

## Past Year: Chapter 1 Representation of Data

May/June 2002

- 2 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

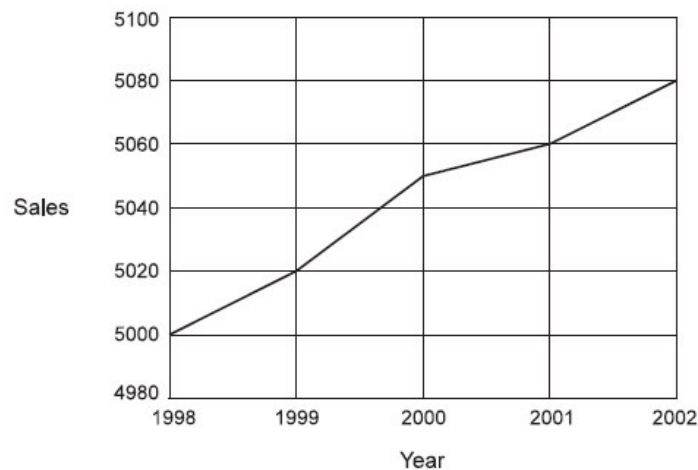
Draw a cumulative frequency graph and use this to estimate the median time and the interquartile range. [6]

- 4 (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]
- (ii) On another day the mean speed of cars on the motorway was found to be  $107.6 \text{ km h}^{-1}$  and the standard deviation was  $13.8 \text{ km h}^{-1}$ . Assuming these speeds follow a normal distribution and that the speed limit is  $110 \text{ km h}^{-1}$ , find what proportion of cars exceed the speed limit. [3]

May/June 2003

- 1 (i)

Sales of Superclene Toothpaste



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]

May/June 2004

- 1 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]
- 2 In a recent survey, 640 people were asked about the length of time each week that they spent watching television. The median time was found to be 20 hours, and the lower and upper quartiles were 15 hours and 35 hours respectively. The least amount of time that anyone spent was 3 hours, and the greatest amount was 60 hours.
- (i) On graph paper, show these results using a fully labelled cumulative frequency graph. [3]
- (ii) Use your graph to estimate how many people watched more than 50 hours of television each week. [2]

May/June 2005

- 2 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]

- 4 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]

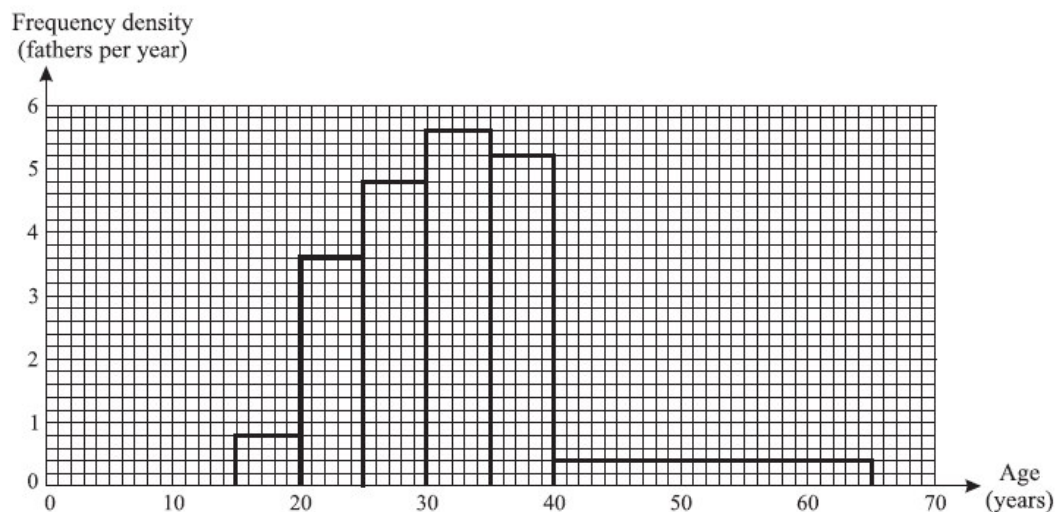
May/June 2006

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

40, 42, 45, 41, 352, 40, 50, 48, 51, 49, 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]

- 5 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.



- (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]

May/June 2007

- 1 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]

- 4 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

- (i) Draw a back-to-back stem-and-leaf diagram to represent the information above. [4]  
 (ii) A new pupil joined the 16-year-old class and swam the distance. The mean time for the class of nine pupils was now 13.6 minutes. Find the new pupil's time to swim the distance. [3]

May/June 2008

- 1 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  $x$ . [2]
- 5 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)	$0.1 \leq t \leq 0.5$	$0.6 \leq t \leq 1.0$	$1.1 \leq t \leq 2.0$	$2.1 \leq t \leq 3.0$	$3.1 \leq t \leq 4.5$
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]

- 6 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]

Oct/Nov 2001

- 1 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]

- 4 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]

- 7 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, ... 4 | 7 | 1 ... 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]

- 1 A computer can generate random numbers which are either 0 or 2. On a particular occasion, it generates a set of numbers which consists of 23 zeros and 17 twos. Find the mean and variance of this set of 40 numbers. [4]
- 2 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]

- 2 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]
- 4 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]

Oct/Nov 2005

- 4 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  $\Sigma x_w^2$  and  $\Sigma x_m^2$ , find the standard deviation for the whole group. [5]

Oct/Nov 2006

- 1 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]

- 3 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]

Oct/Nov 2007

- 1 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]
- 5 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]

Oct/Nov 2008

- 1 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]

- 5 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]

Oct/Nov 2009/11

- 4 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.

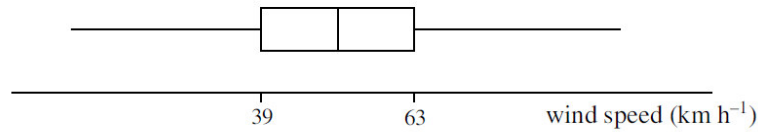
2		1 2 2 2 3 3 4 5 6 6 6 7 9	(13)
3		0 1 1 2 3 4 4 5 6 6 7 7 7 8 8	(15)
4		2 2 3 5 7 7 8 9	(8)

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]



1



Measurements of wind speed on a certain island were taken over a period of one year. A box-and-whisker plot of the data obtained is displayed above, and the values of the quartiles are as shown. It is suggested that wind speed can be modelled approximately by a normal distribution with mean  $\mu$  km h<sup>-1</sup> and standard deviation  $\sigma$  km h<sup>-1</sup>.

- (i) Estimate the value of  $\mu$ . [1]
- (ii) Estimate the value of  $\sigma$ . [3]

6 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]

May/June 2010/61

2 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]

4 The numbers of rides taken by two students, Fei and Graeme, at a fairground are shown in the following table.

	Roller coaster	Water slide	Revolving drum
Fei	4	2	0
Graeme	1	3	6

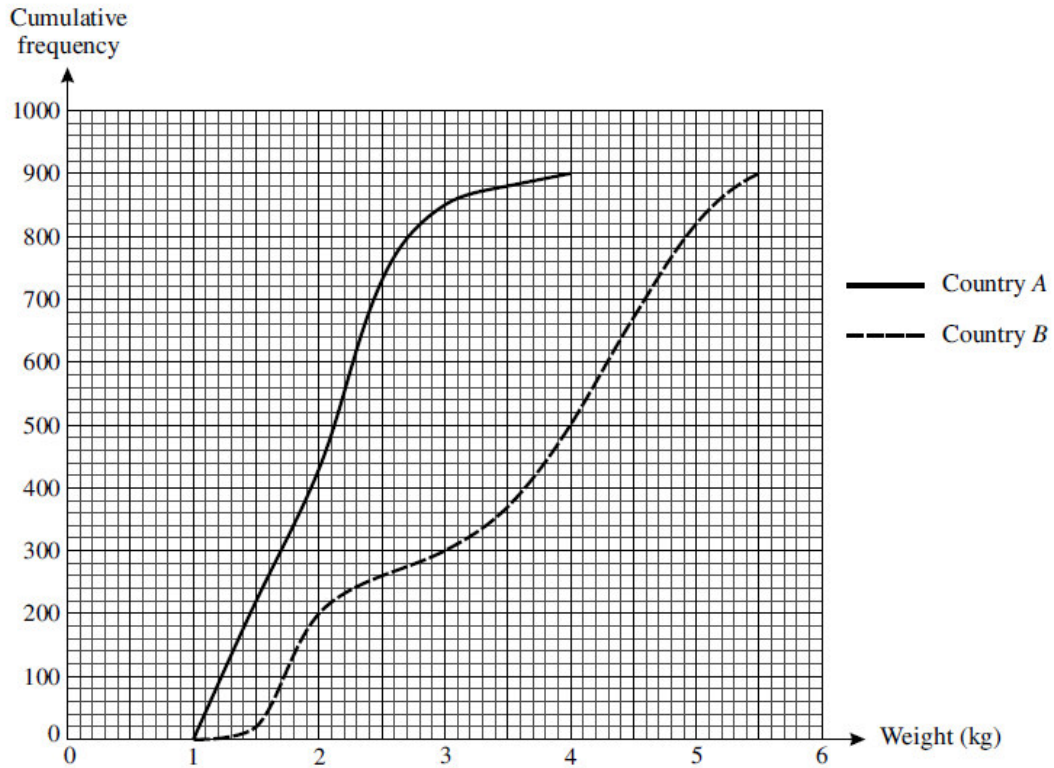
- (i) The mean cost of Fei’s rides is \$2.50 and the standard deviation of the costs of Fei’s rides is \$0. Explain how you can tell that the roller coaster and the water slide each cost \$2.50 per ride. [2]
- (ii) The mean cost of Graeme’s rides is \$3.76. Find the standard deviation of the costs of Graeme’s rides. [5]

- 1 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]

3



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]

- 3 Christa takes her dog for a walk every day. The probability that they go to the park on any day is 0.6. If they go to the park there is a probability of 0.35 that the dog will bark. If they do not go to the park there is a probability of 0.75 that the dog will bark.
- (i) Find the probability that they go to the park on more than 5 of the next 7 days. [2]
  - (ii) Find the probability that the dog barks on any particular day. [2]
  - (iii) Find the variance of the number of times they go to the park in 30 days. [1]

- 6 The lengths of some insects of the same type from two countries,  $X$  and  $Y$ , were measured. The stem-and-leaf diagram shows the results.

Country $X$			Country $Y$	
(10)	9 7 6 6 6 4 4 4 3 2	80		
(18)	8 8 8 7 7 6 6 5 5 5 4 4 3 3 3 2 2 0	81	1 1 2 2 3 3 3 5 5 6 7 8 9	(13)
(16)	9 9 9 8 8 7 7 6 5 5 3 2 2 1 0 0	82	0 0 1 2 3 3 3 $q$ 4 5 6 6 7 8 8	(15)
(16)	8 7 6 5 5 5 3 3 2 2 2 1 1 1 0 0	83	0 1 2 2 4 4 4 4 5 5 6 6 7 7 7 8 9	(17)
(11)	8 7 6 5 5 4 4 3 3 1 1	84	0 0 1 2 4 4 5 5 6 6 7 7 7 8 9	(15)
		85	1 2 $r$ 3 3 5 5 6 6 7 8 8	(12)
		86	0 1 2 2 3 5 5 5 8 9 9	(11)

Key:  $5 | 81 | 3$  means an insect from country  $X$  has length 0.815 cm  
and an insect from country  $Y$  has length 0.813 cm.

- (i) Find the median and interquartile range of the lengths of the insects from country  $X$ . [2]
- (ii) The interquartile range of the lengths of the insects from country  $Y$  is 0.028 cm. Find the values of  $q$  and  $r$ . [2]
- (iii) Represent the data by means of a pair of box-and-whisker plots in a single diagram on graph paper. [4]
- (iv) Compare the lengths of the insects from the two countries. [2]

Oct/Nov 2010/61

- 1 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]
- 4 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]

Oct/Nov 2010/62

- 2 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]
- 4 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]

4 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  $\Sigma(x - 60) = 245$ .

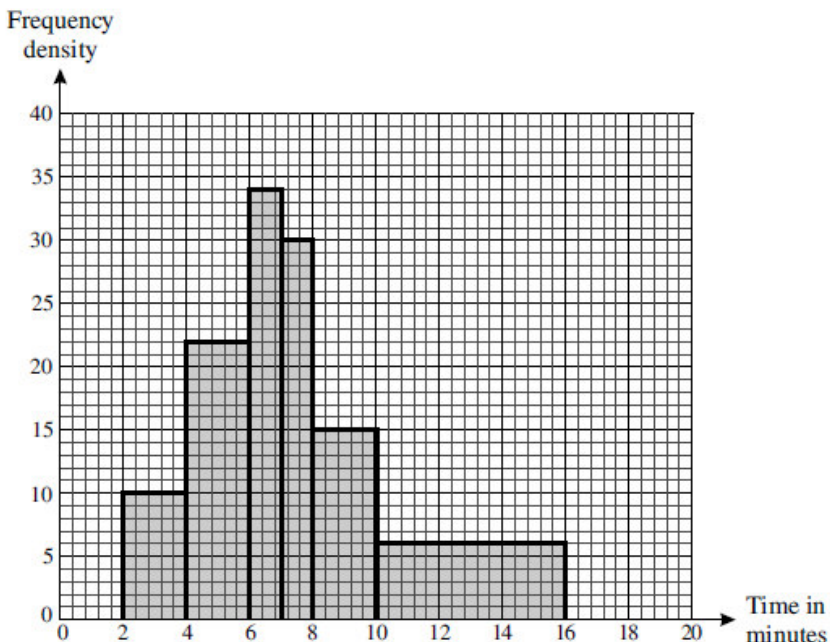
(i) Calculate the mean speed of these 70 cars. [2]

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]

5 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]