

**Statistics GCSE****Paper 1**

Edexcel Foundation - 2026

Foundation Tier

Variant 5

1ST0/1F

**Instructions**

- Write all answers in the spaces provided.
- Answer all questions.
- You must show all your working.
- There may not be enough space to show all your working out.

**Information**

- This is a practise paper to aid your revision for your exams.
- This site, and all that work on it, have no affiliation or relationship with any exam board.
- This site is not endorsed by any company or charity, unless we state otherwise.

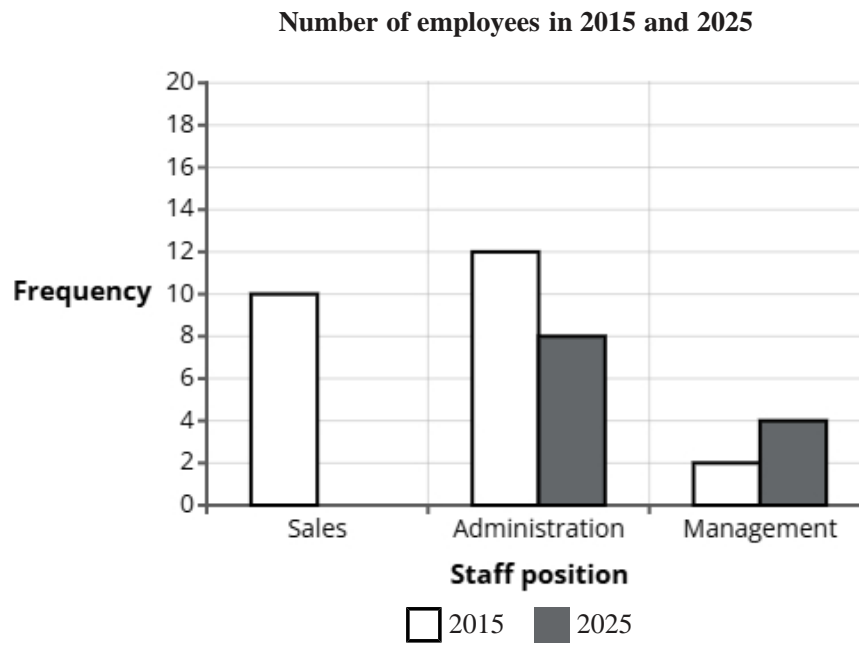
**Copyright**

- Sharing this PDF is strictly forbidden unless you have the author's permission.
- This paper is only authorised to be used by the person who has bought it.
- You may not store this online or any shared area such as an intranet.
- Please contact us if you have any queries.

**Advice**

- You can get support for all these questions at our website: [www.statsgcse.com](http://www.statsgcse.com)
- This paper and more are available on our site with questions that change subtly after each attempt.
- Good luck!

- 1 The comparative bar chart compares the number of employees working in sales, administration and management for Ted's Tiles in 2015 and 2025.

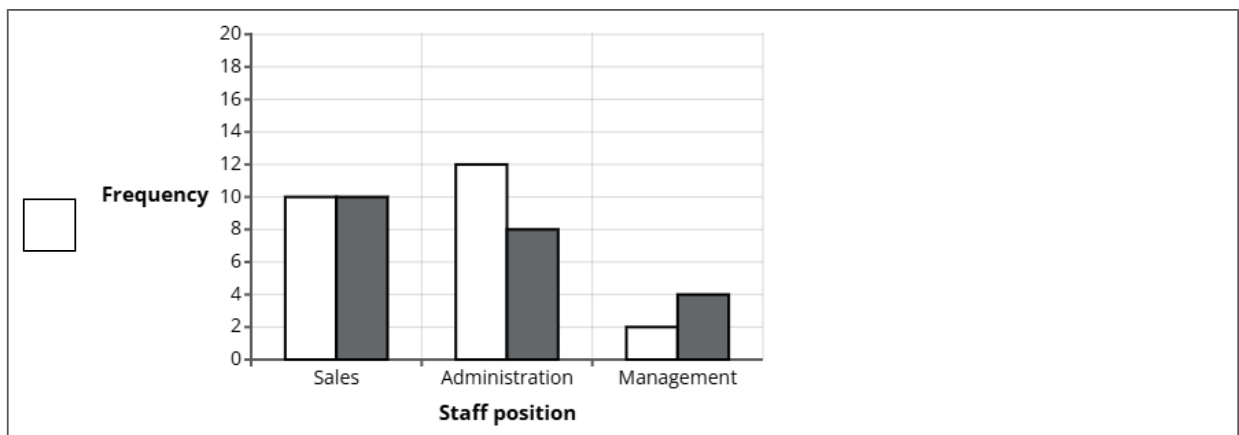
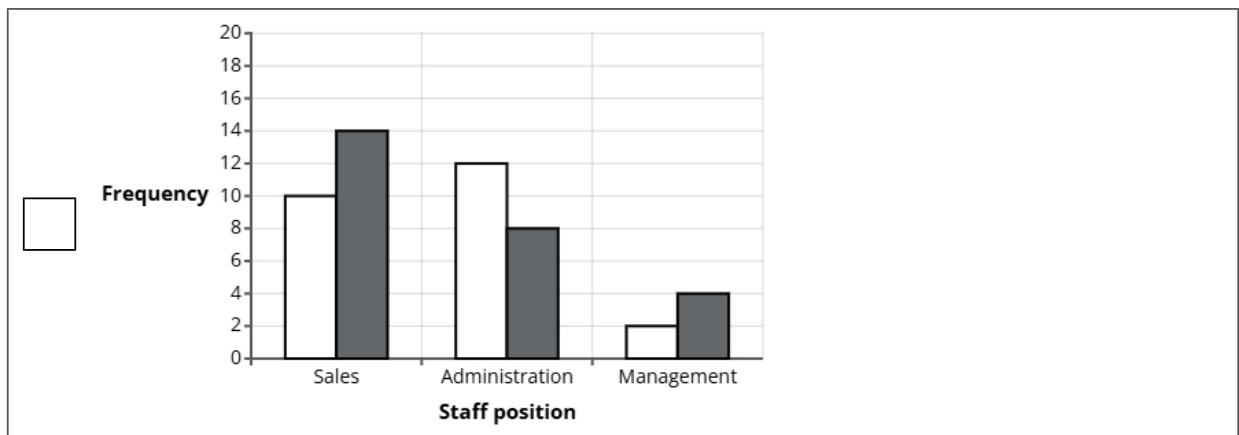
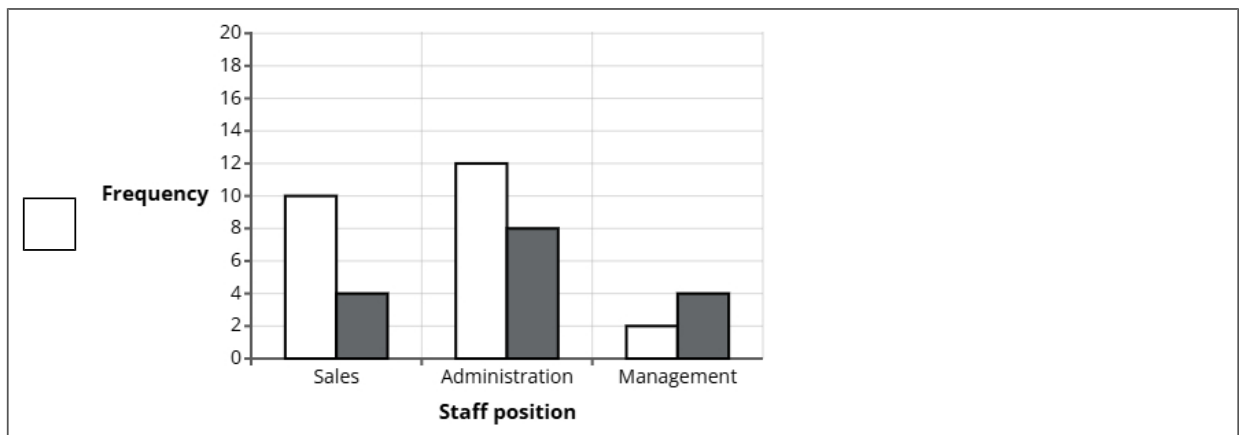
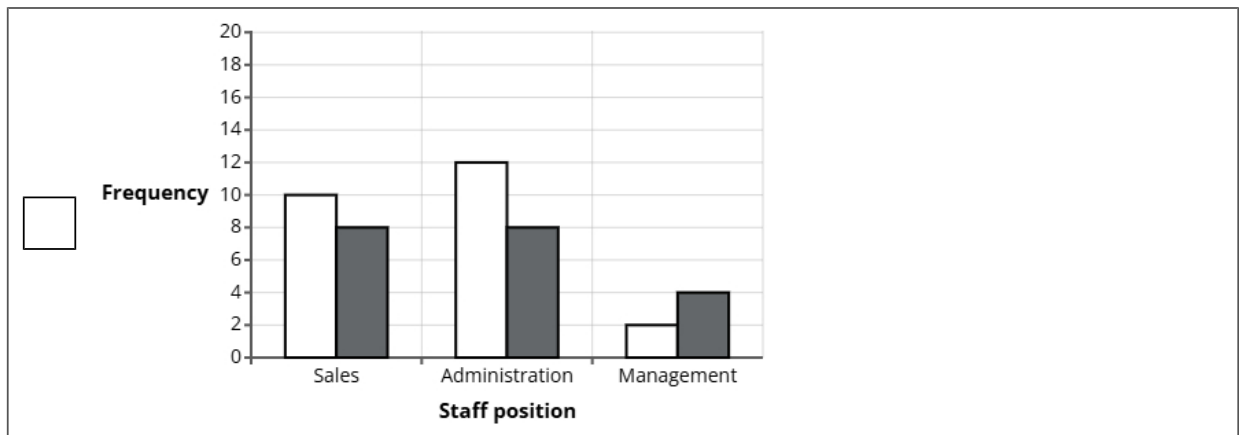


In 2025, there were 14 employees working in sales.

- (a) Complete the comparative bar chart for sales employees.

(1 mark)

Select the correct answer.



(b) Find how many more employees worked in administration than management at Ted's Tiles in 2015.

(2 marks)

2015 is shown with an unshaded bar

(c) Compare the the number of employees who worked in sales, administration and management in 2015.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

- Most staff worked in sales.
- Most staff worked in administration.
- More staff worked in sales than in management.
- More staff worked in management than in sales.

(d) The data displayed in the comparative bar chart is an example of quantitative data.

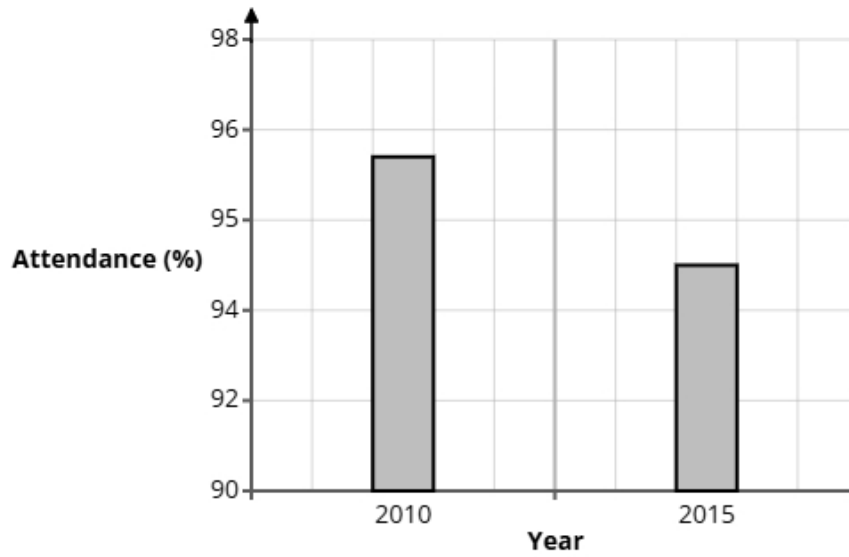
Explain what is meant by quantitative data.

(1 mark)

Select **one** box.

- Quantitative data is data collected through observation without measurement.
- Quantitative data is information that describes qualities using language.
- Quantitative data is information that cannot be analysed statistically.
- Quantitative data is data that represents amounts or measurements using numbers.

- 2 A newspaper columnist gathered data on school attendance rates at a school 2010 and 2015.  
A bar chart is drawn from the information.



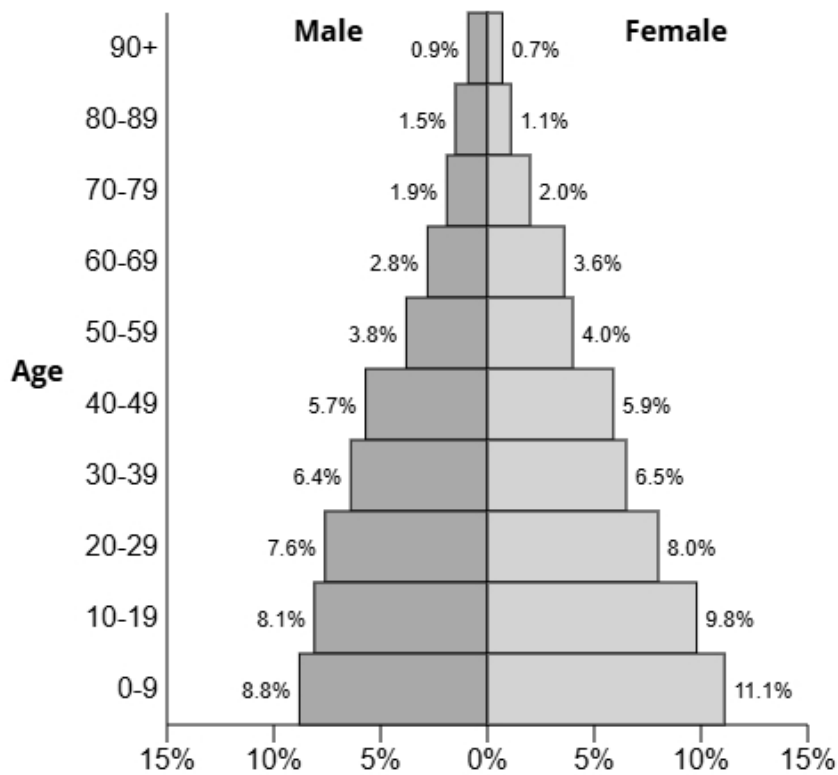
State **two** reasons why the bar chart could be misleading.

(2 marks)

Select **two** boxes.

- The vertical axis on the graph does not begin at 0.
- The heights are wrong.
- The vertical axis scale does not go up in equal steps.
- There is no title for the graph.

- 3 The population pyramid below shows the percentage of males and females in each age group for the town Redmere.



- (a) Write down the percentage of females in the age group 40-49.

(1 mark)

Find 40-49 on the population pyramid and read off the number on the right-hand side (females).

\_\_\_\_\_ %

(b) Find the age group for males that has 2.8% of the population.

(1 mark)

Select **one** box.

60-69

70-79

80-89

50-59

(c) Find the age group that has 15.6% of the population.

(1 mark)

Select **one** box.

0-9

10-19

30-39

20-29

(d) Compare the percentage of the population aged 40-69 between males and females.

(1 mark)

Select **one** box.

There are more males.

They are both the same.

There are more females.

(e) Give a reason why the sum of all the percentages is 100.2% and not 100%.

(1 mark)

Select **one** box.

- Some of the population may not have been counted.
- The figures have been rounded.
- It is incomplete.
- They are percentages not amounts.

4 A theme park is considering building a new roller coaster.

Mia is carrying out a survey to see what all visitors think about the new attraction.

Mia thinks that she should take a sample rather than a census.

(a) Mia has decided to use the electoral register as a sampling frame.

State one problem Mia may have using the electoral register as a sampling frame.

(1 mark)

Select **one** box.

The electoral register would also include people's addresses.

The electoral register may be out of date.

There will be too many names.

Unreliable.

(b) Mia is writing a plan for the investigation into visitors' opinions on the new roller coaster.

Write down what Mia should include in the plan and explain why each of the things is appropriate.

You should include:

- + a sampling method
- + a question Mia should ask in the questionnaire
- + a statistical diagram to show the results of the survey.

(6 marks)

Number the **two** correct statements in the correct order (**two** statements are incorrect).

- Mia should use tick-tock sampling.
- This will ensure that every visitor has an equal chance of being selected.
- This will ensure that the students asked are the most knowledgeable.
- Mia should use random sampling.

Number the **two** correct statements in the correct order (**two** statements are incorrect).

A question could be:

- How would you describe the plans for the new roller coaster?  
 Very Poor  Poor  Neutral  Good  Very Good
- The question is clear and unbiased, avoiding leading students to a particular answer.
- The question is open so will be easier and quicker to analyse.
- A question could be:  
What do you think makes the new roller coaster so amazing?

Number the **two** correct statements in the correct order (**two** statements are incorrect).

- A line chart can be used to display the data.
- This is because it shows frequencies and allows for visual comparisons.
- This is because it can be used also to collect the data.
- A tally chart can be used to display the data.

5 The manager of a supermarket is reviewing the number of checkout staff on duty.

The manager wants feedback from customers.

He plans to give a questionnaire to customers shopping between 10 am and 12 pm on a Friday and ask them to complete it at home and hand it in the following day.

(a) Describe the meaning of the term 'simple random sample'.

(1 mark)

Select **one** box.

- Every person in the population has an identical probability of being included in the sample.
- To ensure that every subgroup in the population is proportionally represented.
- To ensure only the most qualified individuals are included in the sample.
- To make data collection easier by selecting the most accessible individuals.

(b) Assess the supermarket manager's plan to get the opinions of the customers.

(3 marks)

Select **three** boxes.

- No customers will return the questionnaires.
- The customers are not being selected at random.
- Not effective.
- The customers are only being asked on one day.
- Questionnaires are a good way to collect opinions.

(c) Here is an open question that the supermarket manager is considering for the questionnaire.

What do you think about the number of checkout staff?

Give one reason why this is not a good question.

(1 mark)

Select **one** box.

- The responses will not be easy to analyse.
- It is not a closed question.
- The question is not relevant.
- There may be issues with privacy.

(d) Design a suitable closed question for the supermarket manager to use on his questionnaire so that he can decide how many checkouts to open.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

- Good
- Average
- Bad
- Very bad
- 1 to 2 checkouts
- 3 to 4 checkouts
- 5 to 6 checkouts
- 7 or more checkouts
- How many checkouts do you think should be open?
- What do you think of our checkouts?

- (e) When the supermarket manager has designed his questionnaire, he decides to pre-test it by using a pilot survey with a small sample of customers.

Select **two** reasons why he should conduct a pilot survey.

(2 marks)

Select **two** boxes.

- A pilot survey will include more people.
- A pilot survey will check questions are clear.
- A pilot survey will check questions are inoffensive.
- A pilot survey will give more accurate data.

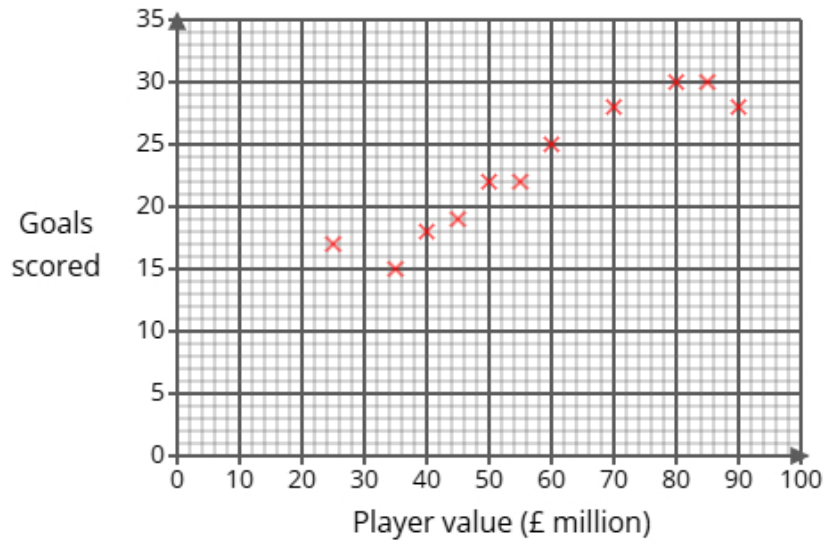
- 6 A school librarian is interested in how much time students spend reading for pleasure each week. He plans to ask a sample of 20 students to record the number of hours they read in one week. Each student will write down their reading time on a piece of paper. Describe one problem the school librarian might face in the statistical enquiry process due to non-response or unexpected results, and explain how he could address this issue.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

- He may need to ask more students.
- Not all students may read for pleasure.
- To address this, before giving out the pieces of paper he could ask if they read for pleasure.
- To address this, he could show the results on a scatter diagram.

- 7 Alice collected data on 11 football strikers, recording the value (in millions of pounds) and the goals scored last season of each player. She represented her findings in the scatter diagram below.



- (a) One of the 11 strikers scored 15 goals last season.

Write down the value of the striker.

(1 mark)

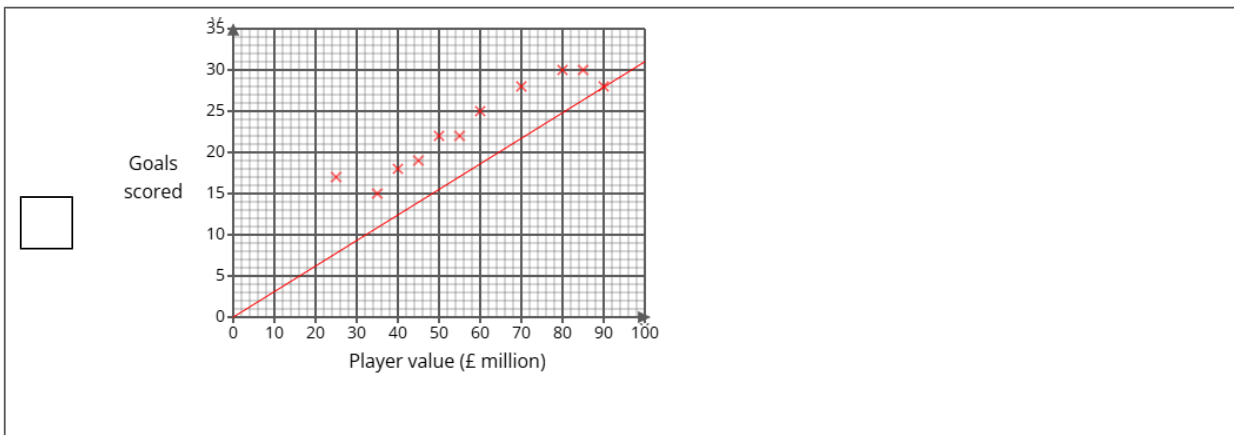
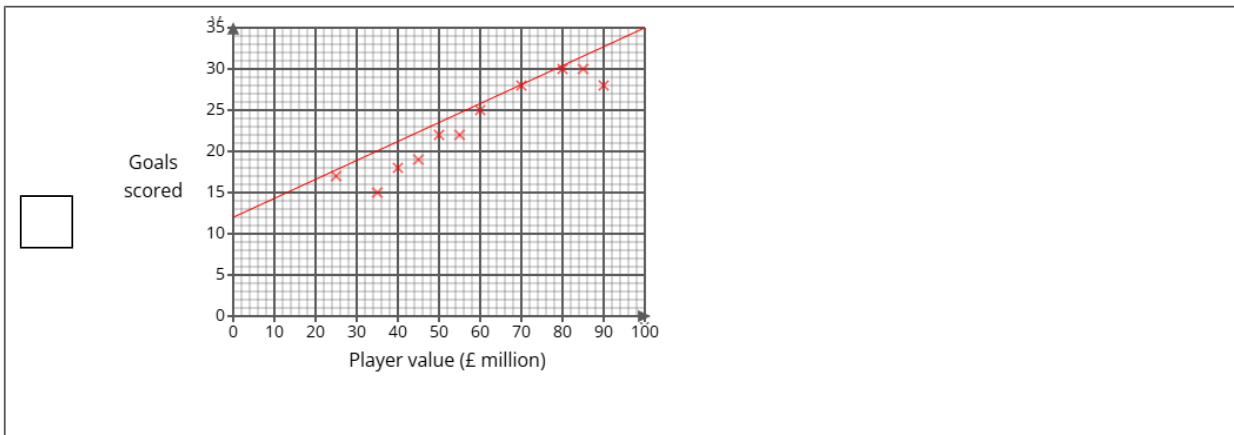
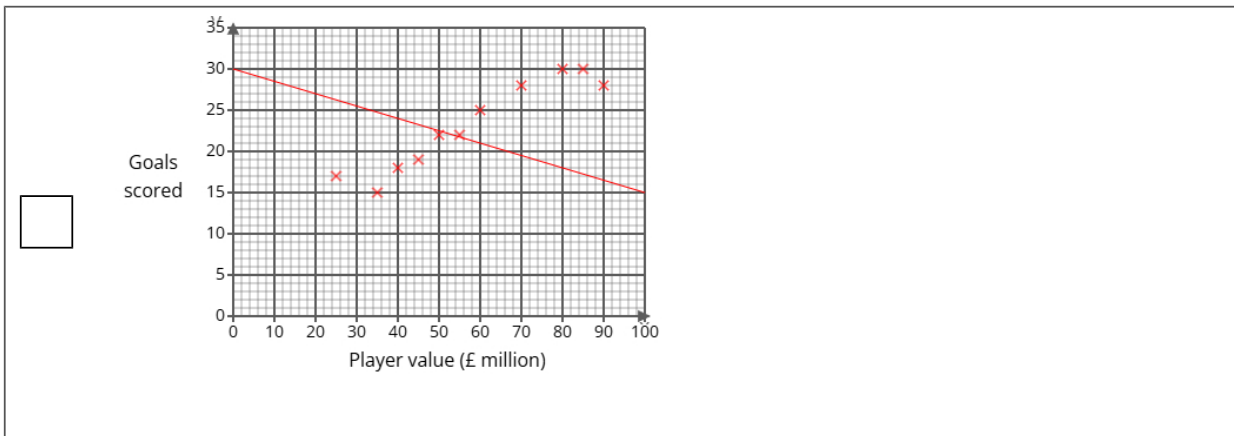
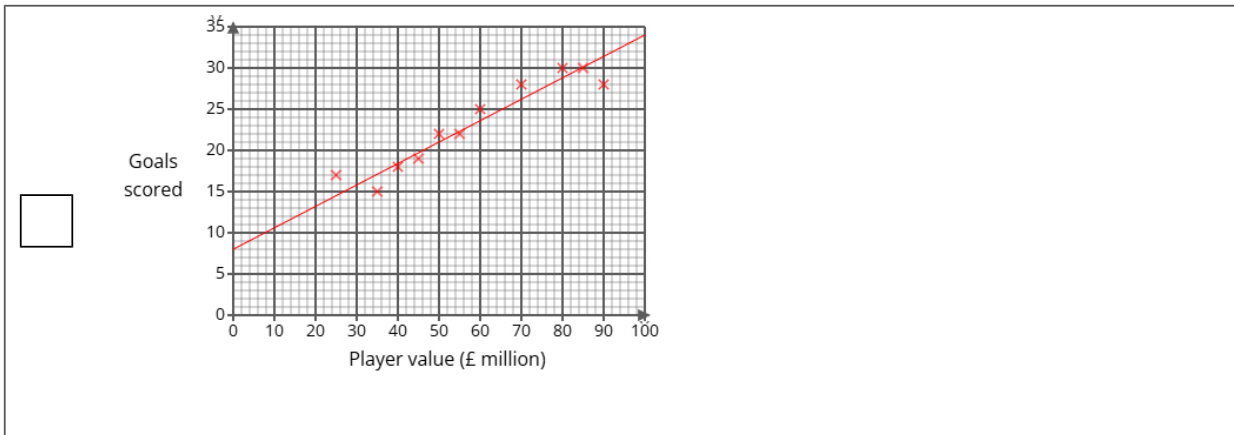
Find the cross on the scatter graph that is at 15 million on the x-axis (the bottom axis), then read off the value from the y-axis (the side axis).

£ \_\_\_\_\_

(b) Draw a line of best fit on the scatter diagram.

(1 mark)

Select the correct answer.



(c) Describe and interpret the type of correlation shown by the scatter diagram.

(3 marks)

Number the **two** correct statements in the correct order (**three** statements are incorrect).

- The correlation is negative and
- There is no correlation but it is
- weak
- strong
- The correlation is positive and

Select **one** box.

- As the value increases the goals scored decreases.
- A striker with a high value will have a low goals scored.
- As the value increases the goals scored increases.
- A striker with a high value will have a high goals scored.

(d) A new strike has just been signed for value of £150 million.

Alice is planning on using the line of best fit on the scatter diagram to predict the goals the strike will score this season.

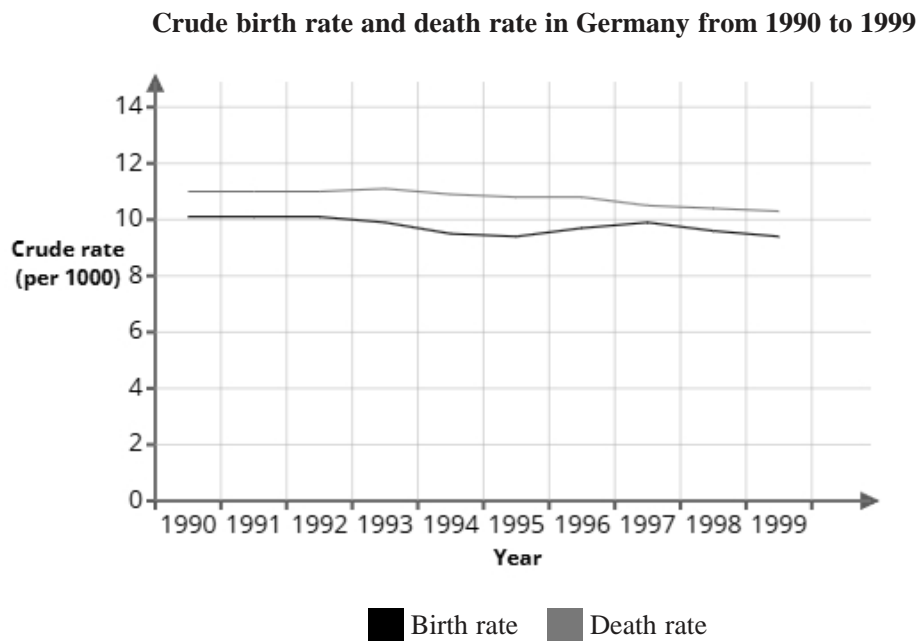
Explain whether or not it is appropriate to use the line of best fit for this prediction.

(2 marks)

Number the **two** correct statements in the correct order (**two** statements are incorrect).

- This is not appropriate
- This is appropriate
- because the point is after the data and the trend may not continue.
- because the trend will continue.

8 The graph shows the crude birth rate and death rate in Germany from 1990 to 1999



Maya uses the information in the graph to conclude:

"The total population in Germany has decreased from 1990 to 1999"

(a) Explain how the information in this graph supports Maya's conclusion.

(1 mark)

Select **one** box.

- In 1990, the death rate is roughly 1 bigger than the birth rate.
- The graph shows that Germany is a large country.
- The crude death rate is higher than the crude birth rate.
- The crude birth rate is higher than the crude death rate.

(b) Give a reason why Maya's conclusion might **not** be correct.

(1 mark)

Select **one** box.

- Maya's conclusion does not take into account where in Germany these births happened.
- The data in the graph may be inaccurate.
- Maya's conclusion does not take into account immigration.
- Maya's conclusion does not take into account the average age of the population.

(c) In 2000, the population of Germany was 82 445 521.

There were 677 054 recorded births.

Using the formula below, calculate the crude birth rate in 2000.

Give your answer correct to 1 decimal place.

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

(2 marks)

Substitute each of the values into the crude birth rate formula.

9 Olivia investigates the number of pages in 170 books from a library.

The pages range from 302 pages to 389 pages.

Olivia considers using one of the two possible grouped frequency tables for the results, Table A or Table B, shown below.

**Table A**

Pages	Frequency
$270 < p \leq 300$	0
$300 < p \leq 330$	42
$330 < p \leq 360$	94
$360 < p \leq 390$	34
$390 < p \leq 420$	0

**Table B**

Pages	Frequency
$300 < p \leq 320$	19
$320 < p \leq 340$	58
$340 < p \leq 360$	59
$360 < p \leq 380$	26
$380 < p \leq 400$	8

(a) Give **two** advantages of using grouped data rather than raw data.

(2 marks)

Select **two** boxes.

- Grouped data gives more precise results.
- Grouped data ensures that there are no gaps in the data.
- Grouped data is easier to process large amounts of data.
- Grouped data doesn't require any calculation, as it displays frequencies directly.
- Grouped data is easier to represent on graphs.

(b) Give **one** disadvantage of using grouped data rather than raw data.

(1 mark)

Select **one** box.

- Grouped data is much more difficult to read.
- Grouped data will lose the detail in the data.
- Grouped data cannot be used with decimals.
- Grouped data cannot show patterns in the data.

(c) Olivia feels that Table B gives more detail than Table A about the results.

Assess the appropriateness of Olivia's claim.

(2 marks)

Select the **three** correct statements (**three** statements are incorrect).

- In Table A, all the data is concentrated into three groups.
- Olivia's claim is justified.
- In Table A, the data goes from 270 to 420, showing a much wider range of data.
- In Table B, the table starts at 300 and the lowest value is 302 pages and ends at 400 with the highest value at 389 pages.
- Olivia's claim is not justified.
- In Table B, some data could have been less than 300 or more than 400, but would not be shown.

(d) Olivia wants to work out the average number of pages in the 170 books from a library.

She decides to use Table B.

Calculate the average number of pages in the 170 books from a library, giving your answer to 1 decimal place.

(3 marks)

Add midpoint and  $fp$  columns onto the table.

Then find the sums of the  $f$  and  $fp$ .

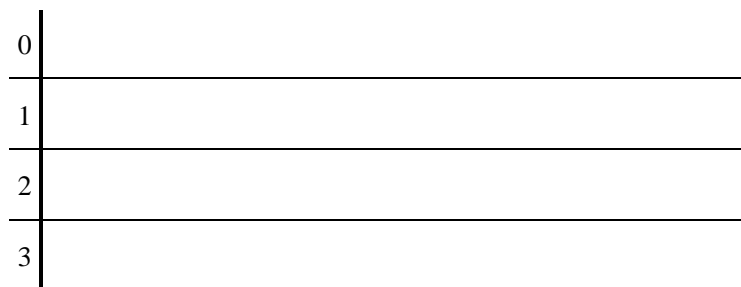
Find the mean by  $\frac{\sum fp}{\sum f}$

**10** 23 male adults were asked to perform a simple dance step repeatedly and the number of completions was measured.

Here are the results.

37	9	10	24	10	27	25	39
8	17	33	39	31	11	10	33
7	9	13	35	23	27	24	

(a) Complete the stem and leaf diagram for the data.



**Key:** 1|0 = 10

(2 marks)

Select the correct answer.

	0	7 8 9 9
	1	3 4 4 5 7 7
<input type="checkbox"/>	2	0 0 0 1 3 7
	3	1 3 3 5 7 9 9

	0	7 8 9 9
	1	0 0 0 1 3 7
<input type="checkbox"/>	2	3 4 4 5 7 7
	3	1 3 3 5 7 9 9

	0	0 0 0 1 3 7
	1	7 8 9 9
<input type="checkbox"/>	2	3 4 4 5 7 7
	3	1 3 3 5 7 9 9

	0	7 8 9 9
	1	0 0 0 1 3 7
<input type="checkbox"/>	2	1 3 3 5 7 9 9
	3	3 4 4 5 7 7

(b) Work out the interquartile range from the data.

(2 marks)

Find the lower quartile using  $\frac{1}{4}(n + 1)^{\text{th}}$

Find the upper quartile using  $\frac{3}{4}(n + 1)^{\text{th}}$

Find the interquartile range

**IQR = upper quartile – lower quartile**

(c) A group of 23 female adults were also measured.

The results from the female adults had a median of 28 and an interquartile range of 27.

David thinks that these results show that females are better at performing a dance step than males.

State whether you agree with David and give reasons why.

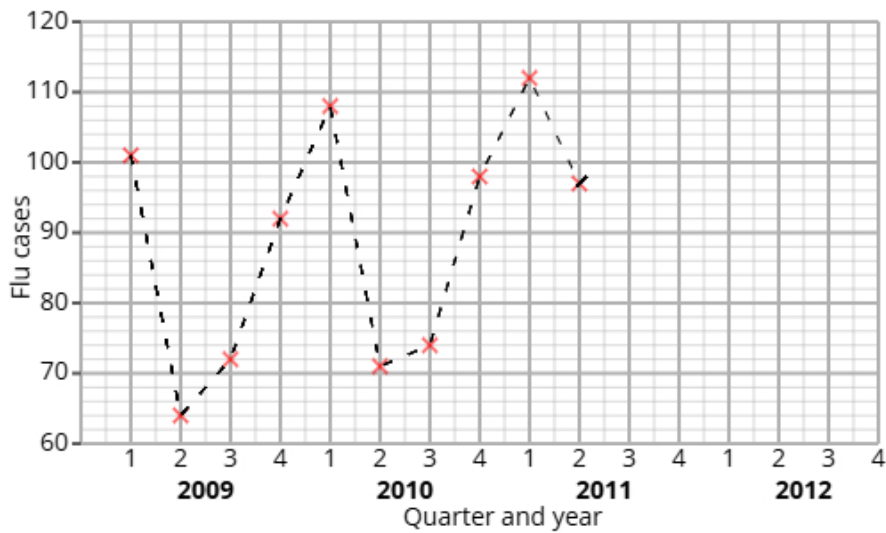
(3 marks)

Median for male adults = \_\_\_\_\_

Select the **two** correct statements (**two** statements are incorrect).

- The median for the females is higher than the median of the males.
- No, David's conclusion is supported by the data.
- The median for the males is higher than the median of the females.
- Yes, David's conclusion is supported by the data.

- 11 The time series graph shows information about the the number of flu cases at a doctor's surgery from 2009 to 2011.



Ethan calculates the 4-point moving averages from the time series graph, which are shown below.

82      84      86      86      88      89      95

- (a) Identify and interpret in context one example of seasonality displayed in the time series graph.

(2 marks)

Number the **two** correct statements in the correct order (**two** statements are incorrect).

- which shows that less people have flu in the winter.
- The lowest values are in Q1
- The greatest values are in Q1
- which shows that more people have flu in the winter.

(b) Explain why a 4-point moving average is appropriate.

(1 mark)

Select **one** box.

- A 4-point moving average captures cyclic patterns that repeat every four quarters.
- It gives more weight to the middle values.
- It shows skew within the data.
- A 4-point moving average gives us more data.

**12** A fair 3-sided spinner is numbered 1, 2, 3.

A fair 4-sided spinner is numbered 1, 2, 3, 4.

The spinners are used to play a game. Both spinners are spun and the total score is recorded.

		4-sided spinner			
		1	2	3	4
3-sided spinner	1	2	3		
	2	3			
	3				

The game is won when the total is at least 6.

Noah plays the game once.

(a) Complete the sample space diagram.

(2 marks)

Select the correct answer.

<input type="checkbox"/>	3-sided spinner	4-sided spinner				
		1	2	3	4	
		1	2	3	3	4
		2	3	3	4	5
3	3	4	5	6		

<input type="checkbox"/>	3-sided spinner	4-sided spinner				
		1	2	3	4	
		1	2	3	4	5
		2	3	4	5	6
3	4	5	6	7		

<input type="checkbox"/>	3-sided spinner	4-sided spinner				
		1	2	3	4	
		1	2	3	3	4
		2	3	4	6	8
3	3	6	9	12		

(b) Find the probability that Noah wins the game.

(2 marks)

Find all the numbers in the table that are 6 or larger

Put this number as the numerator and total amount of numbers as the denominator

$$\text{probability} = \frac{\text{6 or larger}}{\text{total outcomes}}$$

**13** Mia organises two language workshops, French A and French B, to help people learn conversational French.

She wants to compare the two workshops to see which improves conversation skills.

The table shows number of participants who passed or failed the conversation test.

	Passed	Failed	Total
French A	24	16	40
French B	25	50	75

**(i)** Find the relative risk of failing the conversation test having been in French A compared to French B.

**(ii)** Give an interpretation of your answer to part (i).

(4 marks)

Write your answer as a decimal.

\_\_\_\_\_

Select **one** box.

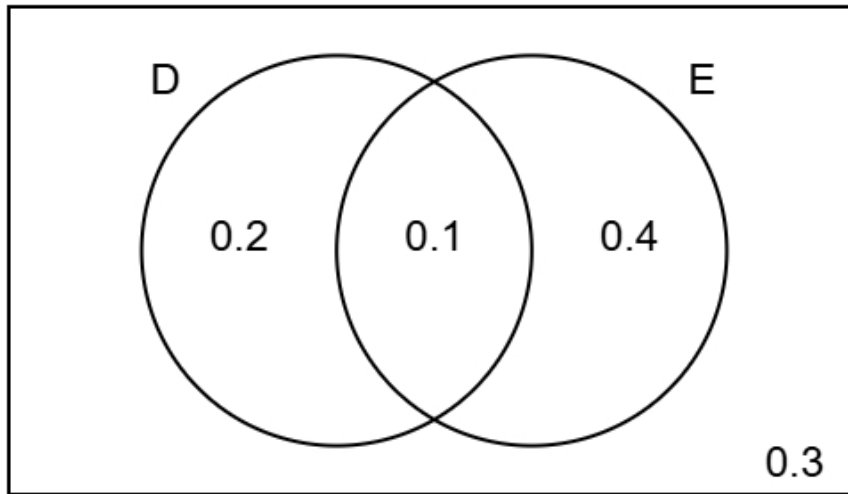
More people failed the conversation test in French A than in French B.

Less people failed the conversation test in French A than in French B.

The risk of failing the conversation test having taken French A is greater than the risk of failing the conversation test having taken French B.

The risk of failing the conversation test having taken French A is lower than the risk of failing the conversation test having taken French B.

- 14 The Venn diagram shows information about the probabilities of two events occurring.  
The events are labelled as D and E.



- (a) Find the probability of event E happening.

(1 mark)

Add the probabilities in the circle marked E together

Leave your answer as a decimal.

\_\_\_\_\_

- (b) Find  $P(D \text{ and } E)$

(1 mark)

$P(D \text{ and } E)$  is shown in the overlap of the Venn diagram

Leave your answer as a decimal.

\_\_\_\_\_

(c) Find  $P(E | D)$

(2 marks)

Use the formula to find  $P(E | D)$

$$P(E | D) = \frac{P(D \text{ and } E)}{P(D)}$$

(d) Two different events events A and B are independent.

$$P(A) = 0.3$$

$$P(B) = 0.9$$

Find  $P(A \text{ and } B)$

(2 marks)

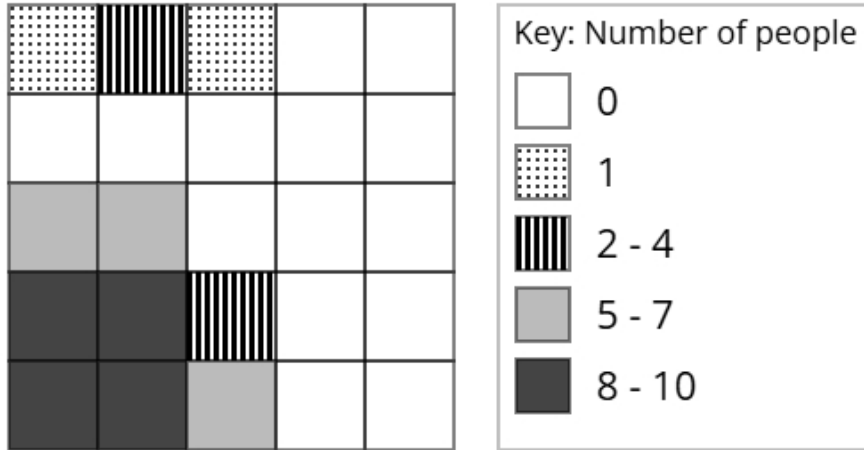
Use the formula for independent events

$$P(A \text{ and } B) = P(A) \times P(B)$$

Leave your answer as a decimal.

\_\_\_\_\_

- 15** The choropleth map below represents a university campus that has been divided into 25 squares of equal area. Fatima has collected data about the popularity of different parts of the university campus. The number of people recorded in each square on one Friday morning is shown.



- (a) Calculate an estimate of the total number of people that were recorded on Friday.

(3 marks)

Find the midpoints for the groups.

Multiply each key with amount of squares and add them up.

(b) Fatima would like to open a bagel stand in the university campus.

After analysing the data, she decides that she should open the bagel stand in the corner of the university campus shown at the bottom left of the choropleth map.

Using the information in the choropleth map, assess the validity of Fatima's conclusion.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

because there were less people at the bottom left of the university campus.

Fatima's comment is not valid

because there were more people at the bottom left of the university campus.

Fatima's comment is valid

(c) Ethan argues that the method used by Fatima to collect the data is not appropriate for reaching a reliable conclusion.

Assess whether Ethan's argument is correct and give a reason.

(1 mark)

Select the **two** correct statements (**two** statements are incorrect).

because the data was only collected on one Friday.

Ethan is correct

Ethan is not correct

because there was a large amount of data collected.