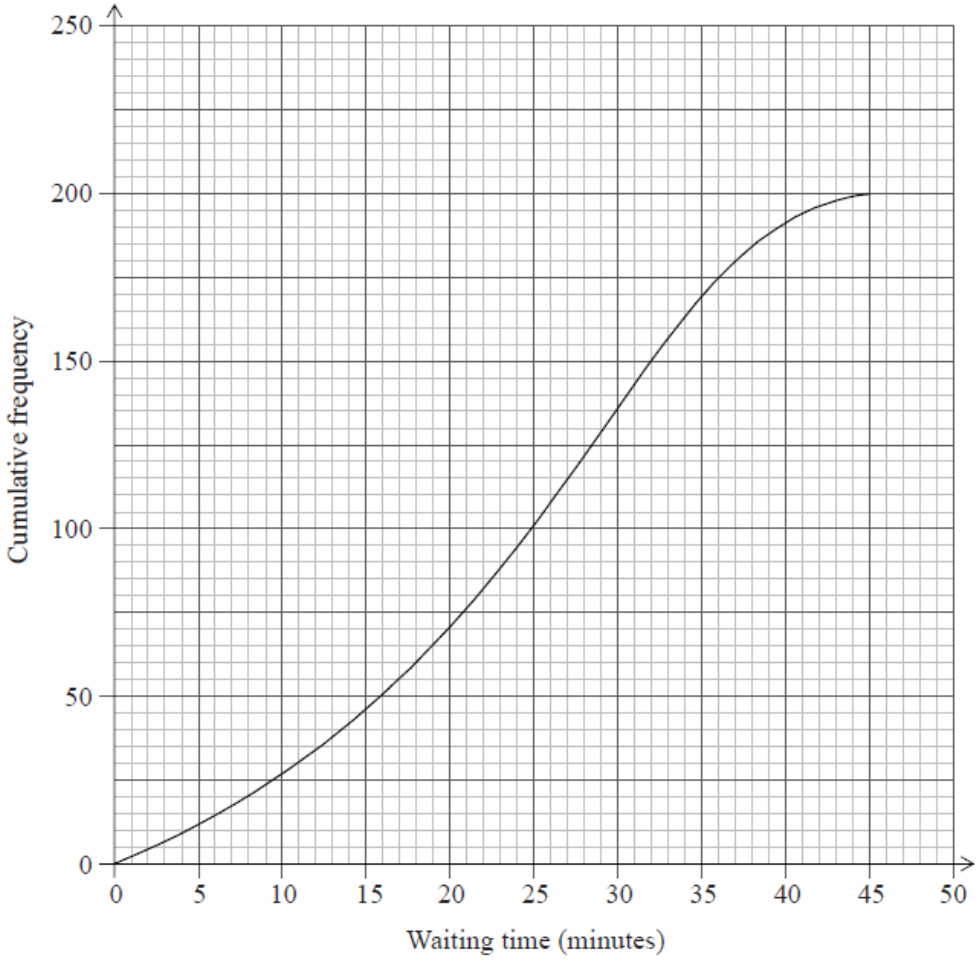
Topic 2 Part 2 *[165 marks]*

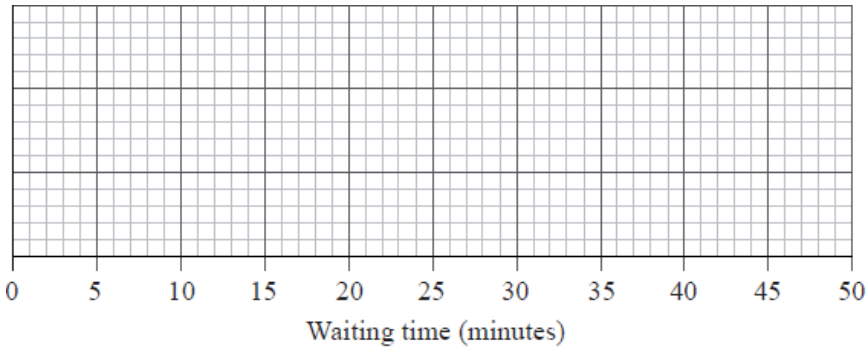
The cumulative frequency graph shows the amount of time in minutes, 200 students spend waiting for their train on a particular morning.



1a. Write down the median waiting time. *[1 mark]*

1b. Find the interquartile range for the waiting time. *[2 marks]*

1c. Draw a box and whisker plot on the grid below to represent this information. *[3 marks]*



The mean of the ten numbers listed below is 6.8.

8, 5, 5, 10, 8, 4, 9, 7, p, q

2a. Write down an equation in terms of p and q. *[2 marks]*

2b. The mode of these ten numbers is five and p is less than q. *[1 mark]*

Write down the value of p.

2c. The mode of these ten numbers is five and p is less than q. *[1 mark]*

Write down the value of q.

2d. Find the median of the ten numbers. *[2 marks]*

In an environmental study of plant diversity around a lake, a biologist collected data about the number of different plant species (y) that were growing at different distances (x) in metres from the lake shore.

Distance (x)	2	5	8	10	13	17	23	35	40
Plant species (y)	35	34	30	29	24	19	15	13	8

3a. Draw a scatter diagram to show the data. Use a scale of 2 cm to represent 10 metres on the x-axis and 2 cm to represent 10 plant species on the y-axis. *[4 marks]*

3b. Using your scatter diagram, describe the correlation between the number of different plant species and the distance from the lake shore. *[1 mark]*

3c. Use your graphic display calculator to write down $\bar{x}$, the mean of the distances from the lake shore. *[1 mark]*

3d. Use your graphic display calculator to write down $\bar{y}$, the mean number of plant species. *[1 mark]*

3e. Plot the point ($\bar{x}$, $\bar{y}$) on your scatter diagram. **Label this point M.** *[2 marks]*

3f. Write down the equation of the regression line y on x for the above data. *[2 marks]*

3g. Draw the regression line y on x on your scatter diagram. *[2 marks]*

3h. Estimate the number of plant species growing 30 metres from the lake shore. *[2 marks]*

Alex and Kris are riding their bicycles together along a bicycle trail and note the following distance markers at the given times.

Time (t hours)	1	2	3	4	5	6	7
Distance (d km)	57	65	72	81	89	97	107

4a. Draw a scatter diagram of the data. Use 1 cm to represent 1 hour and 1 cm to represent 10 km. *[3 marks]*

4b. Write down for this set of data the mean time, $\bar{t}$. *[1 mark]*

4c. Write down for this set of data the mean distance, $\bar{d}$. *[1 mark]*

4d. Mark and label the point $M(\bar{t}, \bar{d})$ on your scatter diagram. *[2 marks]*

4e. Draw the line of best fit on your scatter diagram. *[2 marks]*

4f. **Using your graph**, estimate the time when Alex and Kris pass the 85 km distance marker. Give your answer correct to **one decimal place**. *[2 marks]*

4g. Write down the equation of the regression line for the data given. *[2 marks]*

4h. **Using your equation** calculate the distance marker passed by the cyclists at 10.3 hours. *[2 marks]*

4i. Is this estimate of the distance reliable? Give a reason for your answer. *[2 marks]*

The temperatures in °C, at midday in Geneva, were measured for eight days and the results are recorded below.

7, 4, 5, 4, 8, *T*, 14, 4

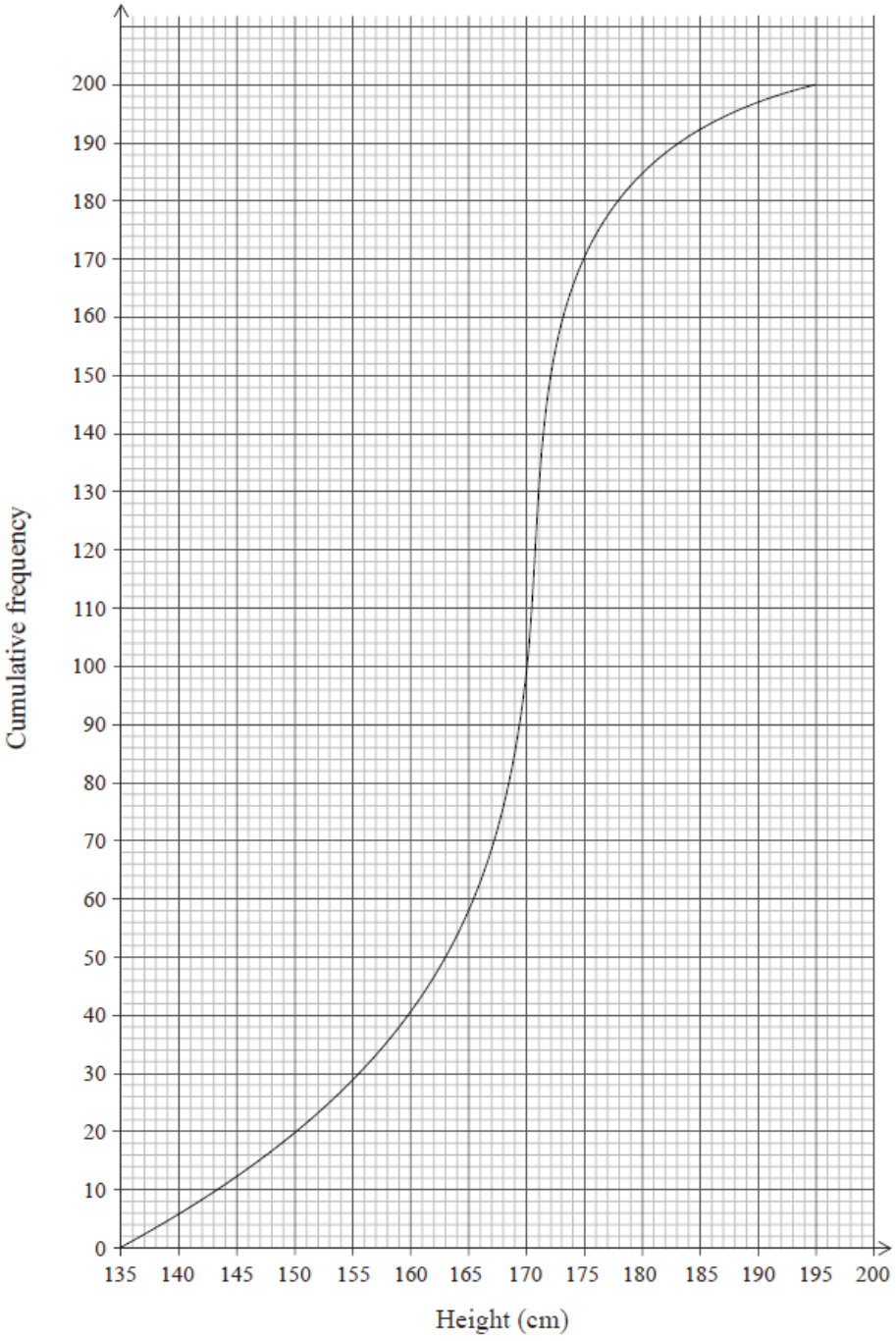
The mean temperature was found to be 7 °C.

5a. Find the value of *T*. *[3 marks]*

5b. Write down the mode. *[1 mark]*

5c. Find the median. *[2 marks]*

A cumulative frequency graph is given below which shows the height of students in a school.



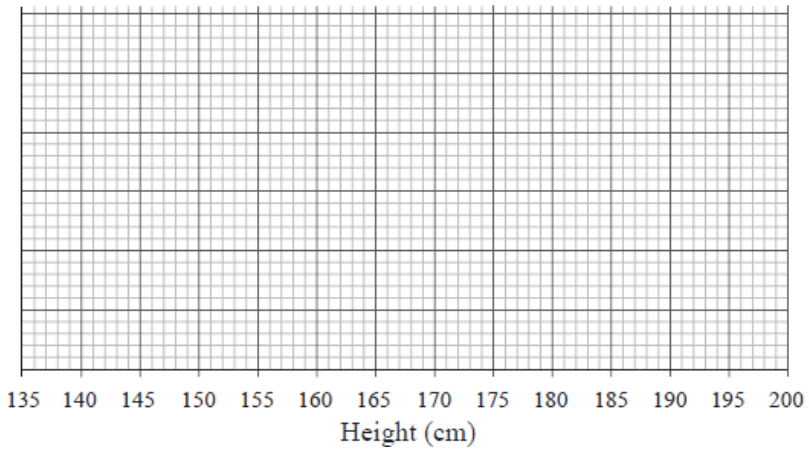
6a. Write down the median height of the students. *[1 mark]*

6b. Write down the 25^{th} percentile. *[1 mark]*

6c. Write down the 75^{th} percentile. *[1 mark]*

6d. The height of the tallest student is 195 cm and the height of the shortest student is 136 cm. *[3 marks]*

Draw a box and whisker plot on the grid below to represent the heights of the students in the school.



The diagram below shows a square based right pyramid. ABCD is a square of side 10 cm. VX is the perpendicular height of 8 cm. M is the midpoint of BC.

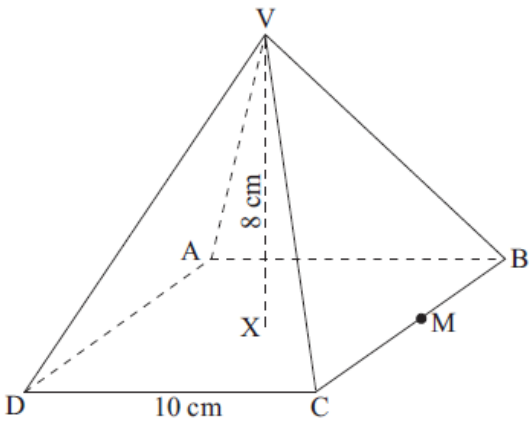


diagram not to scale

7a. Write down the length of XM. *[1 mark]*

In a mountain region there appears to be a relationship between the number of trees growing in the region and the depth of snow in winter. A set of 10 areas was chosen, and in each area the number of trees was counted and the depth of snow measured. The results are given in the table below.

Number of trees (x)	Depth of snow in cm (y)
45	30
75	50
66	40
27	25
44	30
28	5
60	35
35	20
73	45
47	25

7b. Use your graphic display calculator to find the standard deviation of the number of trees. *[1 mark]*

7c. Calculate the length of VM. *[2 marks]*

7d. Calculate the angle between VM and ABCD. *[2 marks]*

A path goes around a forest so that it forms the three sides of a triangle. The lengths of two sides are 550 m and 290 m. These two sides meet at an angle of 115°. A diagram is shown below.

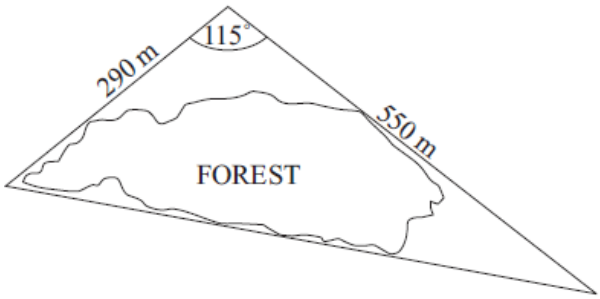


diagram not to scale

7e. Calculate the length of the third side of the triangle. Give your answer correct to the nearest 10 m. *[4 marks]*

7f. Calculate the area enclosed by the path that goes around the forest. *[3 marks]*

7g. Inside the forest a second path forms the three sides of another triangle named ABC. Angle BAC is 53°, AC is 180 m and BC is 230 m. *[4 marks]*

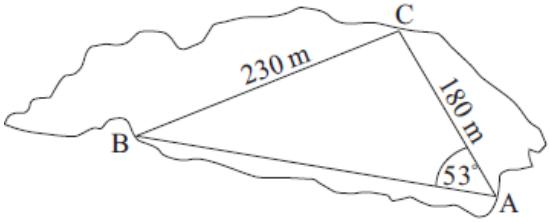


diagram not to scale

Calculate the size of angle ACB.

8a. Use your graphic display calculator to find the mean number of trees. *[1 mark]*

8b. Use your graphic display calculator to find the mean depth of snow. *[1 mark]*

8c. Use your graphic display calculator to find the standard deviation of the depth of snow. *[1 mark]*

8d. The covariance, $S_{xy} = 188.5$. *[2 marks]*

Write down the product-moment correlation coefficient, r.

8e. Write down the equation of the regression line of y on x. *[2 marks]*

8f. If the number of trees in an area is 55, estimate the depth of snow. *[2 marks]*

8g. Use the equation of the regression line to estimate the depth of snow in an area with 100 trees. *[1 mark]*

8h. Decide whether the answer in (e)(i) is a valid estimate of the depth of snow in the area. Give a reason for your answer. *[2 marks]*

In a study on
100 students there seemed to be a difference between males and females in their choice of favourite car colour. The results are given in the table below. A
χ^2 test was conducted.

	Blue	Red	Green
Males	14	6	8
Females	31	24	17

8i. Write down the total number of male students. *[1 mark]*

8j. Show that the expected frequency for males, whose favourite car colour is blue, is 12.6. *[2 marks]*

8k. The calculated value of *[1 mark]*
χ^2 is
1.367 and the critical value of
χ^2 is
5.99 at the
5% significance level.

Write down the null hypothesis for this test.

8l. The calculated value of *[1 mark]*
χ^2 is
1.367 and the critical value of
χ^2 is
5.99 at the
5% significance level.

Write down the number of degrees of freedom.

8m. The calculated value of *[2 marks]*
χ^2 is
1.367 and the critical value of
χ^2 is
5.99 at the
5% significance level.

Determine whether the null hypothesis should be accepted at the
5% significance level. Give a reason for your answer.

Eight houses in a street are inhabited by different numbers of people, as shown in the table below.

House	A	B	C	D	E	F	G	H
Number of inhabitants	5	4	7	6	4	3	6	4

The following statements refer to the number of inhabitants per house. Write down true (T) or false (F) for each.

9a. The mean is *[1 mark]*
5.

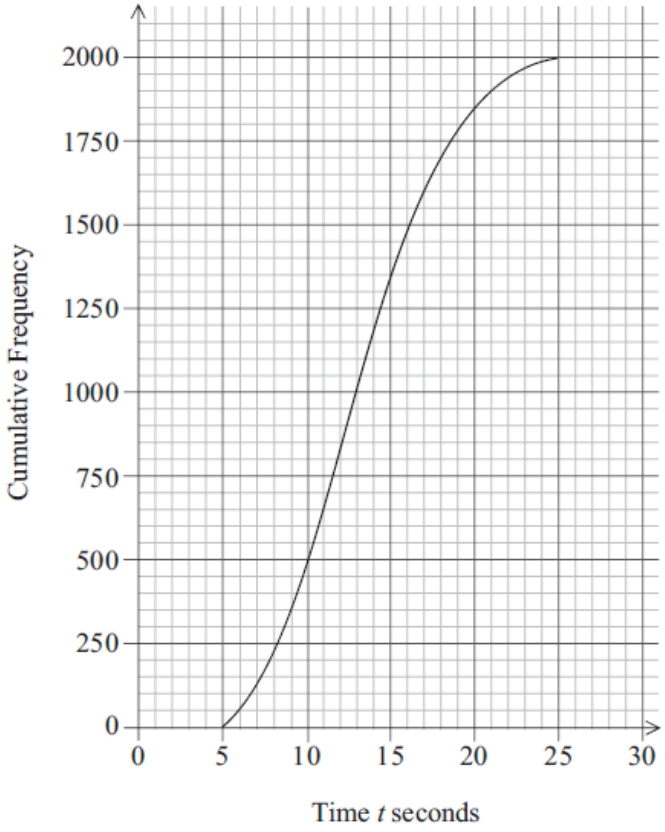
9b. The range is *[1 mark]*
4.

9c. The mode is *[1 mark]*
6.

9d. The standard deviation is 1.4 correct to 2 significant figures. *[1 mark]*

9e. Calculate the interquartile range for the number of inhabitants per house. *[2 marks]*

The diagram shows the cumulative frequency graph for the time t taken to perform a certain task by 2000 men.



10a. Use the diagram to estimate the median time. *[1 mark]*

10b. Use the diagram to estimate the upper quartile and the lower quartile. *[2 marks]*

10c. Use the diagram to estimate the interquartile range. *[1 mark]*

10d. Find the number of men who take **more than** 11 seconds to perform the task. *[3 marks]*

10e. 55 % of the men took less than p seconds to perform the task. Find p. *[2 marks]*

10f. The times taken for the 2000 men were grouped as shown in the table below. *[1 mark]*

Time	Frequency
$5 \le t < 10$	500
$10 \le t < 15$	850
$15 \le t < 20$	a
$20 \le t < 25$	b

Write down the value of a.

10g. The times taken for the 2000 men were grouped as shown in the table below. *[1 mark]*

Write down the value of b.

10h. Use your graphic display calculator to find an estimate of the mean time. *[2 marks]*

10i. Use your graphic display calculator to find an estimate of the standard deviation of the time. *[1 mark]*

10j. Everyone who performs the task in **less than** one standard deviation **below** the mean will receive a bonus. Pedro takes 9.5 seconds to perform the task. *[3 marks]*

Does Pedro receive the bonus? Justify your answer.

The following table shows the number of errors per page in a 100 page document.

Number of errors	0	1	2	3	4
Number of pages	28	24	20	17	11

11a. State whether the data is discrete, continuous or neither. *[1 mark]*

11b. Find the mean number of errors per page. *[2 marks]*

11c. Find the median number of errors per page. *[2 marks]*

11d. Write down the mode. *[1 mark]*

The grades obtained by a group of 20 IB students are listed below:

6	2	5	3	5	5	6	2	6	1
7	6	2	4	2	4	3	4	5	6

12a. Complete the following table for the grades obtained by the students. *[2 marks]*

Grade	Frequency
1	
2	
3	2
4	
5	4
6	
7	1

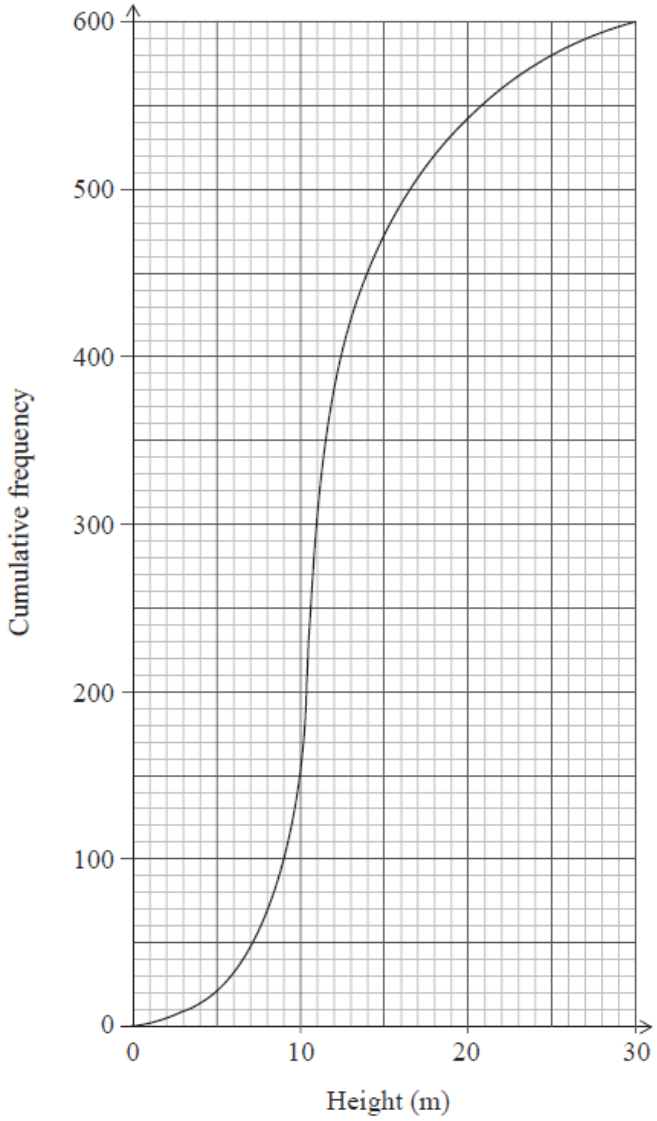
12b. Write down the modal grade obtained by the students. *[1 mark]*

12c. Calculate the median grade obtained by the students. *[2 marks]*

12d. One student is chosen at random from the group. *[1 mark]*

Find the probability that this student obtained either grade 4 or grade 5.

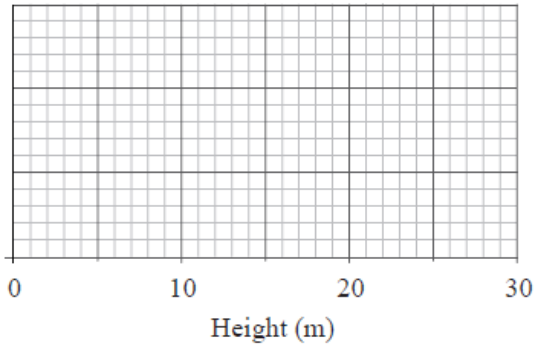
The diagram below shows the cumulative frequency distribution of the heights in metres of 600 trees in a wood.



13a. Write down the median height of the trees. *[1 mark]*

13b. Calculate the interquartile range of the heights of the trees. *[2 marks]*

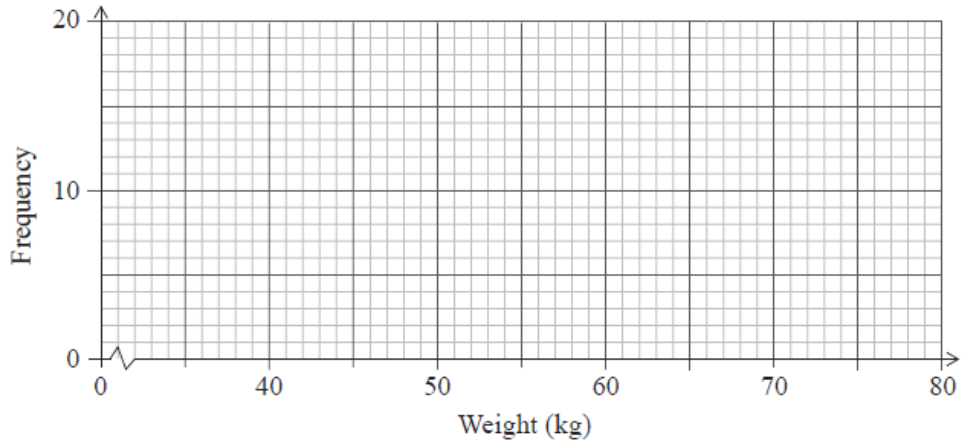
13c. Given that the smallest tree in the wood is 3 m high and the tallest tree is 28 m high, draw the box and whisker plot on the grid below that shows the distribution of trees in the wood. *[3 marks]*



The distribution of the weights, correct to the nearest kilogram, of the members of a football club is shown in the following table.

Weight (kg)	40 – 49	50 – 59	60 – 69	70 – 79
Frequency	6	18	14	4

14a. On the grid below draw a histogram to show the above weight distribution. *[2 marks]*

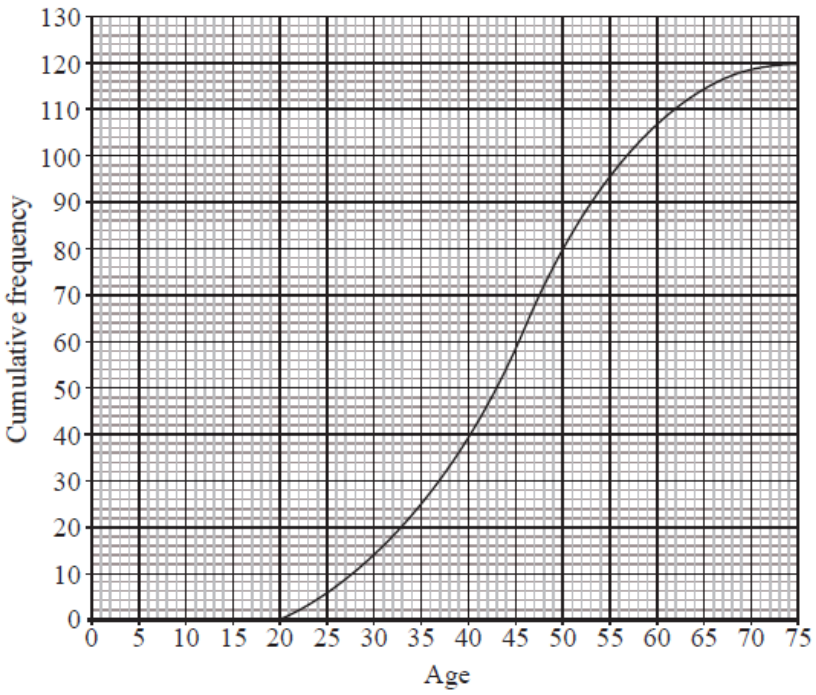


14b. Write down the mid-interval value for the 40 – 49 interval. *[1 mark]*

14c. Find an estimate of the mean weight of the members of the club. *[2 marks]*

14d. Write down an estimate of the standard deviation of their weights. *[1 mark]*

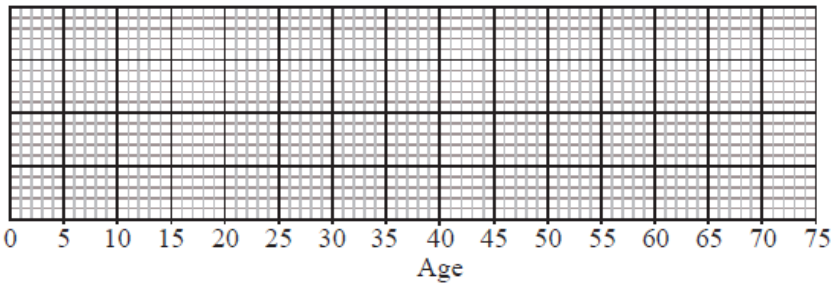
There are 120 teachers in a school. Their ages are represented by the cumulative frequency graph below.



15a. Write down the median age. *[1 mark]*

15b. Find the interquartile range for the ages. *[2 marks]*

15c. Given that the youngest teacher is 21 years old and the oldest is 72 years old, represent the information on a box and whisker plot using the scale below. *[3 marks]*



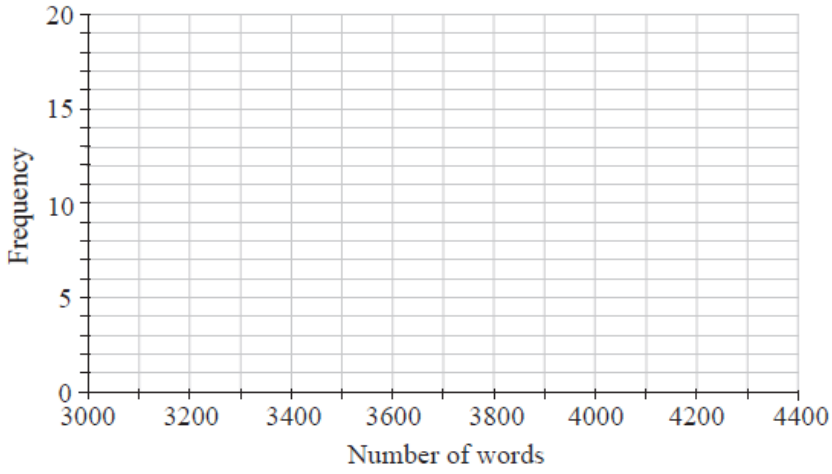
16. Complete the following table of values for the height and weight of seven students. *[4 marks]*

	Values	Mode	Median	Mean	Standard deviation
Height (cm)	151, 158, 171, 163, 184, 148, 171			164	11.7
Weight (kg)	53, 61, 58, 82, 45, 72, 82	82	61		

The table below shows the number of words in the extended essays of an IB class.

Number of words	$3200 \leq w < 3400$	$3400 \leq w < 3600$	$3600 \leq w < 3800$	$3800 \leq w < 4000$	$4000 \leq w < 4200$
Frequency	2	5	8	17	3

17a. Draw a histogram on the grid below for the data in this table. *[3 marks]*

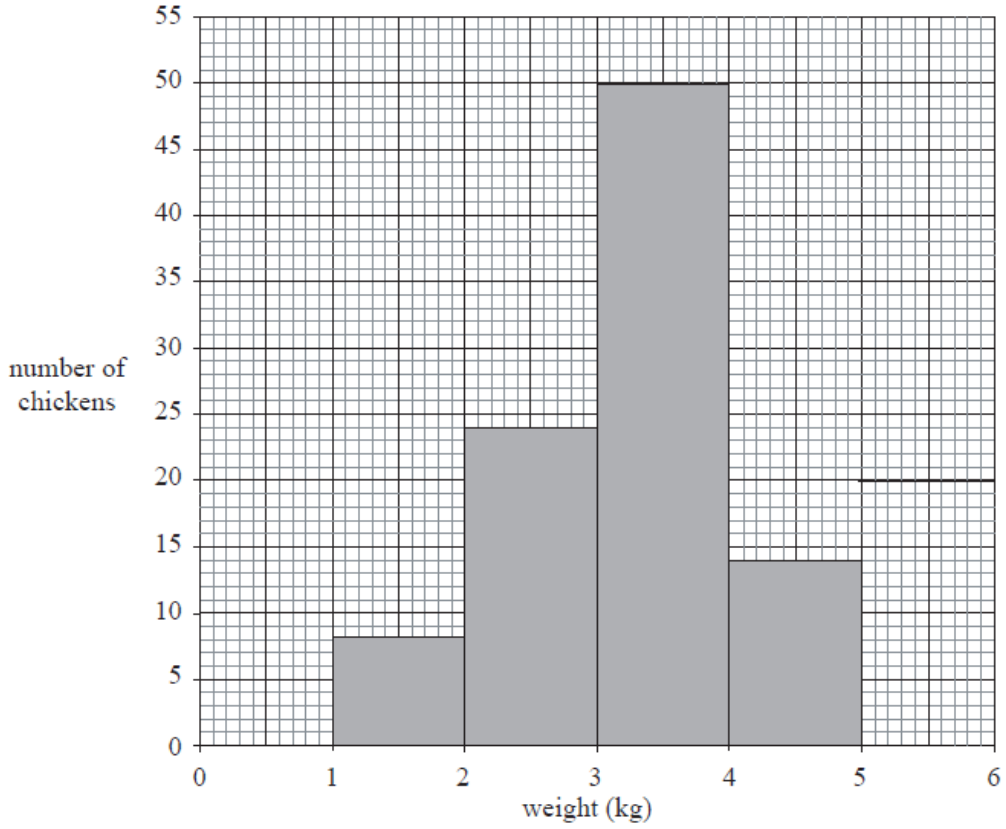


17b. Write down the modal group. *[1 mark]*

17c. The maximum word count is 4000 words. *[2 marks]*

Write down the probability that a student chosen at random is on or over the word count.

The following histogram shows the weights of a number of frozen chickens in a supermarket. The weights are grouped such that
$1 \leqslant \text{weight} < 2$,
$2 \leqslant \text{weight} < 3$ and so on.



18a. Find the total number of chickens. *[1 mark]*

18b. Write down the modal group. *[1 mark]*

18c. Gabriel chooses a chicken at random. *[2 marks]*

Find the probability that this chicken weighs less than 4 kg.

A survey was conducted of the number of bedrooms in 208 randomly chosen houses. The results are shown in the following table.

Number of bedrooms	1	2	3	4	5	6
Number of houses	41	60	52	32	15	8

19a. State whether the data is discrete or continuous. *[1 mark]*

19b. Write down the mean number of bedrooms per house. *[2 marks]*

19c. Write down the standard deviation of the number of bedrooms per house. *[1 mark]*

19d. Find how many houses have a number of bedrooms greater than one standard deviation above the mean. *[2 marks]*