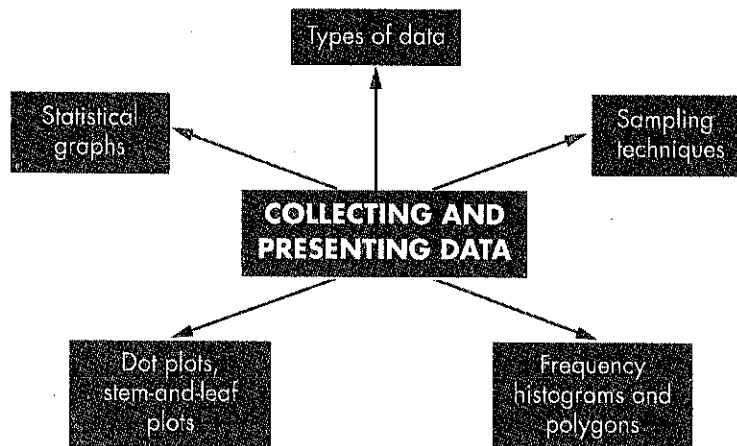This chapter, Collecting and presenting data, looked at the graphing of data, sampling techniques and the methods of data collection. You should be able to interpret and construct various graphs and displays, know the different types of data and samples, and be familiar with statistical terminology.

Make a summary of this topic. Use the chapter outline at the beginning of this chapter and the mind map below as a guide. Use your own words, symbols, diagrams, boxes and reminders. Gain a 'whole picture' view of the topic and identify any weak areas.

Types of data

Statistical graphs

Sampling techniques

COLLECTING AND PRESENTING DATA

Dot plots, stem-and-leaf plots

Frequency histograms and polygons

Collecting & Presenting Data

nese philosopher

nd training.

; a subject.

ig another

ow it all!
ire'.

· reading needed.
ims are more

past exams.

1 This bar graph shows the ten most fatal types of cancer for males in Australia in 2013.

Types of cancer deaths for males
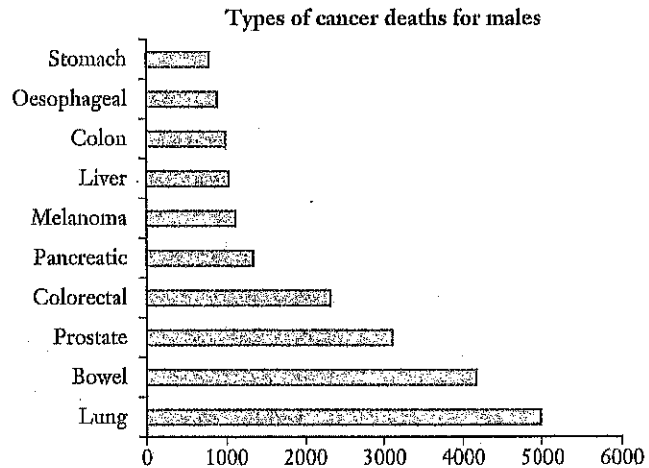
Source: Australian Institute of Health and Welfare 2013

a How many fatal cases of bowel cancer were registered?

b What was the fifth most fatal type of cancer?

c How many types of cancer recorded fewer than 2000 cases each?

d 'There were more lung cancer deaths recorded than bowel and prostate cancer deaths combined.' True or false?

e Why would it be unsuitable to graph this data on a sector graph?

2 A sample of consumers was surveyed about the brand of shower gel they used. This graph of the survey results was used by Green Water in their advertising campaign.

a Describe one way in which this graph is misleading and what a correct graph should look like.

b Name another type of graph that could be used to display this data.

Shower gel used
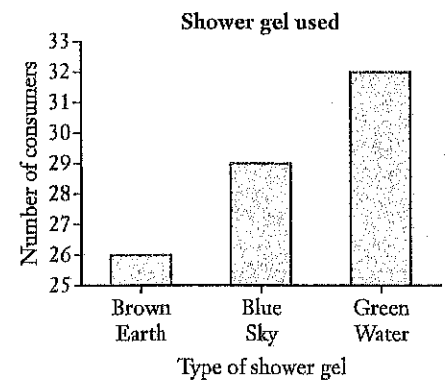
3 Classify each of the following types of data as categorical (C) or numerical (N).

a the number of people watching a State of Origin football match

b the different brands of surfing magazines

c level of university degree

d the temperatures at an airport

e the number of levels in an office building

f the country of birth of each person living in Gosford.

4 For each of the categorical data identified in Question **3**, classify as nominal (N) or ordinal (O).

5 For each of the numerical data identified in Question **3**, classify as discrete (D) or continuous (C).

6 The manufacturers of Burp Cola want to conduct a quality control study to ensure that each bottle of drink made does, in fact, contain 1.25 litres. Which type of sample would be most appropriate: random, systematic, stratified or self-selected? Give a reason for your answer.

7 a To investigate the types of music to play at the school disco, Tasha surveyed a sample of 50 students visiting the canteen. Why might this sample be biased?

b Suppose Tasha uses a stratified sample instead, based on the following numbers of school students in each Year group:

Year	7	8	9	10	11	12
Number of students	114	120	114	128	105	96

If she surveys a sample of 50 students, how many Year 10 students should be in the sample?

8 Explain what is wrong with each survey question below and write a better question each time.
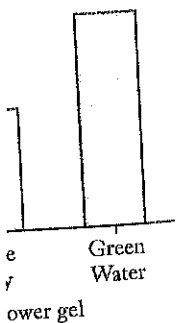
a How often do you visit the doctor?

b What did you like or dislike about the film?

**c** Is it time that Australia grew up and became a republic?

9 A sample of senior citizens was surveyed about the number of countries they have visited. The results are shown in the table.

Number of countries	Number of people
0	3
1–3	8
4–10	32
11–20	28
21–30	17
31+	5

a How many people were in the sample?

b Graph this data on a sector graph.

c What is the angle of the sector representing those who have not travelled outside Australia?

d What percentage of the sample have visited 4 to 10 countries? Answer correct to one decimal place.

e 'More than half of the sample have visited 11 or more countries.' True or false?

f Why would it be inappropriate to graph this data on a line graph?

10 A company surveyed its employees on reasons for arriving late to work, to help assist people to be more punctual. The results are shown on the Pareto chart.

Causes of lateness to work
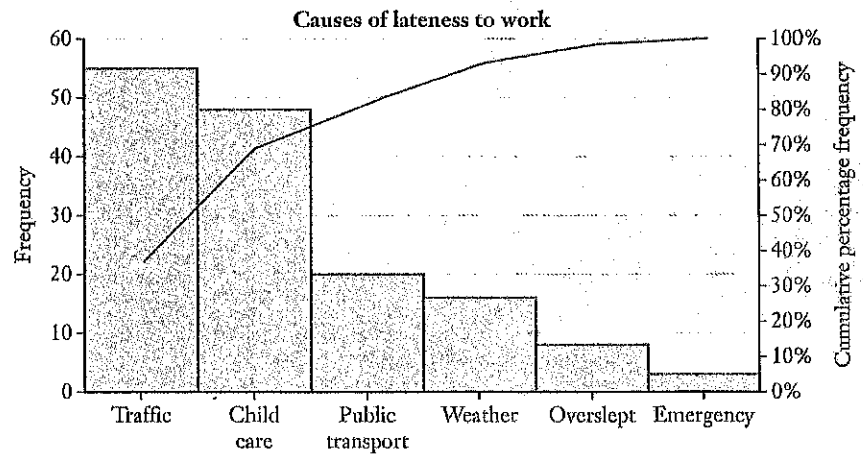
a What are the two main reasons employees give for being late to work?

b What is the least common reason for being late to work?

c How many employees are late due to using public transport?

d What percentage of employees are late due to using public transport?

e What are two possible ways this company might assist its employees to improve their ability to arrive at work on time?

11 The masses (in kilograms) of 40 skydivers were recorded. The results are shown below.

58	63	77	82	53	69	65	80	96	105
79	63	52	90	104	85	65	87	68	105
65	87	109	84	62	75	102	78	93	84
68	105	74	59	68	74	88	66	70	62

a Are people's masses discrete data or continuous data?

b Organise this data into a frequency distribution table with class intervals 50–< 60, 60–< 70 etc.

c Draw a frequency histogram and polygon to represent the data.

d Which class interval had the highest frequency?

e What fraction of skydivers were in the 80–< 90 kg class?

12 The ages of a sample of children at a Wiggles concert are listed below.

| 6 | 0 | 4 | 1 | 5 | 6 | 4 | 2 | 8 | 6 | 4 | 4 |
| 3 | 6 | 5 | 5 | 2 | 5 | 3 | 1 | 5 | 6 | 3 | 3 |

a Draw a dot plot for this data.

b How many children were in the sample?

c What fraction of children were over 5 years old?

d Comment on any clusters.

13 This stem-and-leaf plot shows the ages of visitors entering the Royal Easter Show in a 5-minute period.

a What was the age of the youngest visitor?

b How many visitors entered the show during the 5-minute period?

c What was the most common age of the visitors?

d What percentage (correct to the nearest whole number) of visitors were over 30 years old?

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

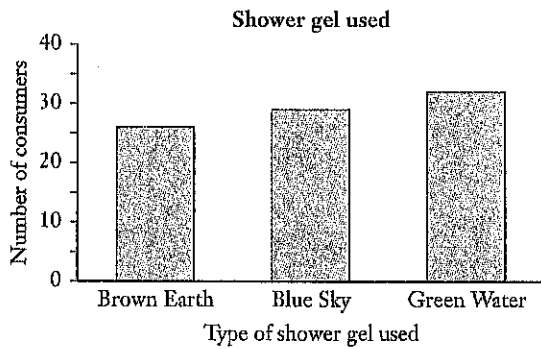Chapter quiz

Solutions

Test yourself 1

1 a 4150   **b** pancreatic   **c** 6
  d false
  e These cancers do not represent parts of a
    whole.
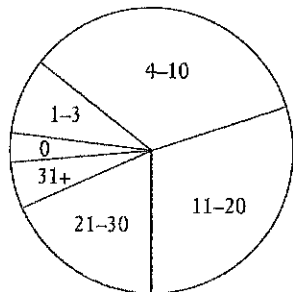**2 a** The vertical scale should start at zero.

Continued next
page

Shower gel used

b sector graph, divided bar graph

3 a N   **b** C   **c** C

   d N   **e** N   **f** C

4 b N   **c** O   **f** N

5 a D   **d** C   **e** D

6 systematic, for example, take every 10th bottle and test

7 a The sample may not necessarily take students from each Year into account.

   b 9

8 Teacher to check

9 a 93

   b Number of countries visited

   c 12°   **d** 34.4%   **e** True

   f The data is not continuous; the class intervals are not equal in size.

10 a traffic, child care

   b emergency

   c 20   **d** 35%

   e provide child care at workplace, change location of company to be near public transport so traffic not a major issue

11 a continuous

Mass (kg)	Frequency
50–< 60	4
60–< 70	12
70–< 80	7
80–< 90	8
90–< 100	3
100–< 110	6
	40

Skydivers' mass

d 60–< 70   **e** $\dfrac{1}{5}$

12
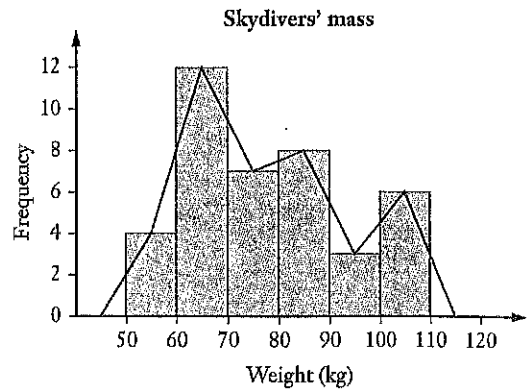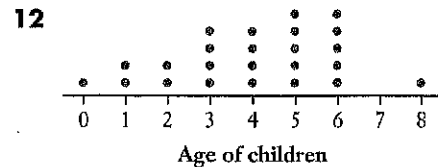
   b 24   **c** $\dfrac{1}{4}$

   d Data are clustered around 3–6 years.

13 a 3 years   **b** 29   **c** 12   **d** 41%

Chapter 2

SkillCheck

1 a −27   **b** 6   **c** −8

   d 25   **e** −4   **f** 2

2 a 15   **b** 3   **c** 16

   d 6   **e** 243   **f** 18

3 a 0   **b** −3   **c** 10

   d −128   **e** 9   **f** 9