

**Statistics GCSE****Paper 2**

Edexcel Foundation - 2026

Foundation Tier

Variant 2

1ST0/2F

**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



Ava 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 1 on it.

(1 mark)

Impossible – This means it cannot happen at all.

Unlikely – This means it could happen, but it probably won't.

Evens – This means it has a 50/50 chance of happening – it's just as likely to happen as not.

Likely – This means it will probably happen, but it's not guaranteed.

Certain – This means it will definitely happen – there's no doubt at all.

Select **one** box.

- unlikely
- impossible
- likely
- evens

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

(1 mark)

Impossible – This means it cannot happen at all.

Unlikely – This means it could happen, but it probably won't.

Evens – This means it has a 50/50 chance of happening – it's just as likely to happen as not.

Likely – This means it will probably happen, but it's not guaranteed.

Certain – This means it will definitely happen – there's no doubt at all.

Select **one** box.

likely

evens

impossible

unlikely

(c)



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

(1 mark)

Select **one** box.

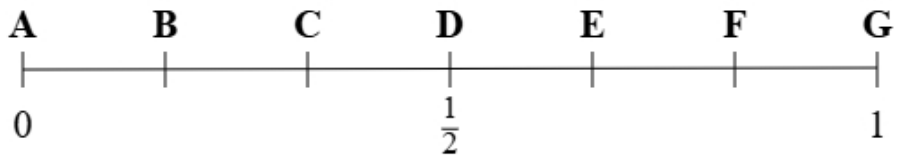
C

B

A

D

(d)



Using the probability scale, write down the letter that shows probability that the tile has a 1 or a 2 on it.

(1 mark)

Select *one* box.

E

F

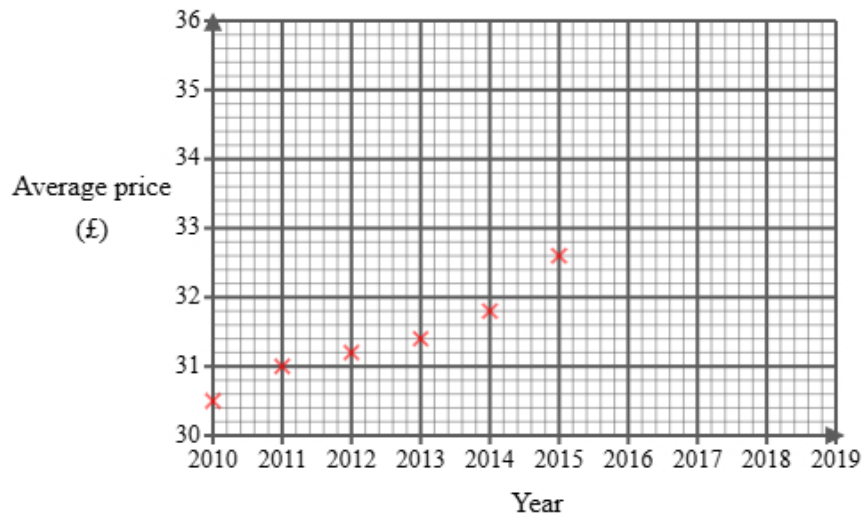
G

D

2 Sofia found the following information about the average price of a gym membership in England.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Average price (£)	30.50	31.00	31.20	31.40	31.80	32.60		33.80	34.00	34.50

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

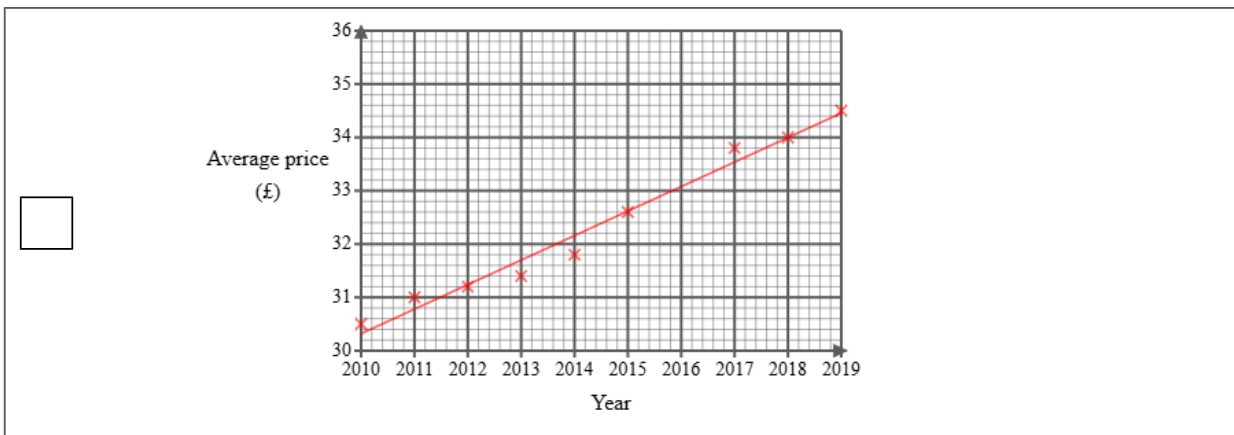
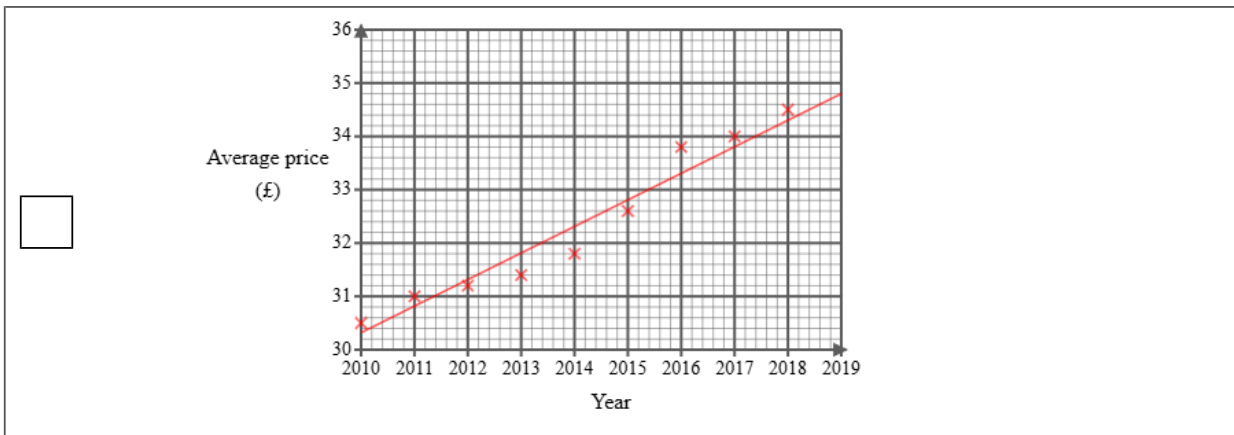
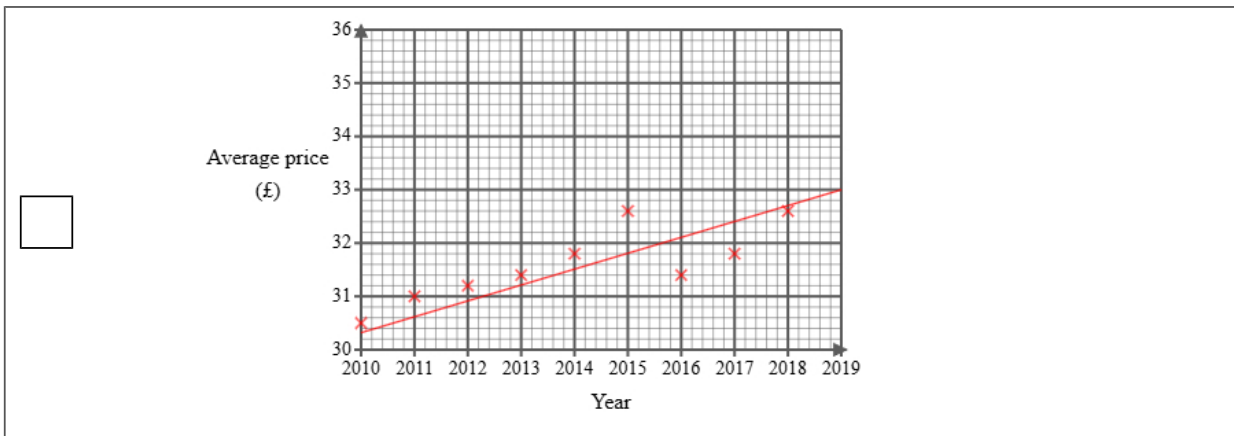
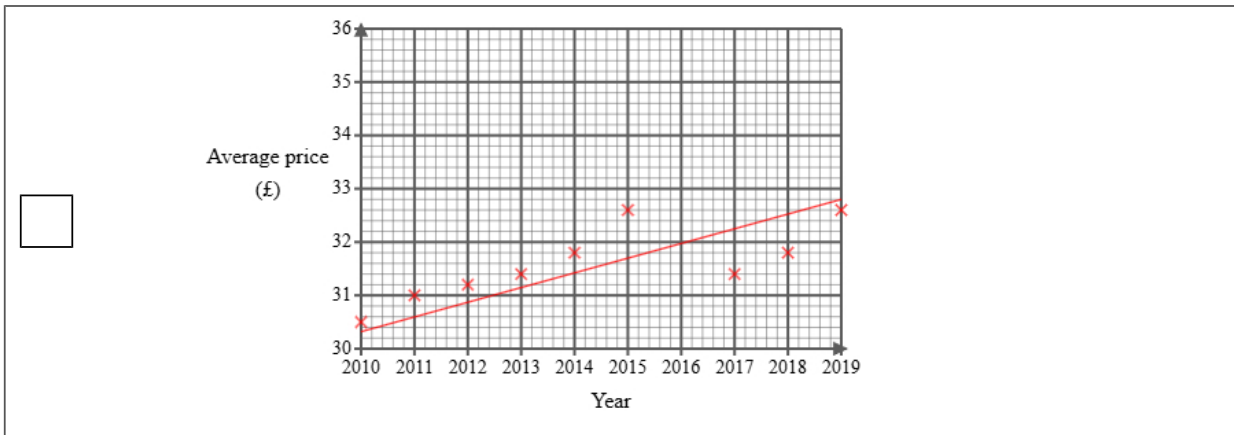


Sofia then used statistical software to calculate the equation for the trend line.

- (a) (i) Plot the average price for each of 2017, 2018 and 2019  
(ii) Draw a trend line for Sofia's data  
(iii) Describe the trend in the average price of a gym membership in the UK from 2010 to 2019

(4 marks)

Select the correct answer.



Select the correct boxes.

- Increasing
- Negative correlation
- Positive correlation
- Decreasing

(b) The gradient of Sofia's trend line is 0.46

Interpret this gradient.

(1 mark)

Select **one** box.

- The average price started at 46 pence in 2010.
- The average price increases per year.
- The average price increases.
- 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 graph could be misleading.
- The graph is not misleading.
- The 'average price' axis does not start from zero.

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

She 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 be reliable because it is within the data.

2020 would be reliable because it is within the data.

2020 would not be reliable because it is outside the range of the data.

2016 would not be reliable because it is outside the range of the data.

- 3 A bookstore monitors the number of novels sold each day of the week. This helps them track which type of books are more popular at different times.

Monday	
Tuesday	
Wednesday	
Thursday	
Friday	

**Key:**

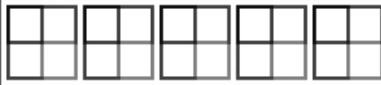





represents 8 novels

- (a) On Thursday, the number of novels sold was 32.  
Show this information on the pictogram.

(1 mark)

Select the correct answer.

<input type="checkbox"/>	Thursday	
<input type="checkbox"/>	Thursday	
<input type="checkbox"/>	Thursday	
<input type="checkbox"/>	Thursday	

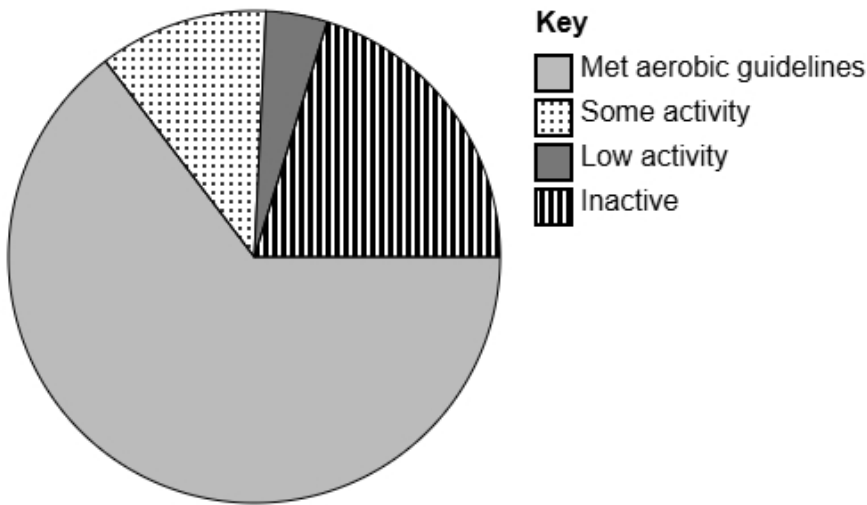
- (b) Kyler suggests redrawing the pictogram using a key with a whole-square representing 5 novels.  
Explain why this key would **not** be suitable.

(1 mark)

Select **one** box.

- If Kyler uses 5 novels for the key this will lead to less sales being shown.
- This would be much better because you can fit more squares on.
- The key must always be an even number.
- Monday shows 16 loaves of bread. This would be very difficult to show because 16 has a remainder 1 when divided by 5.

- 4 The accurately drawn pie chart shows information about how many people in England met the guidelines for aerobic activity levels in 2021.



- (a) Explain how you can tell that most people met the aerobic guidelines in England in 2021 using the pie chart.

(1 mark)

Select **one** box.

- 'Met aerobic guidelines' is the first value in the key.
- 'Met aerobic guidelines' has the largest sector.
- 'Met aerobic guidelines' is the most positive response.
- 'Met aerobic guidelines' is at the bottom of the pie chart.

- (b) The population in England in 2021 was estimated to be 56 million.

Calculate an estimate for the number of people in the UK in 2021 who 'Met aerobic guidelines'.

Round your answer to the nearest million.

(2 marks)

Start by measuring the angle with a protractor.

Next find the amount. Remember that there are 360 degrees in a circle.

\_\_\_\_\_ million

5 Tom owns a sports club.

He wants to collect information about types of sports liked by people in his town.

The following list gives the information he is going to collect about people's favourite sports:

sport type

average match duration

number of players per team

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

(1 mark)

Categorical data can be grouped into non-overlapping categories.

Select **one** box.

sport type

number of players per team

average match duration

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

(1 mark)

Discrete data can take one of a set of certain values.

Select **one** box.

sport type

average match duration

number of players per team

(c) Tom would like to send a questionnaire to 50 of his customers.

He has a list of all 200 of his customers.

Explain how Tom can select a systematic sample of 50 people from his list of customers.

(2 marks)

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

- Select every 4th 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 5th person.

6 Noah is a student and wants to study types of pets owned.

He would like to find out the most common pet in his school.

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

Type of pet	Tally
Dog	
Cat	
Fish	
Hamster	

(a) State the population for this study.

(1 mark)

Select **one** box.

- All students in the UK
- The students he asks
- All the students in Noah's school
- A selection of students in Noah'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 who are easiest to reach
- Sampling people randomly
- Sampling every nth person from a list

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

(1 mark)

Select **one** box.

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

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

(2 marks)

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

- Noah will be able to put the data into a graph very easily.
- Noah may not ask all the students.
- Some students may not have a pet.
- It will make collecting data very difficult.

(e) After collecting the data, Noah 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.
- It is suitable
- It is not suitable
- because the data is qualitative.

7 A tutor wants to get feedback on an online course on coding they ran last week.

112 people attended the course.

The tutor plans to give a questionnaire to a sample of 20 of the people who attended the course.

One of the questions on the questionnaire is:

To what extent do you agree with the statement, the tutor was helpful?

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.

ordinal

continuous

bivariate

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

(3 marks)

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

Select people with the corresponding number from random number list.

Select the people who were assigned the highest number.

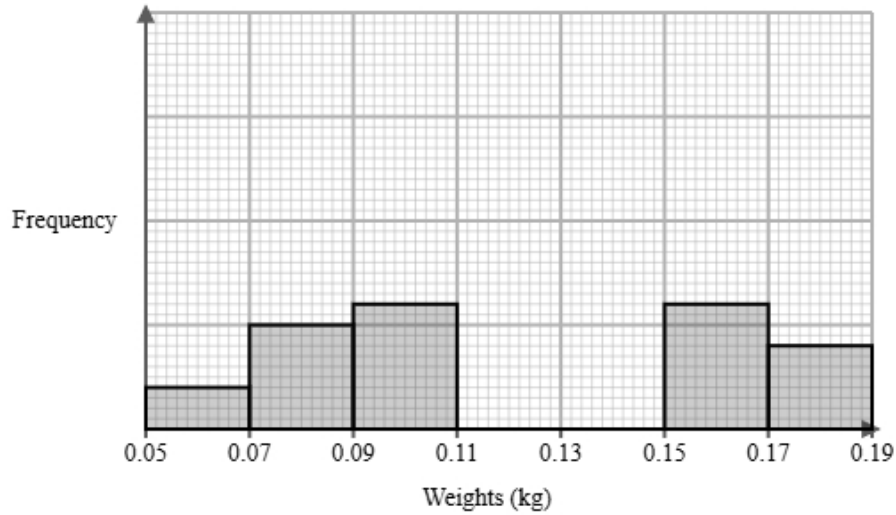
Assign a number to all the people.

List the people in alphabetical order based on their surname.

Select 20 people from a hat.

Select 20 numbers ignoring any numbers that have repeated or are out of range.

- 8 A farm owner measured the weights of Emperor carrots and Danvers carrots in farm plot 1. They recorded the weights after 4 months. The incomplete histogram and grouped frequency table give information about the weights of Emperor carrots in farm plot 1.



Weights $w$ (kg)	Frequency
$0.05 < w \leq 0.07$	2
$0.07 < w \leq 0.09$	5
$0.09 < w \leq 0.11$	6
$0.11 < w \leq 0.13$	12
$0.13 < w \leq 0.15$	8
$0.15 < w \leq 0.17$	
$0.17 < w \leq 0.19$	

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

(2 marks)

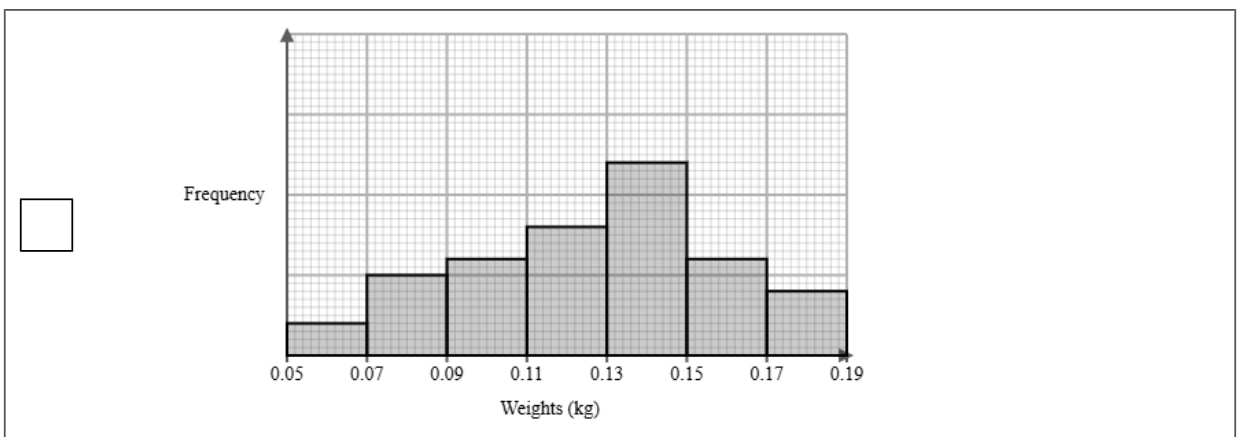
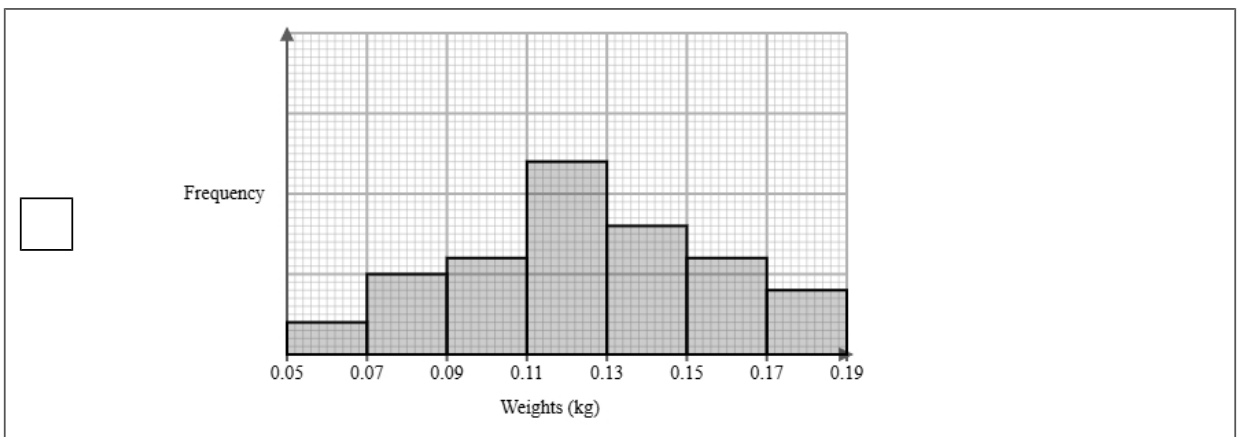
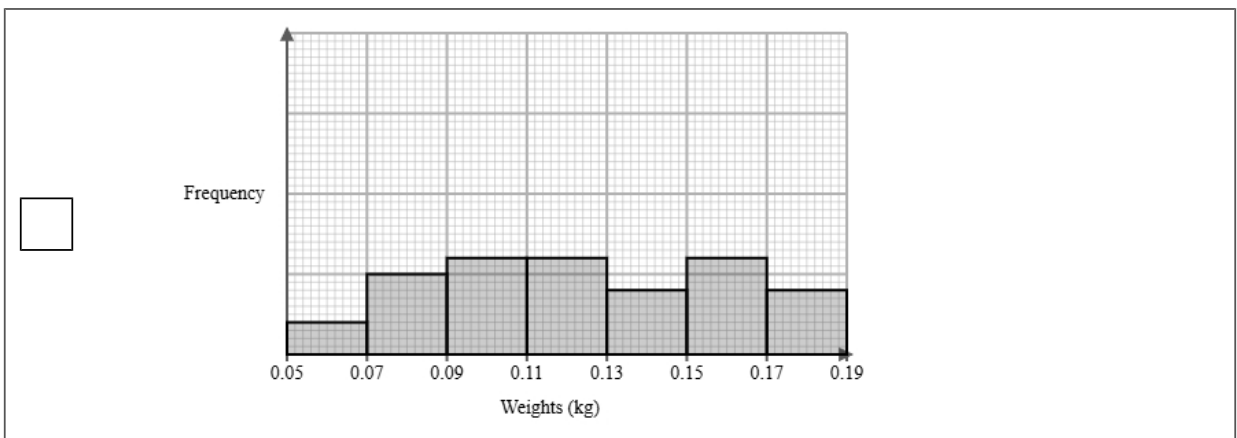
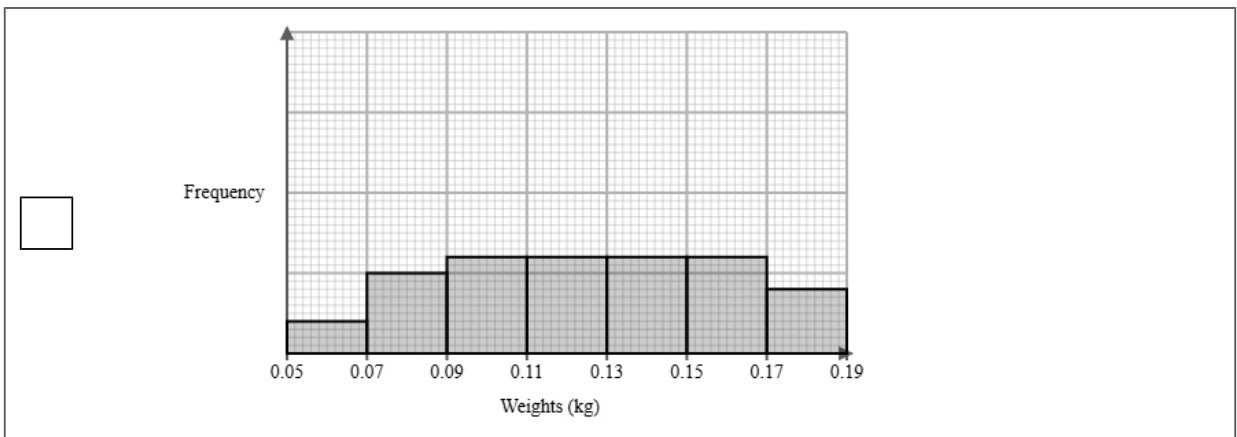
Find the height of  $0.05 < w \leq 0.07$  on the graph and compare it with the frequency to find the scale  
Use this scale to find the missing frequencies

Weights $w$ (kg)	Frequency
$0.05 < w \leq 0.07$	2
$0.07 < w \leq 0.09$	5
$0.09 < w \leq 0.11$	6
$0.11 < w \leq 0.13$	12
$0.13 < w \leq 0.15$	8
$0.15 < w \leq 0.17$	_____
$0.17 < w \leq 0.19$	_____

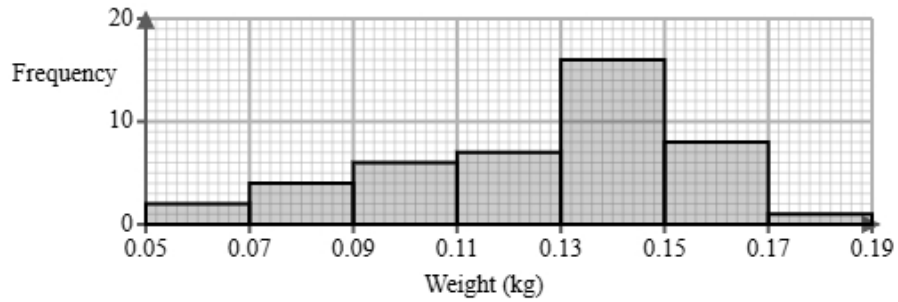
**(b)** Use the information in the table to complete the histogram.

(2 marks)

Select the correct answer.



(c) The histogram below shows data on the weights of Danvers carrots after 4 months.



Identify and interpret the type of skew shown in the histogram for Danvers carrots.

(2 marks)

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

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

- (d) The farm owner also measured the weights of Emperor carrots and Danvers carrots in farm plot 2. The grouped frequency table below gives information about the weights of Emperor carrots and Danvers carrots in farm plot 2.

Weights $w$ (grams)	Frequency	
	Emperor carrots	Danvers carrots
$10 < w \leq 50$	6	2
$50 < w \leq 90$	11	7
$90 < w \leq 130$	16	15
$130 < w \leq 170$	6	3
Total	39	27

The estimate of the mean for Emperor carrots is calculated to be 92.6 g to 1 decimal place.

Eric uses the estimate of the means for Emperor carrots and Danvers carrots to conclude that the Emperor carrots weigh more than Danvers carrots.

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

You must show your working.

(4 marks)

Add a midpoint and  $fw$  column onto the table

Add up the frequencies and the  $fw$  columns

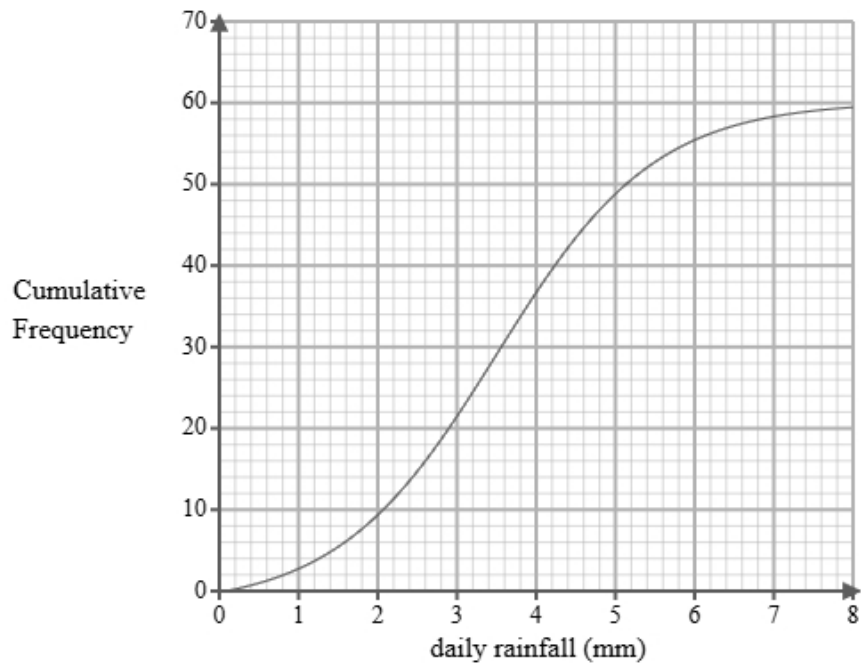
Find the mean by dividing  $\sum fw$  and  $\sum f$

Mean of Danvers carrots = \_\_\_\_\_

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

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

- 9 A researcher measures the daily rainfall, in millimetres, recorded over 60 days in a small town. 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.

<input type="checkbox"/>	Lower quartile	Median	Upper quartile
	1	4.1	5.9

<input type="checkbox"/>	Lower quartile	Median	Upper quartile
	1.8	4.1	5.3

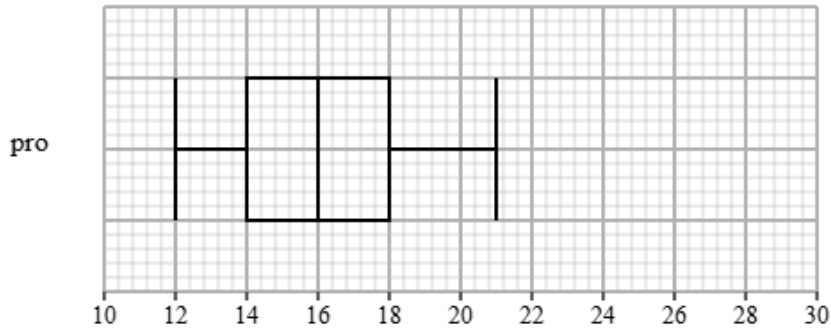
<input type="checkbox"/>	Lower quartile	Median	Upper quartile
	3.1	3.6	4.1

<input type="checkbox"/>	Lower quartile	Median	Upper quartile
	2.5	3.6	4.6

10 Liam recorded the completion times for pro and beginner runners in a 5K race.

Both groups ran the same course.

The box plot presents data on the completion times for the pro runners.



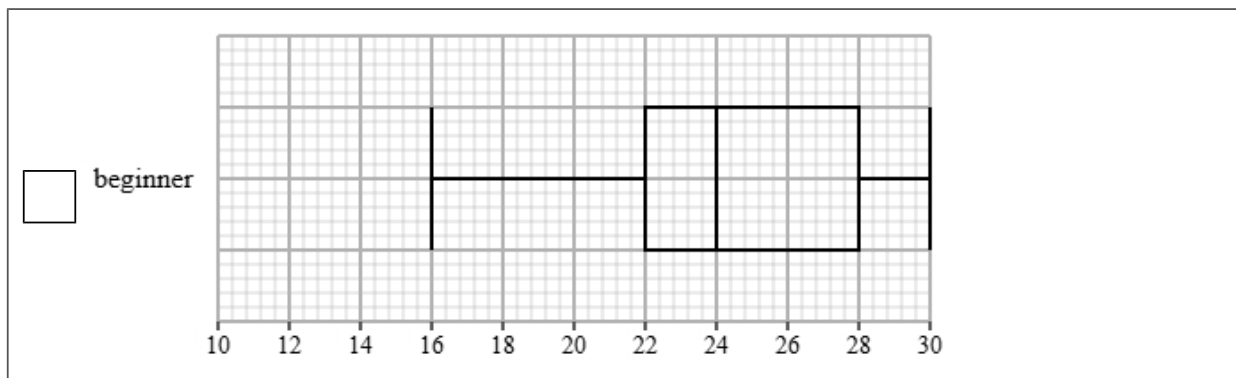
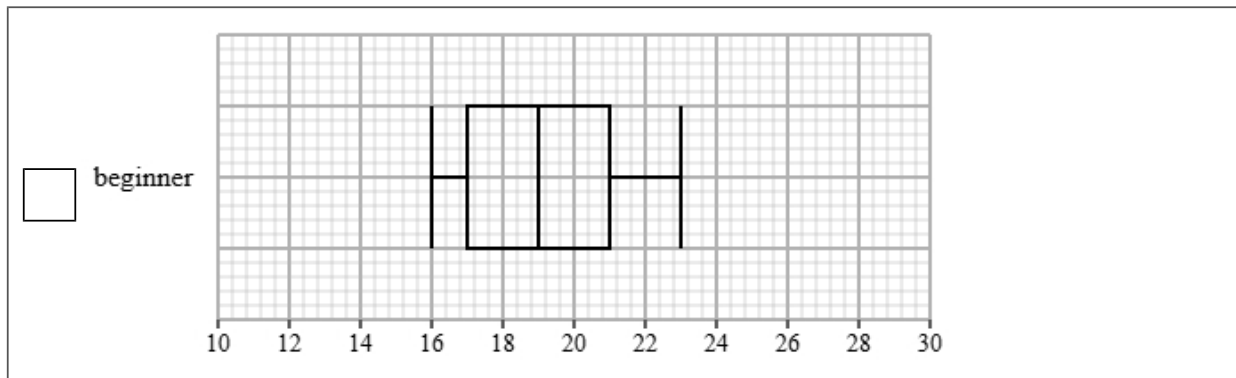
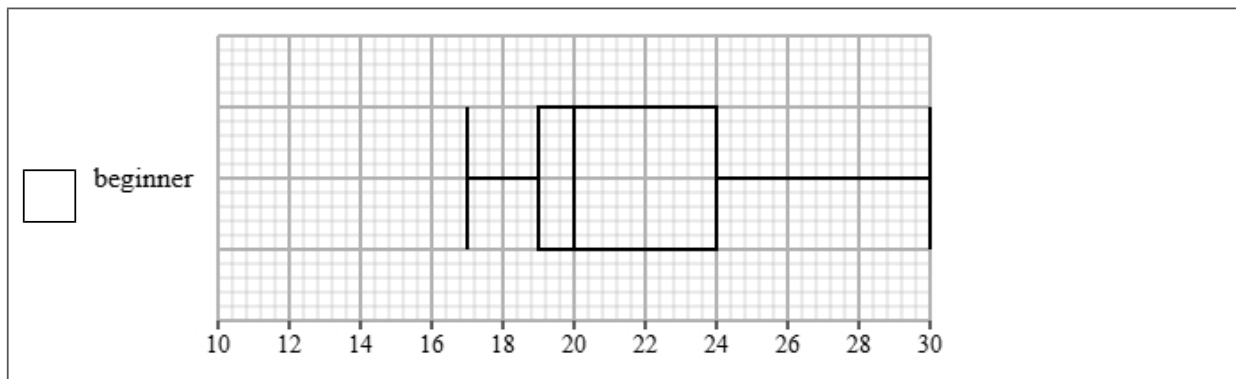
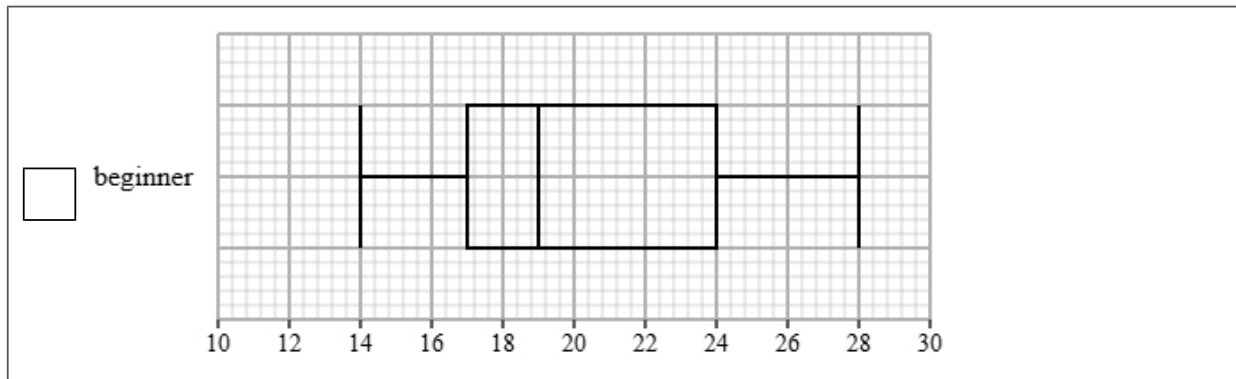
The table gives information about the completion times for the beginner runners.

Least tall	Lower quartile	Median	Upper quartile	Most tall
17	19	20	24	30

(a) Draw a box plot for the completion times for the beginner runners.

(2 marks)

Select the correct answer.



(b) Compare the two distributions of completion times.

Give three comparisons and interpret one of these comparisons.

(4 marks)

Select **one** box.

- The median is bigger.
- The median completion times for pro runners is greater than beginner runners.
- The median completion times for pro runners is lower than beginner runners.
- The median completion times for pro and beginner runners are equal.

Select **one** box.

- The IQR is bigger.
- The IQR for the completion times of the pro runners is greater than beginner runners.
- The IQR for the completion times of the pro and beginner runners are equal.
- The IQR for the completion times of the pro runners is lower than beginner runners.

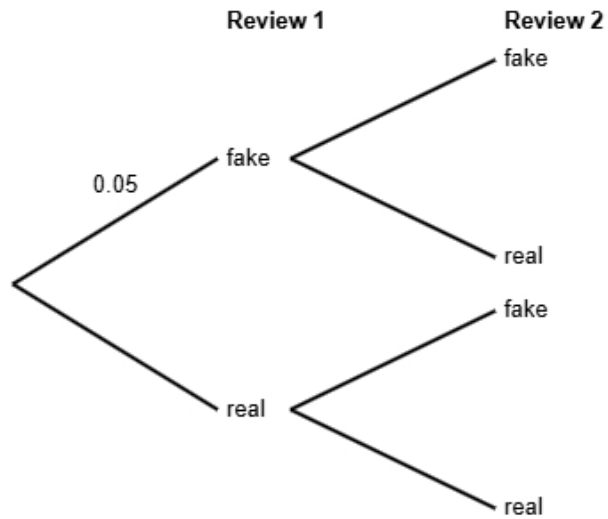
Select **one** box.

- The skews for the completion times of the pro and beginner runners are both positive.
- The skew for the completion times of the pro runners is symmetrical and the skew for the beginner runners is positive.
- The skews for the completion times of the pro and beginner runners are both symmetrical.
- The skew for the completion times of the pro runners is symmetrical and the skew for the beginner runners is negative.

Select **one** box.

- The times for the pro runners are more spread out than the beginner runners.
- The pro runners are on average slower than the beginner runners.
- The pro runners are on average faster than the beginner runners.
- The pro runners are more skewed than beginner runners.

- 11** Research suggests that 5% of online product reviews are fake.  
All other reviews are genuine.  
Emma is reading two reviews for a product.  
She does not know if each review is fake or real.



- (a) Complete the probability tree diagram.

(2 marks)

The branches for each stage must add up to 1.  
Each test is independent so will have the same probabilities.

(b) Find the probability that both of Emma's reviews are real.

(2 marks)

You will need to find  $P(\text{real})$  AND  $P(\text{real})$ .

Remember, AND means  $\times$  in probability.

(c) Emma states that the probability that exactly one reviews is fake is less than 10%

Find out whether or not Emma is correct.

(3 marks)

Find the probability of exactly one reviews is fake (there are two outcomes on the tree diagram).

---

---

---

---

Select **one** box.

- The probability that exactly one reviews is fake is less than 10%, so Emma is not correct.
- The probability that exactly one reviews is fake is more than 10%, so Emma is not correct.
- The probability that exactly one reviews is fake is more than 10%, so Emma is correct.
- The probability that exactly one reviews is fake is less than 10%, so Emma is correct.

12 The table shows information about apartments for rent in Manchester.

number of rooms	number of apartments
1	225
2	180
3	495
4	120
5 or more	180
Total	1200

A researcher wants to investigate the price of these apartments and takes a stratified sample of 80 apartments according to the number of rooms.

(a) The researcher says the mode of the number of rooms for these apartments is 3.

Explain how the researcher knows this.

(1 mark)

Select **one** box.

- 3 rooms has the highest frequency.
- 3 is the middle number.
- 3 is the difference between the largest and smallest number.
- 3 apartments has the highest frequency.

(b) Work out the number of apartments in the sample for each number of rooms.

number of rooms	number of apartments in the sample
1	
2	
3	
4	
5 or more	

(3 marks)

Find the divisor for the stratified sample:  $\frac{\text{total}}{\text{sample size}}$

Divide each frequency by this number to find the required sample in each group

(c) Describe how the 80 apartments in the sample should be selected.

(3 marks)

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

- Ensure that all 1200 apartments are included in the sample.
- Use a sampling frame for each strata.
- Generate random numbers, remove repeats or numbers out of range.
- Complete two of the strata.
- Number each of the apartments, and then use the random numbers to select the required amount of apartments.
- Select the first 80 apartments.

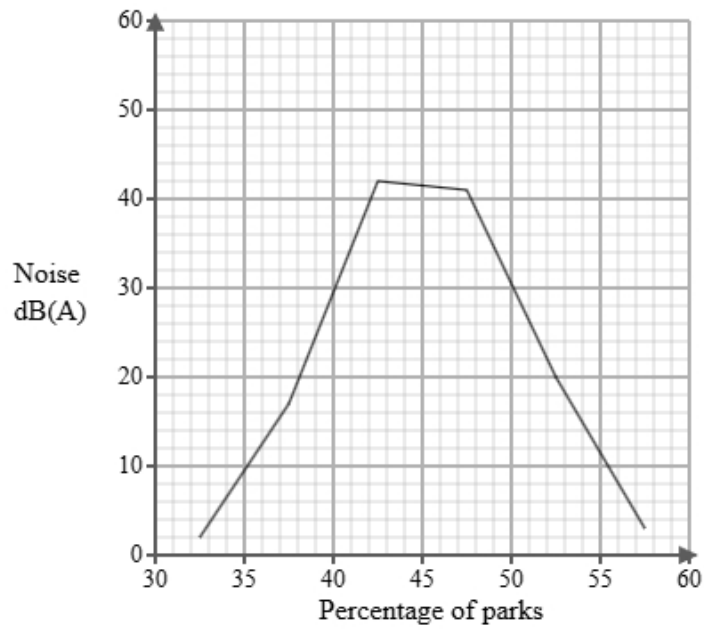
- 13** Luca works for a public health institute. He has been tasked with investigating noise levels in urban parks. Below is a section of the spreadsheet he used to record his findings.

Noise dB(A)	Percentage of parks
$30 < n \leq 35$	8
$35 < n \leq 40$	six
$40 < n \leq 45$	8
$45 < n \leq 50$	111
$50 < n \leq 55$	57
$55 < n \leq 60$	10
Total	100

Luca cleans the data to create the table below.

Noise dB(A)	Percentage of parks
$30 < n \leq 35$	8
$35 < n \leq 40$	6
$40 < n \leq 45$	8
$45 < n \leq 50$	11
$50 < n \leq 55$	57
$55 < n \leq 60$	10
Total	100

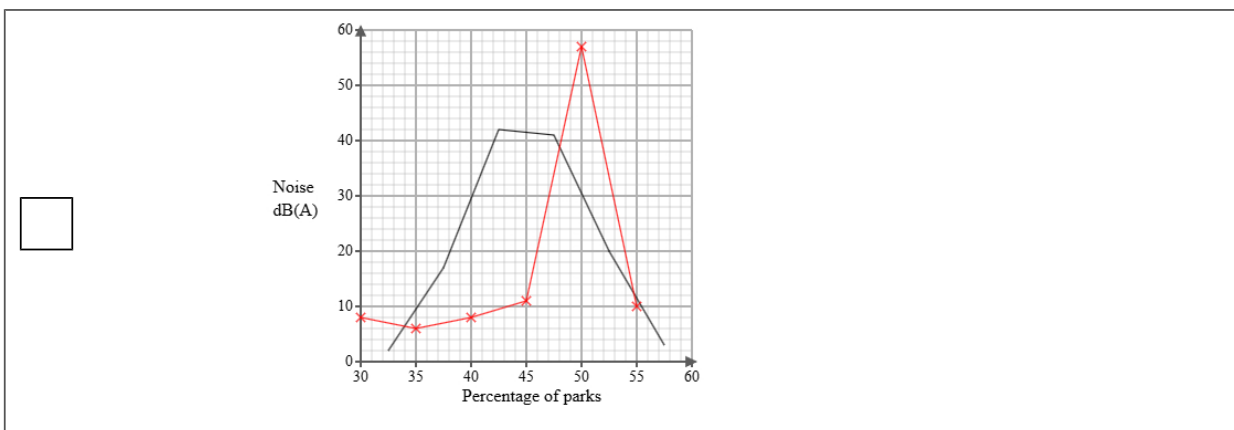
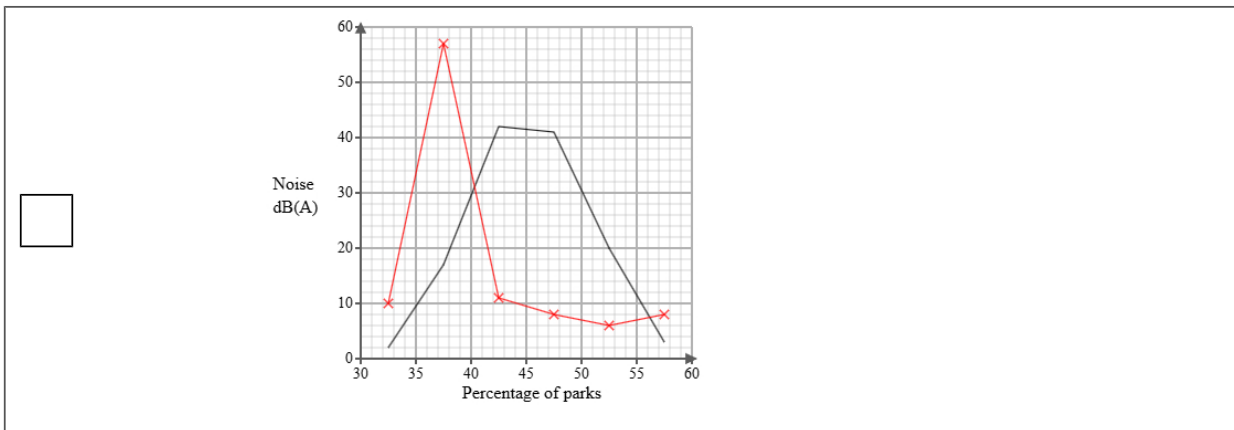
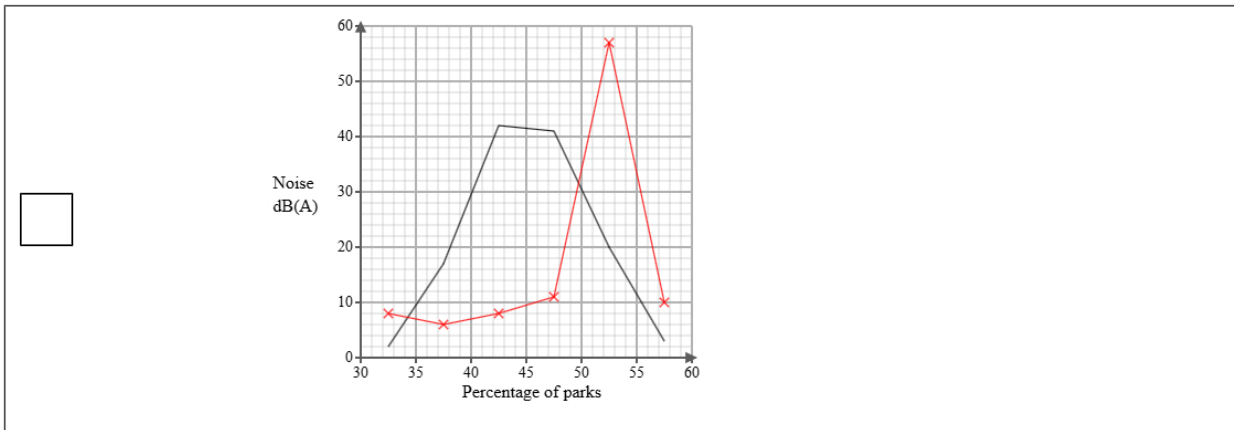
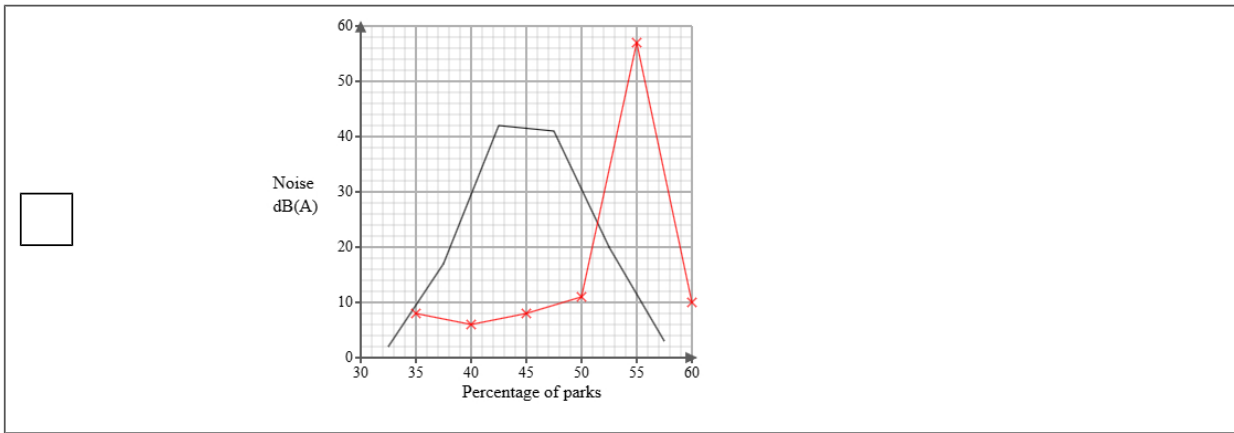
A frequency polygon has been drawn for noise levels in residential streets.



- i) On the same graph, draw the frequency polygon for noise levels in urban parks.
- 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 noise levels in residential streets is positively skewed whereas the distribution of noise levels in urban parks is symmetrical.
- The distribution of noise levels in residential streets is negatively skewed whereas the distribution of noise levels in urban parks is symmetrical.
- This means that in residential areas the noise levels were equally spread out on either side of the median and in urban parks the noise levels were mainly at the upper end of the distribution.
- This means that in residential areas the noise levels were mainly at the upper end of the distribution and in urban parks the noise levels were equally spread out on either side of the median.
- This means that in residential areas the noise levels were mainly at the lower end of the distribution and in urban parks the noise levels were mainly at the upper end of the distribution.
- The distribution of noise levels in residential streets is symmetrical whereas the distribution of noise levels in urban parks is negatively skewed.

- 14 The table shows information about the consumer price index (CPI) and average cost of a book (£) in England for Jan 1995, Jan 2005 and Jan 2015.

	Jan 1995	Jan 2005	Jan 2015
consumer price index	100	116	149
average cost of a book (£)	4.57	5.15	7.35

Describe how the increase in average cost of a book (£) compares with the CPI over the ten years to Jan 2005 and over the twenty years to Jan 2015.

(5 marks)

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

- Between Jan 1995 and Jan 2015 the change in price was more than the CPI.
- Between Jan 1995 and Jan 2005 the change in price was less than the CPI.
- $\frac{5.15}{116} \times 100 = 4$  (nearest integer)
- Between Jan 1995 and Jan 2015 the change in price was less than the CPI.
- $\frac{7.35}{149} \times 100 = 5$  (nearest integer)
- $\frac{5.15}{4.57} \times 100 = 113$  (nearest integer)
- Between Jan 1995 and Jan 2005 the change in price was more than the CPI.
- $\frac{7.35}{4.57} \times 100 = 161$  (nearest integer)

**15** Tom is investigating how the distance travelled in km,  $x$ , affects the resale price (£),  $y$  for two types of bicycles, type A and type B.

He found ten bicycles of each type and recorded their distance travelled and resale price and plotted each on scatter diagrams.

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

Here are the results:

Type	Gradient of line of best fit	y-intercept of line of best fit
A	-15	600
B	-10	750

Interpret and compare these results in context.

(5 marks)

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

- Type A reduces in resale price less than Type B.
- Type A has a greater initial resale price.
- Type B reduces in resale price by £10 per km.
- Both bicycles increase in resale price as the distance travelled increase.
- Both bicycles decrease in resale price as the distance travelled increase.
- Type A changes in resale price by £600 per km.
- Type A reduces in resale price by £15 per km.
- Type A reduces in resale price more per km than Type B.
- Type B changes in resale price by £750 per km.
- Type B has a greater initial resale price.