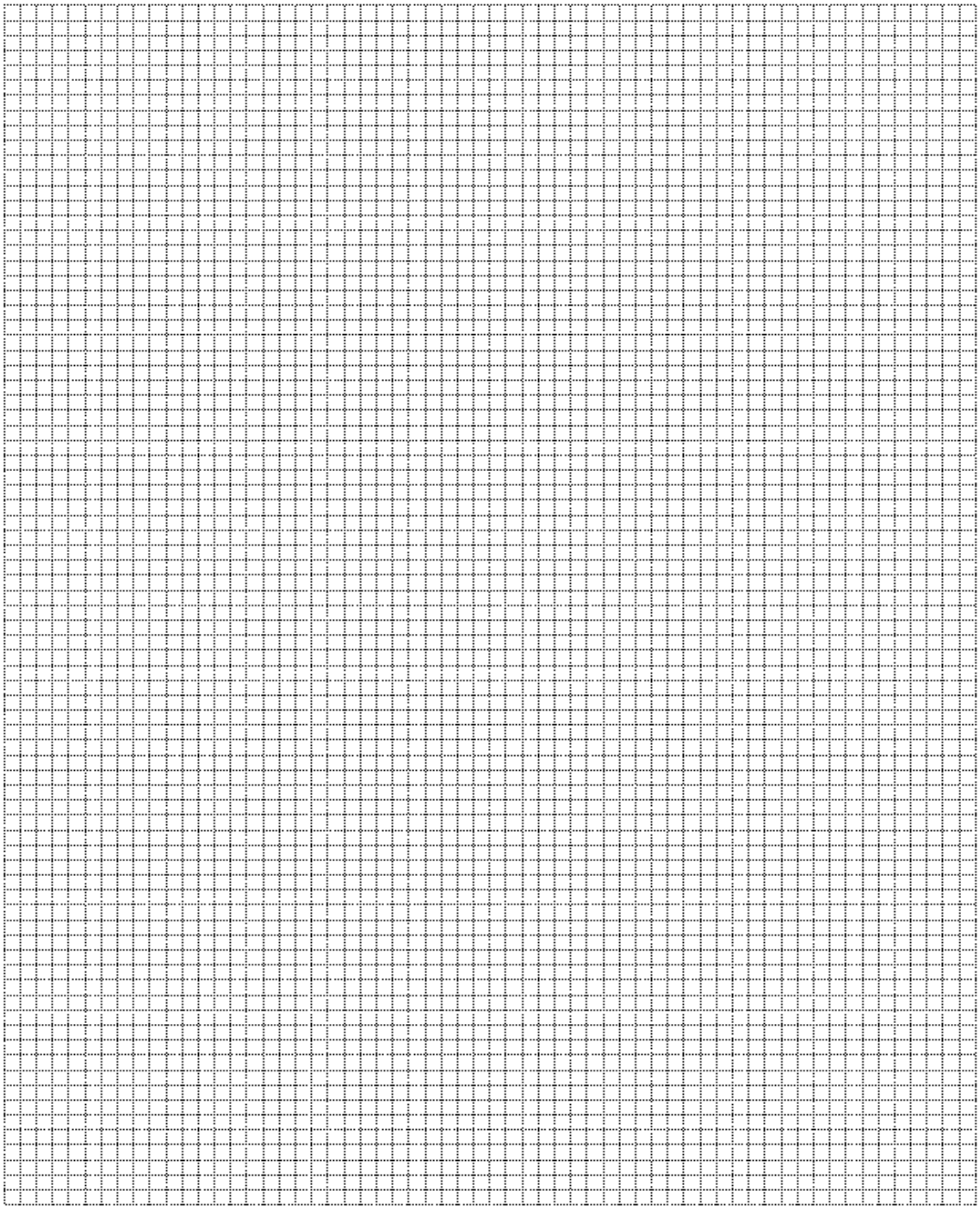
The speeds, in km h^{-1}, of 90 cars as they passed a certain marker on a road were recorded, correct the nearest km h^{-1}. The results are summarised in the following table.

Speed (km h^{-1})	10 – 29	30 – 39	40 – 49	50 – 59	60 – 89
Frequency	10	24	30	14	12

(i) On the grid, draw a histogram to illustrate the data in the table. [



(ii) Calculate an estimate for the mean speed of these 90 cars as they pass the marker. [2]

M/J/2008/Q5

As part of a data collection exercise, members of a certain school year group were asked how long they spent on their Mathematics homework during one particular week. The times are given to the nearest 0.1 hour. The results are displayed in the following table.

Time spent (t hours)	$0.1 \leqslant t \leqslant 0.5$	$0.6 \leqslant t \leqslant 1.0$	$1.1 \leqslant t \leqslant 2.0$	$2.1 \leqslant t \leqslant 3.0$	$3.1 \leqslant t \leqslant 4.5$
Frequency	11	15	18	30	21

(i) Draw, on graph paper, a histogram to illustrate this information. [5]

(ii) Calculate an estimate of the mean time spent on their Mathematics homework by members of this year group. [3]

O/N/2012/Q3

The table summarises the times that 112 people took to travel to work on a particular day.

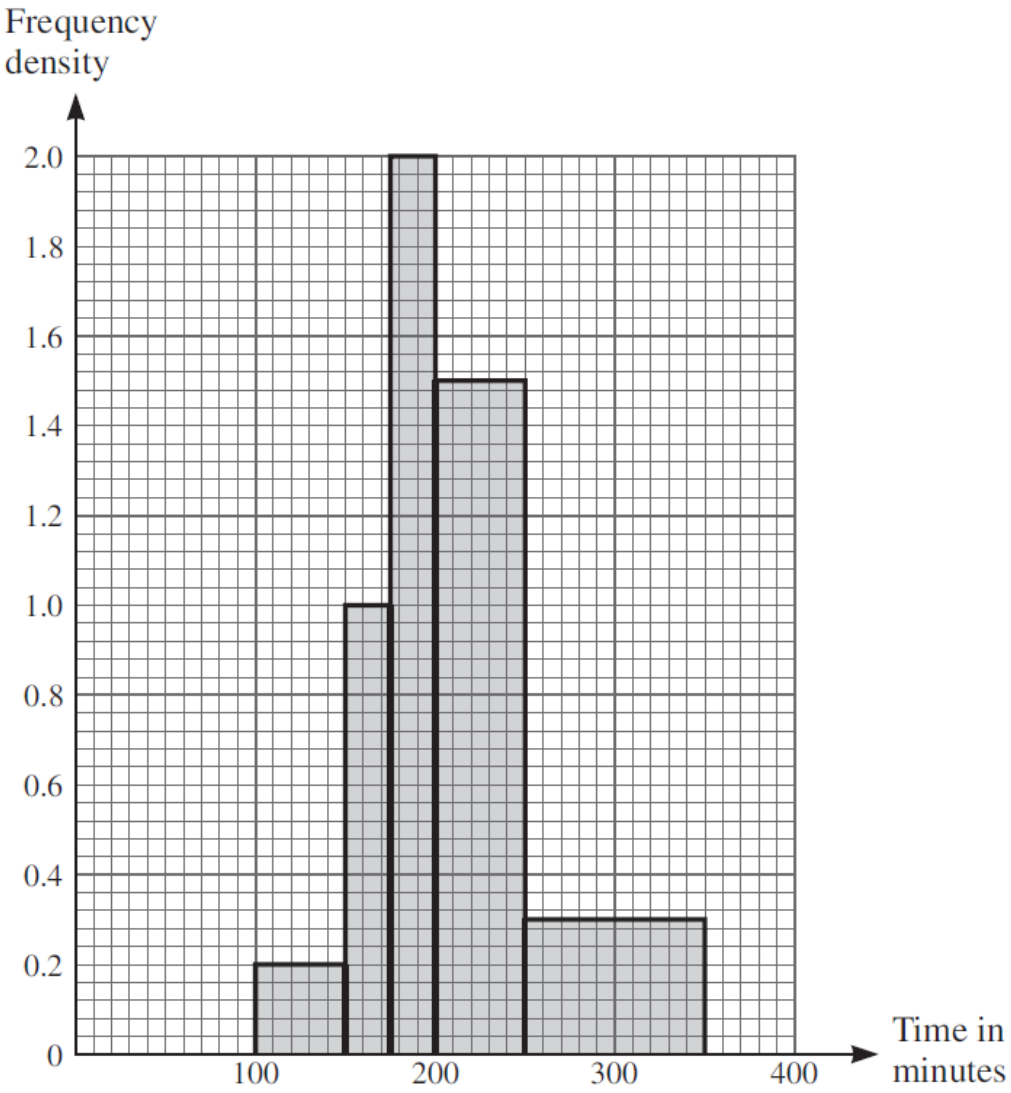
Time to travel to work (t minutes)	$0 < t \leqslant 10$	$10 < t \leqslant 15$	$15 < t \leqslant 20$	$20 < t \leqslant 25$	$25 < t \leqslant 40$	$40 < t \leqslant 60$
Frequency	19	12	28	22	18	13

(i) State which time interval in the table contains the median and which time interval contains the upper quartile. [2]

(ii) On graph paper, draw a histogram to represent the data. [4]

(iii) Calculate an estimate of the mean time to travel to work. [2]

The following histogram summarises the times, in minutes, taken by 190 people to complete a race.



(i) Show that 75 people took between 200 and 250 minutes to complete the race. [1]

(ii) Calculate estimates of the mean and standard deviation of the times of the 190 people. [6]

(iii) Explain why your answers to part **(ii)** are estimates. [1]

O/N/2011/Q4

The weights of 220 sausages are summarised in the following table.

Weight (grams)	<20	<30	<40	<45	<50	<60	<70
Cumulative frequency	0	20	50	100	160	210	220

(i) State which interval the median weight lies in. [1]

(ii) Find the smallest possible value and the largest possible value for the interquartile range. [2]

(iii) State how many sausages weighed between 50 g and 60 g. [1]

(iv) On graph paper, draw a histogram to represent the weights of the sausages. [4]

M/J/2014/Q6

The times taken by 57 athletes to run 100 metres are summarised in the following cumulative frequency table.

Time (seconds)	<10.0	<10.5	<11.0	<12.0	<12.5	<13.5
Cumulative frequency	0	4	10	40	49	57

(i) State how many athletes ran 100 metres in a time between 10.5 and 11.0 seconds. [1]

(ii) Draw a histogram on graph paper to represent the times taken by these athletes to run 100 metres. [4]

(iii) Calculate estimates of the mean and variance of the times taken by these athletes. [4]

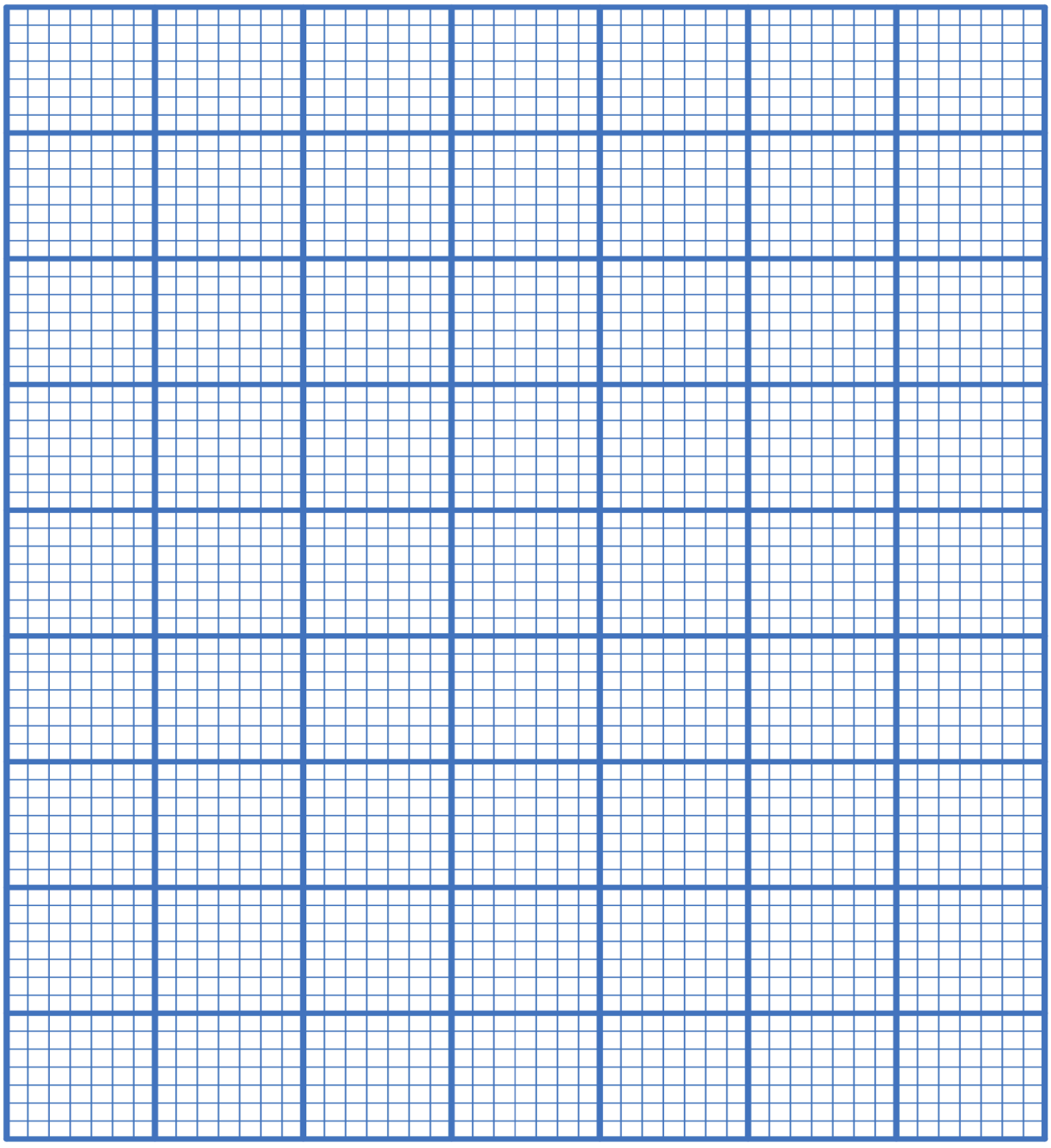
O/N/2007/Q5

The arrival times of 204 trains were noted and the number of minutes, t, that each train was late was recorded. The results are summarised in the table.

Number of minutes late (t)	$-2 \leqslant t < 0$	$0 \leqslant t < 2$	$2 \leqslant t < 4$	$4 \leqslant t < 6$	$6 \leqslant t < 10$
Number of trains	43	51	69	22	19

(i) Explain what $-2 \leqslant t < 0$ means about the arrival times of trains. [1]

(ii) Draw a cumulative frequency graph, and from it estimate the median and the interquartile range of the number of minutes late of these trains. [7]



M/J/2009/Q6

During January the numbers of people entering a store during the first hour after opening were as follows.

Time after opening, x minutes	Frequency	Cumulative frequency
$0 < x \leqslant 10$	210	210
$10 < x \leqslant 20$	134	344
$20 < x \leqslant 30$	78	422
$30 < x \leqslant 40$	72	a
$40 < x \leqslant 60$	b	540

(i) Find the values of a and b. [2]

(ii) Draw a cumulative frequency graph to represent this information. Take a scale of 2 cm for 10 minutes on the horizontal axis and 2 cm for 50 people on the vertical axis. [4]

(iii) Use your graph to estimate the median time after opening that people entered the store. [2]

(iv) Calculate estimates of the mean, m minutes, and standard deviation, s minutes, of the time after opening that people entered the store. [4]

(v) Use your graph to estimate the number of people entering the store between $(m - \frac{1}{2}s)$ and $(m + \frac{1}{2}s)$ minutes after opening. [2]

M/J/2011/Q5

A hotel has 90 rooms. The table summarises information about the number of rooms occupied each day for a period of 200 days.

Number of rooms occupied	1 – 20	21 – 40	41 – 50	51 – 60	61 – 70	71 – 90
Frequency	10	32	62	50	28	18

(i) Draw a cumulative frequency graph on graph paper to illustrate this information. [4]

(ii) Estimate the number of days when over 30 rooms were occupied. [2]

(iii) On 75% of the days at most n rooms were occupied. Estimate the value of n. [2]

O/N/2014/Q6

On a certain day in spring, the heights of 200 daffodils are measured, correct to the nearest centimetre. The frequency distribution is given below.

Height (cm)	4 – 10	11 – 15	16 – 20	21 – 25	26 – 30
Frequency	22	32	78	40	28

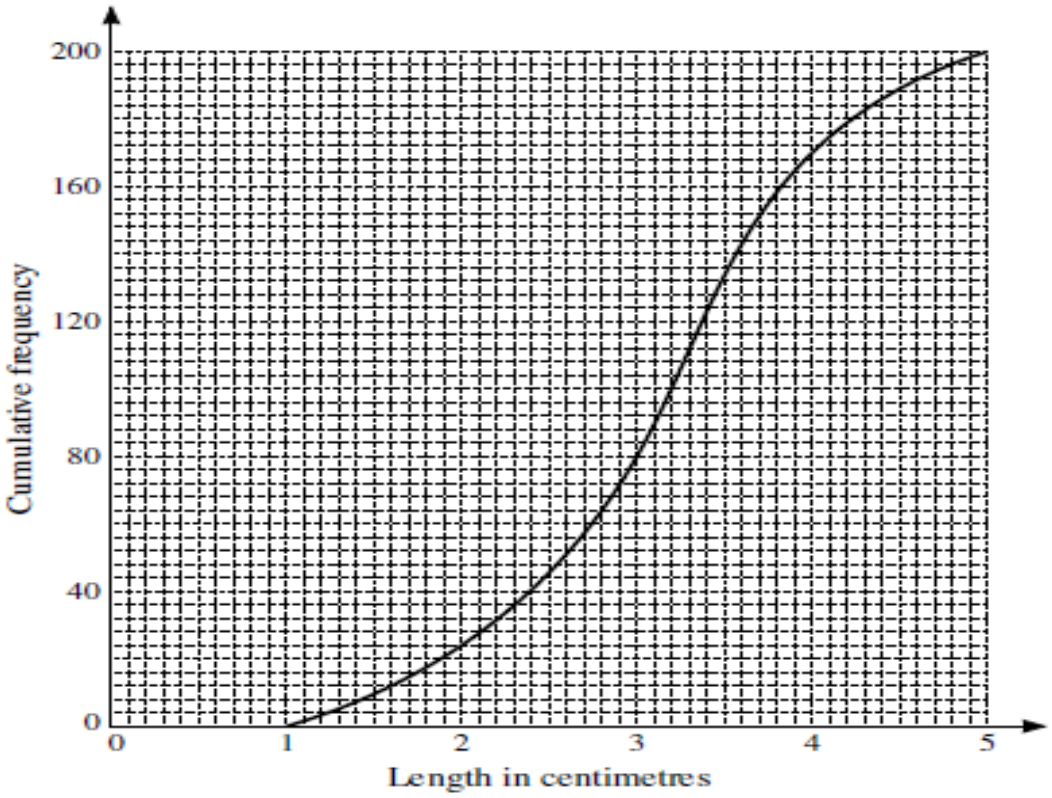
(i) Draw a cumulative frequency graph to illustrate the data. [4]

(ii) 28% of these daffodils are of height h cm or more. Estimate h. [2]

(iii) You are given that the estimate of the mean height of these daffodils, calculated from the table, is 18.39 cm. Calculate an estimate of the standard deviation of the heights of these daffodils. [3]

M/J/2017/Q2

Anabel measured the lengths, in centimetres, of 200 caterpillars. Her results are illustrated in the cumulative frequency graph below.



(i) Estimate the median and the interquartile range of the lengths. [3]

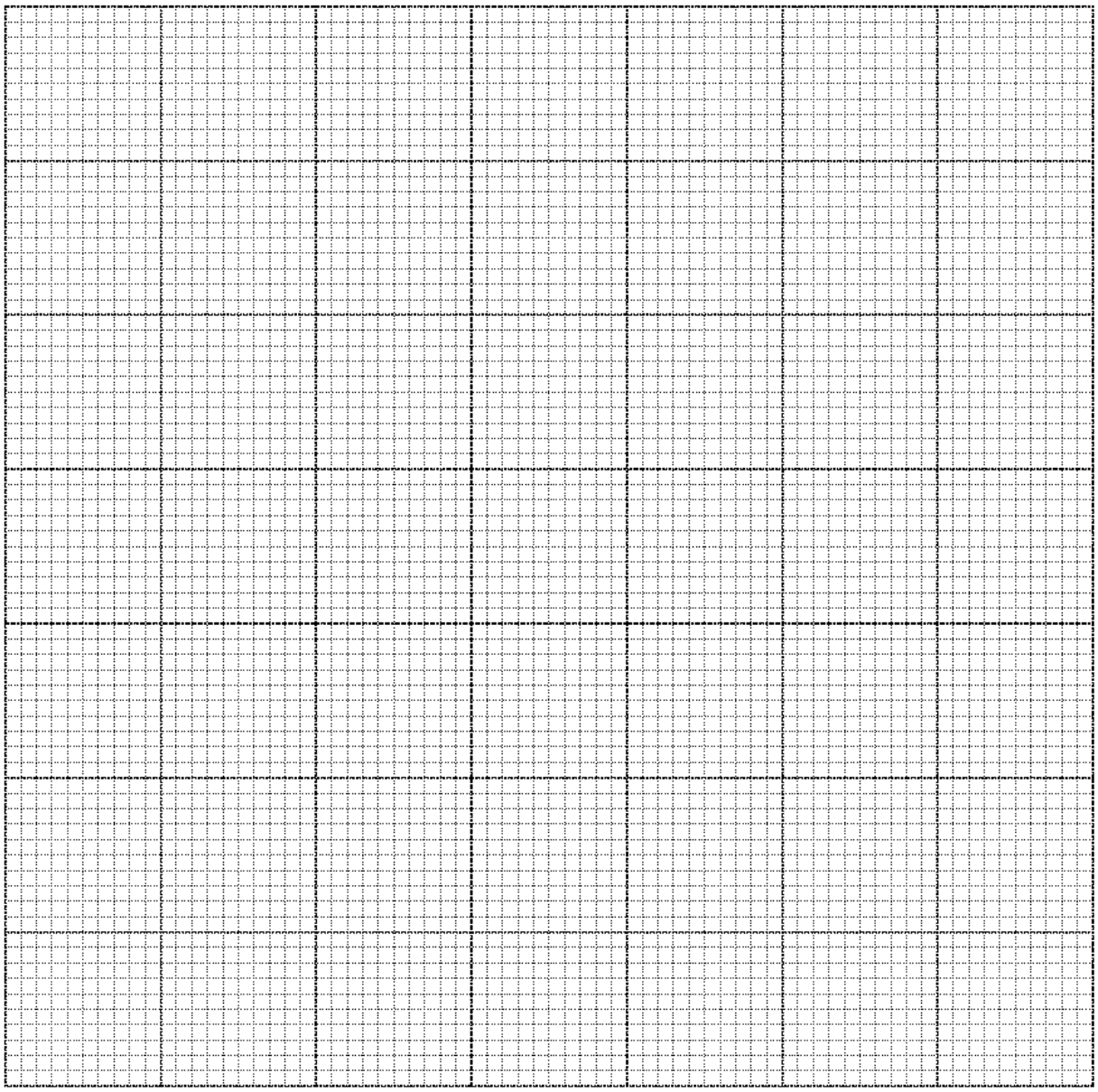
(ii) Estimate how many caterpillars had a length of between 2 and 3.5 cm. [1]

(iii) 6% of caterpillars were of length l centimetres or more. Estimate l. [2]

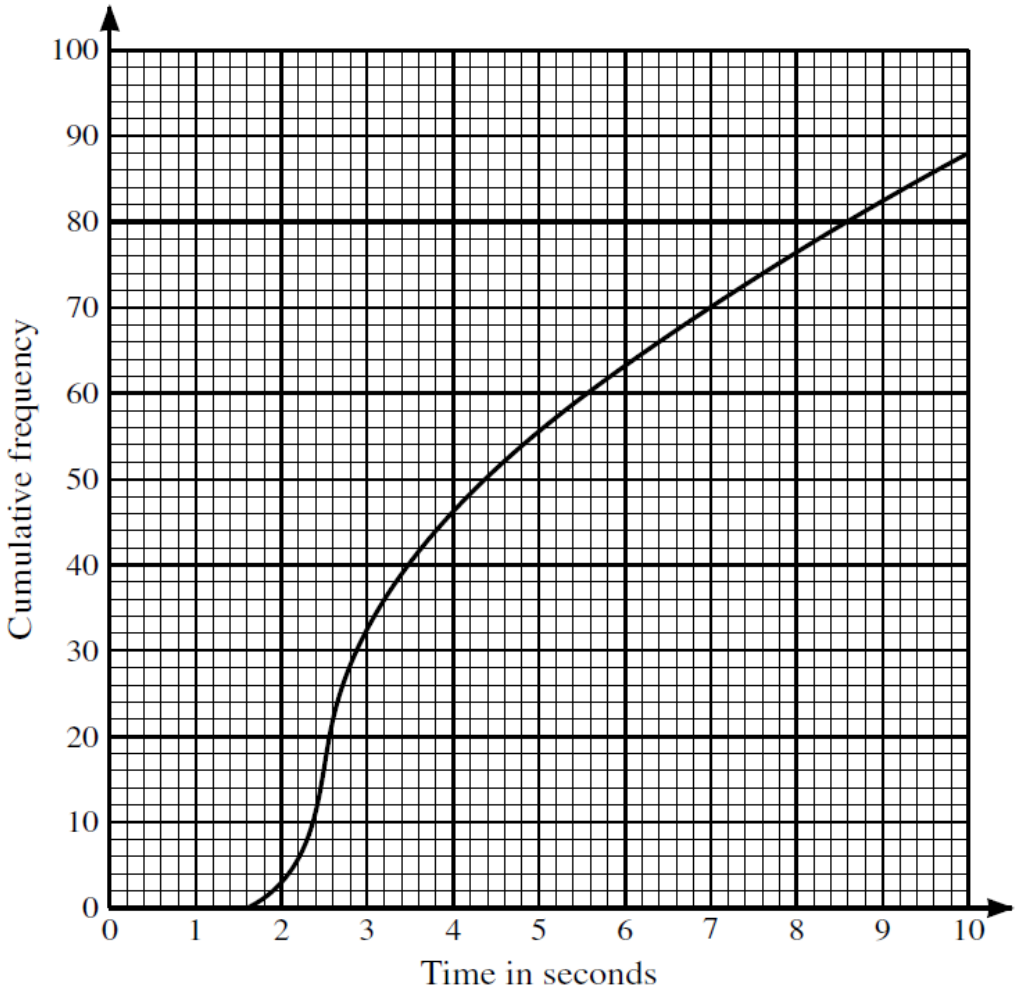
The circumferences, c cm, of some trees in a wood were measured. The results are summarised in the table.

Circumference (c cm)	$40 < c \leqslant 50$	$50 < c \leqslant 80$	$80 < c \leqslant 100$	$100 < c \leqslant 120$
Frequency	14	48	70	8

(i) On the grid, draw a cumulative frequency graph to represent the information. [3]



(ii) Estimate the percentage of trees which have a circumference larger than 75 cm. [2]



In an open-plan office there are 88 computers. The times taken by these 88 computers to access a particular web page are represented in the cumulative frequency diagram.

(i) On graph paper draw a box-and-whisker plot to summarise this information. [4]

An 'outlier' is defined as any data value which is more than 1.5 times the interquartile range above the upper quartile, or more than 1.5 times the interquartile range below the lower quartile.

(ii) Show that there are no outliers. [2]

The following back-to-back stem-and-leaf diagram shows the cholesterol count for a group of 45 people who exercise daily and for another group of 63 who do not exercise. The figures in brackets show the number of people corresponding to each set of leaves.

	People who exercise		People who do not exercise	
(9)	9 8 7 6 4 3 2 2 1	3	1 5 7 7	(4)
(12)	9 8 8 8 7 6 6 5 3 3 2 2	4	2 3 4 4 5 8	(6)
(9)	8 7 7 7 6 5 3 3 1	5	1 2 2 2 3 4 4 5 6 7 8 8 9	(13)
(7)	6 6 6 6 4 3 2	6	1 2 3 3 3 4 5 5 5 7 7 8 9 9	(14)
(3)	8 4 1	7	2 4 5 5 6 6 7 8 8	(9)
(4)	9 5 5 2	8	1 3 3 4 6 7 9 9 9	(9)
(1)	4	9	1 4 5 5 8	(5)
(0)		10	3 3 6	(3)

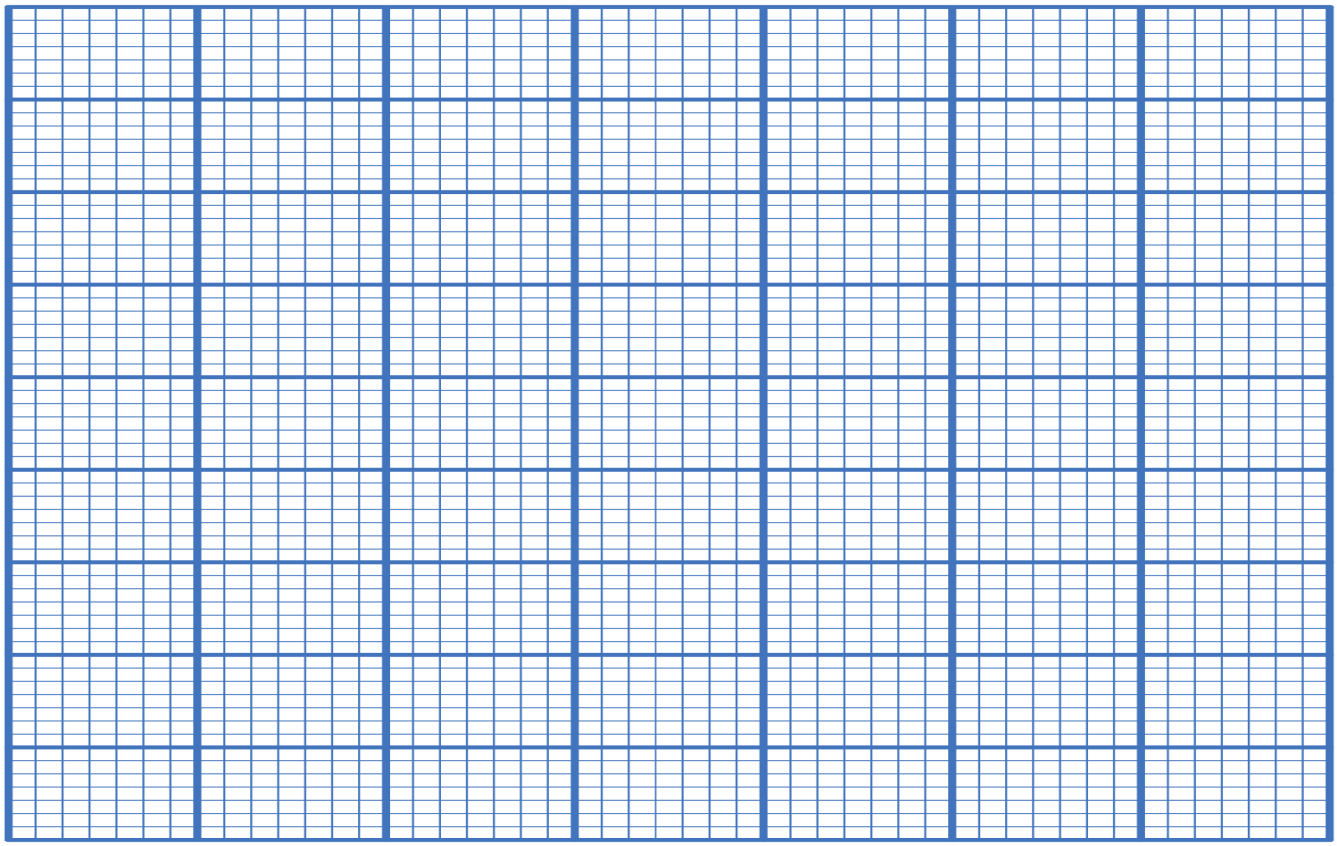
Key: 2 | 8 | 1 represents a cholesterol count of 8.2 in the group who exercise and 8.1 in the group who do not exercise.

(i) Give one useful feature of a stem-and-leaf diagram. [1]

(ii) Find the median and the quartiles of the cholesterol count for the group who do not exercise. [3]

You are given that the lower quartile, median and upper quartile of the cholesterol count for the group who exercise are 4.25, 5.3 and 6.6 respectively.

(iii) On a single diagram on graph paper, draw two box-and-whisker plots to illustrate the data. [4]



O/N/2008/Q5

The pulse rates, in beats per minute, of a random sample of 15 small animals are shown in the following table.

115	120	158	132	125
104	142	160	145	104
162	117	109	124	134

(i) Draw a stem-and-leaf diagram to represent the data. [3]

(ii) Find the median and the quartiles. [2]

(iii) On graph paper, using a scale of 2 cm to represent 10 beats per minute, draw a box-and-whisker plot of the data. [3]

O/N/2009/Q4

A library has many identical shelves. All the shelves are full and the numbers of books on each shelf in a certain section are summarised by the following stem-and-leaf diagram.

3	3 6 9 9	(4)
4	6 7	(2)
5	0 1 2 2	(4)
6	0 0 1 1 2 3 4 4 4 4 4 5 5 6 6 6 7 8 8 9	(20)
7	1 1 3 3 3 5 6 6 7 8 9 9	(12)
8	0 2 4 5 5 6 8	(7)
9	0 0 1 2 4 4 4 4 5 5 6 7 7 8 8 9 9 9	(18)

Key: 3 | 6 represents 36 books

(i) Find the number of shelves in this section of the library. [1]

(ii) Draw a box-and-whisker plot to represent the data. [5]

In another section all the shelves are full and the numbers of books on each shelf are summarised by the following stem-and-leaf diagram.

2	1 2 2 2 3 3 4 5 6 6 6 7 9	(13)
3	0 1 1 2 3 4 4 5 6 6 7 7 7 8 8	(15)
4	2 2 3 5 7 7 8 9	(8)

Key: 3 | 6 represents 36 books

(iii) There are fewer books in this section than in the previous section. State one other difference between the books in this section and the books in the previous section. [1]

M/J/2012/Q4

The back-to-back stem-and-leaf diagram shows the values taken by two variables A and B.

	A		B	
(3)	3 1 0	15	1 3 3 5	(4)
(2)	4 1	16	2 2 3 4 4 5 7 7 7 8	(10)
(3)	8 3 3	17	0 1 3 3 3 4 6 6 7 9 9	(11)
(12)	9 8 8 6 5 5 4 3 2 1 1 0	18	2 4 7	(3)
(8)	9 9 8 8 6 5 4 2	19	1 5	(2)
(5)	9 8 7 1 0	20	4	(1)

Key: 4 | 16 | 7 means $A = 0.164$ and $B = 0.167$.

(i) Find the median and the interquartile range for variable A. [3]

(ii) You are given that, for variable B, the median is 0.171, the upper quartile is 0.179 and the lower quartile is 0.164. Draw box-and-whisker plots for A and B in a single diagram on graph paper. [3]

M/J/2013/Q5

The following are the annual amounts of money spent on clothes, to the nearest \$10, by 27 people.

10	40	60	80	100	130	140	140	140
150	150	150	160	160	160	160	170	180
180	200	210	250	270	280	310	450	570

(i) Construct a stem-and-leaf diagram for the data. [3]

(ii) Find the median and the interquartile range of the data. [3]

An 'outlier' is defined as any data value which is more than 1.5 times the interquartile range above the upper quartile, or more than 1.5 times the interquartile range below the lower quartile.

(iii) List the outliers. [3]

O/N/2018/Q2

The following back-to-back stem-and-leaf diagram shows the reaction times in seconds in an experiment involving two groups of people, A and B.

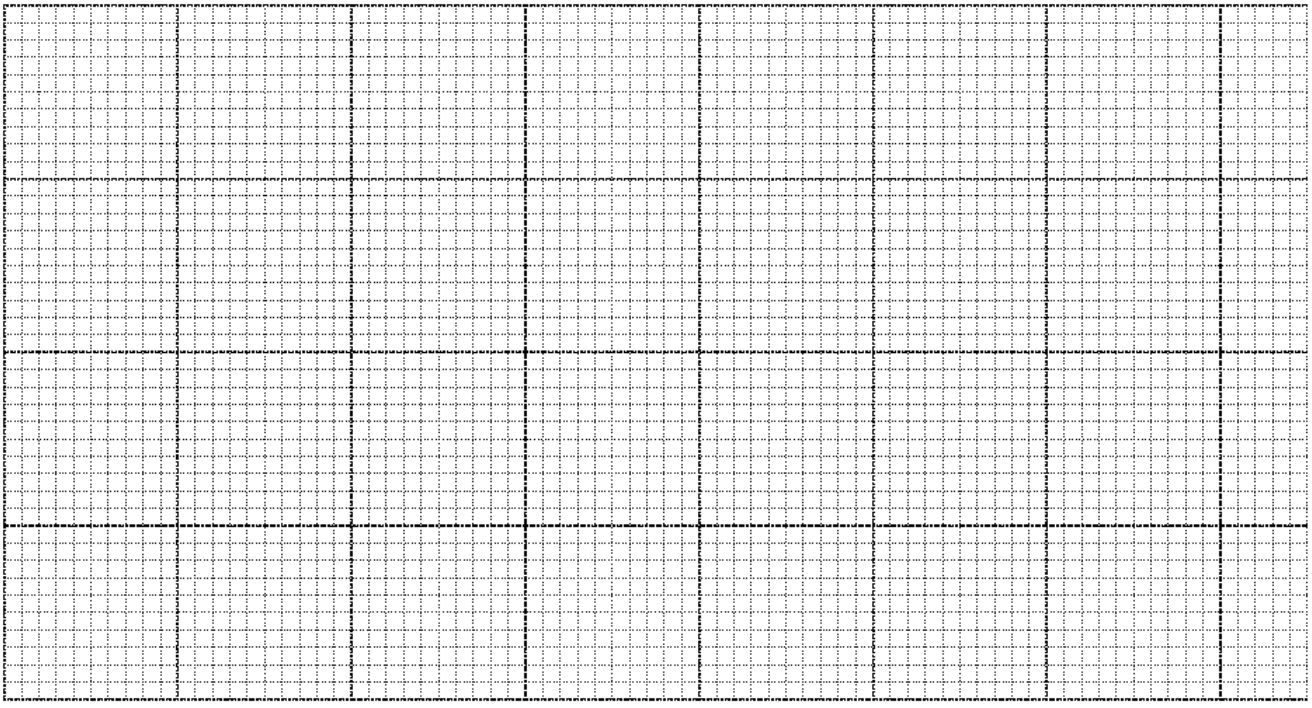
A			B	
(4)	4 2 0 0	20	5 6 7	(3)
(5)	9 8 5 0 0	21	1 2 2 3 7 7	(6)
(8)	9 8 7 5 3 2 2 2	22	1 3 5 6 6 8 9	(7)
(6)	8 7 6 5 2 1	23	4 5 7 8 8 9 9 9	(8)
(3)	8 6 3	24	2 4 5 6 7 8 8	(7)
(1)	0	25	0 2 7 8	(4)

Key: 5 \| 22 \| 6 means a reaction time of 0.225 seconds for A and 0.226 seconds for B

(i) Find the median and the interquartile range for group A. [3]

The median value for group B is 0.235 seconds, the lower quartile is 0.217 seconds and the upper quartile is 0.245 seconds.

(ii) Draw box-and-whisker plots for groups A and B on the grid. [3]



M/J/2019/Q6

(i) Give one advantage and one disadvantage of using a box-and-whisker plot to represent a set of data. [2]

(ii) The times in minutes taken to run a marathon were recorded for a group of 13 marathon runners and were found to be as follows.

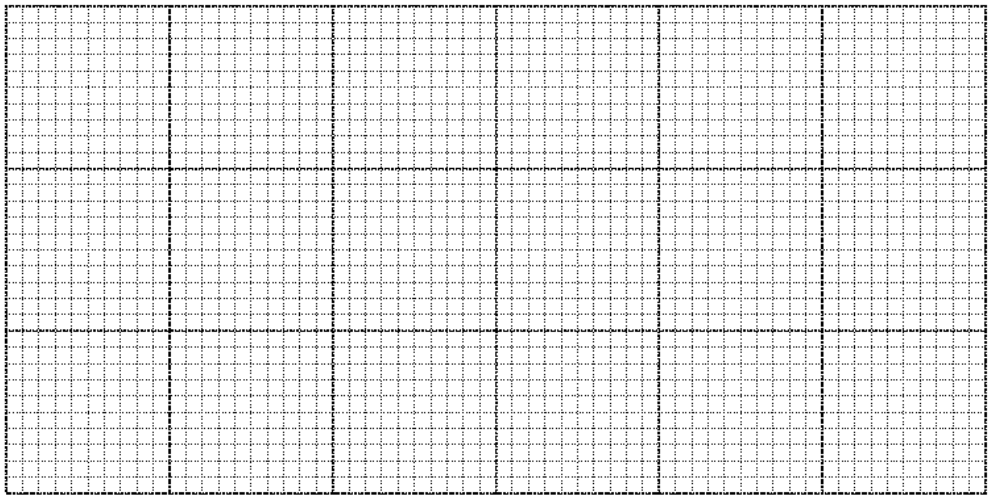
180 275 235 242 311 194 246 229 238 768 332 227 228

State which of the mean, mode or median is most suitable as a measure of central tendency for these times. Explain why the other measures are less suitable. [3]

(iii) Another group of 33 people ran the same marathon and their times in minutes were as follows.

190	203	215	246	249	253	255	254	258	260	261
263	267	269	274	276	280	288	283	287	294	300
307	318	327	331	336	345	351	353	360	368	375

(a) On the grid below, draw a box-and-whisker plot to illustrate the times for these 33 people. [4]



(b) Find the interquartile range of these times. [1]

O/N/2010/Q4

The weights in kilograms of 11 bags of sugar and 7 bags of flour are as follows.

Sugar:	1.961	1.983	2.008	2.014	1.968	1.994	2.011	2.017	1.977	1.984	1.989
Flour:	1.945	1.962	1.949	1.977	1.964	1.941	1.953				

(i) Represent this information on a back-to-back stem-and-leaf diagram with sugar on the left-hand side. [4]

(ii) Find the median and interquartile range of the weights of the bags of sugar. [3]

O/N/2015/Q5

The weights, in kilograms, of the 15 rugby players in each of two teams, A and B, are shown below.

Team A	97	98	104	84	100	109	115	99	122	82	116	96	84	107	91
Team B	75	79	94	101	96	77	111	108	83	84	86	115	82	113	95

(i) Represent the data by drawing a back-to-back stem-and-leaf diagram with team A on the left-hand side of the diagram and team B on the right-hand side. [4]

(ii) Find the interquartile range of the weights of the players in team A. [2]

(iii) A new player joins team B as a substitute. The mean weight of the 16 players in team B is now 93.9 kg. Find the weight of the new player. [3]