statsGCSE

Scaffolded version

Statistics GCSE

Paper 2

Edexcel Foundation - 2026

Foundation Tier

Variant 5

1ST0/2F

Answers

Visit our website for tutorials on each question.

www.statsgcse.com

1

Liam has 6 number tiles shown above.

All of the tiles are placed inside a bag then a random tile is drawn from the bag.

(a) Select the word describes the likelihood that the tile has a 2 on it.

(1 mark)

*Select **one** box.*

- ☐ impossible
- ☐ likely
- ☐ evens
- ■ unlikely

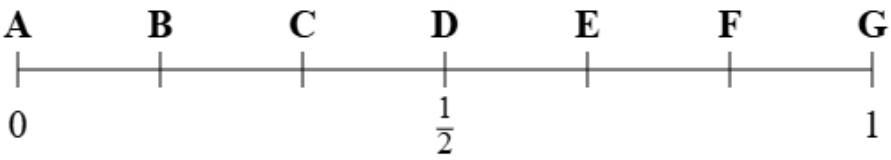
(b) Select the word describes the likelihood that the tile has a number less than 5 on it.

(1 mark)

*Select **one** box.*

- ☐ likely
- ☐ evens
- ☐ unlikely
- ■ certain

(c)



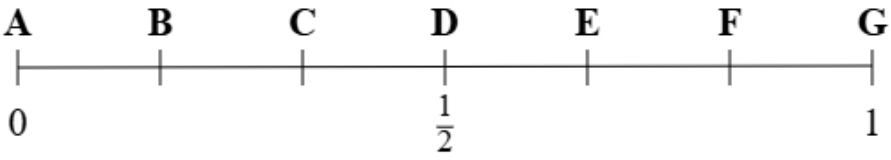
Using the probability scale, write down the letter that shows probability that the tile has a 2 on it.

(1 mark)

*Select **one** box.*

- [x] B
- [ ] A
- [ ] C
- [ ] D

(d)



Using the probability scale, write down the letter that shows probability that the tile has a 1 or a 3 on it.

(1 mark)

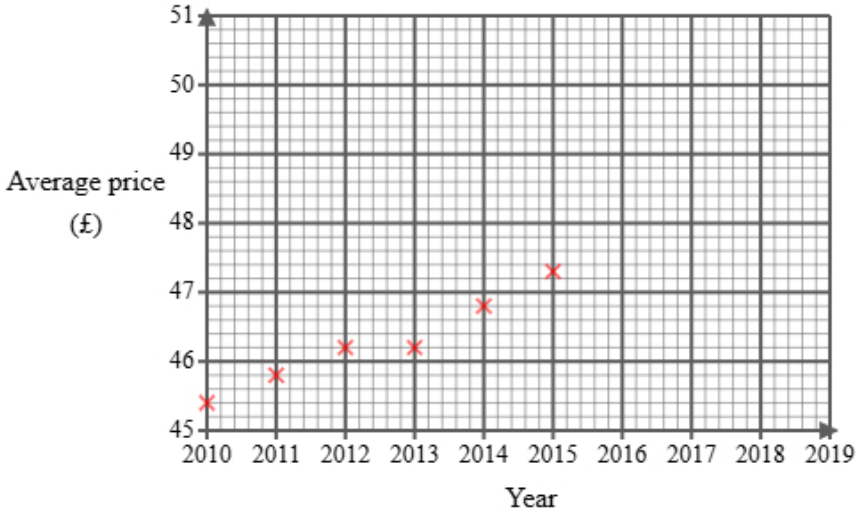
*Select **one** box.*

- [ ] E
- [x] F
- [ ] G
- [ ] D

2 Noah found the following information about the average price of a concert ticket in England.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Average price (£)	45.40	45.80	46.20	46.20	46.80	47.30		47.80	48.00	48.50

He did not find the price for 2016 and has started to draw a graph for the data.



Noah then used statistical software to calculate the equation for the trend line.

(a) (i) Plot the average price for each of 2017, 2018 and 2019

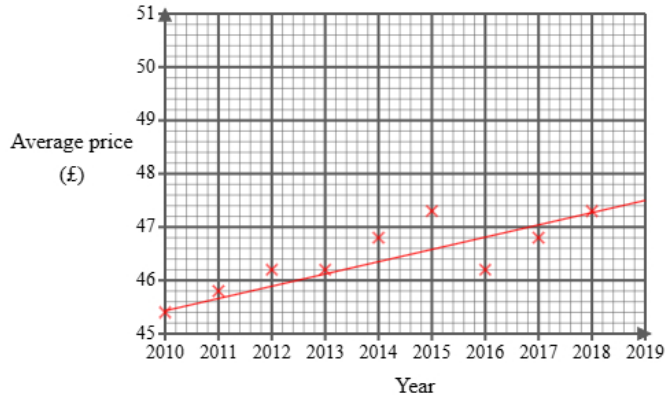
(ii) Draw a trend line for Noah's data

(iii) Describe the trend in the average price of a concert ticket in the UK from 2010 to 2019

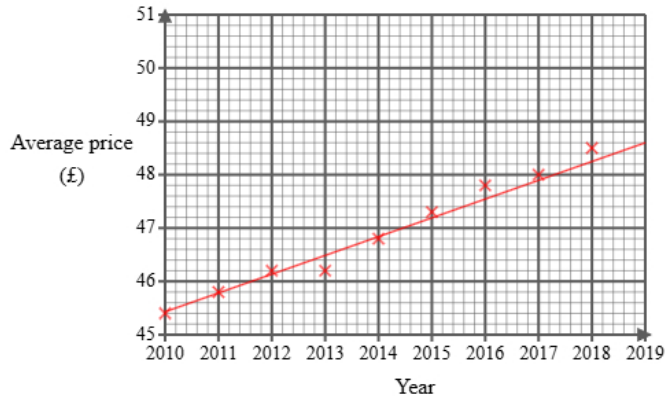
(4 marks)

Select the correct answer.

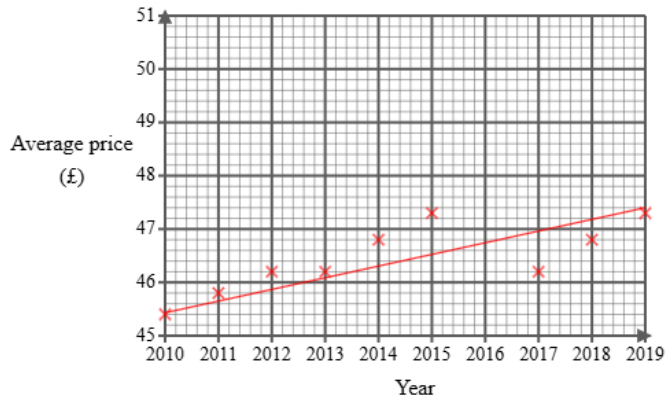
☐



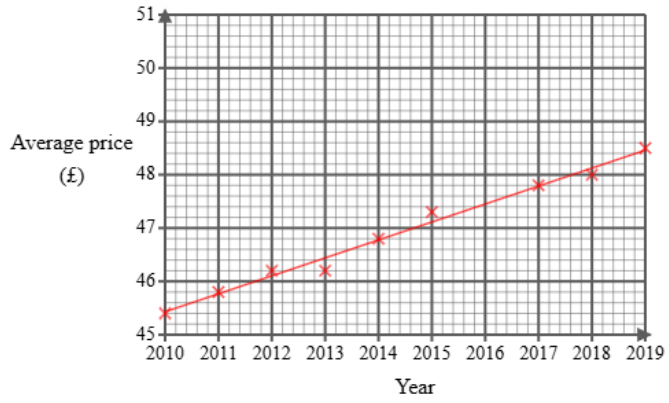
☐



☐



■



Select the correct boxes.

- ☐ Decreasing
- ☐ Negative correlation
- ☐ Positive correlation
- ■ Increasing

(b) The gradient of Noah's trend line is 0.34

Interpret this gradient.

(1 mark)

*Select **one** box.*

- ■ The average price increases per year.
- ☐ The average price increases.
- ☐ The average price started at 34 pence in 2010.
- ☐ The average price decreases.

(c) Explain whether or not the scale used on the 'average price' axis could make the graph misleading.

(2 marks)

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

- ☐ The graph is labelled correctly.
- ■ The 'average price' axis does not start from zero.
- ☐ The graph is not misleading.
- ■ The graph could be misleading.

(d) Noah draws the trend line onto the graph.

He suggests that the trend line could be used to estimate the average price for 2016 **and** 2020

Explain whether each of these estimates would be reliable.

You should **not** work them out.

(3 marks)

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

- [ ] 2016 would not be reliable because it is outside the range of the data.
- [ ] 2020 would be reliable because it is within the data.
- [ ] 2016 would be reliable because it is within the data.
- [ ] 2020 would not be reliable because it is outside the range of the data.

3 An electronics store tracks how many headphones are sold each day. This helps them forecast demand and adjust their inventory levels.

Monday	
Tuesday	
Wednesday	
Thursday	
Friday	

Key:

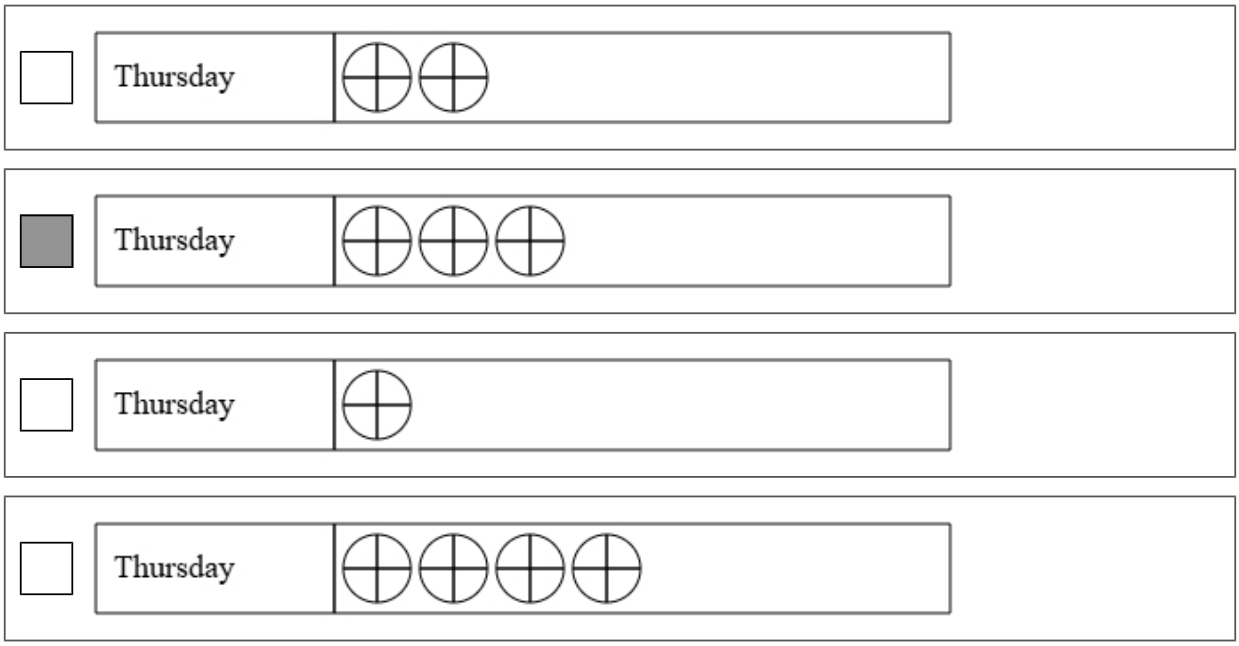
 represents 8 headphones

(a) On Thursday, the number of headphones sold was 24.

Show this information on the pictogram.

(1 mark)

Select the correct answer.



(b) Trent suggests redrawing the pictogram using a key with a whole-circle representing 5 headphones.

Explain why this key would **not** be suitable.

(1 mark)

*Select **one** box.*

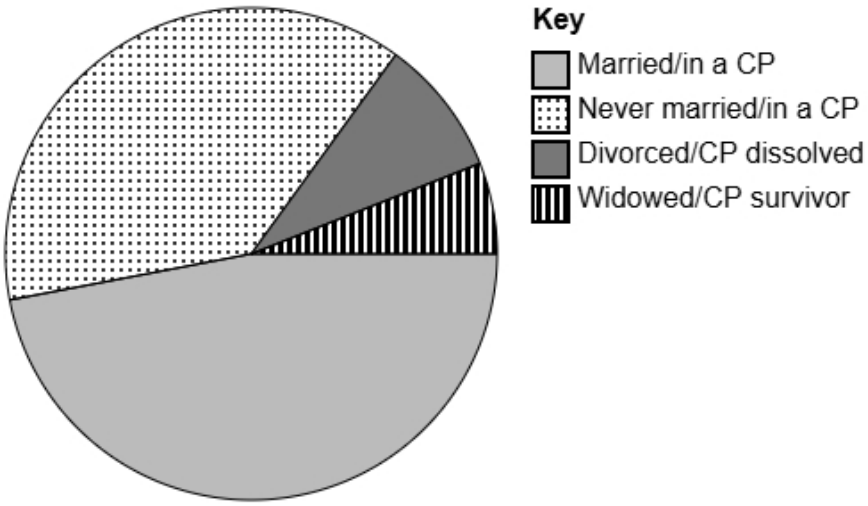
☐ This would be much better because you can fit more circles on.

■ Monday shows 16 loaves of bread. This would be very difficult to show because 16 has a remainder 1 when divided by 5.

☐ The key must always be an even number.

☐ If Trent uses 5 headphones for the key this will lead to less sales being shown.

4 The accurately drawn pie chart shows information about the proportion of adults who are married or in a civil partnership (CP) in England and Wales in 2021.



(a) Explain how you can tell that most adults in 2021 were married or in a civil partnership using the pie chart.

(1 mark)

*Select **one** box.*

☐ 'Married/in a CP' is the first value in the key.

☐ 'Married/in a CP' is the most positive response.

☐ 'Married/in a CP' has the largest sector.

☐ 'Married/in a CP' is at the bottom of the pie chart.

(b) The population in the England and Wales in 2021 was estimated to be 60 million.

Calculate an estimate for the number of adults in England and Wales in 2021 who were 'Never married/in a CP'.

Round your answer to the nearest million.

(2 marks)

23 million

5 Mia owns a computer game store.

She wants to collect information about types of computer games liked by people in her city.

The following list gives the information she is going to collect about people's favourite computer games:

game genre

average play time

number of players

(a) Select the information that is categorical data from the list.

(1 mark)

*Select **one** box.*

- ☐ average play time
- ☐ number of players
- ■ game genre

(b) Select the information that is discrete data from the list.

(1 mark)

*Select **one** box.*

- ☐ game genre
- ☐ average play time
- ■ number of players

(c) Mia would like to send a questionnaire to 40 of her customers.

She has a list of all 200 of her customers.

Explain how Mia can select a systematic sample of 40 people from her list of customers.

(2 marks)

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

- ■ Select every 5th person.
- ☐ Calculate a random number between 1 and 4 to use a starting point.
- ■ Calculate a random number between 1 and 5 to use a starting point.
- ☐ Select every 4th person.

6 Ava is a student and wants to study favourite sports.

She would like to find out the most popular sport in her school.

Ava decides to do convenience sampling outside the school gate an hour after school and uses the data collection sheet shown here:

Sport	Tally
Football	
Basketball	
Tennis	
Swimming	

(a) State the population for this study.

(1 mark)

*Select **one** box.*

- [x] All the students in Ava's school
- [ ] The students she asks
- [ ] All students in the UK
- [ ] A selection of students in Ava's school

(b) Describe what is meant by a convenience sample.

(1 mark)

*Select **one** box.*

- [ ] Sampling people in proportion to a characteristic of the population
- [ ] Sampling people randomly
- [ ] Sampling every nth person from a list
- [x] Sampling people who are easiest to reach

(c) Give **one** disadvantage of convenience sampling.

(1 mark)

*Select **one** box.*

- ☐ It is expensive to do
- ☐ It may be biased
- ☐ It requires a large population
- ☐ It takes too much time

(d) Discuss whether this data collection sheet is appropriate for Ava to collect the data.

(2 marks)

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

- ☐ It will make collecting data very difficult.
- ☐ Ava may not ask enough students.
- ☐ There are too few options.
- ☐ It will make it easy for Ava to analyse the data.

(e) After collecting the data, Ava would like to display the data in a diagram.
Discuss whether or not a stem and leaf diagram would be suitable.

(2 marks)

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

- ☐ because the data is quantitative.
- ☐ because the data is qualitative.
- ☐ It is not suitable
- ☐ It is suitable

7 A speaker wants to get feedback on a seminar on leadership they ran last week.

83 people attended the seminar.

The speaker plans to give a questionnaire to a sample of 20 of the people who attended the seminar.

One of the questions on the questionnaire is:

> To what extent do you agree with the statement, the speaker delivered well?
>
> Use a scale from 0 to 5, where 0 means **strongly disagree** and 5 means **strongly agree**.

(a) Select the word from the list that best describes the type of data collected by this question.

(1 mark)

*Select **one** box.*

- ☐ bivariate
- ☐ continuous
- ■ ordinal

(b) Explain how the speaker could use a list of random numbers to choose a simple random sample of 20 people who attended the seminar.

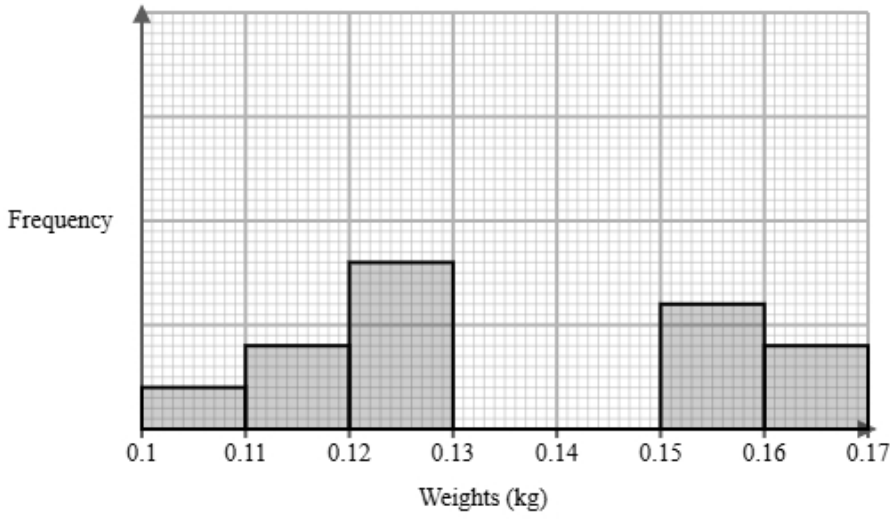
(3 marks)

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

- ■ Assign a number to all the people.
- ☐ Select the people who were assigned the highest number.
- ☐ List the people in alphabetical order based on their surname.
- ■ Select people with the corresponding number from random number list.
- ☐ Select 20 people from a hat.
- ■ Select 20 numbers ignoring any numbers that have repeated or are out of range.

8 A pet shop owner measured the weights of Syrian hamsters and Roborovski hamsters in their pet shop. They recorded the weights after 3 months.

The incomplete histogram and grouped frequency table give information about the weights of Syrian hamsters in their pet shop.



Weights w (kg)	Frequency
$0.10 < w \leqslant 0.11$	2
$0.11 < w \leqslant 0.12$	4
$0.12 < w \leqslant 0.13$	8
$0.13 < w \leqslant 0.14$	12
$0.14 < w \leqslant 0.15$	9
$0.15 < w \leqslant 0.16$	
$0.16 < w \leqslant 0.17$	

(a) Use the information in the histogram to complete the table.

(2 marks)

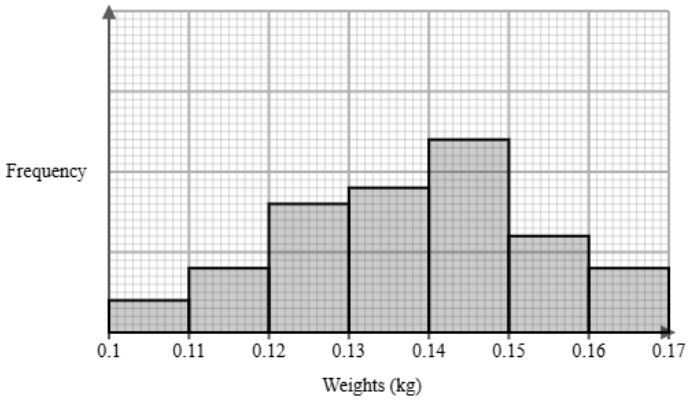
Weights w (kg)	Frequency
$0.10 < w \leqslant 0.11$	2
$0.11 < w \leqslant 0.12$	4
$0.12 < w \leqslant 0.13$	8
$0.13 < w \leqslant 0.14$	12
$0.14 < w \leqslant 0.15$	9
$0.15 < w \leqslant 0.16$	6
$0.16 < w \leqslant 0.17$	4

(b) Use the information in the table to complete the histogram.

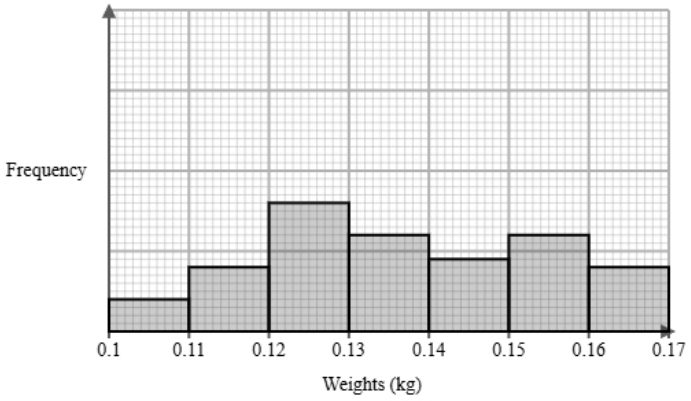
(2 marks)

Select the correct answer.

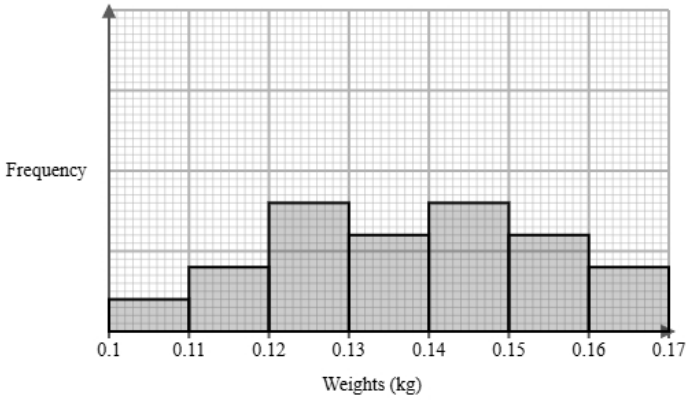
☐



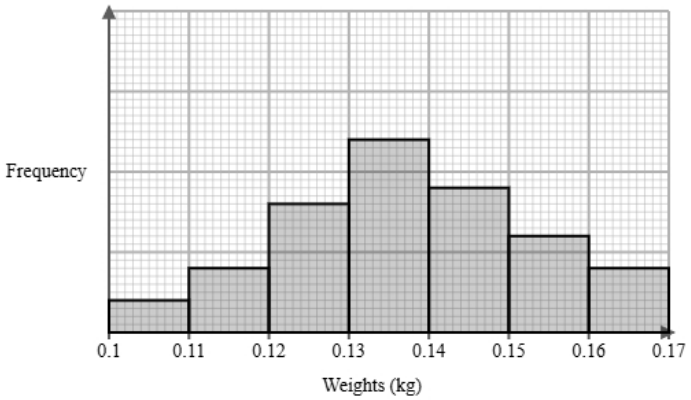
☐



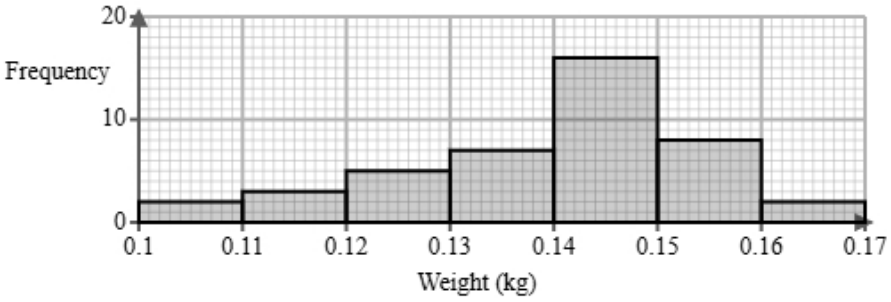
☐



■



(c) The histogram below shows data on the weights of Roborovski hamsters after 3 months.



Identify and interpret the type of skew shown in the histogram for Roborovski hamsters.

(2 marks)

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

- ☐ The mean weight is larger than the median.
- ☐ The mean weight is smaller than the median.
- ☐ Positive skew.
- ☐ Negative skew.

(d) The pet shop owner also measured the weights of Syrian hamsters and Roborovski hamsters in a different pet shop.

The grouped frequency table below gives information about the weights of Syrian hamsters and Roborovski hamsters in a different pet shop.

	Frequency	
Weights w (grams)	Syrian hamsters	Roborovski hamsters
$80 < w \leqslant 100$	3	1
$100 < w \leqslant 120$	10	9
$120 < w \leqslant 140$	17	17
$140 < w \leqslant 160$	5	7
Total	35	34

The estimate of the mean for Syrian hamsters is calculated to be 123.7 g to 1 decimal place.

Omar uses the estimate of the means for Syrian hamsters and Roborovski hamsters to conclude that the Syrian hamsters weigh more than Roborovski hamsters.

Discuss whether or not Omar is correct and give **one** limitation of your conclusion.

You must show your working.

(4 marks)

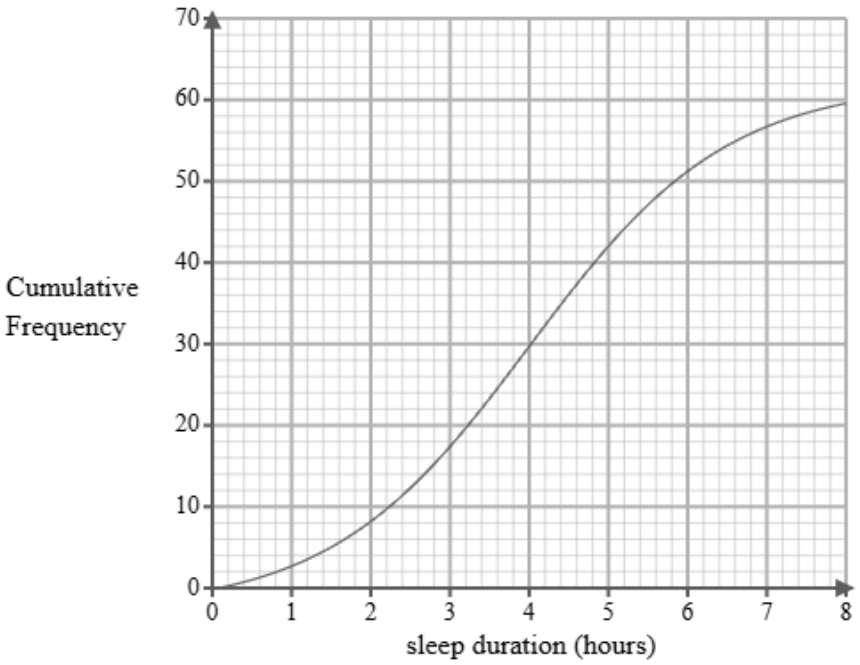
Mean of Roborovski hamsters = 127.6

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

- [ ] Omar is incorrect.
- [ ] We cannot be sure because both means are in the same class interval.
- [ ] Omar is correct.
- [ ] These are only estimates.

9 A researcher measures the number of hours that 60 students sleep on a specific Saturday.

A cumulative frequency diagram is drawn from the data.



Complete the table below from the cumulative frequency diagram.

Lower quartile	Median	Upper quartile

(2 marks)

Select the correct answer.

☐

Lower quartile	Median	Upper quartile
1.5	4.8	6.5

☐

Lower quartile	Median	Upper quartile
3	4	5.2

■

Lower quartile	Median	Upper quartile
2.8	4	5.3

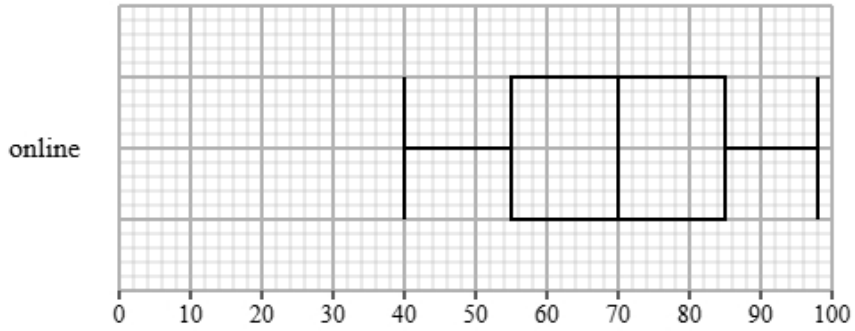
☐

Lower quartile	Median	Upper quartile
2	4.8	5.9

10 Isla collected the satisfaction scores for online and in-store customers.

Both types of shopping experiences were rated by the same number of customers.

The box plot presents data on the satisfaction scores for the online customers.



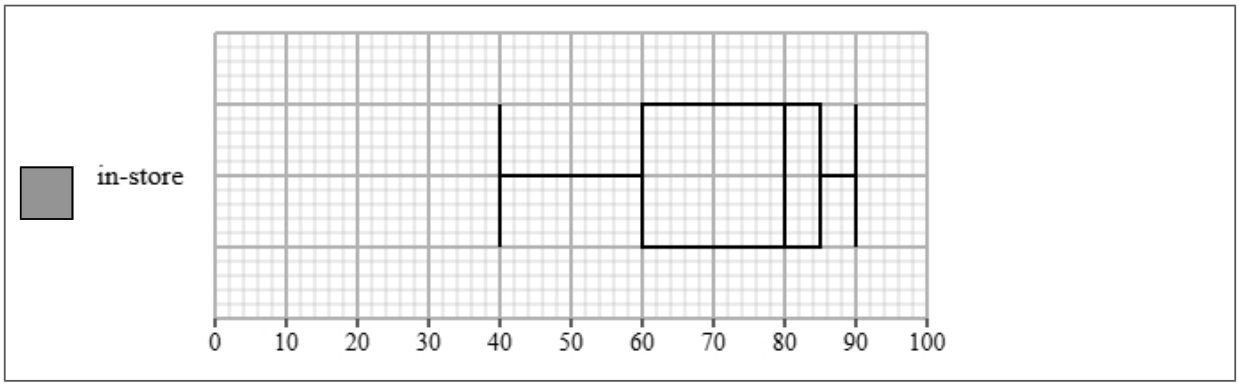
The table gives information about the satisfaction scores for the in-store customers.

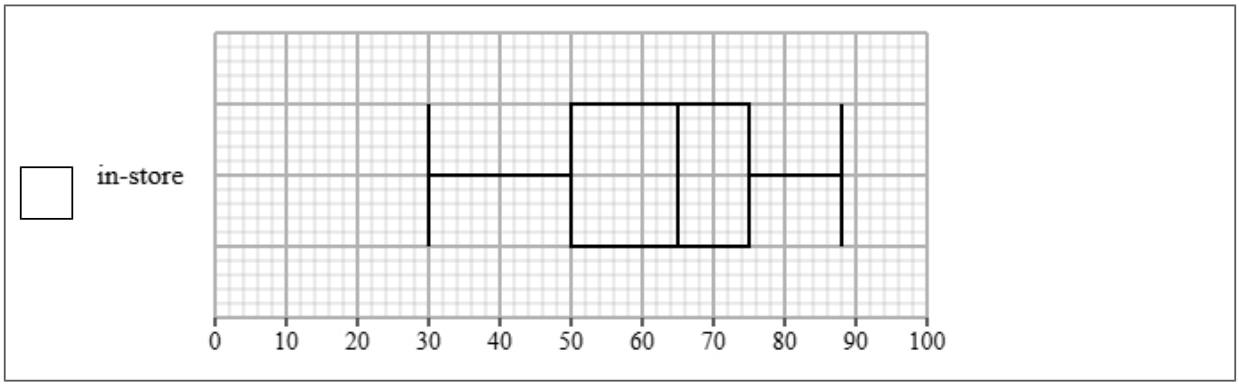
Least tall	Lower quartile	Median	Upper quartile	Most tall
40	60	80	85	90

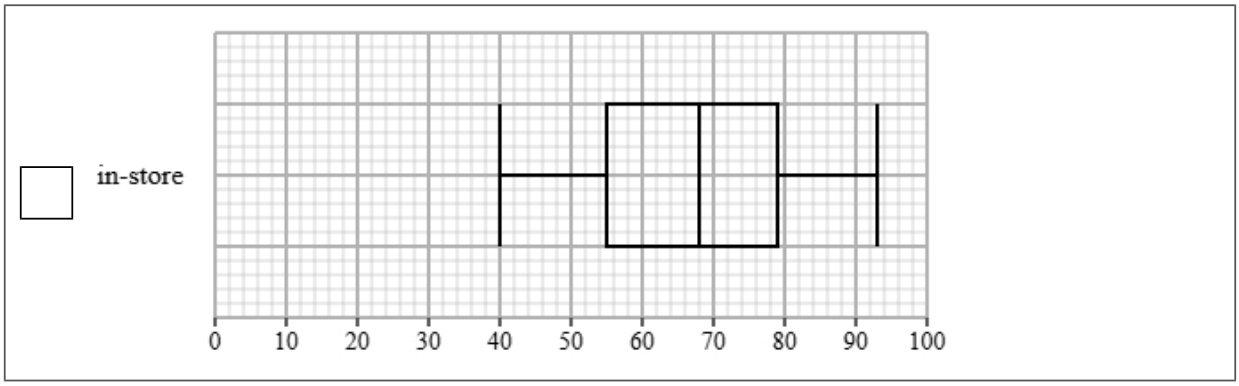
(a) Draw a box plot for the satisfaction scores for the in-store customers.

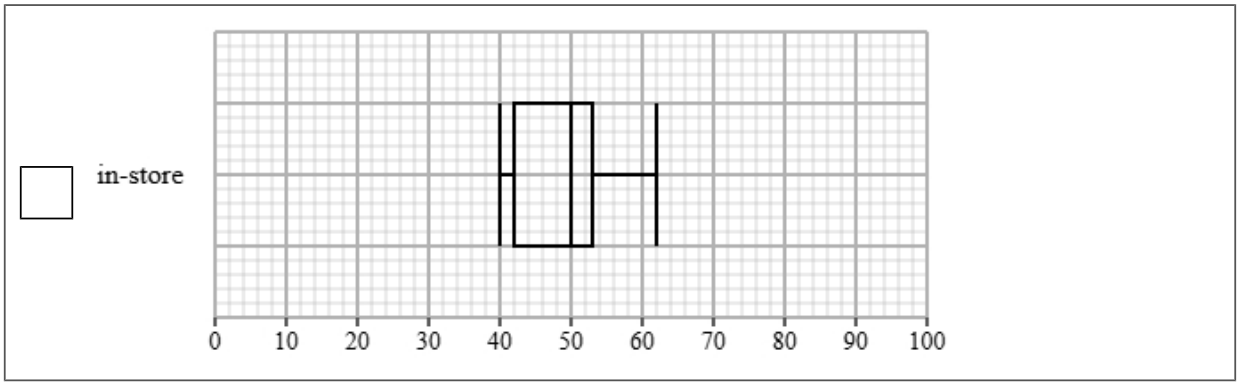
(2 marks)

Select the correct answer.









(b) Compare the two distributions of satisfaction scores.

Give three comparisons and interpret one of these comparisons.

(4 marks)

*Select **one** box.*

- [ ] The median is bigger.
- [ ] The median satisfaction scores for online customers is greater than in-store customers.
- [ ] The median satisfaction scores for online and in-store customers are equal.
- [x] The median satisfaction scores for online customers is lower than in-store customers.

*Select **one** box.*

- [ ] The IQR is bigger.
- [x] The IQR for the satisfaction scores of the online customers is greater than in-store customers.
- [ ] The IQR for the satisfaction scores of the online and in-store customers are equal.
- [ ] The IQR for the satisfaction scores of the online customers is lower than in-store customers.

*Select **one** box.*

- [ ] The skews for the satisfaction scores of the online and in-store customers are both positive.
- [ ] The skews for the satisfaction scores of the online and in-store customers are both symmetrical.
- [x] The skew for the satisfaction scores of the online customers is symmetrical and the skew for the in-store customers is negative.
- [ ] The skew for the satisfaction scores of the online customers is symmetrical and the skew for the in-store customers is positive.

*Select **one** box.*

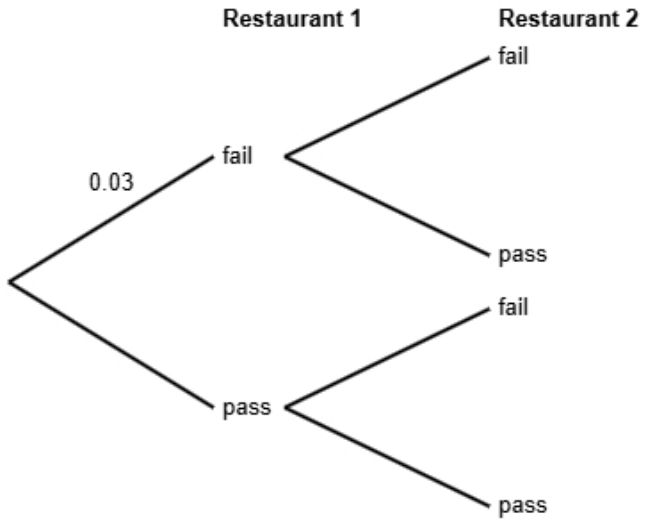
- [ ] The scores for the online customers are less spread out than the in-store customers.
- [ ] The online customers on average have higher satisfaction scores than the in-store customers.
- [ ] The online customers are more skewed than in-store customers.
- [x] The online customers on average have lower satisfaction scores than the in-store customers.

11 Health inspections show that 3% of restaurants in a city fail basic hygiene standards.

The rest pass the required health standards.

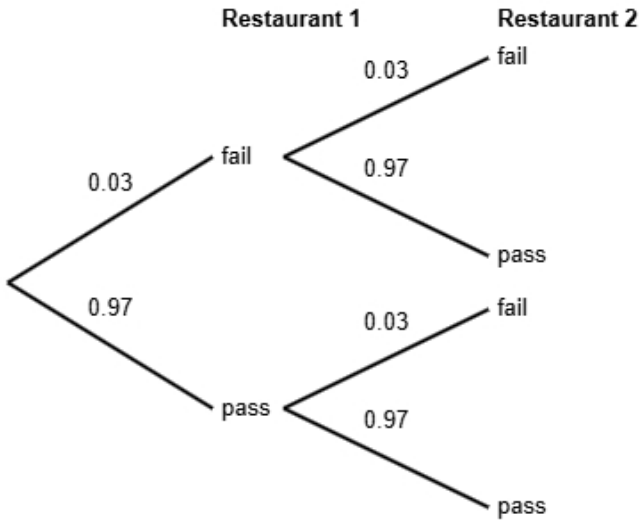
Emma is a food safety officer and will be inspecting two different restaurants.

She does not know if each restaurant will fail or pass.



(a) Complete the probability tree diagram.

(2 marks)



(b) Find the probability that both of Emma's restaurants have an outcome of pass.

(2 marks)

0.9409

(c) Emma states that the probability that exactly one restaurant outcome is fail is less than 6%

Find out whether or not Emma is correct.

(3 marks)

$$\begin{aligned} \text{P(fail AND pass)} &= 0.03 \times 0.97 \\ &= 0.0291 \\ \text{P(exactly one restaurant outcome is fail)} &= 0.0291 \times 2 \\ &= 0.0582 \\ 0.0582 &= 5.82\% \end{aligned}$$

*Select **one** box.*

- [x] The probability that exactly one restaurant outcome is fail is less than 6%, so Emma is correct.
- [ ] The probability that exactly one restaurant outcome is fail is more than 6%, so Emma is not correct.
- [ ] The probability that exactly one restaurant outcome is fail is more than 6%, so Emma is correct.
- [ ] The probability that exactly one restaurant outcome is fail is less than 6%, so Emma is not correct.

12 The table shows information about office spaces in Leeds.

number of desks	number of offices
1	120
2	220
3	740
4	400
5 or more	220
Total	1700

A researcher wants to investigate the price of these office spaces and takes a stratified sample of 85 office spaces according to the number of desks.

(a) The researcher says the mode of the number of desks for these office spaces is 3.

Explain how the researcher knows this.

(1 mark)

*Select **one** box.*

☐ 3 office spaces has the highest frequency.

☐ 3 is the middle number.

☐ 3 is the difference between the largest and smallest number.

☒ 3 desks has the highest frequency.

(b) Work out the number of offices in the sample for each number of desks.

number of desks	number of offices in the sample
1	
2	
3	
4	
5 or more	

(3 marks)

number of desks	number of offices in the sample
1	6
2	11
3	37
4	20
5 or more	11

(c) Describe how the 85 office spaces in the sample should be selected.

(3 marks)

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

- [x] Use a sampling frame for each strata.
- [x] Number each of the office spaces, and then use the random numbers to select the required amount of office spaces.
- [ ] Complete two of the strata.
- [ ] Select the first 85 office spaces.
- [x] Generate random numbers, remove repeats or numbers out of range.
- [ ] Ensure that all 1700 office spaces are included in the sample.

13 Mei works for a marine conservation group. She has been tasked with investigating plastic waste collected by divers.

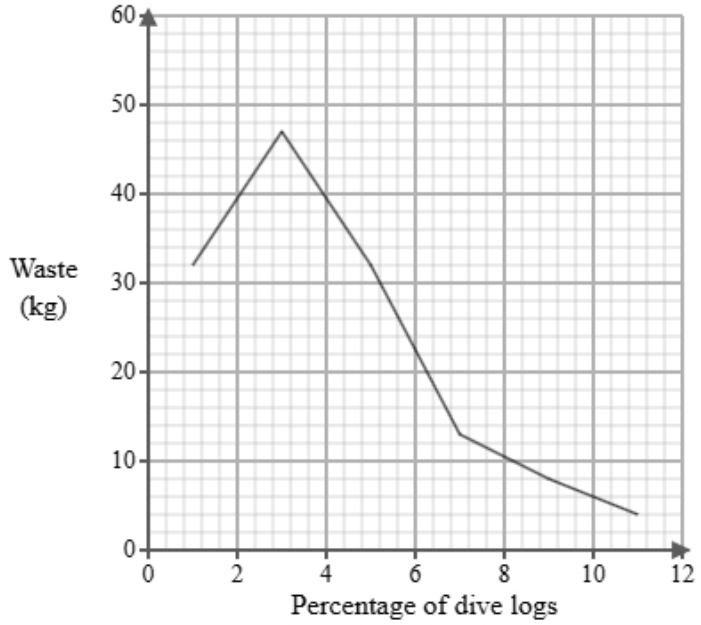
Below is a section of the spreadsheet she used to record her findings.

Waste (kg)	Percentage of dive logs
$0 < p \leqslant 2$	7
$2 < p \leqslant 4$	seven
$4 < p \leqslant 6$	6
$6 < p \leqslant 8$	118
$8 < p \leqslant 10$	47
$10 < p \leqslant 12$	15
Total	100

Mei cleans the data to create the table below.

Waste (kg)	Percentage of dive logs
$0 < p \leqslant 2$	7
$2 < p \leqslant 4$	7
$4 < p \leqslant 6$	6
$6 < p \leqslant 8$	18
$8 < p \leqslant 10$	47
$10 < p \leqslant 12$	15
Total	100

A frequency polygon has been drawn for metal waste collected by divers.



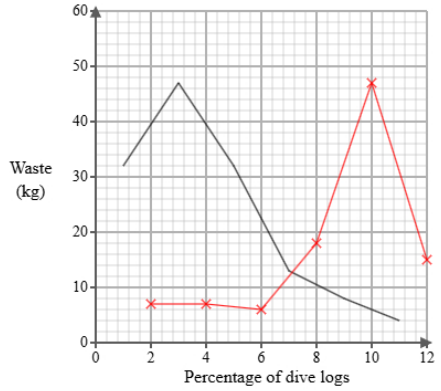
i) On the same graph, draw the frequency polygon for plastic waste collected by divers.

ii) Using the two frequency polygons, compare the skew of the distributions and explain what your comparison means in context.

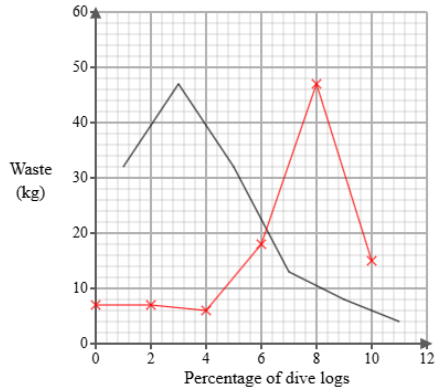
(4 marks)

Select the correct answer.

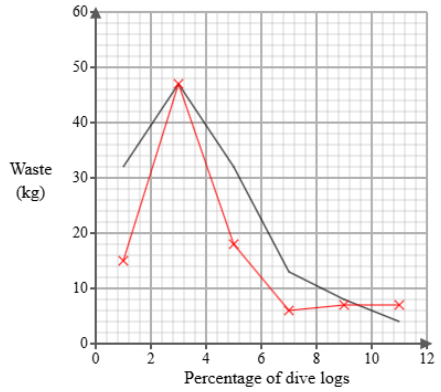
☐



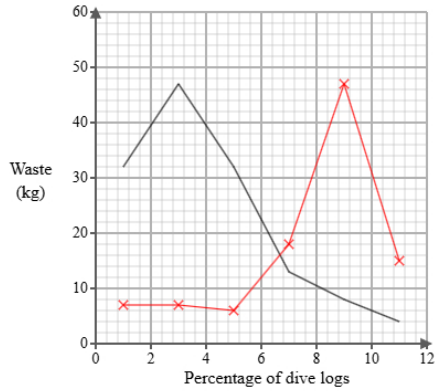
☐



☐



■



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

- [ ] The distribution of metal waste collected by divers is negatively skewed whereas the distribution of plastic waste collected by divers is positively skewed.
- [x] The distribution of metal waste collected by divers is positively skewed whereas the distribution of plastic waste collected by divers is negatively skewed.
- [ ] This means that the metal waste collected by divers was mainly at the upper end of the distribution and the plastic waste collected by divers was mainly at the lower end of the distribution.
- [ ] The distribution of metal waste collected by divers is symmetrical whereas the distribution of plastic waste collected by divers is positively skewed.
- [ ] This means that the metal waste collected by divers was equally spread out on either side of the median and the plastic waste collected by divers was mainly at the upper end of the distribution.
- [x] This means that the metal waste collected by divers was mainly at the lower end of the distribution and the plastic waste collected by divers was mainly at the upper end of the distribution.

14 The table shows information about the consumer price index (CPI) and average cost of bread (pence) in the United Kingdom for Jan 1990, Jan 2000 and Jan 2010.

	Jan 1990	**Jan 2000**	**Jan 2010**
consumer price index	100	130	160
average cost of bread (pence)	65	70	110

Describe how the increase in average cost of bread (pence) compares with the CPI over the ten years to Jan 2000 and over the twenty years to Jan 2010.

(5 marks)

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

- ☐ $\frac{70}{130} \times 100 = 54$ (nearest integer)
- ■ Between Jan 1990 and Jan 2000 the change in price was less than the CPI.
- ☐ $\frac{110}{160} \times 100 = 69$ (nearest integer)
- ☐ Between Jan 1990 and Jan 2010 the change in price was less than the CPI.
- ■ $\frac{110}{65} \times 100 = 169$ (nearest integer)
- ☐ Between Jan 1990 and Jan 2000 the change in price was more than the CPI.
- ■ $\frac{70}{65} \times 100 = 108$ (nearest integer)
- ■ Between Jan 1990 and Jan 2010 the change in price was more than the CPI.

15 Emma is investigating how the engine hours used, x, affects the resale price (£), y for two types of boats, brand A and brand B.

She found ten boats of each type and recorded their engine hours used and resale price and plotted each on scatter diagrams.

She then drew a line of best fit on each diagram and found the gradient and y-intercept of each line.

Here are the results:

Brand	Gradient of line of best fit	y-intercept of line of best fit
A	-200	30000
B	-150	38000

Interpret and compare these results in context.

(5 marks)

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

- [x] Brand A reduces in resale price by £200 per engine hour.
- [ ] Brand A has a greater initial resale price.
- [ ] Brand A changes in resale price by £30000 per engine hour.
- [x] Both boats decrease in resale price as the engine hours used increase.
- [x] Brand B reduces in resale price by £150 per engine hour.
- [ ] Brand B changes in resale price by £38000 per engine hour.
- [ ] Both boats increase in resale price as the engine hours used increase.
- [ ] Brand A reduces in resale price less than Brand B.
- [x] Brand A reduces in resale price more per engine hour than Brand B.
- [x] Brand B has a greater initial resale price.