Collecting and analysing data

This chapter deals with single variable data analysis. After completing this chapter you should be able to:

- identify everyday questions and issues
 involving numerical and categorical
 variables
- describe data using the terms 'symmetric', 'skewed' and 'bimodal'
- construct back-to-back stem-and-leaf plots and histograms
- construct side-by-side histograms and dot plots
- compare data displays using the mean, median and range
- evaluate statistical reports in the media.

NSW Syllabus references: 5.1 S&P Single variable data analysis Outcomes: MA5.1-1WM, MA5.1-2WM, MA5.1-3WM, MA5.1-12SP, MA5.1-13SP STATISTICS & PROBABILITY - ACMSP228, ACMSP253, ACMSP282, ACMSP283

Investigating data

Exercise 3A

Table 1 shows the monthly and annual rainfall for Sydney (Observatory Hill) from 2002 to 2011. Measurements are to the nearest millimetre.

Year	J	F	Μ	Α	Μ	J	J	Α	S	0	Ν	D	Annual
2002	98	348	45	68	93	28	24	20	22	6	32	75	860
2003	14	59	132	192	349	76	58	43	6	103	109	60	1200
2004	51	129	101	33	8	39	44	153	60	234	67	76	995
2005	68	125	154	33	48	79	63	2	51	43	125	25	816
2006	121	51	40	10	40	177	140	86	192	17	45	74	994
2007	45	108	65	180	10	511	67	152	41	27	170	123	1499
2008	57	258	63	147	3	127	90	44	99	67	73	54	1083
2009	25	128	61	153	126	130	53	6	16	180	13	67	956
2010	36	239	51	30	168	147	115	27	42	85	130	83	1154
2011	54	18	192	206	136	94	282	52	72	37	148	78	1369

Table 1: Rainfall for Sydney (mm)

1 In this time period, which year had the:

- **a** highest annual rainfall?
- **2** How much rain fell in:
 - **a** January 2006? **b** May 2007?
- 072
- 3 Which month had the highest rainfall in:a 2004?b 2010?
- 4 Which month had the lowest rainfall in: a 2003? b 2007?
- **5** Which year had the wettest:
 - **a** January? **b** June?
 - **c** December?
- 6 Which year had the driest:
 - **a** February? **b** May?
 - c November?
- 7 Considering winter to be the months June, July and August, which year had the:
 - **a** wettest winter?
 - **b** driest winter?

b lowest annual rainfall?

c November 2011?



Table 2 summarises the rainfall statistics for Sydney for all records kept from 1859 to 2011.

	J	F	Μ	Α	Μ	J	J	Α	S	0	Ν	D	Annual
Mean	102	118	130	127	121	131	98	80	69	77	84	78	1215
Lowest	6	3	8	1	3	4	2	0	2	1	2	3	583
Median	79	94	99	95	91	100	75	55	53	56	67	60	1160
Highest	387	631	521	622	585	643	336	483	356	285	517	402	2194

Table 2: Rainfall statistics for Sydney (mm)

- 8 Consider the annual rainfall statistics for Sydney shown in Table 2. What is the annual: **a** mean?
- **9** a For the month of September, what is the: i mean rainfall?
 - **b** Suggest why the 2000 Olympic Games were held in Sydney in the month of September.
- **10** On average, which month of the year is the: a wettest?
- **11** a What is the least amount of rain that has fallen in any month? **b** In which month did this occur?
- **12** a What is the greatest amount of rain that has fallen in any month? **b** In which month did this occur?
- **13** Which month has the smallest difference between the mean and the median rainfall?

Use tables 1 and 2 to answer questions 14 and 15.

- **14** For the period 2002–2011, in which years was the annual rainfall in Sydney greater than the long-term mean rainfall?
- **15** For the period 2002–2011, what percentage of years had rainfall that was less than the long-term median rainfall?

c Iraq?

Table 3 shows the country of birth of settler arrivals in Australia for the year July 2010 to June 2011.

- **16** How many settlers arrived from: a United Kingdom? b Vietnam?
- **17** The total number of settler arrivals in Australia. from more than 200 countries, was approximately 127 640 from July 2010 to June 2011. Determine the percentage of arrivals who came from:
 - **a** New Zealand **b** China c India
 - **d** South Africa e Sri Lanka.

18 How many more arrivals came from:

- **a** New Zealand than from Iraq?
- **b** the United Kingdom than from Sri Lanka?

Table 3: Country of birth of settler arrivals in Australia (July 2010 to June 2011)

Country of birth	Number of arrivals
New Zealand	25 772
China	14 611
United Kingdom	10 944
India	10 566
Philippines	5 048
South Africa	4 752
Vietnam	3 339
Sri Lanka	3 225
Iraq	2 988

b median?

b driest?

ii median rainfall?

Table 4 shows the top 15 countries of birth of Australian residents in 2006 and 2010 (excluding Australian born).

Table 4:	Country	of birth	of Australian	residents
----------	---------	----------	---------------	-----------

Country of birth	Estimated population 2006	Estimated population 2010
United Kingdom	1 153 000	1 193 000
New Zealand	477 000	544 000
China	203 000	380 000
India	154 000	341 000
Italy	220 000	216 000
Vietnam	180 000	211 000
Philippines	136 000	177 000
South Africa	119 000	156 000
Malaysia	104 000	136 000
Germany	115 000	129 000
Greece	126 000	127 000
South Korea	49 000	100 000
Sri Lanka	71 000	92 <mark>000</mark>
Lebanon	87 000	90 000
Hong Kong	76 000	90 000



19	How many people v	vho were resident Australi	an	in 2010 were l	oorn in	:		
	a New Zealand?	b the Philippines?	c	South Africa?	d	Vietnam?	e	Greece?

- 20 What was the country of birth of the highest proportion of residents in:a 2010?b 2006?
- 21 What was the country of birth of the lowest proportion of residents in:a 2010?b 2006?
- **22** There was a decrease in the resident population from 2006 to 2010 of people born in which country?
- 23 In 2010 approximately 5 994 000 (≈27%) of the Australian population were born outside Australia. What percentage of those born outside Australia came from:
 - a China? b India? c South Korea? d Sri Lanka? e Lebanon?
- 24 In 2010 the resident population of Australia was approximately 22 370 000. What percentage of all Australian residents were born in:
 - a China? b India? c South Korea? d Sri Lanka? e Lebanon?
- 25 From 2006 to 2010, what was the percentage increase in the resident population of people born in:
 a China?
 b India?
 c South Korea?
 d Sri Lanka?
 e Lebanon?

B The shape of displays

As the number of scores in a sample increases, the overall shape of the frequency histogram changes.



The general shape of a distribution can provide information about the scores. Here is the graph of a symmetric distribution, also referred to as a bell-shaped curve or a normal distribution.



If the distribution is not symmetric then it is said to be skewed.

A distribution is **positively skewed** if most of the data is on the left-hand side of the distribution. The data has a 'tail' to the right as shown in the diagram on the left below.

A distribution is **negatively skewed** if most of the data is on the right-hand side of the distribution. The data has a 'tail' to the left as in the diagram on the right below.



The mode is the score with the highest frequency. The mode for the normal and skewed distributions above are shown below.



Some distributions have two modes. This is called a **bimodal** distribution. As long as the distribution has two distinct humps, not necessarily with the same frequency (height), then it is said to be bimodal. Two examples are shown below.





Exercise 3B

1 Describe the shape of the following distributions as symmetric, positively skewed, negatively skewed or bimodal.



EXAMPLE 1 CONTINUED Think Solve Apply Negatively skewed Turn the stem-and-leaf Turn the stem-and-leaf plot on С plot on its side and draw a its side so that the stems are in Leaf smooth curve to fit the tops increasing order, from left to right. 6 of the columns. Draw a smooth curve through the 5 9 The tail is to the left. tops of the columns. Determine 2 5 7 5 from the shape of the distribution 1 2 8 5 4 1 9 4 1 if it is symmetric, positively 8 3 2 4 0 1 skewed, negatively skewed or Stem 0 1 2 3 4 5 bimodal.

2 Describe the shape of each distribution as symmetric, positively skewed, negatively skewed or bimodal.



- **3** a Construct a frequency histogram and polygon for the data in the frequency distribution table below.
 - **b** Describe the shape of the distribution as positively skewed, negatively skewed or bimodal.

Score	Frequency
6	2
7	2
8	3
9	5
10	8
11	6



EXAMPLE 2

Describe the shape of both data sets in the back-to-back stem-and-leaf	Scores on topic test			
plot as symmetric, positively skewed, negatively skewed or bimodal.	Females	Stem	Males	
	9	1	0 5	
	3	2	3 4 4 7	
	54	3	189	

	Solve	Think	Apply
<i>Male scores</i> Bimodal	Leaf 7 9 4 9 8 5 4 8 7 4 0 3 1 2 3 Stem 1 2 3 4 5	For males, turn the plot on its side and draw a curve to fit the columns. For females, flip the data over the stem	Turn each stem-and- leaf plot on its side so that the stems are in increasing order, from left to right. Draw a smooth curve through
Female scores Negatively skewed	Leaf 9 7 8 3 8 5 3 7 9 3 4 2 6 2 Stem 1 2 3 4 5 6 2	and turn it on its side. Draw a curve to fit the columns.	the tops of the columns and describe the shape of the distribution.

4 Describe the shape of each data set in the following distributions as symmetric, positively skewed, negatively skewed or bimodal.

b



5 a This distribution shows the scores for a group of students on a mathematics topic pre-test and post-test. This is called a side-by-side histogram. Describe the data distribution.



Scores on Mathematics test						
9 Blue Stem		9 Red				
998	4	7				
87211	5	3 4 7 8				
432	6	1 5				
54	7	055689				
3 0	8	3 4				
1	9	2				

7332

4 2

98876

4

4

6

2789

3 4

b This distribution shows the weight (in kg) of players in two Rugby League teams. This is called a **back-to-back** histogram. Describe the data distribution for each team.



a

C Comparing like sets of numerical data

EXAMPLE 1

The	times taken for the students from two classes to tra	vel to and from school are given	below.
Cla	ss 9P: 19 49 25 25 22 55 26 38 54 2	2 33 44 15	
	86 31 18 33 67 34 42 49 29 4	5 65 29	
Cla	ss 9W: 22 34 48 18 58 67 74 66 53 3	1 57 25 58	
	49 35 47 50 65 54 49 38 23 5	8 19 42	
a	Draw a back-to-back stem-and-leaf plot for this data	l.	
b	Comment on the shape of each distribution.		
C	Use the mean, median and range to compare the dat	a.	
	Solve	Think	Apply
a	Class 9P Stem Class 9W		Look for the tail to
			determine the shape
	9 9 6 5 5 2 2 2 2 3 5		of the distribution.
	8 4 3 3 1 3 1 4 5 8		The greater the
	99542 4 27899		mean and median,
	5 4 5 0 3 4 7 8 8 8		the longer the travel
			times. The greater
	$\left(\begin{array}{c}7\\8\end{array}\right)$		the range, the
			greater the spread
b	The distribution of times for Class 9P is	For Class 9P the tail is to	of times. The mean
	positively skewed; that is, there are more shorter	the right. For Class 9W the	and range are
	than longer times.	tail is to the left.	greatly influenced
	The distribution of times for Class 9W is		by outliers, so it
	negatively skewed; that is, there are more longer		is often useful to
	than shorter times.		ignore them.
C	For Class QP : Mean $-\frac{955}{2} - 382$	Mean sum of scores	
Č	$101 \text{ cm}^{-1}\text{ cm}^{-1} cm$	number of scores	
	Median = 33, Range = 71	$Median = middle \ score$	
	For Class 9W: Mean $=\frac{1140}{25}=45.6$	Range = highest score $-$	
	Median = 49 Range = 56	lowest score	
	<i>Comment:</i> As expected from the distribution		
	shape, the mean and median of Class 9W are		
	greater than those of Class 9P. In general, the		
	students of Class 9W taker longer than the		
	students of Class 9P to get to and from school.		
	The range of times for Class 9P is much larger		
	than for Class 9W, indicating a greater spread of		
	times for Class 9P. But, if we ignore the outlier,		
	86, the range of Class 9P becomes 52, which is		
	similar to that of Class 9W.		

Exercise 3C

1 Two groups of Year 9 students were asked to unscramble a seven-letter word. Their times, in seconds, are shown below.

 Group 1: 11
 16
 39
 23
 51
 24
 31
 4
 29
 16
 27
 40
 13
 23
 30
 29
 6
 22
 34
 38
 13

 Group 2: 12
 27
 46
 17
 26
 32
 18
 15
 21
 41
 37
 36
 23
 8
 25
 43
 34
 7
 36
 12
 7

- **a** Draw a back-to-back stem-and-leaf plot for this data.
- **b** Comment on the shape of each distribution.
- **c** Use the mean, median and range to compare the data.
- **2** The scores for a class of 16 students on two tests are given below.
 - Test 1: 22 42 34 30 19 39 46 41 38 35 47 39 24 45 27 32
 - Test 2: 13 18 21 6 40 16 26 24 35 12 20 26 31 13 15 19
 - **a** Draw a back-to-back stem-and-leaf plot for this data.
 - **b** Comment on the shape of each distribution.
 - **c** Use the mean, median and range to compare the data.
 - d Suggest a possible reason for the skewness.
- 3 a The table shows the January mean daily maximum temperatures for Sydney and Melbourne over 20 years. Draw a back-to-back stem-and-leaf plot for this data using stems 22 23 24 25 26 27 28 29.
 - **b** Comment on the shape of each distribution.
 - c Use the mean, median and range to compare the data.

	Year	Sydney	Melbourne
	1	26.7	26.9
	2	25.4	28.6
	3	26.2	23.9
	4	25.0	25.0
	5	27.2	25.3
100	6	26.2	24.0
5	7	26.9	28.0
	8	26.9	24.3
1	9	26.3	29.5
	10	26.4	27.4
(11	26.4	24.2
	12	25.4	23.9
	13	26.8	23.8
	14	25.9	23.1
	15	27.0	24.4
0	16	26.8	28.1
No.	17	25.4	25.7
1	18	25.5	26.0
1000	19	28.9	26.0
1	20	25.3	22.4

EXAMPLE 2

The scores of two groups of university students on a mechanical aptitude test are given at right.

- a Draw a back-to-back histogram and a side-by-side histogram for this data.
- **b** Comment on the shape of each distribution.
- c Compare the mean, median and range for each distribution.
- **d** Suggest a possible reason for the differences in these distributions.





2: Engineering students					
Score	Frequency				
4	1				
5	2				
6	3				
7	3				
8	4				
9	2				



STATISTICS & PROBABILITY

4 The scores of female and male students on a class English test are shown in the tables.

remaies					
Score	Frequency				
4	1				
5	1				
6	1				
7	6				
8	4				
9	2				

Famalas

Males					
Score	Frequency				
3	2				
4	4				
5	2				
6	2				
7	4				
8	1				

- **a** For this data, draw:
 - i a back-to-back histogram

ii a side-by-side histogram.

- **b** Describe the shape of each distribution.
- **c** Compare the mean, median and range for each distribution.
- **5** The 25 students in a Year 9 class were given a test in term 1 and a test in term 2. The results are given below.

						_							
Term 1:	8	9	10	10	10	9	8	7	9	8	10	9	10
	9	8	9	3	4	9	9	8	5	5	8	9	
Term 2:	2	3	1	0	3	2	5	7	8	9	8	7	6
	3	4	4	2	3	4	4	0	1	1	2	3	

a Complete the following tables and draw a back-to-back histogram for this data.

Tei	rm 1		Tei	rm 2
Mark	Frequency		Mark	Frequency
3			0	
4			1	
5			2	
6			3	
7			4	
8			5	
9			6	
10			7	
Degewike the she	no of each distrib	ution	8	

- **b** Describe the shape of each distribution.
- c Suggest a possible reason for the skewness.
- **d** Use the mean, median and range to compare the data.
- **6** The parallel dot plot shows the number of goals scored by two soccer teams in a 16 match competition.
 - **a** Describe the shape of each distribution.
 - **b** Use the mean, median and range to compare the data.



9

- 7 This back-to-back dot plot shows the number of saves
 - made by goalkeepers for the two teams in question 6.
 - a Comment on the shape of each distribution.
 - **b** Use the mean, median and range to compare the data.

Team A			N of	umb f goa saveo	er ls 1		Геа	n B	į		
			•	F	9	-					
			٠	-	8	-	•	٠	•	•	•
		٠	٠	-	7	-	٠	٠	٠	٠	
	• •	٠	٠	F	6	-	٠	٠	٠	٠	
		٠	٠	F	5	-	٠	٠			
	• •	٠	٠	F	4	-	٠				
		٠	٠	F	3	-					
				I		1					

D Statistical claims in the media

Exercise 3D

- 1 Write a short evaluation or analysis of the following statistical statements. Is there anything misleading about each of them?
 - a Four out of five dentists recommend Britex toothpaste.
 - **b** A newspaper reported that 'The Citico company has dismissed 10% of its workforce'.
 - **c** Twenty people at a netball game were surveyed and 70% said that they were going to vote for Mrs Brown at the coming state election. (Mrs Brown supports spending more money on sport facilities.)
 - **d** Of the two major political parties, one claimed that unemployment was rising and the other claimed that employment was rising. Could both parties be correct?
- 2 Consider the following two questions. What sort of responses are they likely to get?'Do you think that the taxes we pay should be used to pay people who stay at home and don't work?''Do you think that the government should give financial assistance to people who can't find employment?'

Investigation 1 Statistical claims made in the media

Collect various reports and statistical claims made in the media.

- 1 Investigate and interpret advertising that quotes various statistics.
- 2 Analyse graphical displays to determine features that may have been manipulated to cause a misleading interpretation, or which support a particular point of view.
- **3** Critically review claims linked to data displays.
- 4 Consider the reliability of conclusions from statistical investigations, including factors that may have masked the result, the accuracy of measurements taken and whether the results can be generalised in other situations.



Language in mathematics

- **1** Describe in words what is meant by a data distribution that is:
 - a symmetric b positively skewed

c negatively skewed

d bimodal

- 2 Read the following article about Hanna Neumann. Answer the questions in complete sentences.
 - **a** Where and when was Hanna Neumann born?
 - **b** What were her main areas of interest?
 - **c** For what is she remembered?
 - d Construct a timeline of her life.

Hanna Neumann (1914–1971)

Hanna Neumann was born Hanna von Coemmerer in Berlin. She was the daughter of a historian with teaching qualifications, who was killed in World War 1. As a result, her family was very poor and from the age of 13 she helped support them by tutoring younger children. Hanna became an extremely capable student and commenced studies at the University of Berlin in 1932. Her main area of study was mathematics, but she also had an interest in physics, history and religion.

Here she met her husband Bernhard, who left for England, in 1933, after deciding that living in Germany under the Nazis had become too dangerous. Hanna secretly travelled to England in 1934 to become engaged and then returned to continue her study in Berlin. Eventually, in 1938 after completing her studies and working as a research student at Gottingen University, she travelled to England and married Bernhard.

Hanna worked extremely hard and was recognised as being an excellent teacher. In 1948 she started opening her house in the evenings for others to come and discuss mathematics. She continued to be involved in teaching and studying mathematics and, in 1964, joined Bernhard at the Australian National University in Canberra.

In her position at the university, Hanna was recognised as having an enormous capacity for work and a great concern for those she taught. She held a number of administrative positions within professional mathematical bodies and travelled giving lectures. Unfortunately, at the age of 57, on a lecture tour in Canada, she became ill suddenly and died. She is remembered not only for her mathematical ability, but also for the willingness with which she was prepared to devote her time to teaching and her students.

Terms				
back-to-back	bimodal	distribution	histogram	mean
median	mode	negatively skewed	parallel	parallel dot plot
positively skewed	range	side-by-side	stem-and-leaf plot	symmetric

Check your skills

Use the table on the next page showing the top 10 countries of residence of visitors to Australia for 2010 and 2011 to answer questions **1** to **3**.

1	The country from which n	nost visitors came to Austra	ilia in 2011 was:	
	A New Zealand	B China	C Japan	D United Kingdom
2	From which country was t A New Zealand	the increase in visitors from B China	C Japan	D United Kingdom
3	What percentage of visitor	rs came from Malaysia in 2	010?	
	A 23.6%	B 5.6%	C 25.6%	D 3.6%

Country of residence	Visitors 2010 ('000)	Visitors 2011 ('000)
United Kingdom	646.7	608.3
USA	472.2	456.2
Malaysia	236.9	241.2
Singapore	308.0	318.5
China	453.8	542.0
Hong Kong	163.9	166.3
Japan	398.1	332.7
Korea	214.0	198.0
New Zealand	1161.8	1172.7
Germany	160.1	153.9

Top 10 countries of residence of visitors to Australia for 2010 and 2011

4 The distribution shown by the dot plot is:

A positively skewed **B** negatively skewed

D symmetric

C bimodal



5 Which of the following statements about the shape of the distributions in this side-by-side histogram is correct?

- A Group 1 is symmetric and group 2 is positively skewed.
 B Group 1 is negatively skewed and group 2 is negatively skewed.
- C Group 1 is symmetric and group 2 is negatively skewed.
- **D** Group 1 is negatively skewed and group 2 is positively skewed.

Questions 6 to 10 refer to the back-to-back stem-and-leaf plot showing runs scored by Michael and Ricky.

6	Ricky's lowest score is:			Ru	ns score	d
	A 0	B 9		Michael	Stem	Ricky
	C 19	D 91	-		0	0
				952	1	9
7	Michael's highest score is:			98776	2	38
	A 50	B 42		652	3	78899
	C 24	D 91		2 1	4	1 3 6
					5	2
8	Michael's median score is	:				
	A 82	B 29	C 28		D 27	
9	The mean of Ricky's score	es (to 1 decimal place) is:				
	A 5.6	B 6.1	C 34.1		D 36.9	

- **10** Which of the following statements is true?
 - A Ricky's distribution of scores is slightly positively skewed.
 - **B** Michael's distribution of scores is slightly negatively skewed.
 - **C** The range of Ricky's scores is greater than the range of Michael's scores.
 - **D** All of the above.



If you have any difficulty with these questions, refer to the examples and questions in the sections listed in the table.

Question	1–3	4, 5	6–10
Section	А	В	С

3A Review set

1 The table compares the population density of Australia with that of the 10 most densely populated countries in the world (with a population over 10 million).

Country	Population	Land area (km ²)	Density (population/km ²)
Australia	22 719 000	7 691 951	2.95
Bangladesh	152 518 000	147 750	1034
Taiwan	22 955 000	36 190	634
South Korea	48 456 000	99 538	487
Rwanda	10 718 000	26 338	407
England	53 013 000	130 440	406
Netherlands	16 760 000	41 526	404
India	1 210 193 000	3 287 263	368
Belgium	11 007 000	30 528	361
Japan	127 960 000	377 944	339
Sri Lanka	20 653 000	65 610	315

- a Which country has the greatest:i population?ii land area?iii population density?
- b Which country has the smallest:i population?ii land area?iii population density?
- c Which countries have a population less than Australia's population?
- d Why is Australia's population density so much smaller than that of the other countries listed here?
- 2 Describe the shape of these distributions as symmetric, positively skewed, negatively skewed or bimodal.





- 3 The heights, in centimetres, of the boys and girls in a Year 9 class are shown in the back-to-back stem-and-leaf plot.
 - **a** Comment on the shape of each distribution.
 - **b** Compare the mean, median and range of the heights.

rieignis recorded									
Boys	Stem	Girls							
	14	9							
64	15	267							
95532	16	0 1 2 3 4 7 8 9							
9875410	17	1 2 3							
3	18								

Unights recorded

a

f

Review set

3B

State or territory	Land area (km ²)	Population	Density (population/km²)	% of population in capital city
ACT	2 358	344 200	146.0	99.6%
NSW	800 642	6 967 200	8.7	63%
Vic	227 416	5 297 600	23.3	71%
Qld	1 730 648	4 279 400	2.5	46%
SA	983 482	1 601 800	1.6	73.5%
WA	2 529 875	2 163 200	0.9	73.4%
Tas	68 401	498 200	7.3	41%
NT	1 349 129	219 900	0.2	54%

The table shows estimates of the population density of Australia's states and territories. 1

- **a** Which state or territory has the greatest:
 - i land area?
 - iii population density?
- **b** Which state or territory has the smallest:
 - i land area?
 - iii population density?

- **ii** population?
- iv percentage of population in its capital city?
- **ii** population?
- iv percentage of population in its capital city?
- c Why does the ACT have such a high population density compared with the other states and territories?
- 2 State whether the shape of the following distributions are symmetric, positively skewed, negatively skewed or bimodal.



3 The numbers of goals per match scored by two soccer teams are shown in the tables below. Team B Team A

Number of goals	Number of matches
0	4
1	7
2	3
3	0
4	1



Number of goals	Number of matches
0	2
1	4
2	5
3	3
4	1

a Display this information in:

i a back-to-back histogram

- ii a parallel dot plot.
- **b** Comment on the shape of each distribution.
- Compare the mean, median and range for these two distributions. С
- **d** Which team do you think performed better?

3C Review set

- 1 The following graphs show the results of a survey of the skill levels in literacy and numeracy of 15-74 year olds in Australia in 2011–12. Level 1 is the lowest.
 - a What proportion of those surveyed was assessed at level 2 or less in:

Proportion at each literacy level

- i literacy?
- ii numeracy?
- **b** What proportion was assessed at level 4/5 in:
 - i literacy?
 - ii numeracy?







- c What proportion was in level 3 literacy and:
- ii unemployed? iii out of the labour force? i employed? **d** Which group had the highest proportion of people assessed at level 1 for literacy?
- What proportion was in level 3 numeracy and: e
 - i employed? ii unemployed? iii out of the labour force?
- f Which group had the highest proportion of people assessed at level 1 for numeracy?

2 Draw a neat sketch of a:

- a histogram of a positively skewed distribution
- **b** dot plot of a negatively skewed distribution
- stem-and-leaf plot of a symmetric distribution С
- **d** frequency curve with bimodal distribution.
- **3** A Year 9 class was given a Geography test. The results are given for girls and boys.

Girls:	12	24	26	41	38	9	17	15	20	36	29	28	32	28	9

- 43 31 26 32 32 17 51 49 18 24 49 35 48 46 40 Boys:
- a Draw a back-to-back stem-and-leaf plot for this data.
- b Comment on the shape of each distribution.
- Use the mean, median and range to compare the data. С