EXAM-STYLE QUESTIONS

9 The times taken for 50 students to complete a crossword puzzle are shown in the table.

Time (m minutes)	Frequency
$15 \leq m < 20$	3
$20 \leq m < 25$	7
$25 \leq m < 30$	10
$30 \leq m < 35$	11
$35 \leq m < 40$	12
$40 \leq m < 45$	5
$45 \leq m < 50$	2

Use the midpoint of each class to estimate the mean and the standard deviation of grouped data.

Find an approximation for the mean and standard deviation.

10 The percentage marks obtained for an ITGS (Information Technology for a Global Society) test by the 25 boys and 25 girls at Bright High are shown in the table.

a Calculate an estimated value for the mean and standard deviation for the girls and the boys separately.

b Comment on your findings.

Girls' frequency	Percentage mark	Boys' frequency
0	$0 \leq x < 10$	2
0	$10 \leq x < 20$	1
0	$20 \leq x < 30$	1
3	$30 \leq x < 40$	1
5	$40 \leq x < 50$	5
7	$50 \leq x < 60$	9
8	$60 \leq x < 70$	2
2	$70 \leq x < 80$	0
0	$80 \leq x < 90$	2
0	$90 \leq x < 100$	2

Review exercise

Paper 1 style questions

EXAM-STYLE QUESTIONS

1 The mean of the twelve numbers listed is 6.

3 4 a 8 3 5 9 5 8 6 7 5

a Find the value of a.

b Find the median of these numbers.

2 The mean of the ten numbers listed is 5.

4 3 a 6 8 4 6 6 7 5

a Find the value of a.

b Find the median of these numbers.

EXAM-STYLE QUESTIONS

3 For the set of numbers
3 4 1 7 6 2 9 11 13 6 8 10 6

a calculate the mean
b find the mode
c find the median.

4 The lengths of nine snakes, in meters, are:
6.5 4.6 7.2 5.0 2.4 3.9 12.9 10.3 6.1

a **i** Find the mean length of the snakes.
ii Find the standard deviation of the length of the snakes.
b Find the median length of the snakes.

5 A survey was conducted of the number of bathrooms in 150 randomly chosen houses. The results are shown in the table.

Number of bathrooms	1	2	3	4	5	6
Number of houses	79	31	22	10	5	13

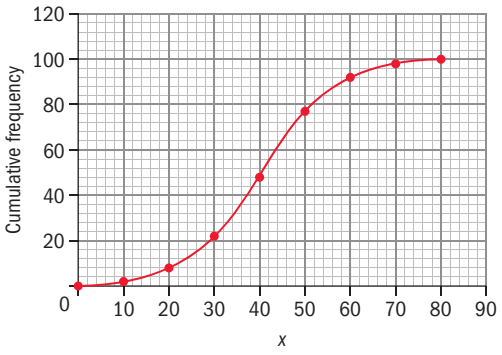
a State whether the data are discrete or continuous.
b Write down the mean number of bathrooms per house.
c Write down the standard deviation of the number of bathrooms per house.

6 The table shows the age distribution of members of a chess club.

Age (years)	**Number of members**
$20 \leq x < 30$	15
$30 \leq x < 40$	23
$40 \leq x < 50$	34
$50 \leq x < 60$	42
$60 \leq x < 70$	13

a Calculate an estimate of the mean age.
b Draw a histogram to represent these data.

7 Using the cumulative frequency graph, write down the value of
a the median
b the lower quartile
c the upper quartile
d the interquartile range.



EXAM-STYLE QUESTION

8 The numbers of horses counted in 35 fields are represented in the table.
Draw a box and whisker graph to represent this information.

Number of horses	Frequency
8	4
10	9
12	7
15	12
21	3

Paper 2 style questions

EXAM-STYLE QUESTIONS

1 Nineteen students carried out an experiment to measure gravitational acceleration in cm s^{-2}.
The results are given to the nearest whole number.

96 97 101 99 100 98 99 94 96 100
97 98 101 98 99 96 96 100 97

a Use these results to find an estimate for
- **i** the mean value for the acceleration
- **ii** the modal value for the acceleration.

b **i** Construct a frequency table for the results.
- **ii** Use the table to find the median value and the interquartile range.

2 A gardener wanted to estimate the number of weeds on the sports field.
He selected at random 100 sample spots, each of area 100 cm^2, and counted the number of weeds in each spot.

The table shows the results of his survey.

Number of weeds	Frequency
0–4	18
5–9	25
10–14	32
15–19	14
20–24	7
25–29	4

a **i** Construct a cumulative frequency table and use it to draw the cumulative frequency curve.
- **ii** Write down the median number of weeds.
- **iii** Find the percentage of spots that have more than 19 weeds.

b **i** Estimate the mean number of weeds per spot.
- **ii** Estimate the standard deviation of the number of weeds per spot.

The area of the field is 8000 m^2.
- **iii** Estimate the total number of weeds on the field.

3 The marks for a test are given in the frequency table.

Mark, x	Frequency
$0 \le x < 10$	3
$10 \le x < 20$	14
$20 \le x < 30$	21
$30 \le x < 40$	35
$40 \le x < 50$	42
$50 \le x < 60$	55
$60 \le x < 70$	43
$70 \le x < 80$	32
$80 \le x < 90$	15
$90 \le x < 100$	10

a Complete a cumulative frequency table and use it to draw the cumulative frequency curve.

b Find the median mark.

c Find the interquartile range.

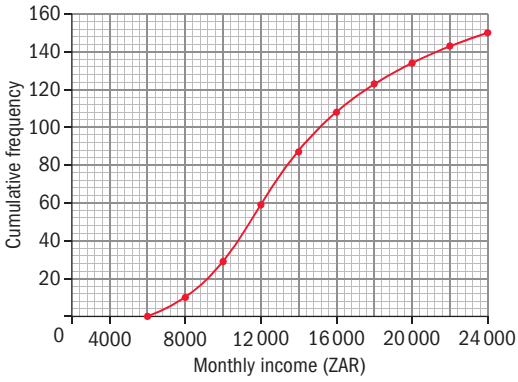
60% of the candidates passed the examination.

d Find the pass mark.

e Given that the lowest mark was 9 and the highest was 98, draw a box and whisker graph to represent the information.

EXAM-STYLE QUESTIONS

4 The cumulative frequency graph shows the monthly incomes, in South African Rand, ZAR, of 150 people.



a Write down the median and find the interquartile range.

b Given that the lowest monthly income is 6000 ZAR and the highest is 23 500 ZAR, draw a box and whisker graph to represent this information.

c Draw a frequency table for the monthly incomes.

d Use your GDC to find an estimate of the mean and standard deviation of the monthly incomes.

5 The weights of 200 female athletes are recorded in the table.

Weight (w kg)	Frequency
$45 \leq w < 50$	4
$50 \leq w < 55$	16
$55 \leq w < 60$	45
$60 \leq w < 65$	58
$65 \leq w < 70$	43
$70 \leq w < 75$	28
$75 \leq w < 80$	6

a Write down the modal group.

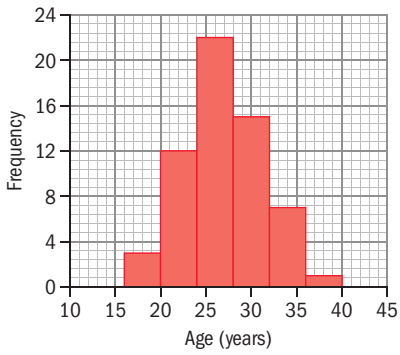
b Calculate an estimate of the mean and the standard deviation.

c Construct a cumulative frequency table and use it to draw the cumulative frequency graph.

d Write down the median, the lower quartile and the upper quartile.

e The lowest weight is 47 kg and the heaviest is 76 kg. Use this information to draw a box and whisker graph.

6 A group of 60 women were asked at what age they had their first child. The information is shown in the histogram.



a Calculate an approximation for the mean and standard deviation.

b Write down the modal class.

c Construct a cumulative frequency table for the data and draw the cumulative frequency curve.

d Use your graph to find the median and interquartile range.

e Given that the youngest age was 16 and the oldest was 39, draw a box and whisker graph to represent the information.

7 The average times, to the nearest second, that 100 participants waited for an elevator are shown in the table.

Time (t seconds)	Frequency
$0 \leq t < 10$	5
$10 \leq t < 20$	19
$20 \leq t < 30$	18
$30 \leq t < 40$	22
$40 \leq t < 50$	16
$50 \leq t < 60$	12
$60 \leq t < 70$	8

a Write down the modal class.

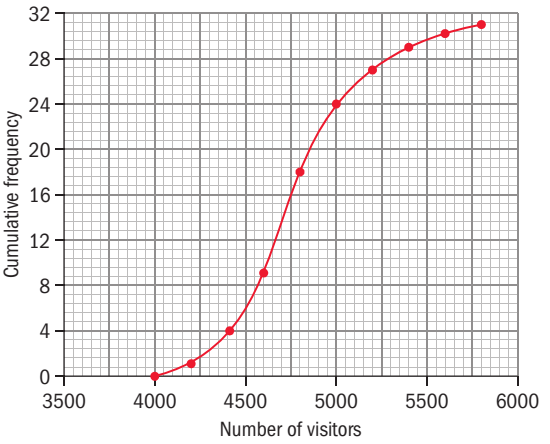
b Calculate an estimate of the mean time and the standard deviation.

c Construct a cumulative frequency table and use it to draw the cumulative frequency graph.

d Write down the median and interquartile range.

EXAM-STYLE QUESTIONS

8 The cumulative frequency graph shows the daily number of visitors to the Mausoleum on Tiananmen Square in the month of January.



a Write down the median, the lower quartile and the upper quartile.

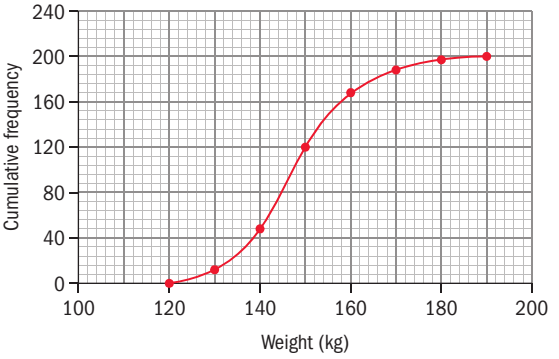
b Given that the least number of visitors was 4000 and the most was 5700, draw a box and whisker graph to represent the information.

c Construct a frequency table for this information.

d Write down the modal class.

e Calculate an estimate of the mean and the standard deviation.

9 The cumulative frequency graph shows the weights, in kg, of 200 professional wrestlers.



a Construct a grouped frequency table for this information.

b Write down the modal class.

c Calculate an estimate of the mean weight.

CHAPTER 2 SUMMARY

Classification of data

- **Discrete data** are either data that can be counted or data that can only take specific values.
- **Continuous data** can be measured. They can take any value within a range.

Grouped discrete or continuous data

- To draw a **frequency histogram**, find the lower and upper boundaries of the classes and draw the bar between these boundaries. There should be no spaces between the bars.

Measures of central tendency

- The **mode** of a data set is the value that occurs most frequently.
- The **median** of a data set is the value that lies in the middle when the data are arranged in size order.
- The **mean** of a data set is the sum of all the values divided by the number of values.
- For data in a frequency table, the **mode** is the entry that has the largest frequency.

Continued on next page

- The **median** is the middle entry as the entries in the table are already in order. For n pieces of data, the median is the $\frac{n+1}{2}$th value.
- The **mean** from a frequency table is:

  $$\text{mean} = \frac{\text{total of } f_i \times x_i}{\text{total frequency}}$$

  where f_i is the frequency of each data value x_i and $i = 1, \ldots, k$, where k is the number of data values.
- For grouped data, the **modal class** is the group or class interval that has the largest frequency.
- To calculate the **mean** from a grouped frequency table, an estimate of the mean is

  $$\frac{\text{total of } f_i \times x_i}{\text{total frequency}}$$

  where f_i is the frequency and x_i is the corresponding midpoint of each class.

Cumulative frequency curves

- The **cumulative frequency** is the sum of all of the frequencies up to and including the new value. To draw a **cumulative frequency curve** you need to construct a cumulative frequency table, with the upper boundary of each class interval in one column and the corresponding cumulative frequency in another. Then plot the upper class boundary on the x-axis and the cumulative frequency on the y-axis.
- To find the **lower quartile**, Q_1, read the value on the curve corresponding to $\frac{n+1}{4}$ on the cumulative frequency axis, where n is the total frequency.
- To find the median, read the value on the curve corresponding to $\frac{n+1}{2}$ on the cumulative frequency axis.
- To find the **upper quartile**, Q_3, read the value on the curve corresponding to $\frac{3(n+1)}{4}$ on the cumulative frequency axis.
- To find the **percentiles**, p%, read the value on the curve corresponding to $\frac{p(n+1)}{100}$ on the cumulative frequency axis.
- To find the **interquartile range** subtract the lower quartile from the upper quartile: $\text{IQR} = Q_3 - Q_1$.

Box and whisker graphs

- To draw a box and whisker graph, five pieces of information are needed: calculate the lower quartile, median and upper quartile for the data. Find the smallest and largest values.

Measures of dispersion

- The **range** is found by subtracting the smallest value from the largest value.
- The **interquartile range** is found by subtracting the lower quartile, Q_1, from the upper quartile, Q_3: $\text{IQR} = Q_3 - Q_1$.
- The standard deviation is often referred to as the 'root-mean-square deviation' because we find the **deviation** of each entry from the mean, then we **square** these values and find the **mean** of the squared values, and, finally, we take the square **root** of this answer.