

A LEVELS STATISTICS (P6)(S1)

SHEET 1: BASICS QUESTION'S

(ii) On graph paper draw a box-and-whisker plot to represent the data. [5]

- 1 The table summarises the lengths in centimetres of 104 dragonflies.

Length (cm)	2.0 – 3.5	3.5 – 4.5	4.5 – 5.5	5.5 – 7.0	7.0 – 9.0
Frequency	8	25	28	31	12

(i) State which class contains the upper quartile. [1]

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

- 2 The table shows the mean and standard deviation of the weights of some turkeys and geese.

	Number of birds	Mean (kg)	Standard deviation (kg)
Turkeys	9	7.1	1.45
Geese	18	5.2	0.96

(i) Find the mean weight of the 27 birds. [2]

(ii) The weights of individual turkeys are denoted by x_t kg and the weights of individual geese by x_g kg. By first finding $\sum x_t^2$ and $\sum x_g^2$, find the standard deviation of the weights of all 27 birds. [5]

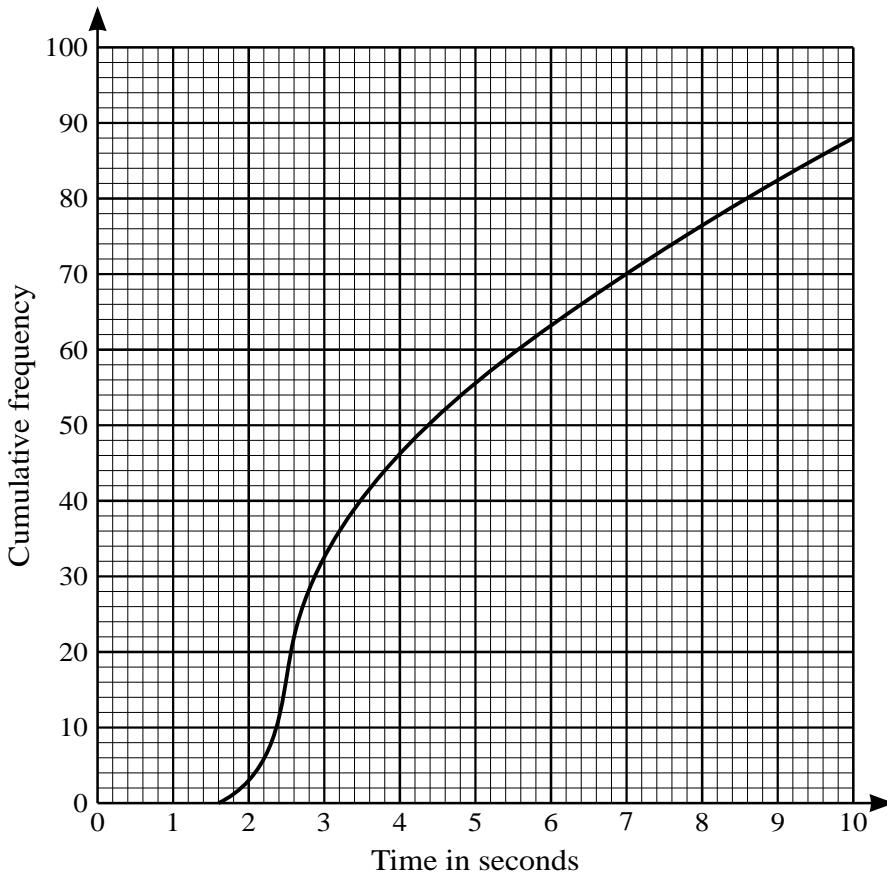
3 120 people were asked to read an article in a newspaper. The times taken, to the nearest second, by the people to read the article are summarised in the following table.

Time (seconds)	1 – 25	26 – 35	36 – 45	46 – 55	56 – 90
Number of people	4	24	38	34	20

Calculate estimates of the mean and standard deviation of the reading times.

[5]

4



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.

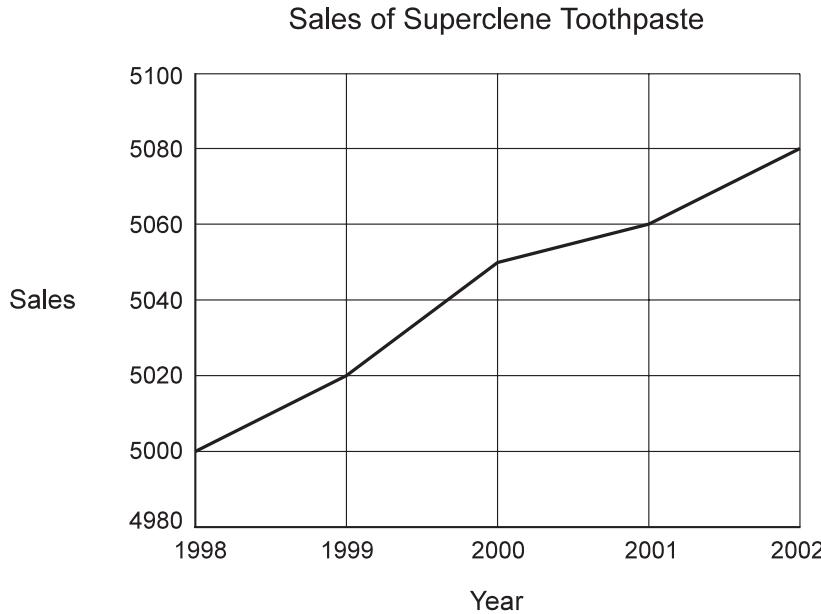
(ii) Show that there are no outliers. [2]

- 5 The weights in kilograms of two groups of 17-year-old males from country P and country Q are displayed in the following back-to-back stem-and-leaf diagram. In the third row of the diagram, $\dots 4 \mid 7 \mid 1 \dots$ denotes weights of 74 kg for a male in country P and 71 kg for a male in country Q .

Country P	Country Q
	5 1 5
	6 2 3 4 8
9 8 7 6 4	7 1 3 4 5 6 7 7 8 8 9
8 8 6 6 5 3	8 2 3 6 7 7 8 8
9 7 7 6 5 5 5 4 2	9 0 2 2 4
5 4 4 3 1	10 4 5

- (i) Find the median and quartile weights for country Q . [3]
- (ii) You are given that the lower quartile, median and upper quartile for country P are 84, 94 and 98 kg respectively. On a single diagram on graph paper, draw two box-and-whisker plots of the data. [4]
- (iii) Make two comments on the weights of the two groups. [2]

6 (i)



The diagram represents the sales of Superclene toothpaste over the last few years. Give a reason why it is misleading. [1]

- (ii) The following data represent the daily ticket sales at a small theatre during three weeks.

52, 73, 34, 85, 62, 79, 89, 50, 45, 83, 84, 91, 85, 84, 87, 44, 86, 41, 35, 73, 86.

- (a) Construct a stem-and-leaf diagram to illustrate the data. [3]
- (b) Use your diagram to find the median of the data. [1]

- 7 A random sample of 97 people who own mobile phones was used to collect data on the amount of time they spent per day on their phones. The results are displayed in the table below.

Time spent per day (t minutes)	$0 \leq t < 5$	$5 \leq t < 10$	$10 \leq t < 20$	$20 \leq t < 30$	$30 \leq t < 40$	$40 \leq t < 70$
Number of people	11	20	32	18	10	6

- (i) Calculate estimates of the mean and standard deviation of the time spent per day on these mobile phones. [5]
- (ii) On graph paper, draw a fully labelled histogram to represent the data. [4]

- 8 The floor areas, $x \text{ m}^2$, of 20 factories are as follows.

150	350	450	578	595	644	722	798	802	904
1000	1330	1533	1561	1778	1960	2167	2330	2433	3231

Represent these data by a histogram on graph paper, using intervals

$0 \leq x < 500$, $500 \leq x < 1000$, $1000 \leq x < 2000$, $2000 \leq x < 3000$, $3000 \leq x < 4000$. [4]

- 9 Two cricket teams kept records of the number of runs scored by their teams in 8 matches. The scores are shown in the following table.

Team A	150	220	77	30	298	118	160	57
Team B	166	142	170	93	111	130	148	86

- (i) Find the mean and standard deviation of the scores for team A. [2]

The mean and standard deviation for team B are 130.75 and 29.63 respectively.

- (ii) State with a reason which team has the more consistent scores. [2]

- 10 The lengths of cars travelling on a car ferry are noted. The data are summarised in the following table.

Length of car (x metres)	Frequency	Frequency density
$2.80 \leq x < 3.00$	17	85
$3.00 \leq x < 3.10$	24	240
$3.10 \leq x < 3.20$	19	190
$3.20 \leq x < 3.40$	8	a

- (i) Find the value of a . [1]
- (ii) Draw a histogram on graph paper to represent the data. [3]
- (iii) Find the probability that a randomly chosen car on the ferry is less than 3.20 m in length. [2]

11 The ages, x years, of 18 people attending an evening class are summarised by the following totals:
 $\Sigma x = 745, \Sigma x^2 = 33951$.

- (i) Calculate the mean and standard deviation of the ages of this group of people. [3]
- (ii) One person leaves the group and the mean age of the remaining 17 people is exactly 41 years.
 Find the age of the person who left and the standard deviation of the ages of the remaining 17 people. [4]
- 12 The following table shows the results of a survey to find the average daily time, in minutes, that a group of schoolchildren spent in internet chat rooms.

Time per day (t minutes)	Frequency
$0 \leq t < 10$	2
$10 \leq t < 20$	f
$20 \leq t < 40$	11
$40 \leq t < 80$	4

The mean time was calculated to be 27.5 minutes.

- (i) Form an equation involving f and hence show that the total number of children in the survey was 26. [4]
- (ii) Find the standard deviation of these times. [2]
- 13 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
(12) 9 8 8 8 7 6 6 5 3 3 2 2	4 2 3 4 4 5 8
(9) 8 7 7 7 6 5 3 3 1	5 1 2 2 2 3 4 4 5 6 7 8 9
(7) 6 6 6 6 4 3 2	6 1 2 3 3 3 4 5 5 5 7 7 8 9 9
(3) 8 4 1	7 2 4 5 5 6 6 7 8 8
(4) 9 5 5 2	8 1 3 3 4 6 7 9 9 9
(1) 4	9 1 4 5 5 8
(0)	10 3 3 6

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]

14 A group of 10 married couples and 3 single men found that the mean age \bar{x}_w of the 10 women was 41.2 years and the standard deviation of the women's ages was 15.1 years. For the 13 men, the mean age \bar{x}_m was 46.3 years and the standard deviation was 12.7 years.

- (i) Find the mean age of the whole group of 23 people. [2]
- (ii) The individual women's ages are denoted by x_w and the individual men's ages by x_m . By first finding $\sum x_w^2$ and $\sum x_m^2$, find the standard deviation for the whole group. [5]

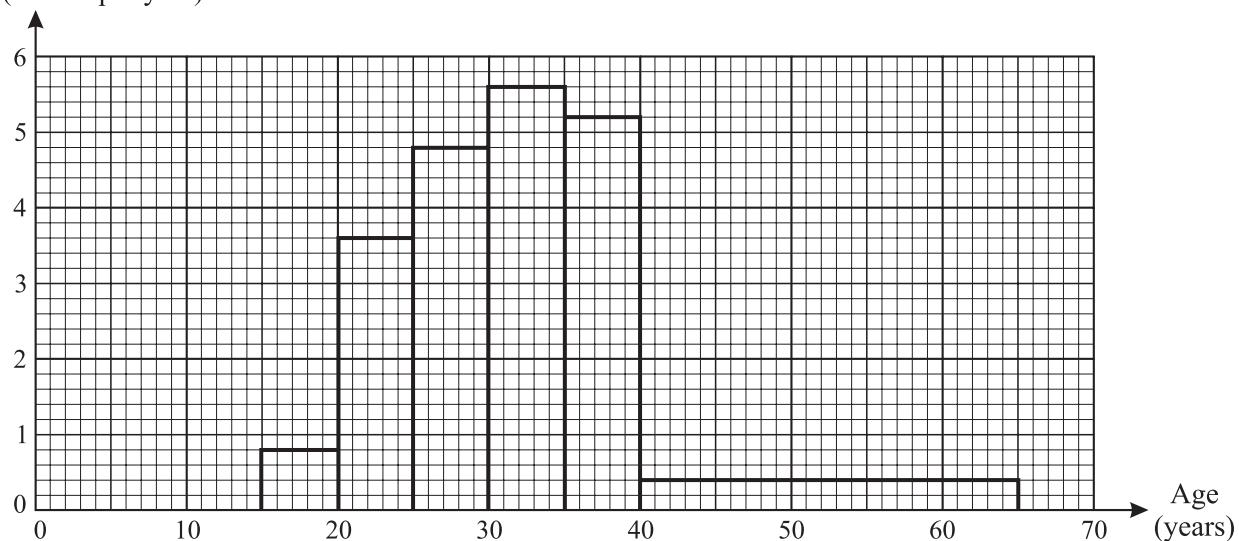
15 The salaries, in thousands of dollars, of 11 people, chosen at random in a certain office were found to be:

$$40, \quad 42, \quad 45, \quad 41, \quad 352, \quad 40, \quad 50, \quad 48, \quad 51, \quad 49, \quad 47.$$

Choose and calculate an appropriate measure of central tendency (mean, mode or median) to summarise these salaries. Explain briefly why the other measures are not suitable. [3]

16 Each father in a random sample of fathers was asked how old he was when his first child was born. The following histogram represents the information.

Frequency density
(fathers per year)



- (i) What is the modal age group? [1]
- (ii) How many fathers were between 25 and 30 years old when their first child was born? [2]
- (iii) How many fathers were in the sample? [2]
- (iv) Find the probability that a father, chosen at random from the group, was between 25 and 30 years old when his first child was born, given that he was older than 25 years. [2]

- 17** The weights of 30 children in a class, to the nearest kilogram, were as follows.

50	45	61	53	55	47	52	49	46	51
60	52	54	47	57	59	42	46	51	53
56	48	50	51	44	52	49	58	55	45

Construct a grouped frequency table for these data such that there are five equal class intervals with the first class having a lower boundary of 41.5 kg and the fifth class having an upper boundary of 61.5 kg. [4]

- 18** In a survey, people were asked how long they took to travel to and from work, on average. The median time was 3 hours 36 minutes, the upper quartile was 4 hours 42 minutes and the interquartile range was 3 hours 48 minutes. The longest time taken was 5 hours 12 minutes and the shortest time was 30 minutes.

- (i) Find the lower quartile. [2]
- (ii) Represent the information by a box-and-whisker plot, using a scale of 2 cm to represent 60 minutes. [4]

- 19** The length of time, t minutes, taken to do the crossword in a certain newspaper was observed on 12 occasions. The results are summarised below.

$$\Sigma(t - 35) = -15 \quad \Sigma(t - 35)^2 = 82.23$$

Calculate the mean and standard deviation of these times taken to do the crossword. [4]

- 20** The lengths of time in minutes to swim a certain distance by the members of a class of twelve 9-year-olds and by the members of a class of eight 16-year-olds are shown below.

9-year-olds:	13.0	16.1	16.0	14.4	15.9	15.1	14.2	13.7	16.7	16.4	15.0	13.2
16-year-olds:	14.8	13.0	11.4	11.7	16.5	13.7	12.8	12.9				

- 21** A summary of 24 observations of x gave the following information:

$$\Sigma(x - a) = -73.2 \quad \text{and} \quad \Sigma(x - a)^2 = 2115.$$

The mean of these values of x is 8.95.

- (i) Find the value of the constant a . [2]
- (ii) Find the standard deviation of these values of x . [2]

- 22** 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)	$-2 \leq t < 0$	$0 \leq t < 2$	$2 \leq t < 4$	$4 \leq t < 6$	$6 \leq t < 10$
Number of trains	43	51	69	22	19

- (i) Explain what $-2 \leq t < 0$ 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]

- 23** The stem-and-leaf diagram below represents data collected for the number of hits on an internet site on each day in March 2007. There is one missing value, denoted by x .

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

Key: 1 | 5 represents 15 hits

- (i) Find the median and lower quartile for the number of hits each day. [2]
- (ii) The interquartile range is 19. Find the value of x .

- 24** 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)	$0.1 \leq t < 0.5$	$0.6 \leq t < 1.0$	$1.1 \leq t < 2.0$	$2.1 \leq t < 3.0$	$3.1 \leq t < 4.5$
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]

- 25** Rachel measured the lengths in millimetres of some of the leaves on a tree. Her results are recorded below.

32 35 45 37 38 44 33 39 36 45

Find the mean and standard deviation of the lengths of these leaves.

[3]

- 26** 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]

- 27** 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
$0 < x \leq 10$	210	210
$10 < x \leq 20$	134	344
$20 < x \leq 30$	78	422
$30 < x \leq 40$	72	a
$40 < x \leq 60$	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]

- 28** 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 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 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]

- 29 The following table gives the marks, out of 75, in a pure mathematics examination taken by 234 students.

Marks	1–20	21–30	31–40	41–50	51–60	61–75
Frequency	40	34	56	54	29	21

- (i) Draw a histogram on graph paper to represent these results. [5]
- (ii) Calculate estimates of the mean mark and the standard deviation. [4]

- 30 The numbers of people travelling on a certain bus at different times of the day are as follows.

17	5	2	23	16	31	8
22	14	25	35	17	27	12
6	23	19	21	23	8	26

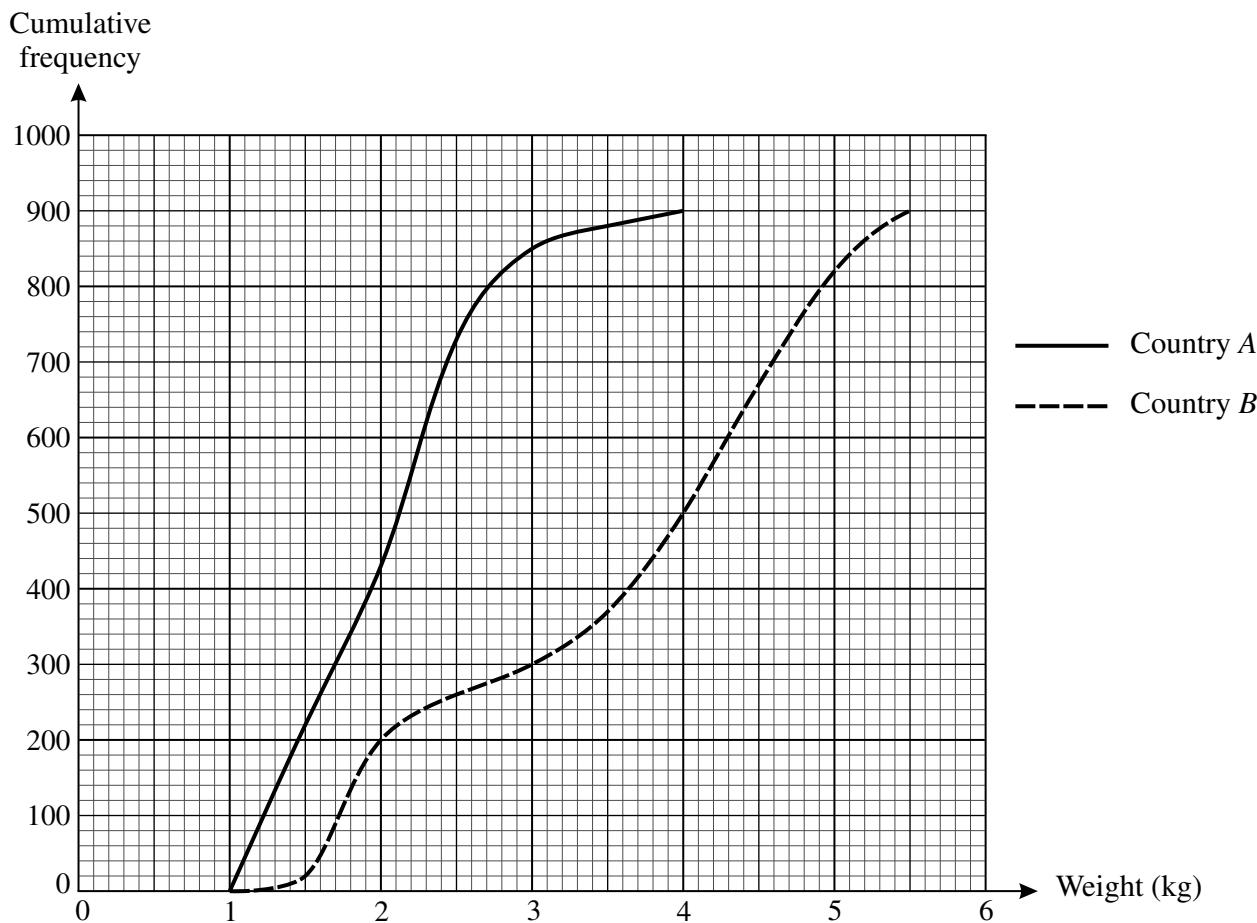
- (i) Draw a stem-and-leaf diagram to illustrate the information given above. [3]
- (ii) Find the median, the lower quartile, the upper quartile and the interquartile range. [3]
- (iii) State, in this case, which of the median and mode is preferable as a measure of central tendency, and why. [1]

- 31** The times in minutes for seven students to become proficient at a new computer game were measured. The results are shown below.

15 10 48 10 19 14 16

- (i) Find the mean and standard deviation of these times. [2]
- (ii) State which of the mean, median or mode you consider would be most appropriate to use as a measure of central tendency to represent the data in this case. [1]
- (iii) For each of the two measures of average you did not choose in part (ii), give a reason why you consider it inappropriate. [2]

32



The birth weights of random samples of 900 babies born in country A and 900 babies born in country B are illustrated in the cumulative frequency graphs. Use suitable data from these graphs to compare the central tendency and spread of the birth weights of the two sets of babies. [6]

- 33** Anita made observations of the maximum temperature, $t^{\circ}\text{C}$, on 50 days. Her results are summarised by $\Sigma t = 910$ and $\Sigma(t - \bar{t})^2 = 876$, where \bar{t} denotes the mean of the 50 observations. Calculate \bar{t} and the standard deviation of the observations. [3]

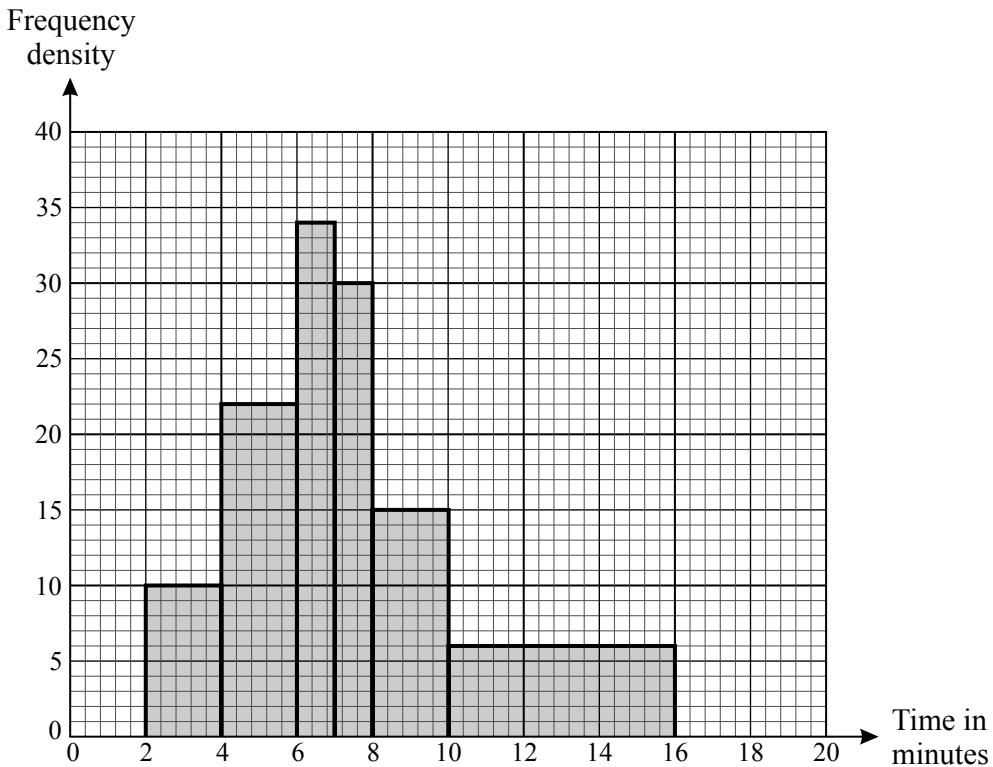
- 34** The weights in grams of a number of stones, measured correct to the nearest gram, are represented in the following table.

Weight (grams)	1 – 10	11 – 20	21 – 25	26 – 30	31 – 50	51 – 70
Frequency	$2x$	$4x$	$3x$	$5x$	$4x$	x

A histogram is drawn with a scale of 1 cm to 1 unit on the vertical axis, which represents frequency density. The 1 – 10 rectangle has height 3 cm.

- (i) Calculate the value of x and the height of the 51 – 70 rectangle. [4]
- (ii) Calculate an estimate of the mean weight of the stones. [3]
- 57** Esme noted the test marks, x , of 16 people in a class. She found that $\Sigma x = 824$ and that the standard deviation of x was 6.5.
- (i) Calculate $\Sigma(x - 50)$ and $\Sigma(x - 50)^2$. [3]
- (ii) One person did the test later and her mark was 72. Calculate the new mean and standard deviation of the marks of all 17 people. [3]
- 56** 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]
- 37** Delip measured the speeds, x km per hour, of 70 cars on a road where the speed limit is 60 km per hour. His results are summarised by $\Sigma(x - 60) = 245$.
- (i) Calculate the mean speed of these 70 cars. [2]
- His friend Sachim used values of $(x - 50)$ to calculate the mean.
- (ii) Find $\Sigma(x - 50)$. [2]
- (iii) The standard deviation of the speeds is 10.6 km per hour. Calculate $\Sigma(x - 50)^2$. [2]

- 38** The following histogram illustrates the distribution of times, in minutes, that some students spent taking a shower.



- (i) Copy and complete the following frequency table for the data. [3]

Time (t minutes)	$2 < t \leq 4$	$4 < t \leq 6$	$6 < t \leq 7$	$7 < t \leq 8$	$8 < t \leq 10$	$10 < t \leq 16$
Frequency						

- (ii) Calculate an estimate of the mean time to take a shower. [2]

- (iii) Two of these students are chosen at random. Find the probability that exactly one takes between 7 and 10 minutes to take a shower. [3]

- 39 There are 5000 schools in a certain country. The cumulative frequency table shows the number of pupils in a school and the corresponding number of schools.

Number of pupils in a school	≤ 100	≤ 150	≤ 200	≤ 250	≤ 350	≤ 450	≤ 600
Cumulative frequency	200	800	1600	2100	4100	4700	5000

- (i) Draw a cumulative frequency graph with a scale of 2 cm to 100 pupils on the horizontal axis and a scale of 2 cm to 1000 schools on the vertical axis. Use your graph to estimate the median number of pupils in a school. [3]

- (ii) 80% of the schools have more than n pupils. Estimate the value of n correct to the nearest ten. [2]

- (iii) Find how many schools have between 201 and 250 (inclusive) pupils. [1]

- (iv) Calculate an estimate of the mean number of pupils per school. [4]

- 40** A sample of 36 data values, x , gave $\Sigma(x - 45) = -148$ and $\Sigma(x - 45)^2 = 3089$.

(i) Find the mean and standard deviation of the 36 values. [3]

(ii) One extra data value of 29 was added to the sample. Find the standard deviation of all 37 values. [4]

- 41** 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]

- 42** Red Street Garage has 9 used cars for sale. Fairwheel Garage has 15 used cars for sale. The mean age of the cars in Red Street Garage is 3.6 years and the standard deviation is 1.925 years. In Fairwheel Garage, $\Sigma x = 64$ and $\Sigma x^2 = 352$, where x is the age of a car in years.

(i) Find the mean age of all 24 cars. [2]

(ii) Find the standard deviation of the ages of all 24 cars. [4]

- 43** The following cumulative frequency table shows the examination marks for 300 candidates in country A and 300 candidates in country B .

Mark	<10	<20	<35	<50	<70	<100
Cumulative frequency, A	25	68	159	234	260	300
Cumulative frequency, B	10	46	72	144	198	300

(i) Without drawing a graph, show that the median for country B is higher than the median for country A . [2]

(ii) Find the number of candidates in country A who scored between 20 and 34 marks inclusive. [1]

(iii) Calculate an estimate of the mean mark for candidates in country A . [4]

- 44** The values, x , in a particular set of data are summarised by

$$\Sigma(x - 25) = 133, \quad \Sigma(x - 25)^2 = 3762.$$

The mean, \bar{x} , is 28.325.

(i) Find the standard deviation of x . [4]

(ii) Find Σx^2 . [2]

- 45 The marks of the pupils in a certain class in a History examination are as follows.

28 33 55 38 42 39 27 48 51 37 57 49 33

The marks of the pupils in a Physics examination are summarised as follows.

Lower quartile: 28, Median: 39, Upper quartile: 67.

The lowest mark was 17 and the highest mark was 74.

- (i) Draw box-and-whisker plots in a single diagram on graph paper to illustrate the marks for History and Physics. [5]
- (ii) State one difference, which can be seen from the diagram, between the marks for History and Physics. [1]

- 46 The following are the times, in minutes, taken by 11 runners to complete a 10 km run.

48.3 55.2 59.9 67.7 60.5 75.6 62.5 57.4 53.4 49.2 64.1

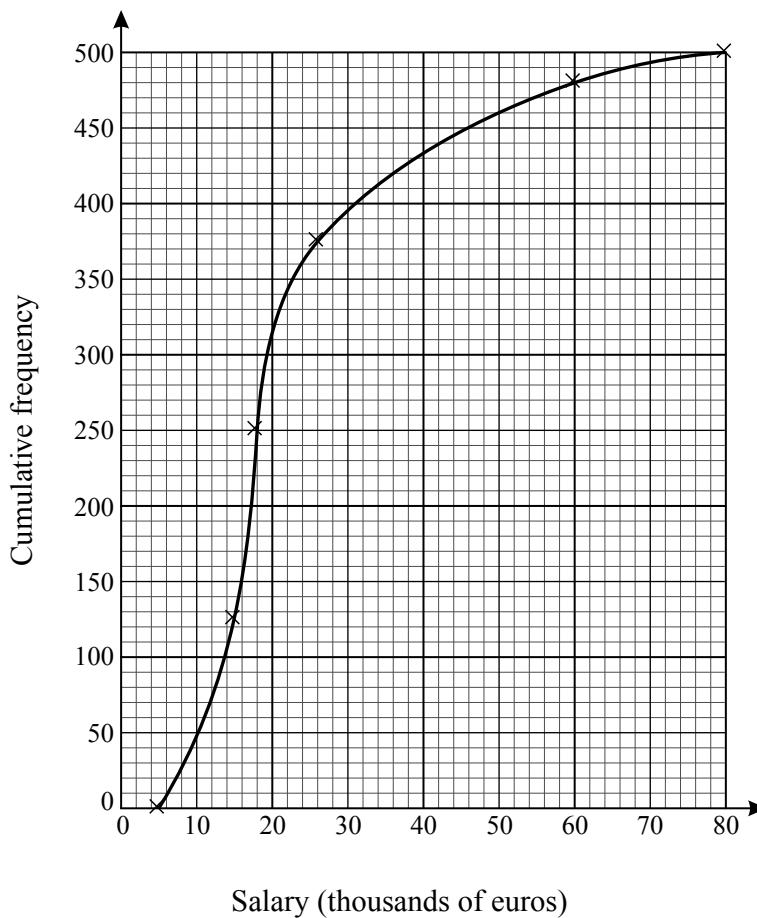
Find the mean and standard deviation of these times.

[3]

- 47 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]

48

The cumulative frequency graph shows the annual salaries, in thousands of euros, of a random sample of 500 adults with jobs, in France. It has been plotted using grouped data. You may assume that the lowest salary is 5000 euros and the highest salary is 80 000 euros.

- (i) On graph paper, draw a box-and-whisker plot to illustrate these salaries. [4]
 - (ii) Comment on the salaries of the people in this sample. [1]
 - (iii) 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.
 - (a) How high must a salary be in order to be classified as an outlier? [3]
 - (b) Show that none of the salaries is low enough to be classified as an outlier. [1]
- 49** The lengths of the diagonals in metres of the 9 most popular flat screen TVs and the 9 most popular conventional TVs are shown below.

Flat screen : 0.85 0.94 0.91 0.96 1.04 0.89 1.07 0.92 0.76

Conventional : 0.69 0.65 0.85 0.77 0.74 0.67 0.71 0.86 0.75

- (i) Represent this information on a back-to-back stem-and-leaf diagram. [4]
- (ii) Find the median and the interquartile range of the lengths of the diagonals of the 9 conventional TVs. [3]
- (iii) Find the mean and standard deviation of the lengths of the diagonals of the 9 flat screen TVs. [2]

- 50** The ages, x years, of 150 cars are summarised by $\Sigma x = 645$ and $\Sigma x^2 = 8287.5$. Find $\Sigma(x - \bar{x})^2$, where \bar{x} denotes the mean of x . [4]

- 51** 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]

- 52** The heights, x cm, of a group of young children are summarised by

$$\Sigma(x - 100) = 72, \quad \Sigma(x - 100)^2 = 499.2.$$

The mean height is 104.8 cm.

- (i)** Find the number of children in the group. [2]
- (ii)** Find $\Sigma(x - 104.8)^2$. [3]

- 53** The amounts of money, x dollars, that 24 people had in their pockets are summarised by $\Sigma(x - 36) = -60$ and $\Sigma(x - 36)^2 = 227.76$. Find Σx and Σx^2 . [5]

- 54** Prices in dollars of 11 caravans in a showroom are as follows.

16 800 18 500 17 700 14 300 15 500 15 300 16 100 16 800 17 300 15 400 16 400

- (i)** Represent these prices by a stem-and-leaf diagram. [3]
- (ii)** Write down the lower quartile of the prices of the caravans in the showroom. [1]
- (iii)** 3 different caravans in the showroom are chosen at random and their prices are noted. Find the probability that 2 of these prices are more than the median and 1 is less than the lower quartile. [3]

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

Time to travel to work (t minutes)	$0 < t \leq 10$	$10 < t \leq 15$	$15 < t \leq 20$	$20 < t \leq 25$	$25 < t \leq 40$	$40 < t \leq 60$
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]

56 In a survey, the percentage of meat in a certain type of take-away meal was found. The results, to the nearest integer, for 193 take-away meals are summarised in the table.

Percentage of meat	1 – 5	6 – 10	11 – 20	21 – 30	31 – 50
Frequency	59	67	38	18	11

- (i) Calculate estimates of the mean and standard deviation of the percentage of meat in these take-away meals. [4]
- (ii) Draw, on graph paper, a histogram to illustrate the information in the table. [5]
- 57** A summary of 30 values of x gave the following information:

$$\Sigma(x - c) = 234, \quad \Sigma(x - c)^2 = 1957.5,$$

where c is a constant.

- (i) Find the standard deviation of these values of x . [2]
- (ii) Given that the mean of these values is 86, find the value of c . [2]
- 58** The following back-to-back stem-and-leaf diagram shows the annual salaries of a group of 39 females and 39 males.

Females										Males										
(4)					5	2	0	0		20	3								(1)	
(9)	9	8	8	7	6	4	0	0	0	21	0	0	7						(3)	
(8)		8	7	5	3	3	1	0	0	22	0	0	4	5	6	6			(6)	
(6)			6	4	2	1	0	0	0	23	0	0	2	3	3	5	6	7	(9)	
(6)				7	5	4	0	0	0	24	0	1	1	2	5	5	6	8	9	(10)
(4)					9	5	0	0		25	3	4	5	7	7	8	9		(7)	
(2)						5	0			26	0	4	6						(3)	

Key: 2 | 20 | 3 means \$20 200 for females and \$20 300 for males.

- (i) Find the median and the quartiles of the females' salaries. [2]

You are given that the median salary of the males is \$24 000, the lower quartile is \$22 600 and the upper quartile is \$25 300.

- (ii) Represent the data by means of a pair of box-and-whisker plots in a single diagram on graph paper. [3]

- 59** A summary of the speeds, x kilometres per hour, of 22 cars passing a certain point gave the following information:

$$\Sigma(x - 50) = 81.4 \quad \text{and} \quad \Sigma(x - 50)^2 = 671.0.$$

Find the variance of the speeds and hence find the value of Σx^2 .

[4]

- 60** 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]

- 61** The weights, x kilograms, of 144 people were recorded. The results are summarised in the cumulative frequency table below.

Weight (x kilograms)	$x < 40$	$x < 50$	$x < 60$	$x < 65$	$x < 70$	$x < 90$
Cumulative frequency	0	12	34	64	92	144

- (i) On graph paper, draw a cumulative frequency graph to represent these results.

[2]

- (ii) 64 people weigh more than c kg. Use your graph to find the value of c .

[2]

- (iii) Calculate estimates of the mean and standard deviation of the weights.

[6]

- 62** Swati measured the lengths, x cm, of 18 stick insects and found that $\Sigma x^2 = 967$. Given that the mean length is $\frac{58}{9}$ cm, find the values of $\Sigma(x - 5)$ and $\Sigma(x - 5)^2$.

[5]

- 63** The following are the house prices in thousands of dollars, arranged in ascending order, for 51 houses from a certain area.

253	270	310	354	386	428	433	468	472	477	485	520	520	524	526	531	535
536	538	541	543	546	548	549	551	554	572	583	590	605	614	638	649	652
666	670	682	684	690	710	725	726	731	734	745	760	800	854	863	957	986

- (i) Draw a box-and-whisker plot to represent the data.

[4]

An expensive house is defined as a house which has a price that is more than 1.5 times the interquartile range above the upper quartile.

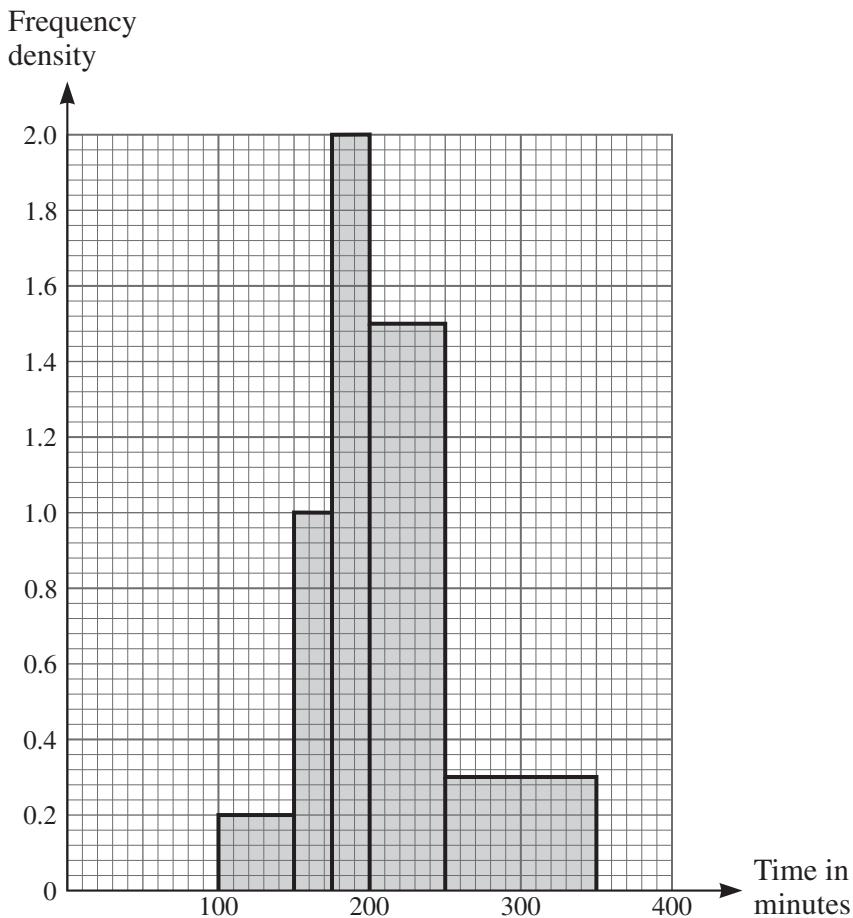
- (ii) For the above data, give the prices of the expensive houses.

[2]

- (iii) Give one disadvantage of using a box-and-whisker plot rather than a stem-and-leaf diagram to represent this set of data.

[1]

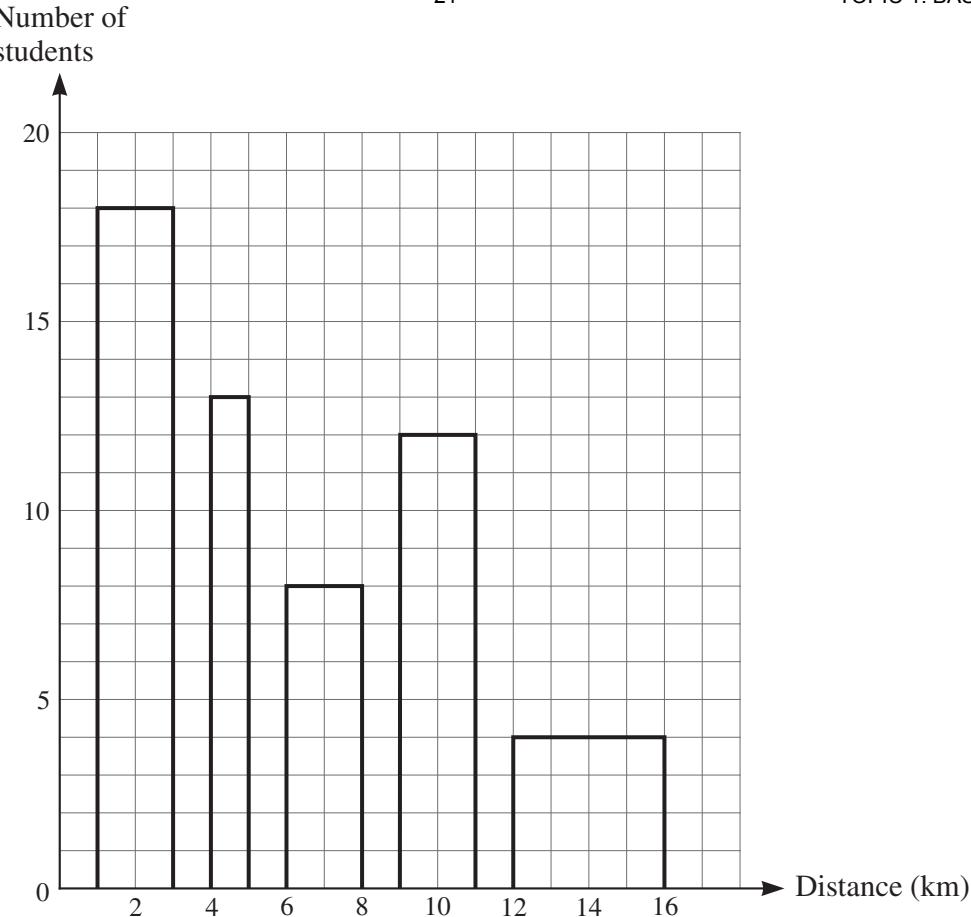
- 64** The following histogram summarises the times, in minutes, taken by 190 people to complete a race.



- (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]
- 65** The distance of a student's home from college, correct to the nearest kilometre, was recorded for each of 55 students. The distances are summarised in the following table.

Distance from college (km)	1 – 3	4 – 5	6 – 8	9 – 11	12 – 16
Number of students	18	13	8	12	4

Dominic is asked to draw a histogram to illustrate the data. Dominic's diagram is shown below.



Give two reasons why this is not a correct histogram.

[2]

- 66 A typing test is taken by 111 people. The numbers of typing errors they make in the test are summarised in the table below.

Number of typing errors	1 – 5	6 – 20	21 – 35	36 – 60	61 – 80
Frequency	24	9	21	15	42

- (i) Draw a histogram on graph paper to represent this information. [5]
- (ii) Calculate an estimate of the mean number of typing errors for these 111 people. [3]
- (iii) State which class contains the lower quartile and which class contains the upper quartile. Hence find the least possible value of the interquartile range. [3]
- 67 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]

- 68** Some adults and some children each tried to estimate, without using a watch, the number of seconds that had elapsed in a fixed time-interval. Their estimates are shown below.

Adults:	55 58 67 74 63 61 63 71 56 53 54 78 73 64 62
Children:	86 95 89 72 61 84 77 92 81 54 43 68 62 67 83

- (i) Draw a back-to-back stem-and-leaf diagram to represent the data. [3]
- (ii) Make two comparisons between the estimates of the adults and the children. [2]
- 69** The heights, x cm, of a group of 28 people were measured. The mean height was found to be 172.6 cm and the standard deviation was found to be 4.58 cm. A person whose height was 161.8 cm left the group.

- (i) Find the mean height of the remaining group of 27 people. [2]
- (ii) Find Σx^2 for the original group of 28 people. Hence find the standard deviation of the heights of the remaining group of 27 people. [4]

- 70** Find the mean and variance of the following data. [3]

5 -2 12 7 -3 2 -6 4 0 8

- 71** The following back-to-back stem-and-leaf diagram shows the times to load an application on 61 smartphones of type *A* and 43 smartphones of type *B*.

	Type A		Type B
(7)	9 7 6 6 4 3 3	2	1 3 5 8 (4)
(7)	5 5 4 4 2 2 2	3	0 4 4 5 6 6 6 6 7 8 8 9 (12)
(13)	9 9 8 8 8 7 6 6 4 3 2 2 0	4	0 1 1 2 3 6 8 8 9 9 (10)
(9)	6 5 5 4 3 2 1 1 0	5	2 5 6 6 9 (5)
(4)	9 7 3 0	6	1 3 8 9 (4)
(6)	8 7 4 4 1 0	7	5 7 (2)
(10)	7 6 6 6 5 3 3 2 1 0	8	1 2 4 4 (4)
(5)	8 6 5 5 5	9	0 6 (2)

Key: 3 | 2 | 1 means 0.23 seconds for type *A* and 0.21 seconds for type *B*.

- (i) Find the median and quartiles for smartphones of type *A*. [3]

You are given that the median, lower quartile and upper quartile for smartphones of type *B* are 0.46 seconds, 0.36 seconds and 0.63 seconds respectively.

- (ii) Represent the data by drawing a pair of box-and-whisker plots in a single diagram on graph paper. [3]

- (iii) Compare the loading times for these two types of smartphone. [1]

- 72 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]
- 73 A traffic camera measured the speeds, x kilometres per hour, of 8 cars travelling along a certain street, with the following results.

62.7 59.6 64.2 61.5 68.3 66.9 62.0 62.3

- (i) Find $\Sigma(x - 62)$. [1]
- (ii) Find $\Sigma(x - 62)^2$. [1]
- (iii) Find the mean and variance of the speeds of the 8 cars. [3]
- 74 A random sample of 25 people recorded the number of glasses of water they drank in a particular week. The results are shown below.

23	19	32	14	25
22	26	36	45	42
47	28	17	38	15
46	18	26	22	41
19	21	28	24	30

- (i) Draw a stem-and-leaf diagram to represent the data. [3]
- 75 The age at which a child first walked (to the nearest month) was recorded for 8 children. The results were as follows.

12 11 16 19 10 12 12 13

Calculate the mean and standard deviation of the data. [3]

- 76 A survey was made of the number of people attending church services on one particular Sunday morning. A random sample of 500 churches was taken. The results are as follows.

Number of people attending	1–20	21–40	41–60	61–100	101–200	201–300
Number of churches	46	110	122	100	86	36

- (i) Draw a histogram on graph paper to represent these results. [5]
- (ii) Find the probability that, in each of 3 churches chosen at random from the sample, the number of people attending was less than 61. [2]

- 77 The manager of a company noted the times spent in 80 meetings. The results were as follows.

Time (t minutes)	$0 < t \leq 15$	$15 < t \leq 30$	$30 < t \leq 60$	$60 < t \leq 90$	$90 < t \leq 120$
Number of meetings	4	7	24	38	7

Draw a cumulative frequency graph and use this to estimate the median time and the interquartile range. [6]

- 78 (i) In a spot check of the speeds $x \text{ km h}^{-1}$ of 30 cars on a motorway, the data were summarised by $\Sigma(x - 110) = -47.2$ and $\Sigma(x - 110)^2 = 5460$. Calculate the mean and standard deviation of these speeds. [4]

- 79 Seventy samples of fertiliser were collected and the nitrogen content was measured for each sample. The cumulative frequency distribution is shown in the table below.

Nitrogen content	≤ 3.5	≤ 3.8	≤ 4.0	≤ 4.2	≤ 4.5	≤ 4.8
Cumulative frequency	0	6	18	41	62	70

- (i) On graph paper draw a cumulative frequency graph to represent the data. [3]
- (ii) Estimate the percentage of samples with a nitrogen content greater than 4.4. [2]
- (iii) Estimate the median. [1]
- (iv) Construct the frequency table for these results and draw a histogram on graph paper. [5]

- 80 For n values of the variable x , it is given that $\Sigma(x - 100) = 216$ and $\Sigma x = 2416$. Find the value of n . [3]

- 81 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]
- 82 The time taken, t hours, to deliver letters on a particular route each day is measured on 250 working days. The mean time taken is 2.8 hours. Given that $\Sigma(t - 2.5)^2 = 96.1$, find the standard deviation of the times taken. [3]

- 83 The heights to the nearest metre of 134 office buildings in a certain city are summarised in the table below.

Height (m)	21 – 40	41 – 45	46 – 50	51 – 60	61 – 80
Frequency	18	15	21	52	28

- (i) Draw a histogram on graph paper to illustrate the data. [4]
- (ii) Calculate estimates of the mean and standard deviation of these heights. [5]