

Please write clearly in block capitals.

Centre number

--	--	--	--	--

Candidate number

--	--	--	--

Surname

Forename(s)

Candidate signature

I declare this is my own work.

GCSE STATISTICS

H

Higher Tier Paper 2

Friday 13 June 2025

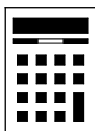
Afternoon

Time allowed: 1 hour 45 minutes

Materials

For this paper you must have:

- a calculator
- mathematical instruments.



Instructions

- Use black ink or black ball-point pen. Draw diagrams in pencil.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- You must answer the questions in the spaces provided. Do not write outside the box around each page or on blank pages.
- If you need extra space for your answer(s), use the lined pages at the end of this book. Write the question number against your answer(s).
- Do all rough work in this book. Cross out any work you do not want to be marked.

Information

- The marks for questions are shown in brackets.
- The maximum mark for this paper is 80.
- You may ask for more answer paper and graph paper. These must be tagged securely to this answer booklet.

For Examiner's Use	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
TOTAL	



J U N 2 5 8 3 8 2 2 H 0 1

Answer **all** questions in the spaces provided.

- 1** What percentage of a large data set does the interquartile range represent?

[1 mark]

Answer _____ %

1

- 2** In 2024 the birth rate of a population was 2.5 and there were 80 births.

Calculate the total population in 2024.

Use the formula,

$$\text{birth rate} = \frac{\text{number of births}}{\text{total population}} \times 1000$$

[2 marks]

Answer _____

2



- 3** 500 people take part in a race.
13 of the people do not finish the race.
Calculate the risk of not finishing the race.

[1 mark]

Answer _____

1

- 4** Andy picks, at random, one of his 200 books.
He has,
40 children's books
and
30 books with red covers.

The probability that he picks a children's book and the probability that he picks a book with a red cover are independent.

Calculate the probability that he picks a children's book with a red cover.

[2 marks]

Answer _____

2

Turn over for the next question**Turn over ►**

5 Thea lives on St Mary's, an island in the Isles of Scilly.

5 (a) Thea is investigating the ages and jobs of the people who live on St Mary's.

State the population of Thea's investigation.

[1 mark]

5 (b) (i) Thea gathers information using a questionnaire.

Thea includes this question,

'How old are you?'

Give a reason why this question may result in Thea getting non-responses.

[1 mark]

5 (b) (ii) Suggest **one** way Thea could reduce the number of non-responses to this question.

[1 mark]



Thea considers two different sampling methods.

Method A

- Stand in the main town on Tuesday morning.
- Ask the first 50 people to complete the questionnaire.

Method B

- Give each of the 1700 people who live on St Mary's a unique number.
- Obtain 50 unique, random numbers between 1 and 1700
- Ask the people whose random numbers come up to complete the questionnaire.

Name each method and give one disadvantage of each sampling method.

State which method Thea should use and give an advantage of that method.

[5 marks]

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. There are no margins, text, or other markings on the paper.

Turn over ►



The frequency polygon represents information about the salaries of 25 people who live in the main town on St Mary's, **Hugh Town**.



- 5 (d) Calculate an estimate for the mean salary of the 25 people living in Hugh Town.

[3 marks]

Answer £ _____



- 5 (e)** The table shows data about the salaries of people who live in **Old Town**, another town on St Mary's.

Salary, x (£ thousands)	Frequency
$0 \leq x < 10$	3
$10 \leq x < 20$	12
$20 \leq x < 30$	9
$30 \leq x < 40$	1
$40 \leq x < 50$	0

On the grid opposite, draw a frequency polygon to represent these data.

[3 marks]

- 5 (f)** Without calculations, write down **two** comparisons of the salary distributions of people who live in Hugh Town and people who live in Old Town.

[2 marks]

Comparison 1 _____

Comparison 2 _____

Question 5 continues on the next page

Turn over ►



- 5 (g)** Thea finds data on the St Mary's Council website about salaries of people living on St Mary's.

The Council's data are different to Thea's data.

Which data are more reliable?

Tick (✓) a box.

Thea's data

☐

Council's data

☐

Give a reason for your answer.

[1 mark]

17



6 A company wants to estimate the number of people who illegally download software.

It uses the following method to collect data from 200 people.

- Each person secretly flips a coin.
- If they get heads, they answer 'Yes'.
- If they get tails, they answer the question truthfully,
'Have you illegally downloaded software?'

6 (a) Why does the company use this method?

[1 mark]

6 (b) 120 of the 200 people answered yes.

Ada says,

"More than 30% of people have illegally downloaded software."

Is Ada likely to be correct?

Tick (✓) a box.

Yes

☐

No

☐

Show working to support your answer.

[2 marks]

3



7 Both Redcrest School and Hazelwood School want to investigate how much homework students complete.

7 (a) Redcrest uses an anonymous survey to collect the following information for every student each week.

- Hours spent completing homework.
- Number of pieces of homework set.
- Number of pieces of homework completed.

Circle the type of data **not** collected by Redcrest.

[1 mark]

Quantitative

Primary

Qualitative

Multivariate

7 (b) (i) Which data collection method samples every member of the population?

[1 mark]

Answer _____

7 (b) (ii) Give one advantage and one disadvantage of collecting data from every student.

[2 marks]

Advantage _____

Disadvantage _____



7 (c) A sample of 60 students is selected from Hazelwood.

7 (c) (i) What is the sample frame for this sample?

[1 mark]

7 (c) (ii) Hazelwood has a total of 600 students.

Hazelwood uses systematic sampling to select the students from a list that is ordered alphabetically by surname followed by forename.

Why is it unlikely that two students with the same surname will **both** be in the sample?

[1 mark]

Question 7 continues on the next page

Turn over ►



7 (d) A teacher at Hazelwood is investigating the hypothesis,

‘When all homework is set on paper, students complete more homework.’

She uses two groups of students.

Group A: 30 students with homework set on paper.

Group B: a control group of 30 students with homework set online, as usual.

7 (d) (i) Give **one** reason to use a control group.

[1 mark]

7 (d) (ii) The teacher finds that over ten weeks,

Group A complete 59% of its homework

Group B complete 67% of its homework.

With reference to the control group, comment on whether these data support the hypothesis.

[2 marks]



- 8** Liv owns a card shop.
She buys a large number of birthday cards to sell.

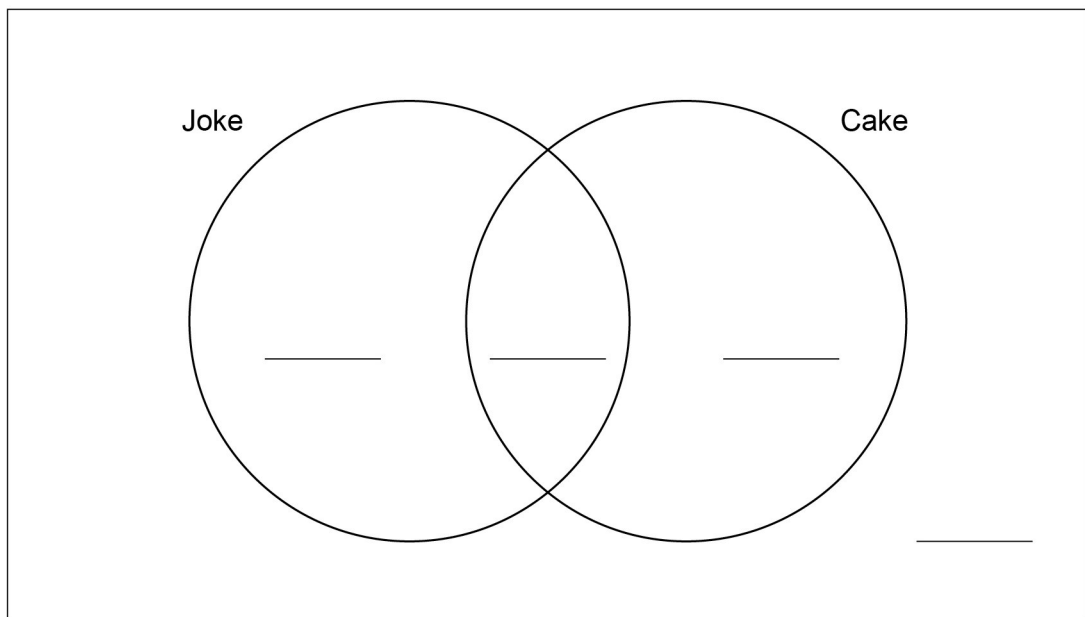
- 8 (a)** Liv looks at all the cards she bought.

On the cards she finds,

62% have a picture of a cake on them

30% have a joke on them

$\frac{1}{4}$ have neither a joke nor a picture of a cake on them.



Complete the Venn diagram.

[3 marks]

Question 8 continues on the next page

Turn over ►



8 (b)

Liv paid £1.20 for each card.

She expects to sell 300 of the cards next week.

8% of these cards are damaged and are priced at £0.75 each.

The rest are priced at £2.50 each.

Liv works out her expected profit using the following method.

$$0.92 \times 300 \times (£2.50 - £1.20) = £358.80$$

$$0.8 \times 300 \times (£1.20 - £0.75) = £108.00$$

$$\text{Expected profit} = £358.80 + £108.00 = £466.80$$

Identify the **two** errors in Liv's method.**[2 marks]**

Error 1 _____

Error 2 _____

5



- 9** Here is information about the daily number of visitors to a park for 90 days.

Number of visitors (n)	Frequency
$0 \leq n < 2000$	5
$2000 \leq n < 4000$	34
$4000 \leq n < 6000$	28
$6000 \leq n < 8000$	10
$8000 \leq n < 10\,000$	13

- 9 (a)** Calculate an estimate of the median number of visitors per day.

[3 marks]

Answer _____

Question 9 continues on the next page

Turn over ►



- 9 (b)** The data for the number of visitors in a year to a different park have,

mean = 6722

median = 5230

standard deviation = 3340

- 9 (b) (i)** Calculate the skew of these data.

Use the formula,

$$\text{skew} = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

[2 marks]

Answer _____

- 9 (b) (ii)** Write down the type of skew and interpret this in context.

[2 marks]

Type of skew _____

Interpretation _____



Turn over for the next question

*Do not write
outside the
box*

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**

Turn over ►



10

The table shows data about how many different words are in songs on a playlist.

Number of different words (w)	Frequency
$20 \leq w < 30$	
$30 \leq w < 35$	23
$35 \leq w < 40$	37
$40 \leq w < 45$	31
$45 \leq w < 55$	20
$55 \leq w < 70$	

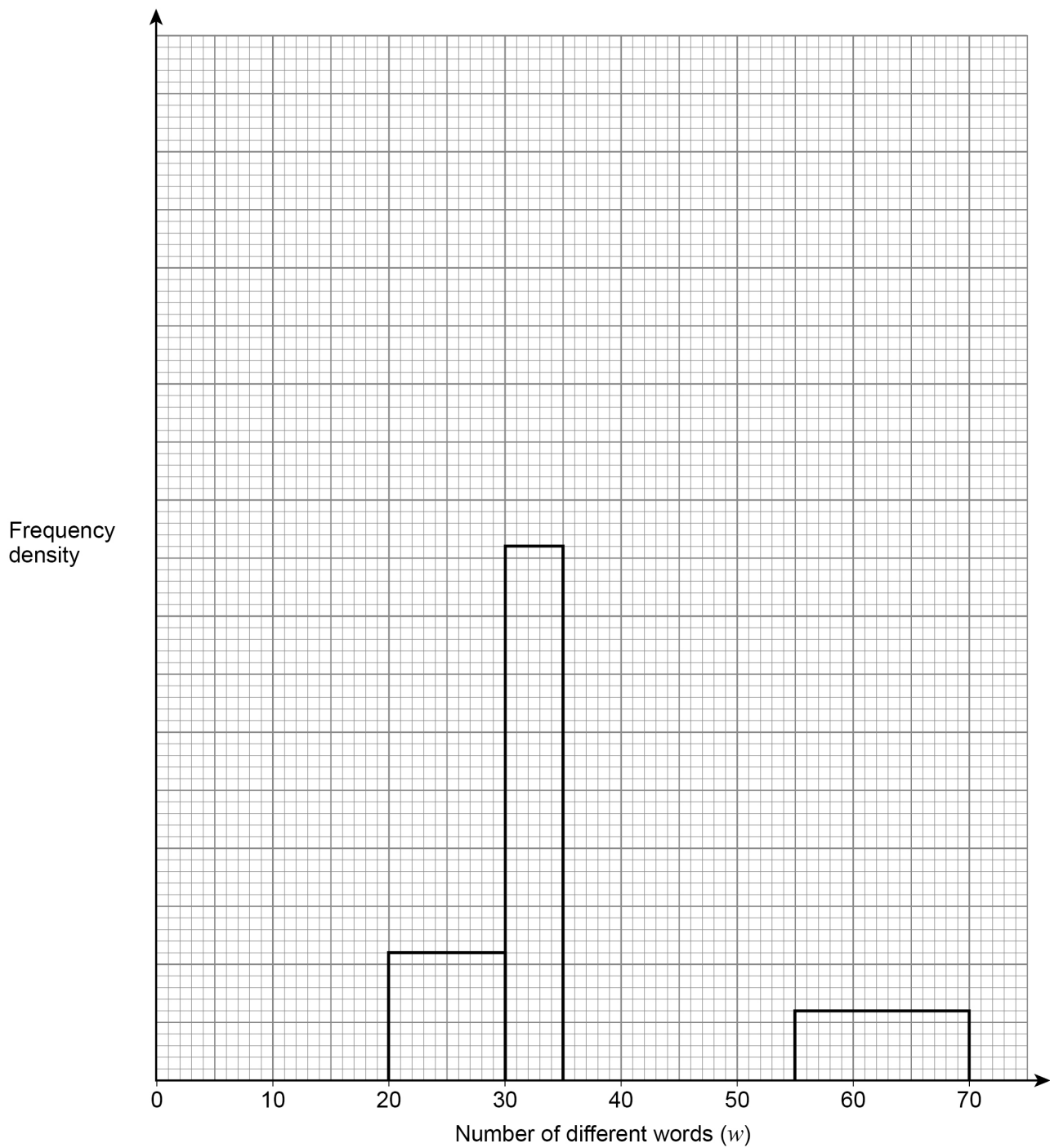
A partially completed histogram is shown opposite.

10 (a) Complete the table and the histogram.

[5 marks]



Different words in songs on a playlist



Question 10 continues on the next page

Turn over ►



10 (b) Nicki says,

“A histogram can **only** be used to describe the distribution of data if it has equal class widths.”

Is Nicki correct?

Tick (✓) a box.

Yes

☐

No

☐

Give a reason for your answer.

[1 mark]

6



Turn over for the next question

*Do not write
outside the
box*

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**

Turn over ►

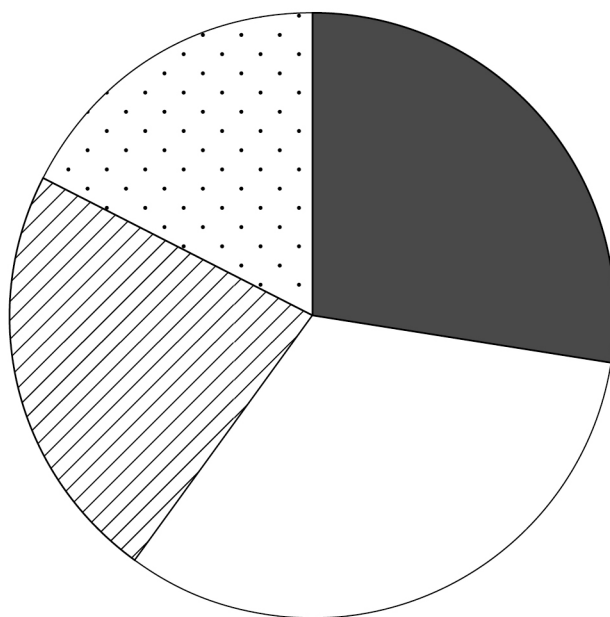


11

Cheryl collected data on the number of people in different sports clubs in two towns.

The two comparative pie charts represent the data.

Town A



Key

Swimming

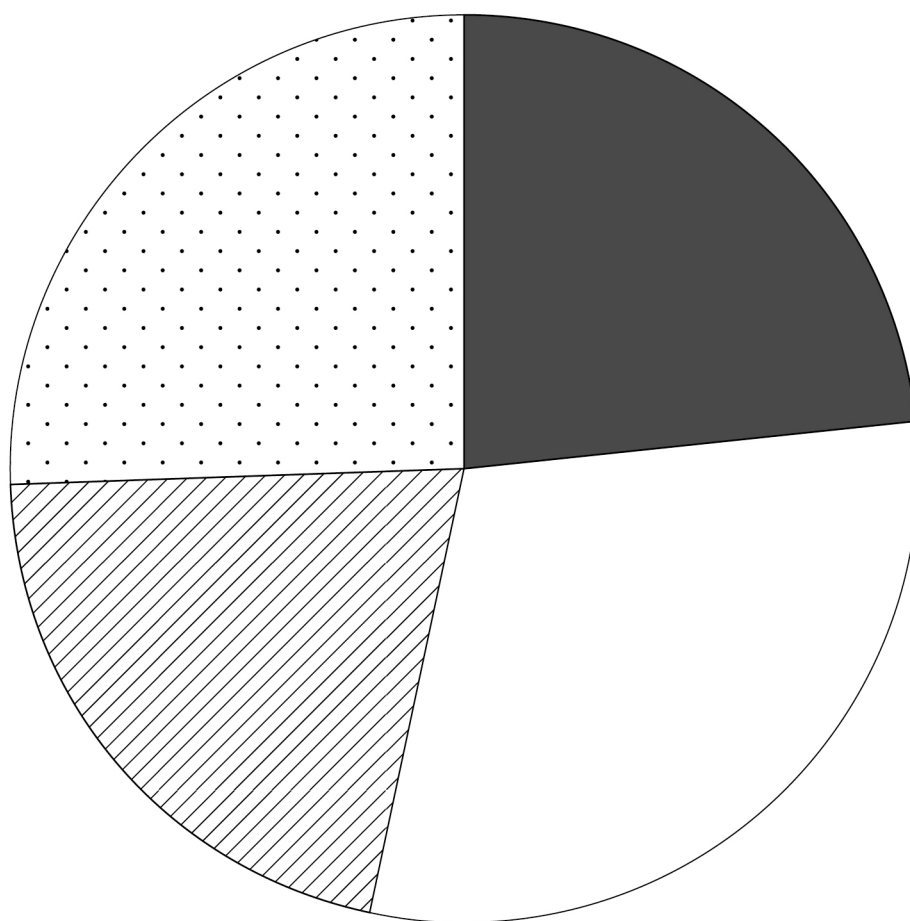
Running

Cycling

Climbing

Radius: 4cm

Town B



Key

Swimming

Running

Cycling

Climbing

Radius: 6cm



- 11 (a)** Give **one** reason why comparative pie charts are a suitable way to compare data between the towns.

[1 mark]

- 11 (b)** The pie chart for town B represents 90 people.
Calculate how many people are represented by the pie chart for town A.

[2 marks]

Answer _____

- 11 (c)** There are 11 people in the swimming club in town A.

Cheryl says,

“The angle for swimming is smaller for town B than for town A.
This means there are fewer people in the swimming club in town B.”

Is Cheryl correct?

Use calculations to support your answer.

[2 marks]



12 Two running teams record the time it takes each of their members to run 5 km, to the nearest minute.

The table below shows some of the data they collected.

	Team A	Team B
2nd decile	21	25
8th decile	33	33

Compare the results of the two teams, using calculations.

[2 marks]

2

13 Russel collects data about content creators on a video streaming service.

For each content creator, Russel records,

- their number of subscribers (n)
- the number of views per year (v)

13 (a) A content creator with 55 subscribers and 213 455 views per year is an outlier.

13 (a) (i) Give one reason why it **may** be appropriate to remove the data.

[1 mark]

13 (a) (ii) Give one reason why it **may not** be appropriate to remove the data.

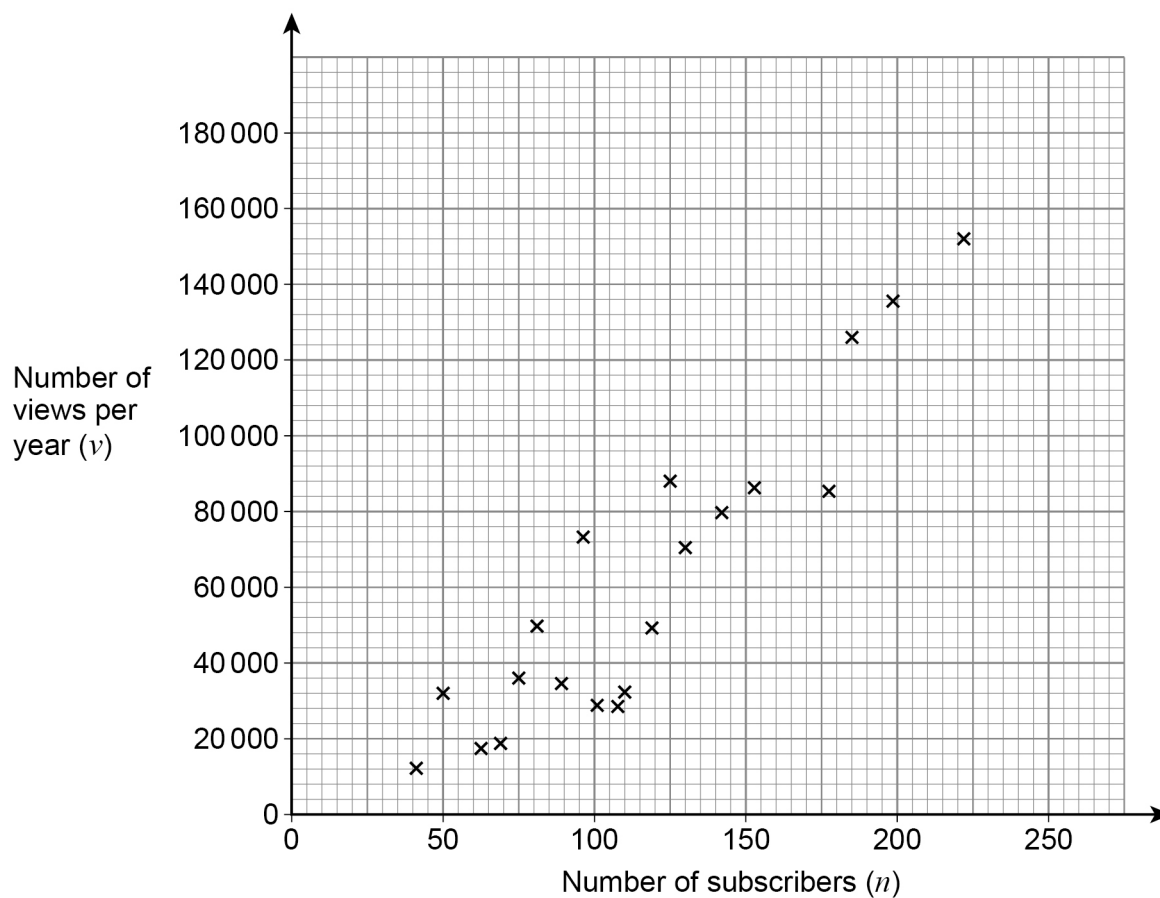
[1 mark]

Question 13 continues on the next page

Turn over ►



Here is a scatter diagram representing Russel's data after he removed the outlier.



13 (b) Which variable is being treated as the explanatory variable?

Tick (✓) a box.

Number of
subscribers

☐

Number of
views per year

☐

Give a reason for your answer.

[1 mark]



- 13 (c) After removing the outlier, the equation of the line of regression is

$$v = 750n - 25\,800$$

where v is the number of views per year, and n is the number of subscribers.

- 13 (c) (i) Plot the line of regression on the graph on the opposite page.

[2 marks]

- 13 (c) (ii) Russel uses the data to estimate the number of subscribers needed to get 250 000 views per year.

Give one reason why this is **not** appropriate.

[1 mark]

- 13 (c) (iii) Explain what the 750 in the equation of the line of regression shows in this context.

[1 mark]

Question 13 continues on the next page

Turn over ►



13 (c) (iv) A content creator had 86 subscribers.

After an increase in the number of subscribers, the same content creator's videos had approximately 58 000 views per year.

Use the equation of the line of regression to estimate how many **more** subscribers they now have.

The equation of the line of regression from **13(c)** is repeated below:

$$v = 750n - 25\,800$$

[3 marks]

Answer _____

13 (c) (v) A content creator increases their number of subscribers.

Russel uses the line of regression to estimate how many new views per year they now have.

The Pearson's product moment correlation coefficient for Russel's data is 0.92

How accurate is this estimate likely to be?

Give a reason for your answer.

[1 mark]



13 (d) For the 20 pairs of data, excluding the outlier, $\sum d^2 = 198$

The equation for Spearman's rank correlation coefficient is

$$1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Calculate the Spearman's rank correlation coefficient for the data.

[2 marks]

Answer _____

Question 13 continues on the next page

Turn over ►



- 13 (e)** The creator with the most subscribers actually has 10000 **more** views per year than recorded in the scatter graph.

State how this would affect each of the correlation coefficients and justify your answer.

[2 marks]

Pearson's product moment correlation coefficient

Tick (✓) a box.

Change

☐

No change

☐

Cannot tell

☐

Spearman's rank correlation coefficient

Tick (✓) a box.

Change

☐

No change

☐

Cannot tell

☐



14

Students complete two tests.

Each test is out of 100 marks and has a different weighting.

Test	Weighting
Test 1	80
Test 2	60

Kya says,

“The student with the greatest weighted mean score **must** have the greatest total marks for the two tests.”

Is Kya correct?

Tick (✓) a box.

Yes

☐

No

☐

Cannot tell

☐

Give a reason for your answer.

[1 mark]

1

Turn over for the next question

Turn over ►



- 15** Events A and B are **not** independent.

$$P(A|B) = \frac{2}{5} \quad \text{and} \quad P(B|A) = \frac{1}{4}$$

How many times greater is $P(A)$ than $P(B)$?

[4 marks]

Answer _____

4

END OF QUESTIONS



There are no questions printed on this page

*Do not write
outside the
box*

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**



*Do not write
outside the
box*

[illegible]

[illegible]

Question number	
	Additional page, if required. Write the question numbers in the left-hand margin.
	Copyright information For confidentiality purposes, all acknowledgements of third-party copyright material are published in a separate booklet. This booklet is published after each live examination series and is available for free download from www.aqa.org.uk. Permission to reproduce all copyright material has been applied for. In some cases, efforts to contact copyright-holders may have been unsuccessful and AQA will be happy to rectify any omissions of acknowledgements. If you have any queries please contact the Copyright Team. Copyright © 2025 AQA and its licensors. All rights reserved.

