1(a). A farmer grows pumpkins.
The farmer records the masses, m kilograms, of 80 of their pumpkins.
The table shows the results.

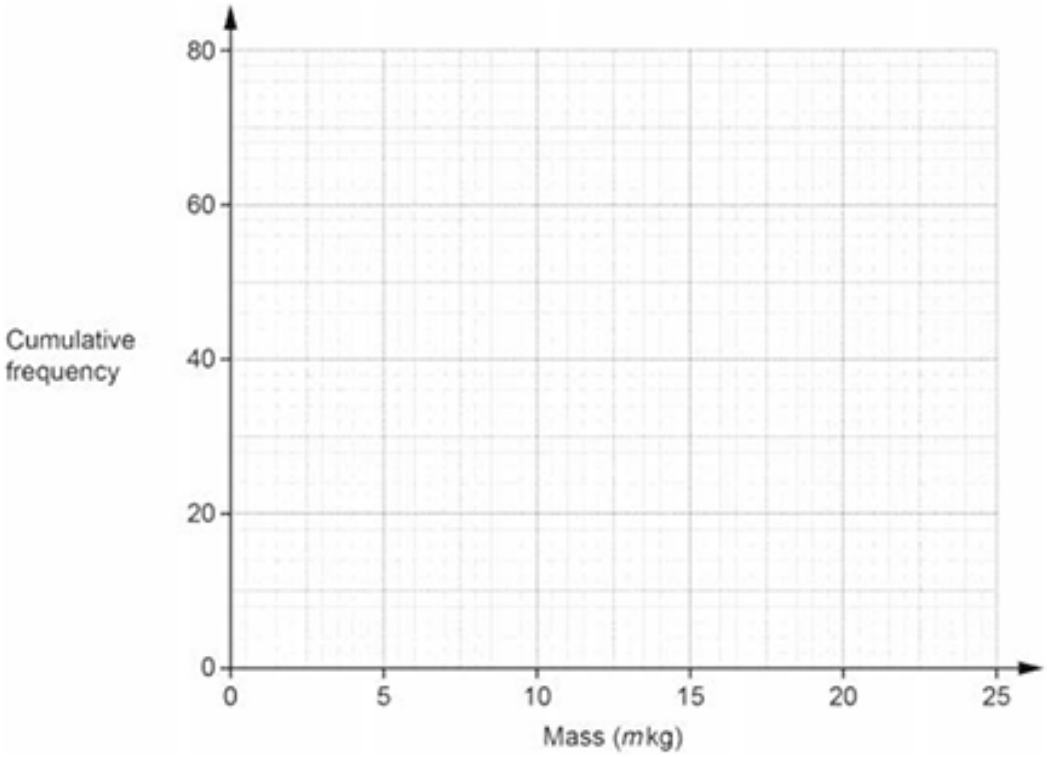
Mass (m kg)	$0 < m \leqslant 5$	$5 < m \leqslant 10$	$10 < m \leqslant 15$	$15 < m \leqslant 20$	$20 < m \leqslant 25$
Frequency	8	12	32	18	10

Complete the cumulative frequency table.

Mass (m kg)	$m \leqslant 5$	$m \leqslant 10$	$m \leqslant 15$	$m \leqslant 20$	$m \leqslant 25$
Cumulative frequency	8	20			

[1]

(b). Draw the cumulative frequency graph to represent these results.



[3]

(c). Write down the lower quartile of the mass of the 80 pumpkins.

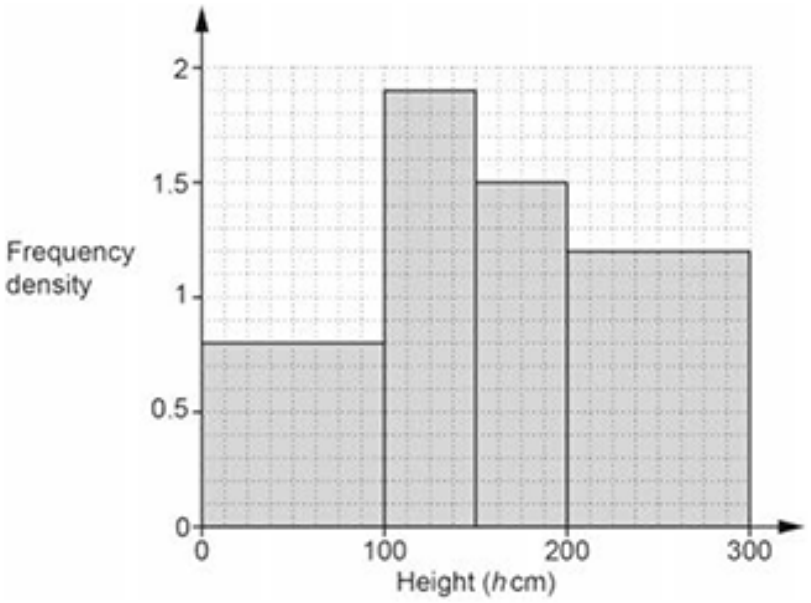
..... kg **[1]**

(d). The farmer picks a pumpkin at random.

Find an estimate for the probability that the pumpkin has a mass greater than 17 kg.

..... **[2]**

2(a). The histogram summarises the heights, h cm, of some plants in a garden centre.



Show that there are 120 plants with a height in the interval $200 < h \leq 300$.

[1]

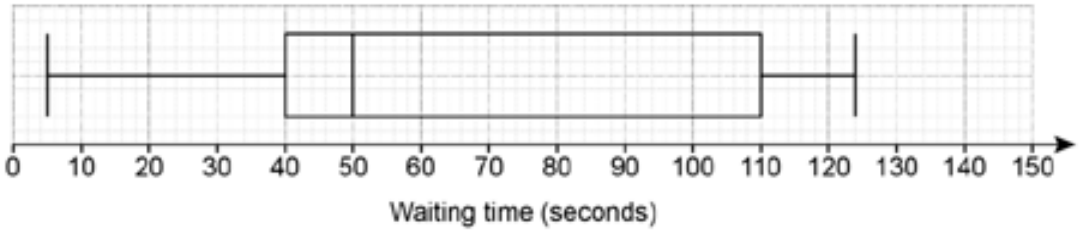
(b). The value, in pounds, of each plant depends on the plant's height.
The table below shows this information.

Height (h cm)	**Value (£)**
$0 < h \leq 100$	3.10
$100 < h \leq 150$	3.60
$150 < h \leq 200$	4.00
$200 < h \leq 300$	5.30

Use this information to find the **total** value of the plants represented in the histogram.

£ ..... **[4]**

3(a). The box plot shows the distribution of the waiting time of cars at a road junction.



Write down the median waiting time.

..... s **[1]**

(b). What percentage of the waiting times were less than 40 seconds?

..... % **[1]**

(c). Given that a randomly chosen car waited for more than 40 seconds, write down the probability that the car waited for more than 110 seconds.

..... **[1]**

4. Four numbers are written, in ascending order, as algebraic expressions

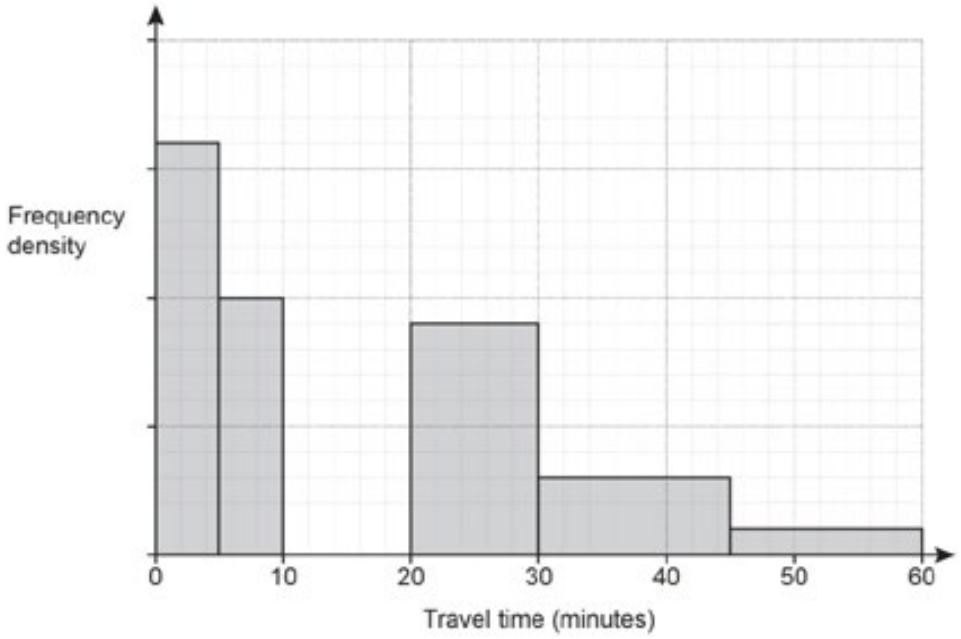
a $a + b$ $a + 2b$ $3a - b$

The mean of these four numbers is 18.
The range of these four numbers is 20.

Find the value of a and the value of b.
You must show your working.

a = ...
b = ... **[5]**

5. A group of students record the time taken to travel to school.
All students in the group took less than an hour to travel to school.
Some of their results are recorded on this histogram.



3 students took more than 45 minutes to travel to school.

How many students took more than 20 minutes to travel to school?

..... students **[4]**

6(a). A farmer grows pumpkins.
The farmer records the masses, m kilograms, of 80 of their pumpkins.
The table shows the results.

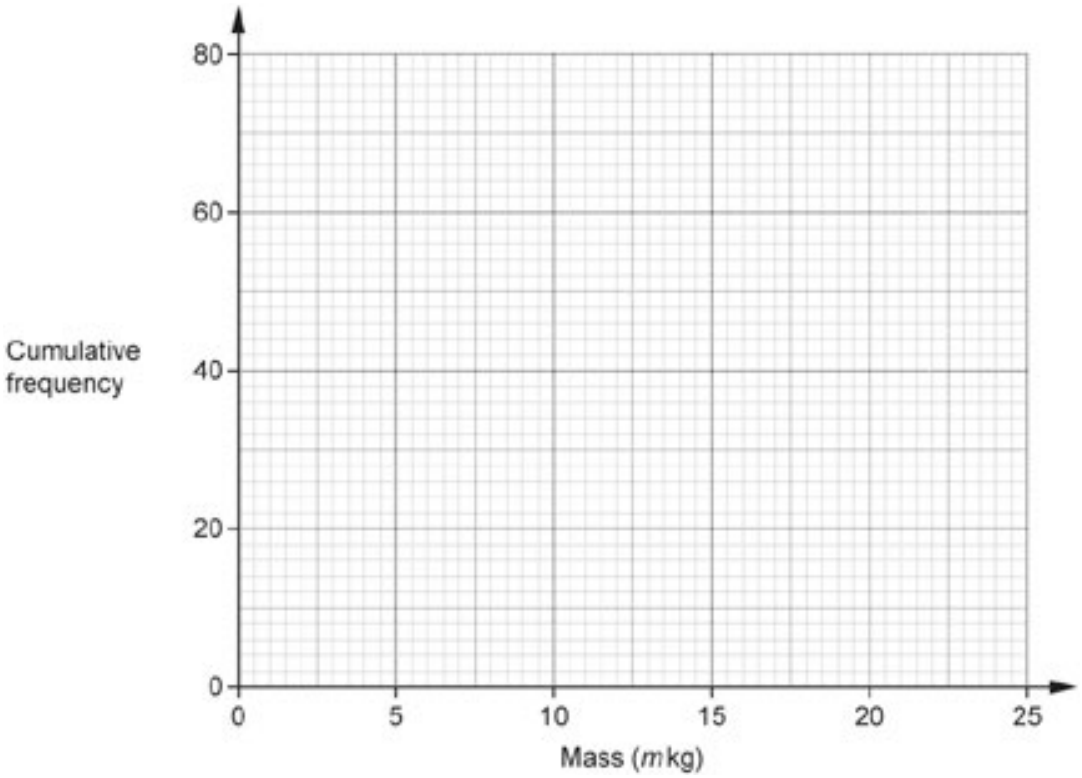
Mass (m kg)	$0 < m \leq 5$	$5 < m \leq 10$	$10 < m \leq 15$	$15 < m \leq 20$	$20 < m \leq 25$
Frequency	10	22	28	14	6

Complete the cumulative frequency table.

Mass (m kg)	$m \leq 5$	$m \leq 10$	$m \leq 15$	$m \leq 20$	$m \leq 25$
Cumulative frequency	10	32			

[1]

(b). Draw the cumulative frequency graph to represent these results.



[3]

(c). Write down the upper quartile of the mass of the 80 pumpkins.

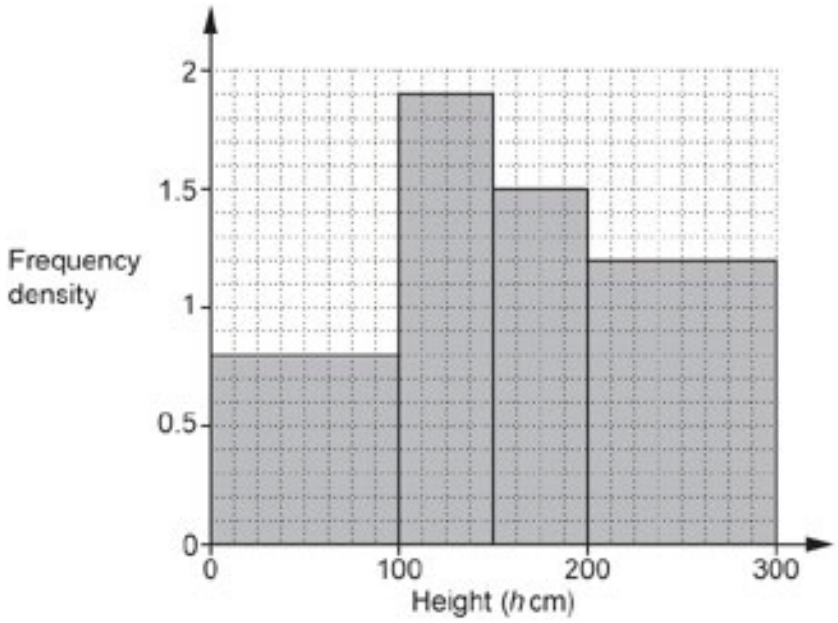
..... kg **[1]**

(d). The farmer picks a pumpkin at random.

Find an estimate for the probability that the pumpkin has a mass greater than 18 kg.

..... **[2]**

7(a). The histogram summarises the heights, h cm, of some plants in a garden centre.



Show that there are 80 plants with a height in the interval $0 < h \leqslant 100$.

[1]

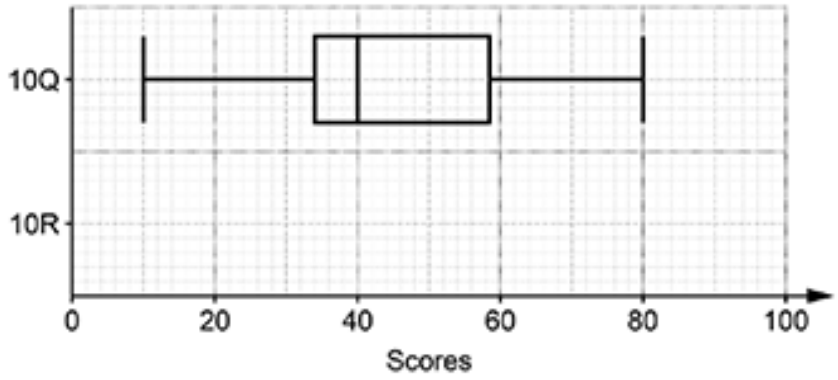
(b). The value, in pounds, of each plant depends on the plant's height.
The table below shows this information.

Height (h cm)	**Value (£)**
$0 < h \leqslant 100$	2.50
$100 < h \leqslant 150$	3.40
$150 < h \leqslant 200$	5.00
$200 < h \leqslant 300$	6.30

Use this information to find the **total** value of the plants represented in the histogram.

£ ..... **[4]**

8(a). The box plot shows the distribution of the scores, in a test, taken by class 10Q.



The scores for the same test, taken by class 10R to complete the same test are summarised below.

- median = 54
- lower quartile = 44
- interquartile range = 16
- range = 70
- highest score = 84

Show the distribution of 10R's scores as a box plot on the diagram above.

[3]

(b). Which class has the more **consistent** scores?
Give a reason for your answer.

_____ because _____

_____ **[1]**

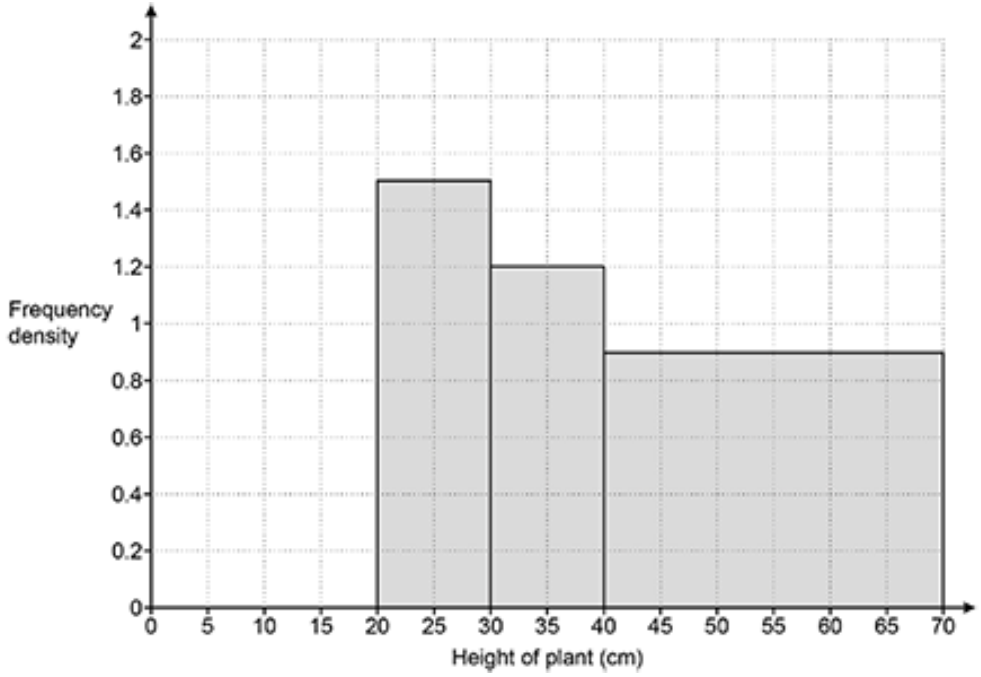
9. A fitness centre records how long each customer spends in the gym.
This **cumulative frequency** table summarises the results.

Time (t minutes)	Cumulative frequency
$t \leq 10$	4
$t \leq 20$	15
$t \leq 30$	32
$t \leq 40$	51
$t \leq 50$	60
$t \leq 60$	72

Calculate an estimate of the mean time the customers spend in the gym.
You must show your working.
You may use the table above to help present your work.

..... minutes **[5]**

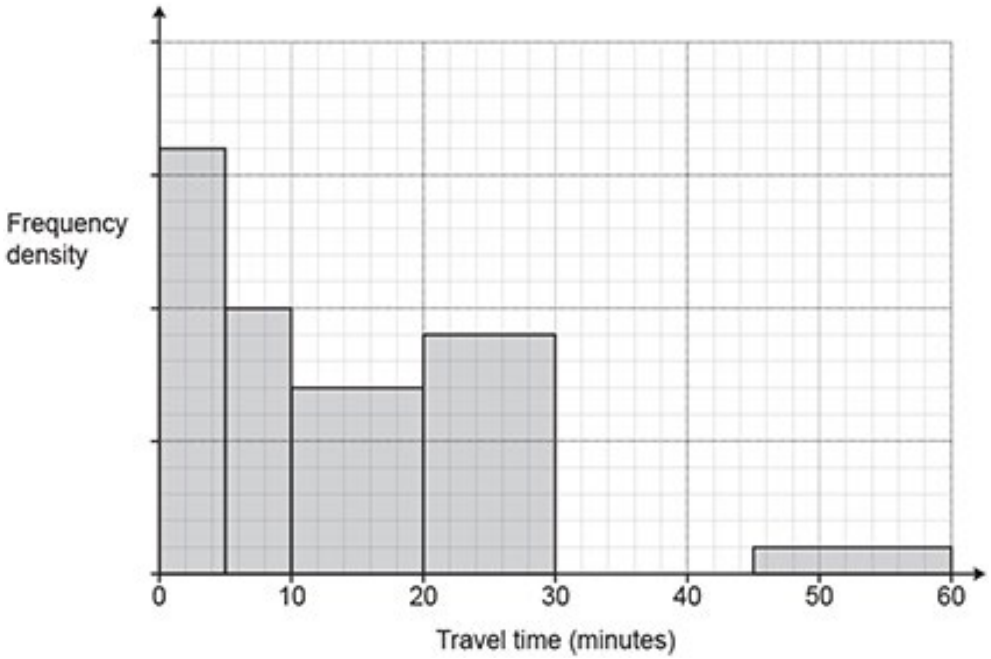
10. The histogram shows the heights of some plants out of a total of 70 plants.



Complete the histogram to show the plants with heights between 0 cm and 20 cm.

[5]

11. A group of students record the time taken to travel to school.
All students in the group took less than an hour to travel to school.
Some of their results are recorded on this histogram.

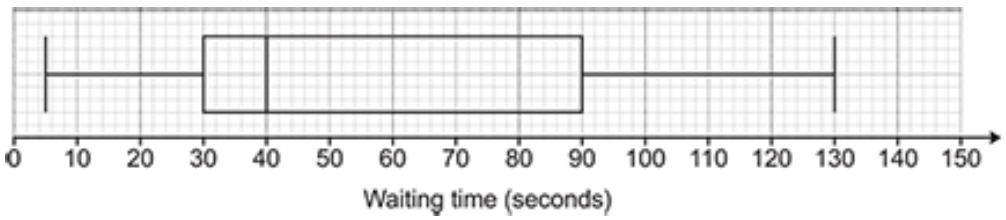


16 students took less than 5 minutes to travel to school.

How many students took less than 20 minutes to travel to school?

.....students **[4]**

12(a). The box plot shows the distribution of the waiting time of cars at a road junction.



Write down the median waiting time.

..... s **[1]**

(b). What percentage of the waiting times were less than 30 seconds?

..... % **[1]**

(c). Given that a randomly chosen car waited for more than 30 seconds, write down the probability that the car waited for more than 90 seconds.

..... **[1]**

13. Four numbers are written, in ascending order, as algebraic expressions.

a $\quad a + b$ $\quad a + 2b$ $\quad 3a - b$

The mean of these four numbers is 27.
The range of these four numbers is 24.

Find the value of a and the value of b.
You must show your working.

a = ...

b = ... **[5]**

14(a). Mia weighs many apples.

The weights of the apples are summarised below.

- heaviest apple = 70 g
- range = 40 g
- median = 50 g
- lower quartile = 45 g
- 50% of the apples weigh between 45 g and 60 g
- mean = 53 g

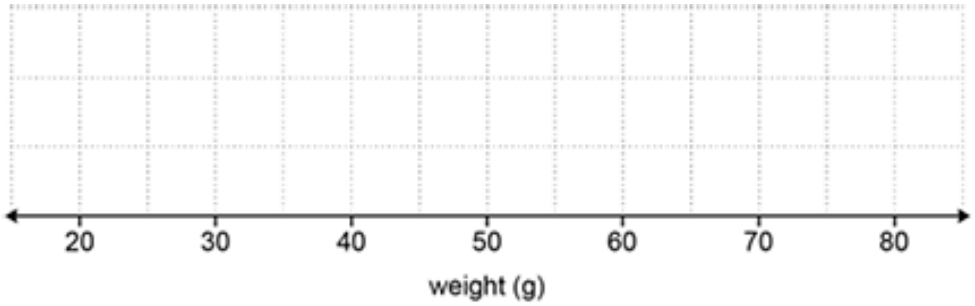
i. Write down the interquartile range for the weights of the apples.

…….……. g **[1]**

ii. Write down the percentage of the apples that weigh between 45 g and 50 g.

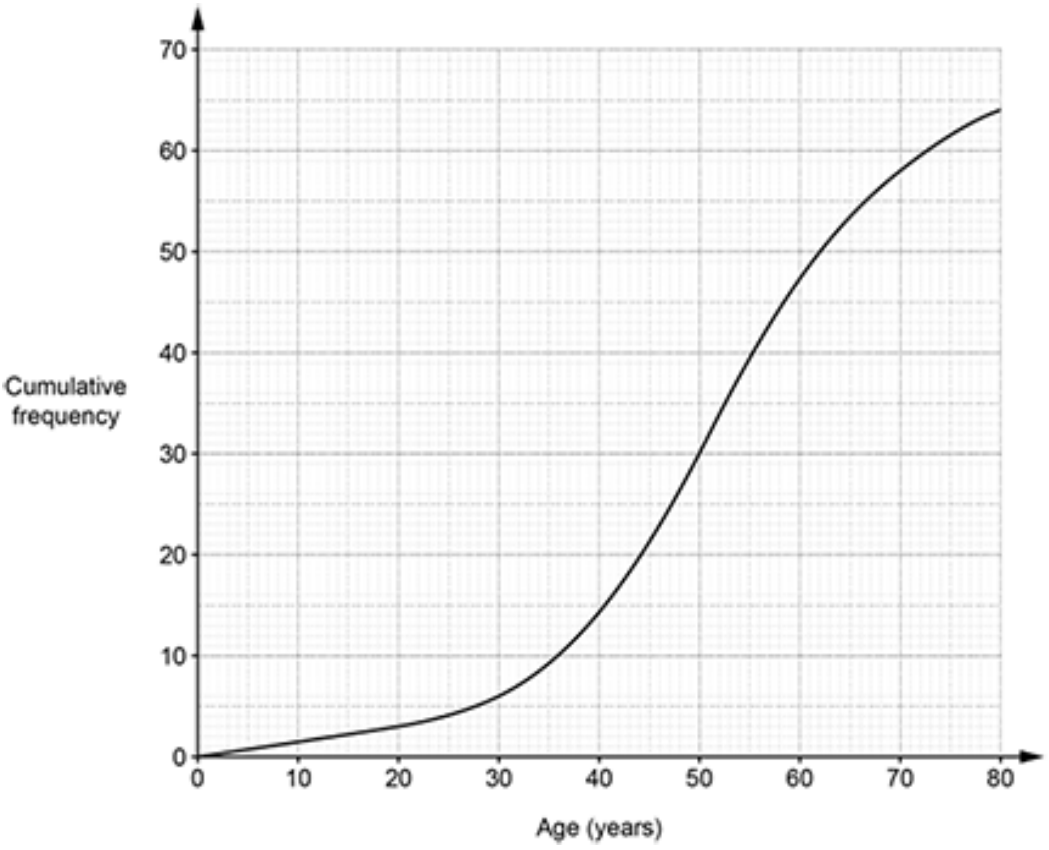
……..…… % **[1]**

(b). Draw a box plot to show the distribution of the weights of the apples.



[3]

15(a). The cumulative frequency graph shows the distribution of the ages of the members of a **tennis** club.



The table summarises the ages of the members of a cycling club.

Age (*a* years)	$0 < a \leq 20$	$20 < a \leq 30$	$30 < a \leq 40$	$40 < a \leq 50$	$50 < a \leq 70$	$70 < a \leq 80$
Frequency	6	15	9	14	16	4

On the graph above, draw the cumulative frequency graph of the ages of the members of the **cycling** club.

[5]

(b). Find out which club has younger members on average.
Give evidence to support your decision.

_____ because _____

_____ **[2]**

16. Mia weighs many apples.

The weights of the apples are summarised below.

- heaviest apple = 70 g
- range = 40 g
- median = 50 g
- lower quartile = 45 g
- 50% of the apples weigh between 45 g and 60 g
- mean = 53 g

Mia eats two of the apples.
The apples that they eat weigh 50 g and 58 g.

Mia says

The mean weight of all the apples was 53 g.
I ate one apple that weighed less than the mean and another apple that weighed more than the mean.
Therefore, the mean of the remaining apples will still be 53 g.

Is Mia correct?
Explain your reasoning.

_____

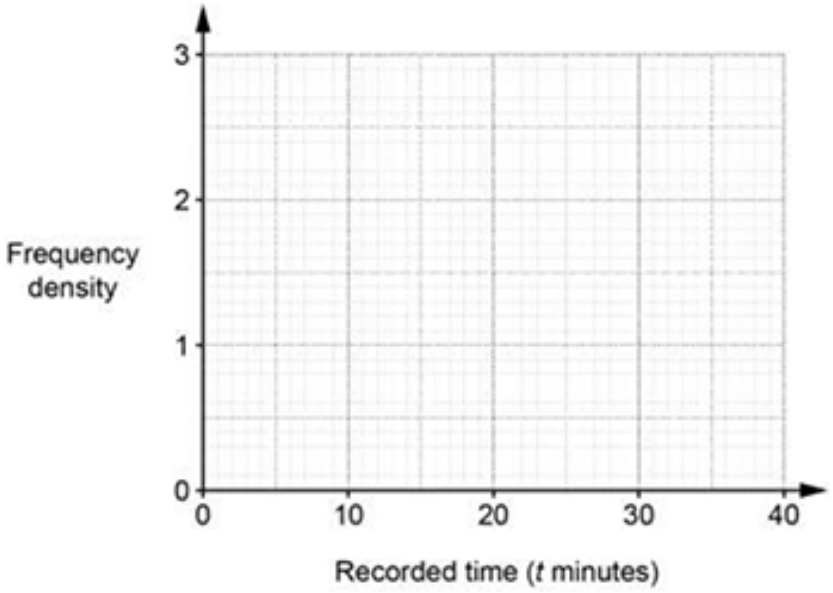
_____

_____ **[2]**

17(a). 70 people each try to solve a puzzle.
The table summarises their recorded times.

Recorded time (t minutes)	**Frequency**
$0 < t \leq 5$	14
$5 < t \leq 15$	23
$15 < t \leq 25$	21
$25 < t \leq 40$	12

Draw a histogram to show this information.



[3]

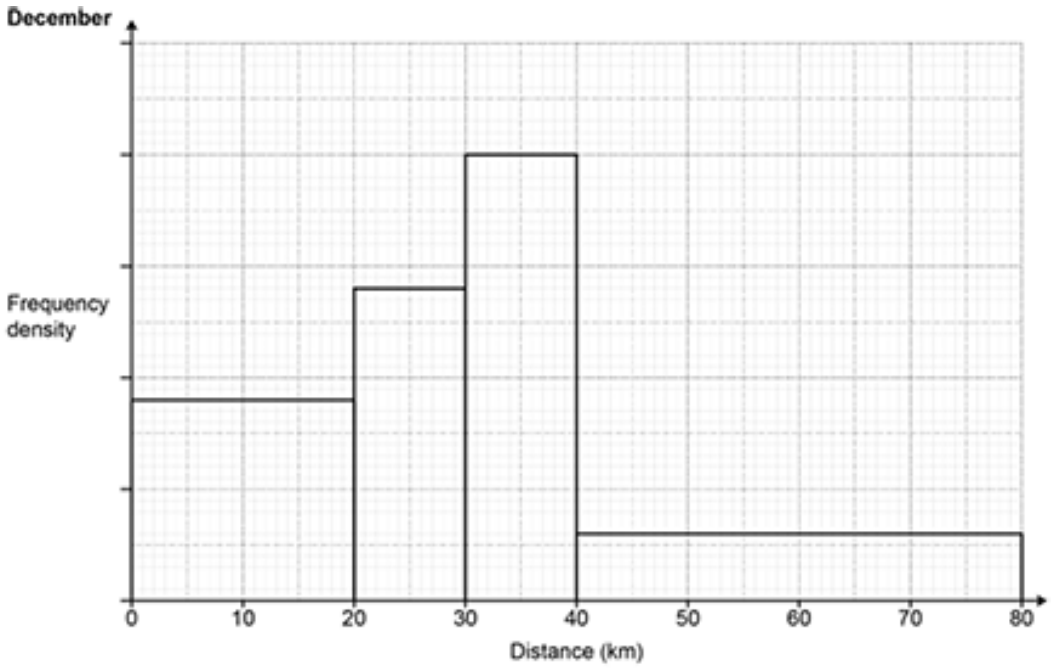
(b). Those people who failed to solve the puzzle within 40 minutes were given a recorded time of 40 minutes.

Mia uses mid-interval values to estimate the mean recorded time of the 70 people.

Explain why Mia's answer is likely to be an under-estimate for the mean of the actual time taken by the 70 people.

---**[1]**

18(a). A running club records the distances run by each member during December.
The results are shown in this histogram.



12 members run more than 40 km.

i. Work out the number of members who run less than 30 km.

…….. **[3]**

ii. Eve says

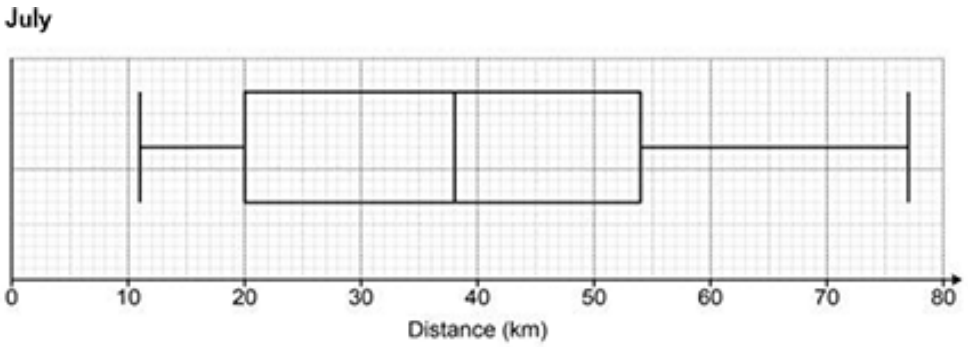
To estimate the range, I subtracted the midpoint of the 0 to 20 class from the midpoint of the 40 to 80 class. So, 60 – 10 = 50 km.

Explain why Eve's method is likely to underestimate the true value of the range.

---

[1]

(b). This box plot shows the distribution of the distance run by each member of the running club during July.



During **December**,

- the median distance run was 30 km
- the interquartile range of the distance run was 20 km.

Make **two** comparisons between the distances run during December and the distances run during July. Include values to support your comparisons.

1 _____

_____

_____

2 _____

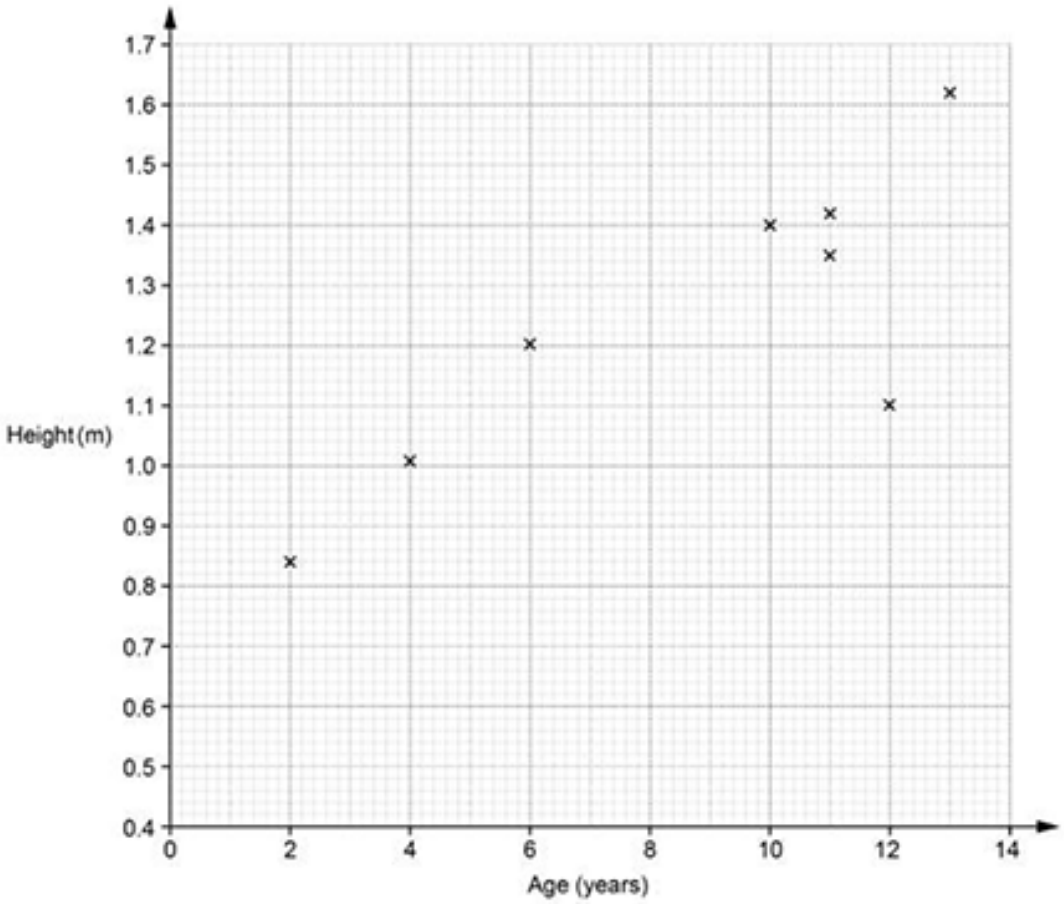
_____

[4]

19. The table shows the ages and heights of 12 children.

Age (years)	2	4	12	6	10	11	13	11	5	7	9	14
Height (m)	0.84	1.01	1.1	1.2	1.4	1.35	1.62	1.42	1.16	1.24	1.28	1.66

The points for the first eight children (shaded in the table above) are plotted on the scatter diagram.



Explain why using this data to estimate the height of a child that is 16 years old may be unreliable.

---

[1]

20(a). 80 cyclists take part in a race.
A summary of their times is shown in the table.

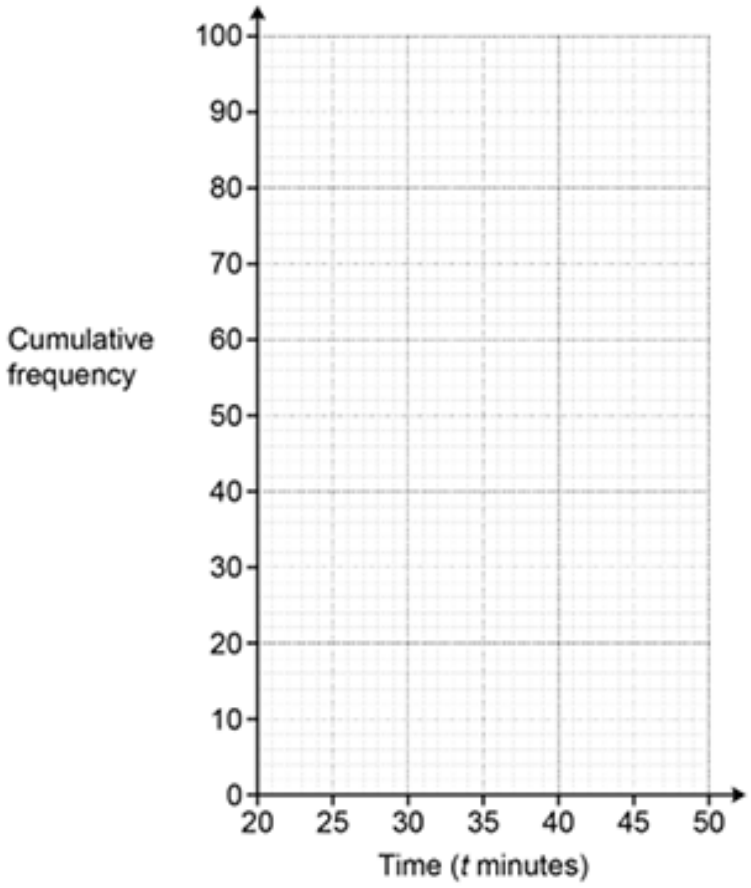
Time (*t* minutes)	**Frequency**
$20 < t \leqslant 25$	6
$25 < t \leqslant 30$	12
$30 < t \leqslant 35$	20
$35 < t \leqslant 40$	22
$40 < t \leqslant 45$	12
$45 < t \leqslant 50$	8

Complete the cumulative frequency table.

Time (*t* minutes)	**Cumulative frequency**
$t \leqslant 25$	6
$t \leqslant 30$	
$t \leqslant 35$	
$t \leqslant 40$	
$t \leqslant 45$	
$t \leqslant 50$	

[2]

(b). Draw the cumulative frequency graph to show the information.



[3]

(c). Amir makes two comments about the times taken to complete the race.

For each comment, decide if Amir is right or wrong and give a reason for your answer.

i. $\frac{1}{4}$ of the 80 cyclists took more than 40 minutes to complete the race.

Amir is _____ because _____

_____

_____ **[2]**

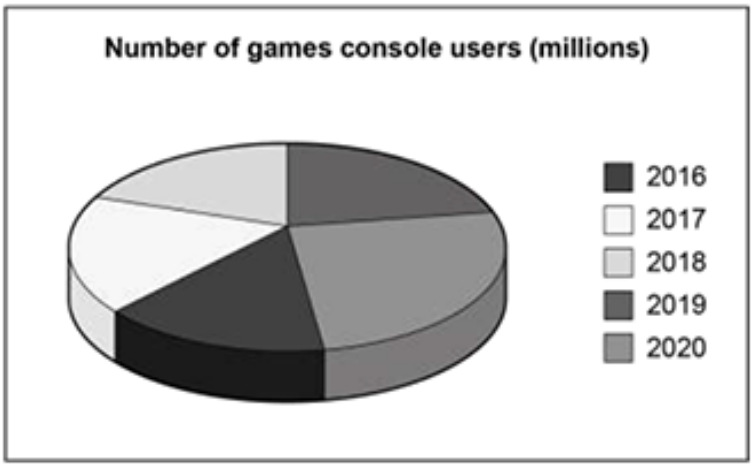
ii. The longest time that any of the 80 cyclists took to complete the race must have been 50 minutes.

Amir is _____ because _____

_____ **[1]**

21(a). Two pupils are given data that shows the estimated number of games console users in a country from 2016 to 2020.

Morgan creates this pie chart to show the data.



Write down two reasons why Morgan's pie chart is not suitable to represent the data.

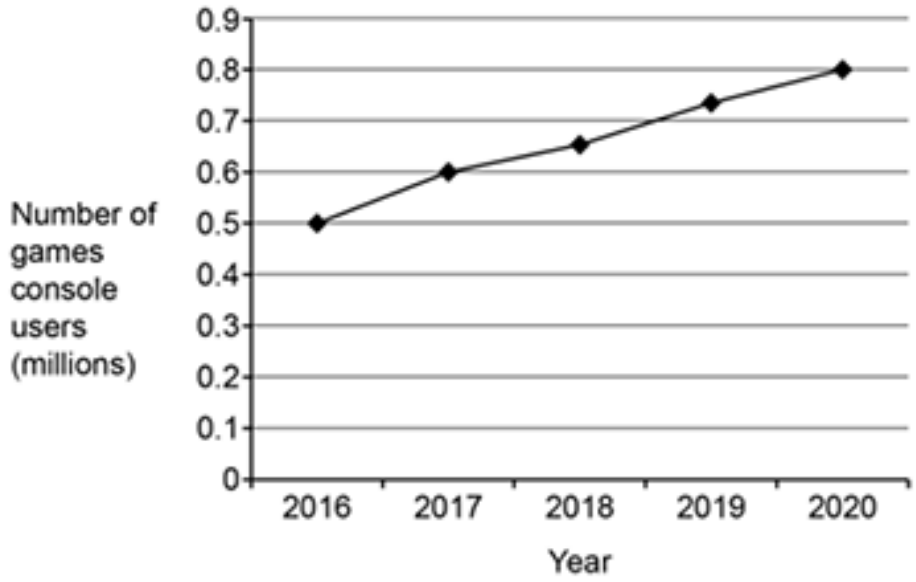
1 _____

_____

2 _____

_____ **[2]**

(b). Charlie creates this line graph to show the same data.



Work out the percentage increase in the number of games console users from 2016 to 2020.

…… % **[4]**

22(a). A student is researching the difference in how much exercise adults and children do.
To collect their data, the student interviews the first 20 people found in the library at 2 pm on one Monday afternoon.

Make **three** different criticisms of the student's method of collecting data.

1 _____

_____

_____

2 _____

_____

_____

3 _____

_____

[3]

(b). Here is the data collection table that the student used.

Hours exercised in a week (h)	Adult tally	Child tally
$0 \leqslant h \leqslant 3$		
$3 \leqslant h \leqslant 6$		
$6 \leqslant h \leqslant 9$		
$9 \leqslant h \leqslant 12$		
$12 \leqslant h \leqslant 18$		

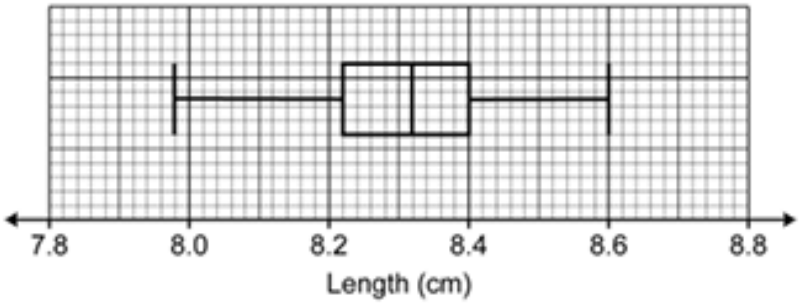
Make **one** criticism of the student's table.

_____

_____

[1]

23(a). The box plot shows the distribution of the lengths, in cm, of 60 full-grown hamsters owned by a pet shop.



Find the range.

…….…….…….. cm **[2]**

(b). Work out the number of these hamsters that have a length of at least 8.4 cm.

…….…….…….. **[2]**

(c). Azmi owns 5 full-grown hamsters.
Azmi picks the third longest hamster and measures its length.
Azmi then looks at the box plot.

Azmi says

This hamster is 8.25 cm long.
Therefore, the hamsters I own are shorter than the full-grown hamsters owned by the pet shop.

i. Give a mathematical reason to support Azmi's conclusion.

_____

---

[1]

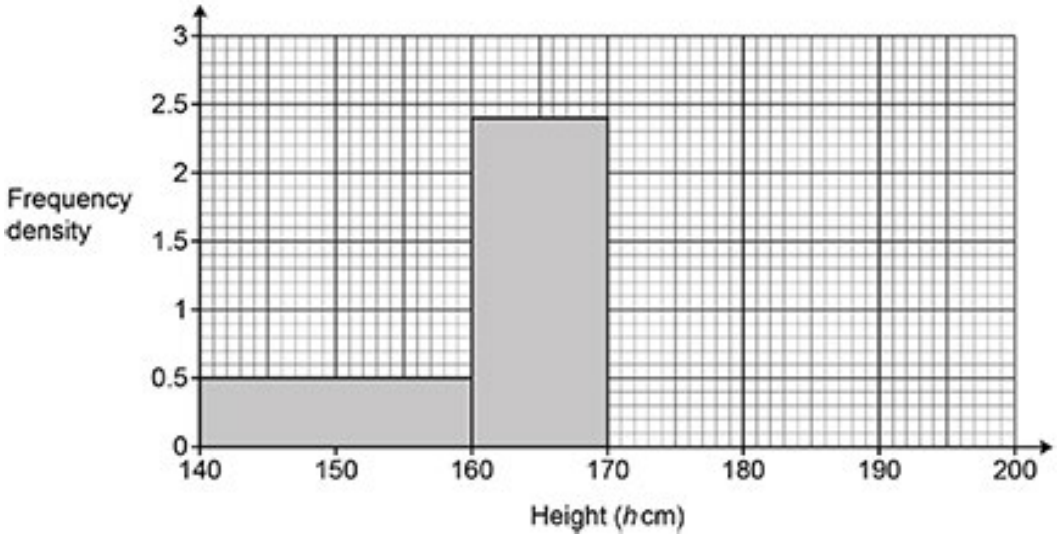
ii. Give a mathematical reason why Azmi's conclusion may be unreliable.

_____

---

[1]

24. The height, *h* cm, of each member of a tennis club is recorded.
The histogram shows some of the results.



40% of the members have a height in the interval $160 \leq h < 170$.

30% of the members have a height in the interval $170 \leq h < 180$.

100% of the members have a height in the interval $140 \leq h < 200$.

Complete the histogram for the intervals $170 \leq h < 180$ and $180 \leq h < 200$.

[6]

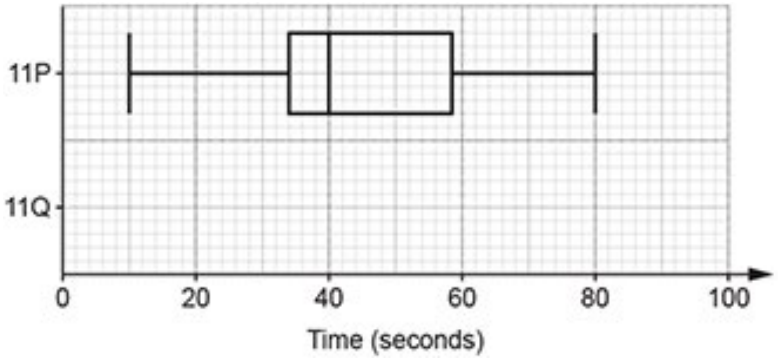
25. A fitness centre records how long each customer spends in the gym.
This **cumulative frequency** table summarises the results.

Time (*t* minutes)	**Cumulative frequency**
$t \leq 10$	6
$t \leq 20$	24
$t \leq 30$	35
$t \leq 40$	48
$t \leq 50$	60
$t \leq 60$	74

Calculate an estimate of the mean time the customers spend in the gym.
You must show your working.
You may use the table above to help present your work.

..... minutes **[5]**

26(a). The box plot shows the distribution of the times, in seconds, taken by class 11P to complete a problem.



The times, in seconds, taken by class 11Q to complete the same problem are summarised below.

- median = 52
- lower quartile = 42
- interquartile range = 18
- range = 70
- highest score = 86

Show the distribution of 11Q's times as a box plot on the diagram above.

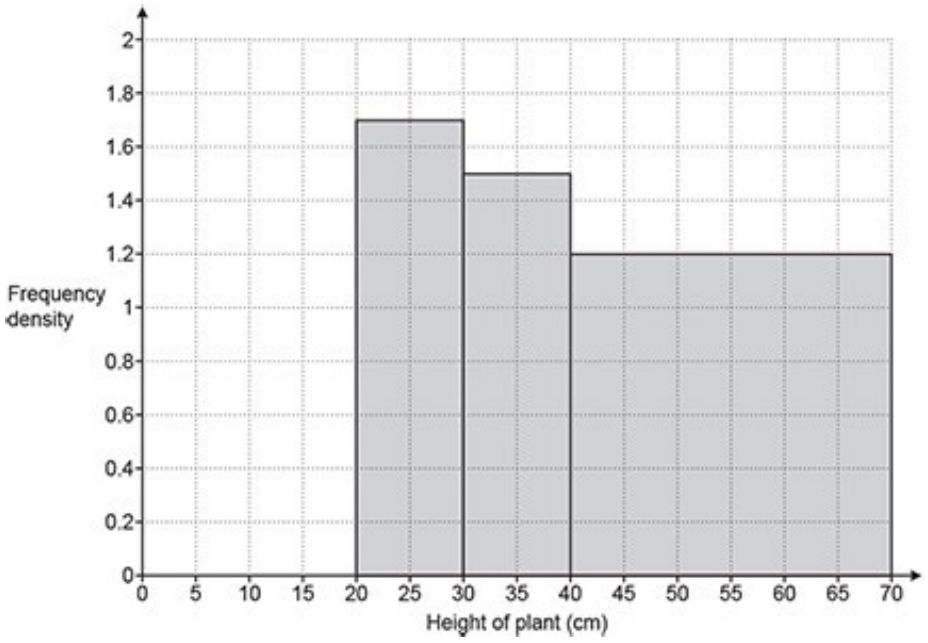
[3]

(b). Which class has the more **consistent** times?
Give a reason for your answer.

_____ because _____

_____ **[1]**

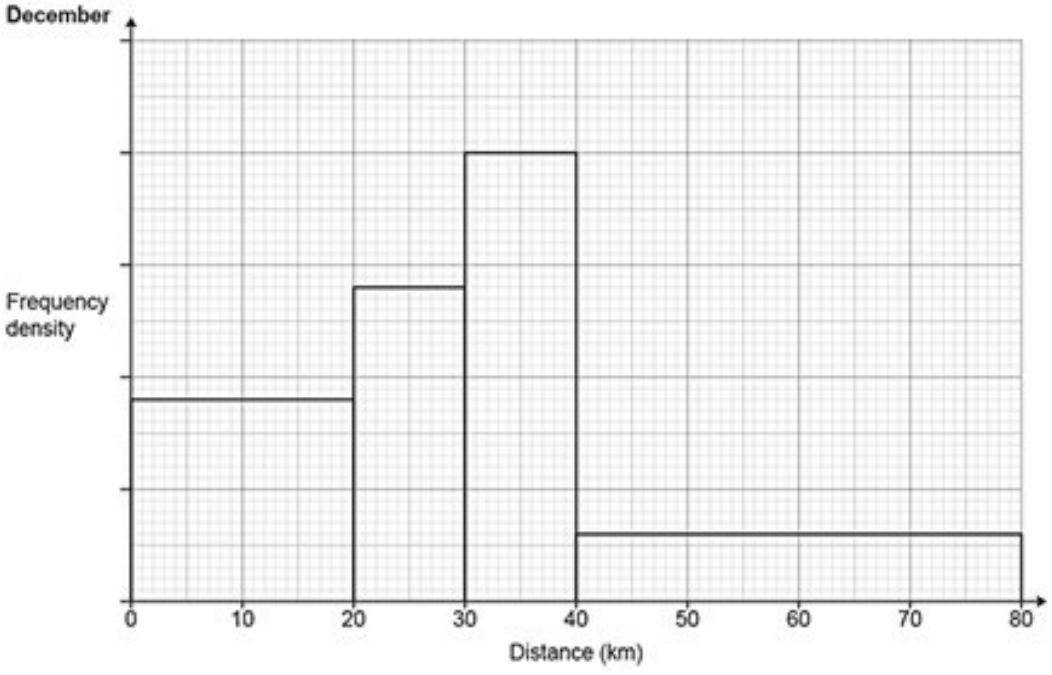
27. The histogram shows the heights of some plants out of a total of 80 plants.



Complete the histogram to show the plants with heights between 0 cm and 20 cm.

[5]

28(a). A running club records the distances run by each member during December.
The results are shown in this histogram.



18 members run less than 20 km.

i. Work out the number of members who run more than 30 km.

..... **[3]**

ii. Finley says

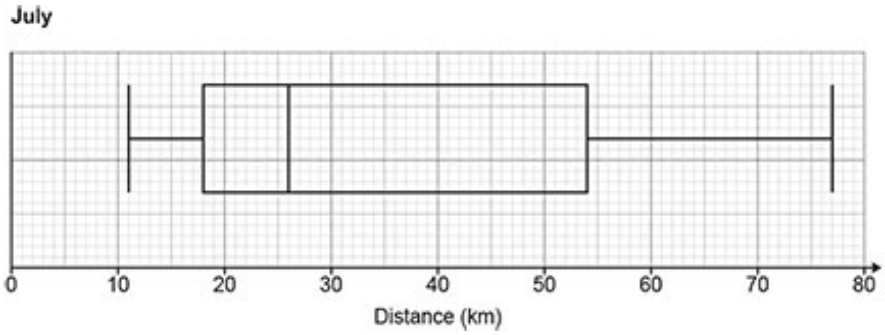
To estimate the range, I subtracted the smallest possible value from the largest possible value. So, 80 – 0 = 80 km.

Explain why Finley's method is likely to overestimate the true value of the range.

_____

[1]

(b). This box plot shows the distribution of the distance run by each member of the running club during July



During **December**,

- the median distance run was 30 km
- the interquartile range of the distance run was 20 km.

Make **two** comparisons between the distances run during December and the distances run during July. Include values to support your comparisons.

1 _____

_____

_____

2 _____

_____

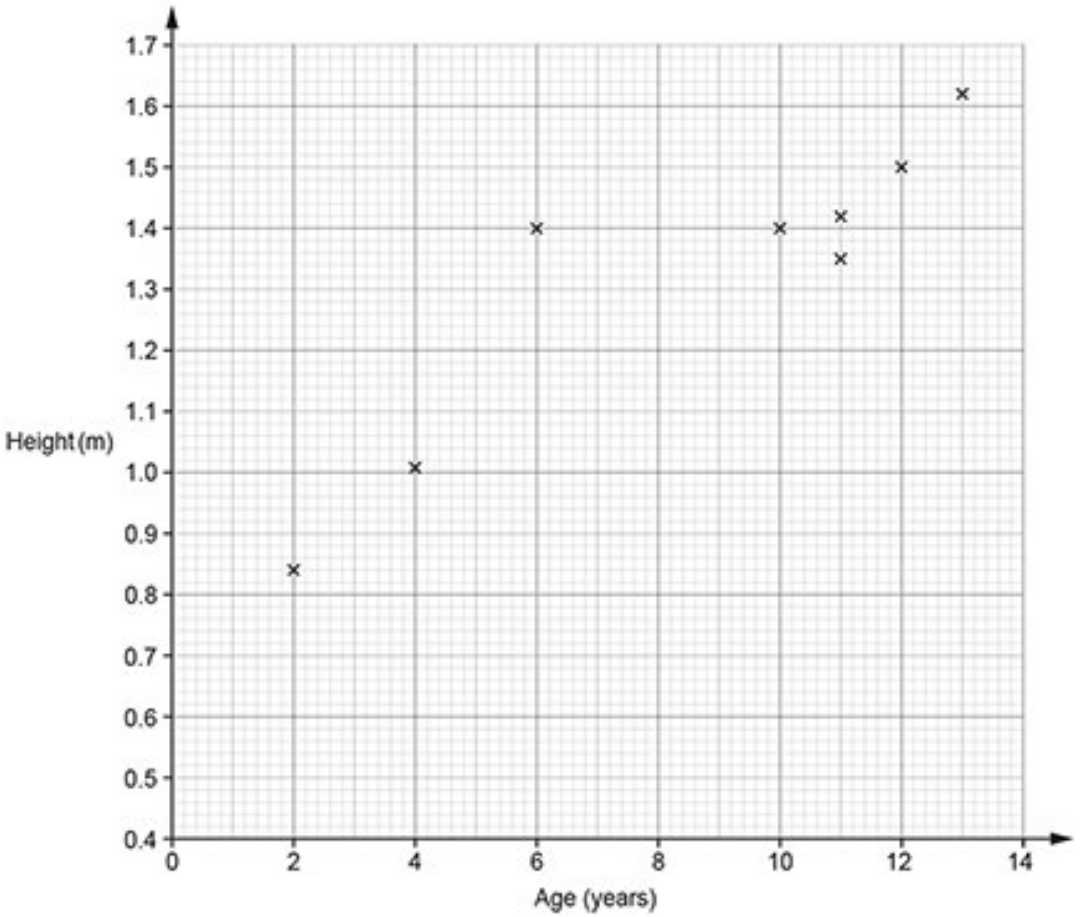
_____

[4]

29. The table shows the ages and heights of 12 children.

Age (years)	2	4	12	6	10	11	13	11	5	7	9	14
Height (m)	0.84	1.01	1.5	1.4	1.4	1.35	1.62	1.42	1.14	1.24	1.26	1.68

The points for the first eight children (shaded in the table above) are plotted on the scatter diagram.



Explain why using this data to estimate the height of a child that is 17 years old may be unreliable.

_____

_____ **[1]**

30(a). Charlie weighs many apples.

The weights of the apples are summarised below.

- heaviest apple = 75 g
- range = 50 g
- median = 60 g
- lower quartile = 45 g
- 50% of the apples weigh between 45 g and 65 g
- mean = 63 g

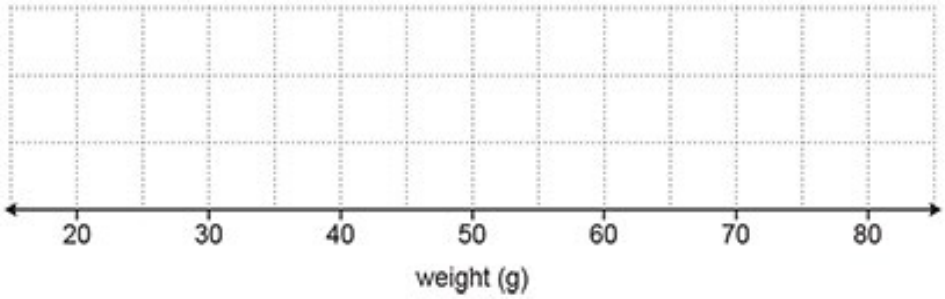
i. Write down the interquartile range for the weights of the apples.

..... g **[1]**

ii. Write down the percentage of the apples that weigh between 45 g and 60 g.

..... % **[1]**

(b). Draw a box plot to show the distribution of the weights of the apples.



[3]

(c). Charlie eats two of the apples.
The apples that they eat weigh 58 g and 66 g.

Charlie says

The mean weight of all the apples was 63 g.
I ate one apple that weighed less than the mean and another apple that weighed more than the mean.
Therefore, the mean of the remaining apples will still be 63 g.

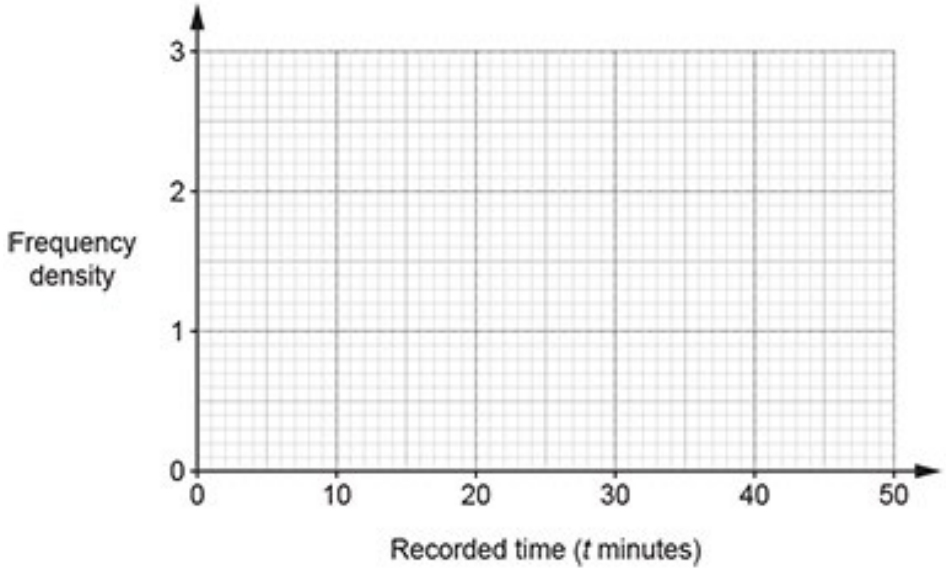
Is Charlie correct?
Explain your reasoning.

[2]

31(a). 60 people each try to solve a puzzle.
The table summarises their recorded times.

Recorded time (t minutes)	Frequency
$0 < t \leq 5$	12
$5 < t \leq 15$	19
$15 < t \leq 30$	18
$30 < t \leq 50$	11

Draw a histogram to show this information.



[3]

(b). Those people who failed to solve the puzzle within 50 minutes were given a recorded time of 50 minutes.

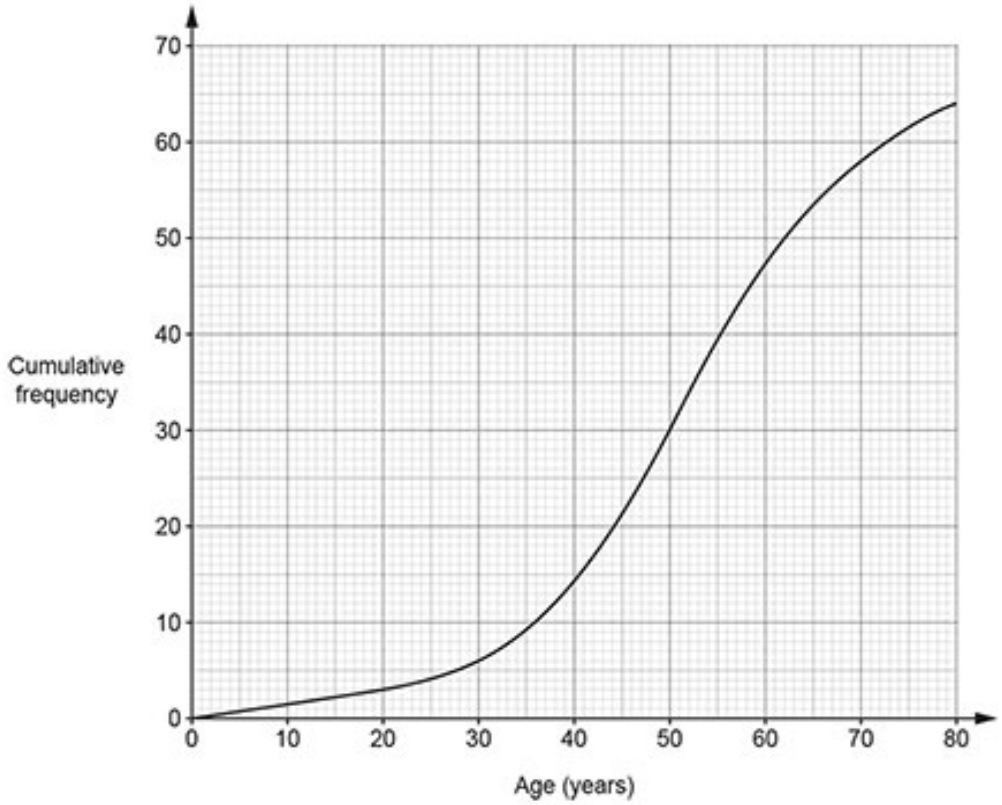
Nina uses mid-interval values to estimate the mean recorded time of the 60 people.

Explain why Nina's answer is likely to be an under-estimate for the mean of the actual time taken by the 60 people.

---

[1]

32(a). The cumulative frequency graph shows the distribution of the ages of the members of a **tennis** club.



The table summarises the ages of the members of a **cycling** club.

Age (*a* years)	$0 < a \leq 20$	$20 < a \leq 30$	$30 < a \leq 40$	$40 < a \leq 50$	$50 < a \leq 70$	$70 < a \leq 80$
Frequency	8	14	8	12	17	5

On the graph above, draw the cumulative frequency graph of the ages of the members of the **cycling** club.

[5]

(b). Find out which club has younger members on average.
Give evidence to support your decision.

---------- because ----------

---------- **[2]**

33(a). A student is researching the difference in how much exercise adults and children do.To collect their data, the student interviews the first 25 people found in the High Street at 11 am on one Monday morning.

Make **three** different criticisms of the student's method of collecting data.

1 _____

_____

_____

2 _____

_____

_____

3 _____

_____

_____

[3]

(b). Here is the data collection table that the student used.

Hours exercised in a week (h)	**Adult tally**	**Child tally**
$0 \leq h \leq 2$		
$2 \leq h \leq 4$		
$4 \leq h \leq 8$		
$8 \leq h \leq 12$		
$12 \leq h \leq 20$		

Make **one** criticism of the student's table.

_____

_____

[1]

34(a). 80 cyclists take part in a race.
A summary of their times is shown in the table.

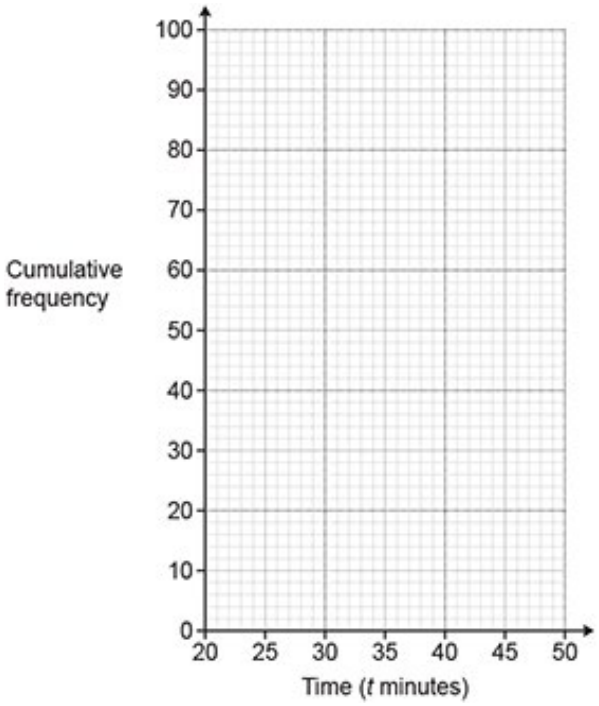
Time (*t* minutes)	Frequency
$20 < t \leq 25$	5
$25 < t \leq 30$	15
$30 < t \leq 35$	24
$35 < t \leq 40$	25
$40 < t \leq 45$	7
$45 < t \leq 50$	4

Complete the cumulative frequency table.

Time (*t* minutes)	Cumulative frequency
$t \leq 25$	5
$t \leq 30$	
$t \leq 35$	
$t \leq 40$	
$t \leq 45$	
$t \leq 50$	

[2]

(b). Draw the cumulative frequency graph to show the information.



[3]

(c). Reece makes two comments about the times taken to complete the race.

For each comment, decide if Reece is right or wrong and give a reason for your answer.

i. $\frac{3}{4}$ of the 80 cyclists took more than 30 minutes to complete the race.

Reece is _____ because _____

_____

_____ **[2]**

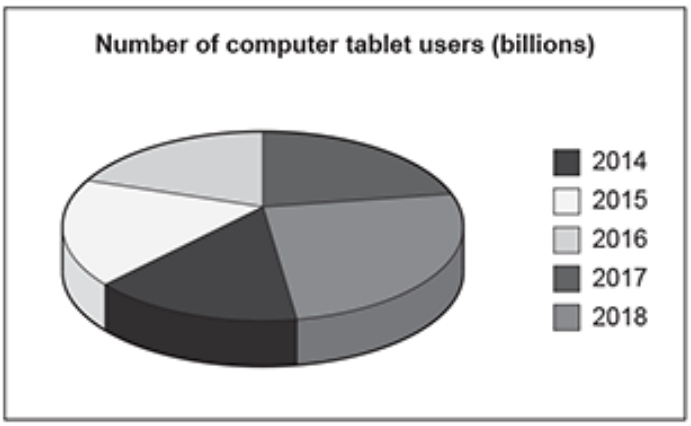
ii. The longest time that any of the 80 cyclists took to complete the race must have been 50 minutes.

Reece is _____ because _____

_____ **[1]**

35(a). Two pupils are given data that shows the estimated number of computer tablet users worldwide from 2014 to 2018.

Li creates this pie chart to show the data.



Write down two reasons why Li's pie chart is not suitable to represent the data.

1 _____

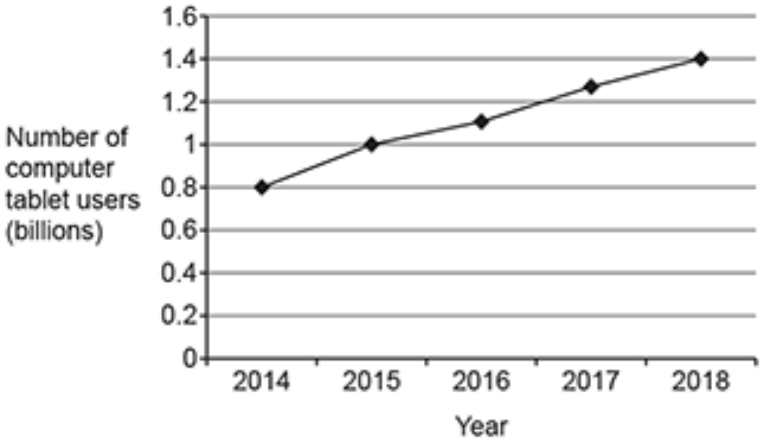
_____

2 _____

_____

[2]

(b). Amaya creates this line graph to show the same data.



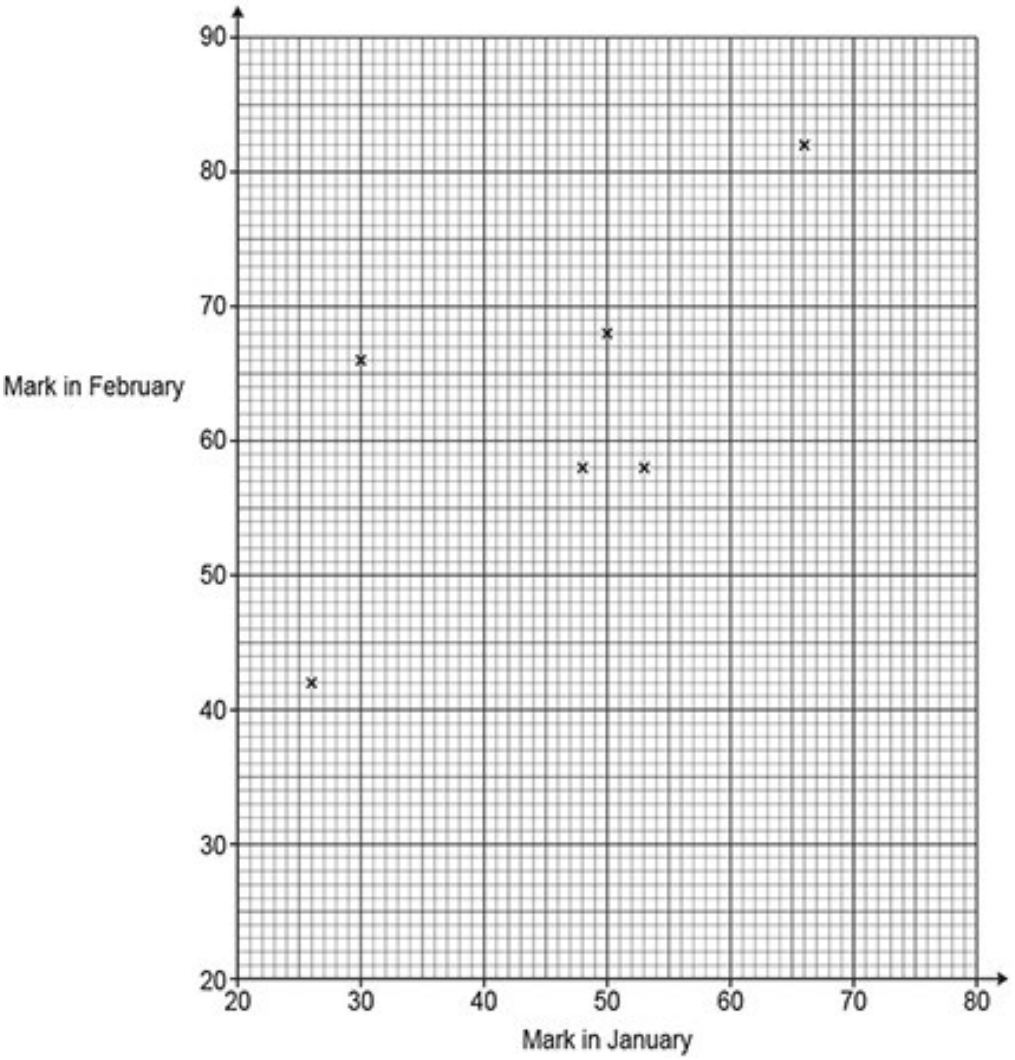
Work out the percentage increase in the number of computer tablet users from 2014 to 2018.

..... % **[4]**

36(a). The table shows the marks obtained by 10 students in spelling tests in January and February.

Mark in January	26	53	50	48	30	66	70	44	37	38
Mark in February	42	58	68	58	66	82	86	60	48	50

The marks for the first six students are plotted on the scatter diagram.



Plot the marks for the remaining four students. **[2]**

(b). Describe the type of correlation shown in the completed scatter diagram.

[1]

(c).

i. On the scatter diagram, **circle** the student that made the greatest improvement in their marks from January to February.

[1]

ii. Work out the percentage change in this student's marks from January to February.

(ii) ..... % **[3]**

(d). Another student, Kai, scored 79 marks in the test in January but was absent for the test in February.

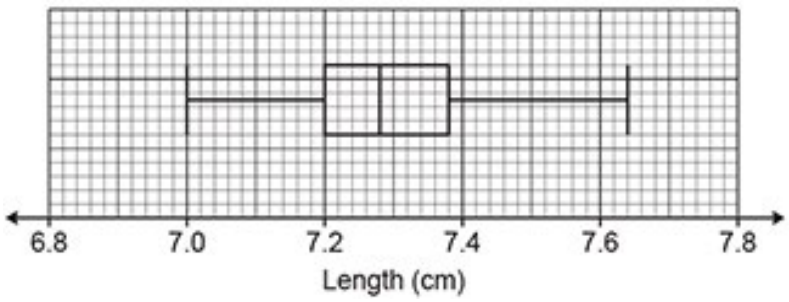
Kai says

I could use a line of best fit on the scatter diagram to estimate the marks I may have achieved in the test in February.

Is Kai's method reliable?
Give a reason for your answer.

[1]

37(a). The box plot shows the distribution of the lengths, in cm, of 60 full-grown mice owned by a pet shop.



Length (cm)

Find the range.

.....cm **[2]**

(b). Work out the number of these mice that have a length of at least 7.2 cm.

..... **[2]**

(c). Sam owns 5 full-grown mice.
Sam picks the third longest mouse and measures its length.
Sam then looks at the box plot.

Sam says

This mouse is 7.35 cm long.
Therefore, the mice I own are longer than the full-grown mice owned by the pet shop.

i. Give a mathematical reason to support Sam's conclusion.

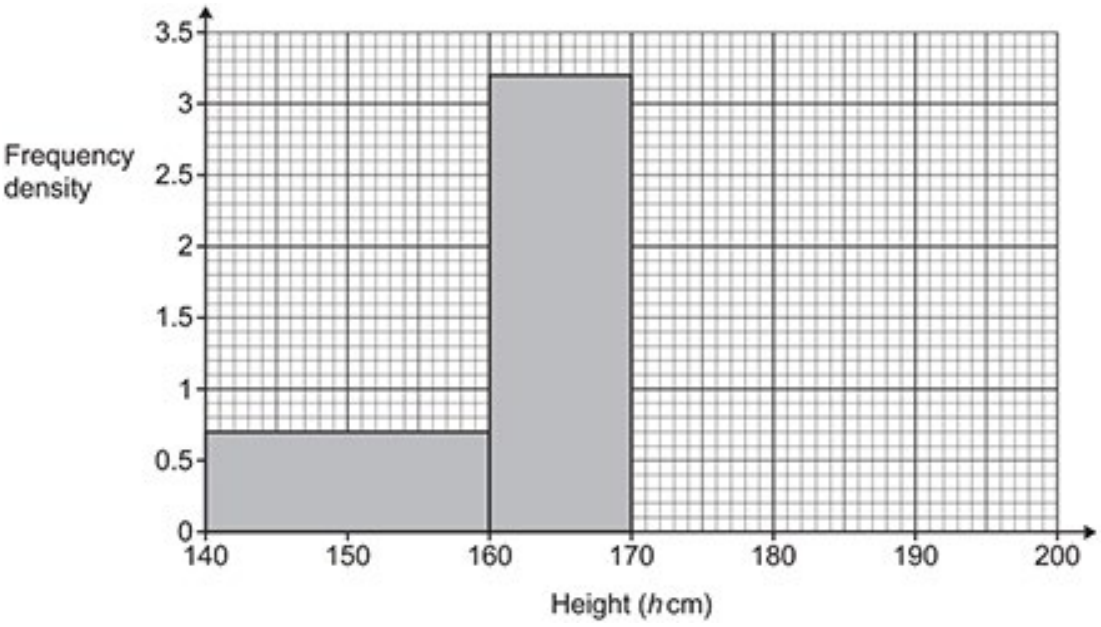
[1]

ii. Give a mathematical reason why Sam's conclusion may be unreliable.

[1]

38. The height, h cm, of each member of a tennis club is recorded.

The histogram shows some of the results.



40% of the members have a height in the interval $160 \leqslant h < 170$.
30% of the members have a height in the interval $170 \leqslant h < 180$.
100% of the members have a height in the interval $140 \leqslant h < 200$.

Complete the histogram for the intervals $170 \leqslant h < 180$ and $180 \leqslant h < 200$. **[6]**

END OF QUESTION PAPER