Part E

Chapter 5

Descriptive statistics

EXERCISE 5D

1 The initial weights of students (in kg) in a fitness class were:

69 90 61 58 94 68 77 80 64 87 59 81 73 56 72 66
62 75 78 75 87 73 81 70 73 58 85 73 76 92 82 86
82 71 79 80 85 70 60 75

a Using classes 50 - , 60 - , 70 - , 80 - , 90 - , tabulate the data using columns of score, tally, frequency.

b How many students are in the 70 - class?

c How many students weighed less than 80 kg?

d Find the percentage of students who weighed 70 kg or more.

2 A group of young athletes were invited to participate in a javelin throwing competition. The following results were obtained:

Distance (metres)	0 -	10 -	20 -	30 -	40 -
No. of athletes	3	26	41	14	6

a How many athletes threw less than 20 metres?

b What percentage of the athletes were able to throw at least 30 metres?

3

Height (mm)	Frequency
300 -	12
325 -	18
350 -	42
375 -	28
400 -	14
425 -	6

A plant inspector takes a random sample of two week old seedlings from a nursery and measures their height to the nearest mm.

The results are shown in the table alongside.

a How many of the seedlings are 400 mm or more?

b What percentage of the seedlings are between 349 and 400 mm?

c The total number of seedlings in the nursery is 1462. Estimate the number of seedlings which measure:

i less than 400 mm   ii between 374 and 425 mm.

E   FREQUENCY DISTRIBUTION TABLES

Frequency tables may include other columns which show **relative frequency**, **cumulative frequency**, and **cumulative relative frequencies**.

The **relative frequency** of an event is the frequency of that event expressed as a fraction (or decimal equivalent) of the total frequency.

The **cumulative frequency** of an event is the accumulation (sum) of the frequencies up to and including that event.

The **cumulative relative frequency** of an event is the accumulation (sum) of the relative frequencies up to and including that event.

The data listed below gives the lengths of steel rods in metres as produced by a machine:

1.95	3.5	2.5	2.8	1.25	4.75	5.5	5.5	5.75	5.75	4.5	2.25	4.75	0.65	3.5
2.8	4.5	4.75	5.5	5.5	4.75	4.5	4.75	1.25	2.25	4.5	4.75	5.5	5.5	5.75
5.5	5.5	4.75	2.8	2.5	3.5	1.4	4.75	5.5	5.75	5.75	5.75	5.5	5.5	4.75
4.5	2.25	1.4	5.75	5.75										

More sense can be made of this data when it is organised according to size and frequencies and presented in a tabular format.

The various rod sizes can be listed in the table from lowest to highest and a tally column can be used to record the number of times each value (length) occurs.

See **Example 3** below.

CLASS INTERVAL

Another way that this data could be organised is by grouping the rod lengths together into various ranges of lengths called class intervals.

These class intervals are usually of **equal length** and must be constructed so that each value is assigned to one class interval only.

The size of the class intervals devised will depend on the data being organised, but it is generally thought that about 6 or 12 class intervals is the right number for a frequency table.

See **Example 4** on the next page.

Example 3

Prepare a frequency distribution table for the steel rod data. Include columns for length, tally, frequency, cumulative frequency, relative frequency and cumulative relative frequency.

Frequency distribution table for the Steel Rod production

Rod Length (m)	Tally	Frequency	Cumulative Frequency	Relative Frequency	Cumu. Rel. Frequency
0.65			1	1	0.02	0.02
1.25	||	2	3	0.04	0.06
1.40	||	2	5	0.04	0.10
1.95			1	6	0.02	0.12
2.25	|||	3	9	0.06	0.18
2.50	||	2	11	0.04	0.22
2.80	|||	3	14	0.06	0.28
3.50	|||	3	17	0.06	0.34
4.50	ЖЖ	5	22	0.10	0.44
4.75	ЖЖ ||||	9	31	0.18	0.62
5.50	ЖЖ ЖЖ		11	42	0.22	0.84
5.75	ЖЖ |||	8	50	0.16	1.00
Total		50			

Example 4

Rework the frequency distribution table using the class intervals:
0-0.99, 1.00-1.99, 2.00-2.99, etc.

Rod Length (metres)	Tally	Frequency	Cumulative Frequency	Relative Frequency	Cumulative Relative Frequency
0 - 0.99	|	1	1	0.02	0.02
1.00 - 1.99	||||	5	6	0.10	0.12
2.00 - 2.99	|||| |||	8	14	0.16	0.28
3.00 - 3.99	|||	3	17	0.06	0.34
4.00 - 4.99	|||| |||| ||||	14	31	0.28	0.62
5.00 - 5.99	|||| |||| |||| ||||	19	50	0.38	1.00
Total		50			

HISTOGRAMS

Once data has been summarised in a frequency table, it can be represented diagrammatically
by a **histogram**.

Example 5

Draw a frequency histogram of
the steel rod data from
Example 4.

Histogram of 'Rod length' data
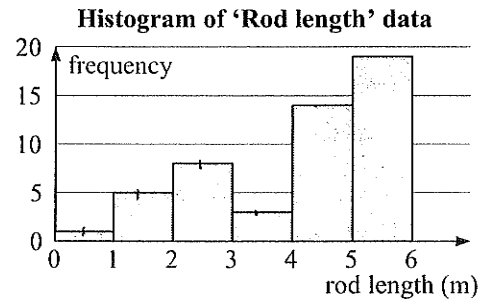
FREQUENCY POLYGON

A **frequency polygon** is a line graph which, like the histogram, gives a good visual appreci-
ation of the shape of the frequency distribution.

Instead of drawing the bars, the **midpoint** of each bar is found and is used to represent the
whole interval. These points are then joined by straight lines.

Example 6

Draw a frequency polygon of
the steel rod data from
Example 4.

Frequency polygon of 'Rod length' data
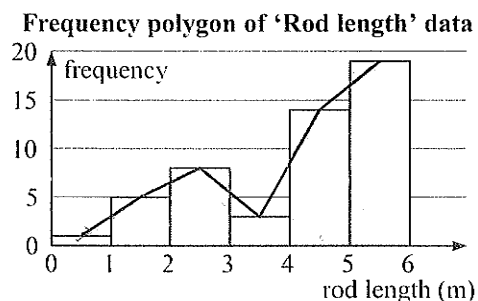
CUMULATIVE FREQUENCY GRAPH (POLYGON)

A **cumulative frequency graph (polygon)** is used to represent **cumulative frequencies**.

Since cumulative frequency is the sum of the frequencies up to and including the value being examined, we plot the cumulative frequencies against the **upper end points** of each class interval. These points are then joined by straight lines to produce the cumulative frequency graph.

The **cumulative frequency graph** is based on **increasing accumulations** and one is illustrated in the example which follows. The graph begins at the lowest boundary.

In this case the curve slopes upwards as we look from left to right on the graph. It gives all points less than any particular value. For instance, on the cumulative frequency graph which follows, if you wanted to know how many rods were less than, say, 4 metres, you could easily find the answer by reading off the graph. There are 17 such rods less than 4 metres.

Example 7

Draw a cumulative frequency graph for the steel rod data from **Example 4**.

Rod length (metres)	Upper End Point	Cumulative Frequency
0 - 0.99	0.995	1
1 - 1.99	1.995	6
2 - 2.99	2.995	14
3 - 3.99	3.995	17
4 - 4.99	4.995	31
5 - 5.99	5.995	50

Cumulative frequency graph of 'Steel rod' data
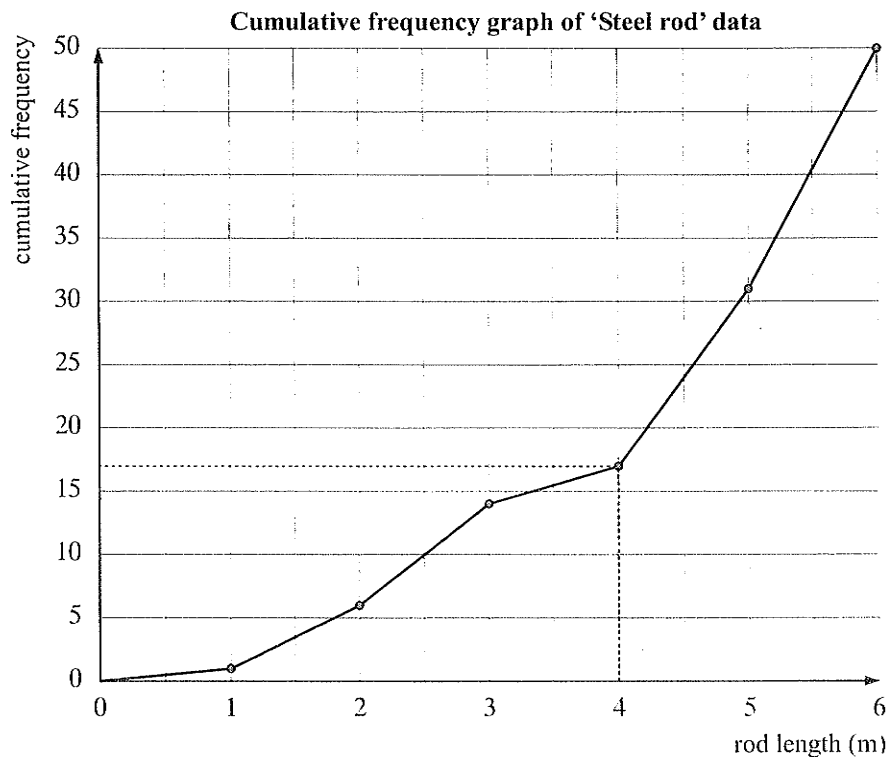
EXERCISE 5E

1   Copy and complete the table alongside including columns of cumulative frequency, relative frequency and cumulative relative frequency.

Value	Frequency
32	3
35	7
36	8
39	11
41	15
44	12
45	8
46	7
49	5
50	4
total	80

2   A bottle manufacturer records the following number of rejected bottles each day over a 30 day period:

112 127   92 147 134 131 104   99 116 122
125 118 109   96 142 127 106 100 132 138
113 123 118 131 115   94 144 124 117 124

a   Organise the data into six class intervals starting with 91 - 100 and ending with 141 - 150. Hence, complete a frequency table including columns of cumulative frequency, relative frequency and cumulative relative frequency.

b   From your table find:

i   how many times there were less than 120 rejections

ii   what percentage of time there were between 91 and 100 (inclusive) rejections

iii   what percentage of time there were less than 130 rejections.

c   Construct a histogram and a cumulative frequency graph from your table.

3   A country motel has room for 80 people.
The manager keeps records of the number of guests staying at the motel over the summer period of 90 days. The results are shown alongside:

No. of guests	Freq. (days)
1 - 10	8
11 - 20	11
21 - 30	14
31 - 40	17
41 - 50	20
51 - 60	9
61 - 70	7
71 - 80	4

a   Construct a frequency table for this data and show cumulative frequencies for each class interval.

b   Construct a histogram and frequency polygon of the data.

c   Prepare a cumulative frequency polygon of the data and hence find:

i   the number of nights when the motel had 30 or less guests

ii   the number of nights when the motel had less than 45 guests.

d   The motel will 'break even' if it has less than 30 guests for no more than 25% of the time. Did it break even?

4 A market research company conducts a survey on the number of times people eat out at restaurants each year. The results from the 100 people surveyed are as follows:

No. of times	Freq.
1 - 20	15
21 - 40	21
41 - 60	24
61 - 80	18
81 - 100	12
101 - 120	6
121 - 140	4

  a Construct a frequency table which also shows cumulative frequencies.

  b Construct a frequency polygon.

  c Construct a cumulative frequency polygon and use it to answer the following:

  i What percentage of people ate at restaurants less than 81 times a year?

  ii If a person was in the lower 25% of the distribution, what is the maximum number of times they would have eaten at restaurants during the year?
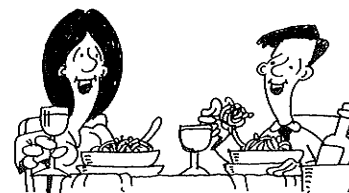
5 The manager of a fast food restaurant is concerned that customers are waiting too long for their food. She decides to gather some statistics on customer waiting times and the following times (in minutes) are recorded:

```
1.25   2.5   8.5   1.8   4.6   10.5   3.4   7.0   6.25   4.10   5.15   5.95   7.35   5.8   2.9
0.9    3.4   6.0   8.8   2.7   10.2   4.5   5.2   4.1    2.5    7.7    3.8    2.1    5.5   6.25
4.3    1.8   8.0   9.9   3.7   4.4    6.2   3.3   7.2    8.6    3.45   6.55   2.85   9.4   4.25
5.6    11.9  6.4   4.8   5.8
```

  a Group this data into classes of 0 - 1.99, 2 - 3.99, etc. and construct a frequency table which also shows cumulative frequencies.

  b Construct a histogram of the data.

  c Construct a cumulative frequency graph and answer the following questions:

  i How many customers have to wait less than 4 minutes for their food?

  ii What percentage of customers have to wait more than 5 minutes for their food?

  iii If the restaurant's goal is for 90% of the customers to be given their food within 8 minutes, are they achieving this goal?

F   SUMMARISING THE DATA

MEASURES OF THE MIDDLE OF A DISTRIBUTION

After collecting and presenting statistical data, you can now attempt to interpret the data. One way of doing this is to find the value of the **centre or middle** of the distribution.

There are three commonly used measures for the middle of a distribution; the **mean**, the **mode** and the **median**.

However, before proceeding further we will define some of the terms that will be used from now on:

• **Ungrouped data** comprises of single values which have not been put into groups or classes.

    For example, the heights of five children are 1.23 m, 1.56 m, 1.34 m, 1.09 m, 1.71 m.