

- 1 The stem and leaf diagram illustrates the heights in metres of 25 young oak trees.

3		4	6	7	8	9	9		
4		0	2	2	3	4	6	8	9
5		0	1	3	5	8			
6		2	4	5					
7		4	6						
8		1							

Key: 4 | 2 represents 4.2

- (i) State the type of skewness of the distribution. [1]
- (ii) Use your calculator to find the mean and standard deviation of these data. [3]
- (iii) Determine whether there are any outliers. [4]
- 2 The mean daily maximum temperatures at a research station over a 12-month period, measured to the nearest degree Celsius, are given below.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
8	15	25	29	31	31	34	36	34	26	15	8

- (i) Construct a sorted stem and leaf diagram to represent these data, taking stem values of 0, 10, ... . [4]
- (ii) Write down the median of these data. [1]
- (iii) The mean of these data is 24.3. Would the mean or the median be a better measure of central tendency of the data? Briefly explain your answer. [2]



- 5 In a traffic survey, the number of people in each car passing the survey point is recorded. The results are given in the following frequency table.

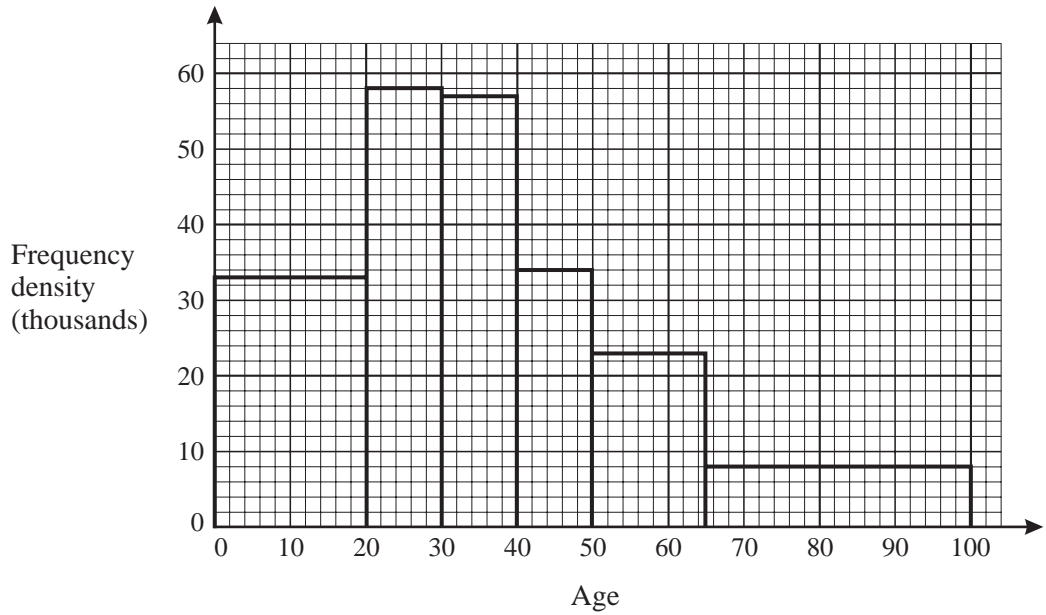
Number of people	1	2	3	4
Frequency	50	31	16	5

- (i) Write down the median and mode of these data. [2]
- (ii) Draw a vertical line diagram for these data. [2]
- (iii) State the type of skewness of the distribution. [1]
- 6 A supermarket chain buys a batch of 10 000 scratchcard draw tickets for sale in its stores. 50 of these tickets have a £10 prize, 20 of them have a £100 prize, one of them has a £5000 prize and all of the rest have no prize. This information is summarised in the frequency table below.

Prize money	£0	£10	£100	£5000
Frequency	9929	50	20	1

- (i) Find the mean and standard deviation of the prize money per ticket. [4]
- (ii) I buy two of these tickets at random. Find the probability that I win either two £10 prizes or two £100 prizes. [3]

7 The histogram shows the age distribution of people living in Inner London in 2001.



Data sourced from the 2001 Census, [www.statistics.gov.uk](http://www.statistics.gov.uk)

- (i) State the type of skewness shown by the distribution. [1]
- (ii) Use the histogram to estimate the number of people aged under 25. [3]
- (iii) The table below shows the cumulative frequency distribution.

Age	20	30	40	50	65	100
Cumulative frequency (thousands)	660	1240	1810	$a$	2490	2770

- (A) Use the histogram to find the value of  $a$ . [2]
- (B) Use the table to calculate an estimate of the median age of these people. [3]

The ages of people living in Outer London in 2001 are summarised below.

Age ( $x$ years)	$0 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 40$	$40 \leq x < 50$	$50 \leq x < 65$	$65 \leq x < 100$
Frequency (thousands)	1120	650	770	590	680	610

(iv) Illustrate these data by means of a histogram. [5]

(v) Make two brief comments on the differences between the age distributions of the populations of Inner London and Outer London. [2]

(vi) The data given in the table for Outer London are used to calculate the following estimates.

Mean 38.5, median 35.7, midrange 50, standard deviation 23.7, interquartile range 34.4.

The final group in the table assumes that the maximum age of any resident is 100 years. These estimates are to be recalculated, based on a maximum age of 105, rather than 100. For each of the five estimates, state whether it would increase, decrease or be unchanged. [4]