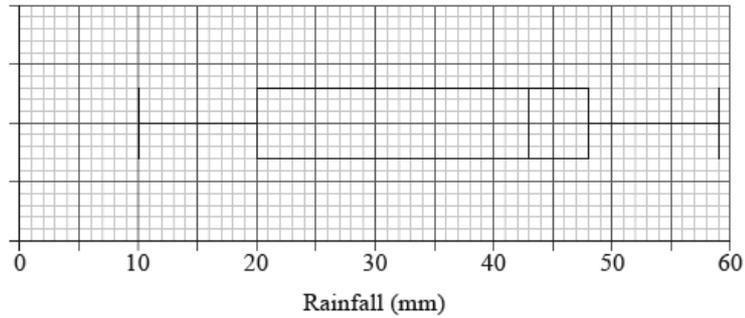


## Topic 2 Part 1 [195 marks]

The distribution of rainfall in a town over 80 days is displayed on the following box-and-whisker diagram.



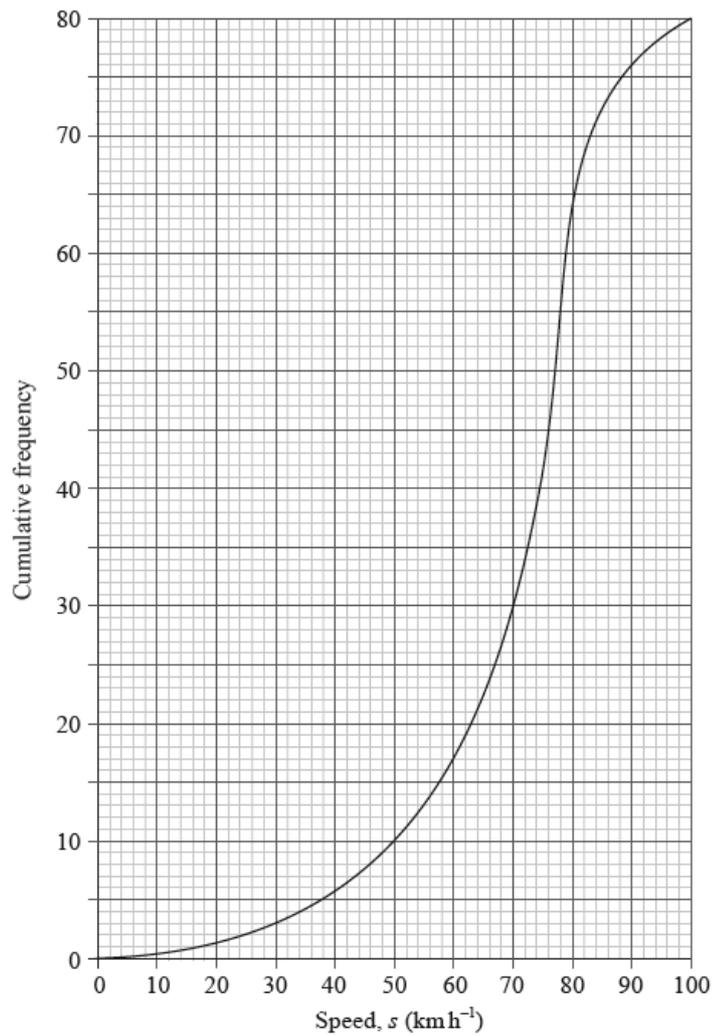
- 1a. Write down the median rainfall. [1 mark]
- 1b. Write down the minimum rainfall. [1 mark]
- 1c. Find the interquartile range. [2 marks]
- 1d. Write down the number of days the rainfall will be [2 marks]
- (i) between 43 mm and 48 mm;
  - (ii) between 20 mm and 59 mm.

A class of 13 Mathematics students received the following grades in their final IB examination.

3 5 3 4 7 3 2 7 5 6 5 3 4

- 2a. For these grades, find the mode; [1 mark]
- 2b. For these grades, find the median; [2 marks]
- 2c. For these grades, find the upper quartile; [1 mark]
- 2d. For these grades, find the interquartile range. [2 marks]

The cumulative frequency graph represents the speed,  $s$ , in  $\text{km h}^{-1}$ , of 80 cars passing a speed camera.



3a. Write down the number of cars passing the camera with speed of less than or equal to 50  $\text{km h}^{-1}$ .

[1 mark]

3b. Complete the following grouped frequency table for  $s$ , the speed of the cars passing the camera.

[1 mark]

$s$ ( $\text{km h}^{-1}$ )	$0 < s \leq 50$	$50 < s \leq 70$	$70 < s \leq 80$	$80 < s \leq 90$	$90 < s \leq 100$
Frequency			34		4

3c. Write down the mid-interval value of the  $50 < s \leq 70$  interval.

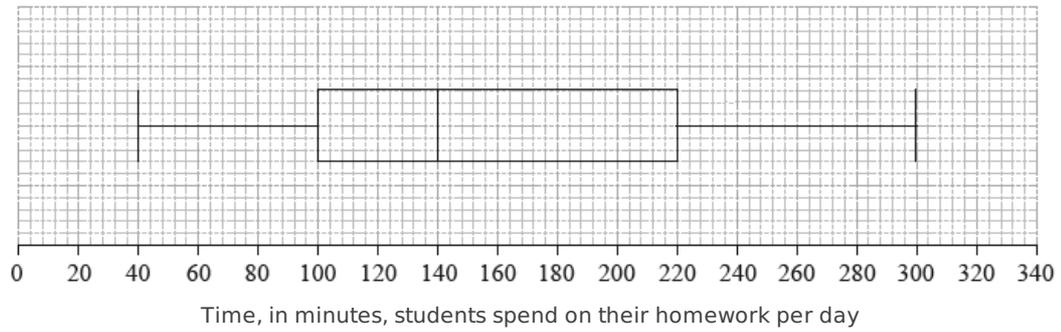
[1 mark]

3d. Use your graphic display calculator to find an estimate of

[3 marks]

- the mean speed of the cars passing the camera;
- the standard deviation of the speed of the cars passing the camera.

The time, in minutes, that students in a school spend on their homework per day is presented in the following box-and-whisker diagram.



4a. Find [3 marks]

- (i) the longest amount of time spent on homework per day;
- (ii) the interquartile range.

.....

.....

.....

4b. State the statistical term corresponding to the value of 140 minutes. [1 mark]

.....

.....

.....

4c. Find the percentage of students who spend

[2 marks]

- (i) between 100 and 140 minutes per day on their homework;
- (ii) more than 100 minutes per day on their homework.

.....

.....

.....

In a particular week, the number of eggs laid by each hen on a farm was counted. The results are summarized in the following table.

<b>Number of eggs</b>	1	2	3	4	5	6
<b>Frequency</b>	4	7	12	10	14	13

5a. State whether these data are discrete or continuous.

[1 mark]

.....

.....

.....

5b. Write down

[2 marks]

- (i) the number of hens on the farm;
- (ii) the modal number of eggs laid.

.....

.....

.....

5c. Calculate

[3 marks]

- (i) the mean number of eggs laid;
- (ii) the standard deviation.

.....

.....

.....

The IB grades attained by a group of students are listed as follows.

6 4 5 3 7 3 5 4 2 5

6a. Find the median grade.

[2 marks]

.....

.....

.....

6b. Calculate the interquartile range.

[2 marks]

.....

.....

.....

6c. Find the probability that a student chosen at random from the group scored at least a grade 4.

[2 marks]

.....

.....

.....

Two groups of 40 students were asked how many books they have read in the last two months. The results for **the first group** are shown in the following table.

Number of books read	Frequency
2	5
3	8
4	13
5	7
6	4
7	2
8	1

The quartiles for these results are 3 and 5.

7a. Write down the value of the median for these results.

[1 mark]

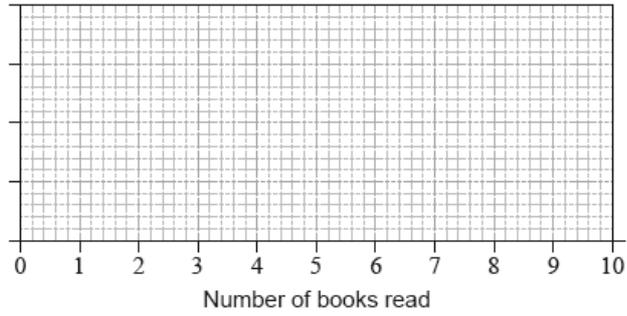
.....

.....

.....

7b. Draw a box-and-whisker diagram for these results on the following grid.

[3 marks]



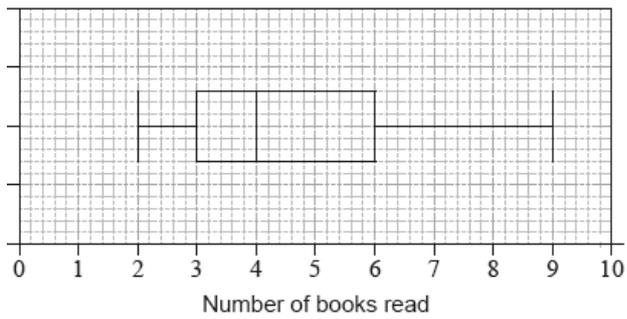
.....

.....

.....

7c. The results for **the second group** of 40 students are shown in the following box-and-whisker diagram.

[2 marks]



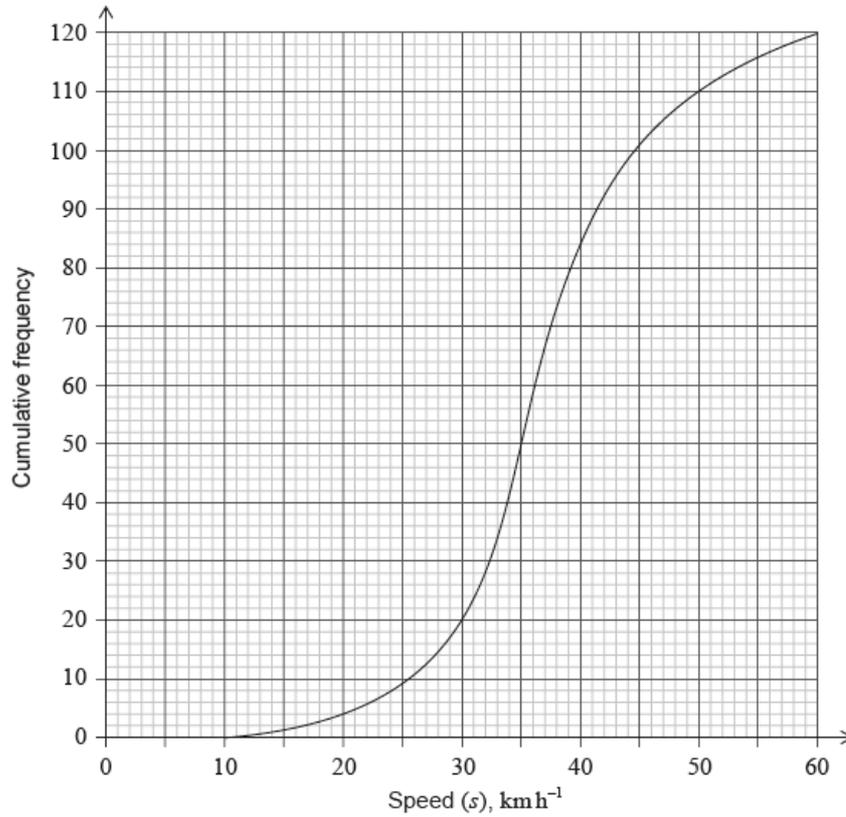
Estimate the number of students **in the second group** who have read at least 6 books.

.....

.....

.....

The cumulative frequency graph shows the speed,  $s$ , in  $\text{km h}^{-1}$ , of 120 vehicles passing a hospital gate.



8a. Estimate the minimum possible speed of one of these vehicles passing the hospital gate. [1 mark]

.....

.....

.....

8b. Find the median speed of the vehicles. [2 marks]

.....

.....

.....

8c. Write down the 75<sup>th</sup> percentile.

[1 mark]

.....

.....

.....

8d. Calculate the interquartile range.

[2 marks]

.....

.....

.....

8e. The speed limit past the hospital gate is  $50 \text{ km h}^{-1}$ .

[2 marks]

Find the number of these vehicles that exceed the speed limit.

.....

.....

.....

8f. The table shows the speeds of these vehicles travelling past the hospital gate.

[2 marks]

Speed of Vehicles	Number of Vehicles
$0 < s \leq 10$	0
$10 < s \leq 20$	$p$
$20 < s \leq 30$	16
$30 < s \leq 40$	64
$40 < s \leq 50$	26
$50 < s \leq 60$	$q$

Find the value of  $p$  and of  $q$ .

.....

.....

.....

8g. The table shows the speeds of these vehicles travelling past the hospital gate.

[2 marks]

Speed of Vehicles	Number of Vehicles
$0 < s \leq 10$	0
$10 < s \leq 20$	$p$
$20 < s \leq 30$	16
$30 < s \leq 40$	64
$40 < s \leq 50$	26
$50 < s \leq 60$	$q$

- (i) Write down the modal class.
- (ii) Write down the mid-interval value for this class.

.....

.....

.....

8h. The table shows the speeds of these vehicles travelling past the hospital gate.

[3 marks]

Speed of Vehicles	Number of Vehicles
$0 < s \leq 10$	0
$10 < s \leq 20$	$p$
$20 < s \leq 30$	16
$30 < s \leq 40$	64
$40 < s \leq 50$	26
$50 < s \leq 60$	$q$

Use your graphic display calculator to calculate an estimate of

- (i) the mean speed of these vehicles;
- (ii) the standard deviation.

.....

.....

.....

8i. It is proposed that the speed limit past the hospital gate is reduced to  $40 \text{ km h}^{-1}$  from the current  $50 \text{ km h}^{-1}$ . [2 marks]

Find the percentage of these vehicles passing the hospital gate that **do not** exceed the current speed limit but **would** exceed the new speed limit.

.....

.....

.....

Daniel grows apples and chooses at random a sample of 100 apples from his harvest.

He measures the diameters of the apples to the nearest cm. The following table shows the distribution of the diameters.

<b>Diameter (to the nearest cm)</b>	5	6	7	8	9
<b>Frequency</b>	15	27	33	17	8

9a. Using your graphic display calculator, write down the value of

[3 marks]

- (i) the mean of the diameters in this sample;
- (ii) the standard deviation of the diameters in this sample.

.....

.....

.....

9b. Daniel assumes that the diameters of all of the apples from his harvest are normally distributed with a mean of 7 [3 marks] cm and a standard deviation of 1.2 cm. He classifies the apples according to their diameters as shown in the following table.

<b>Classification</b>	<b>Diameter (cm)</b>
Small	Diameter < 6.5
Medium	$6.5 \leq \text{Diameter} < a$
Large	Diameter $\geq a$

Calculate the percentage of **small** apples in Daniel's harvest.

.....

.....

.....

- 9c. Daniel assumes that the diameters of all of the apples from his harvest are normally distributed with a mean of 7 [2 marks] cm and a standard deviation of 1.2 cm. He classifies the apples according to their diameters as shown in the following table.

Classification	Diameter (cm)
Small	Diameter < 6.5
Medium	$6.5 \leq \text{Diameter} < a$
Large	Diameter $\geq a$

Of the apples harvested, 5% are **large** apples.

Find the value of  $a$ .

.....

.....

.....

- 9d. Daniel assumes that the diameters of all of the apples from his harvest are normally distributed with a mean of 7 [2 marks] cm and a standard deviation of 1.2 cm. He classifies the apples according to their diameters as shown in the following table.

Classification	Diameter (cm)
Small	Diameter < 6.5
Medium	$6.5 \leq \text{Diameter} < a$
Large	Diameter $\geq a$

Find the percentage of **medium** apples.

.....

.....

.....

- 9e. Daniel assumes that the diameters of all of the apples from his harvest are normally distributed with a mean of 7 [2 marks] cm and a standard deviation of 1.2 cm. He classifies the apples according to their diameters as shown in the following table.

Classification	Diameter (cm)
Small	Diameter < 6.5
Medium	$6.5 \leq \text{Diameter} < a$
Large	Diameter $\geq a$

This year, Daniel estimates that he will grow 100 000 apples.

Estimate the number of **large** apples that Daniel will grow this year.

.....

.....

.....

The table shows the number of bicycles owned by 50 households.

Number of bicycles per household	Frequency (number of households)	Cumulative frequency
0	3	3
1	7	10
2	12	22
3	14	36
4	4	40
5	$t$	$w$
6	2	50

- 10a. Write down the value of

[2 marks]

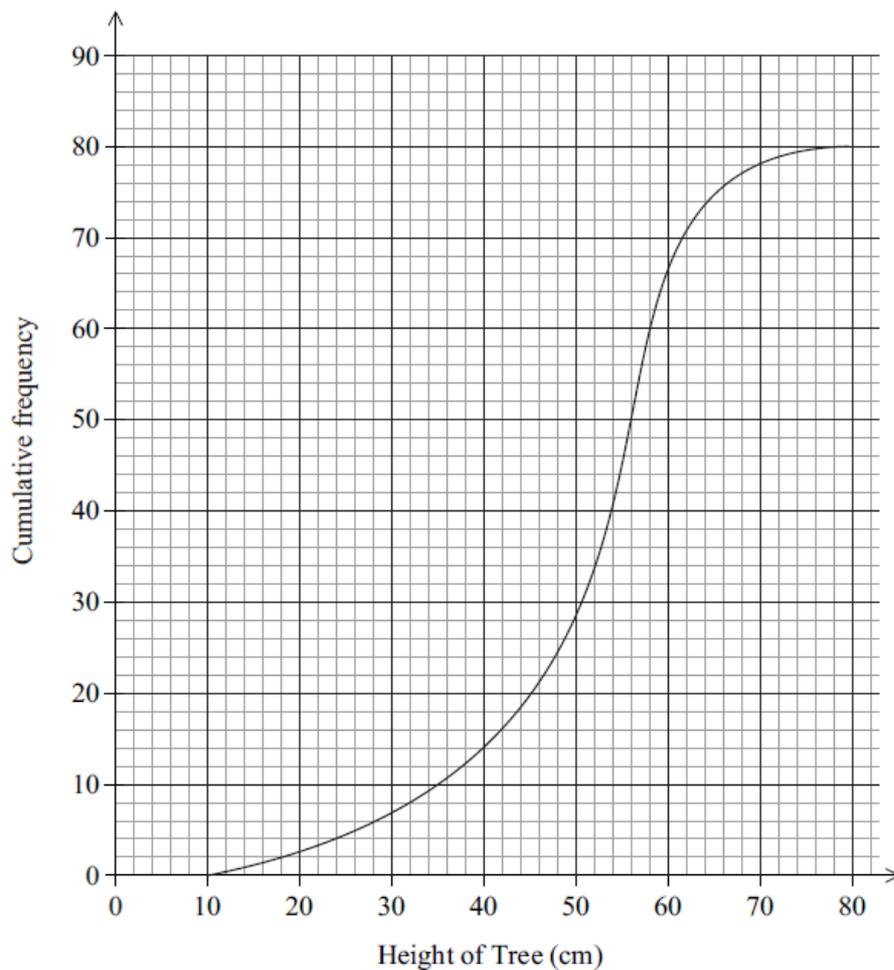
- (i)  $t$  ;  
(ii)  $w$  .

- 10b. Indicate with a tick (✓) whether the following statements are True or False.

[4 marks]

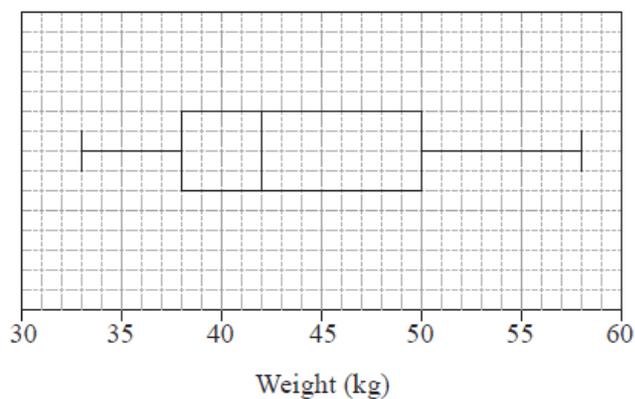
Statement	True	False
Every household owns at least 1 bicycle.		
The median number of bicycles per household is 3.		
The 25 <sup>th</sup> percentile is 1 bicycle per household.		
There are 10 households with at most 1 bicycle.		

The cumulative frequency graph shows the heights, in cm, of 80 young trees.



- 11a. Write down the median height of the trees. [1 mark]
- 11b. Write down the 75<sup>th</sup> percentile. [1 mark]
- 11c. Find the interquartile range. [2 marks]
- 11d. Estimate the number of trees that are more than 40 cm in height. [2 marks]

The weights, in kg, of 60 adolescent females were collected and are summarized in the box and whisker diagram shown below.



- 12a. Write down the median weight of the females. [1 mark]

12b. Calculate the range.

[2 marks]

12c. Estimate the probability that the weight of a randomly chosen female is more than 50 kg.

[1 mark]

12d. Use the box and whisker diagram to determine if the mean weight of the females is less than the median weight. Give a reason [2 marks] for your answer.

The number of passengers in the first ten carriages of a train is listed below.

6, 8, 6, 3, 8, 4, 8, 5,  $p$ ,  $p$

The mean number of passengers per carriage is 5.6.

13a. Calculate the value of  $p$ .

[2 marks]

13b. Find the median number of passengers per carriage.

[2 marks]

13c. If the passengers in the eleventh carriage are also included, the mean number of passengers per carriage increases to 6.0.

[2 marks]

Determine the number of passengers in the eleventh carriage of the train.

A survey was conducted to determine the length of time,

$t$ , in minutes, people took to drink their coffee in a café. The information is shown in the following grouped frequency table.

Time, $t$ (minutes)	Number of People
$0 < t \leq 5$	3
$5 < t \leq 10$	5
$10 < t \leq 15$	12
$15 < t \leq 20$	14
$20 < t \leq 25$	16
$25 < t \leq 30$	10

14a. Write down the total number of people who were surveyed.

[1 mark]

14b. Write down the mid-interval value for the

$10 < t \leq 15$  group.

[1 mark]

14c. Find an estimate of the mean time people took to drink their coffee.

[2 marks]

14d. The information above has been rewritten as a cumulative frequency table.

[2 marks]

Time, $t$ (minutes)	$t \leq 5$	$t \leq 10$	$t \leq 15$	$t \leq 20$	$t \leq 25$	$t \leq 30$
Cumulative frequency	3	8	20	$a$	50	$b$

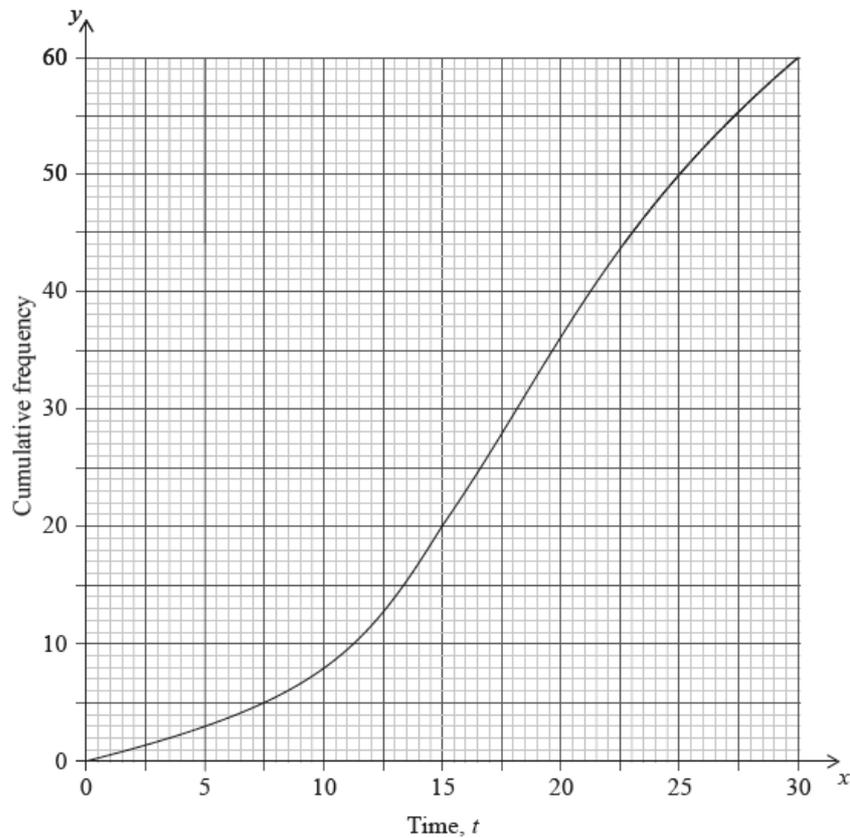
Write down the value of

$a$  and the value of

$b$ .

14e. This information is shown in the following cumulative frequency graph.

[4 marks]



For the people who were surveyed, use the graph to estimate

- (i) the time taken for the first 40 people to drink their coffee;
- (ii) the number of people who take less than 8 minutes to drink their coffee;
- (iii) the number of people who take more than 23 minutes to drink their coffee.

A survey was carried out on a road to determine the number of passengers in each car (excluding the driver). The table shows the results of the survey.

<b>Number of passengers</b>	0	1	2	3	4
<b>Number of cars</b>	37	23	36	15	9

15a. State whether the data is discrete or continuous.

[1 mark]

15b. Write down the mode.

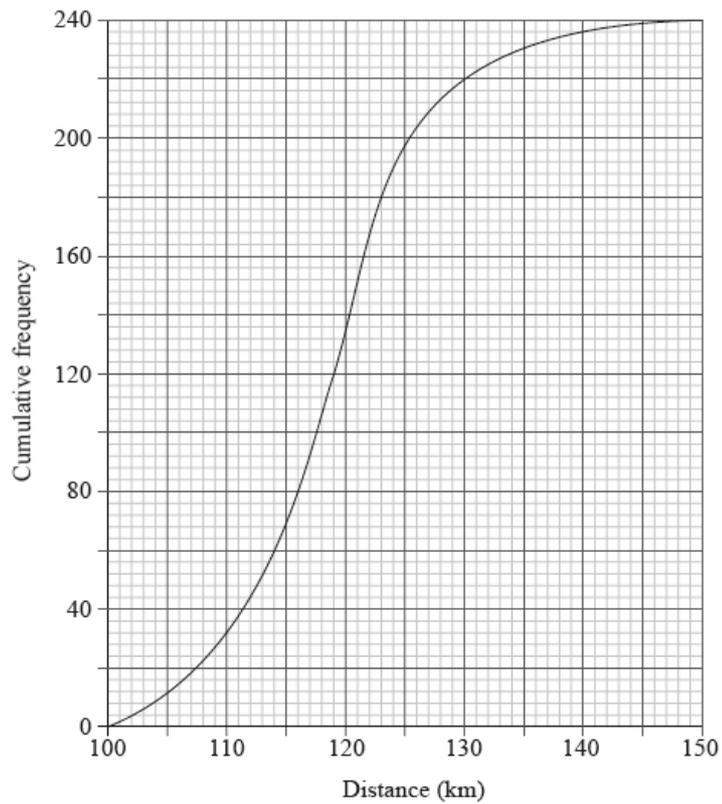
[1 mark]

15c. Use your graphic display calculator to find

[4 marks]

- (i) the mean number of passengers per car;
- (ii) the median number of passengers per car;
- (iii) the standard deviation.

240 cars were tested to see how far they travelled on 10 litres of fuel. The graph shows the cumulative frequency distribution of the results.



16a. Find the median distance travelled by the cars.

[2 marks]

16b. Calculate the interquartile range of the distance travelled by the cars.

[2 marks]

16c. Find the number of cars that travelled more than 130 km.

[2 marks]

The table shows the distance, in km, of eight regional railway stations from a city centre terminus and the price, in \$, of a return ticket from each regional station to the terminus.

<b>Distance in km (<math>x</math>)</b>	3	15	23	42	56	62	74	93
<b>Price in \$ (<math>y</math>)</b>	5	24	43	56	68	74	86	100

17a. Draw a scatter diagram for the above data. Use a scale of

[4 marks]

1 cm to represent  
10 km on the  
 $x$ -axis and  
1 cm to represent  
\$10 on the  
 $y$ -axis.

- 17b. Use your graphic display calculator to find [2 marks]
- (i)  $\bar{x}$ , the mean of the distances;
  - (ii)  $\bar{y}$ , the mean of the prices.

- 17c. Plot and label the point [1 mark]
- M ( $\bar{x}$ ,  $\bar{y}$ ) on your scatter diagram.

- 17d. Use your graphic display calculator to find [3 marks]
- (i) the product-moment correlation coefficient,  $r$ ;
  - (ii) the equation of the regression line  $y$  on  $x$ .

- 17e. Draw the regression line [2 marks]
- $y$  on  $x$  on your scatter diagram.

- 17f. A ninth regional station is [3 marks]
- 76 km from the city centre terminus.
- Use the equation of the regression line to estimate the price of a return ticket to the city centre terminus from this regional station. **Give your answer correct to the nearest \$.**

- 17g. Give a reason why it is valid to use your regression line to estimate the price of this return ticket. [1 mark]

- 17h. The actual price of the return ticket is [2 marks]
- \$80.
- Using your answer to part (f)**, calculate the percentage error in the estimated price of the ticket.

As part of his IB Biology field work, Barry was asked to measure the circumference of trees, in centimetres, that were growing at different distances, in metres, from a river bank. His results are summarized in the following table.

<b>Distance, <math>x</math> (metres)</b>	5	12	17	21	24	30	34	44	47
<b>Circumference, <math>y</math> (centimetres)</b>	82	76	70	68	67	60	62	50	50

- 18a. State whether *distance from the river bank* is a continuous **or** discrete variable. [1 mark]

- 18b. **On graph paper**, draw a scatter diagram to show Barry's results. Use a scale of 1 cm to represent 5 m on the  $x$ -axis and 1 cm to represent 10 cm on the  $y$ -axis. [4 marks]

- 18c. Write down [2 marks]
- (i) the mean distance,  $\bar{x}$ , of the trees from the river bank;
  - (ii) the mean circumference,  $\bar{y}$ , of the trees.

18d. Plot and label the point  $M(\bar{x}, \bar{y})$  on your graph. [2 marks]

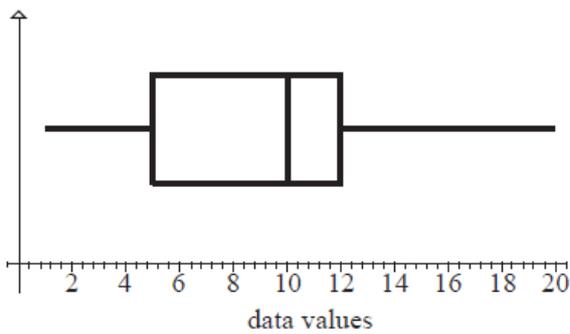
18e. Write down [4 marks]  
(i) the Pearson's product-moment correlation coefficient,  $r$ , for Barry's results;  
(ii) the equation of the regression line  $y$  on  $x$ , for Barry's results.

18f. Draw the regression line  $y$  on  $x$  on your graph. [2 marks]

18g. Use the equation of the regression line  $y$  on  $x$  to estimate the circumference of a tree that is 40 m from the river bank. [2 marks]

19a. State which of the following sets of data are discrete. [2 marks]  
(i) Speeds of cars travelling along a road.  
(ii) Numbers of members in families.  
(iii) Maximum daily temperatures.  
(iv) Heights of people in a class measured to the nearest cm.  
(v) Daily intake of protein by members of a sporting team.

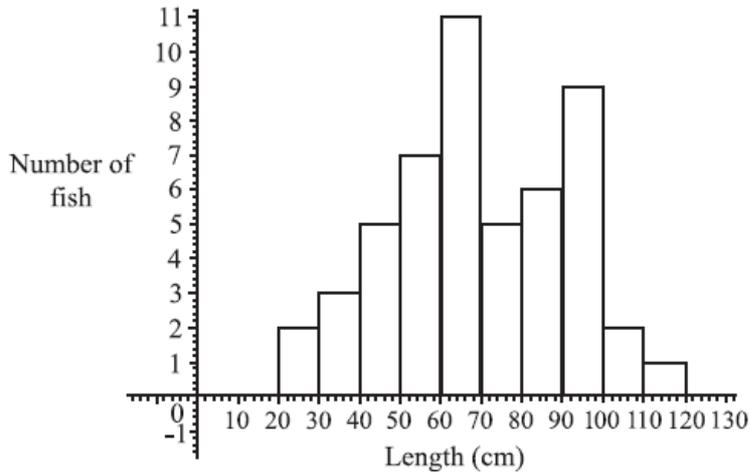
19b. The boxplot below shows the statistics for a set of data. [3 marks]



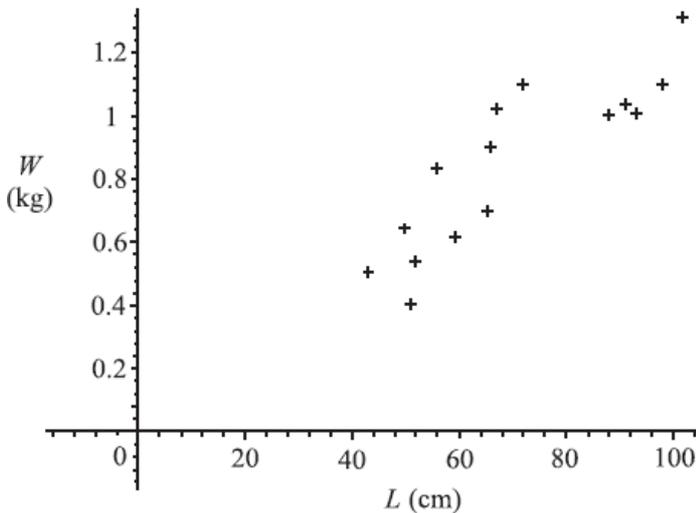
For this data set write down the value of  
(i) the median  
(ii) the upper quartile  
(iii) the minimum value present

19c. Write down three different integers whose mean is 10. [1 mark]

The figure below shows the lengths in centimetres of fish found in the net of a small trawler.



- 20a. Find the total number of fish in the net. [2 marks]
- 20b. Find (i) the modal length interval, [5 marks]  
(ii) the interval containing the median length,  
(iii) an estimate of the mean length.
- 20c. (i) Write down an estimate for the standard deviation of the lengths. [3 marks]  
(ii) How many fish (if any) have length **greater than** three standard deviations **above** the mean?
- 20d. The fishing company must pay a fine if more than 10% of the catch have lengths less than 40cm. [2 marks]  
Do a calculation to decide whether the company is fined.
- 20e. A sample of 15 of the fish was weighed. The weight,  $W$  was plotted against length,  $L$  as shown below. [2 marks]

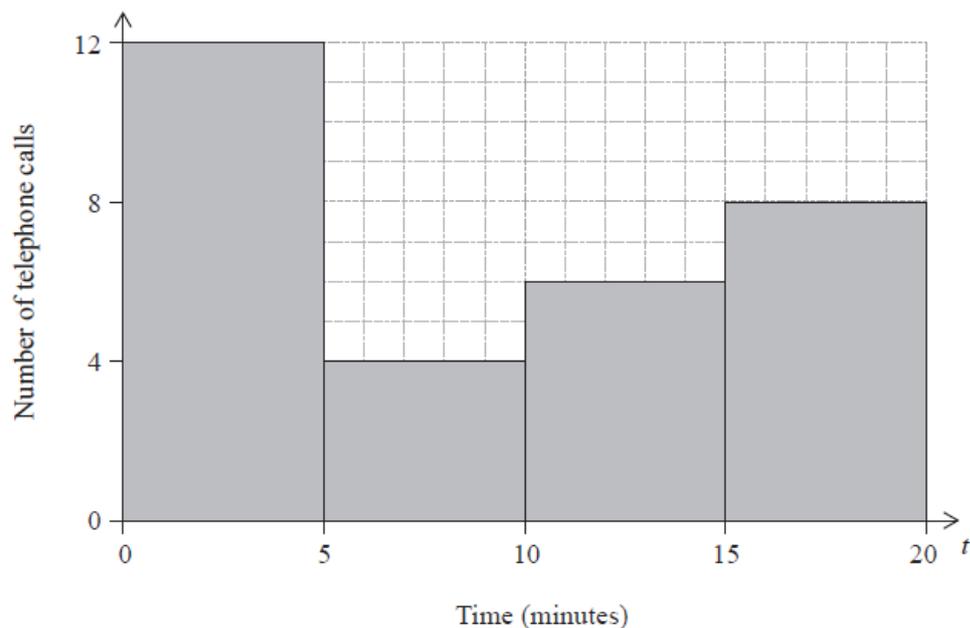


Exactly **two** of the following statements about the plot could be correct. Identify the two correct statements.

**Note:** You do **not** need to enter data in a GDC **or** to calculate  $r$  exactly.

- (i) The value of  $r$ , the correlation coefficient, is approximately 0.871.
- (ii) There is an exact linear relation between  $W$  and  $L$ .
- (iii) The line of regression of  $W$  on  $L$  has equation  $W = 0.012L + 0.008$ .
- (iv) There is negative correlation between the length and weight.
- (v) The value of  $r$ , the correlation coefficient, is approximately 0.998.
- (vi) The line of regression of  $W$  on  $L$  has equation  $W = 63.5L + 16.5$ .

Consider the frequency histogram for the distribution of the time,  $t$ , in minutes of telephone calls that Helen made last week.



21a. Complete the frequency table for this distribution.

[2 marks]

Time (minutes)	Number of telephone calls
$0 < t \leq 5$	
$5 < t \leq 10$	
$10 < t \leq 15$	
$15 < t \leq 20$	

21b. Write down the modal class.

[1 mark]

21c. Write down the mid interval value of the  $10 < t \leq 15$  class.

[1 mark]

21d. Use your graphic display calculator to find an estimate for the mean time.

[2 marks]

Five pipes labelled, “6 metres in length”, were delivered to a building site. The contractor measured each pipe to check its length (in metres) and recorded the following;

5.96, 5.95, 6.02, 5.95, 5.99.

22a. (i) Find the mean of the contractor’s measurements.

[3 marks]

(ii) Calculate the percentage error between the mean and the stated, **approximate** length of 6 metres.

22b. Calculate  $\sqrt{3.87^5 - 8.73^{-0.5}}$ , giving your answer

[3 marks]

(i) correct to the nearest integer,

(ii) in the form

$a \times 10^k$ , where  $1 \leq a < 10$ ,

$k \in \mathbb{Z}$ .

23a. Write down the following numbers in increasing order.

[3 marks]

- 3.5,
- $1.6 \times 10^{-19}$ ,
- 60730,
- $6.073 \times 10^5$ ,
- $0.006073 \times 10^6$ ,
- $\pi$ ,
- $9.8 \times 10^{-18}$ .

23b. Write down the median of the numbers in part (a).

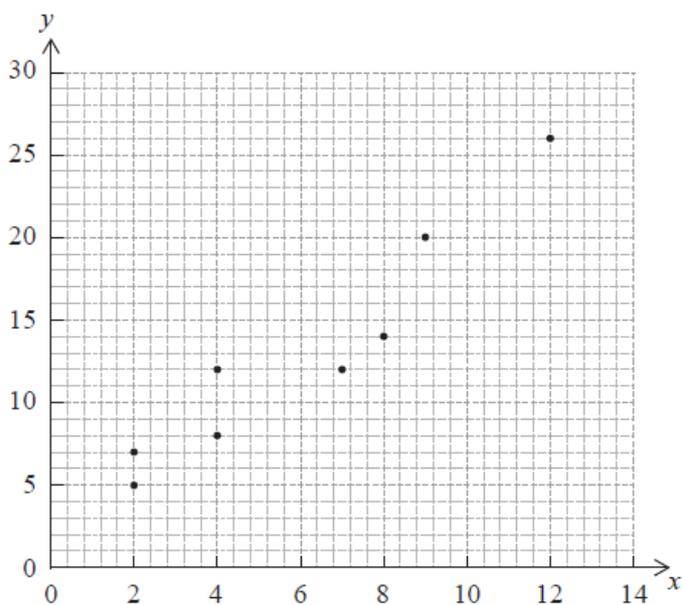
[1 mark]

23c. State which of the numbers in part (a) is irrational.

[1 mark]

Consider the following set of data which is plotted on the scatter diagram below.

$x$	2	4	7	12	4	8	9	2
$y$	5	8	12	26	12	14	20	7



24a. Write down the coordinates of the mean point  $(\bar{x}, \bar{y})$ .

[2 marks]

24b. Write down the value of  $r$ , the Pearson's product-moment correlation coefficient for this set of data.

[2 marks]

24c. Draw the regression line for  $y$  on  $x$  on the set of axes above.

[2 marks]