1 Ungrouped Data

Learning Objectives

In this chapter you will learn:

- How to summarise and display data, and about the importance and usefulness of doing this
- How to interpret data presented in various forms and to extract useful information
- About the advantages, disadvantages and appropriateness of various forms of summary and display
- How to make relevant comparisons between the items in a set of data

Introduction

Data are facts and statistics collected together for reference or analysis. Before they are processed for use they are referred to as **raw data**. The processing of raw data may involve summarising in ordered lists and tables, or displaying in diagrams and graphs. There is never just one way of doing this – there is always a choice – and each method has its own advantages and disadvantages, which also depend on the type and subject of the data, how much of it there is, and the purpose for which the summary or display is being made. In **ungrouped** data, individual values can be seen separately from other values. Once processed in an appropriate way, calculations can be performed, and the data can be analysed.

1.1 Summary and Display of Data

When summarising or displaying data, a very important consideration is that the presentation should be clear and easy to understand. Who is it being presented to, and for what purpose is it being done?

Two-way Tables

One method of summarising data is to enter it into a table – data can be **tabulated**.

The columns and rows of a table should have headings, and totals can be included when appropriate.

Numbers and proportions (percentages, fractions and ratios) can then be compared and interpreted.

Examples

Table 1.1 shows the results of a geography test that was taken by a group of 45 students.

Table 1.1: Geography test results

	Pass	Fail	Totals
Boys	22	4	26
Girls	17	2	19
Totals	39	6	45

From Table 1.1 we can see that:

- 26 boys took the test
- 2 girls failed the test
- 39 students passed the test
- $\frac{17}{19}$ of the girls passed the test
- $\frac{1}{3}$ or 33.3% of those who failed were girls
- the ratio of passes to failures for boys was $11:2$.

Exercise 1A

1 Table 1.2 gives information about 100 students who took an examination in physics.

Table 1.2: Physics examination results

	Pass	Fail	Totals
Boys	45	15	60
Girls	28	12	40
Totals	73	27	100

i How many students passed?
ii What fraction of the students failed?
iii What fraction of the girls passed?
iv Calculate:
   a the percentage pass rate for boys,
   b the percentage pass rate for girls.
v Which group performed better, boys or girls?

2 A woman sells red and yellow T-shirts at a market. She has small, medium and large in each colour. There are 120 T-shirts altogether, and 75 of them are red. She has 20 small red T-shirts, 14 medium yellow T-shirts and 48 large T-shirts. The ratio of large red T-shirts to large yellow T-shirts is $2:1$.

i Copy and complete Table 1.3.

Table 1.3: T-shirt stocks

	Small	Medium	Large	Totals
Red	20			75
Yellow		14		
Total			48	120

ii Find the fraction of the T-shirts that are:
   a yellow,   **b** medium.
iii What fraction of the red T-shirts are large?
iv What percentage of the small T-shirts are yellow? Answer to 1 decimal place.

3 On sports day a girl decided to sell two types of canned drinks: Cherry-Fizz and Orange-Wizz. She had 15 bottles of both types of drink in small and medium sizes. She had 13 small bottles altogether, and 7 of the medium bottles contained Orange-Wizz.

i Tabulate the data with the numbers of each of the four items.
ii What fraction of the bottles were small Cherry-Fizzes?
iii How many of the bottles were neither small nor contained Orange-Wizz?

4 At lunchtime 100 students were each asked to choose one main course from either stew or hot pot. The students were also asked to choose one dessert from either ice cream or tart. Altogether 58 chose stew, 24 chose hot pot and ice cream, and 44 chose tart. All of the students chose one main course and one dessert.

i Illustrate the data in a fully labelled table.
ii What fraction of the students chose both hot pot and tart?

5 Details about the 50 motor cars being sold by a second-hand car dealer are given in Table 1.4.

Table 1.4: Second-hand cars available

	Saloon		Hatchback	
	Petrol	Diesel	Petrol	Diesel
Manual	7	4	12	7
Automatic	5	3	9	3

i What percentage of the motor cars are automatic?
ii What fraction of the hatchbacks use diesel?
iii What percentage of the automatic saloons use petrol?
iv Which has the highest proportion of motor cars that use diesel, saloons or hatchbacks?

6 A boy has 60 discs: some are black and some are white. The discs are either plastic or wooden.

Forty percent of the discs are black, and $\frac{5}{9}$ of the white discs are plastic.

Of the wooden discs, there is one less black than white.

i Draw up a table, with headings, showing the number of each of the four types of disc.
ii What proportion of the discs is not a black plastic disc?

7 A survey gave the data shown in Table 1.5 on the numbers of adults and the numbers of employed adults in each of the households in a particular street.

Table 1.5: Adult employment numbers

No. employed adults	No. adults		
	1	2	3
0	6	3	0
1	4	5	2
2	—	5	4
3	—	—	1

i In how many households was there more than one adult?
ii In how many households were all the adults employed?
iii How many households had just one unemployed adult?
iv How many employed adults were there in all these households together?
v Explain why there are no values entered into three of the cells of Table 1.5.

8 Table 1.6 gives information about the method used by some employees to get to work today, and how they intend to get to work tomorrow.

Table 1.6: Commuting methods

		Today	Tomorrow
Bus	Male	10	8
	Female	9	10
Taxi	Male	3	3
	Female	4	3
Walk	Male	15	17
	Female	12	12

i How many employees are represented in Table 1.6?
ii How many of the employees are male?
iii How many of the employees walked to work today?
iv How many of the employees plan to take a bus to work tomorrow?
v Give details about Dora, the only female intending to travel by a different method tomorrow.
vi Explain why it is not necessarily true that only two males intend to travel by a different method tomorrow.

9 Joan wants to tabulate data showing the numbers of students that passed or failed each of the three final papers in geography at her school last year and this year. Altogether 322 students sat for each of the three papers last year, which was 23 fewer than this year. The numbers passing Papers 1, 2 and 3 last year were 310, 303 and 305, respectively. Fifteen more students failed Paper 2 this year than last year; equal numbers passed Paper 1 in both years, and three times as many students failed Paper 3 this year than last year.

i Copy and complete Table 1.7, in which these data are tabulated.

Table 1.7: Geography paper results

	Last year		This year	
	Pass	**Fail**	**Pass**	**Fail**
Paper 1	310			
Paper 2	303			
Paper 3	305			

ii What is the greatest possible number of students that failed all three papers this year?
iii What is the least possible number of students that failed all three papers last year?

10 The 124 employees at a company are classified by gender, employment status and skill. The diagram below shows, for example, that there are 25 skilled full-time females, and 6 unskilled part-time males.

Figure 1.1: Full- and part-time employees

Find:

i the number of skilled employees,
ii the number of male employees,
iii the number of unskilled full-time female employees,
iv what fraction of the full-time employees are skilled,
v what percentage of the part-time employees are male,
vi what fraction of the skilled males work part time,
vii what percentage of the unskilled part-timers are female.

Frequency distributions

A small amount of data can be written conveniently as a list of numbers or words. We can list ten students' scores in a quiz, or the favourite colours of ten people. However, if we have to list the quiz scores of 150 students or the favourite colours of 300 people, then it would take a long time and the lists would not convey much useful information to anyone who looked at them. In cases like this, it would be more appropriate to summarise the data as a **frequency distribution**, where the data are shown in a frequency table.

A frequency distribution is ungrouped when individual values of the variable are listed.

If the variable takes numerical values then the values should be ranked, usually in ascending order.

The letter *f* is used to represent the **frequency**, which indicates the number of times that a particular value occurs in a set of data.

1 Ungrouped Data

Examples

1 The frequency distribution in Table 1.8 shows the 50 values of a variable X.

Table 1.8: Values of variable X

Values of X	0	1	2	3	
Frequency of values (f)	5	10	20	15	$\Sigma f = 50$

Table 1.8 is a simple alternative to making an ordered list of the 50 values of X.

The list of 50 numbers would begin with five 0's and end with fifteen 3's.

If, for example, X represents the number of goals scored by a football team in each of its 50 games, we could see very easily that the team scored:

- 0 goals in each of 5 games
- 1 goal in each of 10 games
- 2 goals in each of 20 games
- 3 goals in each of 15 games.

So we can calculate that the team scored $(0 \times 5) + (1 \times 10) + (2 \times 20) + (3 \times 15) = 95$ goals in 50 games.

2 A survey of a particular street was made to find the number of occupants in each house. The results are displayed as a frequency distribution in Table 1.9.

Table 1.9: Frequency table for number of house occupants

Number of occupants (X)	Number of houses(f)	Number of houses $\times$ number of occupants (fX)
0	0	0
1	4	4
2	7	14
3	9	27
4	13	52
5	7	35
	$\Sigma f = 40$	$\Sigma fX = 132$

X represents the numbers of occupants, which is the variable.

f represents the number of houses, which is the frequency.

$\Sigma f = 40$ is the total number of houses.

$\Sigma fX = 132$ is the total number of occupants.

There are:

- No empty houses
- 4 houses each with 1 occupant = $4 \times 1 = 4$ occupants
- 7 houses each with 2 occupants = $7 \times 2 = 14$ occupants
- 9 houses each with 3 occupants = $9 \times 3 = 27$ occupants
- 13 houses each with 4 occupants = $13 \times 4 = 52$ occupants
- 7 houses each with 5 occupants = $7 \times 5 = 35$ occupants.

Altogether there are 40 houses with 132 occupants.

Exercise 1B

1 The numbers of late arrivals at a certain school during each consecutive school day in June were:

5, 3, 4, 0, 1, 2, 7, 3, 3, 1, 4, 3, 3, 2, 4, 3, 0, 5, 3, 4, 2

  i Show these data as an ungrouped frequency distribution in a table with three columns, as in Table 1.10. Include two totals in the last row of the table.

Table 1.10: Late arrival frequencies

Number of late arrivals (X)	Number of days (f)	fX
0	2	0

  ii Explain what each of the two totals in your table represents.
  iii Find:
     a the most common number of late arrivals,
     b on how many days there were fewer than 4 late arrivals,
     c on how many days there were 5 or more late arrivals,
     d on what proportion of the days there were 3 late arrivals.

Albert arrived late in June more frequently than any other student.

  iv What is the greatest possible number of days on which Albert could have arrived late in June?

2 Some female doctors were asked, 'How many children do you have?' Their responses were:

2, 4, 2, 0, 1, 0, 0, 2, 1, 3, 0, 2, 5, 1, 3, 1, 0, 2, 1, 4, 1, 0, 2, 0, 3

  i Show these data as an ungrouped frequency distribution in a table.
  ii Copy and complete the sentence:

   'These . . . female doctors have . . . children altogether.'

  iii How many of the female doctors have more than 1 but fewer than 5 children?

3 The frequency table in Table 1.11 shows the ages of the members of a cycling club.

Table 1.11: Cycling club membership

Age (years)	under 12	12	13	14	15	16	17	18	over 18
Number of children (f)	0	3	5	12	7	2	0	1	0

  i Find:
     a the most common age of the members of the club,
     b how many children are members of this club,
     c the proportion of the membership that is 16 years old,
     d the number of members who are under 14 years of age,
     e the percentage of members who are 12 years old.
  ii Find the total of the ages of the 10 youngest members.
  iii Calculate the total of all the members' ages.

7

4   The shoe sizes of a group of children are shown in Table 1.12.

Table 1.12: Shoe sizes

Shoe size	No. children (*f*)
$3\frac{1}{2}$	7
4	11
$4\frac{1}{2}$	13
5	26
$5\frac{1}{2}$	31
6	22
$6\frac{1}{2}$	14
7	6

i   How many children are represented in Table 1.12?
ii Find the number of children that wear:
    a   sizes 4 to 5 inclusive,
    b   sizes 5 to 7 exclusive.
iii What fraction of the children wear size $4\frac{1}{2}$?
iv What percentage of the children wear size 5?
v   Of those that wear size 5 or larger, what fraction wear size 6 or smaller?

5   Table 1.13 shows the number of aircraft that took off from an airfield every day for 30 days.

Table 1.13: Daily flights

Number of aircraft	5	6	7	8	9	10	11	12	
Number of days (*f*)	2	4	7	9	4	3	0	1	Total = 30

i   What is the most common number of aircraft that took off per day?
ii On how many days did more than 9 aircraft take off?
iii How many aircraft took off in these 30 days altogether?
iv On what percentage of the days did at most 6 aircraft take off?

6   Table 1.14 shows the angles that are to be used in a pie chart illustrating the number of days' sick leave taken last month by each of the employees at a computer software company. Just one employee took 6 days' sick leave.

Table 1.14: Sick leave

No. days sick leave	0	1	2	3	4	5	6	
Sector angle	189°	72°	39°	33°	15°	9°	3°	Total = 360°

i   Find the number of employees that took 5 days' sick leave.
ii How many employees are there at the company?
iii What proportion of the employees took sick leave last month?
iv Show the data in a table as an ungrouped frequency distribution.
v   Calculate the total number of days' sick leave that were taken.

1.2 Pictorial Representation of Data

A variety of diagrams and graphs can be used to illustrate ungrouped data.

Diagrams usually appeal to people more than lists or tables, as they impart information at a glance.

The diagrams that are frequently used are pictograms, bar charts, line graphs, pie charts and Venn diagrams.

Pictograms

Data are illustrated by the use of symbols with a key to indicate what each symbol represents.

Care must be taken when drawing symbols: if two symbols are different in shape or size, they will not represent the same item or the same number of items. Symbols with simple shapes are recommended, as it may sometimes be necessary to draw fractions of one.

Examples

Three neighbours have grown cabbages in their gardens. The pictogram in Figure 1.2 shows how many cabbages each of them grew.

	represent 4 cabbages
David	:: :: :: :: :: :: :: ::
Samuel	:: :: :: :: :: ::
Bonolo	:: :: :: :: :: :: :

Figure 1.2: Cabbages grown

The three neighbours grew a total of 84 cabbages.

David grew 32; Samuel grew 24 and Bonolo grew 28.

$\frac{28}{84} = \frac{1}{3}$ of the cabbages were grown by Bonolo.

i Express the number of cabbages grown by David to the number grown by Samuel as a simple ratio.

ii What percentage of the cabbages were not grown by Samuel? Give your answer correct to 1 decimal place.

Exercise 1C

1 Forty-five students were asked which sports they play. All students play at least one sport.
The data were collected, and are illustrated in the pictogram in Figure 1.3.

Sport	⇨ **Represents two students**
Softball	⇨ ⇨ ⇨ ⇨ ⇨ ⇨
Football	⇨ ⇨ ⇨ ⇨ ⇨ ⇨ ⇨
Volleyball	⇨ ⇨ ⇨ ⇨ ⇨
Table tennis	⇨ ⇨ ⇨
Badminton	⇨
Tennis	⇨ ⇨

Figure 1.3: Numbers playing different sports

i Find:
  a the most popular sport, b the least popular sport,
  c the fraction that play volleyball, d the percentage that play tennis.
ii Explain why there appear to be more than 45 students represented in Figure 1.3.
iii Express the number that play softball to the number that do not play table tennis as a simple ratio.

2 A farmer grew five different vegetables on his farm. The number of each produced is given in Table 1.15.

Table 1.15: Vegetable yields

Vegetable	No. produced
Cabbage	32
Carrot	64
Onion	48
Tomato	80
Potato	40

i Draw a symbol that would be suitable for showing these data in a pictogram, and state what it would represent.
ii What is the total number of vegetables produced?
iii What fraction of the total number of vegetables are onions?
iv What percentage, to 1 decimal place, are tomatoes?
v Illustrate the data in a pictogram.

9

3 The different types of accommodation that a number of families inhabit are illustrated in the pictogram in Figure 1.4.

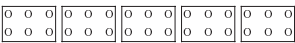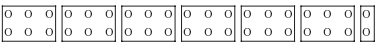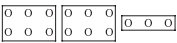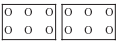|	Represents 6 families
Detached house	
Semi-detached house	
Terraced house	
Bungalow	
Flat	
Houseboat	

Figure 1.4: Family housing

i Explain why you think this particular symbol was chosen to represent 6 families.
ii How many more families live in semi-detached houses than in detached houses?
iii How many fewer families live in a flat than a bungalow?
iv How many families are represented altogether?
**v** What fraction of the families lives in a terraced house?

Bar Charts

Bars or columns of equal width are drawn to the heights of the numbers of items, which are the frequencies. Equal-width gaps may be left between bars, and all bars should be unshaded, or all shaded in the same way – avoid using many bright colours or random patterns, as this may give a false impression of the relative size of the bars.

Advantages of displaying data in a bar chart are that most people understand them, and the frequencies are easily seen and compared. A bar chart is appropriate if there are a fairly small number of values to display, usually not more than eight.

Examples

1 A child asked each of a group of adults, 'What is your favourite drink?'

The adults' responses were: 10 tea, 7 coffee, 5 juice, 4 milk, 3 water and 1 lemonade, as depicted in Figure 1.5.
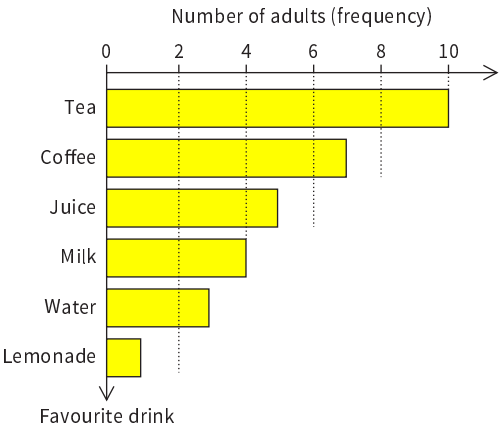
i How many adults were in the group?
ii What fraction of the adults prefer coffee?
iii What percentage of the adults do not prefer water?

Figure 1.5: Drink preferences

2   Each pupil in a class was asked to write down how many times they had checked their emails in the past week. The results are shown in the bar chart in Figure 1.6.
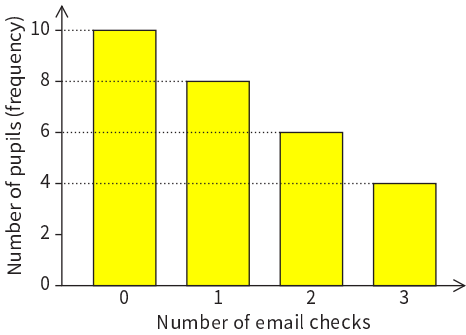

Figure 1.6: Email checking frequency

Table 1.16: Email checking frequency distribution

Number of email checks (X)	Number of pupils (f)
0	10
1	8
2	6
3	4

The data in the bar chart can be shown as a frequency distribution, as in Table 1.16.

i   How many pupils are in this class?   **ii** How many email checks were made altogether?

iii What fraction of the pupils did   **iv** What percentage of the pupils did not check their
check their emails?   emails at all?

Exercise 1D

1   The bar chart in Figure 1.7 shows the methods of transport used by some students when travelling to school.

Figure 1.7: Getting to school
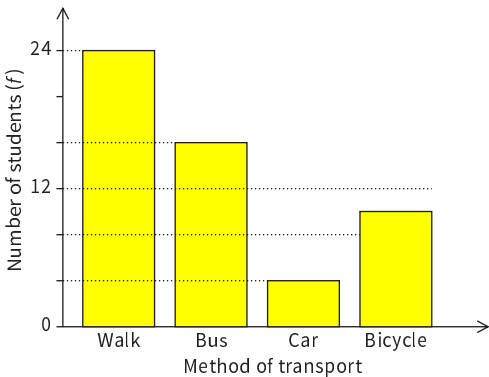
i   Name the least common method of transport.

ii How many more students walk than use the bus?

iii Express as a simple ratio the number of students who use the bus to the number who cycle.

iv What percentage of the students travel in a car? Answer correct to 1 decimal place.

v What fraction of the students do not walk to school?

2   The owner of a shop recorded the number of bags of potatoes sold on each of five days last week, as shown in Table 1.17.

Table 1.17: Bags of potatoes sold

Day	Monday	Tuesday	Wednesday	Thursday	Friday
No. bags	40	23	31	25	47

i   Illustrate the data in a bar chart.

ii How many bags of potatoes were sold last week?

iii Between which two consecutive days did the largest change in the number of bags sold occur?

iv Is it necessarily true that more potatoes were sold on Thursday than on Tuesday? Explain.