

Section A**Specific instructions for Section A**

Section A consists of 13 questions.

Answer **all** questions in this section.

A correct answer scores 1, an incorrect answer scores 0. No mark will be given for a question if two or more letters are circled for that question. Marks will not be deducted for incorrect answers and you should attempt every question in this section.

Core : Data analysis**Question 1**

The stem plot shows the heights (in cm) of 20 football players.

Stem	Leaf
15	0
16	
17	2 3 5
18	0 1 2 4 5 7 9
19	1 2 3 4 6 7 8 9 9

The distribution of data could be described as

- A negatively skewed
- B negatively skewed with one outlier
- C positively skewed
- D positively skewed with one outlier
- E symmetric

Question 2

The mean Study Score for Further Mathematics is 30 with a standard deviation of 7. If the total number of students in 2004 doing Further Mathematics is 22 000 then the number of students who get a study score of 44 or higher is likely to be

- A 33 students
- B 550 students
- C 1100 students
- D 3520 students
- E 7480 students

Question 3

Which of the following is an example of continuous numerical data?

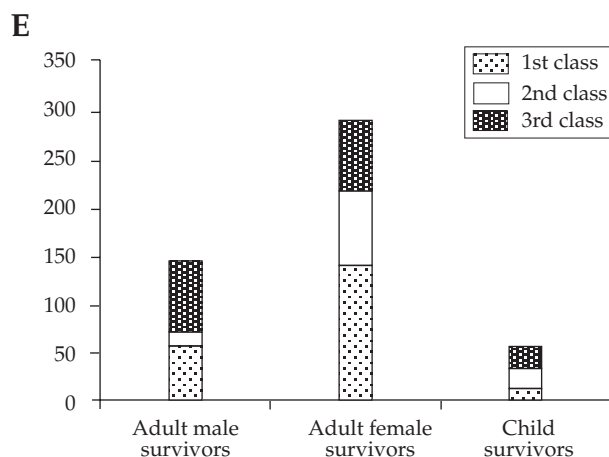
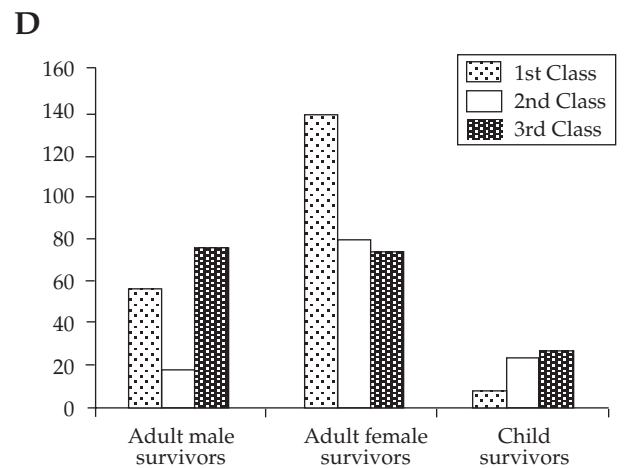
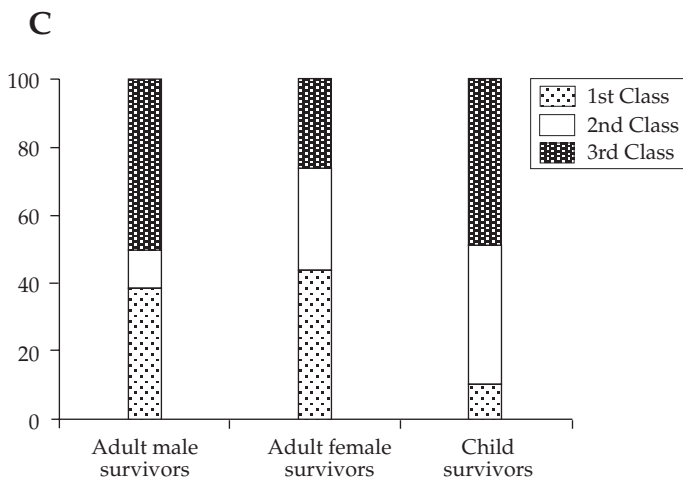
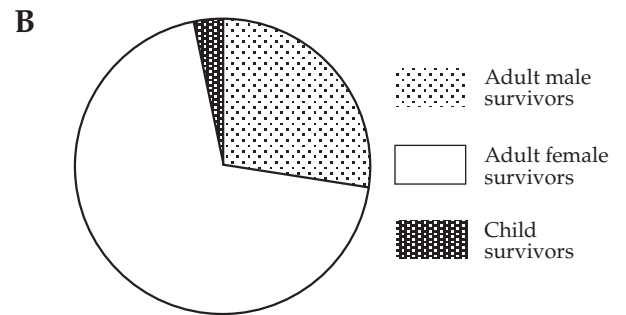
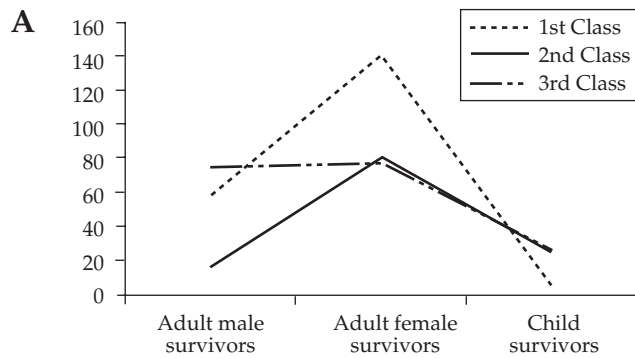
- A shoe size
- B number of runs made by a cricket player
- C your favourite school year level in the past six years
- D Labor/Liberal preference of first 100 people surveyed at a shopping mall
- E speed of a car captured by a speed camera

Question 4

The following is the data of the number of passenger survivals and their class aboard the Titanic when it sank on 15th April 1912.

	Adult male survivors	Adult female survivors	Child survivors
1st Class	57	140	6
2nd Class	14	80	24
3rd Class	75	76	27
Total Survivors	146	296	57

The most appropriate visual display for a two-way frequency table to investigate the above data for a relationship between survival by gender and class is likely to be



Question 5

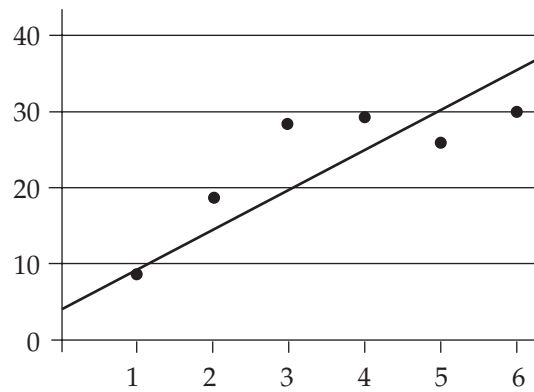
A set of data comparing blood alcohol level (BAL) and driver control ability is found to have a coefficient of determination of 0.49. A competent driver would score high in driver control ability.

The Pearson's correlation coefficient for this data set would be closest to

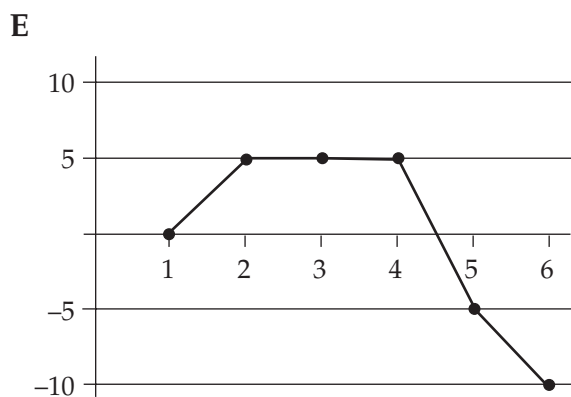
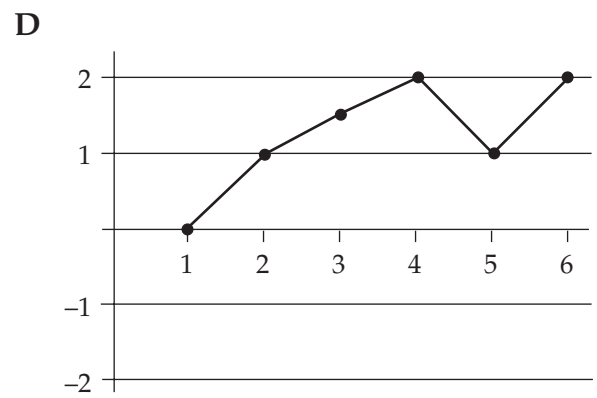
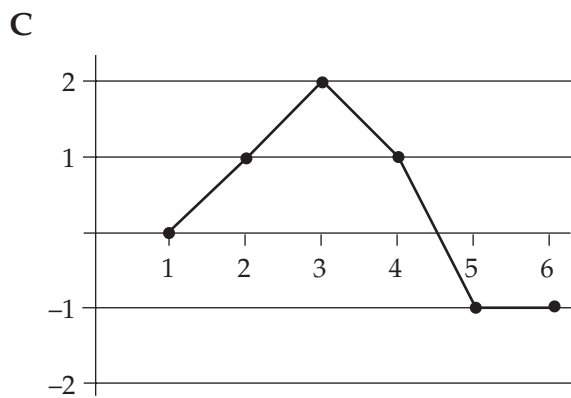
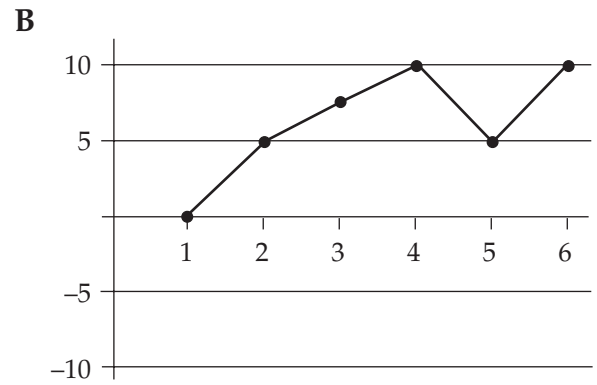
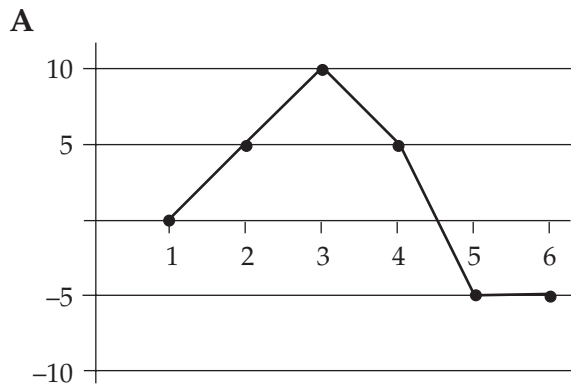
- A 0.51
- B -0.25
- C 0.25
- D -0.7
- E 0.7

Question 6

A least squares regression line is fitted to the 6 points shown in the following figure.



Which of the following looks most similar to the plot of residuals?



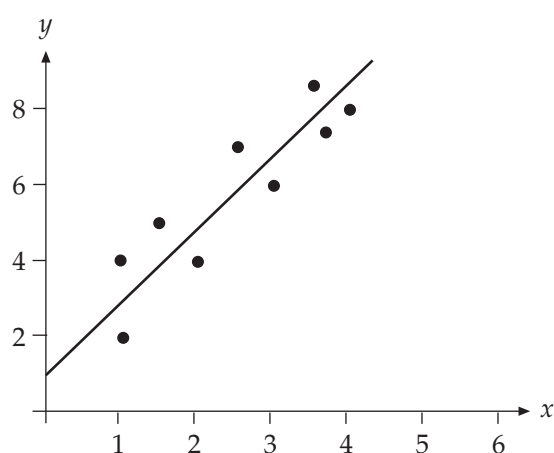
Question 7

A 3-median regression fit yielded the equation $y = -2.5 + 3.2x$. The value of y when $x = 3$ is

- A -4.3
- B 16.5
- C 18.9
- D 0.7
- E 7.1

Question 8

Consider the following figure.



In the least squares regression, the line goes through the point $(3, 7)$. The best estimate of the regression equation is

- A $y = 2 + 2x$
- B $y = 1 + x$
- C $y = 4 + 3x$
- D $y = 1 + 2x$
- E $y = 4x - 1$

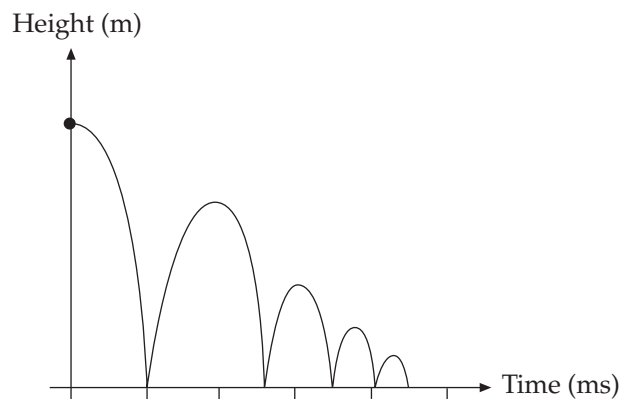
Question 9

The correlation between two variables, x and y , is -0.75 . Which of the following statements is true?

- A As x increases, y tends to increase.
- B As x increases, y tends to decrease.
- C There is a weak relationship between x and y .
- D An increase in x causes y to decrease.
- E An increase in x causes y to increase.

Question 10

The height to which a ball bounces is plotted on this graph.



The trend is probably

- A seasonal only
- B seasonal with a trend
- C cyclical with a trend
- D cyclical only
- E random

Question 11

The table shows the number of photos taken, per second, of celebrities at award ceremonies.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Number of photos '000s	21	23	22	27	32	31	36	36	36	37	35	39

The last two points in the 3-median smoothed trend line are

- A 35 and 39
- B 22 and 22.5
- C 22 and 24
- D 36 and 37
- E 36.25 and 36.5

Question 12

When performing a seasonal smoothing of data, the sum of seasonal indices is equal to

- A the number of seasons
- B the number of years
- C the number of cycles
- D 12
- E 4

Question 13

In the following set of data, two values were accidentally omitted.

5 6 7 7 8 8 9 10

If the original set of ten points gave a mean of 7.8 and a median of 7.5, then the two missing values are

- A 9 and 9
- B 7 and 8
- C 7 and 11
- D 8 and 10
- E 7.8 and 7.5