Scaffolded version

Statistics GCSE

Paper 1

Edexcel Foundation - 2026

Foundation Tier

Variant 2

1ST0/1F

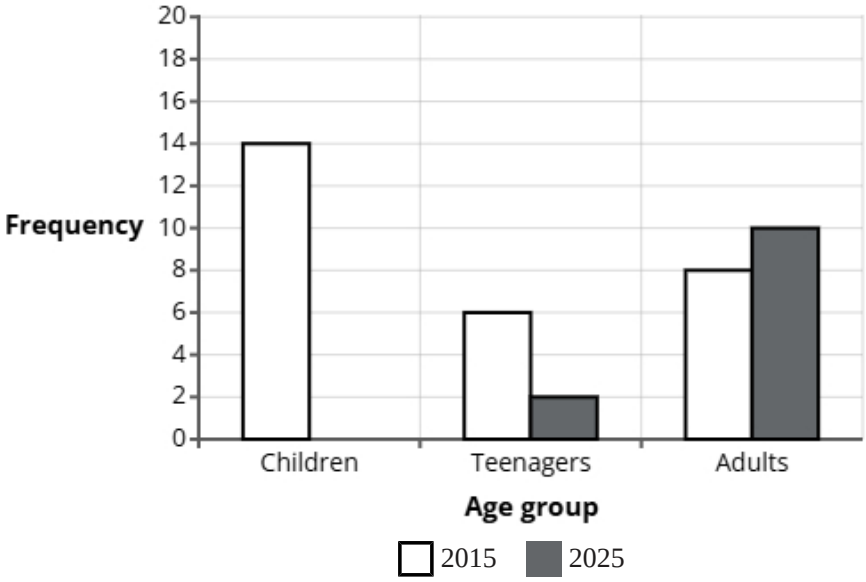
Answers

Visit our website for tutorials on each question.

www.statsgcse.com

1 The comparative bar chart compares the average number of books borrowed from a library by children, teenagers and adults in 2015 and 2025.

Average annual library book loans in 2015 and 2025



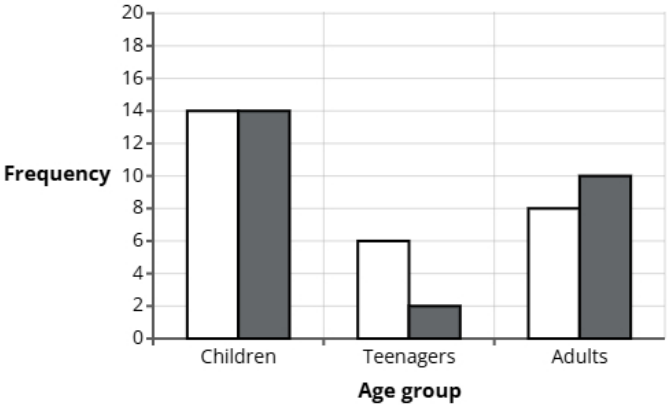
In 2025, children borrowed an average of 16 books.

(a) Complete the comparative bar chart for children.

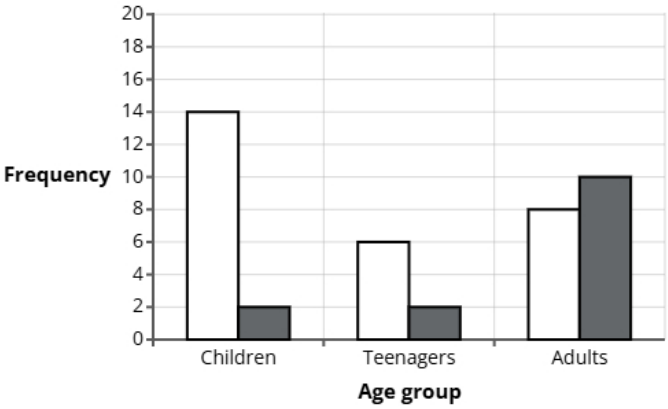
(1 mark)

Select the correct answer.

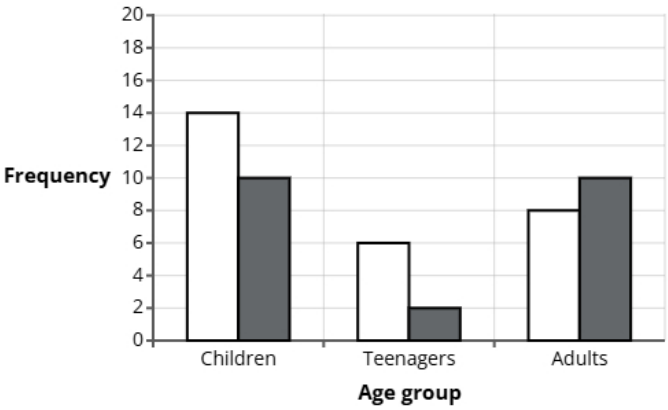
☐



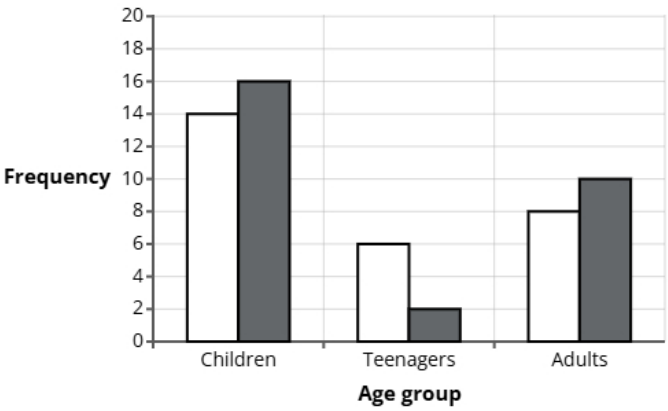
☐



☐



☐



(b) Find how many more library books were borrowed by adults than teenagers in 2015.

(2 marks)

2

(c) Compare the number of library books borrowed by children, teenagers and adults in 2015.

(2 marks)

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

- [ ] Adults borrowed more books than children.
- [x] Teenagers borrowed the least books.
- [x] Children borrowed more books than adults.
- [ ] Adults borrowed the least books.

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

Explain what is meant by quantitative data.

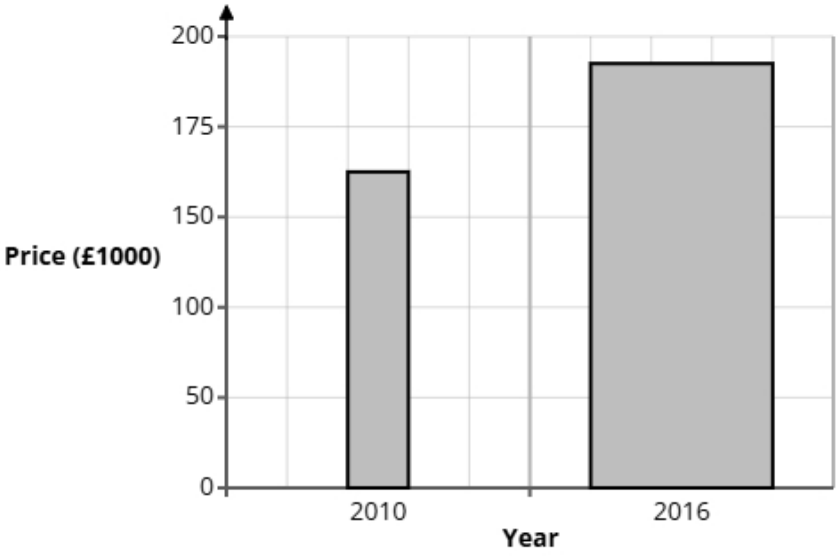
(1 mark)

*Select **one** box.*

- [x] Quantitative data is numerical information that can be measured or counted.
- [ ] Quantitative data is descriptive information recorded in sentences.
- [ ] Quantitative data is data that explains why people behave in certain ways.
- [ ] Quantitative data is data that is always subjective and based on personal interpretation.

2 A newspaper journalist compared house prices in 2010 and 2016 to support an article on the cost of living.

A bar chart is drawn from the information.



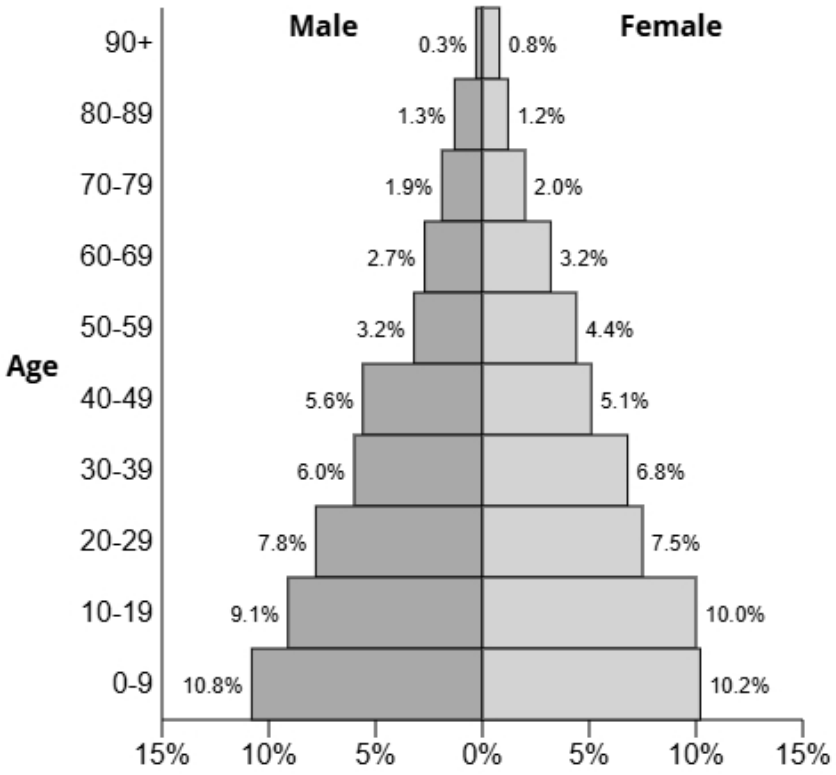
State **two** reasons why the bar chart could be misleading.

(2 marks)

*Select **two** boxes.*

- [x] The vertical axis scale does not go up in equal steps.
- [x] The bar width for 2016 is bigger than 2010.
- [ ] There is no title for the graph.
- [ ] The bars have different areas.

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



(a) Write down the percentage of females in the age group 30-39.

(1 mark)

6.8%

(b) Find the age group for males that has 3.2% of the population.

(1 mark)

*Select **one** box.*

- [ ] 20-29
- [ ] 40-49
- [x] 50-59
- [ ] 30-39

(c) Find the age group that has 19.1% of the population.

(1 mark)

*Select **one** box.*

- [ ] 30-39
- [ ] 10-19
- [ ] 20-29
- [ ] 40-49

(d) Compare the percentage of the population aged 20-49 between males and females.

(1 mark)

*Select **one** box.*

- [ ] There are more males.
- [ ] There are more females.
- [ ] They are both the same.

(e) Give a reason why the sum of all the percentages is 99.9% and not 100%.

(1 mark)

*Select **one** box.*

- [ ] It is incomplete.
- [ ] There are gaps in the ages.
- [ ] The figures are wrong.
- [ ] They have been rounded.

4 A city council is considering adding more public transport routes.

Ethan wants to conduct a survey to learn what all the residents in the city think about the plan.

Ethan thinks that he should take a sample rather than a census.

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

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

(1 mark)

*Select **one** box.*

- [ ] There will be too many names.
- [ ] The electoral register would also include people's addresses.
- [x] Only those registered to vote would be included.
- [ ] Unreliable.

(b) Ethan is writing a plan for the investigation into residents' opinions on the new transport routes.

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

You should include:

+ a sampling method

+ a question Ethan 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).*

☐ Ethan should use first-past-the-post sampling.

1 Ethan should use random sampling.

☐ This will ensure that the students asked are the most knowledgeable.

2 This will ensure that every resident has an equal chance of being selected.

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

☐ A question could be:
Why do you think the bus routes need to improve?

A question could be:

1 How satisfied are you with the current bus service?
☐ Very Satisfied ☐ Satisfied ☐ Neutral ☐ Unsatisfied ☐ Very Unsatisfied

☐ The question is open so will be easier and quicker to analyse.

2 The question is clear and unbiased, avoiding leading students to a particular answer.

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

1 A line chart can be used to display the data.

2 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 owner of a café is reviewing the variety of cakes they offer.

The owner wants feedback from customers.

He plans to give a questionnaire to customers who visit the café on Saturday afternoon and ask them to bring it back next time they come to the café.

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

(1 mark)

*Select **one** box.*

- [ ] Individuals are chosen based on specific characteristics to ensure diversity.
- [ ] Individuals are chosen in order of their appearance in a list.
- [ ] Individuals are selected in a way that guarantees equal representation from all subgroups.
- [x] Individuals are selected completely at random, with every member of the population has the same probability of being selected

(b) Assess the café owner's plan to get the opinions of the customers.

(3 marks)

*Select **three** boxes.*

- [ ] No customers will return the questionnaires.
- [x] Many customers will not return the questionnaire.
- [x] The customers are only being asked on one day.
- [ ] Questionnaires are not good.
- [x] The customers are not being selected at random.

(c) Here is an open question that the café owner is considering for the questionnaire.

What do you think about the variety of cakes we offer?

Give one reason why this is not a good question.

(1 mark)

*Select **one** box.*

☐ The question is not relevant.

☐ It is an open question.

☐ The responses will not be easy to analyse.

☐ There may be issues with privacy.

(d) Design a suitable closed question for the café owner to use on his questionnaire so that he can decide how many cakes to offer customers.

(2 marks)

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

☐ How amazing was your service today?

☐ Excellent
☐ Great
☐ Good
☐ Not bad

☐ 1 to 3 types
☐ 4 to 6 types
☐ 7 to 9 types
☐ 10 or more types

☐ How many different types of cakes should the café offer to customers?

(e) When the café owner 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.
- [x] A pilot survey will test questions are understood.
- [x] A pilot survey will give an idea of response rate.
- [ ] A pilot survey will give more accurate data.

6 A teacher is interested in how long students spend revising at home each week.

He plans to ask a sample of 30 students in Year 10 to record the number of hours they spend revising in one week.

Each student will write their total revision time on a sheet of paper.

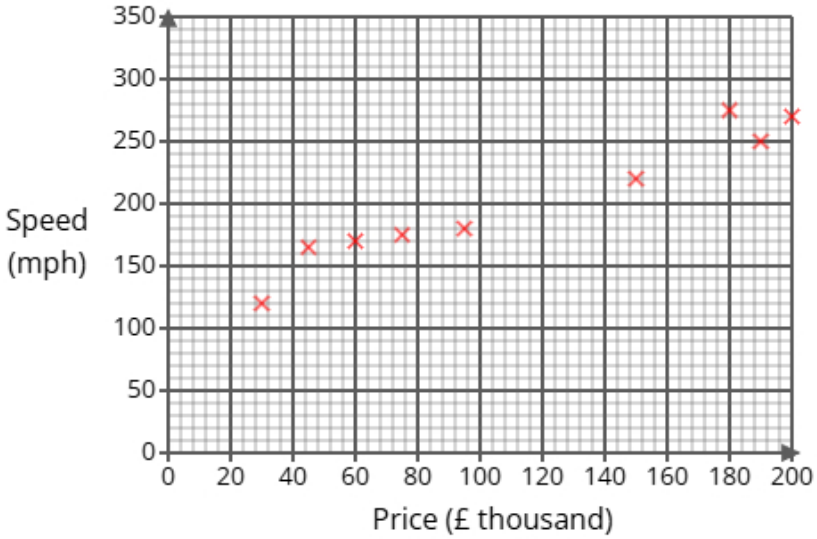
Describe one problem the teacher might face in the statistical enquiry process due to non-response or unexpected results, and explain how he could address this issue.

(2 marks)

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

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

7 Emma recorded data on 11 cars, recording their top speed (in miles per hour) and their price (in thousands of pounds). She represented her findings in the scatter diagram below.



(a) One of the 11 cars has a top speed of 225[s038].

For this car, write down its price.

(1 mark)

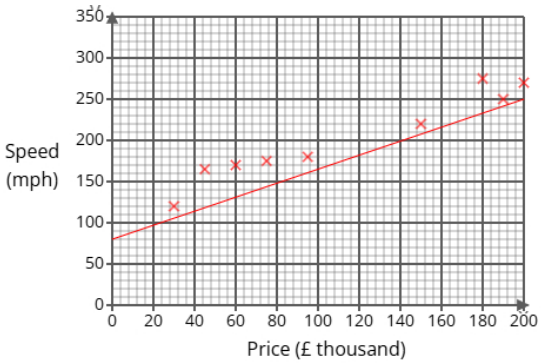
£ 150000

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

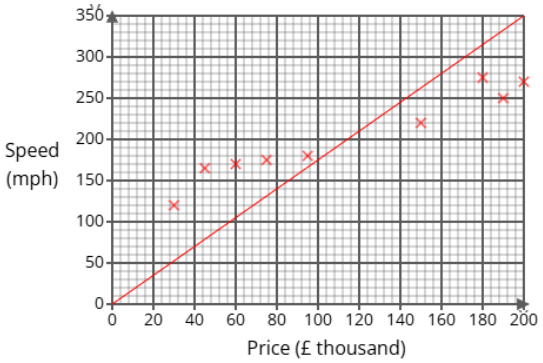
(1 mark)

Select the correct answer.

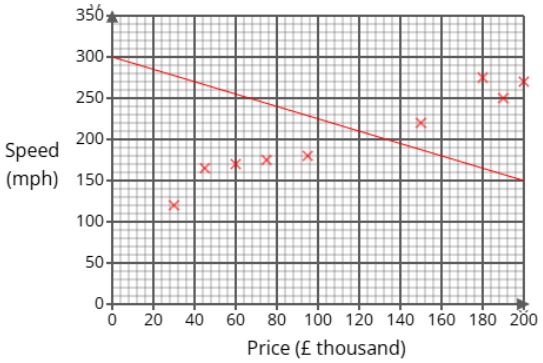
☐



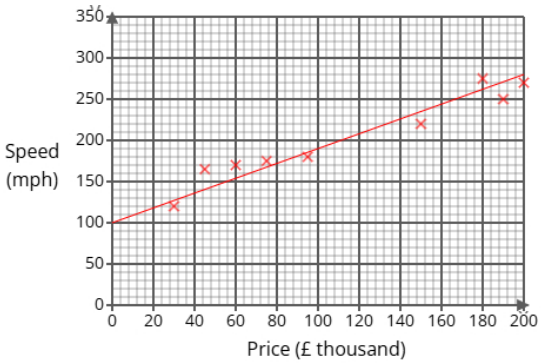
☐



☐



■



(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).*

- [1] The correlation is positive and
- [ ] There is no correlation but it is
- [ ] weak
- [2] strong
- [ ] The correlation is negative and

*Select **one** box.*

- [ ] As the price increases the top speed decreases.
- [ ] A high price car will have a low top speed.
- [■] As the price increases the top speed increases.
- [ ] A high price car will have a high top speed.

(d) A new car will be releasing soon with a price of £250,000.

Emma is planning on using the line of best fit on the scatter diagram to predict the top speed of the new car.

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

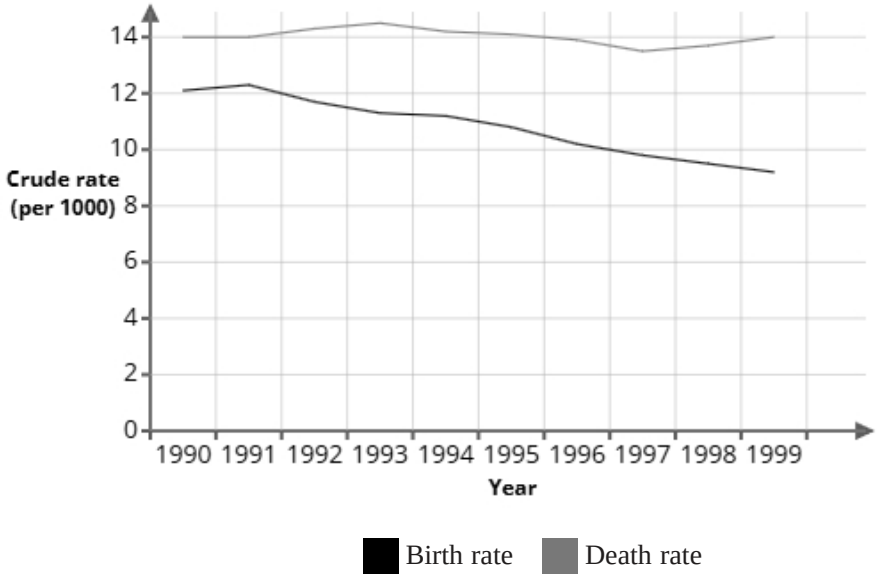
(2 marks)

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

- [ ] because the point is inside the range of the data.
- [2] because the point is outside the range of the data.
- [ ] This is appropriate
- [1] This is not appropriate

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

Crude birth rate and death rate in Hungary from 1990 to 1999



Victor uses the information in the graph to conclude:

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

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

(1 mark)

*Select **one** box.*

- [ ] There are more births than deaths.
- [ ] The graph shows that Hungary is a large country.
- [x] There are more deaths than births.
- [ ] In 1990, the death rate is roughly 2 bigger than the birth rate.

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

(1 mark)

*Select **one** box.*

- [ ] Victor's conclusion does not take into account the average age of the population.
- [ ] The data in the graph may be inaccurate.
- [ ] Victor's conclusion does not take into account where in Hungary these births happened.
- [x] Victor's conclusion does not take into account people moving into Hungary.

(c) In 2000, the population of Hungary was 10 222 254.

There were 135 664 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)

13.3

9 Liam investigates the weights of 180 dogs at a pet shelter.

The weights range from 32 kg to 78 kg.

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

Table A

Weight (w kg)	**Frequency**
$0 < w \leqslant 30$	0
$30 < w \leqslant 60$	87
$60 < w \leqslant 90$	93
$90 < w \leqslant 120$	0
$120 < w \leqslant 150$	0

Table B

Weight (w kg)	**Frequency**
$30 < w \leqslant 40$	5
$40 < w \leqslant 50$	28
$50 < w \leqslant 60$	54
$60 < w \leqslant 70$	59
$70 < w \leqslant 80$	34

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

(2 marks)

*Select **two** boxes.*

- [ ] Grouped data eliminates the need to use any statistical methods.
- [ ] Grouped data keeps the precision in the data.
- [ ] Grouped data makes it easier to identify outliers.
- [x] Grouped data is easier to read.
- [x] Grouped data makes it easier to process large amounts of data.

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

(1 mark)

*Select **one** box.*

- [ ] Grouped data cannot be used for statistical tests.
- [ ] Grouped data may introduce too much complexity when analysing trends.
- [x] Grouped data will lose the accuracy in the data.
- [ ] Grouped data cannot be used to identify patterns.

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

Assess the appropriateness of Liam's claim.

(2 marks)

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

- [ ] In Table A, there are no dogs in three of the groups, so there is a much better spread of data.
- [ ] in Table B, each group has a smaller class width, so detail is lost.
- [x] Liam's claim is justified.
- [x] In Table A, there are no dogs in three of the groups, so these groups are not needed.
- [ ] Liam's claim is not justified.
- [x] in Table B, each group has a smaller class width, showing more detail.

(d) Liam wants to work out the average weight of the 180 dogs at the pet shelter.

He decides to use Table B.

Calculate the average weight of the 180 dogs at the pet shelter, giving your answer to 1 decimal place.

(3 marks)

59.9 kg

10 23 female participants were asked to stack cups before the tower falls and the number of cups stacked was measured.

Here are the results.

34	17	1	35	10	39	15	36
7	36	11	34	10	17	8	35
34	31	17	9	8	26	32	

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

0	
1	
2	
3	

Key: 1|0 = 10

(2 marks)

Select the correct answer.

☐

0	1 7 8 8 9
1	6
2	0 0 1 5 7 7 7
3	1 2 4 4 4 5 5 6 6 9

■

0	1 7 8 8 9
1	0 0 1 5 7 7 7
2	6
3	1 2 4 4 4 5 5 6 6 9

☐

0	0 0 1 5 7 7 7
1	1 7 8 8 9
2	6
3	1 2 4 4 4 5 5 6 6 9

☐

0	1 7 8 8 9
1	0 0 1 5 7 7 7
2	1 2 4 4 4 5 5 6 6 9
3	6

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

(2 marks)

24

(c) A group of 23 male participants were also measured.

The results from the male participants had a median of 15 and an interquartile range of 21.

Emily thinks that these results show that females are better at stacking cups than males.

State whether you agree with Emily and give reasons why.

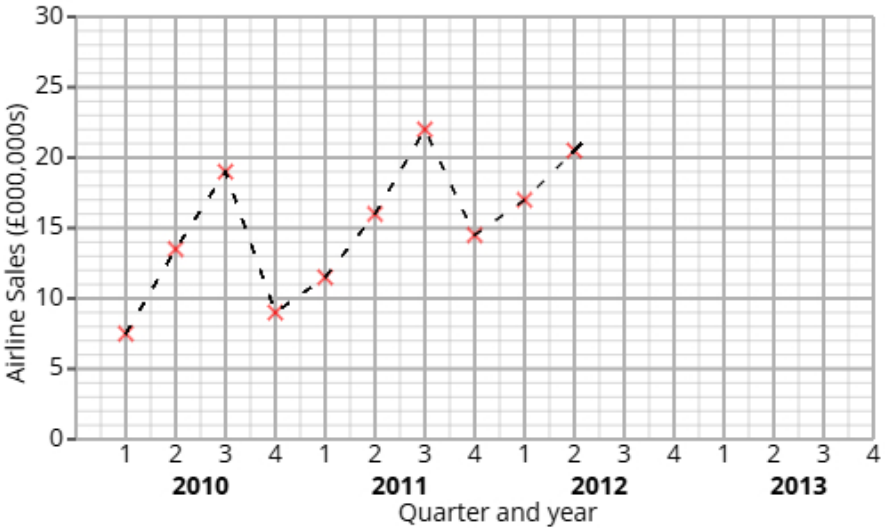
(3 marks)

Median for female participants = 17

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

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

11 The time series graph shows information about the the sales for an airline from 2010 to 2012.



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

12.5 13.5 14 14.5 16 17.5 18.5

(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 fly in the summer.
2	which shows that more people fly in the summer.
	The lowest values are in Q2
1	The greatest values are in Q3

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

(1 mark)

*Select **one** box.*

- [ ] 4-points is more detailed than annual data.
- [ ] It shows correlation.
- [ ] It allows us to predict future values.
- [x] The pattern in the data repeats every four quarters.

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

A fair 5-sided spinner is numbered 1, 2, 3, 4, 5.

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

		5-sided spinner				
		1	**2**	**3**	**4**	**5**
	1	2	3			
3-sided spinner	**2**	3				
	3					

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

Maya plays the game once.

(a) Complete the sample space diagram.

(2 marks)

Select the correct answer.

☐

		5-sided spinner				
		1	**2**	**3**	**4**	**5**
	1	2	3	5	6	7
3-sided spinner	**2**	3	5	6	7	8
	3	5	6	7	8	9

☐

		5-sided spinner				
		1	**2**	**3**	**4**	**5**
	1	2	3	4	5	6
3-sided spinner	**2**	3	4	6	8	10
	3	4	6	9	12	15

■

		5-sided spinner				
		1	**2**	**3**	**4**	**5**
	1	2	3	4	5	6
3-sided spinner	**2**	3	4	5	6	7
	3	4	5	6	7	8

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

(2 marks)

$\frac{2}{5}$

13 Liam organises two fitness programs, Program A and Program B, to help people improve their stamina.

He wants to compare the two programs to see which improves stamina better.

The table shows number of participants who passed and failed the fitness test for each program.

	Passed	Failed	Total
Program A	10	15	25
Program B	12	36	48

(i) Find the relative risk of failing the fitness test having taken Program A compared to Program B.

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

(4 marks)

Write your answer as a decimal.

0.8

*Select **one** box.*

☐ More people failed the fitness test in Program A than in Program B.

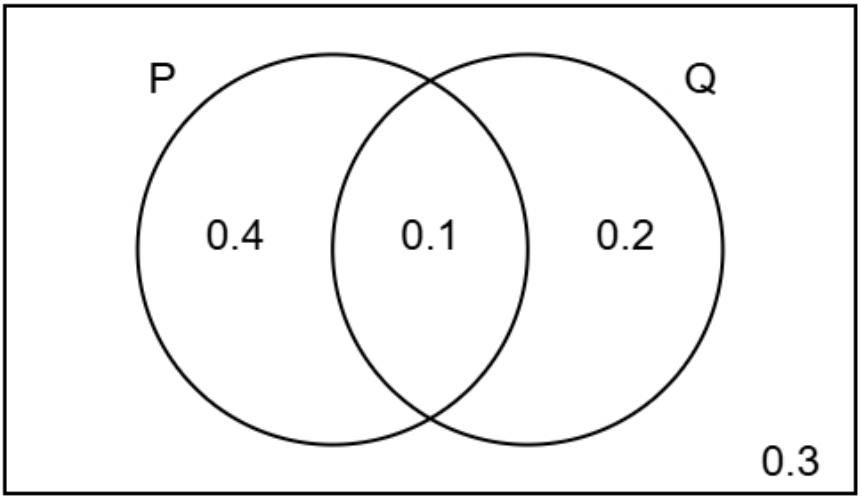
☒ The risk of failing the fitness test having taken Program A is lower than the risk of failing the fitness test having taken Program B.

☐ Less people failed the fitness test in Program A than in Program B.

☐ The risk of failing the fitness test having taken Program A is greater than the risk of failing the fitness test having taken Program B.

14 The Venn diagram shows information about the probabilities of two events occurring.

The events are labelled as P and Q.



(a) Find the probability of event Q happening.

(1 mark)

Leave your answer as a decimal.

0.3

(b) Find P(P and Q)

(1 mark)

Leave your answer as a decimal.

0.1

(c) Find $P(Q \mid P)$

(2 marks)

$\frac{1}{5}$

(d) Two different events events X and Y are independent.

$P(X) = 0.5$

$P(Y) = 0.9$

Find $P(X \text{ and } Y)$

(2 marks)

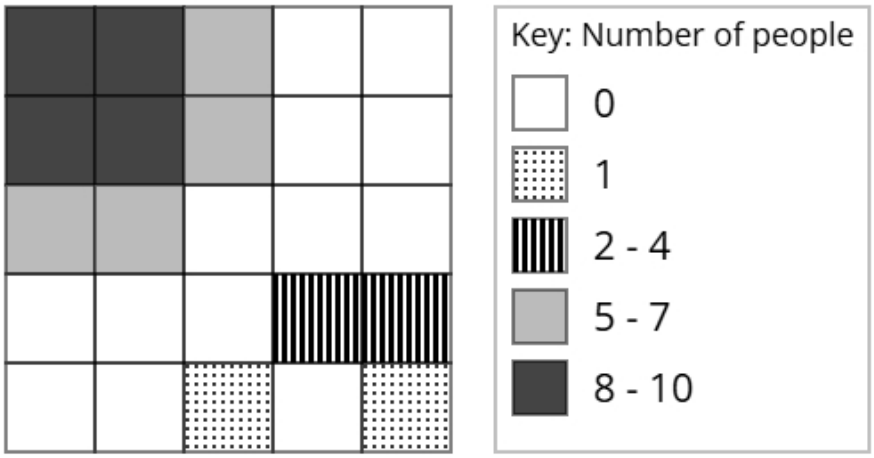
Leave your answer as a decimal.

0.45

15 The choropleth map below represents a shopping centre that has been divided into 25 squares of equal area.

Dave has collected data about the popularity of different parts of the shopping centre.

The number of people recorded in each square on one Tuesday morning is shown.



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

(3 marks)

68

(b) Dave would like to open a smoothie cart in the shopping centre.

After analysing the data, he decides that he should open the smoothie cart in the corner of the shopping centre shown at the top left of the choropleth map.

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

(2 marks)

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

- [ ] because there were less people at the top left of the shopping centre.
- [x] Dave's comment is valid
- [ ] Dave's comment is not valid
- [x] because there were more people at the top left of the shopping centre.

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

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

(1 mark)

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

- [x] because the data was only collected on one Tuesday.
- [x] Aisha is correct
- [ ] Aisha is not correct
- [ ] because there was a large amount of data collected.