

MATHEMATICS

Stage 5

Statistics and Probability, Data Analysis

Bivariate Data

Box Plots and 5-Number Summary

Standard Deviation

Statistics and Probability, Probability

Multi-Stage Events

Relative Frequency and Venn Diagrams

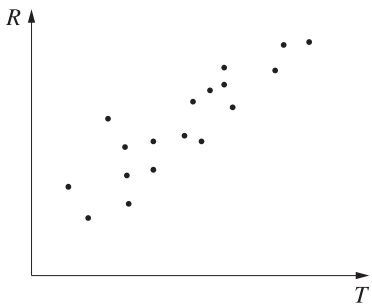
Teacher: Kyall Duncan

Exam Equivalent Time: 109.5 minutes *(based on allocation ~1.25 min/mark)*

Questions

1. C Bivariate Data (014201) Core B HSC

A scatterplot is shown.

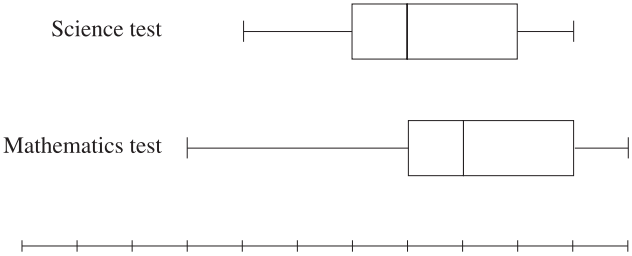


Which of the following best describes the correlation between *R* and *T*?

- A. Positive
- B. Negative
- C. Positively skewed
- D. Negatively skewed

2. C Box Plots and 5-Number Summary (014192) Core A HSC

The marks for a Science test and a Mathematics test are presented in box-and-whisker plots.



Which measure must be the same for both tests?

- A. Mean
- B. Range
- C. Median
- D. Interquartile range

3. C Bivariate Data (015107) Core B HSC

Which of the following would be most likely to have a positive correlation?

- A. The population of a town and the number of schools in that town
- B. The price of petrol per litre and the number of litres of petrol sold
- C. The hours training for a marathon and the time taken to complete the marathon
- D. The number of dogs per household and the number of televisions per household

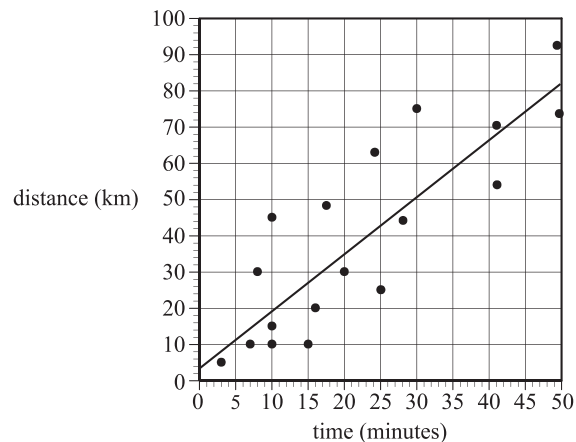
4. C Multi-Stage Events (029225) Core A HSC

The probability of winning a game is $\frac{7}{10}$.

Which expression represents the probability of winning two consecutive games?

- A. $\frac{7}{10} \times \frac{6}{9}$
- B. $\frac{7}{10} \times \frac{6}{10}$
- C. $\frac{7}{10} \times \frac{7}{9}$
- D. $\frac{7}{10} \times \frac{7}{10}$

5. **B** Bivariate Data (262967) **Core B**



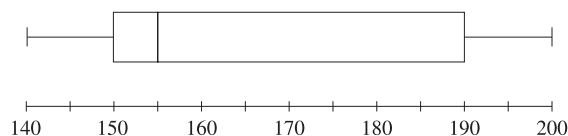
A line of best fit has been fitted to the scatterplot above to enable *distance*, in kilometres, to be predicted from *time*, in minutes.

The equation of this line is closest to

- A. $distance = 3.5 + 1.6 \times time$
- B. $time = 3.5 + 1.6 \times distance$
- C. $distance = 1.6 + 3.5 \times time$
- D. $time = 1.8 + 3.5 \times distance$

6. **B** Box Plots and 5-Number Summary (002595) **Core A** **HSC**

A set of data is displayed in this box-and-whisker plot.



Which of the following best describes this set of data?

- A. Symmetrical
- B. Positively skewed
- C. Negatively skewed
- D. Normally distributed

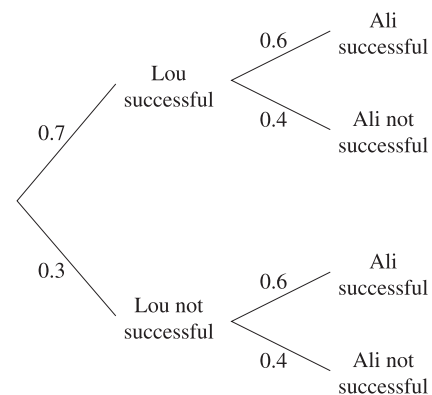
7. **B+** Bivariate Data (002211) **Core B** **HSC**

Which of the following relationships would most likely show a negative correlation?

- A. The population of a town and the number of hospitals in that town.
- B. The hours spent training for a race and the time taken to complete the race.
- C. The price per litre of petrol and the number of people riding bicycles to work.
- D. The number of pets per household and the number of computers per household.

8. **B+** Multi-Stage Events (004350) **Core A** **HSC**

Lou and Ali are on a fitness program for one month. The probability that Lou will finish the program successfully is 0.7 while the probability that Ali will finish successfully is 0.6. The probability tree shows this information

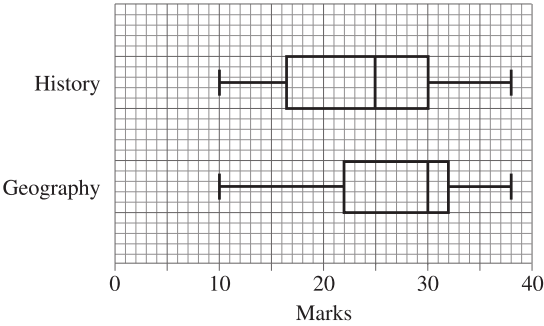


What is the probability that only one of them will be successful ?

- A. 0.18
- B. 0.28
- C. 0.42
- D. 0.46

9. **A** Box Plots and 5-Number Summary (060352) **Core A** **HSC**

The box-and-whisker plots show the results of a History test and a Geography test.



In History, 112 students completed the test. The number of students who scored above 30 marks was the same for the History test and the Geography test.
How many students completed the Geography test?

- A. 8
- B. 50
- C. 56
- D. 112

10. **A** Multi-Stage Events (010151) **Core A** **HSC**

In Mathsville, there are on average eight rainy days in October.

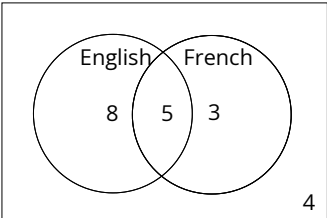
Which expression could be used to find a value for the probability that it will rain on two consecutive days in October in Mathsville?

- A. $\frac{8}{31} \times \frac{7}{30}$
- B. $\frac{8}{31} \times \frac{7}{31}$
- C. $\frac{8}{31} \times \frac{8}{30}$
- D. $\frac{8}{31} \times \frac{8}{31}$

11. **C** Relative Frequency and Venn Diagrams (251080) **Path B**

A group of 20 museum visitors were surveyed about what languages they could speak fluently.

They were surveyed on whether they could speak English or French and the results were recorded in the Venn diagram below.



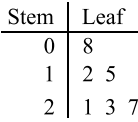
Record this information in the partially completed 2-way table below. (2 marks)

	English	No English
French		
No French	8	

12. **C** Standard Deviation (262472) **Core A**

In a cricket test match, a scorebook recorded the number of runs scored by England's top six batsman.

The scores are summarised in the Stem and Leaf plot below.



Determine the standard deviation of the scores, giving your answer correct to one decimal place. (2 marks)

.....

.....

.....

.....

13. **c** Relative Frequency and Venn Diagrams (250997) **Path B**

A group of 125 people were asked if they wear a watch or not.
This table shows the results.

	Wears Watch	No Watch	Total
Men	25	35	60
Women	15	50	65
Total	40	85	125

A man was selected at random.
What is the exact probability that he wears a watch? *(2 marks)*

.....

.....

.....

.....

14. **c** Standard Deviation (262468) **Core A**

Ms Arnott has seven students in her Ethics class. The results of the most recent exam, completed by the whole class, is summarised in the Stem and Leaf plot below.

Stem	Leaf
5	6
6	3 9
7	2 7 8
8	4

5 | 6 means 56

Determine the standard deviation of the exam results, giving your answer correct to one decimal place.
(2 marks)

.....

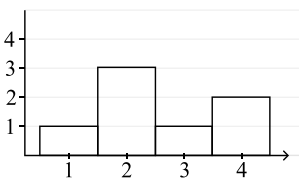
.....

.....

.....

15. **c** Standard Deviation (262529) **Core A**

Seven 40-year old males are asked how many children they have in a survey.
The results are summarised in the histogram drawn below.



i. Determine the mean number of children of the group surveyed, giving your answer to two decimal places. *(1 mark)*

.....

.....

ii. Calculate the standard deviation of this data set, giving your answer to two decimal places. *(1 mark)*

.....

.....

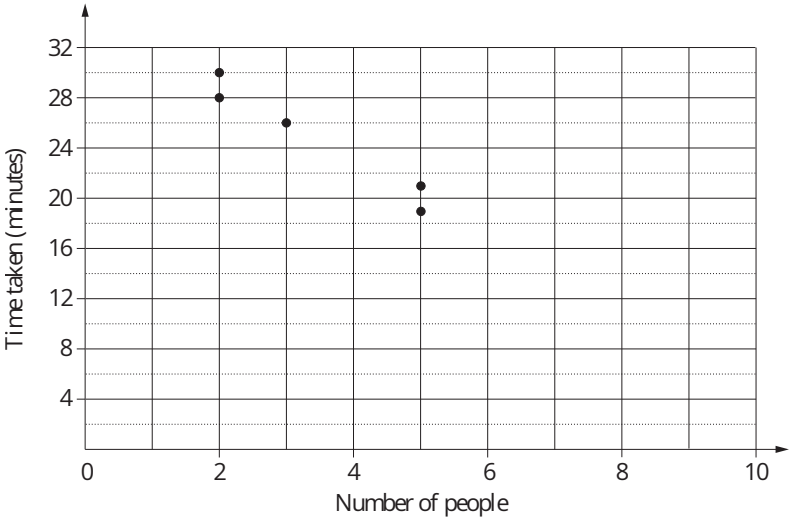
16. **C** Bivariate Data (262725) **Core B**

People are placed into groups to complete a puzzle. There are 9 different groups.

The table shows the number of people in each group and the amount of time, in minutes, for each group to complete the puzzle.

Number of people	2	2	3	5	5	7	7	7	8
Time taken (min)	28	30	26	19	21	12	13	11	8

a. Complete the scatterplot by adding the last four points from the table. *(1 mark)*



b. Add a line of best fit by eye to the graph in part (a). *(1 mark)*

c. The graph in part (a) shows the association between the time to complete the puzzle and the number of people in the group.

Identify the form (linear or non-linear), the direction and the strength of the association. *(2 marks)*

.....

.....

.....

17. **B** Box Plots and 5-Number Summary (261221) **Core A**

Le Bron plays basketball and records the number of free throws he shoots in the first nine games of the season.

12, 8, 4, 15, 13, 7, 8, 3, 10

Using a 5-number summary of this dataset, or otherwise, determine

i. the median *(1 mark)*

.....

.....

ii. the interquartile range *(2 marks)*

.....

.....

.....

.....

.....

.....

18. **B** Multi-Stage Events (020615) **Core A** **HSC**

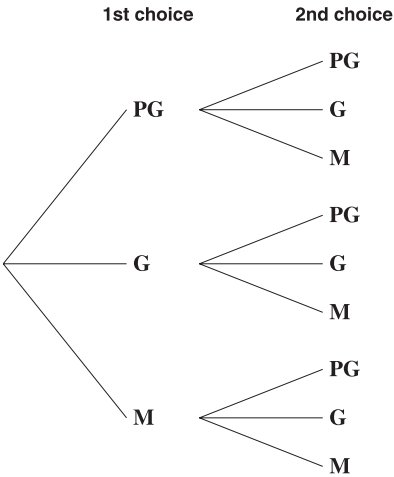
In a stack of 10 DVDs, there are 5 rated PG, 3 rated G and 2 rated M.

i. A DVD is selected at random. What is the probability that it is rated M? *(1 mark)*

.....

.....

Grant chooses two DVDs at random from the stack. Copy or trace the tree diagram into your writing booklet.



ii. Complete the tree diagram by writing the correct probability on each branch. *(2 marks)*

iii. Calculate the probability that Grant chooses two DVDs with the same rating. *(2 marks)*

.....

.....

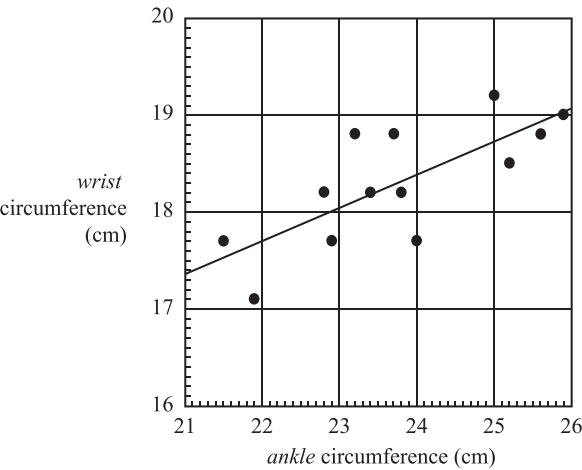
.....

.....

19. **B** Bivariate Data (263015) **Core B**

The scatterplot below shows the *wrist* circumference and *ankle* circumference, both in centimetres, of 13 people.

A line of best fit been drawn with ankle circumference as the independent variable.



i. If a person has a wrist circumference of 18.5 centimetres, estimate the ankle circumference that is predicted by the line of best fit. *(1 mark)*

.....

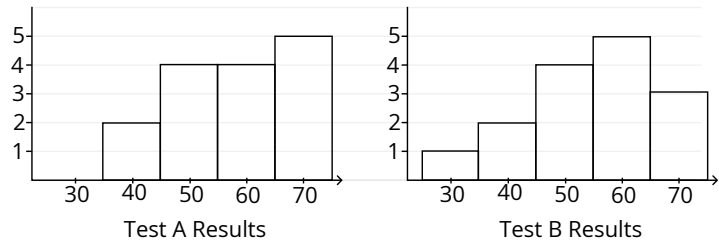
ii. Explain why the *y*-intercept of this graph has no meaning in this context. *(1 mark)*

.....

.....

20. **B** Standard Deviation (262487) **Core A**

Mr Ralph has a History class that has completed two separate exams during the term. The results of each exam are summarised in the histograms below.



Without using calculations, explain which test will have the highest standard deviation. (2 marks)

.....

.....

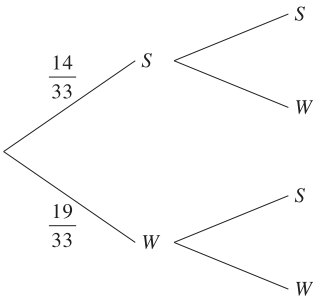
.....

.....

21. **B+** Multi-Stage Events (003169) **Core A** **HSC**

A box contains 33 scarves made from two different fabrics. There are 14 scarves made from silk (*S*) and 19 made from wool (*W*). Two girls each select, at random, a scarf to wear from the box.

i. Complete the probability tree diagram below. (2 marks)



ii. Calculate the probability that the two scarves selected are made from silk. (1 mark)

.....

.....

iii. Calculate the probability that the two scarves selected are made from different fabrics. (2 marks)

.....

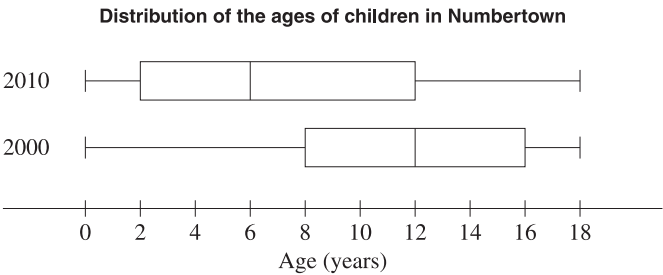
.....

.....

.....

22. **B+** Box Plots and 5-Number Summary (005376) **Core A** **HSC**

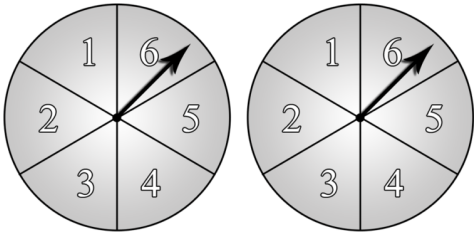
The graphs show the distribution of the ages of children in Numbertown in 2000 and 2010.



- i. In 2000 there were 1750 children aged 0–18 years.
How many children were aged 12–18 years in 2000? *(1 mark)*
-
- ii. The number of children aged 12–18 years is the same in both 2000 and 2010.
How many children aged 0–12 years are there in 2010? *(1 mark)*
-
- iii. Identify TWO changes in the distribution of ages between 2000 and 2010. In your answer, refer to measures of location or spread or the shape of the distributions. *(2 marks)*
-
-
-
-
-
- iv. What would be ONE possible implication for government planning, as a consequence of this change in the distribution of ages? *(1 mark)*
-
-

23. **B+** Multi-Stage Events (250900) **Core A**

Jon spins each pointer 50 times.



Each time he added the numbers that the pointers landed on.
His results are shown below.

Sum of numbers	2	3	4	5	6	7	8	9	10	11	12	Total
Number of spins	1	2	3	6	8	9	7	5	4	4	1	50

What percentage of the spins resulted in a sum of 9? *(2 marks)*

.....

.....

.....

.....

24. **B+** Standard Deviation (262338) **Core A**

The ages of nine students were recorded in the table below.

Ages		
12	11	16
14	16	15
14	15	14

i. What is the standard deviation, correct to two decimal places? *(2 marks)*

.....

.....

.....

.....

ii. Briefly explain what is meant by the term *standard deviation*. *(1 mark)*

.....

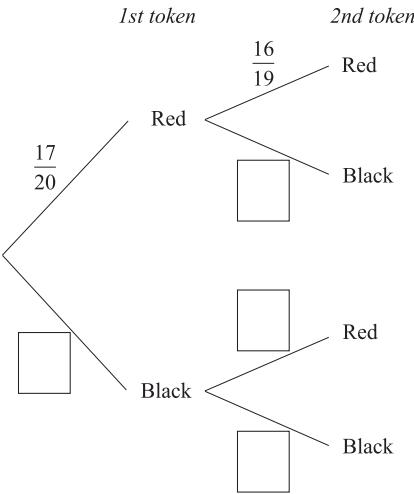
.....

.....

25. **B+** Multi-Stage Events (250979) **Path B**

A game consists of two tokens being drawn at random from a barrel containing 20 tokens. There are 17 red tokens and 3 black tokens. The player keeps the two tokens drawn.

i. Complete the probability tree by writing the missing probabilities in the boxes. *(2 marks)*



ii. What is the probability that a player draws at least one red token? Give your answer in exact form. *(2 marks)*

.....

.....

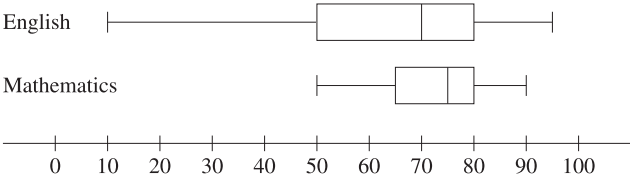
.....

.....

.....

26. **A** Box Plots and 5-Number Summary (261404) **Core A**

The test results in English and Mathematics for a class were recorded and displayed in the box plots.



i. What is the interquartile range for English? *(1 mark)*

.....

ii. Compare and contrast the two data sets by referring to the skewness of the distributions and the measures of their centres and spread. *(3 marks)*

.....
.....
.....
.....
.....
.....

27. **A** Multi-Stage Events (004531) **Path B** **HSC**

In each of three raffles, 100 tickets are sold and one prize is awarded.

Mary buys two tickets in one raffle. Jane buys one ticket in each of the other two raffles.

Determine who has the better chance of winning at least one prize. Justify your response using probability calculations. *(4 marks)*

.....
.....
.....
.....
.....
.....
.....
.....

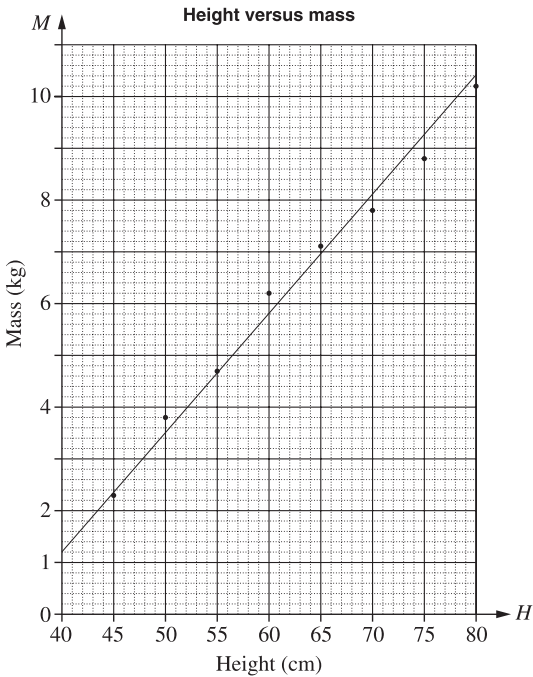
.....

28. **A** Bivariate Data (004584) **Core B** **HSC**

The height and mass of a child are measured and recorded over its first two years.

Height (cm), H	45	50	55	60	65	70	75	80
Mass (kg), M	2.3	3.8	4.7	6.2	7.1	7.8	8.8	10.2

This information is displayed in a scatter graph.



i. Describe the correlation between the height and mass of this child, as shown in the graph. (1 mark)

.....

.....

ii. A line of best fit has been drawn on the graph.

Find the equation of this line. (2 marks)

.....

.....

.....

.....

29. **A** Standard Deviation (262573) **Core A**

Dataset 1 has mean $\bar{x}_1$ and standard deviation σ_1 .

Dataset 2 has mean $\bar{x}_2$ and standard deviation σ_2 .

Consider the following statement: If $\bar{x}_1 < \bar{x}_2$, then $\sigma_1 < \sigma_2$.

Is this statement correct? Explain your answer. (2 marks)

.....

.....

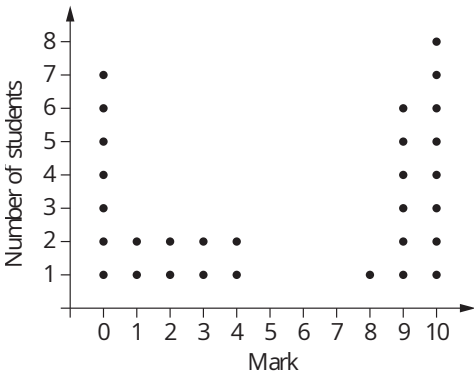
.....

.....

30. **A** Standard Deviation (129518) **Core A**

All the students in a class of 30 did a test.

The marks, out of 10, are shown in the dot plot.



i. Find the median test mark. (1 mark)

.....

ii. The mean test mark is 5.4. The standard deviation of the test marks is 4.22.

Using the dot plot, calculate the percentage of the marks which lie within one standard deviation of the mean. (2 marks)

.....

.....

.....

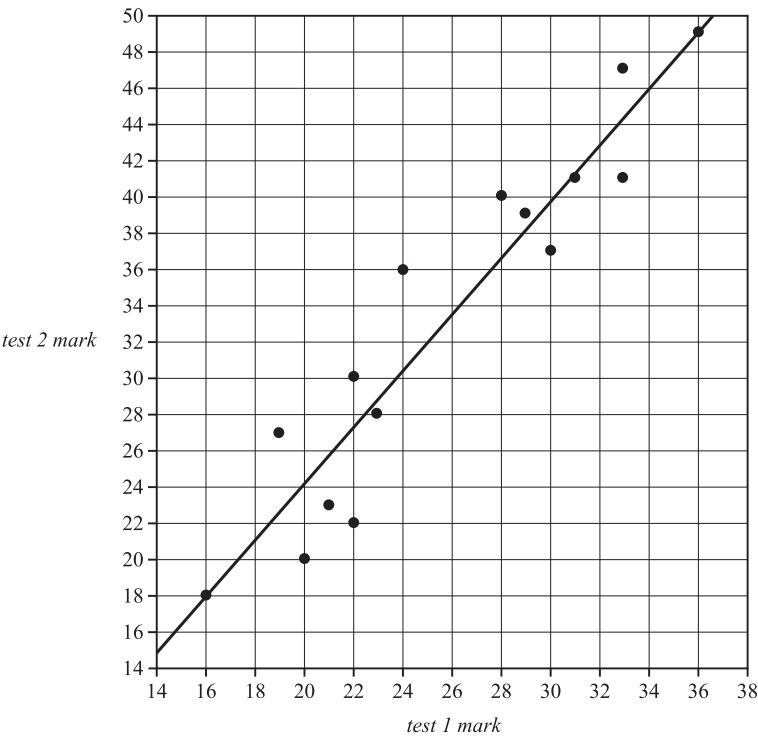
.....

31. **A** Bivariate Data (263106) **Core B**

A teacher analysed the class marks of 15 students who sat two tests.

The *test 1 mark* and *test 2 mark*, all whole number values, are shown in the scatterplot below.

A line of best fit has been fitted to the scatterplot.



i. If a student scored 34 in the first test, what is their expected mark in the second test. (1 mark)

.....

ii. The line of best fit shows the predicted *test 2 mark* for each student based on their *test 1 mark*.

Determine the number of students whose actual *test 2 mark* was within two marks of that predicted.

(1 mark)

.....

.....