A_n - 1200$

B $A_0 = 180000,\ A_{n+1} = 1.00324 \times A_n - 1200$

C $A_0 = 180000,\ A_{n+1} = 1.27 \times A_n - 1200$

D $A_0 = 180000,\ A_{n+1} = 1.027 \times A_n - 1200$

E $A_0 = 180000,\ A_{n+1} = 1.0027 \times A_n - 1200$

8 A repayment schedule for the first two repayments of a reducing-balance loan is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	40 000.00
1	400.00	160.00	240.00	39 760.00
2	400.00	159.04	240.96	39 519.04

Which one of the following is the next line of this repayment schedule?

A

3	400.00	158.08	241.92	39 277.12

B

3	400.00	159.04	240.96	39 278.08

C

3	400.00	240.96	159.04	39 360.00

D

3	400.00	241.92	158.08	39 360.96

E

3	400.00	243.21	156.79	39 362.25

9 A reducing-balance loan has a current balance of $65 000. Interest is charged at the rate of 8.28% per annum, compounding monthly. Repayments of $1500 are made each month.

How many repayments does it take to reduce the balance of this loan below $60 000 for the first time?

A 4 **B** 5 **C** 6 **D** 7 **E** 8

10 The balance of a reducing-balance loan with principal $350 000 has been reduced to $250 031.74 after six years of monthly repayments. If each repayment had value $2400, the annual percentage rate of interest is closest to:

A 3% **B** 3.5% **C** 4% **D** 4.5% **E** 5%

ISBN 978-1-108-46154-2

Cambridge University Press

11 Which one of the following recurrence relations could be a model for an annuity in withdrawal phase with principal \$256000, earning interest at an annual percentage rate of interest of 7.6% compounding quarterly and with quarterly payments of \$7500?

A $A_0 = 256000,\ A_{n+1} = 0.76 \times A_n - 7500$
B $A_0 = 256000,\ A_{n+1} = 1.019A_n - 7500$
C $A_0 = 256000,\ A_{n+1} = 1.076 \times A_n - 7500$
D $A_0 = 256000,\ A_{n+1} = 1.19 \times A_n - 7500$
E $A_0 = 256000,\ A_{n+1} = 1.76 \times A_n - 7500$

12 Elvira has inherited \$100000 and will invest this money into an annuity from which she will withdraw monthly payments. Interest will be earned at the rate of 4.68% per annum, compounding monthly. If the balance of Elvira's investment was \$93463.21 after five payments have been withdrawn, what is the value of Elvira's monthly payment?

A \$1024.22 **B** \$1307.36 **C** \$1396.75 **D** \$1687.20 **E** \$6536.79

13 Which one of the following recurrence relations could be a model for a perpetuity?

A $A_0 = 85000,\ A_{n+1} = 1.035 \times A_n - 3570$
B $A_0 = 85000,\ A_{n+1} = 1.039 \times A_n - 3570$
C $A_0 = 85000,\ A_{n+1} = 1.041 \times A_n - 3570$
D $A_0 = 85000,\ A_{n+1} = 1.042 \times A_n - 3570$
E $A_0 = 85000,\ A_{n+1} = 1.065 \times A_n - 3570$

14 Arthur has invested \$40000 in an annuity. His investment will earn interest at the rate of 7.44% per annum, compounding monthly. Arthur will withdraw \$1200 a month from this annuity.
How many payments of \$1200 can Arthur expect from this annuity?

A 37 **B** 38 **C** 39 **D** 40 **E** 41

15 A perpetuity will be set up to provide an annual prize of \$800 to the winner of a mathematics competition. Interest will be earned on the principal of the investment at the rate of 4% per annum and this will be used to pay the prize money every year. The amount that must be invested is:

A \$2000 **B** \$3200 **C** \$20000 **D** \$80000 **E** \$320000

ISBN 978-1-108-46154-2

Cambridge University Press

Short-answer questions

Simple familiar questions

1 The following recurrence relation can be used to model a compound interest investment of \$10000 earning interest at the rate of 7.68% per annum, compounding monthly.

$$A_0 = 10000, A_{n+1} = 1.0064 \times A_n$$

In this recurrence relation, A_0 is the balance of the investment after n months.

a Apply the recurrence relation to find the balance of the investment after one, two and three months.

b How many months will it take for the balance of this investment to first exceed \$10500?

2 Erica has invested \$14500 into an account that pays compound interest at the rate of 4.8% per annum, compounding monthly.

a Construct a recurrence relation model for the balance of Erica's investment after n months.

b Apply the recurrence relation to determine the balance of Erica's investment after three months.

c Using the compound interest formula, what will the balance of Erica's investment be after two years?

3 Jack has borrowed \$4500 to buy furniture for his home. He will be charged compound interest at the rate of 10.2% per annum, compounding monthly.
Let A_n be the balance of Jack's loan after n months.

a What is the monthly percentage rate of interest for this loan?

b Construct a recurrence relation that models the balance of Jack's loan.

c If Jack will pay the principal and all interest charged after one year, how much money will he have to repay?

4 Hugh has invested \$25000 in an account that will pay compound interest every month, at the annual percentage rate of interest of 3.84%.

a Use the compound interest rule to determine the balance of Hugh's investment after three years.

b How much interest has been earned in total after three years?

5 Dorothy borrowed some money to pay for a travelling holiday. The annual percentage rate of interest for her loan was 8.76%, compounding monthly. Dorothy repaid the principal and all interest charged. a total sum of \$13057.59, after travelling for six months.
Use the compound interest rule to determine the principal amount that Dorothy borrowed.

ISBN 978-1-108-46154-2

6 A reducing-balance loan is modelled using the recurrence relation shown below.

$$A_0 = 3000, A_{n+1} = 1.0024 \times A_n - 150$$

In this recurrence relation, A_n is the balance of the loan after n weekly repayments.

a What is the principal of this loan?

b What is the value of the weekly repayments?

c What is the annual percentage rate of interest?

d What is the balance of this loan after five repayments?

7 Celia is considering borrowing \$50 000 to buy a caravan. Her bank will charge interest at the rate of 7.08% per annum, compounding monthly. Celia can afford to make monthly repayments of \$500 to repay the loan.

Let A_n be the balance of Celia's loan after n months.

a What is the monthly percentage rate of interest for Celia's loan?

b Construct a recurrence relation model for the balance of Celia's loan after n months.

c Apply the recurrence relation to determine the balance of Celia's loan after six months.

d How many months will it take for Celia's loan to have a balance that is below \$48 000 for the first time?

8 Gracie has a loan of \$12 000 that she will repay with monthly repayments of \$450. Interest is charged at the percentage annual interest rate of 6.24%.

a Determine the balance of Gracie's loan after six repayments have been made.

b How much interest has been paid in total after six repayments have been made?

9 A loan of \$5000 is repaid with monthly repayments of \$250. Interest is charged at the annual percentage interest rate of 7.32%.

a Use the annuities formula to determine the balance of this loan after five repayments have been made.

b Calculate the total interest that has been paid after five repayments have been made.

10 A little while ago, Trent borrowed some money from a bank. The bank charges interest at the annual percentage rate of 4.68% and Trent has been repaying his loan with monthly repayments of \$400. Trent still owes the bank \$21 385.76 after three years of monthly repayments.

Use the annuities formula to determine the amount of money that Trent initially borrowed.

ISBN 978-1-108-46154-2

Revision

11 The balance of an annuity, A_n, after n monthly payments have been received is modelled by the recurrence relation below.

$$A_0 = 250000, A_{n+1} = 1.0031 \times A_n - 1800$$

a What is the percentage annual rate of interest for this annuity?

b How much is left invested in this annuity after five payments have been received?

c Calculate the total amount of interest that has been earned after five payments have been received.

12 Georgina has \$145 000 to invest in an annuity. Interest will be paid at the annual percentage interest rate of 4.08%, compounding monthly. Georgina will withdraw a payment of \$2500 each month from the investment.
Let A_n be the balance of Georgina's annuity after n monthly payments have been withdrawn.

a Construct a recurrence relation model for the balance of Georgina's investment after n payment withdrawals.

b Apply the recurrence relation to calculate the amount remaining in Georgina's investment after five payments have been withdrawn.

c Calculate the total interest that Georgina has earnt after five payments have been withdrawn.

13 Use the annuities formula to determine the balance of an annuity with principal \$120 000 earning interest at the rate of 8.4% per annum, compounding quarterly, and with quarterly withdrawals of \$6000 after 5 years.

14 An advertising agency has invested \$140 000 in a perpetuity. The interest earned each month by this investment will pay for a monthly award to a high performing employee.

a If the interest on the investment is paid at the rate of 5.16% per annum, what is the value of the prize awarded each month?

b If the prize has value \$800 per month, what is the annual percentage rate of interest for the investment? Round your answer to two decimal places.

ISBN 978-1-108-46154-2

Cambridge University Press

Revision

Extended-response questions

Complex familiar questions

1 Brian has $35 000 to invest. He has two investment options:

- Bank A offers to pay 4.68% per annum compounding monthly
- Bank B offers to pay 4.56% per annum compounding fortnightly

Brian would like to withdraw his money and all of the interest it has earned after a period of three years.

a Which of the two investment options would earn Brian the most interest after one year? Explain how you compared the two investment options.

b Write a letter to Brian explaining the comparison of the two investment options, showing him the calculations for the total amount he could withdraw after three years.

2 Edward has borrowed $5000. He will be charged interest at the rate of 11.4% per annum, compounding monthly.

a Let A_n be the balance of Edward's loan after n months.

i Construct a recurrence relation that models the value of A_n from month to month.

ii Write down a rule for A_n in terms of n only.

b Edward is planning to repay his loan and all interest charged after eight months. After the fifth month of the loan, the interest rate was increased to 11.88% per annum, compounding monthly.
Calculate the total amount that Edward must repay. Round your answer to the nearest dollar.

3 Cleo borrowed $120 000 and has been charged compound interest at the annual percentage interest rate of 7.68%, compounding monthly. After two years of repayments, Cleo will still owe a balance of $107 342.34 on her loan.

a Determine the monthly repayment that Cleo has been paying.

b How much will Cleo owe after a further three years of repayments?

After the first two years of repayments, Cleo increases her repayment amount so that her loan will be fully repaid after four further years.

c Determine the new monthly repayment for this loan. Round your answer to the nearest cent.

ISBN 978-1-108-46154-2

Cambridge University Press

Revision

4 Fillipe would like to buy an apartment and he will need to borrow \$250 000 to pay for this. Interest will be charged at the annual percentage interest rate of 3.96%.

a Fillipe plans to repay his loan over a period of 25 years.

i Calculate the monthly repayment amount required to achieve this aim. Round your answer to the nearest cent.

ii Using the rounded repayment amount, calculate the balance of the loan after 25 years.

iii This amount is positive. Explain the significance of this amount.

b After four years of repayments (48 repayments), Fillipe will make a lump sum repayment of \$50 000. How many further repayments will be required to repay the loan? Round your answer to the nearest whole number.

5 Luther receives monthly payments of \$5400 from an annuity that is earning interest at the rate of 6.12% per annum, compounding monthly. The balance of Luther's investment is \$326 296.83 after four years of investment.

a What is the principal amount of Luther's investment?

b How much interest has Luther earned after four years of investment?

c How many more payments of \$5400 can Luther withdraw?

6 Vusa has recently retired and will invest his superannuation payment into an annuity that will earn interest at the rate of 6.8% per annum, compounding quarterly. The principal amount of this investment will be \$396 000.

a If Vusa withdraws a payment of \$20 000 per quarter, how many payments in total can he expect to withdraw?

b What is the balance of Vusa's investment after three years of repayments?

Vusa decided to increase his monthly payment after three years of investment. He will now withdraw monthly payments of \$25 000 until his investment is fully exhausted.

c How many payments of \$25 000 can he expect?

d His final payment will be smaller than \$25 000. How much will this final payment be?

ISBN 978-1-108-46154-2

Cambridge University Press

▶ Complex unfamiliar questions

7 Leanne currently owes $138 500 on her home loan. She pays interest at the annual percentage interest rate of 4.32% per annum and repays the loan with monthly repayments of $1250.
After six months, the interest rate of Leanne's loan increased to 4.44% per annum, compounding monthly. Leanne decided to increase her payments to $1500 per month. How much will Leanne owe on this loan after a further 12 months?

8 The details of two different home loans with principal $320 000 are shown in the table below.

	Interest rate	Term	Compounding period	Extra repayments	Repayment
Loan 1	3.12%	20 years	Monthly	Allowed	$1800
Loan 2	3.38%	18 years	Fortnightly	Not allowed	$900

Amanda is trying to decide between the two loans above. She believes that she can afford repayments of $1900 per month.

a If no extra repayments are made, which of the two loans would be best for Amanda? Justify your decision by explaining your mathematical reasoning.

b If Amanda could make larger repayments, does this change your decision? Justify your answer by explaining your mathematical reasoning.

ISBN 978-1-108-46154-2

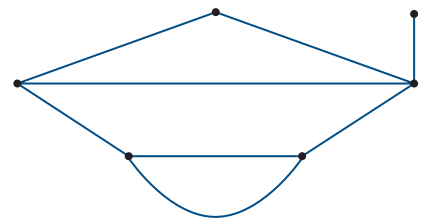
Cambridge University Press

Revision

13B Topic 2 Graphs and networks

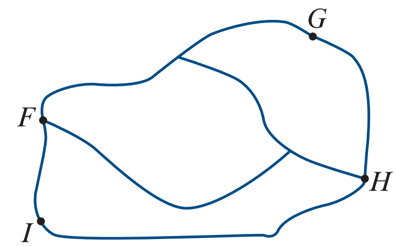
Multiple-choice questions

1 The sum of the degrees of all the vertices in the network opposite is:

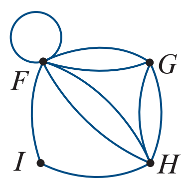


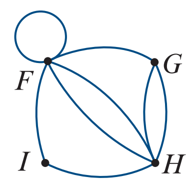
A 6 **B** 7
C 8 **D** 15
E 16

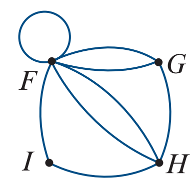
2 The diagram shows a map of the roads between four towns: *F*, *G*, *H* and *I*.

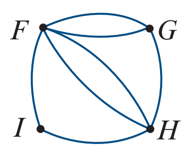


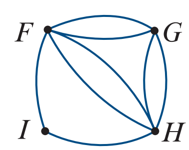
A network diagram that represents the connections between towns on the map is:

A


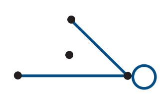
B


C


D


E


3 An adjacency matrix for the graph opposite could be:



A $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ **B** $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ **C** $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

D $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ **E** $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

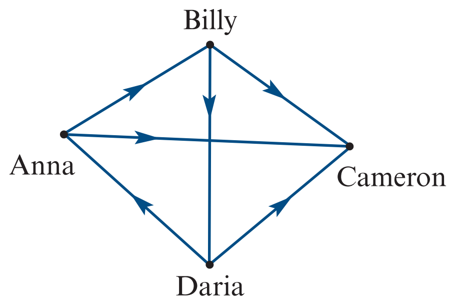
4 A connected graph with 12 edges divides a plane into four faces. The number of vertices in this graph will be:

A 6 **B** 10 **C** 12 **D** 13 **E** 14

ISBN 978-1-108-46154-2

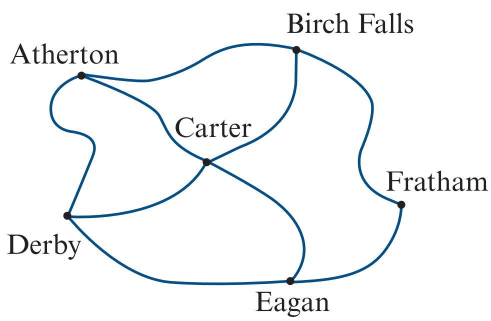
Cambridge University Press

5 The directed graph below shows the results of a chess competition between four competitors, Anna, Billy, Cameron and Daria. The arc on the graph represents the game between the people at the vertices connected by that arc. The arrow points to the loser of the game.
Which one of the following statements about this chess competition is **not** true?



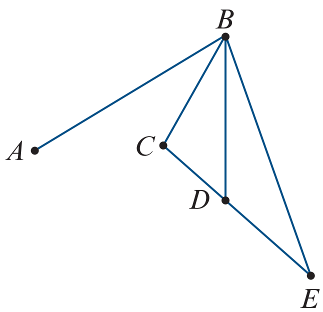
A Billy beat Daria.
B Daria lost more than one game.
C Anna won two of the three games she played.
D Cameron did not win any of his games.
E Nobody won all of their games.

6 The graph on the right shows six cities represented by vertices and the railway lines between those cities represented by edges.
Which one of the following walks along the railway lines is also a trail?



A Birch Falls – Atherton – Carter – Derby – Atherton – Carter – Eagan
B Eagan – Carter – Derby – Eagan – Fratham
C Fratham – Brich Falls – Atherton – Carter – Birch Falls – Fratham
D Carter – Derby – Carter – Atherton – Derby
E Derby – Carter – Eagan – Derby – Carter – Atherton

7 The graph on the right is best described as:
A bipartite
B Hamiltonian
C semi-Eulerian
D Eulerian
E complete



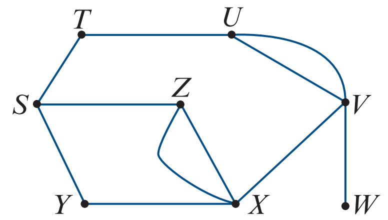
ISBN 978-1-108-46154-2

Cambridge University Press

Revision

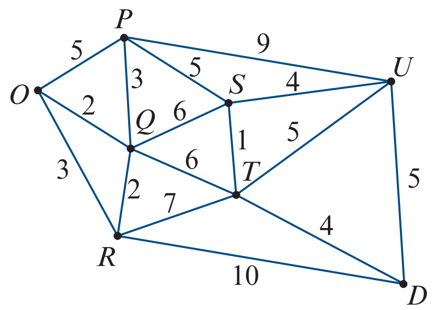
8 Adding which one of the following edges to the graph makes an eulerian trail possible?

A ST **B** SU **C** SX
D XW **E** ZY

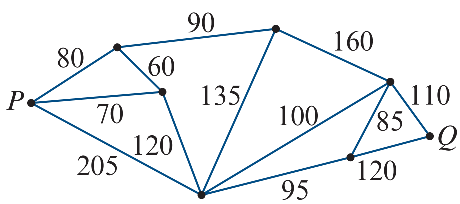


9 The length of the shortest path between the origin, O, and destination, D, in the network shown here is:

A 11 **B** 12 **C** 13
D 14 **E** 15



10 The graph below shows towns in a particular region represented as vertices and the roads between them represented as edges. The weight on an edge shows the travel time, in minutes, between the vertices connected by that edge.
A salesman is about to travel from town P to town Q.



The shortest time it could take him, in minutes, is:

A 400 **B** 405 **C** 410 **D** 420 **E** 440

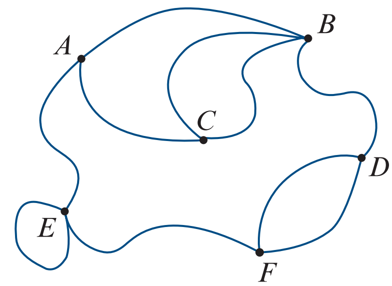
ISBN 978-1-108-46154-2

Cambridge University Press

Short-answer questions

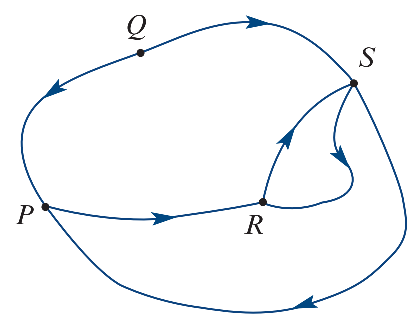
► Simple familiar questions

1 Consider the graph on the right.



- **a** What is the degree of vertex C?
- **b** Which vertex has a loop?
- **c** How many vertices in this graph have an even degree?
- **d** Which vertices are immediately adjacent to vertex B?
- **e** Which pairs of vertices have multiple edges between them?

2 Consider the directed graph shown on the right.



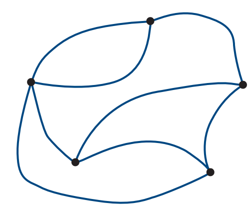
- **a** Which vertex is the only one that can be reached from vertex P?
- **b** Which vertex cannot be reached from any vertex?
- **c** Construct an adjacency matrix for this directed graph.

3 The adjacency matrix for a graph is shown on the right.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}\begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

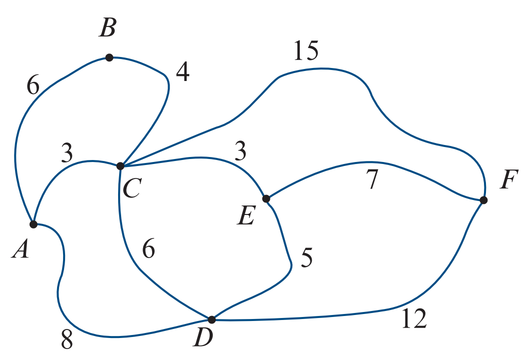
- **a** How many edges are there between vertex A and vertex D?
- **b** How many loops are there in the graph for this adjacency matrix?
- **c** Draw the graph represented by this adjacency matrix.

4 Consider the graph on the right.



- **a** How many edges does this graph have?
- **b** How many faces does this graph have?
- **c** How many vertices does this graph have?
- **d** Verify Euler's rule for this graph.

5 The vertices of the network on the right represent camping sites within a national park. The arcs of the network represent the walking tracks between each camping site and the numbers on these arcs show the distances along the tracks, in kilometres.



- **a** A hiker is at campsite A. If she walks to campsite F, via campsite D, how far would she walk?
- **b** What is the shortest distance from campsite A to campsite F?

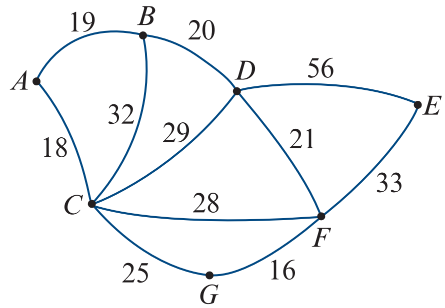
ISBN 978-1-108-46154-2

Revision

Extended-response questions

Complex familiar questions

1 Seven towns (A, B, C, D, E, F, G) are represented by vertices on the graph to the right. The edges of the graph represent the road connections between the towns. The weights on the edges represent the distance along the road connections, in kilometres.



a This graph is planar. Explain what this means.

b Verify that Euler's formula for this graph.

c An inspector or roads will begin at town B and inspect each of the roads only once.

  i Where will the inspector end this inspection?

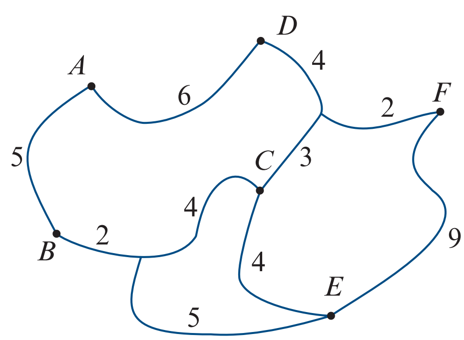
  ii What is the name given to the walk that the inspector completes through this graph?

d Each of the towns have a branch of a bank. A bank manager from the branch at town C must visit all of the other branches in the region and then return to his own branch at town C

  i Write down one walk that the bank manager could take.

  ii What is the name given to this walk?

2 The graph below shows six towns, A, B, C, D, E, F represented by vertices. The towns are connected by roads shown as edges on the graph. The weights on the edges show the length, in kilometres, along each section of road.



a A person can drive from E to B directly. How many kilometres will this journey be?

b How many extra kilometres is the journey from E to B via C compared to the direct journey?

c Which towns are adjacent to town A on this graph?

d Complete the following for graphs.

  i Draw a graph that shows the shortest direct distances between each of the towns.

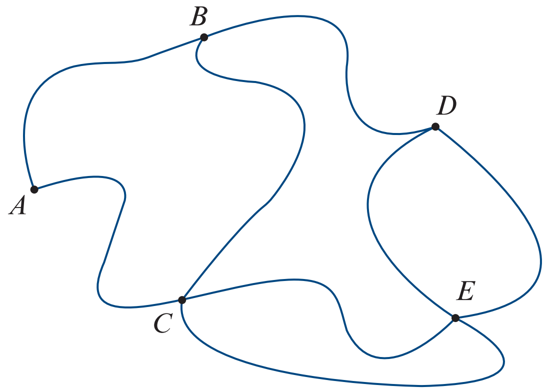
  ii Construct the adjacency matrix for this graph.

ISBN 978-1-108-46154-2

Cambridge University Press

Complex unfamiliar questions

3 A family is visiting a theme park and will visit five rides. A map of the theme park is shown below with the vertices representing the rides and the edges representing the paths connecting the rides.

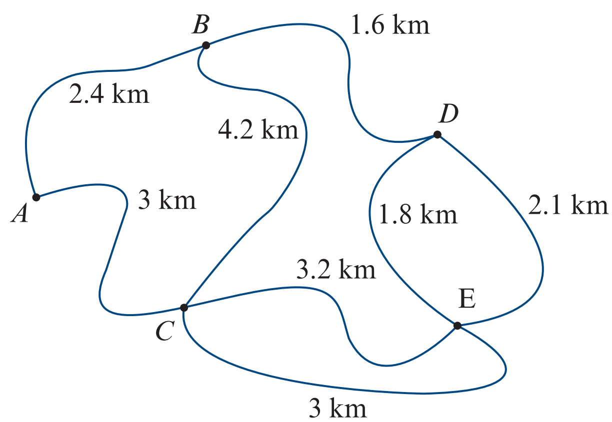


a Determine a Hamilton cycle that the family can follow.

b Make a list of any paths (edges) that the family does not follow.

c Determine an Eulerian trail that the family can follow. Explain how you determined this trail.

The network below shows the lengths of the paths that join the rides.



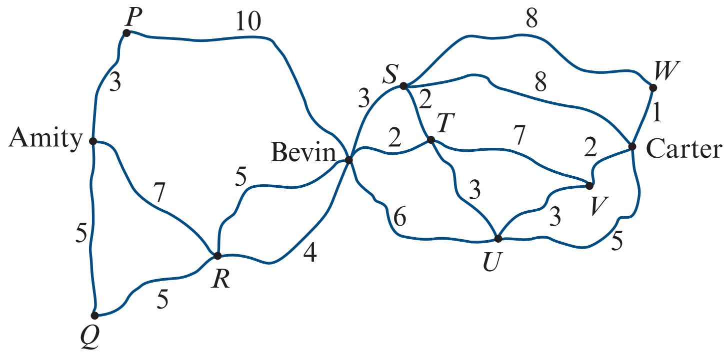
The family decide to visit the rides by following an Eulerian trail. Assume that the family can walk at a speed of 3 km/hr. The theme park will close at 5 p.m.

d What is the latest time that the family can enter the theme park?

ISBN 978-1-108-46154-2

Cambridge University Press

4 The diagram below shows the roads that connect the towns of Amity, Bevin and Carter represented as edges of a network. The vertices of the network, labelled P, Q, R, S, T, U, V and W, are checkpoints for the Amity Cycling Club road race. The numbers on the edges of the network are the lengths, in kilometres, of the roads between the checkpoints and the towns.



a Find, by inspection, the length of the shortest path from Amity to Bevin.

The road race covers the full length of every road on the network in any order or direction chosen by the riders. A rider may pass through each checkpoint more than once, but must travel along each road exactly once.

b One competitor claims this cannot be done. Explain why it *is* possible to travel every road once only during this race.

c If the race begins at Amity, where must this race finish?

d One of the competitors is following this path: Amity–P–Bevin–T–S–Bevin. Which checkpoint should *not* be visited next by this competitor? Explain why.

A road race for junior riders begins at Amity and ends at Carter. Participants are allowed to take any route they prefer.

e Find the shortest path from Bevin to Carter.

f Using your answers to parts **a** and **e**, what is the shortest distance from Amity to Carter?

The Water Authority wants to lay water mains along the roads in order to put a fire hydrant at the locations of the checkpoints in the diagram above. A minimal spanning tree will be used for these water mains.

g Draw the minimum spanning tree for the diagram above.

based on VCAA (2002)

ISBN 978-1-108-46154-2

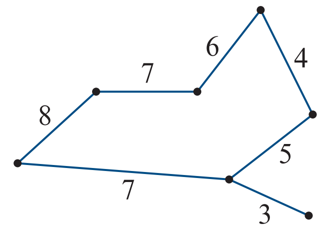
Cambridge University Press

Revision

13C Topic 3 Networks and decision mathematics

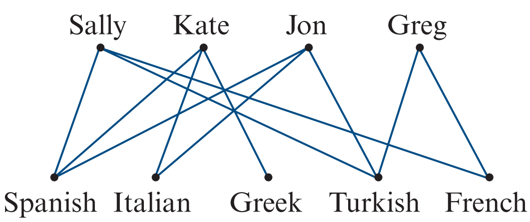
Multiple-choice questions

1 The sum of the weights of the minimum spanning tree of the weighted graph is:



A 2 **B** 30 **C** 32
D 33 **E** 35

2 Which one of the following statements is not implied by this bipartite graph?



A There are more translators of French than Greek.
B Sally and Kate can translate five languages between them.
C Jon and Greg can translate four languages between them.
D Kate and Jon can translate more languages between them than can Sally and Greg.
E Sally and Jon can translate more languages between them than can Kate and Greg.

3 Five people are to be each allocated one of five tasks (A, B, C, D, E). The table shows the time, in hours, that each person takes to complete the tasks. The tasks must be completed in the least possible total amount of time. If no person can help another, Francis should be allocated task:

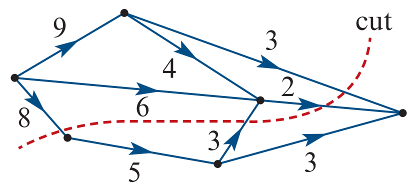
Name	A	B	C	D	E
Francis	12	15	99	10	14
David	10	9	10	7	12
Herman	99	10	11	6	12
Indira	8	8	12	9	99
Natalie	8	99	9	8	11

A A **B** B **C** C **D** D **E** E

ISBN 978-1-108-46154-2

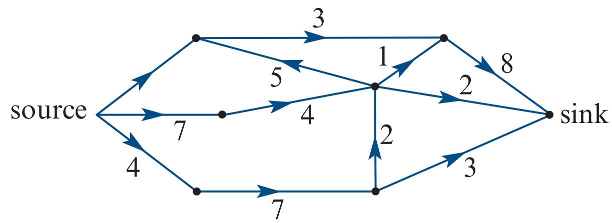
Cambridge University Press

4 The capacity of the cut in the flow network shown is:



A 0 **B** 2 **C** 10
D 13 **E** 16

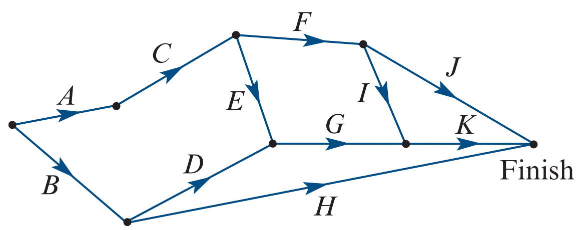
5 The flow network on the right shows connected water pipes represented as edges. The arrows show the direction of flow of the water through the pipes. The weights on the edges represent the maximum flow of water, in kilolitres per minute, through each pipe.



What is the maximum rate of flow of water that is possible from the source to the sink?

A 2 **B** 3 **C** 9 **D** 13 **E** 14

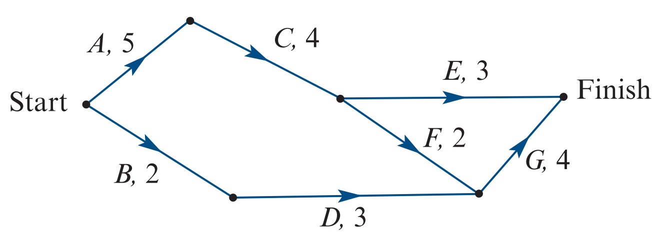
6 An activity network for a particular project is shown on the right. Activity C is an immediate predecessor of:



A activity E only
B activity F only
C activities E and F
D activities E, F and G
E activities E, F, G and I

Use the following information to answer questions 7, 8 and 9.

The activity network for a particular project is shown below. The duration of each activity, in hours, is shown on the network.



7 The earliest time that activity G can begin is:

A 4 **B** 5 **C** 9 **D** 11 **E** 13

8 What is the shortest number of days it will take to complete this project?

A 9 **B** 12 **C** 15 **D** 18 **E** 19

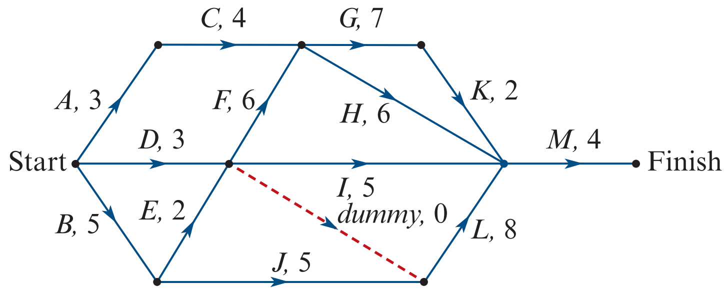
ISBN 978-1-108-46154-2

Cambridge University Press

9 What is the latest possible time that activity E can begin without delaying the completion of the entire project?

A 3 **B** 9 **C** 11 **D** 12 **E** 13

10 The activity network for a particular project is shown below.



What is the critical path for this project?

A $A - C - H - M$

B $A - C - G - K - M$

C $B - E - dummy - L - M$

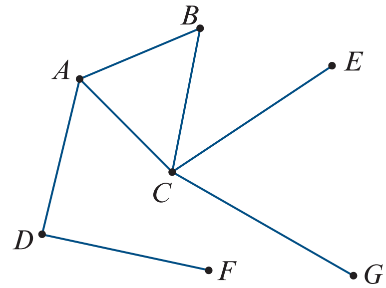
D $B - E - F - H - M$

E $B - E - F - G - K - M$

Short-answer questions

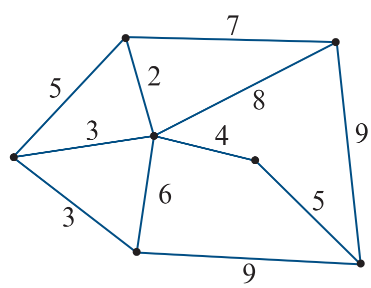
Simple familiar questions

1 Consider the graph shown below.
One edge can be removed from this graph in order to leave a spanning tree.
Between which pair(s) of vertices is this edge?



2 **a** Apply Prim's algorithm to draw the minimum spanning tree for the graph below.

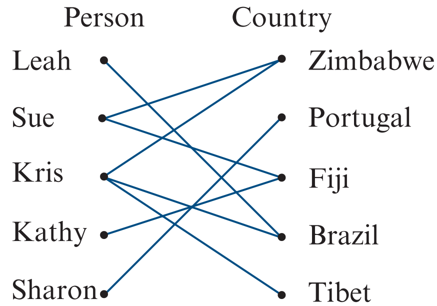
b What is the weight of this minimum spanning tree?



ISBN 978-1-108-46154-2

Cambridge University Press

3 The bipartite graph on the right shows the countries (Zimbabwe, Portugal, Fiji, Brazil and Tibet) that have been visited by five people (Leah, Sue, Kris, Kathy and Sharon).



- **a** How many countries has Leah visited?
- **b** Which person is the only one to have visited Portugal?
- **c** Each of the people will give a speech about their travels to only one country at a travel show. Make a list of each person and the country about which they will speak.

4 Roy runs a catering business with four employees, Ahmet, Beryl, Cynthia and Dorian. Each of these employees will be responsible for preparing one of the courses, canapes, starter, main or desert, for a dinner party.
The time that each employee is expected to take to prepare each of the courses is shown in the table below.

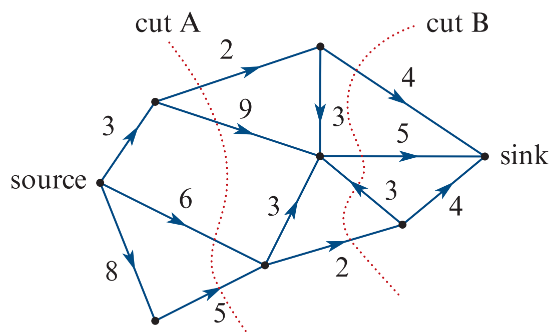
	Canapes	Starter	Main	Desert
Ahmet	1	2	4	4
Beryl	5	3	5	3
Cynthia	3	3	3	2
Dario	6	4	4	7

- **a** Which one of the employees is the quickest to prepare Canapes?
- **b** Apply the Hungarian algorithm to determine the allocation of employee to course so that the total time taken to prepare the dinner is as small as possible.
- **c** How long will Roy and his employees take to prepare the dinner if:
  - **i** the courses are prepared one after the other?
  - **ii** preparation of all courses starts at the same time?

ISBN 978-1-108-46154-2

Cambridge University Press

5 In the flow network on the right, the arcs represent pipes through which water can flow. The numbers on the arcs show the maximum rate at which water can flow through each pipe, in kilolitres per minute. Two cuts are shown on this network.

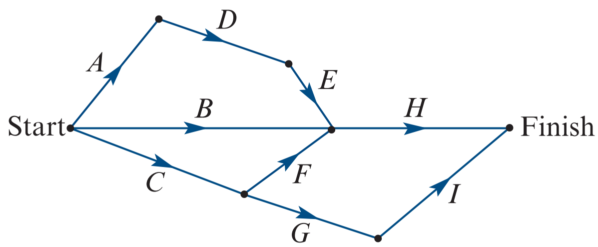


a Calculate the capacity of:

i cut A **ii** cut B

b The maximum rate at which water can flow from the source to the sink is 8 kilolitres per minute. Draw the cut with this minimum capacity on the network.

6 An activity network for a particular project is shown below.



a How many activities are required by this project?

b How many immediate predecessors does activity *H* have?

c Write down the activities for which activity *C* is an immediate predecessor.

7 A precedence table for a project is shown below.

Activity	Immediate predecessors	Duration (hours)
A	–	3
B	*A*	2
C	*A*	5
D	*B*	2
E	*C*	3
F	*E*	6
G	*D*, *F*	2

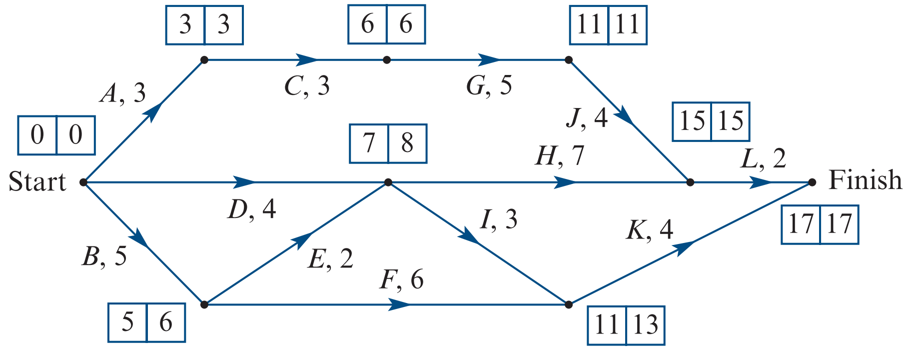
a Draw an activity network for this project.

b Perform the forward scanning process and determine the shortest time in which this project could be completed.

ISBN 978-1-108-46154-2

Cambridge University Press

8 The activity network for a project is shown below. All durations in this network are in hours.



The forward scanning and backward scanning processes have already been completed.

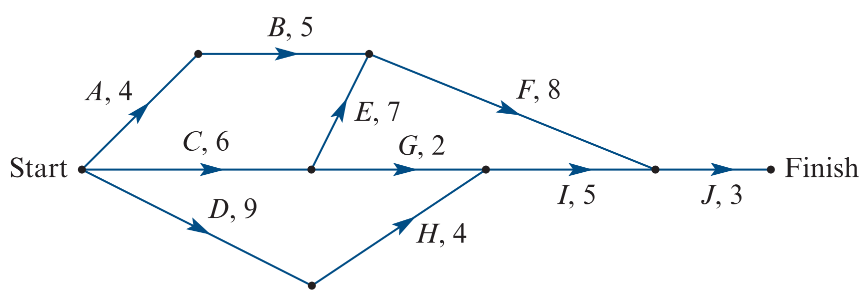
a How many hours should it take for this project to be completed?

b What is the earliest starting time for activity *H*?

c Calculate the float time for activity *E*.

d Write down the critical path for this project.

9 The activity network for a particular project is shown below. All durations in this network are in days.



a Construct a precedence table for this project.

b Complete the forward scanning process to determine the shortest time in which this project can be completed.

c Complete the backward scanning process to determine the critical path for this project.

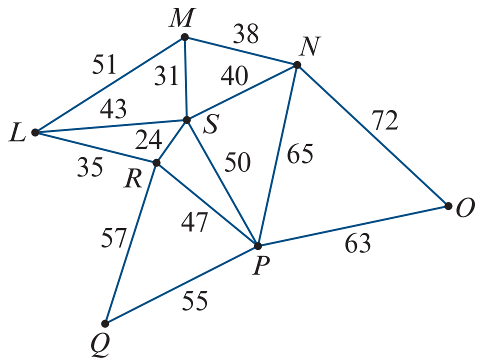
ISBN 978-1-108-46154-2

Cambridge University Press

Extended-response questions

Complex familiar questions

1 A number of towns need to be linked by pipelines to a natural gas supply. In the network shown, the existing road links between towns L, M, N, O, P, Q and R and to the supply point, S, are shown as edges. The towns and the gas supply are shown as vertices. The distances along roads are given in kilometres.



a What is the shortest distance along roads from the gas supply point S to the town O?

b The gas company decides to run the gas lines along the existing roads. To ensure that all nodes on the network are linked, the company does not need to place pipes along all the roads in the network.

i What is the usual name given to the network within a graph (here, the road system) which links all nodes (towns and supply) and which gives the shortest total length?

ii Sketch this network.

iii What is the minimum length of gas pipeline the company can use to supply all the towns by running the pipes along the existing roads?

2 Margaret has four grandchildren, Tyson, Emma, Gregory and Rose. Margaret has four chocolate bars (Flakey, Cherry Chomp, Honey Crunch and Snacker) and will give one to each of her grandchildren.

- Tyson likes Flakey and Snacker
- Emma only likes Flakey
- Gregory likes all of the chocolates
- Rose likes every chocolate except Cherry Chomp

a Draw a bipartite graph with the grandchildren on the left and the chocolate bars on the right.

b Which grandchild will receive the Snacker?

c How will Margaret distribute the chocolate bars so that every grandchild receives one they like?

ISBN 978-1-108-46154-2

Cambridge University Press

Revision

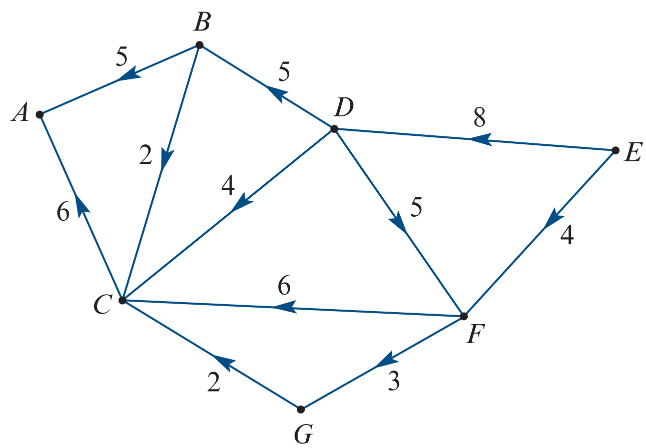
3 Camp sites A, B, C and D are to be supplied with food. Four local residents, W, X, Y and Z, offer to supply one campsite each. The cost in dollars of supplying one load of food from each resident to each campsite is tabulated.

Camp site	W	X	Y	Z
A	30	70	60	20
B	40	30	50	80
C	50	40	60	50
D	60	70	30	70

a Find the two possible matchings between campsites and residents so that the total cost is a minimum.

b State this minimum cost.

4 A reservoir at E pumps water through pipes along the flow network shown below.



The capacity of each pipe, in megalitres per day, is shown as weights on the arcs in the network.

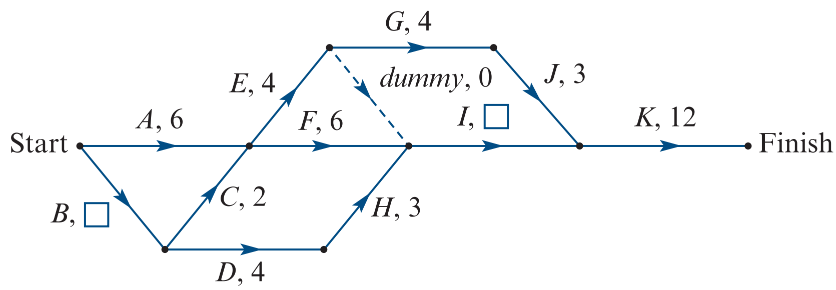
a What is the maximum flow of water that can reach the sink at A from the source at E?

b How many litres of water will flow into the sink over a three-day period?

ISBN 978-1-108-46154-2

Cambridge University Press

5 A particular project requires 11 individual activities to complete.
The activity network below shows these activities represented as arcs and the durations of the activities represented as weights on those arcs. Some of the activity durations are missing.



a Explain what the dummy activity indicates for this project.

b Complete the table of values below.

Activity	Duration (days)	EST	LFT	LST	Float
A	6	0	7	1	1
B	□	0	5	0	0
C	2	5	7	5	□
D	4	5	13	9	4
E	4	□	11	7	0
F	6	7	16	10	3
G	4	11	15	□	0
H	3	9	16	13	□
I	□	13	18	16	3
J	3	15	18	15	0
K	12	18	30	18	0

c What is the critical path for this activity?

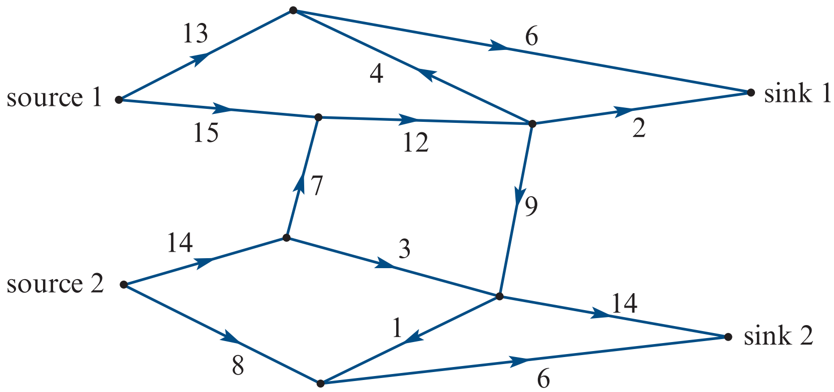
d Explain what the critical path means for this entire project.

ISBN 978-1-108-46154-2

Cambridge University Press

Complex unfamiliar questions

6 Water pipes of different capacities are connected to two water sources and two sinks. The network of water pipes is shown in the diagram below. The numbers on the edges represent the capacities, in kilolitres per minute, of the pipes.

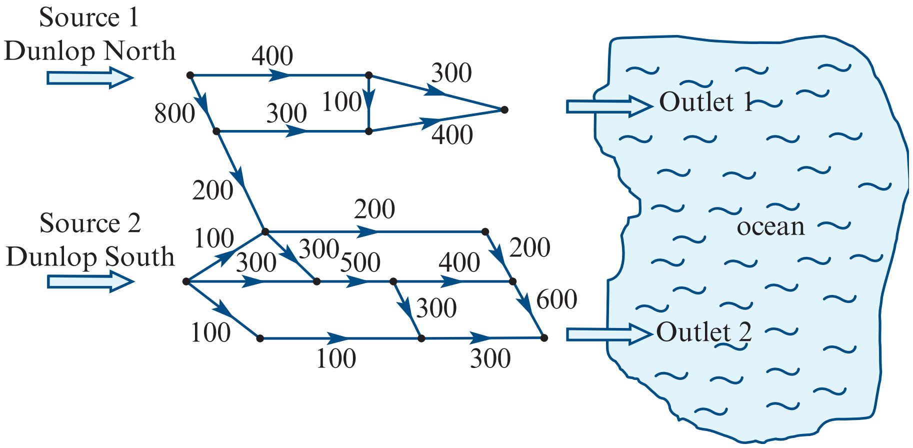


Find the maximum flow, in kilolitres per minute, to each of the sinks in this network.

7 Storm water enters a network of pipes at either Dunlop North (Source 1) or Dunlop South (Source 2) and flows into the ocean at either Outlet 1 or Outlet 2.

On the network diagram below, the pipes are represented by straight lines with arrows that indicate the direction of the flow of water. Water cannot flow through a pipe in the opposite direction.

The numbers next to the arrows represent the maximum rate, in kilolitres per minute, at which storm water can flow through each pipe.



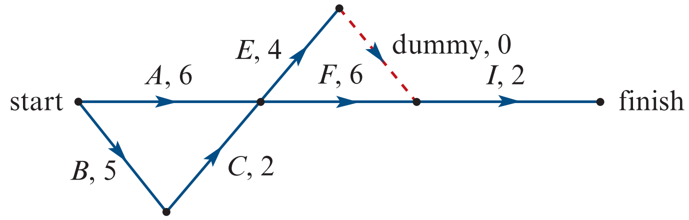
Determine the maximum rate, in kilolitres per minute, at which water can flow from these pipes into the ocean at Outlet 1 and Outlet 2.

©VCAA (Further Maths, 2, 2011)

ISBN 978-1-108-46154-2

Cambridge University Press

8 The activity network for a project is shown in the diagram below. The duration for each activity is in hours.



a Determine the shortest time in which this project can be completed.

b Write down the critical path for this project.

c The time it takes to complete activity A can be reduced by one hour at a cost of \$50. Explain why this will not affect the completion time of this project.

Activity B can be reduced in time at a cost of \$100 per hour. Activity F can be reduced in time at a cost of \$50 per hour.

d What is the cost of reducing the completion time of this project as much as possible?

based on VCAA (2004)

ISBN 978-1-108-46154-2

Cambridge University Press

13D List of Unit 4 and Units 3 & 4 assessment and examination practice online items

These assessment practice items can be found in the interactive textbook and in the teacher resources of the online teaching suite.

Interactive Textbook

For student and teacher access:

1 IA3: A practice internal examination on Unit 4

2 EA: a practice external examination on Units 3&4

Online Teaching Suite

For teacher access:

1 IA3: An internal examination on Unit 4

2 EA: a practice external examination on Units 3&4

Assessment items for Unit 3 are listed at the end of Chapter 6.

ISBN 978-1-108-46154-2

Cambridge University Press

Answers

Chapter 1

Exercise 1A

1 a Categorical
b Categorical
c Numerical
d Numerical
e Categorical
f Numerical
g Numerical
h Categorical
i Categorical

2 a Two categorical variables
b One categorical and one numerical
c Two numerical
d Two categorical

3 a EV: *colour*; RV: *toxicity*
b EV: *type of diet*; RV: *weight loss*
c EV: *age*; RV: *price*
d EV: *fuel*; RV: *cost*
e EV: *location*; RV: *house price*

4 a *Age*
b *Years of education*
c *Temperature*
d *Time of year*
e *Age group*
f *State of residence*

5 a *Sex* – categorical, EV; *attitude to lowering the drinking age* – categorical, RV
b *Hours of study* – numerical; *hours spent using social media* – numerical. Either variable could be the EV or RV, it would depend on the question asked.
c *Gestation time* – numerical, EV; *birth weight* – numerical, RV
d *Sex* – categorical, EV; *hours spent using social media* – numerical, RV
e *Voting preference* – categorical, *support for tax cuts* – categorical. Either variable could be the EV or RV, it would depend on the question asked.

Exercise 1B

1 a *Sex* is the EV, *intention to go to university* is the RV.

b

		Sex		
		F	M	Total
Intend to go to university	No	4	4	8
	Yes	8	4	12
Total		12	8	20

2 a *Age group* is the EV, *reduce fees* is the RV.

b

	Age group			
Reduce fees	17–18	19–25	26 or more	Total
No	3	3	4	10
Yes	8	6	6	20
Total	11	9	10	30

c

	Age group		
Reduce fees	17–18	18–25	26 or more
No	27.3%	33.3%	40.0%
Yes	72.7%	66.7%	60.0%
Total	100.0%	100.0%	100.0%

3 a *Enrolment status*
b No. The percentage of full-time and part-time students who drank alcohol is similar: 80.5% to 81.8%. This indicates that drinking behaviour is not related to enrolment status.

ISBN 978-1-108-46154-2

Cambridge University Press

4 a *Sex* is the EV, *are you satisfied with your life overall?* is the RV.

b

	Sex of respondent	
Are you satisfied with your life overall?	Female	Male
Yes	86.4%	84.7%
No	13.6%	15.3%
Total	100.0%	100.0%

There is no association between satisfaction with life overall and the sex of the respondent. Very similar percentages of males and females report that they are satisfied with their life overall (Females: 86.4%, Males: 84.7%).

5 a *Handedness*

b

	Sex	
Handedness	Male	Female
Left	9.0%	9.8%
Right	91.0%	90.2%
Total	100.0%	100.0%

c No, there is little difference in the percentage of males and females who are left handed, 9.0% compared to 9.8%.

6 a *Sex*

b 54.9%

c There are several ways you can answer the question.
Focusing on the category 'rarely'.
Yes; the percentage of males who rarely exercised (28.8%) was significantly less than the percentage of females who rarely exercised (39.2%).
or
Yes; the percentage of males who exercised regularly (18.6%) was significantly higher than the percentage of females who exercised regularly (5.9%).
Note: For the category 'sometimes', there is no association between level of exercise and sex.

7 a *Satisfaction with job* is the EV, *satisfaction with life* is the RV.

	Satisfaction with job	
Satisfaction with life	Dissatisfied	Satisfied
Dissatisfied	75.0%	22.6%
Satisfied	25.0%	77.4%
Total	100.0%	100.0%

b The data supports the contention that those people who are satisfied with their jobs are more likely to be satisfied with their lives. Of those people who are satisfied with their job, 77.4% were satisfied with life, while a much lower percentage 25.0% of people who were dissatisfied with their job were satisfied with their lives.

8 *Class* is the EV, *exam grade* is the RV.

	Class	
Exam grade	Dr Evans	Dr Smith
Fail	11.1%	9.4%
Pass	61.1%	62.5%
Credit or above	27.8%	28.1%
Total	100.0%	100.0%

The data does not support the suggestion that Dr Evans's maths class achieves higher grades than Dr Smith's maths class. Both classes achieved similar percentages in each group at all exam grade levels, with 61.1% of Dr Evans's class achieving a pass, which is very close to the 62.5% who achieved a pass in Dr Smith's class.

9 *Type of treatment* is the EV, *outcome of treatment* is the RV.

	Type of treatment	
Outcome of treatment	Drug	Pillow
Complete cure	9.8%	31.3%
Partial cure	26.8%	37.5%
No improvement	63.4%	31.3%
Total	100.0%	100.0%

ISBN 978-1-108-46154-2

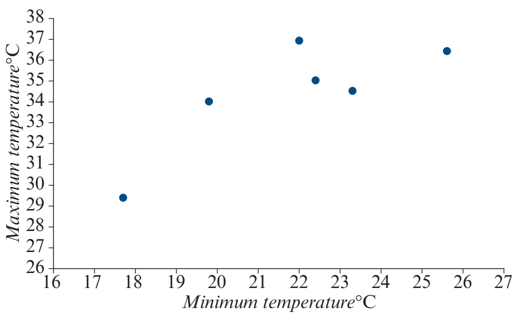
Cambridge University Press

The data supports the contention that the special pillow would be more effective in the treatment of snoring than the treatment with drugs. A much higher percentage of those using the pillow experienced a complete cure compared to those using the drug treatment (31.3% compared to only 9.8%).

Exercise 1C

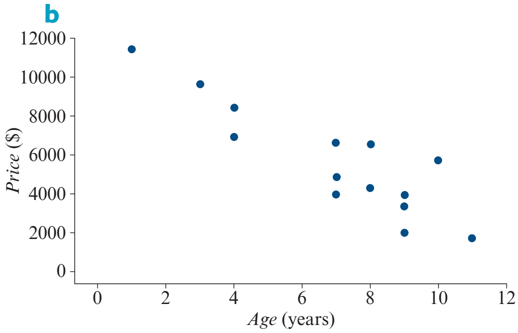
1 **a** Number of seats **b** Numerical
c 8 aircraft **d** Around 800 km/h

2



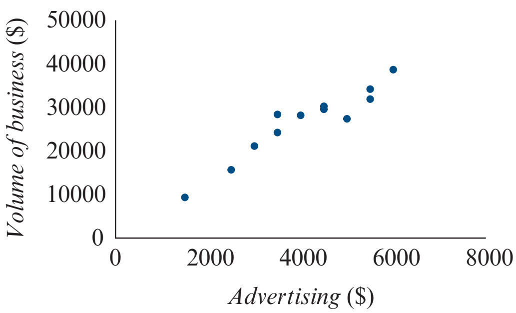
3 **a** *Age* is the EV and *price* is the RV.

b

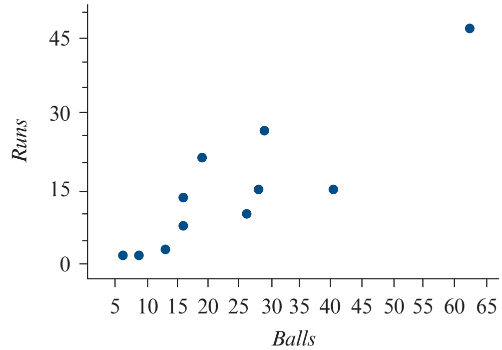


4 **a** *Advertising* is the EV and *volume of business* is the RV.

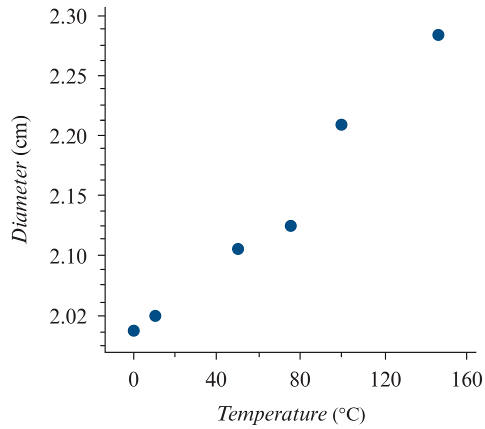
b



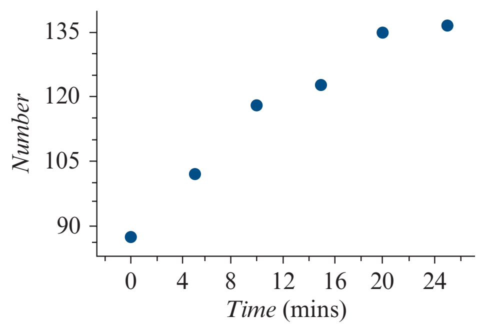
5



6



7



Exercise 1D

1 Note: There are no absolute right or wrong answers to these questions as answering them requires a degree of personal judgement.
a No association **b** Yes, positive
c Yes, positive **d** Yes, positive
e Yes, negative **f** Yes, negative

2 **a, b, c** **i** Positive, linear, moderate
ii Negative, linear, weak
iii Positive, linear, strong
iv No association

ISBN 978-1-108-46154-2

Cambridge University Press

3 a Positive **b** Linear
c Hard to classify with so little data

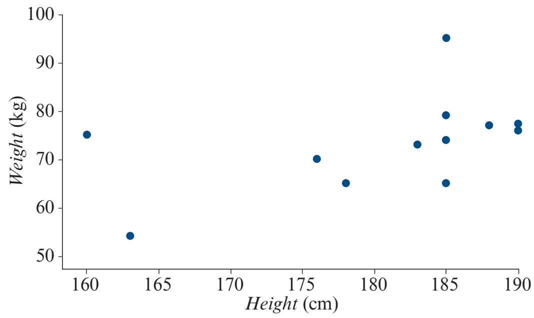
4 a Positive **b** Linear **c** Strong

5 a Negative **b** Linear **c** Strong

6 a

b Positive, with an outlier at (25, 35)
c Linear **d** Moderate

7 a



b Positive, with an outlier at (160, 75)
c Linear, although it's hard to tell with no data in the range 160–175 cm.
d Moderate

Exercise 1E

1 a A: strong, positive, non-linear relationship with no outliers
B: strong, negative, linear relationship with an outlier
C: weak, negative, linear relationship with no outliers
b A: non-linear B: outlier

2 0.5477 **3** 0.7475

4 0.2869 **5** 0.73

6 0.972

7 a–c Answers given in question.

8 a −0.889 **b** Strong

9 a 0.607 **b** Moderate

10 a 0.482 **b** Weak

Exercise 1F

1 a 45.6% **b** 11.9%
c 32.1% **d** 45.3%
e 1.5%

2 a $r = 0.906$ **b** $r = -0.353$

3 a The coefficient of determination is $r^2 = (-0.611)^2 = 0.373$ or 37.3%; that is, 37.3% of the variation observed in hearing test scores can be explained by variation in age.
b The coefficient of determination is $r^2 = (0.716)^2 = 0.513$ or 51.3%; that is, 51.3% of the variation observed in mortality rates can be explained by variation in smoking rates.
c The coefficient of determination is $r^2 = (-0.807)^2 = 0.651$ or 65.1%; that is, 65.1% of the variation observed in life expectancies can be explained by variation in birth rates.
d The coefficient of determination is $r^2 = (0.818)^2 = 0.669$ or 66.9%; that is, 66.9% of the variation observed in daily maximum temperature is explained by the variability in daily minimum temperatures.
e The coefficient of determination is $r^2 = (0.8782)^2 = 0.771$ or 77.1%; that is, 77.1% of the variation in the runs scored by a batsman is explained by the variability in the number of balls they face.

4 a $R^2 = 36\%$, thus 36% of variation in the weekly expenditure on food can be explained by the variation in income.
b $R^2 = 25\%$, thus 25% of variation in the weekly expenditure on leisure can be explained by the variation in income.
c People's incomes are associated with their expenditure on both leisure and food, but it is more powerful in explanation of expenditure on food.

Chapter 1 review

Multiple-choice questions

1 A **2** D **3** B **4** D **5** C
6 B **7** D **8** E **9** C **10** C
11 C **12** D **13** A

ISBN 978-1-108-46154-2

Cambridge University Press

Short-answer questions

1 **a** Two categorical
b One categorical and one numerical
c Two numerical

2 **a** EV: sex, RV: belief in Astrology
b EV: Sex, RV: weight
c EV: hrs in paid employment, RV: exam mark

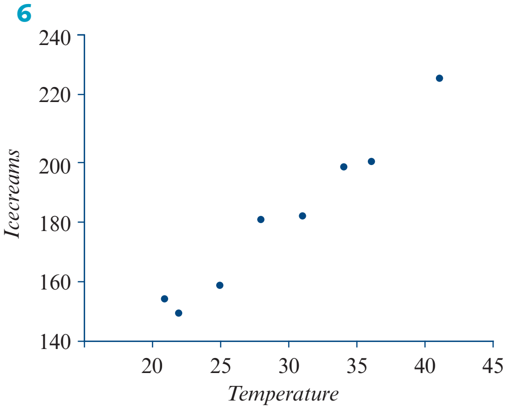
3 Both variables are categorical, EV; Sex is labelling columns, RV: Intend to go to university is labelling rows.

4

Intend going to university?	Male	Female
Yes	67%	77%
No	33%	23%
Total	100%	100%

5 In this sample of 300 year 12 students we see than females are more likely to intend to go to university than males (77% of females, 67% of males).

6



7 **a** $r = 0.984$ **b** strong

8 **a** There is a moderate, negative, linear relationship between government expenditure on health and infant mortality. Those countries which spend more on health tend to have lower infant mortality dates. There is one country (14, 36) which is a possible outlier - this country seems to have a higher infant mortality than is indicated by the health expenditure.
b Yes, because both variables are numerical, and relationship is linear

9 $r = 1$

10 covariance = 4.328

11 **a** $R^2 = 59.3\%$
b 59.3% of the variation in hearing test scores in explained by the variation in age

12 **a** There is a moderate, positive, linear relationship between hours of sunlight and height of the seedlings. Those seedlings which get more sunlight tend to be taller. There is one plant (4.8, 11.5) which is a possible outlier - this plant seems to be taller that would be indicated by the number of hours of sunlight.
b The value of r will be closer to 1, indicating a stronger correlation.

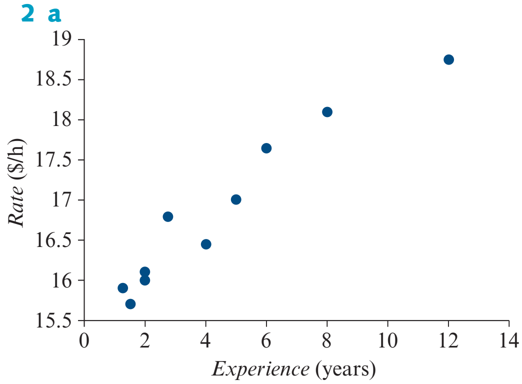
Extended-response questions

1 **a** *Number of accidents* and *age*; both categorical variables
b RV: *Number of accidents*; EV: *age*
c 470
d

Number of accidents	Age < 30	Age ⩾ 30
At most one accident	21.7%	42.5%
More than one accident	78.3%	57.5%

e The statement is correct. Of drivers aged less than 30, 78.3% had more than one accident compared to only 57.5% of drivers in the older category.

2 **a**



'Rate' is the response variable.
b There is a strong, positive, linear relationship; that is, people with more experience are generally being paid a higher starting pay rate. There are no apparent outliers. $r \approx +0.9$.
c 0.967
d Coefficient of determination = 0.934; that is, 93.4% of variation in pay rate is explained by the variation in experience.

ISBN 978-1-108-46154-2

Cambridge University Press

3 a 11.9%

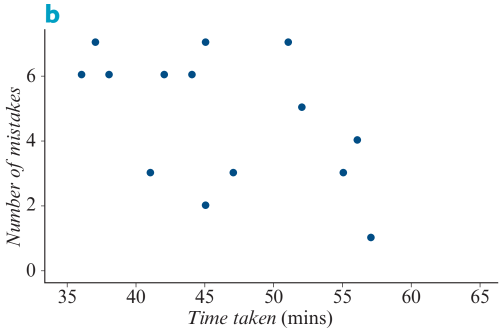
b 52.3%

c Marital status

d Yes. There are several ways that this can be seen. For example, by comparing the married and widowed groups, we can see that a smaller percentage of those widowed found life exciting (33.8%) compared to those who were married (47.6%). Or: a bigger percentage of widowed people found life pretty routine (54.3% to 48.7%) and dull (11.9% to 3.7%) compared to those who were married.

4 a *Time taken* is EV, *number of mistakes* is RV.

b



There is a moderate, negative, linear relationship between time taken and the number of mistakes – those who take longer tend to make fewer mistakes.

c Yes, because both variables are numerical, and the relationship is linear.

d −0.635

e 40.3% of the variation in the number of mistakes made can be explained by the variation in time taken.

Chapter 2

Exercise 2A

1 A residual is the difference between a data value and its value predicted by a regression line.

2 C

3 The data is numerical; the association is linear; there are no clear outliers.

4 a RV: *pollution level*; EV: *traffic volume*

b *pollution level* = −330 + 49 × *traffic volume*

5 a RV: *life expectancy*; EV: *birth rate*

b *life expectancy* = 110 − 1.5 × *birth rate*

6 a RV: *distance travelled*; EV: *age*

b *distance travelled* = 16 + 11 × *age*

7 a r is also negative.

b Slope is zero: regression line is horizontal.

c Intercept = $\bar{y}$ (mean of RV)

8–9 Answers given in question.

10 a Answer given in question.

b *runs* = −2.6 + 0.73 × *balls faced*

11 a RV: *number of TVs*

b Answer given in question.

c *number of TVs* = 61.2 + 0.930 × *number of cars*

Exercise 2B

1 *mark* = 80 − 4.3 × *days absent*

2 A: clear curved pattern in the residuals (not random)

C: clear curved pattern in the residuals (not random)

3 a *height* **b** 0.33, 2.9

c 55.7 **d** 2.8

4 a *fuel consumption* **b** 0.01, − 0.1

c 9.7 **d** −0.8

5 a 14.7, 27.8 **b** 14.7

c 0.87 **d** 75.7, fat content

e i 145.4

ii −13.4

6 a −0.278: the slope predicts that success rate decreases by 27.8% for each additional metre the golfer is from the hole.

b 73.5 **c** 3.54 m **d** −0.705

e 49.7%: 49.7% of the variation in success rate in putting is explained by the variation in the distance the golfer is from the hole.

7 a Yes, linear relationship

b 0.9351 or 93.5%

c 93.5%

d *pay rate* = 8.56 + 0.289 × *experience*

e The pay rate for a worker with no experience

f On average, the pay rate increases by 29 cents per hour for each additional year of experience.

g i $10.87 **ii** $0.33

h Yes; no clear pattern in the residual plot indicating that there are no further underlying trends.

ISBN 978-1-108-46154-2

8 a $r = -0.608$

b 37% of the variation in the hearing test score is explained by the variation in age.

c *hearing test score* = 4.9 − 0.043 × *age*

d −0.043; the hearing test score, on average, decreases by 0.043 as age increases by 1.

e i 4.04 **ii** −2.04

f i 0.3 **ii** −0.4

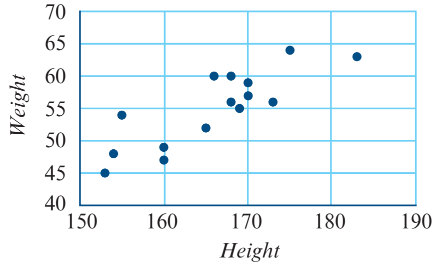
g Yes; no clear pattern in the residual plot indicating that there are no further underlying trends.

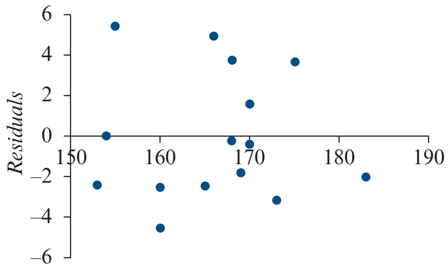
9 a Yes, as the variables are both numerical and the relationship is linear. This can be seen in both the scatterplot and the residual plot.

b The y-intercept would give the estimated height of a person with an arm span of 0 cm, which is not meaningful.

c From the scatterplot, we can see that there is a strong, positive, linear relationship between height and arm span: $r = 0.951$. There are no obvious outliers. The equation of the least squares regression line is *height* = 32.97 + 0.807 × *arm span*
The slope of the regression line predicts an increase of 0.81 cm in height for each 1 cm increase in arm span.
The coefficient of determination indicates that for this sample 90.4% of the variation in height is explained by the variation in arm span.

10 a





b From the scatterplot, we can see that there is a strong, positive, linear relationship between height and weight: $r = 0.840$.
There are no obvious outliers. The equation of the least squares regression line is *weight* = −42.35 + 0.587 × *height*.
The slope of the regression line predicts an increase of 0.59 kg in weight for each 1 cm increase in height.
The coefficient of determination indicates that for this sample 70.5% of the variation in weight is explained by the variation in height.

Exercise 2C

Note: These answers are for guidance only. Alternative explanations for the source of an association may be equally acceptable as the variables suggested.

1 Not necessarily. In general, older children are taller and have been learning mathematics longer. Therefore they tend to do better on mathematics tests. Age is the probable common cause for this association.

2 Not necessarily. While one possible explanation is that religion is encouraging people to drink, a better explanation might be that towns with large numbers of churches also have large populations, thus explaining the larger amount of alcohol consumed. Town size is the probable common cause for this association.

3 Probably not. The amount of ice cream consumed and the number of drownings would both be affected by weather conditions. Weather conditions are the probable common cause.

4 Maybe but not necessarily. Bigger hospitals tend to treat more people with serious illnesses and these require longer hospital stays.
A common cause could be the type of patients treated at the hospital.

Exercise 2D

1 Scatterplot shows relationship is linear. For the total data set, $r = 0.871$. For the females $r = 0.876$, for the males $r = 0.614$, so the relationship is stronger for females than males.

2 In 1990 agreement with this statement for both males and females was similar in Australia and the UK (males Aust 23%, males UK 30%, females Aust 25%, females UK 24%) , but both much lower than that of both males and females in Europe (males Europe 37%, females Europe 28%). In 2010 agreement with

ISBN 978-1-108-46154-2

Cambridge University Press

the statement has lowered considerably for both sexes across all three geographies, but the relationship had changed for males, with males in Australia much less likely to agree than those in both the UK and Europe (males Aust 14%, males UK 25%, males Europe 30%).

3&4 Answers will vary.

Chapter 2 review

Multiple-choice questions

1 C	**2** D	**3** A
4 B	**5** B	**6** D
7 C	**8** A	**9** A
10 A	**11** D	**12** E
13 A	**14** C	**15** A
16 E	**17** C	

Short-answer questions

1 Both variables are numerical, the association is linear, there are no outliers.

2 $a = 91.725$, $b = 2.625$

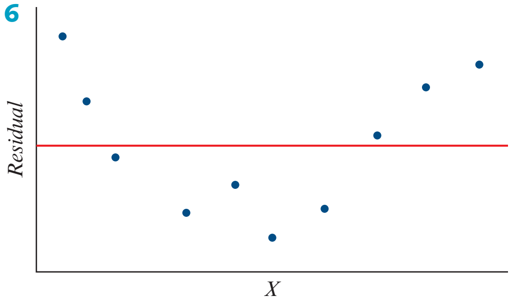
3 **a** Slope = 0, the line is horizontal.

b Intercept = $\bar{y}$, the mean of the RV.

4 Number of ice creams = 58.2 + 4.1 × temperature

5 Residual = 600

6



7 **a** On average, price decreases by \$5675 per year.

b On average, the price of a new caravan is \$87 500.

8 **a** There is a weak positive correlation between educational attainment and the amount spent on education ($r = 0.23$). Those countries which spend more on education also tend to have higher educational attainment.
There is a moderate negative correlation between educational attainment and the student : teacher ratio ($r = -0.34$). Those countries with a higher student : teacher ratio tend to have lower educational attainment.

b Student: teacher ratio explains 11.6% of the variation in educational attainment, making it a much more important explanatory variable than amount spent on education, which explains only 5.3%.

9 **a** 3.91 secs **b** 3.19 secs

c Predicting for a person 55 years of age is interpolation, and we can be reasonably confident that this prediction will be reliable. Predicting for a person 35 years of age is extrapolation, and we cannot be confident of this prediction.

10 **a** score

b slope is 0.34, intercept is 32

c 70.1 **d** 7.9

11 Correlation implies that two variables have been observed to vary together, either one increasing as the other increases (positive correlation) or one increasing as the other decreasing (negative correlation). It may be that this is because there is a causal relationship between the variables (for example, the further we drive the less fuel there will be in the tank of the car) or it could be because there is a third variable which is affecting both (e.g. age will affect both height and scores on and IQ test, meaning that although there might be a high correlation between height and IQ scores in children, we do not think taller people are smarter).

12 Causal relation does not exist. It may be that only mature people are likely to be prepared to have children. It is also possible that the underlying cause is age. Higher age leads to both having children and higher maturity levels.

Extended-response questions

1 **a** days of rain **b** −6.88, 2850

c 2024 **d** decrease, 6.88

e −0.696 **f** 48.4 , days of rain

g **i** 1873 **ii** −483

h interpolation

2 **a** *Cost*

b Answer given in question.

c **i** \$182.30, interpolating

ii \$125.60, extrapolating

d **i** 81.5: The fixed costs of preparing meals is \$81.50.

ii \$2.10: The slope of the regression line predicts that, on average, meal preparation costs increase by \$2.10 for each additional meal produced.

ISBN 978-1-108-46154-2

e 0.956; 95.6% of the variation in the cost of preparing meals is explained by the variation in the number of meals produced.

3 a *Male income*

b \$350

c **i** \$18 250

ii Making the prediction involves going well beyond the data used to determine the regression equation (extrapolation). We have no way of knowing whether the same relationship between male and female incomes still applies outside of this data.

4 a RV: *height*; EV: *femur length*

b *height* = 36.3 + 5.35 × *femur length*

c On average, height increases by 5.35 cm for each 1 cm increase in femur length.

d $r^2 = 0.988$; that is, 98.8% of the variation in height is explained by the variation in femur length.

5 a RV: *height*; EV: *age*

b Answer given in question.

c 83 cm, extrapolation

d On average, height increases by 6.4 cm for each extra year.

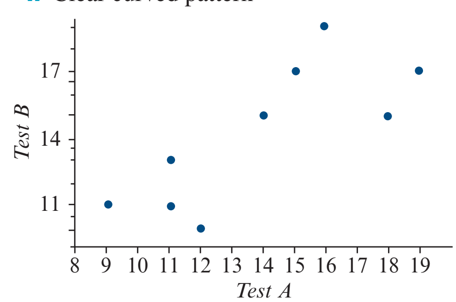
e $r^2 = 0.995$; that is, 99.5% of the variation in height is explained by the variation in age.

f **i** 140.3 cm **ii** −0.7 cm

g **i** Answer given in question.

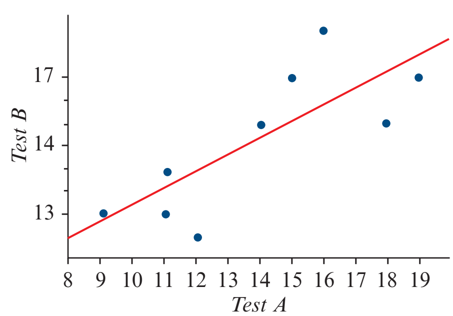
ii Clear curved pattern

6 a

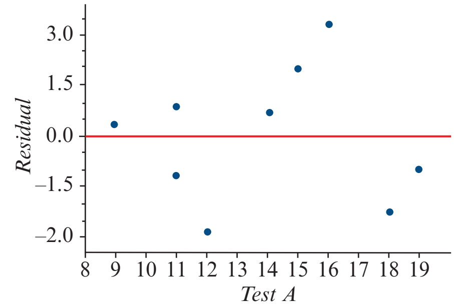


b *test B score* = 4.2 + 0.72 × *test A score*, $r = 0.78$, $R^2 = 0.61$

c

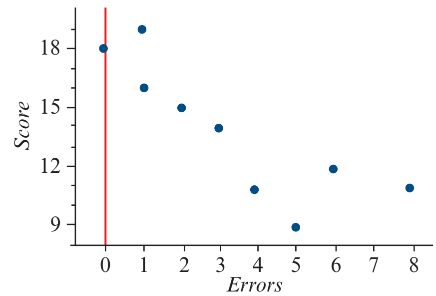


d



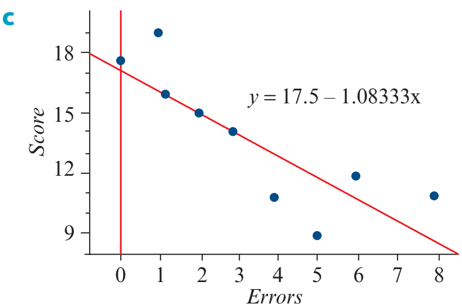
e Answers will vary.

7 a

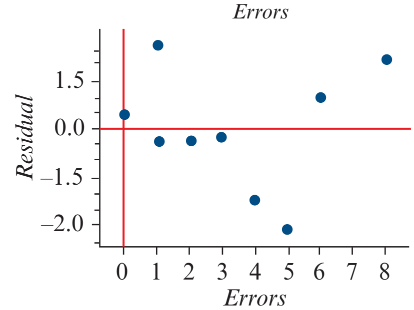


b *score* = 17.5 − 1.08 × *error*, $r = -0.841$, $R^2 = 0.707$

c



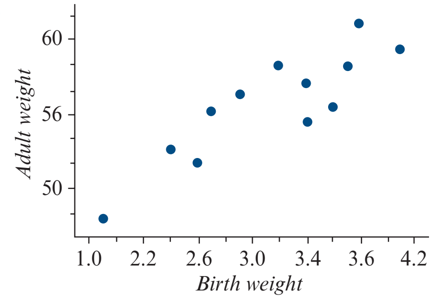
d



e Answers will vary.

8 a RV: *adult weight*; EV: *birth weight*

b



ISBN 978-1-108-46154-2

Cambridge University Press

c **i** Strong, positive, linear association with no outliers

ii Your judgement

d *adult weight* = 38.4 + 5.86 × *birth weight*, $R^2 = 0.765$, $r = 0.875$

e 76.5% of the variation in the adult weight is explained by the variation in birth weight.

f On average, adult weight increases by 5.9 kg for each additional kilogram of birth weight.

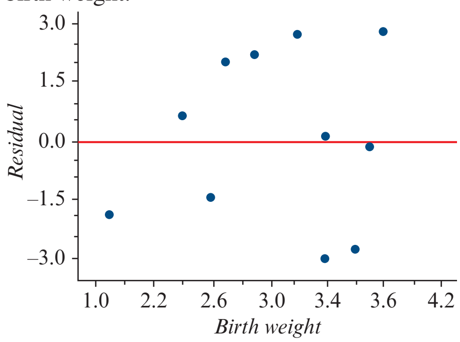
g **i** 56.0

ii 53.1

iii 61.3

h Yes. 76.5% of the variation in the adult weight is explained by the variation in birth weight.

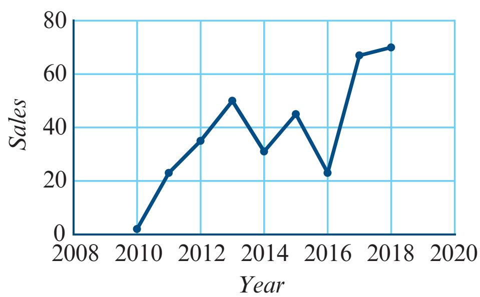
i



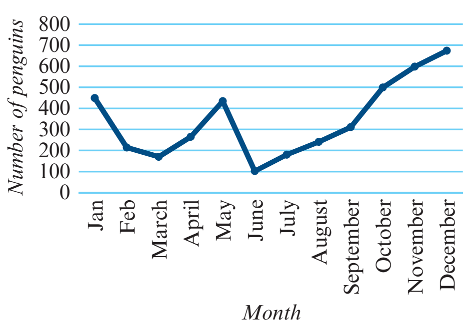
The lack of a clear pattern in the residual plot supports the assumption that the relationship between adult weight and birth weight is linear.

Exercise 3A

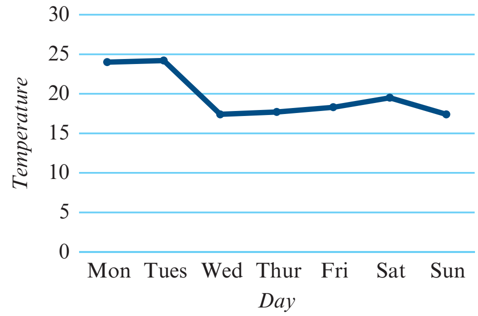
1



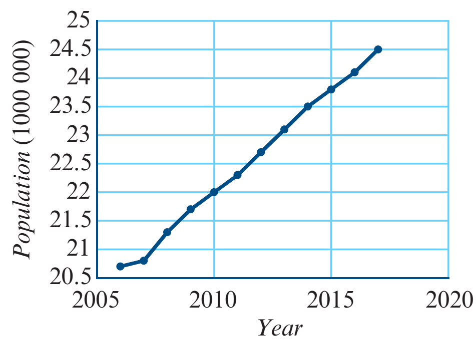
2



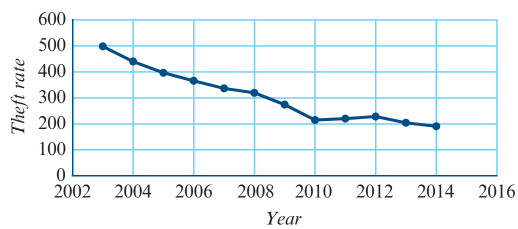
3



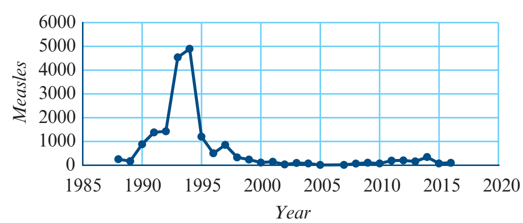
4



5



6



7

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 3B

1

Feature	Plot A	Plot B	Plot C
Irregular fluctuations	✓	✓	✓
Increasing trend			✓
Decreasing trend	✓		
Cycles	✓		
Outlier			✓

2

Feature	Plot A	Plot B	Plot C
Irregular fluctuations	✓	✓	✓
Increasing trend			✓
Decreasing trend	✓		
Cycles	✓	✓	
Seasonality	✓		✓

3

Characteristic	Plot A	Plot B	Plot C
Irregular fluctuations	✓		✓
Structural change	✓		
Increasing trend	✓		
Decreasing trend	✓		
Seasonality			✓

4 **a** The percentage of males who smoke has consistently decreased since 1945, while the percentage of females who smoke increased from 1945 to 1975 but then decreased at a similar rate to males over the period 1975–1992.

b Decrease

5 The number of whales caught increased rapidly between 1920 and 1930 but levelled off during the 1930s. In the period 1940–1945 there was a rapid decrease in the number of whales caught and numbers fell to below those of the 1920 catch. In the period 1945–1960 the numbers increased but then fell again from 1960–1985 when numbers were back to around the 1920 level.

6 The population of Australia shows a steady increasing trend over the years 2006–2017.

7 The motor vehicle theft rate shows steady decreasing trend over the years 2003–2010. From 2010 to 2014, the motor vehicle theft rate shows some variation from year-to-year, but overall remains fairly steady.

8 The number of measles cases in Australia started to increase in 1990, showing a peak in 1993–1994. In 1995 the number of cases reported started dropping back to pre-1990 levels, and since 2000 have remained steady at a low level. Note that there was a slight increase in the number of cases in 2013–2014.

9 The levels of carbon intensity vary randomly from year-to-year. Over the years 1960–1977, the overall levels seem quite steady, then there was a marked increase in carbon intensity between 1977 and 1980. From 1980–1990 the levels remained high. In 1990 levels dropped but not as low as pre-1980 levels. Levels remained steady from 1990 to 2012. Since 2012, levels appear to be reducing again.

Exercise 3C

1 C **2** A **3** B **4** D **5** A

6 **a** 3 **b** 1 **c** 4 **d** 3.2
e 1.2 **f** 2.2 **g** 3.75 **h** 2.0
i 3.25 **j** 1.5

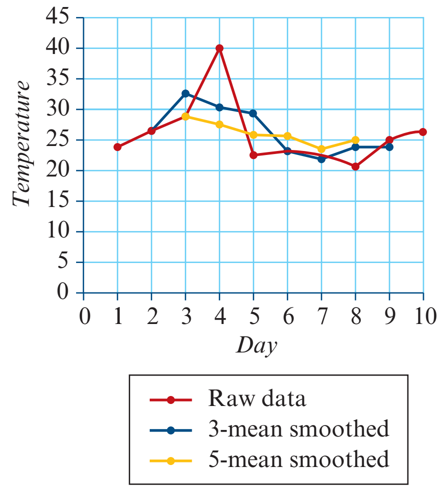
7

t	1	2	3	4	5	6	7	8	9
y	10	12	8	4	12	8	10	18	2
3-mean	–	10	8	8	8	10	12	10	–
5-mean	–	–	9.2	8.8	8.4	10.4	10	–	–

ISBN 978-1-108-46154-2

Cambridge University Press

8 a, c

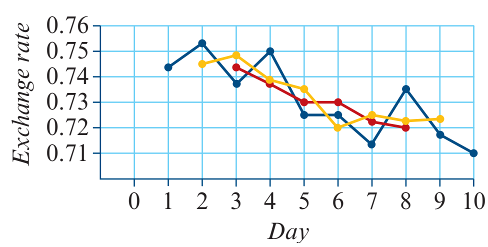


The smoothed plots show that the ‘average’ maximum temperature changes relatively slowly over the 10-day period (the 5-day average varies by only 5°) when compared to the daily maximum, which can vary quite widely (for example, nearly 20° between the fourth and fifth day) over the same period of time.

b

Day	Temperature (C°)	3-moving mean	5-moving mean
1	24	–	–
2	27	26.3	–
3	28	31.7	28.2
4	40	30.0	28.0
5	22	28.3	27.0
6	23	22.3	25.6
7	22	22.0	22.6
8	21	22.7	23.4
9	25	24.0	–
10	26	–	–

9 a, c



- exchange rate
- 3-moving mean exchange rate
- 5-moving mean exchange rate

The exchange rate has a downward trend over the 10-day period. This is most obvious from the smoothed plots, particularly the 5-moving mean plot.

b

Day	Exchange rate	3-moving mean	5-moving mean
1	0.743	–	–
2	0.754	0.745	–
3	0.737	0.747	0.742
4	0.751	0.737	0.738
5	0.724	0.733	0.730
6	0.724	0.720	0.729
7	0.712	0.724	0.722
8	0.735	0.721	0.720
9	0.716	0.721	–
10	0.711	–	–

10

Month	Number of births	2-moving mean	2-moving mean with centring
January	10		
		11	
February	12		10
		9	
March	6		7.25
		5.5	
April	5		9.5
		13.5	
May	22		16.75
		20	
June	18		17.75
		15.5	
July	13		12.75
		10	
August	7		9
		8	

ISBN 978-1-108-46154-2

Cambridge University Press

Month	Number of births	2-moving mean	2-moving mean with centring
September	9		8.75
		9.5	
October	10		9.25
		9	
November	8		10.25
		11.5	
December	15		

11

Month	Internet usage	4-moving mean	4-moving mean with centring
April	21		
May	40		
		38.75	
June	52		43.375
		48	
July	42		52.875
		57.75	
August	58		61.375
		65	
September	79		66.5
		68	
October	81		67
		66	
November	54		
December	50		

Exercise 3D

1 C **2** B **3** E **4** B
5 C **6** D **7** B **8** C
9 D **10** D **11** A **12** B

13

Number of students:	56	125	126	96
Deseasonalised numbers:	112	125	97	80
Seasonal index	0.5	1.0	1.3	1.2

14 a, c

Deseasonalised	152	142	148	153
Seasonal index	1.30	1.02	0.58	1.1

b In quarter 1 the restaurant chain employs 30% more waiters than the number employed in an average quarter.

15

Q1	Q2	Q3	Q4
0.89	0.83	1.12	1.16

16

Jan	Feb	Mar	April	May	June
0.89	0.96	1.04	1.26	1.33	1.11

July	Aug	Sept	Oct	Nov	Dec
0.67	0.74	0.59	0.81	1.11	1.48

17

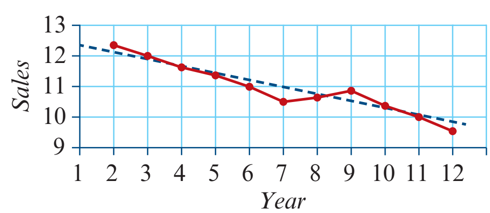
Jan	Feb	Mar	April	May	June
1.1	0.95	1.25	1.15	1	0.9

July	Aug	Sept	Oct	Nov	Dec
1	0.75	0.7	0.55	1.15	1.5

Exercise 3E

1 a *population* = −692.74 + 0.3556 × *year*
b 27.0 million

2 a, d



b General decreasing trend in the percentage of retail sales made in department stores

c *sales* = 12.5 − 0.258 × *year* (to 3 sig. figs)
The percentage of total retail sales that are made in department stores is decreasing by approximately 0.3% per year.

e 8.6%

ISBN 978-1-108-46154-2

Cambridge University Press

3 a The number of university students in Australia has increased steadily from 2006 to 2016.
number of university students (000's) = −72 247.92 + 36.373 × *year*

b The slope tells us that on average the number of university students is increasing by 36 373 per year.

c 1 407 400 to the nearest 100.

4 a deseasonalised number = 50.9 + 1.59 × quarter number

b deseasonalised number = 76.34
reseasonalised (actual) number = 90 (to the nearest whole number)

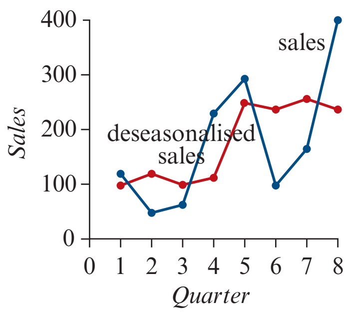
5 a *number of international visitors* (000's) = −276 641 + 140.627 × *year*

b 8 831 800 to the nearest 100. It may not be reliable because we are extrapolating well outside the range of the data and the trend may no longer be linear.

6 a

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	122	128	118	130
2	250	245	263	236

b



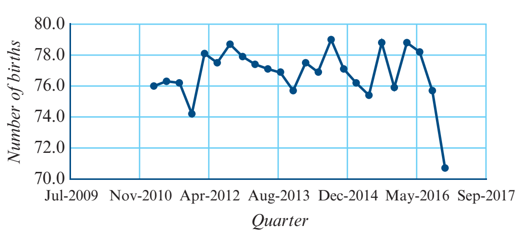
The deseasonalised sales appear to show an increasing trend over time.

c deseasonalised sales = 80.8 + 23.5 × quarter

d forecasted actual sales = 386 × 1.13 = 436

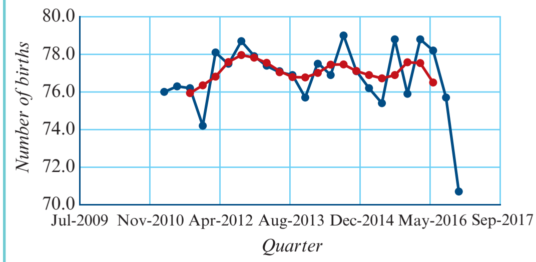
Exercise 3F

1 a



The number of births shows mainly random variation from Q1 2011 to Q3 2016. There seems to have been a drop in Q4 2016.

b



Graph still shows some variation but no clear trend.

c Using the years 2013–2015

Q1	Q2	Q3	Q4
1.000	0.993	1.016	0.990

Four-point moving average has been chosen to remove any quarterly seasonality.

d There is not a strong seasonal pattern in the number of births, although we do see a slightly higher birth number than average in Q3, and slightly lower in Q1 and Q4.

e If quarters are numbered 1–24, then the *number of births* (000's) = 77.175 − 0.0334 × *quarter*. Predicted number for Q3 2020 is 75 900 to the nearest 100. The intercept predicts the number of births in quarter 0, which would be Q4 2010. The slope indicates that the number of births is decreasing by about 33 each quarter.

2 a i

The plot shows both a clear increasing trend as well as quite a lot of month-to-month variation. There is also some evidence of seasonality, with the number of employed appearing to drop each January.

ii Using 2014 and 2015:

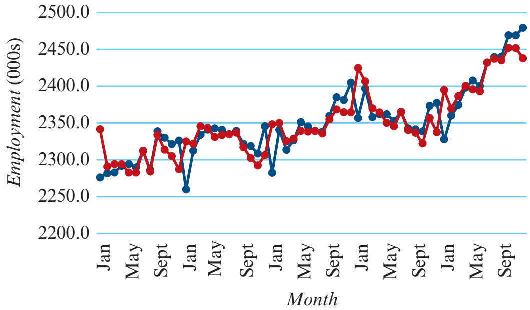
Jan	Feb	March	April	May	June
0.972	0.996	0.995	0.999	1.005	1.003
July	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**
1.000	1.001	1.002	1.007	1.004	1.017

ISBN 978-1-108-46154-2

Cambridge University Press

iii As suggested by the plot, the seasonality indexes confirm that employment is lowest in January, and then increases steadily to average month levels by May. It then remains fairly steady, peaking in December before dropping again in January.

iv



Overlaying the deseasonalised data in the plot we can see that some of the variation in the data has been removed.

v With month numbered 1–60:
$employed(000\text{'s}) = 2285.53 + 2.0999 \times month$
Intercept (2 285 530) gives the estimated number of employed in Month 0, which would be December 2012. Slope indicates that on average employment is growing by 2100 per month.

b i

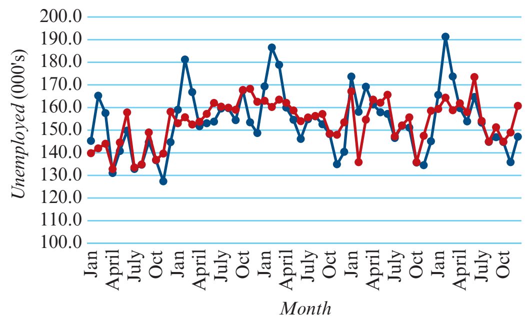
The plot shows month-to-month variation, with no clear trend. There is evidence of seasonality, with unemployment to peak in February and be at its lowest in November each year.

ii Using 2014 and 2015

Jan	Feb	March	April	May	June
1.039	1.164	1.094	0.987	0.974	0.949
July	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**
0.995	0.999	0.971	0.999	0.912	0.915

iii As suggested by the plot, unemployment is highest in the period January to March, lower but reasonably steady from April to October, and then reducing in November and December.

iv



Again, deseasonalisation of the data has removed some of the variation.

v With month numbered 1–60:
$unemployed(000\text{'s}) = 148.68 + 0.1693 \times month$
Intercept (148 680) gives the estimated number of employed in Month 0, which would be December 2012. Slope indicates that on average unemployment is growing by about 170 people per month.

c A scatterplot of unemployed against employed shows that, contrary to what we might expect, there is no relationship between employed and unemployed. This is confirmed by the value of the correlation coefficient, $r = -0.126$.

Chapter 3 review

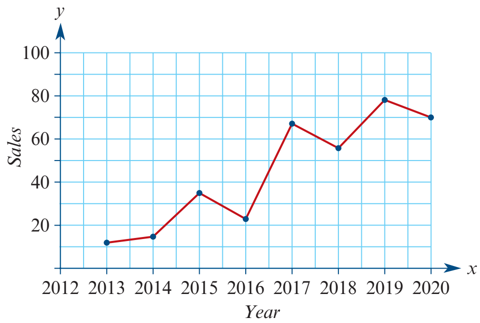
Multiple-choice questions

1 C	**2** A	**3** B	**4** E
5 C	**6** C	**7** B	**8** A
9 E	**10** E	**11** D	**12** C
13 A	**14** A	**15** D	**16** A

 ISBN 978-1-108-46154-2 Cambridge University Press

Short-answer questions

1



2 Trend is when there is an overall increase or decrease in the time series.
Seasonality is when there is a periodic movement in the time series that is calendar related, such as weekly or monthly.
A cycle is also a periodic movement in the time-series, but this is generally used for time period of greater than one year.

3 The time series plot shows an increasing trend over the 12 years, with cycle of 4 years within this.

4 Smoothing is used to eliminate some of the random variation in a time series plot, so that the features such as trend can be more easily identified.

5 a 22.3 **b** 32.8

6 0.7071

7 0.7053

8 2214.4

9 a 1.86

b On Saturday the ice cream sales are 86% higher than they are on an average day.

10

Q1	Q2	Q3	Q4
0.64	1.78	1.18	0.42

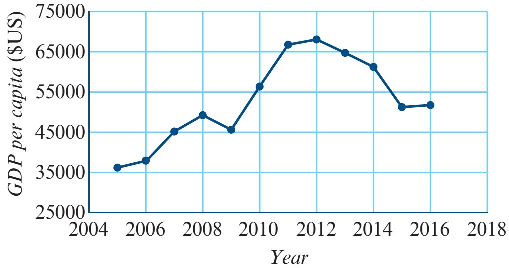
11 a 000's of visitors $= -571\,149.46 + 286.75 \times$ year

b 9519289 – may be unreliable as we are extrapolating.

12 205 to the nearest whole person.

Extended-response questions

1 a

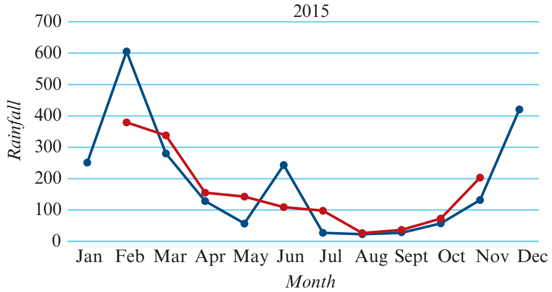


b The GDP per capita increased steadily between 2005 and 2012, with the exception of 2009 where there was a significant drop. Since 2012, Australia has seen a steady decrease in GDP per capita, although improved results in 2016 may indicate that the GDP is now steady.

c $GDP = -3\,874\,825 + 1953.58 \times year$

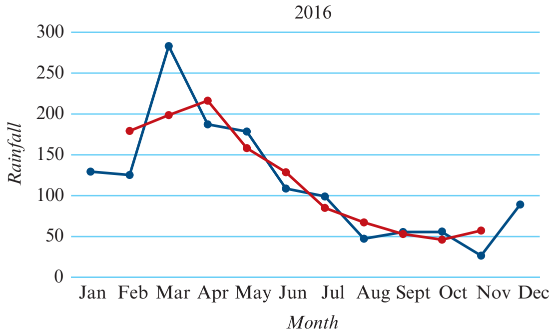
d Predicted 2017 = \$65 546, error = \$9411

2 a & b



In 2015 the highest rainfall in Cairns is in the months from February to April, and then a steady decease in rainfall from April to November.

c & d



In 2016 the highest rainfall in Cairns is in the months from February to April, and then a steady decease in rainfall from April to November.

e Both charts show the highest rainfall in Cairns in the months February to April, and then a steady decease in rainfall from April to October, when rainfall then starts to increase again.

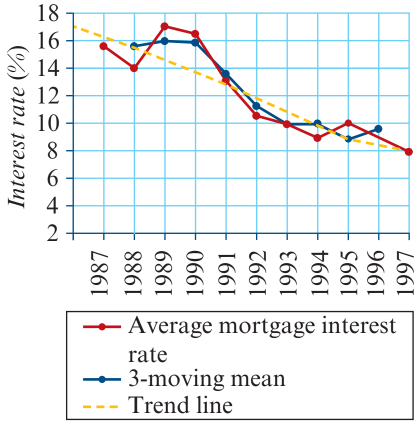
3 a 3-mean smoothed value for 1992

$$= \frac{13.0 + 10.5 + 9.5}{3} = \frac{33}{3} = 11\%$$

ISBN 978-1-108-46154-2

Cambridge University Press

b



c Both the raw data and the smoothed data reveal a steadily decreasing trend.

Chapter 4

Exercise 4A

1 a 2, 8, 14, 20, 26 **b** 5, 2, –1, –4, –7

c 10, 5, 2.5, 1.25, 0.625 **d** 6, 12, 24, 48, 96

2 a 4, 6, 8, 10, 12 **b** 50, 10, 2, 0.4, 0.08

c 24, 20, 16, 12, 8 **d** 5, 15, 45, 135, 405

e 2, 10, 50, 250, 1250

Exercise 4B

1 3, 7, 11, 15, 19 **2** 15, 11, 7, 3, –1

3 a 4 **b** 40

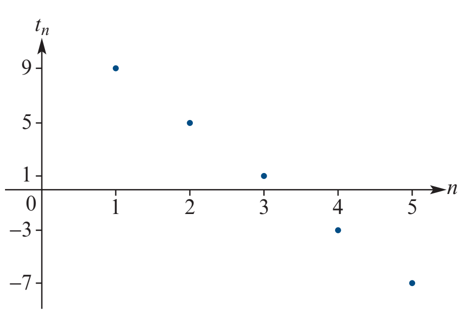
c 7 times, $t_8 = 48$ **d** 12 times, $t_{13} = 68$

4 a –2 **b** –5

c 6, $t_7 = -7$ **d** –13, –93

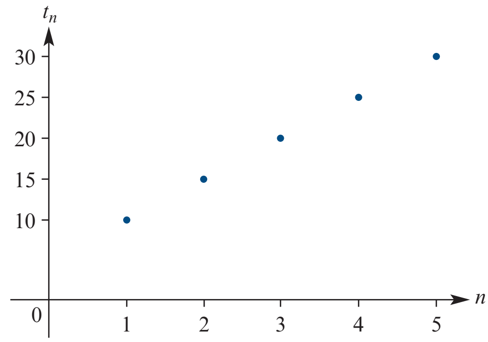
5

n	1	2	3	4	5
t_n	9	5	1	–3	–7



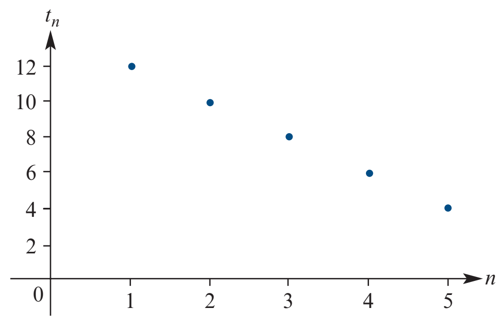
6

n	1	2	3	4	5
t_n	10	15	20	25	30



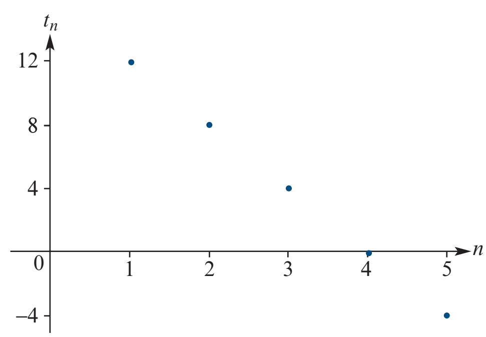
7

n	1	2	3	4	5
t_n	12	10	8	6	4



8

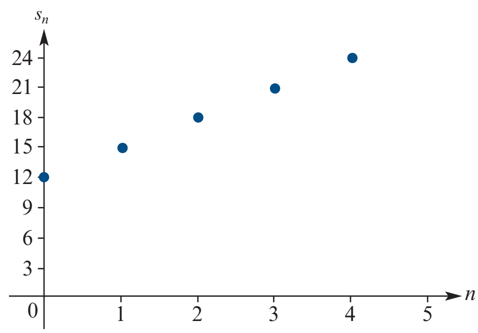
n	1	2	3	4	5
t_n	12	8	4	0	–4



ISBN 978-1-108-46154-2

9

n	0	1	2	3	4
s_n	12	15	18	21	24



Exercise 4C

1 a 220 **b** 290 **c** 1700

2 $t_n = 7n - 1$

3 $t_n = 59 - 11n$

4 $t_n = 11n - 3$

5 $t_n = 1020 - 20n$

6 a 55 **b** –16
c 0.36 **d** 101
e –5.5 **f** –528
g –1

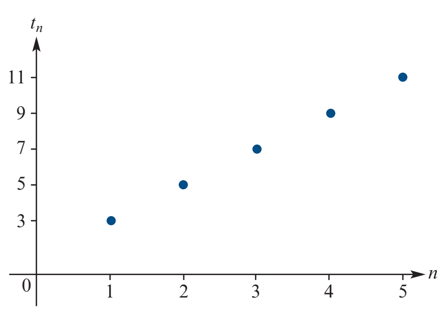
7 4

8 8

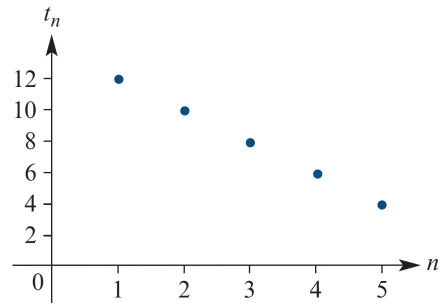
9 a 7, 12, 17 **b** 1, 4, 7
c 20, 16, 12 **d** 99, 104, 109
e 20, 16, 12 **f** 120, 110, 100
g –25, –20, –15

10 a 12 **b** 19
c 9 **d** 8
e 7 **f** 10
g 27

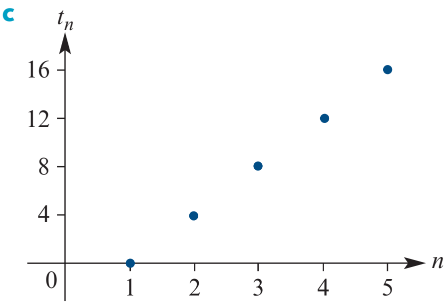
11 a



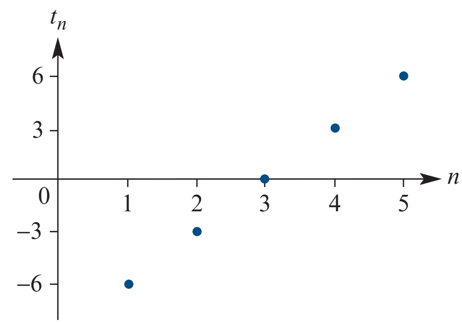
b



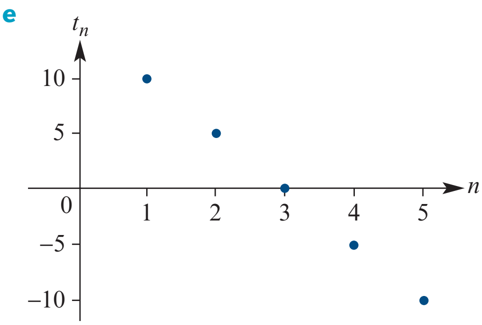
c



d



e



12 A: $a = 25, d = -4, t_n = 29 - 4n$
B: $a = 5, d = 5, t_n = 5n$
C: $a = 18, d = -1, t_n = 19 - n$
D: $a = 1, d = 6, t_n = 6n - 5$

13 a $t_n = 2n$ **b** 2, 4, 6, 8

14 a $t_n = 6n + 5$ **b** 11, 17, 23, 29

15 a $t_n = 59 - 3n$ **b** 56, 53, 50, 47

ISBN 978-1-108-46154-2

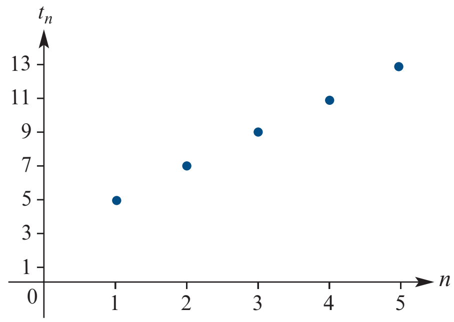
Cambridge University Press

16 a $t_n = 2n + 3$

b

n	1	2	3	4	5
t_n	5	7	9	11	13

c

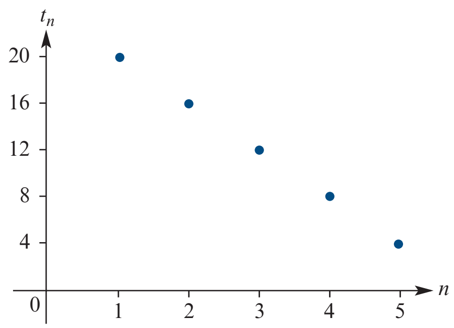


17 a $t_n = 24 - 4n$

b

n	1	2	3	4	5
t_n	20	16	12	8	4

c

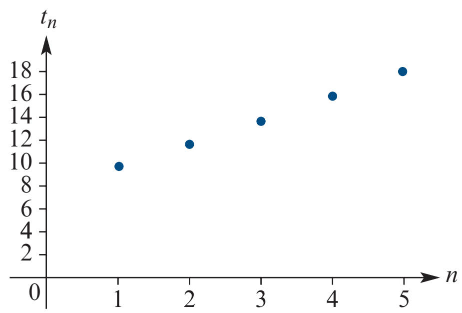


18 a $t_n = 2n + 8$

b

n	1	2	3	4	5
t_n	10	12	14	16	18

c



19 $t_1 = 7, d = 5$

20 $t_1 = 2, d = -2$

21 $t_1 = 18, d = 7; t_1 = 18, t_{n+1} = t_n + 7$

22 $t_1 = 3, d = -5; t_1 = 3, t_{n+1} = t_n - 5$

Exercise 4D

1 a \$375

b

Year	1	2	3	4	5
Value (\$)	5375	5750	6125	6500	6875

c $A = 5000 + 375n$ dollars

d i \$10625 **ii** \$14375

2 a \$4500

b

End of year	1	2	3	4	5
Amount (\$)	54500	59000	63500	68000	72500

c $A = 50000 + 4500n$

d i \$117500 **ii** \$162500

3 a $A = 60000 + 2700n$

b \$73500

c 8 years

4 a $A = 2000 + 76n$

b \$2456

c 14 years

5 a $A = 7000 + 518n$

b \$10108

c 6 years

6 a $A = 100000 - 4000n$ **b** \$80000

7 a $V = 48000 - \frac{n}{5}$

b \$45000

c 25000 km

8 a $C_n = 2.25 + 0.02n$

b i \$2.45 **ii** \$2.95

c i 80 pages **ii** 200 pages

9 a $g_n = 86 + 40n$

b i 686 kg **ii** 1166 kg

c 35

10 a $a_n = 15 - 0.2n$

b i 10.4 litres **ii** 6 litres

c 67

11 a $s_n = 20500 + 450(n - 1)$

b i \$22300 **ii** \$23650

c Start of 67th year

12 a $m_n = 430 - 25.75(n - 1)$

b i \$275.50 **ii** \$121

c 17 days

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 4E

1 3, 12, 48, 192, 768

2 15, 30, 60, 120, 240

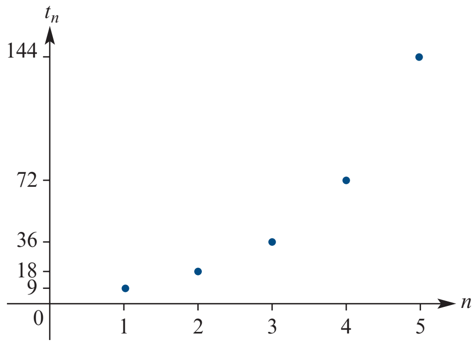
3 **a** 10 **b** 20000 **c** 4 **d** $14, 2 \times 10^{14}$

4 **a** $\frac{1}{4}$ **b** 4 **c** $6, \frac{1}{4}$ **d** $0.00390625, 3.23 \times 10^{-27}$

5 **a** –5 **b** –625 **c** 5, 3125 **d** 14, –6103515625

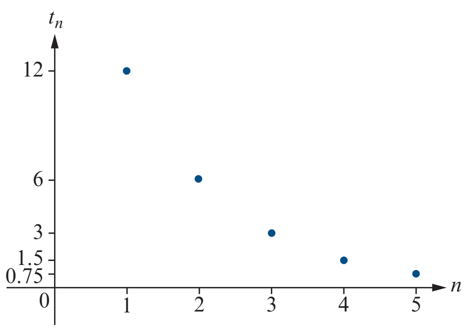
6

n	1	2	3	4	5
t_n	9	18	36	72	144



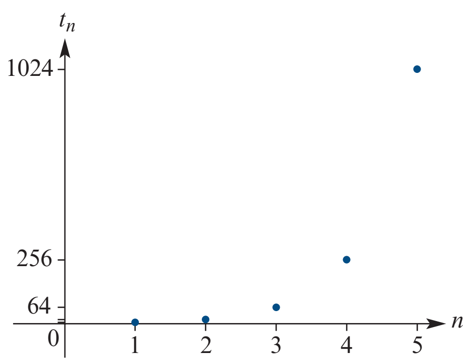
7

n	1	2	3	4	5
t_n	12	6	3	1.5	0.75



8

n	1	2	3	4	5
t_n	4	16	64	256	1024



Exercise 4F

1 **a** 16 **b** 4096 **c** 268435456

2 $t_n = 8 \times 5^{n-1}$

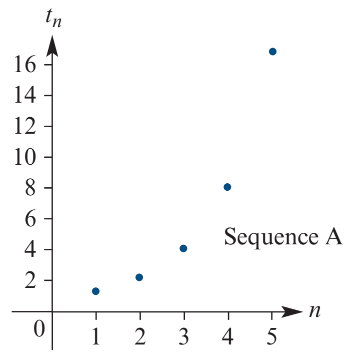
3 $t_n = 5 \times 0.6^{n-1}$

4 **a** 15625 **b** 0.128 **c** 512 **d** –131220 **e** 0.015625 **f** 161.051 **g** 0.015625

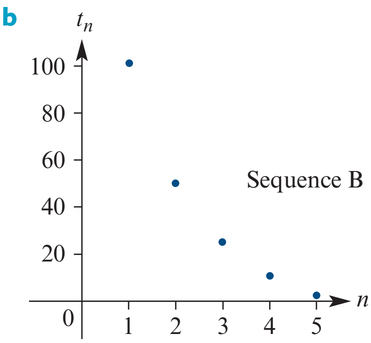
5 **a** 1, 3, 9
b 20000, 10000, 5000
c 10, –20, 40
d 128, 160, 200

6 **a** 8 **b** 9 **c** 12 **d** 10 **e** 22 **f** 16 **g** 15

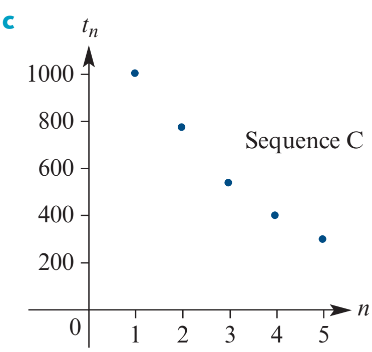
7 **a**



b



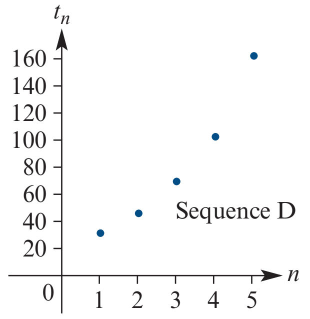
c



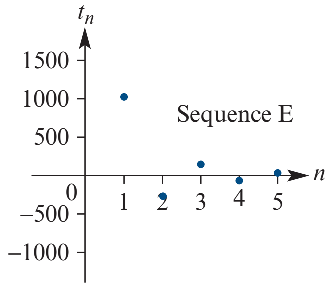
ISBN 978-1-108-46154-2

Cambridge University Press

d



e



8 Sequence A

a 25 **b** $0 < r < 1$ **c** $r \approx 0.6$

Sequence B

a 10 **b** $r > 1$ **c** $r \approx 1.2$

Sequence C

a 10 **b** $0 < r < 1$ **c** $r \approx 0.5$

9 a 2, 4, 8, 16 **b** 2

Exercise 4G

1 a $A = P \times 0.96^{12m}$

b $A = 30000 \times 0.965^n$

2 a $A = 2700 \times 1.05^n$

b 3446 **c** 15 years

d 971 **e** 18 years

3 $A = 6000 \times (1.042)^n$

4 a $A = 20000 \times (1.063)^n$

b 7 years

c $A = 18000 \times (1.094)^n$

5 a $A = 9800 \times (0.965)^n$

b \$8201 **c** \$319

6 a $A = 8000 \times (1.125)^n$

b \$11 390.63 **c** \$3390.63 **d** \$1265.63

7 a $A = 3300 \times (1.075)^n$

b \$6801.40 **c** \$3501.40 **d** \$474.52

8 a $A = 1200 \times (0.88)^n$

b \$490.41

9 a Yes, the ratio of height of bounce to previous bounce is constant $= 0.75$ (small discrepancies in value of r can be attributed to accuracy of original measurements).

b i 113.91 cm

ii $h_n = 360(0.75)^n$

iii 4.81 cm

10 a 20

b

1	2	3
10	20	40

c $b_n = 10 \times 2^{n-1}$

d i 160 **ii** 640 **iii** 40 960

11 a 1400

b

1	2	3	4
1000	1400	1960	2744

c $F_n = 1000(1.4)^{n-1}$

d i 7530 **ii** 28 925 **iii** 836 683

12 a \$18 500

b

1	2	3	4
20 000	18 500	17 113	15 829

c $V_n = 20000(0.925)^{n-1}$

d i \$12 528 (to the nearest dollar)

ii \$10 719 (to the nearest dollar)

iii \$6211 (to the nearest dollar)

13 a \$27 000

b

1	2	3
12 000	27 000	60 750

c $V_n = 12000(2.25)^{n-1}$

d i \$307 547 (to the nearest dollar)

ii \$1 556 956 (to the nearest dollar)

iii \$39 903 081 (to the nearest dollar)

Chapter 4 review

Multiple-choice questions

1 C **2** A **3** B **4** B **5** C
6 A **7** D **8** D **9** C **10** C
11 D **12** B **13** D **14** A **15** C
16 C **17** B **18** B **19** A **20** D

Short-answer questions

1 46 **2** 56 **3** 320
4 7.8125 **5** 72 **6** $t_n = 17 - 4n$
7 1, 6, 11 **8** 10, 20, 40 **9** 10
10 9 **11** 6, 12, 18 **12** 1000, 500, 250
13 $t_n = 13 - 5n$ **14** $t_n = 3^{n-1}$ **15** 4374
16 a $t_n = 2 \times 6^{n-1}$ **b** 2592
c Approx. $3.385331889 \times 10^{13}$

ISBN 978-1-108-46154-2

Cambridge University Press

17 **a** $t_n = 120 \times 0.25^{n-1}$
b 7.5
c Approximately 0.000002 ($1.788139343261719 \times 10^{-6}$)

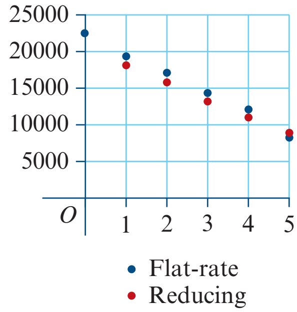
18 **a** $A = 38500 \times 0.905^n$
b \$23 372.42 **c** \$15 127.58

19 **a** $A = 20000 + 1880n$
b \$29 400 **c** 11 years

20 **a** 20 cents **b** $A = 1650 - 0.2n$
c \$1250

Extended-response questions

1 **a i** $A = 22500 - 2700n$ **ii** \$9000
b i $B = 22500 \times 0.84^n$ **ii** \$9409.77
c



2 **a** \$6575 **b** \$6777.89 **c** 7.1%

3 **a**

Day	1	2	3	4	5	6	7	8	9	10
Emission each day	1500	1370	1240	1110	980	850	720	590	460	330

b 162.5 kg
c i 0.75
ii

Day	1	2	3	4	5	6	7	8	9	10
Emission each day	1500.00	1125.00	843.75	632.81	474.61	355.96	266.97	200.23	200.00	200.00

d Day 8
e 3350.67 kg
f 20.1%

Chapter 5

Exercise 5A

1 **a** 32°27′ **b** 43°12′ **c** 122°27.6′
d 91°7.2′ (minutes to one decimal place) **e** 45′

2 **a** 32.75° **b** 15.5833° **c** 7.36667° **d** 142.733° **e** 67.25°

3 **a** 50.27 cm **b** 87.96 m **c** 282.74 mm **d** 37.70 mm **e** 43.98 m

4 **a** $\frac{1}{4}$ **b** $\frac{3}{4}$ **c** $\frac{1}{12}$ **d** $\frac{1}{3}$ **e** $\frac{1}{6}$ **f** $\frac{5}{12}$

5 **a** 7.85 cm **b** 10.47 cm **c** 26.18 cm **d** 23.56 cm **e** 36.65 cm **f** 57.60 cm

6 **a** 13.09 cm **b** 5.24 cm **c** 78.54 cm **d** 37.70 cm **e** 122.17 cm **f** 109.96 cm

ISBN 978-1-108-46154-2

7 45.81 cm
8 **a** 95.5° **b** 47.75°
9 6.20 cm

Exercise 5B

1 **a** Bucharest, Romania
b Oslo, Norway
c Houston, USA
d Singapore
e Cape Town, South Africa
2 **a** 30.04°N, 31.24°E
b 19.43°N, 99.13°W
c 34.60°S, 58.38°W
d 19.08°N, 72.88°E
e 6.5°N, 3.37°E
f 17.83°S, 31.03°E
3 **a** 28.04°S, 148.59°E
b 26.80°S, 153.13°E
c 21.14°S, 149.18°E
d 12.65°S, 141.85°E
e 24.87°S, 152.35°E
f 20.08°S, 146.26°E
g 138°E
4 **a** (65°N, 0°) **b** (0°, 75°E)
5 **a** (18°N, 24°E) **b** (38°N, 44°E)
c (17°S, 23°E) **d** (43°N, 37°W)
6 **a** **i** Mexico City, Yangon
ii Brisbane
b **i** Brisbane **ii** Plymouth
iii London **iv** Zurich
c Marseilles, Mexico City, Plymouth, Zurich, London, Yangon
d Plymouth, Mexico City, Lima, London (just)
e Mexico City, Yangon
f Cooktown, Melbourne
g London
h Wellington
7 6338 km
8 2313 km
9 1779 km
10 34.27°S
11 **a** 3892 km **b** 3002 km **c** 4893 km
d 4448 km **e** 3336 km
12 3558 km
13 4003 km
14 **a** 5560 km **b** 7784 km **c** 3336 km
d 6672 km **e** 16 902 km
15 5.4°
16 **a** 3336 km **b** 6672 km **c** 13 344 km
17 **a** 4226 km **b** 5782 km **c** 14 234 km
18 **a** 4176 km **b** 14 184 km **c** 5832 km
19 **a** Equator 4670 km; North Pole 5338 km; South Pole 14 678 km
b Equator 6116 km; North Pole 3892 km; South Pole 16 124 km
c Equator 1668 km; North Pole 11 676 km; South Pole 8340 km
d Equator 1557 km; North Pole 11 565 km; South Pole 8451 km
20 **a** 6154 km **b** 5517 km
c 4505 km **d** 3186 km
21 60°N, 60°S
22 **a** 1134 km **b** 2829 km **c** 5439 km
23 10 045 km
24 **a** 1371 km **b** 6783 km
c 4689 km **d** 2391 km
25 **a** 9129 km **b** 12 485 km
c 13 232 km **d** 9898 km
26 **a** 445 km **b** 645 km **c** 1090 km
27 **a** 2335 km **b** 2340 km **c** 4675 km
28 Answers will vary

Exercise 5C

1 **a** X is 2 hours ahead of Y
b X is 5 hours ahead of Y
c Y is 5 hours ahead of X
d Y is 1 hour 48 minutes ahead of X
2 **a** **i** 1 hour 28 minutes behind
ii 9 hours 52 minutes ahead
iii 3 hours 48 minutes ahead
iv 10 hours 12 minutes ahead
v 16 hours 48 minutes ahead
b **i** 7:28 a.m.
ii 8:08 p.m. the day before
iii 2:12 a.m.
iv 7:48 p.m. the day before
v 1:12 p.m. the day before
3 **a** 150°W **b** 2 p.m.
c 7:30 p.m. the day before
4 **a** 5:30 p.m. **b** 7:30 a.m.
5 **a** 45° **b** 3 hours
c 2:45 p.m.
6 **a** 3:26 a.m. Sunday
b 5:34 p.m. Saturday
c 6:30 p.m. Saturday
d 3:22 a.m. Sunday
e 10:26 a.m. Saturday
f 7:22 a.m. Saturday

ISBN 978-1-108-46154-2

Cambridge University Press

7 a 3 a.m. the same day b 3 p.m. the next day
8 a 3 a.m. the same day b 5 a.m. the next day
9 a 4 p.m. Monday b 14 hours
c 13 hours 55 minutes
10 a 12 midday same day
b 11:30 a.m. same day
c 10 a.m. same day
d 12 midday same day
e 2 p.m. same day f 1 p.m. same day
g 2 p.m. same day
h 4 p.m. the day before
i 1 p.m. same day

Chapter 5 review

Multiple-choice questions

1 C 2 B 3 A 4 E
5 E 6 B 7 C

Short-answer questions

1 3336 km
2 3892 km
3 a 52°
b 3 hours 28 minutes
c 6:13 a.m.
4 (0°, 160°W)
5 19 905 km
6 3 hours 8 minutes
7 8 a.m.
8 5 a.m.
9 7 p.m.
10 6 a.m. Saturday
11 a 4 p.m. Tuesday b 2 p.m. Wednesday
12 a 15 hours 28 minutes
b 11:28 a.m. Tuesday c 8:32 a.m. Thursday

Extended-response questions

1 a 5543 km b 1451 km
c 6400 km d 1 km
2 a 5226.4 km b 5842.58 km
c 16 622.26 km

Chapter 6

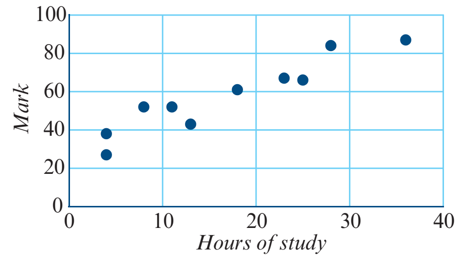
6A Topic 1 Bivariate data analysis

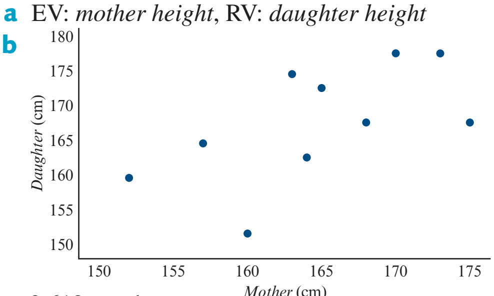
Multiple-choice questions

1 C 2 E 3 B 4 E
5 E 6 D 7 C 8 A
9 D 10 B 11 A 12 B
13 E 14 B 15 D 16 A

Short-answer questions

Simple familiar questions

1 a Numerical & numerical, scatterplot
b Categorical & numerical, parallel boxplots
c Categorical & categorical, two-way frequency table
2 a EV: *time of observation*, RV: *activity*
b 38%
c Yes. Only 6.7% of dolphins were observed feeding in the afternoon, less than the percentage observed feeding in the morning 38%, and very much less than the percentage observed feeding in the evening 70.9%.
3 a EV: *hours*, RV: *mark*
b


c There is a strong, positive linear relationship between hours of study and mark. Those who studied more hours obtained higher marks in the examination.
4 a 0.966. There is a strong, positive linear relationship between male and female life expectancies. Those countries that have high male life expectancies also tend to have high female life expectancies.
b The variables are both numerical and the relationship is linear.
5 a EV: *mother height*, RV: *daughter height*
b


c 0.612, moderate
d *daughter height* = 50.23 + 0.715 × *mother height*
Intercept: mother's height of 0 cm predicts a daughter's height of 50.23 cm, which is meaningless. Slope: An increase of 1 cm in mother's height predicts an increase of 0.715 cm in daughter's height.

ISBN 978-1-108-46154-2

Cambridge University Press

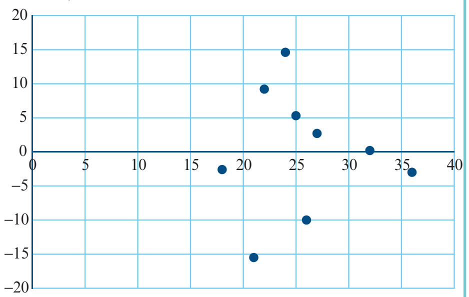
e $R^2 = 37.5\%$. 37.5% of the variation in daughter's height can be explained by the variation in mother's height.

f 171.8 cm

g Interpolation, as 170 cm is within the range of the data used to determine the equation of the regression line.

6 a −15.5, 2.7

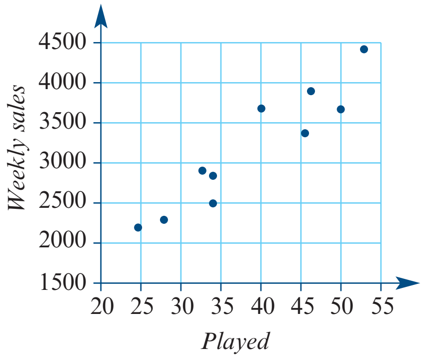
b



The residual plot shows only random variation, so the linearity assumption holds.

7 a EV: *played*, RV: *weekly sales*

b



c $r = 0.9485$

d Strong, positive, linear relationship

e $\textit{weekly sales} = 293 + 74.3 \times \textit{played}$

f Slope: on average, the number of downloads increases by 74.3 for each additional time the song is played on the radio in the previous week.

Intercept: predicts 293 downloads of the song if it is not played on radio in the previous week.

g 7723

h Extrapolating

8 a EV = *height*, RV = *weight*

b Yes, scatterplot shows a linear relationship.

c $\textit{weight} = -79.78 + 0.878 \times \textit{height}$

d Slope = 0.878. On average, each additional centimetre in height adds an additional 0.878 kg to a player's weight.

e A player will never be 0 cm tall.

f 91.4 kg

g $R^2 = 78.1\%$. 78.1% of the variation in player's weight is explained by the variation in their height.

Complex familiar questions

9 a $\textit{cost} = 165.63 + 5.97 \times \textit{number of meals}$

b Base cost = $165.63

c Additional cost per meal = $5.97

10 A study was conducted to investigate the relationship between femur length and radius length. Data were collected from a sample of 10 people.

From the scatterplot of radius versus femur, we can see that there is a strong, positive, linear relationship between femur length and radius length: $r = 0.988$. That is, those people with a longer femur also tended to have a longer radius. There are no obvious outliers, and the linearity assumption is confirmed by the residual plot. The equation of the least squares regression line is:

$$\textit{radius} = -7.25 + 0.739 \times \textit{femur}$$

The slope of the regression line predicts an increase of 0.739 cm in radius for each 1 cm increase in femur.

The coefficient of determination tells us that 97.5% of the variation in radius is explained by the variation in femur.

11 a Birth rate 76.7%, births attended by skilled staff 51.0%, exclusive breastfeeding 2.9%, health expenditure 24.5%, literacy rate female 64.0%, literacy rate male 72.1%, safe sanitation 40.4%, safe drinking water 66.7%.

b Of the variables listed, only exclusive breastfeeding does not appear to be related to infant mortality. The variation in infant mortality is strongly related to birth rate, and parent's educational levels (as indicated by literacy rate, particularly their fathers). Also important are access to clean water, skilled health staff, and safe sanitation (in that order). Expenditure on health is also indicated to be able explain a reasonable percentage of the variation in mortality rate.

12 There is an association between preference for type of instruction and faculty. A higher level of preference for online study is seen in Business students (58.6%), reducing to similar percentages for Arts and Science students (36.3% and 36.0% respectively).

ISBN 978-1-108-46154-2

Cambridge University Press

Extended-response questions

Complex unfamiliar questions

1 a The value of the correlation coefficient would be closer to -1.

b The value of the slope would decrease (become close to -2).

2 a \$22.90 per pair of jeans

b \$147.90

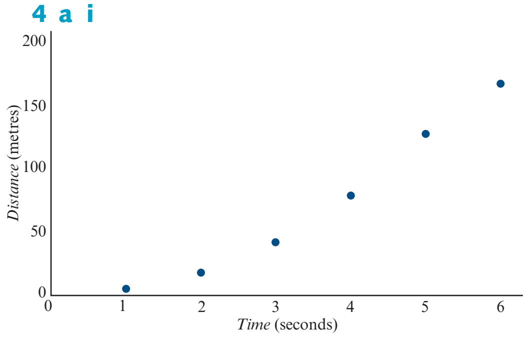
c i \$258.83 **ii** \$83.83

3 a There is a relationship between political affiliation and attitude to a republic for males, with 28.2% of those who identify as Liberal in favour of retaining the Queen, compared to only 8.1% of those who identify as Labor wanting to retain the Queen.

b There is a relationship between political affiliation and attitude to a republic for females, with 36.7% of those who identify as Liberal in favour of retaining the Queen, compared to only 16.3% of those who identify as Labor wanting to retain the Queen.

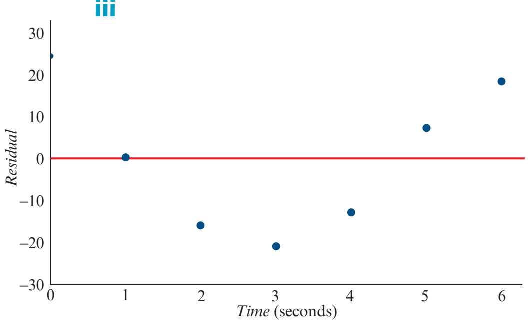
c For both males and females, the relationship between support for a republic and political affiliation is the same, with both groups showing a stronger preference for retaining the Queen in Liberal supporters than in Labor supporters.

4 a i



ii $distance = -23.96 + 28.950 \times time$

iii

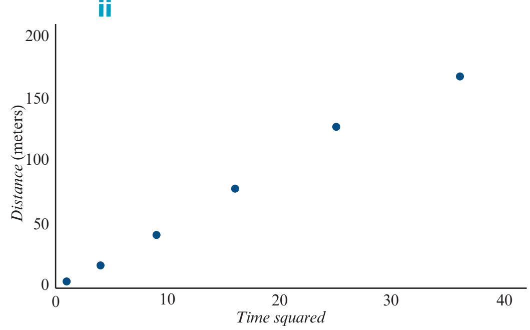


Residual plot shows that the relationship is not linear as points are not evenly distributed above and below the zero line and indicate another underlying trend.

b i

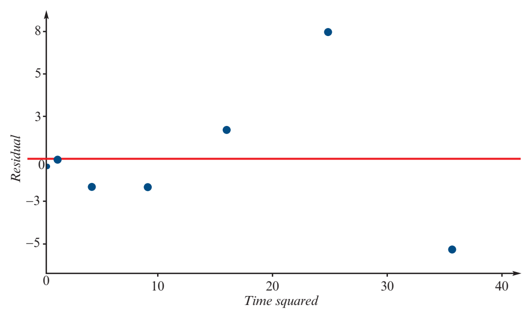
Time (seconds)	$(\text{Time})^2$	Distance (metres)
0	0	0
1	1	5.2
2	4	18
3	9	42
4	16	79
5	25	128
6	36	168

ii



iii $distance = 0.45 + 4.803 \times time^2$

iv



Linearity assumption is better met.

Problem-solving and modelling

5 Answers will vary.

6 Answers will vary.

6B Topic 2 Time series analysis

Multiple-choice questions

1 A	**2** A	**3** A	**4** B	**5** E
6 B	**7** E	**8** B	**9** D	**10** E
11 E	**12** B	**13** E	**14** D	**15** A

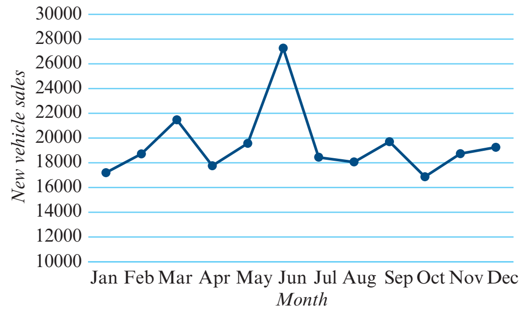
ISBN 978-1-108-46154-2

Cambridge University Press

Extended-response questions

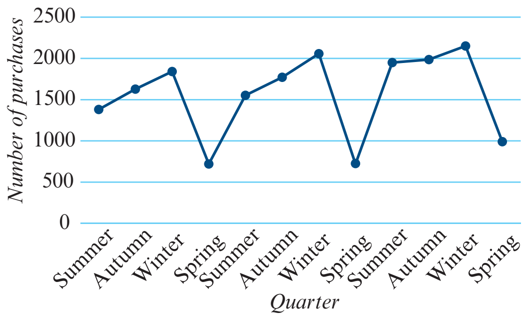
Simple familiar questions

1



New vehicle sales in Queensland in 2016 varied between about 16 500 and 19 500 in most months, with a small increase in sales in March, and a much larger increase in sales in June (27 270).

2 a

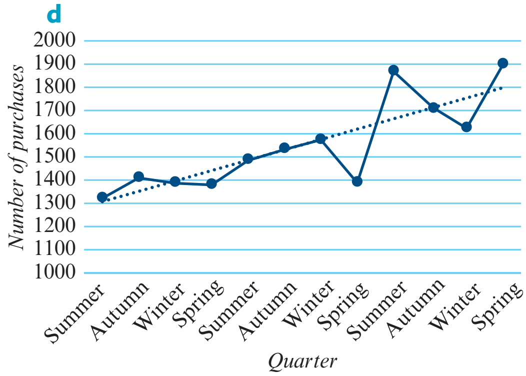


b i Summer = 1.04 **ii** Spring = 0.52

c

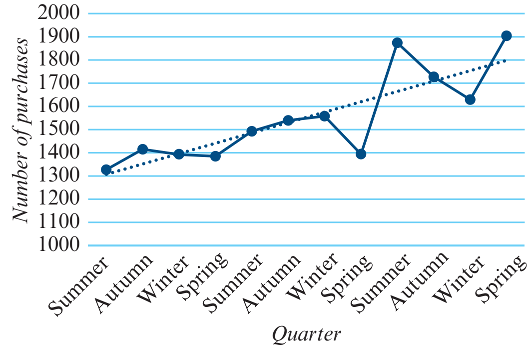
	Summer	Autumn	Winter	Spring
2016	1327	1415	1415	1385
2017	1492	1539	1582	1394
2018	1874	1727	1654	1904

d



e *number of purchases* = 1266.4 + 45.01 × *quarter*

f

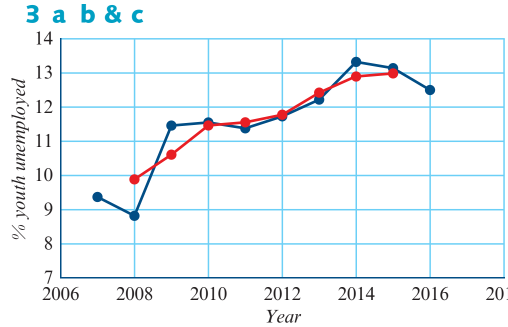


g Gradient = 45.01. On average, the number of purchases increases by about 45 each quarter.

h Seasonalised predicted purchases for summer 2020 = 2212.9

i R^2 = 70.1%. 70.1% of the variation in the purchases is explained by the linear relationship between the number of purchases and quarter.

3 a b & c



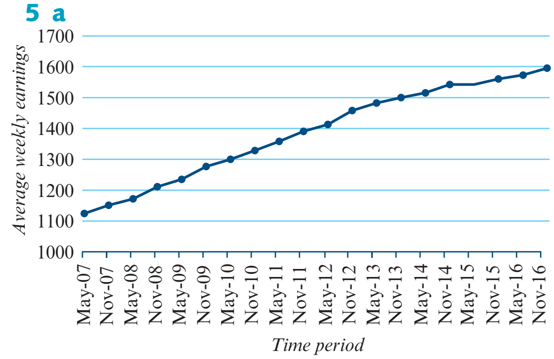
d From the smoothed graph we can see that the % of youth unemployment has increased steadily over the years 2007–2016.

4 a \$4600 **b** \$5080

c i 0.860 **ii** \$3500

Complex familiar question

5 a



ISBN 978-1-108-46154-2

Cambridge University Press

b

	Centred 2-point moving average
May-07	
Nov-07	1149.4
May-08	1176.2
Nov-08	1207.0
May-09	1239.3
Nov-09	1272.2
May-10	1301.5
Nov-10	1328.8
May-11	1358.7
Nov-11	1388.0
May-12	1418.7
Nov-12	1452.9
May-13	1480.8
Nov-13	1499.6
May-14	1518.5
Nov-14	1535.5
May-15	1546.5
Nov-15	1559.0
May-16	1575.7
Nov-16	

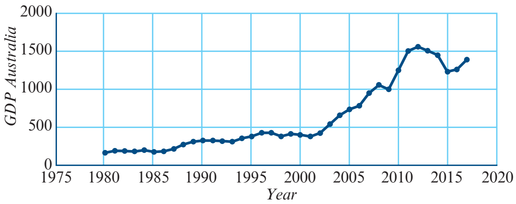
c *average weekly earnings* = 1115.69 + + 25.786 × *time period*

d Predicted = \$1657.20, which overestimates the actual value by \$51.60.

e From the graph we can see that average weekly earnings appeared to increase at a steady rate from May 07 to Nov 13. Fitting a regression line to this data showed an average increase each period of \$29.84. From May 14 to Sept 16 the rate of increase slowed. Fitting a regression line to this data confirmed this change, showing an average increase of \$14.58 per period.

Complex unfamiliar question

6 a



From the plot we can see that GDP rose slowly but steadily over the years 1980 to 2002. From here, GDP started to rise rapidly until 2013, when it began to decrease. In 2017, GDP started to rise again.

b **i** *GDP* = −23 848.3 + 12.121 × *year*

ii

Year	Percentage error
1990	18.9%
2000	1.4%
2010	142.7%
2015	113.7%
2017	131.8%

iii The equation provides a reasonable prediction over the years where the relationship has the same linear pattern as that of the data used to determine the equation.

c **i** *GDP* = −164 131 + 82.226 × *year*

ii

Year	Percentage error
2010	9.3%
2015	−20.9%
2017	−19.1%

iii Again, the equation provides a reasonable prediction over the years where the relationship has the same linear pattern as that of the data used to determine the equation.

d **i** *GDP* = −50 493.2 + 25.527 × *year*

ii

Year	Percentage error
2010	53.1%
2015	30.3%
2017	39.7%

ISBN 978-1-108-46154-2

Cambridge University Press

iii This equation is not a good predictor because the relationship is not linear over the range of the data used.

e Comparison of each of the models shows that the first two models are not unreasonable if used only for short term predictions but becomes very unreliable when used to extrapolate too far from the data used for determining the equation. The third model is not reliable at all because the relationship does not follow the same linear pattern over the range of the data used.

Problem-solving and modelling

7 Answers will vary.
8 Answers will vary.

6C Topic 3 Growth and decay in sequences

Multiple-choice questions

1 C **2** D **3** E **4** B **5** D

Short-answer questions

Simple familiar questions

1 51 **2** 310 **3** 4096
4 162 **5** 98911
6 $t_n = 525 - 25n$
7 $t_1 = 5, t_2 = 13, t_3 = 21$
8 $t_1 = 2, t_2 = 10, t_3 = 50$
9 $n = 29$
10 $t_1 = 8, t_2 = 19.25, t_3 = 30.5$
11 $t_n = 815 - 15n$
12 $t_n = 2 \times 6^{n-1}$
13 \$249445

Complex familiar questions

14 –15
15 20
16 **a** $V_n = 48000 \times 0.92^n$
b \$29105
c \$16364
17 **a** $V_n = 25550(1 + 0.16n)$
b \$37740
c 7 years
18 **a** $C_n = 5000 + 0.02n$
b \$7400
c 250000

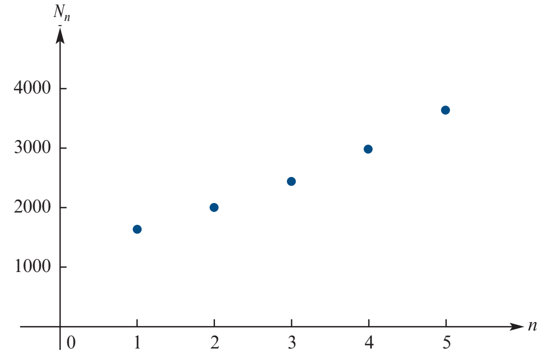
Extended-response questions

Complex unfamiliar questions

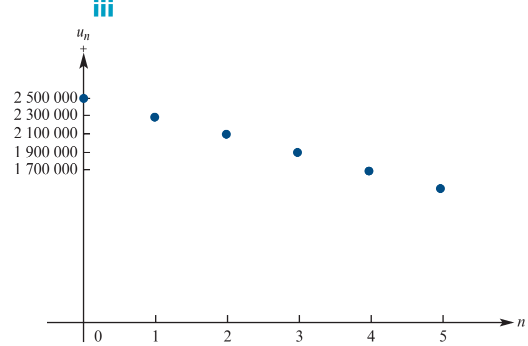
1 **a** $L_{n+1} = L_n + 5, L_1 = 20$
b $L_n = 15 + 5n$ **c** 65 metres
2 **a** $N_{n+1} = 1.22N_n, N_1 = 1654$
b

Start of year	1	2	3	4	5
Number of deer	1654	2018	2462	3004	3665

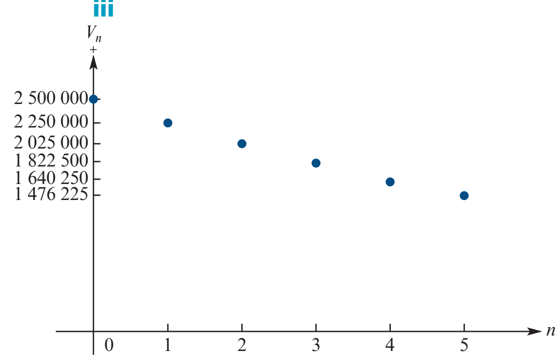
c



d 7 years
3 **a** **i** $V_n = 2500000 - 200000n$
ii 2300000, 2100000, 1900000, 1700000, 1500000
iii



b **i** $U_n = 2500000(0.9)^n$
ii 2250000, 2025000, 1822500, 1640250, 1476225
iii



c **i** Method 1
ii Method 2

ISBN 978-1-108-46154-2

Cambridge University Press

6D Topic 4 Earth geometry and time zones

Multiple-choice questions

1 B **2** B **3** E **4** E

Short-answer questions

Simple familiar questions

1 4114 km
2 15 038 km
3 **a** 264° **b** 17.6 hours
c 11:36 a.m. the next day
4 **a** 5727 km **b** 4281 km
c 15 735 km

Complex familiar questions

5 140°E, 22°N
6 4838 km
7 **a** 35 670 km **b** 25 190 km
8 **a** 11 120 km **b** 15 330 km
c 4210 km
9 0°, 26.5°W

Extended-response questions

Complex unfamiliar questions

1 **a** 556 km **b** 481.5 km
c 26°N or 26°S
2 **a** 7690 km **b** 7760 km **c** 70 km
d 3 hours 40 minutes. *A* is ahead of *B*

Chapter 7

Exercise 7A

1 **a** 2, 8, 14, 20, 26
b 5, 2, –1, –4, –7
c 1, 4, 16, 64, 256
d 10, 5, 2.5, 1.25, 0.625
e 6, 14, 30, 62, 126
f 12, 9, 7.5, 6.75, 6.375
2 **a** 4, 6, 8, 10, 12
b 24, 20, 16, 12, 8
c 2, 6, 18, 54, 162
d 50, 10, 2, 0.4, 0.08
e 5, 13, 29, 61, 125
f 18, 16.4, 15.12, 14.096, 13.2768
3 **a** 2, 5, 8, 11, 14
b 50, 45, 40, 35, 30
c 1, 3, 9, 27, 81
d 3, –6, 12, –24, 48
e 5, 9, 17, 33, 65
f 2, 7, 17, 37, 77
g –2, –1, 2, 11, 38
h –10, 35, –100, 305, –910
4 **a** 12, 57, 327, 1947, 11 667
b 20, 85, 280, 865, 2620
c 2, 11, 47, 191, 767
d 64, 15, 2.75, –0.3125, –1.078125
e 48 000, 45 000, 42 000, 39 000, 36 000
f 25 000, 21 950, 19 205, 16 734.5, 14 511.05
5 **a** $A_0 = 4, A_{n+1} = A_n + 2$
b $A_0 = 24, A_{n+1} = A_n - 4$
c $A_0 = 2, A_{n+1} = 3A_n$
d $A_0 = 50, A_{n+1} = \frac{A_n}{5}$
e $A_0 = 5, A_{n+1} = 2A_n + 3$
f $A_0 = 18, A_{n+1} = 0.8A_n + 2$
6 4 **7** 8

Exercise 7B

1 **a** $A_0 = 2000, A_{n+1} = A_n + 50$
b $A_0 = 6000, A_{n+1} = A_n + 252$
c $A_0 = 25\,000, A_{n+1} = A_n + 1600$
2 **a** 0.006 **b** 0.0093 **c** 0.0031
d 0.0052 **e** 0.019 **f** 0.0261
3 **a** 3.75% **b** 6.4% **c** 9.6%
d 11.28% **e** 6.12% **f** 10.92%
4 **a** $A_0 = 5000, A_{n+1} = 1.084A_n$
b $A_0 = 8500, A_{n+1} = 1.042A_n$
c $A_0 = 26\,000, A_{n+1} = 1.126A_n$
5 **a** $A_0 = 2000, A_{n+1} = 1.007A_n$
b $A_0 = 24\,000, A_{n+1} = 1.0064A_n$
c $A_0 = 16\,000, A_{n+1} = 1.0168A_n$
d $A_0 = 2800, A_{n+1} = 1.0124A_n$
e $A_0 = 34\,000, A_{n+1} = 1.0018A_n$
f $A_0 = 18\,000, A_{n+1} = 1.0024A_n$
6 **a** 1 year: $2076, 2 years: $2152, 3 years: $2228
b 14 years
7 **a** 1 year: $7518, 2 years: $8036, 3 years: $8554
b 6 years
8 **a** 1 month: $6028.80, 2 months: $6057.74, 3 months: $6086.82
b $86.82 **c** $28.94 **d** 7 months
9 **a** 1 month: $8488.20, 2 months: $8577.33, 3 months: $8667.39
b $88.20 **c** $89.13 **d** $267.39
e 11 months
10 **a** 0.5%
b $A_0 = 7600, A_{n+1} = 1.005A_n$
c $7830.87
11 **a** 2%
b $A_0 = 3500, A_{n+1} = 1.02A_n$
c $3788.51
d $312.30

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 7C

1 a

Principal: $12 000 Annual interest rate: 5.2%	
Compounds per year	Balance after 1 year
1	$12 624.00
2	$12 632.11
4	$12 636.27
12	$12 639.09
52	$12 640.18

b

Principal: $25 000 Annual interest rate: 9.4%	
Compounds per year	Balance after 1 year
1	$27 350.00
2	$27 405.23
4	$27 434.14
12	$27 453.94
52	$27 461.66

c

Principal: $67 500 Annual interest rate: 5.7%	
Compounds per year	Balance after 1 year
1	$71 347.50
2	$71 402.33
4	$71 430.52
12	$71 449.62
52	$71 457.04

d

Principal: $180 000 Annual interest rate: 3.47%	
Compounds per year	Balance after 1 year
1	$186 246.00
2	$186 300.18
4	$186 327.75
12	$186 346.30
52	$186 353.48

2 a Weekly **b** $1.46
3 a Quarterly **b** $9.13

Exercise 7D

1 a 6.38% **b** 8.76% **c** 4.91%
d 13.10% **e** 7.64%
2 a More compounds earn more interest
b 4.68%
c 4.70%
d More frequent compounds (monthly) has a higher effective interest rate.
3 a Fewer compounds charge less interest
b 8.25%
c 8.24%
d Less frequent compounds (monthly) has a lower effective interest rate.
4 a A: 8.62% B: 8.11%
b A: $3018.10 B: $2837.08
c Luke should choose loan B as he will pay less interest than with loan A.
5 a A: 5.43% B: 5.61%
b A: $7602.92 B: $7860.27
c Sharon should choose investment B as she will earn more interest than with investment A.

Exercise 7E

1 a $2771.79 **b** $16 751.89
c $6377.17 **d** $32 709.21
e $9785.98 **f** $16 563.11
2 a $8500 **b** $24 000 **c** $35 000
d $6400 **e** $31 000
3 a 4.14% **b** 6.12% **c** 6.12%
d 5.28% **e** 4.92%
4 6 years
5 4.90%
6 50 years
7 48 quarters
8 4.00%
9 6.25%

Chapter 7 review

Multiple-choice questions

1 B **2** C **3** C **4** A
5 B **6** D **7** E **8** D
9 B **10** C

ISBN 978-1-108-46154-2

Cambridge University Press

Short-answer questions

1 a $A_0 = 12500, A_{n+1} = 1.0065A_n$

b \$12828.18

2 a 1 month: \$5832.48

2 months: \$5865.14

3 months: \$5897.99

b 7 months

3 a \$1600

b 18.2%

c One week: \$1605.60, two weeks: \$1611.22, three weeks: \$1616.86

d 9 weeks

4 a 3.8%

b $A_0 = 2200, A_{n+1} = 1.038A_n$

c \$2553.95

5 a \$16411.74

b \$1411.74

6 a \$1528.33

b \$28.33

7 18.6%

Extended-response questions

1 a Bank A: 6.45% per annum Bank B: 6.27% per annum

b Bank A. It has the largest effective rate of interest and will earn more interest in one year than Bank B.

c \$9

2 a Bank A. It has the largest effective rate of interest (8.37% per annum compared to 8.35% per annum) and will earn more interest in the time of the investment.

b \$7

3 a i \$15701.99

ii \$15704.73

iii \$15705.80

b Effective annual rate of interest

c The weekly compounding investment has the highest effective annual rate of interest. It will earn more interest than the other compounding frequencies and so will have the most benefit to Eva. This can be seen by the fact that her account will be the highest after one year if she has weekly compounding.

4 a \$2300.00

b 17.68%

c i \$2428.36

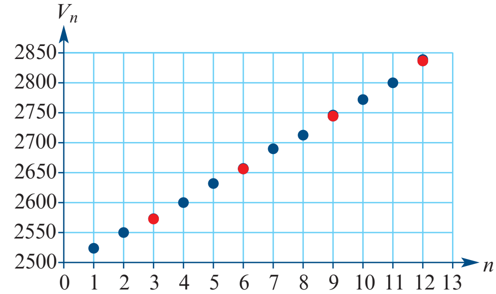
ii \$128.36

iii \$443.98

5 a

Month	Quarterly compounds	Monthly compounds
1		\$2526.00
2		\$2552.27
3	\$2578.00	\$2578.81
4		\$2605.63
5		\$2632.73
6	\$2658.43	\$2660.11
7		\$2687.78
8		\$2715.73
9	\$2741.38	\$2743.97
10		\$2772.51
11		\$2801.35
12	\$2826.91	\$2830.48

b i



ii The table shows that Lucille will pay more with monthly compounds compared to quarterly compounds. The graph of monthly compounds is higher on the axes than that for quarterly compounds.

c \$3.57

ISBN 978-1-108-46154-2

Cambridge University Press

Chapter 8

Exercise 8A

1 a i $A_0 = 6500, A_{n+1} = 1.14A_n - 1800$
ii $3438.76
b i $A_0 = 14000, A_{n+1} = 1.028A_n - 2000$
ii $9039.67
c i $A_0 = 22000, A_{n+1} = 1.018A_n - 1000$
ii $20155.19
d i $A_0 = 85000, A_{n+1} = 1.0067A_n - 1800$
ii $81283.71
e i $A_0 = 150000, A_{n+1} = 1.0054A_n - 1700$
ii $147315.56
f i $A_0 = 245000, A_{n+1} = 1.0016A_n - 1200$
ii $242572.12

2 a $600 **b** $1201.44
c $501.44

3 a $860 **b** $868.19
c $168.19

4 a 8.64% **b** $10282.98
c $282.98

5 a 11.4% **b** $4717.02
c $217.02

6 a $A_0 = 20000, A_{n+1} = 1.0059A_n - 600$
i $17561.39
ii $561.39
b $A_0 = 20000, A_{n+1} = 1.0059A_n - 800$
i $16549.52
ii $549.52
c $11.87

Exercise 8B

1 a

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	8400.00
1	500	79.80	420.20	7979.80
2	500	75.81	424.19	7555.61
3	500	71.78	428.22	7127.39
4	500	67.71	432.29	6695.10
5	500	63.60	436.40	6258.70

b $358.70

2 a

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	2000.00
1	100.00	7.00	93.00	1907.00
2	100.00	6.67	93.33	1813.67
3	100.00	6.35	93.65	1720.02
4	100.00	6.02	93.98	1626.04
5	100.00	5.69	94.31	1531.73

b i $20.02 **ii** $26.04

3

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	4600.00
1	300.00	18.40	281.60	4318.40
2	300.00	17.27	282.73	4035.67
3	300.00	16.14	283.86	3751.81
4	450.00	15.01	434.99	3316.82
5	450.00	13.27	436.73	2880.09

Balance after five repayments = $2880.09

4

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	7500.00
1	$300.00	71.25	228.75	7271.25
2	$300.00	69.08	230.92	7040.33
3	$600.00	66.88	533.12	6507.21
4	$600.00	61.82	538.18	5969.03
5	$600.00	56.71	543.29	5425.74

Balance after five repayments = $5425.74

5 a i $14680.68
ii Too much would be paid with the final payment at the end of the 11th month. It must be less than normal at $2086.82
iii $516.82
b i $1194.53 **ii** $494.53
iii $22.29
c Better to be made as repayment number 1. It would save the most interest because the payment was made earlier in the term of the loan.

ISBN 978-1-108-46154-2

Exercise 8C

1 a $6061.91 b $12 095.13
  c $168 519.40 d $45 196.78
  e $33 735.99
2 a $33.16 b $7674.57
  c $19 088.26 d $58 652.52
  e $26 369.97
3 a $857.09 b $88 076.01
4 a $2616.00 b 56 months
5 a 55 months b $3124.12
6 a $349.43 b $32 437.90 c $418.66

Chapter 8 review

Multiple-choice questions

1 B 2 A 3 E 4 B 5 C
6 B 7 C 8 E 9 A 10 D

Short-answer questions

1 a $9500
  b $250.00
  c $8188.07
2 a 0.49%
  b $A_0 = 250000, A_{n+1} = 1.0049A_n - 2400$
  c $242 863.07
  d 9
3 a $3500
  b $600.00
  c i 0.92% ii 11.04%
  d $613.87
4 a

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	125 000.00
1	1000.00	387.50	612.50	124 387.50
2	1000.00	385.60	614.40	123 773.10
3	1000.00	383.70	616.30	123 156.80
4	1000.00	381.79	618.21	122 538.59

  b $109 763.84
5 a $72 923.58
  b $923.58
6 3.95%

Extended-response questions

1 a $1500 b $11 943.00
  c $2463.37
2 a $51 800.72 b $4200.72
  c $29 134.15
3 a 307 months b 145
4 a 159 months b $1327.00
5 a i $135 481.70 ii $34 741.70
  b i $113 840.36 ii $21 641.34
   iii Nicholas will be charged interest on the outstanding balance of his loan every month. He should reduce the outstanding balance of his loan as soon as possible in order to minimise the interest that he is charged.
  c $1058.37 d $1070.48
  e The balance of Nicholas' loan after each year of the loan is shown in the table below.

n	interest rate for year n	balance after n years
1	4.5	$156 654.27
2	4.55	$148 003.11
3	4.6	$139 023.50
4	4.65	$129 690.86
5	4.7	$119 914.88

If Nicholas chooses the fixed interest home loan, with the maximum affordable repayments of $1300 per month, his balance after 10 years would be $63 537.41.
If he chooses the variable interest home loan, with maximum affordable repayments of $1300 per month and with annual increases of 0.05% in the interest rate each year, the balance would be $64 671.89.
If Nicholas was not intending to make lump sum payments, the best loan for him to take would be the fixed rate loan as he would pay the loan out after a shorter time and pay less interest.

Chapter 9

Exercise 9A

1 a i $A_0 = 0, A_{n+1} = 1.025 \times A_n + 5000$
   ii $15 378.13
  b i $A_0 = 0, A_{n+1} = 1.016 \times A_n + 6500$
   ii $19 813.66

ISBN 978-1-108-46154-2

Cambridge University Press

c **i** $A_0 = 320000, A_{n+1} = 1.009 \times A_n + 8000$
ii $352934.64
d **i** $A_0 = 460000, A_{n+1} = 1.0058 \times A_n + 4200$
ii $480723.73
e **i** $A_0 = 845000, A_{n+1} = 1.0041 \times A_n + 7500$
ii $878028.55
f **i** $A_0 = 1250000, A_{n+1} = 1.0012 \times A_n + 2700$
ii $1262615.13

2 **a** **i** $A_0 = 120500, A_{n+1} = 1.028 \times A_n - 8000$
ii $106229.79
b **i** $A_0 = 276000, A_{n+1} = 1.0126 \times A_n - 4600$
ii $272590.20
c **i** $A_0 = 358000, A_{n+1} = 1.0143 \times A_n - 25000$
ii $297501.26
d **i** $A_0 = 440000, A_{n+1} = 1.0036 \times A_n - 5000$
ii $429715.06
e **i** $A_0 = 845000, A_{n+1} = 1.0067 \times A_n - 9600$
ii $833105.16
f **i** $A_0 = 1360000, A_{n+1} = 1.0015 \times A_n - 2900$
ii $1357416.13

3 **a** The recurrence relation has a value subtracted at the end.
b $1030.00 **c** $2030.50 **d** $120.50

4 **a** The recurrence relation has a positive value added at the end.
b $300 **c** $7666.58 **d** $166.58

5 **a** **i** 2.16% **ii** $62942.44
b **i** $7500 **ii** $62942.44
iii $25875.72
c **i** $442.44 **ii** $433.28
iii $875.72

6 **a** **i** 3.72% **ii** $269729.13
b **i** $1800 **ii** $269729.13
iii $266828.64
c **i** $2479.13 **ii** $2499.51
iii $4978.64

7 **a** $A_0 = 125\,000, A_{n+1} = 1.0013 \times A_n + 695$
b 6
c $1502.75

8 **a** $A_0 = 32000, A_{n+1} = 1.0036 \times A_n - 3500$
b 4
c 9
d $1094.43
e $1098.37

9 **a** $A_0 = 54000, A_{n+1} = 1.0064 \times A_n + 1500$
b $65252.30
c $A_0 = 65252.30, A_{n+1} = 1.0063 \times A_n - 1800$
d $4586.48

Exercise 9B

1 **a**

Deposit number	Deposit	Interest	Principal increase	Balance of annuity
0	0	0	0.00	50000.00
1	4000.00	400.00	4400.00	54400.00
2	4000.00	435.20	4435.20	58835.20
3	4000.00	470.68	4470.68	63305.88
4	4000.00	506.45	4506.45	67812.33
5	4000.00	542.50	4542.50	72354.83

b

Deposit number	Deposit	Interest	Principal increase	Balance of annuity
0	0	0	0.00	135000.00
1	1200.00	486.00	1686.00	136686.00
2	1200.00	492.07	1692.07	138378.07
3	1200.00	498.16	1698.16	140076.23
4	1200.00	504.27	1704.27	141780.50
5	1200.00	510.41	1710.41	143490.91

2 **a**

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	25000.00
1	1000.00	143.75	856.25	24143.75
2	1000.00	138.83	861.17	23282.58
3	1000.00	133.87	866.13	22416.45
4	1000.00	128.89	871.11	21545.34
5	1000.00	123.89	876.11	20669.23

b

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	380000.00
1	12000.00	4560.00	7440.00	372560.00
2	12000.00	4470.72	7529.28	365030.72
3	12000.00	4380.37	7619.63	357411.09
4	12000.00	4288.93	7711.07	349700.02
5	12000.00	4196.40	7803.60	341896.42

ISBN 978-1-108-46154-2

3 **a** \$164000.00 **b** 5.8%
c \$155822.47 **d** \$747.95
e \$2765.26
f **i** 3805.16 **ii** \$3805.16
g

6	3500.00	721.46	2778.54	147526.62
7	3500.00	708.13	2791.87	144734.75

Exercise 9C

1 **a** \$6477.82 **b** \$12118.63
c \$101846.82 **d** \$39905.39
e \$11058.58
2 **a** \$144226.52 **b** \$295234.80
c \$327661.26 **d** \$428306.28
e \$234134.63
3 **a** 58 **b** \$430.43
c \$10380.43
4 21
5 **a** 120
b **i** \$474.81 **ii** \$474.29
c \$16976.68
6 **a** \$412000.00 **b** \$65662.50
7 **a** 5.27% **b** \$396300.89
c 487 **d** 5

Exercise 9D

1 **a** \$4791.67 **b** \$4791.67
2 **a** \$9790.50 **b** \$642 000
3 **a** Verified **b** 3.125%

$$P = \frac{d}{i} = \frac{2500}{0.025} = 100000$$

4 **a** \$18518.52 **b** 4.17%
c Verified
5 **a** \$10000
b **i** \$87.50 **ii** 0.875% **iii** 3.5%
c No impact

Chapter 9 review

Multiple-choice questions

1 A **2** B **3** D **4** E **5** C
6 C **7** B **8** D **9** D **10** C

Short-answer questions

1 **a** \$624000 **b** \$2500 **c** \$613834.21
2 **a** $A_0 = 345000$, $A_{n+1} = 1.0115 \times A_n - 12000$
b \$295397.96 **c** \$22397.96
3 **a** \$84000.00 **b** \$14500.00
c **i** 0.47% **ii** 5.64%
d \$12888.23
4 \$16528.65
5 **a** \$97222.22 **b** 5.83%

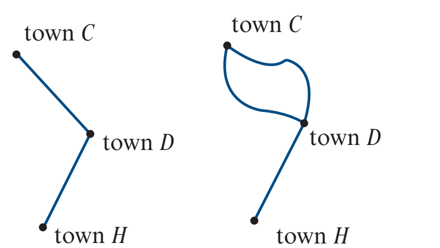
Extended-response questions

1 **a** \$5250
b \$441
c **i** \$5335.88 **ii** \$57934.44
2 **a** \$2460
b \$191.88
c **i** \$27519.87 **ii** \$60960.18
d **i** \$7732.89 **ii** \$2649.34
3 **a** \$720.00
b **i** $A_0 = 150000$, $A_{n+1} = 1.0048 \times A_n - 2000$
ii \$142227.25 **iii** 16
c **i** 36 **ii** \$1850.89
4 **a** \$496000.00 **b** \$69623.14
c 86 **d** \$2047.67
5 **a** 112 **b** \$334651.73
c 89 **d** \$1554.55
6 **a** \$2941.59
b 8% of \$1800 is more than 8.2% of \$1600 and so Byron will have more deposited into his superannuation account each fortnight with Option 1.
c **i** \$102107.91 **ii** \$22275.91
7 **a** \$314296.20 **b** \$7029.52

Chapter 10

Exercise 10A

1 **a** **i** 3 **ii** 2 **iii** 1
b **i** 7 **ii** 14
iii Verified as 14
c Two answers are possible, as shown below.



d H and D

ISBN 978-1-108-46154-2

Cambridge University Press

2 a Chelsea **b** Wayne

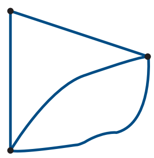
c Samantha

d i deg(Samantha) = 4

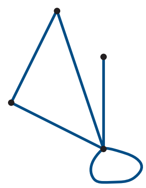
ii deg(Eli) = 4

3 Many different answers are possible. One possibility for each is shown.

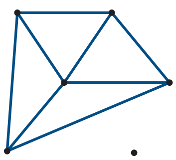
a



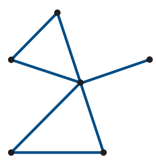
b



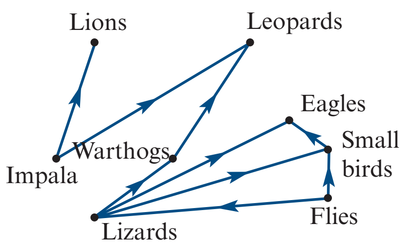
c



d

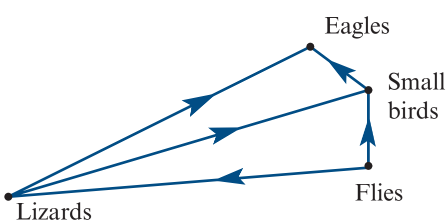


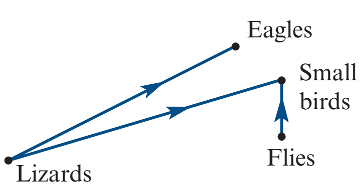
4 a



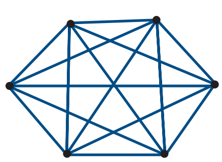
b deg(Warthog) = 2

c Multiple answers are possible. Two of these are shown below.





5 a



b 15

Exercise 10B

1 a

	A	B	C	D
A	0	1	1	0
B	1	0	1	1
C	1	1	0	0
D	0	1	0	0

b

	A	B	C	D
A	0	1	1	0
B	1	0	0	1
C	1	0	0	1
D	0	1	1	0

c

	A	B	C	D
A	0	1	0	0
B	1	0	0	0
C	0	0	0	1
D	0	0	1	0

d

	A	B	C	D
A	0	1	1	1
B	1	0	1	1
C	1	1	0	1
D	1	1	1	0

e

	A	B	C	D	E	F
A	0	1	1	0	0	0
B	1	0	0	1	0	0
C	1	0	0	1	0	0
D	0	1	1	0	0	0
E	0	0	0	0	0	1
F	0	0	0	0	1	0

f

	A	B	C	D
A	0	0	0	0
B	0	0	0	1
C	0	0	0	2
D	0	1	2	0

2 a

	A	B	C
A	0	1	1
B	1	0	1
C	0	0	0

b

	A	B	C	D
A	0	0	1	0
B	1	0	0	1
C	0	1	0	0
D	0	1	1	0

c

	M	N	P	Q
M	0	1	0	1
N	1	0	0	0
P	0	1	0	0
Q	0	0	1	1

d

	E	F	G	H
E	0	1	0	1
F	1	0	1	0
G	0	1	0	1
H	0	0	1	0

ISBN 978-1-108-46154-2

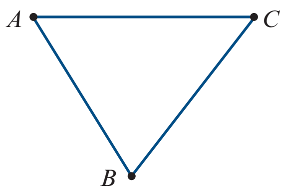
e

	A	B	C	D	E
A	0	1	1	0	0
B	0	1	0	1	0
C	1	1	0	1	0
D	0	0	0	0	1
E	0	0	1	1	0

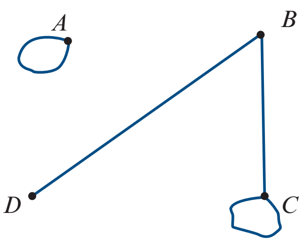
f

	A	B	C	D	E	F
A	0	1	0	0	0	0
B	0	0	1	1	0	0
C	1	0	0	1	0	0
D	0	0	0	0	0	0
E	0	0	1	1	0	0
F	0	0	0	0	1	0

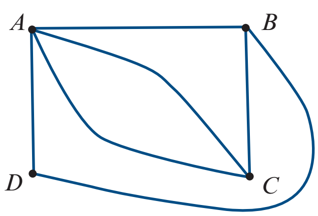
3 a



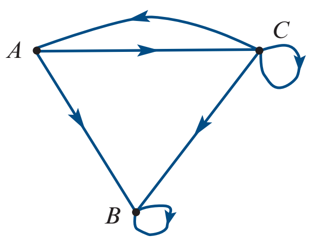
b



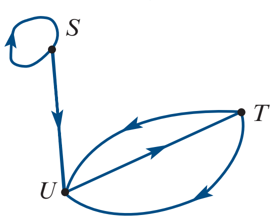
c



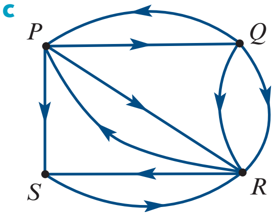
4 a



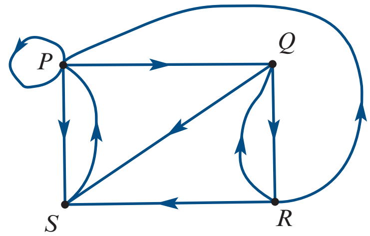
b



c



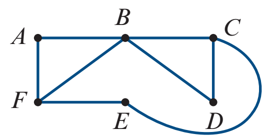
d



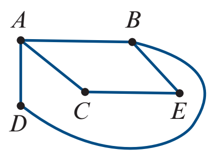
Exercise 10C

1 a i **b** ii **c** ii

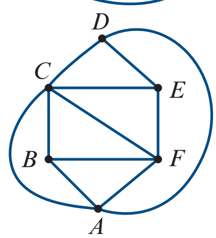
2 a



b



c



d not possible

3 a $v = 8, e = 12, f = 6$

b $v = 6, e = 12, f = 8$

c $v = 7, e = 12, f = 7$

4 a $f = 4$ **b** $v = 12$

c $f = 11$ **d** $e = 19$

Exercise 10D

1 a path **b** trail

c path **d** closed walk

e trail **f** path

2 a closed path or cycle **b** open walk only

c open walk only **d** trail

e closed walk

f closed path or cycle

ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 10E

1 a i Semi-Eulerian
ii *E–A–B–E–D–B–C–D–A*
b i Neither
c i Semi-Eulerian
ii *A–C–E–C–B–D–E–F*
d i Eulerian
ii *A–B–C–E–D–C–A*
e i Eulerian
ii *E–F–D–E–A–B–D–C–B–E*

2 a *A*:4, *B*:2, *C*:5, *D*:2, *E*:4, *F*:4, *G*:3
b i *C* or *G*
ii *G* or *C*
c i *C* and *G*
ii Eulerian trail

Exercise 10F

1 a i *A–B–C–F–I–H–E–G–D*
ii *E–G–D–A–B–C–F–I–H–E*
b i *A–B–C–D–E–F*
ii *E–F–A–B–C–D–E*
c i *A–B–D–C–E*
ii *E–A–B–D–C–E*

2 *F–A–B–C–D–E–H–G*

3 a 5
b i Hamilton path
ii *E–W–D–C–B–A–F*

Exercise 10G

1 a *D* and *E*
b 17 minutes
c 8 minutes
d 36 minutes (*A–B–C–D–E*)

2 11

3 a 34 km
b 56 km
c Two answers are possible: *A–E–F–G–I* or *A–C–F–G–I*

4 19 km

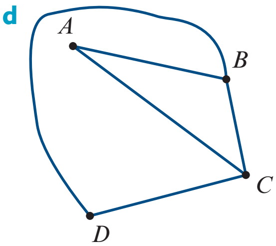
Chapter 10 review

Multiple-choice questions

1 A	**2** C	**3** D	**4** A	**5** D
6 C	**7** A	**8** B	**9** B	**10** B

Short-answer questions

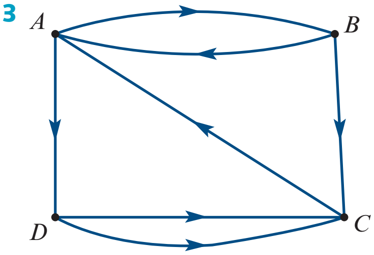
1 a 3 **b** 2 **c** *A* and *D*
d



2

$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

3



4 a Mackay Masters and Gladstone Gladiators
b 2
c Bundaberg Braves

5 a 4 **b** 6
c 8 **d** 4 + 6 – 8 = 2

6 a *C* and *D* **b** 12 hours
c 13 hours

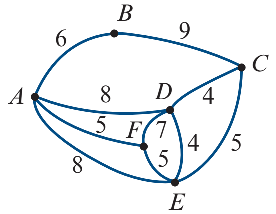
Extended-response questions

1 a It can be drawn so that none of the edges in the graph cross over each other.
b $v + f + e = 9 + 7 - 14 = 2$
c 750 m
d i Yes, all vertices have even degrees and so the graph is Eulerian.
ii Multiple answers possible. One is: Office–C5–C7–C8–C6–C5–C4–C3–C2–C4–C1–C2–C8–C1–office
e i Hamiltonian cycle
ii C7 to park office. Other answers possible.
iii Office–C1–C2–C3–C4–C5–C6–C8–C7–office, or same route in reverse order. Other answers possible.

ISBN 978-1-108-46154-2

Cambridge University Press

2 a i



ii

	A	B	C	D	E	F
A	0	1	0	1	1	1
B	1	0	1	0	0	0
C	0	1	0	1	1	0
D	1	0	1	0	1	1
E	1	0	1	1	0	1
F	1	0	0	1	1	0

b i 45 km (at minimum)

ii Some vertices are visited more than once.

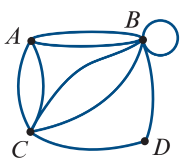
iii $F - E - D - C - B - A - F$

iv 33 km (for route above; other answers possible)

c F and C

3 a 7 **b** 2

c



d Vertices are not all even.

4 a Park Entrance – Information Centre – Boathouse – Campsite – Lookout – Park Entrance.

Park Entrance – Lookout – Campsite – Boathouse – Information Centre – Park Entrance.

b Yes. There are exactly two vertices that have an odd degree.

c i 30 minutes

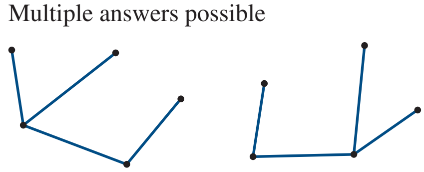
ii 2 hours and 45 minutes

Chapter 11

Exercise 11A

1 a 11 **b** 9

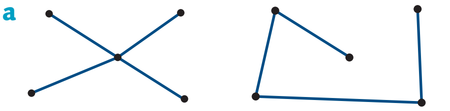
c Multiple answers possible



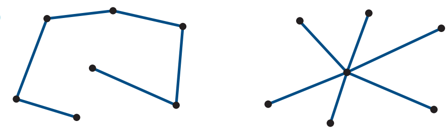
2 A, B, D

3 Multiple answers possible.

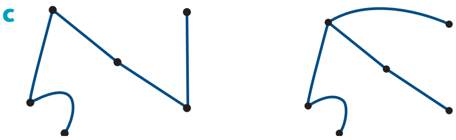
a



b

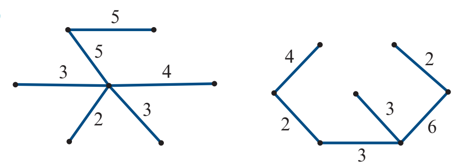


c



4 a 6

b

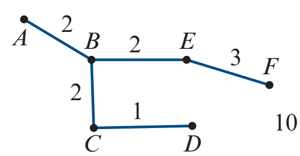


Note: other answers are possible

c 22, 20

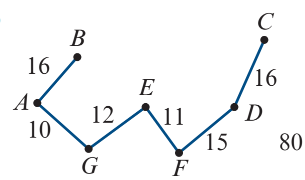
Note: other answers are possible

5 a



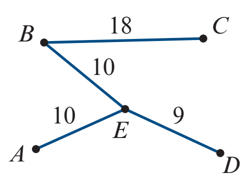
10

b



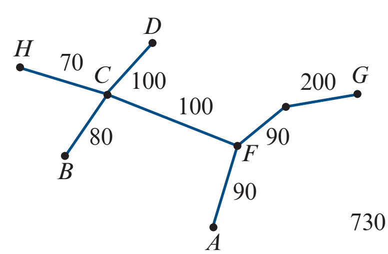
80

c



47

d



730

6 44

7 94

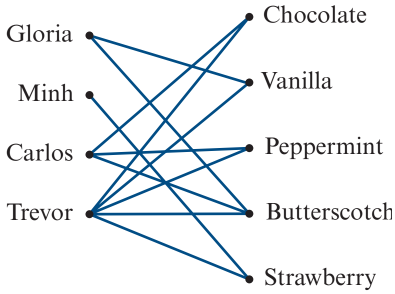
ISBN 978-1-108-46154-2

Cambridge University Press

Exercise 11B

1 a There are two distinct groups for vertices, people and ice-creams that must be matched together.

b

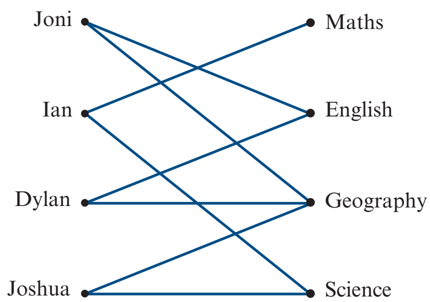


c 5

d Gloria

e Gloria – Vanilla, Minh – Strawberry, Carlos – Peppermint, Trevor – Chocolate

2 a

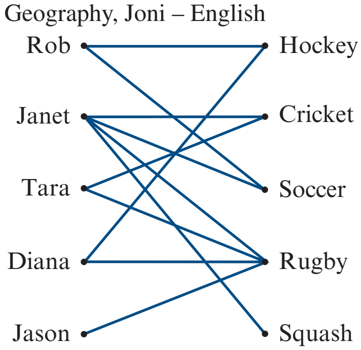


b Ian is the only teacher who can teach Maths and so he cannot teach Science. Joshua is the only other teacher who can teach Science and so he must take this class.

c Ian – Maths, Joshua – Science, Dylan – English, Joni – Geography

Ian – Maths, Joshua – Science, Dylan – Geography, Joni – English

3 a



b Jason can only coach Rugby and so Diana cannot. The only other sport Diana can coach is Hockey.

c Jason – Rugby, Diana – Hockey, Rob – Soccer, Janet – Squash, Tara – Cricket

4 Joe – C, Meg – A, Ali – B

5 Stars – Away, Champs – Home, Wests – Neutral

Stars – Away, Champs – Neutral, Wests – Home

Cost = $20 000

6 $A - Y, B - Z, C - X, D - W$

7 a $W - D, X - A, Y - B, Z - C$

b Minimum cost = 11, one possible allocation is W – B, X – C, Y – D, Z – A

8 A – Karla, B – Raj, C – Mark, D – Jess

A – Mark, B – Karla, C – Raj, D – Jess

Distance = 55 km

9 Dimitri – 800 m, John – 400 m, Carol – 100 m, Elizabeth – 1500 m

Exercise 11C

1 Cut C_1 – 14, Cut C_2 – 12, Cut C_3 – 21

2 Cut C_1 – 12, Cut C_2 – 16, Cut C_3 – 16

3 a 9 **b** 11 **c** 8 **d** 18

4 a Cut A – 14, Cut B – 23, Cut C – 12, Cut D – 16, Cut E – not a cut

b It does not completely separate the source from the sink.

c 12

Chapter 11 review

Multiple-choice questions

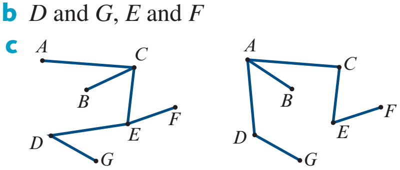
1 B	**2** C	**3** A	**4** D	**5** C
6 E	**7** D	**8** A	**9** C	**10** A

Short-answer questions

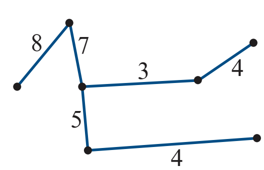
1 a 3

b D and G, E and F

c



2



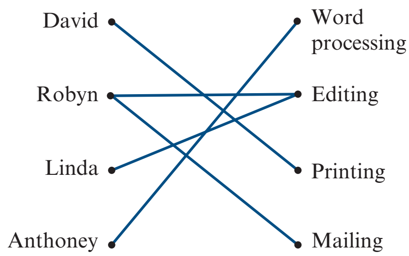
ISBN 978-1-108-46154-2

Cambridge University Press

3 **a** 2 **b** Audrey

c Audrey – Natives, Brian – Aloes, Cameron – Cactus, Daphne – Grasses

4 **a**



b David – Printing, Robyn – Mailing, Linda – Editing, Anthony – Word processing

5 **a** Julia **b** Mario

c Julia – Assembling, Mario – Painting, Sylvana – Preparing, George – Cutting

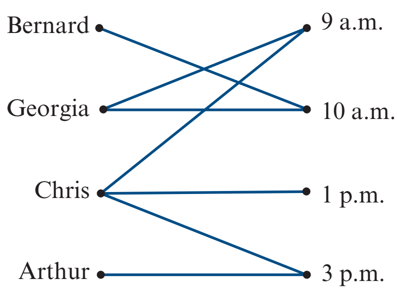
6 **a** 26 **b** 20

Extended-response questions

1 490

2 8

a



b Chris

c Bernard – 10 a.m., Georgia – 9 a.m., Chris – 1 p.m., Arthur – 3 p.m.

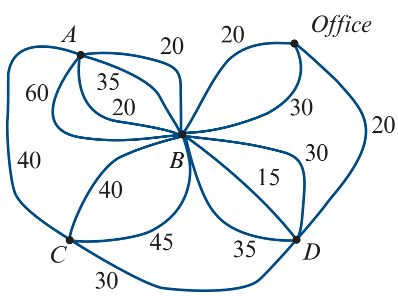
3 Ann – *D*, Bianca – *B*, Con – *C*, David – *E*

4 **a** 600 litres per minute

b 72 000 L **c** 4.5 min

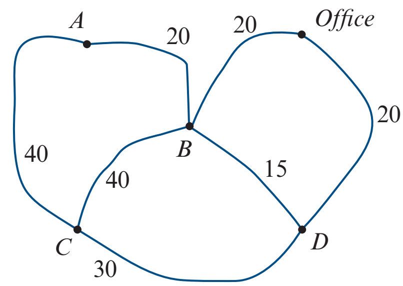
5 **a** **i** 4

ii



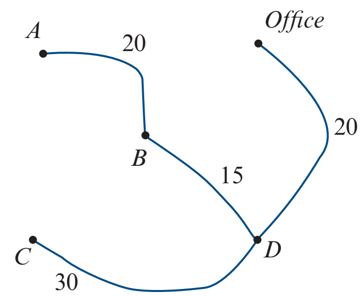
iii *A*, *B*, *D*

b **i**

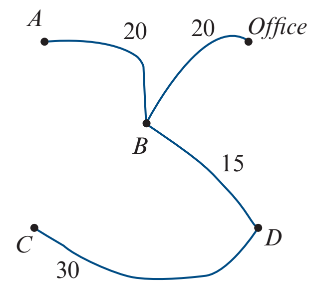


ii $44 000

c **i**



or



ii $20 400

d **i** 2.5 minutes

ii 3.5 minutes

Chapter 12

Exercise 12A

1 **a**

Activity	Immediate Predecessors
A	–
B	–
C	*A*
D	*A*
E	*B*, *C*
F	*D*
G	*E*

ISBN 978-1-108-46154-2

b

Activity	Immediate Predecessors
P	–
Q	P
R	P
S	Q
T	Q
U	S, V
V	R
W	R
X	T, U

c

Activity	Immediate Predecessors
J	–
K	–
L	J
M	N
N	K
O	K
P	N
Q	L, M
R	P
S	O, R
T	Q

d

Activity	Immediate Predecessors
A	–
B	–
C	A
D	A
E	B, D
F	C, E
G	D, B
H	B

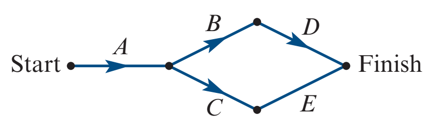
e

Activity	Immediate Predecessors
P	–
Q	–
R	P
S	P
T	Q
U	R
V	S
W	S, T
X	U
Y	W
Z	V, X, Y

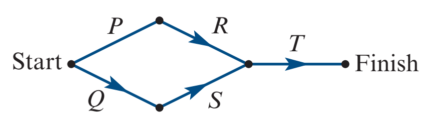
f

Activity	Immediate Predecessors
A	–
B	A
C	A
D	A
E	B
F	C, D
G	D
H	E, F, G
I	G
J	I
K	H

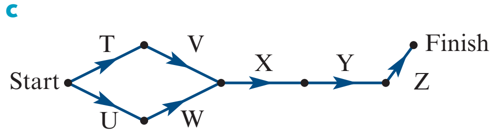
2 a



b



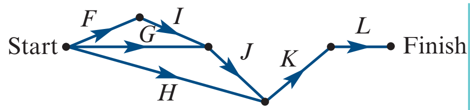
c



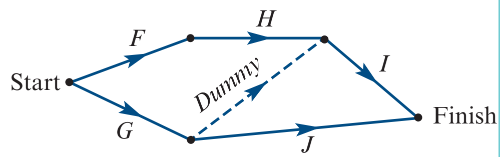
ISBN 978-1-108-46154-2

Cambridge University Press

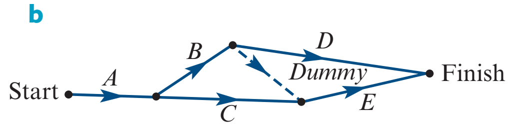
d



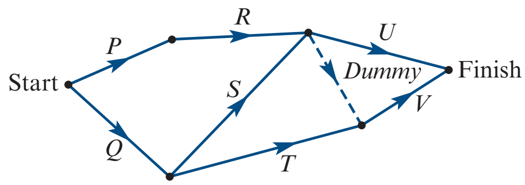
3 a



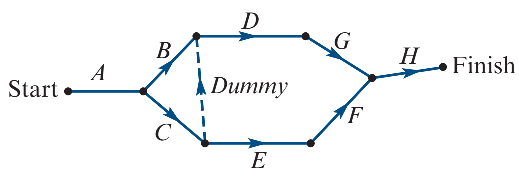
b



c



d

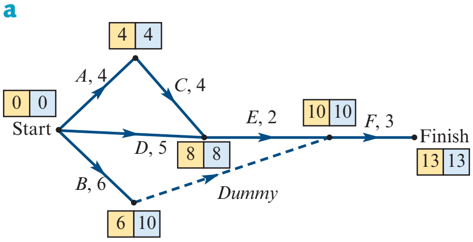


Exercise 12B

1 a 38
b 8
c 9
d 26
e i 18 **ii** EST is not the same as LST
f A

2 a 3 **b** 5 **c** 2
d 13 **e** $A - C$

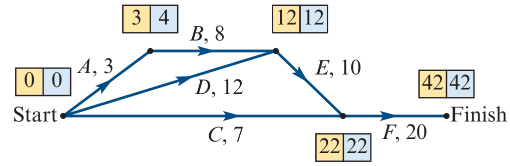
3 a



b 13
c $A - C - E - F$

d i B: 4, D: 3
ii B: 4, D: 3

4 a



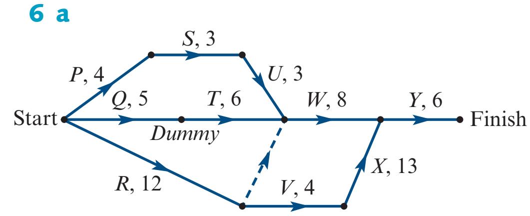
b 42
c $D - E - F$
d i A: 1, B: 4, C: 15
ii B: 1, A: 1, C: 15

5 a

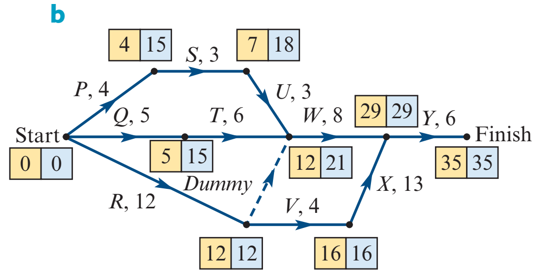
Activity	Duration	EST	LFT	LST	Float
A	10	**0**	11	1	1
B	**9**	0	9	**0**	**0**
C	3	0	**17**	14	14
D	2	**10**	13	11	**1**
E	4	9	**13**	9	0
F	**1**	13	14	**13**	0
G	1	**13**	15	14	1
H	**3**	14	**17**	14	**0**
I	2	14	17	**15**	1
J	1	**17**	18	17	0

b $B - E - F - H - J$

6 a



b



c 35 weeks
d $R - V - X - Y$

ISBN 978-1-108-46154-2

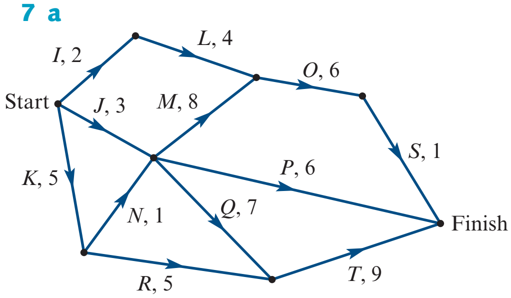
Cambridge University Press

e

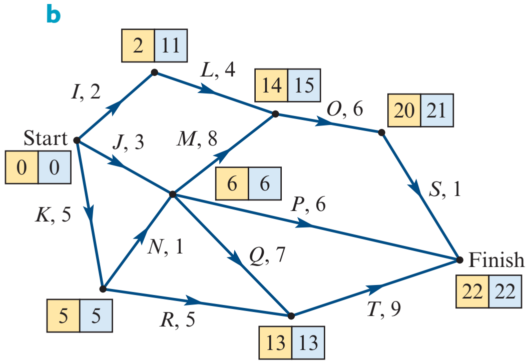
Activity	Duration	EST	LFT	LST	Float
P	4	0	15	11	11
Q	5	0	15	10	10
R	12	0	12	0	0
S	3	4	18	15	11
T	6	5	21	15	10
U	3	7	21	18	11
V	4	12	16	12	0
W	8	12	29	21	9
X	13	16	29	16	0
Y	6	29	35	29	0

f Verified. The activities with zero floats correspond to the path, thus it is verified.

7 a



b



c 22 weeks

d $K - N - Q - T$

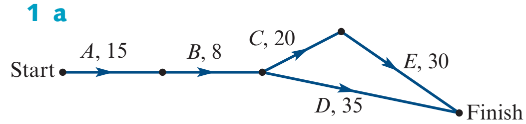
e

Activity	Duration	EST	LFT	LST	Float
I	2	0	11	9	9
J	3	0	6	3	3
K	5	0	5	0	0
L	4	2	15	11	9
M	8	6	15	7	1
N	1	5	6	5	0
O	6	14	21	15	1
P	6	6	22	16	10
Q	7	6	13	6	0
R	5	5	13	8	3
S	1	20	22	21	1
T	9	13	22	13	0

f Verified

Exercise 12C

1 a



b 73 minutes

c i

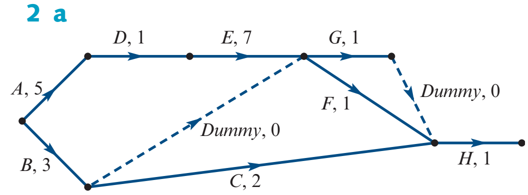
Activity	Duration	EST	LFT	LST	Float
A	15	0	15	0	0
B	8	15	23	15	0
C	20	23	43	23	0
D	35	23	73	38	15
E	30	43	73	43	0

ii Cleaning the interior

iii $A - B - C - E$

d 38 minutes

2 a



b 15 days

ISBN 978-1-108-46154-2

Cambridge University Press

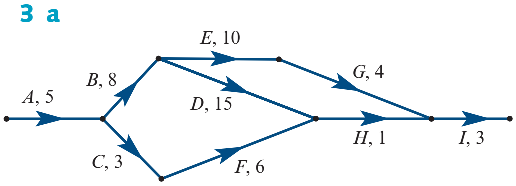
c i

Activity	Duration	EST	LFT	LST	Float
A	5	0	5	0	0
B	3	0	12	10	10
C	2	3	14	12	9
D	1	5	6	5	0
E	7	6	13	6	0
F	1	13	14	13	0
G	1	13	14	13	0
H	1	14	15	14	0

ii $A-D-E-F-H$ and $A-D-E-G-H$

d This would delay the project by 2 days.

3 a



b 32 weeks

c i

Activity	Duration	EST	LFT	LST	Float
A	5	0	5	0	0
B	8	5	13	5	0
C	3	5	22	19	14
D	15	13	28	13	0
E	10	13	25	15	2
F	6	8	28	22	14
G	4	23	29	25	2
H	1	28	29	28	0
I	3	29	32	29	0

ii $A-B-D-H-I$

d i The project would be completed in a minimum of 30 weeks.

ii Nothing. This activity has a float of 14 and so it could be extended in duration by 14 weeks.

iii The project would be completed in a minimum of 37 weeks.

Chapter 12 review

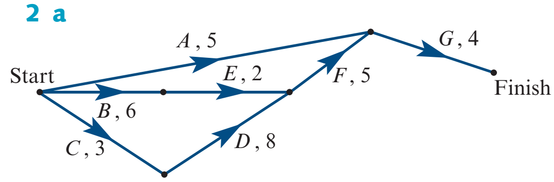
Multiple-choice questions

1 B **2** E **3** D **4** E **5** E
6 D **7** C **8** B **9** A **10** E

Short-answer questions

1 a *C* **b** *F* and *G* **c** 2

2 a



b 20 weeks

3 a 30 days **b** 13 days **c** 1 day

d $B-C-E-G-J-K$

4 a

Activity	Immediate predecessors
A	–
B	–
C	–
D	*A*
E	*C*
F	*B*, *E*
G	*B*, *E*
H	*B*, *E*
I	*G*
J	*D*, *F*
K	*D*, *F*
L	*J*
M	*H*, *K*
N	*I*

b 26 hours

c $B-F-J-L$

Extended-response questions

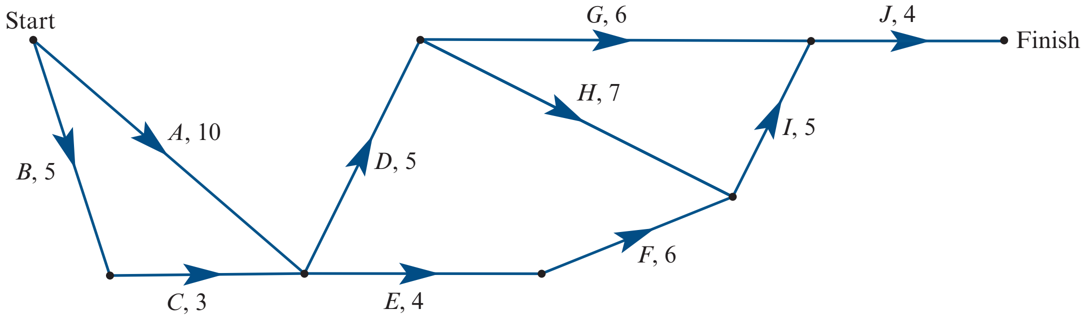
1 a 5 hours

b 24 hours

c 7 hours

ISBN 978-1-108-46154-2

2 a



b 31 days

c

Activity	Duration	EST	LFT	LST	Float
A	10	0	10	0	0
B	5	0	7	2	2
C	3	5	10	7	2
D	5	10	15	10	0
E	4	10	16	12	2
F	6	14	22	16	2
G	6	15	27	21	6
H	7	15	22	15	0
I	5	22	27	22	0
J	4	27	31	27	0

d If an activity is on the critical path, it is an activity that cannot be delayed or extended in duration without affecting the overall minimum completion time of the project.

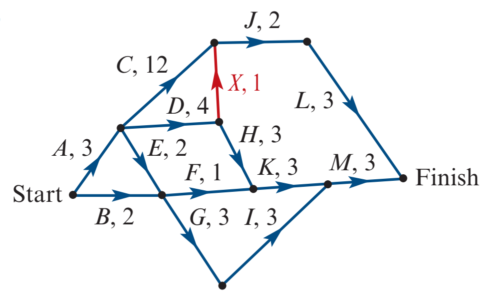
e The activities that have float equal to zero.

f $A - D - H - I - J$

g Nothing. C has a float time of 2, which means it can be delayed by up to 2 days without affecting the overall completion time of the project.

3 a Immediate predecessor of E is A. EST for l is 8. EST for M is 13.

b



c i 16 hours

ii Critical path is the sequence of activities that cannot be delayed without delaying the entire project.

d i 6 hours

ii $A - C - J - L$

iii 8 hours

4 a $B - E - H - J$ **b** 2 hours

c 6 hours **d** 14 hours

5 a A, B, C

b LST for B is 1, EST for E is 10, LST for I is 18

c i $A - D - F - I - J$

ii 27 months

d i $B - C - D - F - I - J$ and $A - D - F - I - J$

ii 24 months

Chapter 13

13A Topic 1 Loans, investments and annuities

Multiple-choice questions

1 E	**2** E	**3** D	**4** E	**5** C
6 A	**7** E	**8** A	**9** B	**10** C
11 B	**12** D	**13** D	**14** A	**15** C

ISBN 978-1-108-46154-2

Cambridge University Press

Short-answer questions

Simple familiar questions

1 a $A_1 = \$10064.00$; $A_2 = \$10128.41$; $A_3 = \$10193.23$
b 8
2 a $A_0 = 14500, A_{n+1} = 1.004 \times A_n$
b \$14674.70 **c** \$15957.95
3 a 0.85%
b $A_0 = 4500, A_{n+1} = 1.0085 \times A_n$
c \$4981.08
4 a \$28047.29 **b** \$3047.29
5 \$12500.00
6 a \$3000 **b** \$150
c 12.48% **d** \$2282.56
7 a 0.59%
b $A_0 = 50000, A_{n+1} = 1.0059 \times A_n - 500$
c \$48751.71 **d** 10
8 a \$9643.96 **b** \$343.96
9 a \$3889.03 **b** \$139.03
10 \$32000
11 a 3.72% **b** \$244843.13 **c** \$3843.13
12 a $A_0 = 145000, A_{n+1} = 1.0034 \times A_n - 2500$
b \$134896.53 **c** \$2396.53
13 \$34598.05
14 a \$602 **b** 6.86%

Extended-response questions

Complex familiar questions

1 a Bank A, it has the highest effective rate of interest
b Multiple answers possible.
Total amount withdrawn from Bank A is \$40264.69
Total amount withdrawn from Bank B is \$40126.15
2 a i $A_0 = 5000, A_{n+1} = 1.0095 \times A_n$
ii $A_n = 5000 \times (1.0095)^n$
b \$5399
3 a \$1257.62
b \$84322.76
c \$2045.92
4 a i \$1314.08
ii \$1.62
iii This is the amount that has been over-paid. The bank must refund this.
b 176
5 a \$485000 **b** \$100496.83 **c** 72
6 a 24 **b** \$221020.56
c 9 **d** \$16584.36

Complex unfamiliar questions

7 \$121649.93
8 a Loan 2. Amanda will pay less interest in total with this loan.
b No. Even with payments of \$1900, Loan 2 still results in less interest overall.

13B Topic 2 Graphs and networks

Multiple-choice questions

1 E	**2** B	**3** B	**4** B	**5** E
6 B	**7** C	**8** B	**9** B	**10** A

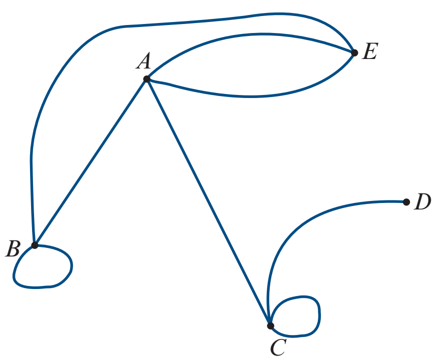
Short-answer questions

Simple familiar questions

1 a 3 **b** E **c** 2
d A, C, D **e** $B - C, F - D$
2 a R **b** Q
c

	P	Q	R	S
P	0	0	1	0
Q	1	0	0	1
R	0	0	0	1
S	1	0	1	0

3 a None
b 2
c



4 a 8 **b** 5
c 5 **d** $5 + 5 - 8 = 2$
5 a 20 km **b** 13 km

Extended-response questions

Complex familiar questions

1 a It is drawn so that no edges cross over each other.
b $v = 7, f = 6, e = 11, 7 + 6 - 11 = 2$

ISBN 978-1-108-46154-2

Cambridge University Press

c **i** C

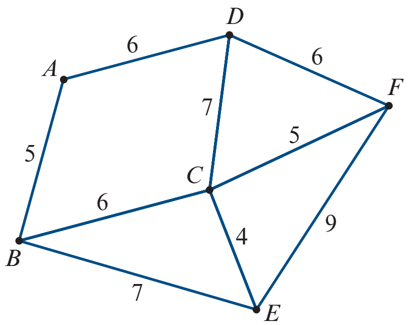
ii (open) trail

d **i** $C - B - D - E - F - G - C$

ii Hamiltonian cycle

2 **a** 7 km **b** 3 km **c** B and D

d **i**



ii

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Complex unfamiliar questions

3 **a** Multiple answers are possible. One is $A - B - D - E - C - A$

b BC, DE, CE

c $B - C - E - D - E - C - A - B - D$. Must start and end at odd-degree vertex.

d 9:54 a.m.

4 **a** 11 km

b There are exactly two odd-degree vertices in the network.

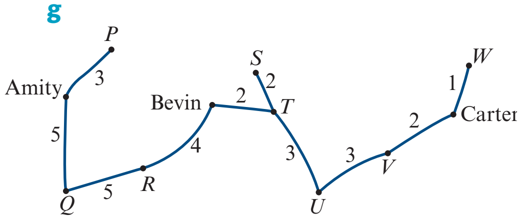
c Checkpoint V.

d Checkpoint U. If they do, they will have to travel one of the roads to Bevin a second time.

e Bevin $- T - U - V -$ Carter or Bevin $- T - U -$ Carter

f 21 km

g



13C Topic 3 Networks and decision mathematics

Multiple-choice questions

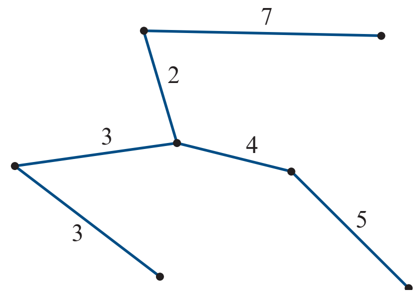
1 C	**2** E	**3** E	**4** D	**5** C
6 C	**7** D	**8** C	**9** D	**10** E

Short-answer questions

Simple familiar questions

1 $A - B$ or $C - B$ or $A - C$

2 **a**



b 24

3 **a** 1

b Sharon

c Leah – Brazil, Sharon – Portugal, Kris – Tibet, Sue – Zimbabwe, Kathy – Fiji

4 **a** Ahmet

b Ahmet – Canapes, Beryl – Starter, Cynthia – Desert, Dario – Main

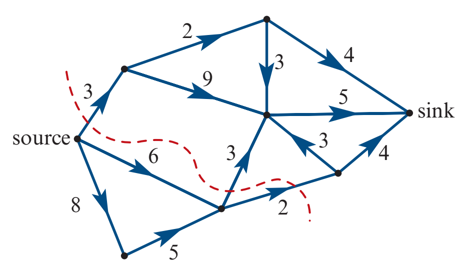
c **i** 10 h

ii 4 h

5 **a** **i** 22

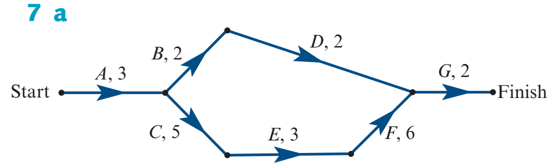
ii 11

b



6 **a** 9 **b** 3 **c** F and G

7 **a**



b 19 hours

8 **a** 17 **b** 7 **c** 1

d $A - C - G - J - L$

ISBN 978-1-108-46154-2

Cambridge University Press

9 a

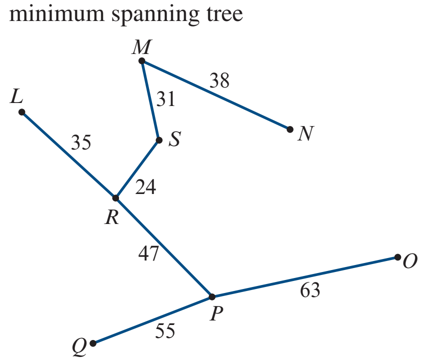
Activity	Immediate predecessors
A	–
B	*A*
C	–
D	–
E	*C*
F	*B*, *E*
G	*C*
H	*D*
I	*G*, *H*
J	*F*, *I*

b 24 days
c $C - E - F - J$

Extended-response questions

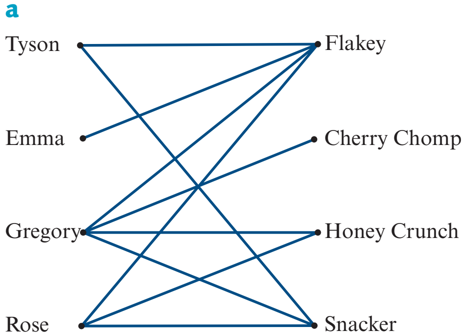
Complex familiar questions

1 a 112 km
b i minimum spanning tree
ii



iii 293 km

2 a



b Tyson
c Tyson – Snacker, Emma – Flakey, Gregory – Cherry Chomp, Rose – Honey Crunch

3 a $A - Z$, $B - W$, $C - X$, $D - Y$, or $A - Z$, $B - X$, $C - W$, $D - Y$
b $130

4 a 11 megalitres per day
b 33 megalitres

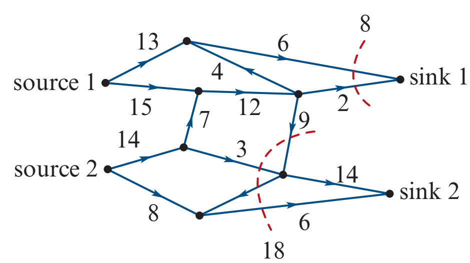
5 a The dummy shows that activity *I* has immediate predecessor *E*.
b

Activity	Duration (days)	EST	LFT	LST	Float
A	6	0	7	1	1
B	5	0	5	0	0
C	2	5	7	5	0
D	4	5	13	9	4
E	4	7	11	7	0
F	6	7	16	10	3
G	4	11	15	11	0
H	3	9	16	13	4
I	2	13	18	16	3
J	3	15	18	15	0
K	12	18	30	18	0

c $B - C - E - G - J - K$
d The critical path contains all of the activities that cannot be delayed or extended without affecting the overall completion time of the project.

Complex unfamiliar questions

6 sink 1 = 8, sink 2 = 18



7 700 kilolitres per minute for each outlet.

8 a 15 hours
b $B - C - F - I$
c *A* is not on the critical path. It already has slack time and reducing it further has no effect.
d $200

ISBN 978-1-108-46154-2

Cambridge University Press

Glossary

A

Activity (CPA) [p. 473] A task to be completed as part of a project. Activities are represented by the edges in the project diagram.

Activity network [p. 473] An activity network is a weighted directed graph that shows the required order of completion of the activities that make up a project. The weights indicate the durations of the activities they represent.

Adjacency matrix [p. 399] A square matrix showing the number of edges joining each pair of vertices in a graph.

Algorithm [p. 441] A step-by-step procedure for solving a particular problem that involves applying the same process repeatedly. Examples include Prim's algorithm and the Hungarian algorithm.

Allocation [p. 451] Allocation is the process of assigning a series of tasks to different members of a group in a way that enables the tasks to be completed for the minimum time or cost.

Annuity [p. 351] An annuity is a compound interest investment from which regular payments are made.

Arc [p. 219] The part of a circle between two given points on the circle. The length of the arc of a circle is given by $s = r\left(\frac{\theta}{180}\right)\pi$, where r is the radius of the circle and θ is the angle in degrees subtended by the arc at the centre of the circle.

Arithmetic sequences [p. 166] A sequence is arithmetic if it satisfies the recurrence relation: $t_{n+1} = t_n + d$ and a starting point usually t_1. Arithmetic sequences are used to model linear growth and linear decay situations. The rule for the nth term of an arithmetic sequence is: $t_n = t_1 + (n - 1)d = a + (n - 1)d$, where $a = t_1$ is the starting value.

B

Backward scanning [p. 485] Backward scanning is the process of determining the LST for each activity in a project activity network.

Balance [p. 324] The balance of a loan or investment is the amount owed or accrued after a period of time.

Bipartite graph [p. 446] A graph whose set of vertices can be split into two subsets, X and Y, in such a way that each edge of the graph joins a vertex in X and a vertex in Y.

Bivariate data [p. 2] Data in which each observation involves recording information about two variables for the same person or thing. An example would be data recording the height and weight of the children in a preschool.

Bridge [p. 394] An edge in a **connected graph** that, if removed, would leave the graph no longer connected.

C

Capacities (flow network) [p. 455] The weights of the directed edges in a flow network are called **capacities**. They give the maximum amount that can move between the two points in the flow network represented by these vertices in a particular time interval. This could be, for example,

ISBN 978-1-108-46154-2

Cambridge University Press

the maximum amount of water in litres per minute or the maximum number of cars per hour.

Categorical variable [p. 2] A variable used to represent characteristics of individuals, for example place of birth, house number. Categorical variables come in types, nominal and ordinal.

Causal relationship [p. 84] When a change in the **explanatory variable** leads to a change in the **response variable**, this is known as a causal relationship.

Centring [p. 124] If smoothing takes place over an even number of data values, the smoothed values do not align with an original data value. A second stage of smoothing is carried out to centre the smoothed values at an original data value.

Coefficient of determination (R^2) [p. 45] A coefficient which gives a measure of the predictive power of a regression line. It gives the percentage of variation in the RV that can be explained by the variation in the EV.

Complete graph [p. 394] A graph with edges connecting all pairs of vertices.

Compound interest [p. 199] Where the interest paid on a loan or investment is added to the principal and subsequent interest is calculated on the total.

Compounding period [p. 290] The compounding period is the time period for the calculation of interest for an investment or loan. Typical compounding periods are yearly, quarterly, monthly or daily.

Connected graph [p. 394] A **connected graph** is a graph that has no isolated vertices and no separate parts.

Continuous variable [p. 3] A variable representing a quantity that is measured rather than counted, for example the weights of people in kilograms.

Coordinated Universal Time (UTC) [p. 241] A measure of time used to regulate time across the world. Equivalent to **GMT**.

Correlation coefficient *r* [p. 36] A statistical measure of the strength of the linear association between two numerical variables.

Cost matrix [p. 448] A cost matrix is a table that contains the cost of allocating objects from one group, such as people, to objects from another group, such as tasks. The cost can be money, or other factors such as the time taken to complete the project.

Critical path [p. 483] The project path that has the longest completion time.

Critical path analysis [p. 483] A project planning method in which activity durations are known with certainty.

Cut [p. 457] A line dividing a directed (flow) graph into two parts in a way that separates all 'sinks' from their 'sources'.

Cut capacity [p. 457] The capacity of a cut is the sum of the capacities of the cuts passing through the cut that represents flow from the source to the sink. Edges that represent flow from the sink to the source do not contribute to the capacity of the cut.

Cycle (graphs) [p. 413] A **walk** with no repeated vertices that starts and ends at the same vertex. *See also* **circuit**.

Cycle (time series) [p. 115] Periodic movement in a time series but over a period greater than a year.

D

Degree of a vertex (deg(A)) [p. 391] The number of edges attached to the vertex. The degree of vertex A is written as deg(A).

Deseasonalise [p. 130] The process of removing seasonality in time series data.

Directed graph (digraph) [p. 395] A graph or network in which directions are associated with each of the edges.

Discrete variable [p. 3] A variable representing a quantity that is determined by counting, for example, the number of people waiting in a queue.

Dummy activity [p. 477] An artificial activity of zero time duration added to a project diagram to ensure that all predecessor activities are properly accounted for.

E

Earliest starting time (EST) [p. 484] The earliest time an activity in a project can be started.

ISBN 978-1-108-46154-2

Cambridge University Press

Edge [p. 391] A line joining one vertex in a graph or network to another vertex or itself (a loop).

Effective annual rate of interest [p. 303] Used to compare the interest paid on loans (or investments) with the same annual nominal interest rate r but with different compounding periods (daily, monthly, quarterly, annually, other).

Elements [p. 450] The numbers or symbols displayed in a matrix.

Equivalent graph [p. 405] *see* **isomorphic graphs**.

Eulerian trail [p. 416] A walk in a graph or network that includes every edge just once (but does not start and finish at the same vertex). To have an eulerian walk (but not an eulerian circuit), a network must be connected and have exactly two vertices of odd degree, with the remaining vertices having even degree.

Euler's formula [p. 407] The formula $v - e + f = 2$, which relates the number of vertices, edges and faces in a connected graph.

Explanatory variable [p. 4] When investigating associations in **bivariate data**, the explanatory variable (EV) is the variable used to explain or predict the value of the **response variable** (RV).

Extrapolation [p. 75] Using a mathematical model to make a prediction *outside* the range of data used to construct the model.

F

Face [p. 406] An area in a graph or network that can only be reached by crossing an edge. One such area is always the area surrounding a graph.

Flat-rate depreciation [p. 180] Depreciation where the value of an item is reduced by the same amount each year. Flat-rate depreciation is equivalent, but opposite, to simple interest.

Float (slack) time [p. 483] The amount of time available to complete a particular activity that does not increase the total time taken to complete the project.

Flow [p. 455] Flow is the transfer of material, such as water, gas or traffic through a directed network.

Flow network [p. 455] A **flow network** occurs where the directed edges of the graph represent the flow of material from one vertex to another. The weight of an edge of a flow network is called the **capacity** of that edge.

Forward scanning [p. 484] Forward scanning is the process of determining the EST for each activity in a project activity network.

G

Geometric decay [p. 188] When a recurrence rule involves multiplying by a factor less than one, the terms in the resulting sequence are said to decay geometrically.

Geometric growth [p. 188] When a recurrence rule involves multiplying by a factor greater than one, the terms in the resulting sequence are said to grow geometrically.

Geometric sequences [p. 187] A sequence is geometric if it satisfies the recurrence relation: $t_{n+1} = r \times t_n$ and a starting point usually t_1. Geometric sequences are used to model geometric growth and linear decay situations. The rule for the nth term of a geometric sequence is $t_n = r^{n-1}t_1 = r^{n-1}a$, where $a = t_1$ is the starting value.

Graph or network [pp. 391, 500] A collection of points called vertices and a set of connecting lines called edges.

Great circle [p. 223] A circle on a sphere whose plane passes through the centre of the sphere. The shortest distance between two points on a sphere is along an arc of the great circle passing through the two points. *See also* **small circle**.

Greenwich Mean Time (GMT) [p. 241] Equivalent to **UTC**, this is a measure of time centred around Greenwich, England and is used across the world.

H

Hamiltonian cycle [p. 420] A hamiltonian path that starts and finishes at the same vertex.

Hamiltonian graph [p. 420] A Hamiltonian graph is a graph that contains a **Hamiltonian cycle**.

ISBN 978-1-108-46154-2

Cambridge University Press

Hamiltonian path [p. 420] A path through a graph or network that passes through each vertex exactly once. It may or may not start and finish at the same vertex.

Hungarian algorithm [p. 448] An algorithm for solving allocation (assignment) problems.

I

Immediate predecessor [p. 473] An activity that must be completed immediately before another one can start.

Intercept (of a straight line) [p. 60] Where the regression line cuts across the y-axis.

Interest [p. 178] The amount of money paid (earned) for borrowing (lending) money over a period of time.

Interest rate [p. 115] The rate at which interest is charged or paid. Usually expressed as a percentage of the money owed or lent.

International Date Line [p. 242] An imaginary line through the Pacific Ocean that corresponds to 180° **longitude**.

Interpolation [p. 75] Using a regression line to make a prediction *within* the range of values of the explanatory variable.

Irregular (random) fluctuations [p. 117] Unpredictable fluctuations in a time series. Always present in any real world time series plot.

Isolated vertex [p. 393] A vertex that is not connected to any other vertex. Its degree is zero.

Isomorphic graphs [p. 405] Equivalent graphs. Graphs that have the same number of edges and vertices that are identically connected.

L

Latest start time (LST) [p. 487] The latest time an activity in a project can begin, without affecting the overall completion time for the project.

Latitude [p. 224] The angle or angular distance north or south of the equator.

Least squares method [p. 56] One way of finding the equation of a regression line. It minimises the sum of the squares of the residuals. It works best when there are no outliers.

Linear decay [p. 169] When a recurrence rule involves subtracting a fixed amount, the terms in the resulting sequence are said to decay linearly.

Linear growth [p. 169] When a recurrence rule involves adding a fixed amount, the terms in the resulting sequence are said to grow linearly.

Linear regression [p. 58] The process of fitting a straight line to bivariate data.

Longitude [p. 224] The angle or angular distance east or west of the **prime meridian**.

Loop [p. 391] An edge in a graph or network that joins a vertex to itself.

M

Matrix [p. 399] A rectangular array of numbers or symbols set out in rows and columns within square brackets (*pl*: matrices).

Maximum flow (graph) [p. 455] The capacity of the ‘minimum’ cut.

Mean ($\bar{x}$) [p. 60] The balance point of a data distribution. The mean is given by $\bar{x} = \frac{\sum x}{n}$, where $\sum x$ is the sum of the data values and n is the number of data values. Best used for symmetric distributions.

Median [p. 89] The median (M) is the middle value in a data distribution. It is the midpoint of a distribution dividing an ordered data set into two equal parts. Can be used for skewed or symmetric distributions.

Meridian [p. 224] Semi-great circles that pass through north and south poles.

Meridians of longitude [p. 224] Semi-great circles which pass through the north and south poles.

Minimum cut (graph) [p. 458] The cut through a graph or network with the minimum capacity.

ISBN 978-1-108-46154-2

Cambridge University Press

Minimum spanning tree [p. 441] The spanning tree of minimum length. For a given connected graph, there may be more than one minimum spanning tree.

Modelling [pp. 198, 285] Mathematical modelling is the use of a mathematical rule or formula to represent real-life situations.

Moving mean smoothing [p. 121] In three-moving mean smoothing, each original data value is replaced by the mean of itself and the value on either side. In five-moving mean smoothing, each original data value is replaced by the mean of itself and the two values on either side.

Multiple edge [p. 393] Where more than one edge connects the same two vertices in a graph.

N

Network [pp. 424, 526] A set of points called vertices and connecting lines called edges, enclosing and surrounded by areas called faces.

Nominal interest rate [p. 290] The annual interest rate for a loan or investment that assumes the compounding period is 1 year. If the compounding period is less than a year, for example monthly, the actual or **effective interest rate** will be greater than r.

Numerical variable [p. 2] A variable used to represent quantities that are counted or measured. For example, the number of people in a queue, the heights of these people in cm. Numerical variables come in types: discrete and continuous.

O

Outliers [pp. 26, 117] Data values that appear to stand out from the main body of a data set.

P

Parallels of latitude [p. 224] Small circles whose planes are parallel to that of the equator.

Path [p. 412] A **walk** with no repeated vertices. *See also* **trail**.

Percentage frequency [p. 11] Frequency expressed as a percentage.

Perpetuity [p. 376] An investment where an equal amount is paid out on a regular basis forever.

Planar graph [p. 405] A graph that can be drawn in such a way that no two edges intersect, except at the vertices.

Precedence table [p. 474] A table that records the activities of a project, their immediate predecessors and often the duration of each activity.

Prim's algorithm [p. 441] An algorithm for determining a minimum spanning tree in a connected graph.

Prime meridian [p. 225] The meridian located at 0° which passes through Greenwich, England.

Principal (*P*) [p. 178] The initial amount borrowed, lent or invested.

R

Radius [p. 219] The distance from the centre to any point on the circle (sphere). Half the diameter.

Recurrence relation [pp. 167, 281] A relation that enables the value of the next term in a sequence to be obtained by one or more current terms. Examples include 'to find the next term, add two to the current term' and 'to find the next term, multiply the current term by three and subtract five'.

Reducing-balance depreciation [p. 201] When the value of an item is reduced by the same percentage each year. Reducing-balance depreciation is equivalent to, but opposite to, compound interest.

Reducing-balance loan [p. 324] A loan that attracts compound interest, but where regular repayments are also made. In most instances the repayments are calculated so that the amount of the loan and the interest are eventually repaid in full.

Reseasonalise [p. 130] The process of converting seasonal data back into its original form.

ISBN 978-1-108-46154-2

Residual [p. 59] The vertical distance from a data point to a straight line fitted to a scatterplot is called a residual:

residual = actual value − predicted value

Residuals are sometimes called *errors of prediction*.

Residual plot [p. 68] A plot of the residuals against the explanatory variable. Residual plots can be used to investigate the linearity assumption.

Response variable [p. 4] The variable of primary interest in a statistical investigation.

S

Scatterplot [p. 20] A statistical graph used for displaying bivariate data. Data pairs are represented by points on a coordinate plane, the EV is plotted on the horizontal axis and the RV is plotted on the vertical axis.

Seasonal indices [p. 130] Indices calculated when the data shows seasonal variation. Seasonal indices quantify seasonal variation. A seasonal index is defined by the formula:

$$\text{seasonal index} = \frac{\text{value for season}}{\text{seasonal average}}$$

For seasonal indices, the average is 1 (or 100%).

Seasonality [p. 115] The tendency for values in the time series to follow a seasonal pattern, increasing or decreasing predictably according to time periods such as time of day, day of the week, month, or quarter.

Sequence [p. 163] A list of numbers or symbols written down in succession, for example 5, 15, 25, . . .

Shortest path [p. 424] The path through a graph or network with minimum length.

Simple graph [p. 394] A graph with no loops or multiple edges.

Simple interest [pp. 178, 285] Interest that is calculated for an agreed period and paid only on the original amount invested or borrowed.

Sink [p. 455] *See* **sink and source**.

Sink and source [p. 455] In a flow network, a **source** generates flow while a **sink** absorbs the flow.

Slope (of a straight line) [p. 60] The slope of a straight line is defined to be: slope $= \frac{rise}{run}$.

The slope is also known as the gradient.

Small circle [p. 223] Any circle on a sphere whose plane does not pass through the centre of the sphere. *See also* **great circle**.

Smoothing [p. 121] A technique used to eliminate some of the variation in a time series plot so that features such as seasonality or trend are more easily identified.

Source [p. 455] *See* **sink and source**.

Spanning tree [p. 439] A subgraph of a connected graph that contains all the vertices of the original graph, but without any multiple edges, circuits or loops.

Standard deviation (*s*) [p. 60] A summary statistic that measures the spread of the data values around the mean. The standard deviation is given by $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$

Strength of a linear relationship [p. 36] Classified as weak, moderate or strong. Determined by observing the degree of scatter in a scatterplot or calculating a correlation coefficient.

Structural change (time series) [p. 116] A sudden change in the established pattern of a time series plot.

Subgraph [p. 395] Part of a graph that is also a graph in its own right.

T

Time series data [p. 107] A collection of data values along with the times (in order) at which they were recorded.

Time series plot [p. 107] A line graph where the values of the response variable are plotted in time order.

Time zone [p. 241] A region of the Earth that has a uniform standard time or local time. There are 24 time zones in total.

ISBN 978-1-108-46154-2

Trail [p. 412] A walk with no repeated edges. *See also* **path**.

Tree [p. 439] A **connected graph** with no circuits, multiple edges or loops.

Trend [p. 113] The tendency for values in the time series to generally increase or decrease over a significant period of time.

Trend line forecasting [p. 143] Using a line fitted to an increasing or decreasing time series to predict future values.

Two-way frequency table [p. 8] A frequency table in which subjects are classified according to two categorical variables. Two-way frequency tables are commonly used to investigate the associations between two categorical variables.

U

Unit-cost depreciation [p. 180] Depreciation based on how many units have been produced or consumed by the object being depreciated. For example, a machine filling bottles of drink may be depreciated by 0.001 cents per bottle it fills.

V

Variable [p. 2] A symbol used to represent a number or group of numbers.

Vertex (graph) [p. 391] The points in a graph or network (*pl* vertices).

W

Walk [p. 411] Any continuous sequence of edges, linking successive vertices, that connects two different vertices in a graph. *See also* **trail** and **path**.

Weighted graph [p. 424] A graph in which a number representing the size of some quantity is associated with each edge. These numbers are called weights.

ISBN 978-1-108-46154-2

Cambridge University Press