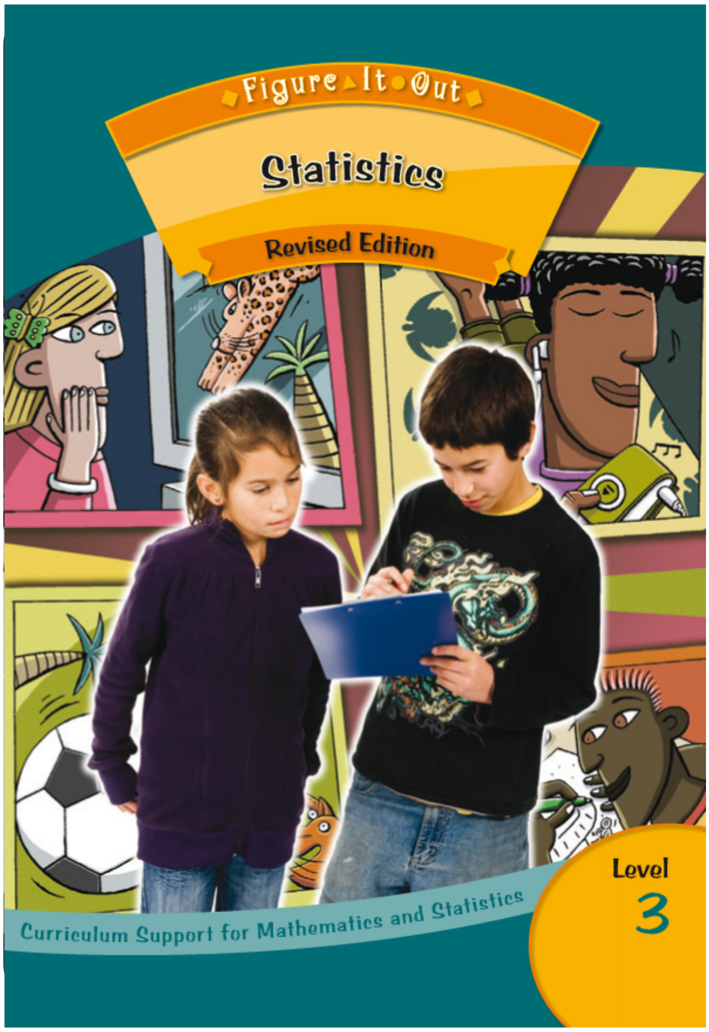
Answers and Teachers' Notes



MINISTRY OF EDUCATION
Te Tāhuhu o te Mātauranga

Contents

Introduction

The books in the Figure It Out series are issued by the Ministry of Education to provide support material for use in New Zealand classrooms. In recent years, much of the Figure It Out student material has been aligned with Numeracy Development Project strategies, which are reflected in the *Answers and Teachers' Notes* where appropriate.

The level 3 Figure It Out *Statistics* book that these *Answers and Teachers' Notes* relate to is an extensive revision, in line with the achievement objectives of the mathematics and statistics learning area of *The New Zealand Curriculum*, of the level 3 *Statistics* book published in 2000.

Student books

The activities in the Figure It Out student books are written for New Zealand students and are set in meaningful contexts, including real-life and imaginary scenarios. The level 3 contexts reflect the ethnic and cultural diversity and life experiences that are meaningful to students in year 5. However, teachers should use their judgment as to whether to use the level 3 book with older or younger students who are also working at this level.

The activities can be used as the focus for teacher-led lessons, for students working in groups, or for independent activities. You can also use the activities to fill knowledge gaps (hot spots), to reinforce knowledge that has just been taught, to help students develop mental strategies, or to provide further opportunities for students moving between strategy stages of the Number Framework.

Answers and Teachers' Notes

The Answers section of the *Answers and Teachers' Notes* for the revised *Statistics* book includes full answers and explanatory notes. Students can use this section for self-marking, or you can use it for teacher-directed marking. The teachers' notes for each activity, game, or investigation include relevant achievement objectives, comments on mathematical ideas, processes, and principles, and suggestions on teaching approaches. The *Answers and Teachers' Notes* are also available on Te Kete Ipurangi (TKI) at www.tki.org.nz/r/maths/curriculum/figure/

Using Figure It Out in the classroom

Where applicable, each page of the students' book starts with a list of equipment that the students will need in order to do the activities. Encourage them to be responsible for collecting the equipment they need and returning it at the end of the session.

Many of the activities suggest different ways of recording the solution to the problem. Encourage your students to write down as much as they can about how they did investigations or found solutions, including drawing diagrams. Discussion and oral presentation of answers is encouraged in many activities, and you may wish to ask the students to do this even where the suggested instruction is to write down the answer.

Students will have various ways of solving problems or presenting the process they have used and the solution. You should acknowledge successful ways of solving questions or problems, and where more effective or efficient processes can be used, encourage the students to consider other ways of solving a particular problem.

Figure It Out

Statistics

Level 3: Revised Edition

Answers

Page 1: Wild about Juice

Activity

1. a. Discussion will vary. Possible comments include:

   i. Based on the data, Amiri is correct that overall, Orange is the least popular juice (although not on Monday). It may be, however, that the canteen ran out of Orange on Monday and that this is the reason why there were no sales on any of the other days in the week.

   ii. There were 41 sales on Thursday and 21 on Friday, so this is a fair statement.

   iii. Sales were 28, 29, 31, 41, 21, which balances out to about 30 a day. However, Nina would have been more accurate if she had said, "On average, they sold about 30 a day."

   iv. Luke's statement is true. Only one other drink, Tropical (on Thursday), matched the lowest sale of Orange & Mango (on Friday).

   v. There were more *sales* of Wild Berry (18) than of Pineapple (16) over the course of the week, but the same people may have bought those flavours on different days, so stating that "more people bought Wild Berry" may not be correct.

   vi. You would have to know the school roll and who actually purchased juice to know if the data supports Rua's statement.

   vii. With sales of 28, 29, and 31, the data does support Liam's statement.

   viii. Based on this week's figures (27 sales of Tropical and 57 sales of Orange & Mango), the data supports Rawiri's statement.

   b. i. Statements will vary. Some examples include:

   "Friday is the least popular day for buying juice."

   "Green Apple is just as popular as Tropical."

   "The canteen may have run out of Orange on Monday."

   ii. Discussion and conclusions will vary, but the conclusions must be able to be backed up from the data.

2. Discussion and reasons for the decisions will vary, but one possible purchase order could be:

Flavour	Boxes
Tropical	3
Wild Berry	2
Orange & Mango	6
Pineapple	2
Green Apple	3
Orange	1

Any decision should be based on the relative popularity of each flavour for the week covered by the data. You will need to decide what to do about Orange. If you assume that the canteen ran out of Orange on Monday, you might suggest 3 cartons of Orange on the grounds that Orange and Green Apple were equally popular on Monday. Then again, maybe they ran out of Orange part-way through Monday and would have sold more than 5 …

Pages 2–3: Petty Differences

Activity One

1. Discussion will vary. Unnecessary information would include "my sister", the dogs that were dead now or had been given away, the sick bird. (The fact that the mice are female could be relevant if they turn out to be pregnant and increase the pet population!)

   Data that is difficult to interpret: You might wonder if the "cat in Aussie" really counts as a pet and whether to include the "2 goldfish in a bowl" as the owner doesn't seem to regard them as pets. You also need to decide whether to include the cat that is "my mum's".

2. **a.** A possible tally chart (taking into account the comments in question **1**) is:

	Number of pets	Frequency
Fish	𝍸 𝍸 𝍸 𝍸 IIII	24
Cat	𝍸 𝍸 𝍸 I	16
Dog	𝍸 𝍸IIII	14
None	𝍸 I	6
Bird	𝍸	5
Mouse	𝍸 I	6
Horse	III	3
Turtle	II	2
Axolotl	I	1
Unsure	III	3

   b. Fish (because this category has the most tally marks [frequency])

3. **a.** (Taking into account the comments in question **1**)

	Number of people	Frequency
Cat	𝍸 II	7
Dog	𝍸 II	7
None	𝍸 I	6
Bird	III	3
Fish	III	3
Mice	II	2
Turtle	I	1
Horse	I	1
Axolotl	I	1
Unsure	III	3

   b. Cats (because this category has the most tally marks [frequency])

4. Comments will vary. The first tally chart tells Wiremu how many different pets there are. The second tally chart tells him how many people have those pets – so if Wiremu wants to know how many people like certain sorts of pets (regardless of number), then the second tally chart is the better one. The first tally chart also creates a somewhat false picture by counting each fish as a separate pet.

Activity Two

1. **a.** 30 (on the day that the survey was conducted)

   b.–d. Practical activity

   e. How many people chose each particular kind of animal as their favourite pet

2. Choices will vary, but most people will prefer to use the bar graph because it shows at a glance how many people prefer each pet.

Pages 4–6: All in the Family

Activity One

1. **Table A**

Family	Boys	Girls	Total
1	4	1	5
2	0	1	1
3	2	1	3
4	1	1	2
5	0	2	2
6	2	2	4
7	0	4	4
8	1	1	2
9	1	2	3
10	0	1	1
11	2	1	3
12	0	2	2
13	2	1	3
14	1	1	2
15	3	1	4

16	0	2	2
17	1	1	2
18	1	2	3
19	2	0	2
20	2	0	2
21	1	0	1
22	2	1	3
23	1	1	2
24	2	1	3
25	1	2	3
26	0	2	2
27	1	1	2
28	2	2	4
29	2	1	3
30	1	0	1
31	0	2	2

Table B

Number of children per family	Number of families
1	4
2	13
3	9
4	4
5	1

2. **a.** The line graph because this data is discrete, not continuous. You can't, for example, have between 1 and 2 children in a family (for example, $1\frac{1}{2}$ children).

   b. The pie chart. It doesn't show the frequency (number of families).

3. **a.** Yes. 14 have 2 or more siblings, while 17 have 1 or no siblings. So you could say this is "about half".

   b. Either the bar graph or the pie chart. The relevant information is very clear in the bar graph, and you can easily total the frequencies. It is also very clear in the pie chart that families with 3 or more children account for just under half of the total.

4. Comments will vary, but should be based on a reasonable argument. The data collected does not include age, so some of the children that have been counted in the table may be regarded as "adults" and therefore not eligible for entry as "children" on a family pass. This may be what Evan had in mind.

Activity Two

1.

Girls \ Boys	0	1	2	3	4
0		II	II		
1	II	卌 I	卌 I	I	I
2	卌	III	II		
3					
4	I				

2. **i.** True. There are no tally marks in any of the columns for showing 3 girls.

   ii. False. There are 6 families with 1 boy and 1 girl and 2 families with 2 boys and 2 girls. That is 8 families in all.

   iii. True. There are 8 families with no boys and 4 families with no girls.

   iv. True. Of the 31 families, 16 have only 1 girl.

   v. False. 13 of the 31 families have just 2 children. This is less than half, not "most".

   vi. True. 10 families have 2 boys, and 10 families have 2 girls.

Activity Three

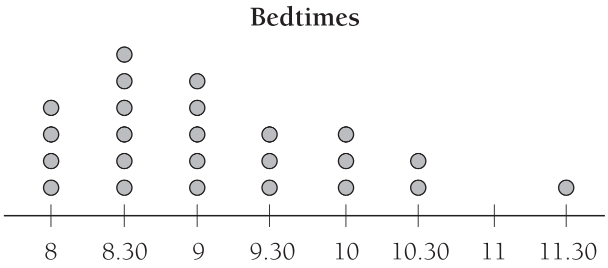
1.–3. Data and conclusions will vary.

Page 7: It's on the Cards

Activity

1. Questions will vary. For example, "I wonder if more boys than girls list playing sport as their favourite spare-time activity?" or "I wonder how many people go to bed at 9 p.m. or later?"

2. Practical activity. Results will vary. For example, a dot plot could be used to show bedtimes and answer the question “I wonder what time most people went to bed?” It might look like this:



This dot plot shows that 14 out of 24 people went to bed at 9 p.m. or later.

3. Practical activity. Questions and data will vary, but you need to give full statements and back them up with data and graphs.

Pages 8–9: Television Times

Activity One

1. 30 (unless someone was absent that day) because there are 30 “leaves” on the graph.

2. a. Ryan: 3 hours 30 minutes. Bridie: no time (0.00)

   b. Suggestions will vary. For example, Bridie’s family may not have a TV; there may have been nothing on that she wanted to watch; she might have been engrossed in an exciting book; she might have gone to the movies, been out with her family, been at swimming or basketball practice, or been involved in some other necessary or preferred activity.

3. i. True because, of the 30 students, only 9 watched 1 hour or less on Wednesday.

   ii. False because, apart from the fact that 5 students watched no more than 30 minutes of TV, the data relates to a single evening (a different picture might be revealed if data was collected over a week).

   iii. The statement is false if you take it to mean “the number who watched 30 minutes but no more than 30 minutes” (because only 3 students watched exactly 30 minutes of TV). The statement is true if you (reasonably) assume that everyone who watched 30 minutes or more of TV is to be included. (Note that the statement is also false if you thought because 12 boxes had 30 in them that this meant that 12 students watched TV for 30 minutes – that thinking ignores the “stem” part of the stem-and-leaf graph.)

   iv. True, if by “most” we mean “more than half”. 16 out of 30 is more than half.

   v. True: Ryan was the student. (He watched for 3 hours 30 minutes.)

   vi. True because all the students except for one watched TV on Wednesday.

   vii. False because 200 minutes is (3 × 60) + 20 or 3 hours 20 minutes. Only one student (Ryan) watched more than this amount.

4. Discussion will vary. To answer the question about their viewing, students had to try and remember what they did the previous evening. They may have found it difficult to remember. People may also minimise or exaggerate their viewing, perhaps because they are embarrassed to admit how much time they spent in front of the TV or they want to show off.

Activity Two

1.–2. Answers will vary.

3. a. Questions will vary. You might want to investigate what time your classmates start watching TV or what their favourite programmes are.

   b. Answers will vary. For example, you might need to ask your classmates to keep a log of their viewing for a week or to write down (in order) a list of their favourite programmes. (You don’t actually need to do this investigation, although you could!)

Pages 10–11: Well Weathered

Activity

1. a. The average daily temperature is higher in Kaitaia than in Tekapō. This is true for all months of the year. The difference is least in summer (about 3–4 degrees in December and January) and greatest in winter (about 10 degrees in July). The main reason for

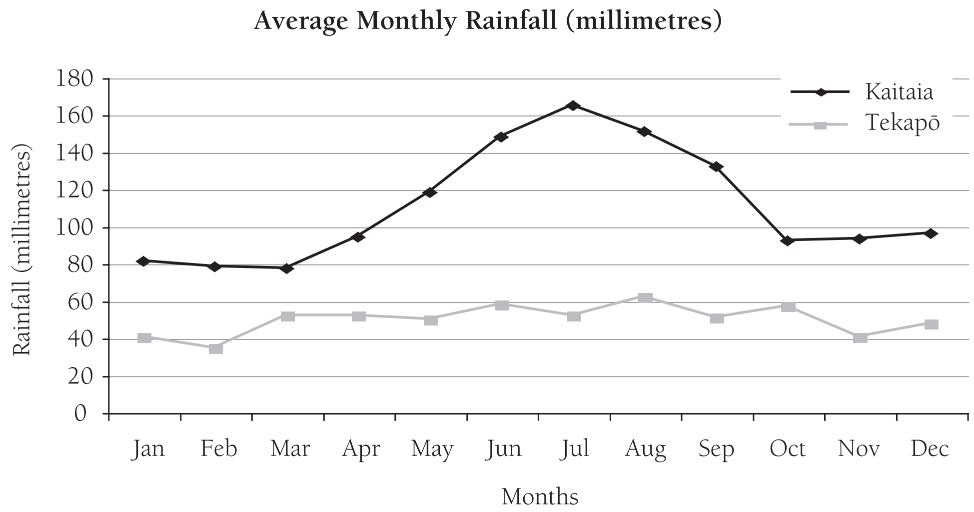
the difference is that Kaitaia is much further north (and therefore closer to the equator) than Tekapō.

b. In spring and summer (from September to February), Tekapō has markedly more sunshine than Kaitaia (as much as another 50 hours per month). The simple reason for this is that when rain-bearing clouds from the west (the prevailing wind for that season) meet the Southern Alps, they drop their rain, leaving the skies clear over Tekapō. In the depths of winter (June and July), it has markedly less (see explanation in **1a**). In autumn (March to May) and early spring (August), the hours are very nearly the same. See question **4** for related data.

2. The graphs do not show how much the temperature varies within a day. The range may be much greater in Tekapō than in Kaitaia. The graphs also do not show how consistent (predictable) temperature is throughout a month.

3. On average, Kaitaia is much warmer than Tekapō, but for all but 2 months, Tekapō has about the same or more sunshine hours than Kaitaia.

4. a. If you use a graphing program, your graph will look something like this. (You may have drawn a similar graph by hand.)



b. Kaitaia has a lot more rain than Tekapō. March, April, and October are the only months when Kaitaia's average rainfall is not at least twice that of Tekapō. In July, the average rainfall in Kaitaia is over three times that of Tekapō.

Investigation

Investigations will vary. Your paragraph for question **3** should refer closely to your graphs. For question **5**, if you agree with any suggestions from your classmate for improving your graphs, revise your displays accordingly.

Pages 12–13: Winning Dots

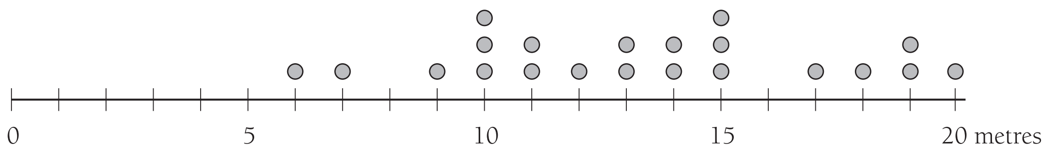
Activity One

1. a. Descriptions will vary. There is a cluster of three students whose throws were 10 to 11 metres. The great majority of the students threw the ball 14 to 19 metres. This fact shows up on the dot plot as a big cluster. The 3 metre and 7 metre throws are short throws. The 7 metre throw may have been from someone who genuinely couldn't throw very far. It is unlikely that someone could manage to throw a tennis ball only 3 metres, so this outlier may represent a mis-throw.

b. Most of the students in Room 1 can throw a reasonable distance. Only one or two are what you could call poor throwers.

2. a.

Tennis Ball Throw: Room 2



b. The main cluster is between 10 and 15 metres; a smaller cluster is between 17 and 20 metres. There are 2 short throws of 6 metres and 7 metres.

c. Fewer students in Room 2 are very good throwers, but Room 2's best throw is a metre further than Room 1's best throw.

3. i. True. The best throw is 20 metres, from Room 2.

ii. Although most Room 1 students (14 out of 19) threw 14 metres or more, fewer than half of Room 2's students (10 out of 21) did. Taking the two classes together, 24 out of 40 threw 14 metres or more. If by "most", "more than half" is meant, then the statement is true. If by "most", "considerably more than half" is meant, the statement is false.

iii. False. The range for Room 1 is 19 – 3 = 16 metres and for Room 2 is 20 – 6 = 14 metres.

iv. False. 17 students from Room 1 and 18 from Room 2 did throws of 10 metres or more.

v. True. Room 1 has 19 students and Room 2 has 21.

vi. False. 6 students from Room 1 and 5 from Room 2 did throws of at least 17 metres.

vii. True

viii. It depends on your definition of throwing "really well". If you define it as 16 metres or more, the statement is true (Room 1: 10 throws; Room 2: 5 throws).

4. Reasons and discussion will vary. You will need to have considered longest throws, your definition of "really good throws", and how the throws are clustered.

Activity Two

Data and findings will vary.

Pages 14–15: Old Enough?

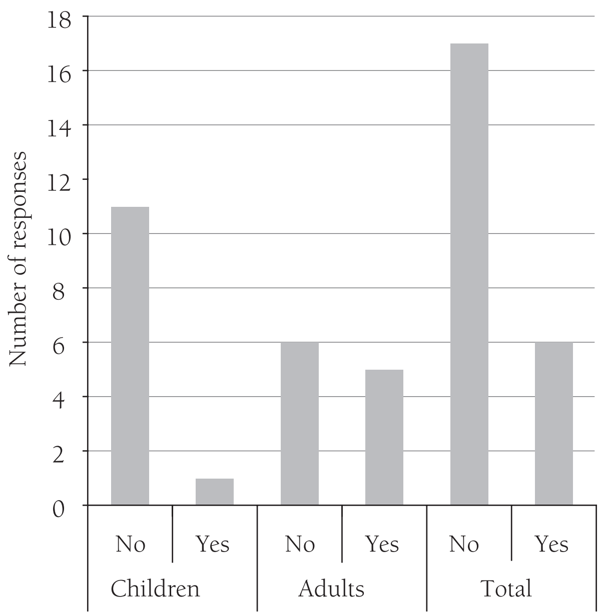
Activity

1. a.

Is 8 too young to receive pocket money?					
Children		Adults		Total	
No	Yes	No	Yes	No	Yes
11	1	6	5	17	6

b. Your bar graph should look similar to this:

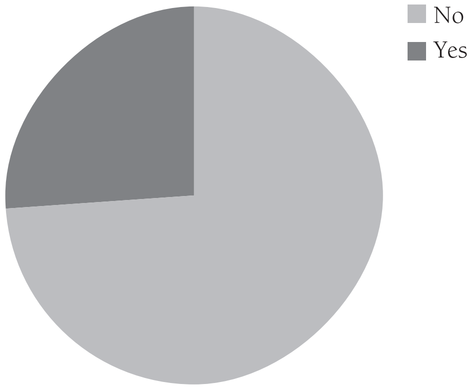
Is 8 too young to receive pocket money?



c. Statements will vary. For example, "Most children think 8 is not too young to receive pocket money. Just over half of the adults think 8 is not too young."

d. Your pie chart should look similar to this:

Is 8 too young to receive pocket money?



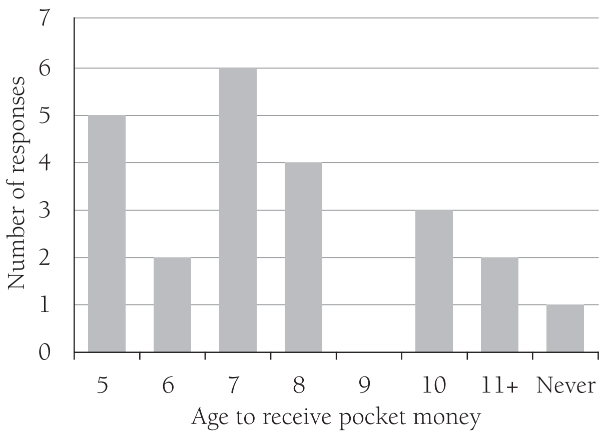
e. Comments will vary. The pie chart shows that there are a lot more "no" answers, but you can't tell the exact number (unless you put numbers around the outside of the pie).

2\. a.

At what age should children start receiving pocket money?								
Age	5	6	7	8	9	10	11+	Never
Votes	5	2	6	4	0	3	2	1

b. Your bar graph should look similar to this:

At what age should children start receiving pocket money?



3\. Discussion will vary. Choices may be:

i. The pie chart

ii. The bar graph in **2b**

iii. The bar graph in **1b**.

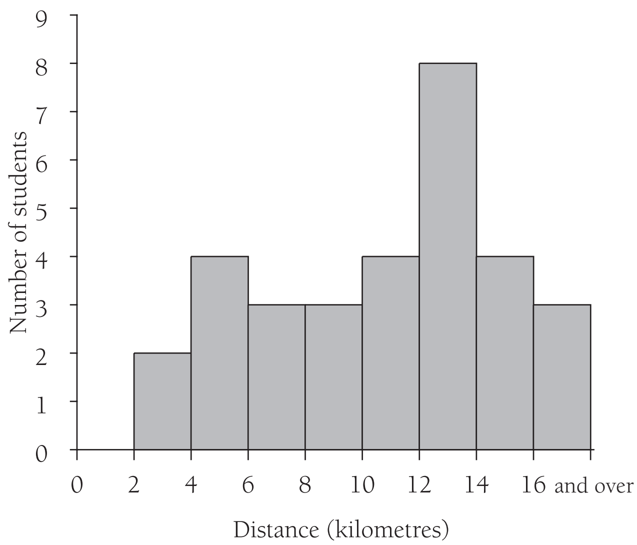
4\. Marika is quite correct in asking the question. Although she could have got that information from her second question (only the people who wrote down an age of more than 8 in response to Marika's second question answered her first question with a "yes"), asking the first question gives her clear yes/no responses that she can count up. Note that her data also includes the neighbour who didn't think that children should get pocket money at any age. She could choose to discount this response as an outlier.

Pages 16–17: Way to Go

Activity One

1\. Practical activity. Your graph should be similar to this:

Distance to School (Otago)



2\. a. All the Auckland students live less than 6 kilometres from school, whereas most of the Otago students live further than 6 kilometres from school.

b. Answers will vary, but city students tend to go to schools that are near their homes, while many students at rural schools live on farms, which can be a long way from school.

3\. Because the distance range for the rural school is much greater than that for the city school, you would end up with a very long histogram for the rural school if the same scale were used. The differences still show up very clearly when the scales are different.

Activity Two

Practical activity. Results and comparisons will vary.

Pages 18–19: Scratch 'n' Win

Activity

1. **a.–d.** Practical activity. Results will vary. But you will probably decide that it's not easy to get a win! If you tried a large number of scratch 'n' wins, you would find that about one-quarter would be "wins".

   e. It's not possible to specify a set number to guarantee a match. The result depends on probability, so any card can deliver, or not deliver, a win.

2. Sunita is right. With a 4-symbol card, the probability of choosing matching symbols is 1 in 4. With a 5-symbol card, the probability is 1 in 5.

Pages 20–21: Superbeans

Activity One

1. **a.**

		Coin A	
		H	T
Coin B	H	HH	TH
	T	HT	TT

   b. 4 outcomes (shaded)

   c. 1 out of 4 or one-quarter ($\frac{1}{4}$)

   d. 2 out of 4 or one-half ($\frac{2}{4} = \frac{1}{2}$)

Activity Two

1.–2. Practical activity. Results will vary. A possible statement for **2a** is: "The three Superbeans were all of a different colour in 6 of the 20 trials."

3. **a.–b.** Results and comments will vary. The bigger the data set, the more accurate the patterns should be. Ngaio might expect 2 of the same colour when she takes 3 Superbeans from her packet.

4. **a.** YYY, YYR, YRY, RYY, YYB, YBY, BYY, RRR, RRY, RYR, YRR, RRB, RBR, BRR, BBB, BBY, BYB, YBB, BBR, BRB, RBB, YRB, RBY, BRY, YBR, RYB, BYR

   b. **i.** 3. (YYY, BBB, RRR)

   ii. 6. (YRB, RBY, BRY, YBR, RYB, BYR)

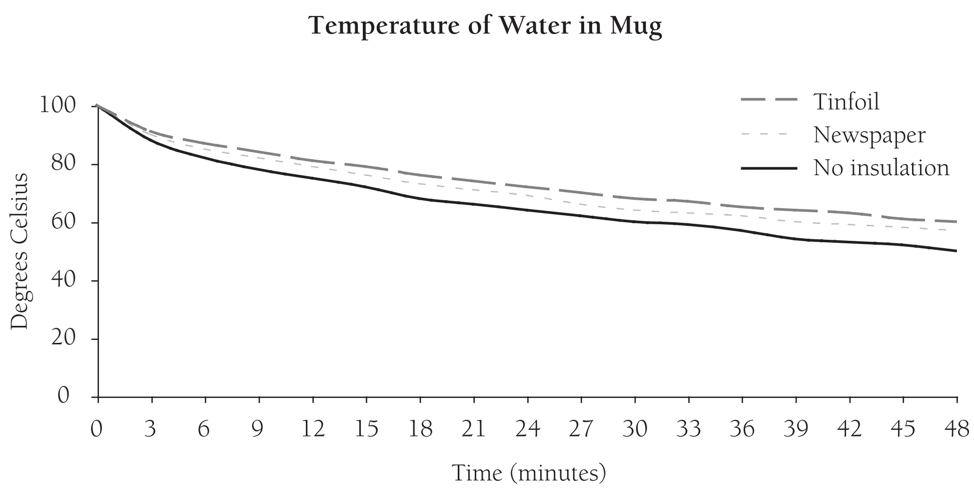
   iii. 18. (6 for yellow: YYR, YYB, BYY, RYY, YRY, YBY; and similar combinations for the other two colours)

   c. The "least likely" possibility is that all 3 Superbeans are the same colour (only 3 of the 27 outcomes). The "most likely" outcome is that two Superbeans will be of the same colour (18 out of 27 outcomes).

Pages 22–23: Cold Coffee

Activity One

1.



2. a. Descriptions will vary but should report that, over time, the water in all three mugs cooled, but more slowly in the mug wrapped in newspaper and even more slowly in the mug wrapped in tinfoil. (Alternatively, both the newspaper and the tinfoil helped keep the water hot longer, but the tinfoil was more effective than the newspaper.) Regardless of the mug, the temperature changed most quickly in the first few minutes and then more and more slowly. It could also be noted that, even after 48 minutes, there was only up to 10 degrees difference in the temperatures.

   b. Probably not. Even with no insulation, the temperature was still 72 degrees after 15 minutes. Most people would be able to drink their coffee in this time (unless they forget it's there!).

   c. Advice will vary. Craig may suggest that Ms O'Connor use an insulated cup, which would be at least as effective as tinfoil.

Activity Two

1.–2. Practical activity. Results will vary. Paragraphs need to relate to the graphs.

3. The temperature will continue to increase, but more and more slowly, until it reaches room temperature.

4. For example, the room temperature, how quickly you get the water into the mugs, having exactly the same volume of water in each mug, whether measurements are done at exactly the right time and in the same way, whether the thermometers used all read temperature accurately, whether the mugs experience the same environmental conditions (for example, one isn't in a draught or in full sun).

5. Discussion will vary. Workers who buy hot drinks from street vendors need to be sure that their drinks will still be hot when they get to work and while they drink them. Some people use insulated mugs with lids for maximum heat retention.

Page 24: Dicey Differences

Activity

1. In round 1, differences of 0, 1, or 2 occurred twice as often as differences of 3, 4, or 5 (in fact, 5 didn't occur at all!). In round 2, the differences of 0, 1, or 2 occurred three times as often as the differences of 3, 4, or 5. Over the two rounds, Maaka had 17 points to Whina's 7, so it is not surprising that Whina thinks the game is unfair.

2.–3. Practical activity. Results will vary, but they will almost certainly favour by a big margin the player who gets points for a difference of 0, 1, or 2.

4. Comments will vary, but it should be clear that the lower differences turn up much more frequently than the higher differences, with 1 being the most common difference of all.

   If you create a difference table like the one below, it will be clear that the probability of a difference of 5 is 2 in 36 (1 in 18), while the probability of a difference of 1 is 10 in 36 (5 in 18). In other words, the probability of getting a difference of 1 is 5 times greater than the probability of getting a difference of 5.

Difference Table

Dice 2

Dice 1	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

You could also shade the 3, 4, and 5 squares on your table and use these to discuss how much more likely it is to get a 0, 1, or 2 difference.

Figure It Out

Statistics

Revised Edition

Teachers' Notes

Overview of Level 3

Title	Focus	Page in students' book	Page in teachers' notes
Wild about Juice	Interpreting data and using it to make decisions	1	21
Petty Differences	Cleaning and collating data and communicating findings	2–3	22
All in the Family	Collating, displaying, and interpreting data	4–6	24
It's on the Cards	Planning and conducting a survey using data cards	7	26
Television Times	Exploring an issue and evaluating the findings	8–9	27
Well Weathered	Creating and interpreting time-series graphs	10–11	29
Winning Dots	Creating and interpreting dot plots	12–13	31
Old Enough?	Evaluating responses to a survey	14–15	32
Way to Go	Graphing continuous data and interpreting the results	16–17	34
Scratch 'n' Win	Exploring probability	18–19	36
Superbeans	Exploring probability experimentally and by listing outcomes	20–21	37
Cold Coffee	Evaluating experimental data with the help of time-series graphs	22–23	40
Dicey Differences	Investigating probabilities in a game of chance	24	41

What Is Statistics About?

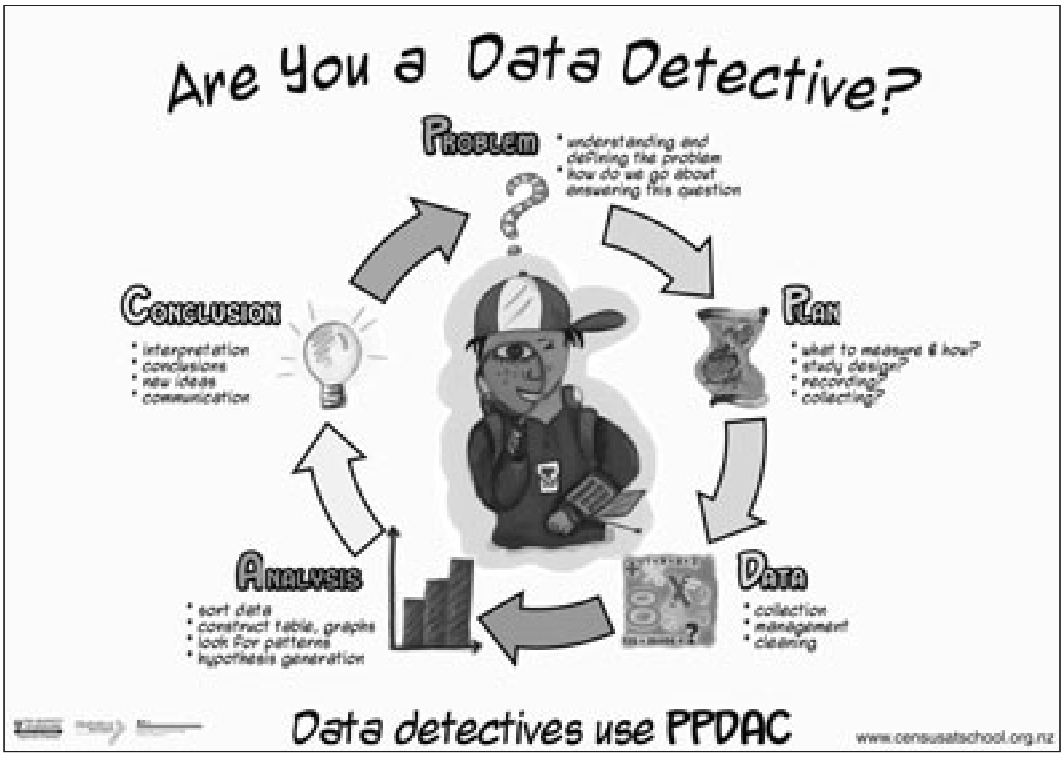
Statistics is defined in *The New Zealand Curriculum* as "the exploration and use of patterns and relationships in data". Like mathematics, it aims to equip students with "effective means for investigating, interpreting, explaining, and making sense of the world in which they live".

The New Zealand Curriculum goes on to say:

> Mathematicians and statisticians use symbols, graphs, and diagrams to help them find and communicate patterns and relationships, and they create models to represent both real-life and hypothetical situations. These situations are drawn from a wide range of social, cultural, scientific, technological, health, environmental, and economic contexts …
>
> Statistics involves identifying problems that can be explored by the use of appropriate data, designing investigations, collecting data, exploring and using patterns and relationships in data, solving problems, and communicating findings. Statistics also involves interpreting statistical information, evaluating data-based arguments, and dealing with uncertainty and variation.
>
> page 26

The **PPDAC** (Problem, Plan, Data, Analysis, Conclusion) statistical investigation cycle used for the CensusAtSchool New Zealand resources (see www.censusatschool.org.nz) provides an ideal model for statistical investigation. This approach is used in the revised level 3 Figure It Out *Statistics* book and in the *Answers and Teachers' Notes* that accompanies it.

CensusAtSchool New Zealand makes available two posters (aimed at different age levels) for the PPDAC cycle. One version is:



The five steps in this model are:

- Problem – deciding what to investigate, and why, and how to go about it;
- Planning – determining how to gather the necessary data;
- Data – collecting, managing, and preparing the data for analysis;
- Analysis – exploring the data with the help of graphs and statistical tools and asking what it says;
- Conclusion – determining how the data answers the original problem and deciding what to do next.

CensusAtSchool New Zealand provides this information in the form of a downloadable PDF.

Much of the information in the following sections is adapted (by permission) from information available on CensusAtSchool New Zealand and Statistics New Zealand (www.stats.govt.nz).

Types of Data

Category data classifies data according to a non-numeric attribute, such as gender, colour, style, model, opinion, type, feel, and so on. For example, foods could be sorted into categories such as meat, fish, vegetables, fruit, and cereal.

Numeric data classifies data according to an attribute that can be counted or measured. Numeric data may be either discrete or continuous. **Discrete data** is whole-number, countable data. **Continuous data** is data obtained using measurement, for example, time, height, area, mass, age. When continuous data is rounded to the nearest whole unit, it is effectively treated as discrete. **Time-series data** is data that is collected from a series of observations over time, with a view to discerning time-related trends.

A **variable** is an attribute or factor that can take on different values, for example, time, colour, length, favourite author, number of items, cost, age, temperature. When a number of pieces of data are collected for a single object or person, the result is a **bivariate** or **multivariate** data set. A list of movies by length is a **univariate** data set (that is, there is a single variable, time); a list of movies by length, genre, and country of origin is a multivariate data set. Multivariate data sets have much greater potential for exploration than univariate data sets.

Graphs

Graphs or charts? *Graph* is the more common usage in New Zealand (except in the case of pie chart) but *chart* is the term used by most graphing programs. In statistical contexts, these two terms are used virtually interchangeably.

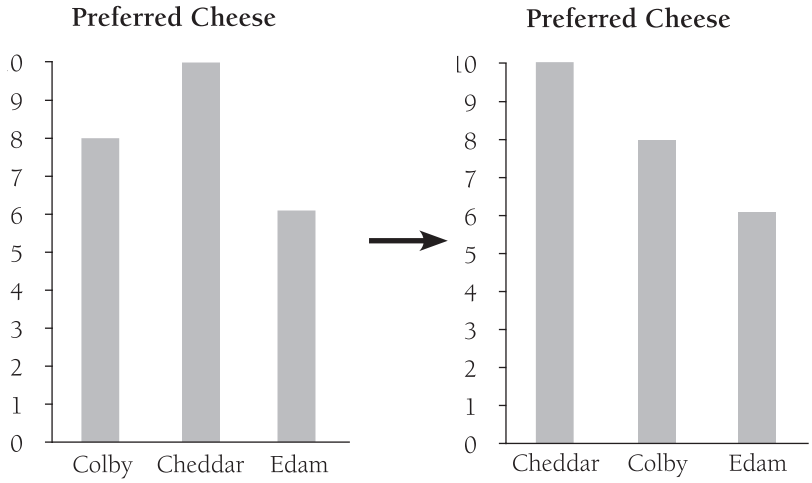
The activities in the students' book promote graphs as a means of exploring data and communicating findings. It is important that students learn to "read" graphs, question what they read or see, find the stories in the data, and ask more questions. Especially in the early years, students can devise their own graphical representations. As they learn more, they need to become familiar with the standard types of graphs and associated conventions. All standard graphs should have a title that states the intent, axes (if used) should be labelled clearly, and the measures used should be consistent throughout. These basic conventions are designed to enhance the communicative power of graphs.

Bar graphs are used to show the frequency of category data or discrete numeric data. Unlike dot plots and strip graphs, they have two axes, one labelled with the category and the other with the frequency. There is always a gap between bars, showing that the categories are quite separate.

The bars are normally vertical and, for category data, may be coloured or shaded differently.

On a well-constructed and labelled bar graph, it is easy to see which of the categories is most "popular" and to compare categories. Differences that appear insignificant in a pie chart or strip graph typically show up clearly in a bar graph.

Unless there is a good reason not to, the bars for **category** data are usually arranged in order of height, as in the example below.



In the students' book, bar graphs feature in the activities on pages **2–3**, **4–5**, and **14–15**.

Pie charts and **strip graphs** show the relative size of the categories that make up a whole (whatever the whole may be). The categories are always labelled. The percentage value (and sometimes the actual data value) may also be shown on or alongside each region. Unlike bar graphs, pie charts and strip graphs do not show categories that contain zero data. Students find pie charts difficult to create by hand but easy to create in most graphing programs.

Pie charts or strip graphs feature in several of the activities in the students' book, including those on pages **2–3**, **4–5**, and **14–15**. While they have their place, these graphs can only be used for a single variable and can only tell the simplest of stories.

Stem-and-leaf graphs, explored on pages **8–9** of the students' book, are a convenient means of organising and displaying discrete numeric data. Each individual data value retains its identity at the same time as overall patterns emerge.

A stem-and-leaf graph is made by arranging numeric data in a display, using the first part of the number as the stem and the last digit as the leaf. For example, for 16 in the data set {16, 31, 25, 33, 27, 24, 14, 26, 31}:

- the tens digit (1) is the stem and appears to the left of the vertical line
- the ones digit (6) is the leaf and appears to the right of the vertical line.

Stem	Leaf
1	4 6
2	4 5 6 7
3	1 1 3

It is usual to sort the leaf data in number order, from least to greatest. This is often best done as a second step, particularly where there is quite a lot of data.

When constructing stem-and-leaf graphs, students should ensure that they space the digits equally. This is important because it makes it possible to observe features such as general shape, symmetry, gaps, and clusters. It also makes it easy to track down the median (and, in a larger data set, the quartiles).

If the data collected is three-digit numbers, such as height in centimetres, the hundreds and tens digits make the stem of the graph, with the ones digits as the leaves. For example, if graphing the following heights recorded in centimetres, 114, 122, 142, 116, 125, 127, 142, 144, the stem represents the hundreds and tens digits:

Stem	Leaf		
11	4	6	
12	2	5	7
13			
14	2	2	4

In Scandinavian countries, stem-and-leaf graphs are used for bus timetables. For example:

Hour	Minute					
06	00	25	40			
07	05	23	37	48	55	
08	00	14	26	33	47	52
09	02	13	46			

You could discuss with your students the suitability of this system for New Zealand. What system is currently used for our bus timetables?

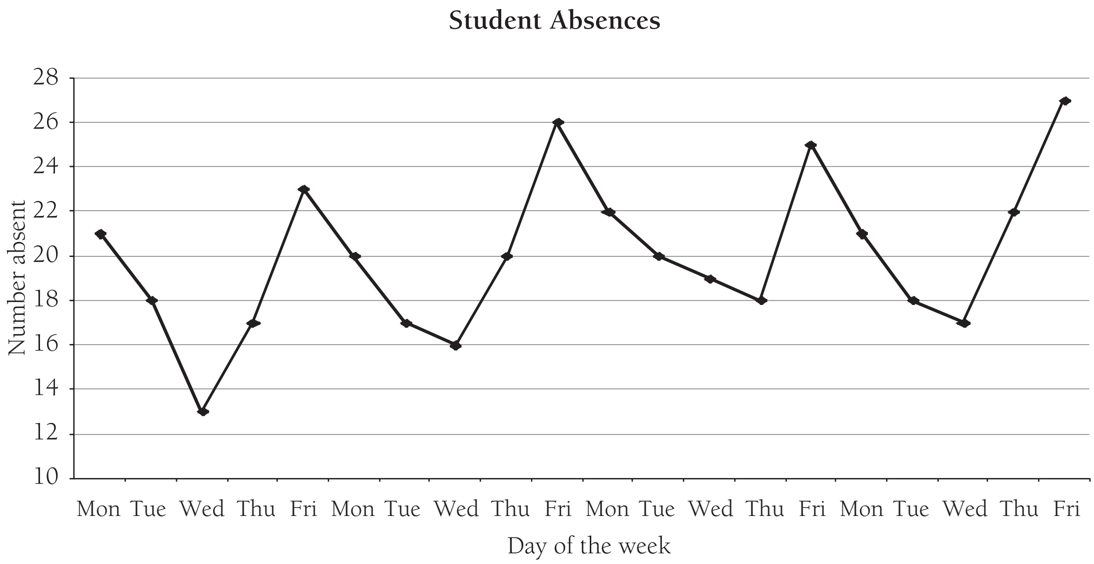
Two sets of numeric data can be displayed in a **back-to-back stem-and-leaf graph**, in which the stem is shared. This makes it possible to compare the shapes of the two distributions without losing the individual values. This graph is also useful for finding summary statistics such as the median and quartiles.

Girls' Arm Span (cm)						**Boys' Arm Span (cm)**		
				8	11			
					12	5	8	
			0	8	13			
			0	7	14	4	4	4
0	5	5	6	9	15	4		
		2	4	5	16	3	3	
				5	17	5		
					18	2		

Students should take care with the order of the leaves on back-to-back stem-and-leaf graphs; they must be consistent throughout.

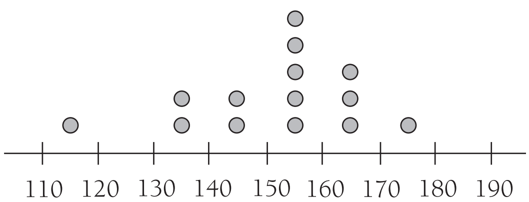
Line graphs, which appear on pages **4–5** and **10–11** of the students' book, compare two variables, one of which is plotted on the horizontal axis and the other on the vertical axis. Line graphs are useful for showing one variable in relation to another (for example, tree growth in relation to rainfall), or making predictions about the results of data that has not yet been decided or recorded (**extrapolations**), for example, visitor numbers to Abel Tasman National Park for the summer of 2012, based on previous visitor numbers for the same season and the general trend visible in several years' worth of data. Line graphs should not be used for category data.

Line graphs are excellent for showing how something changes over time. A **time-series graph** (see pages **10–11** in the students' book) is a line graph in which time is measured on the horizontal axis and the variable being observed is measured on the vertical axis. For example:

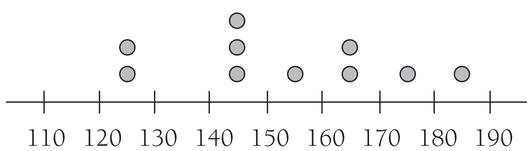


Dot plots, used on pages **12–13** and **24** of the students' book, are very easy to construct and clearly show the spread of the data involved (that is, the way in which it is distributed and/or grouped). They suit discrete numeric data: each dot represents a single piece of data. Continuous (measurement) data is normally rounded to the unit used on the scale (for example, the nearest centimetre). The beginning and end of the scale are dictated by the least and greatest data value.

Data can also be grouped, as in the following dot plots, which are similar to a histogram.



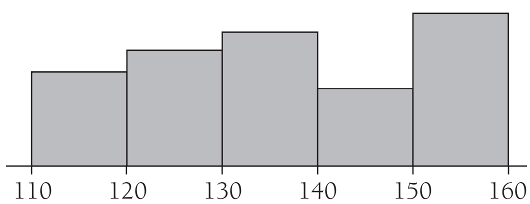
Girls' Arm Span Length (cm)



Boys' Arm Span Length (cm)

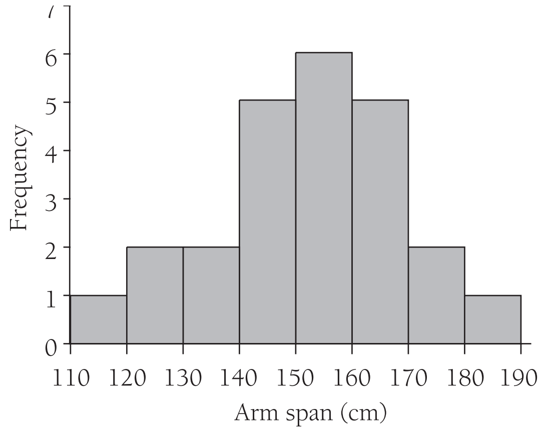
A **histogram** looks similar to a bar graph but, in this case, the bars touch. Histograms are used for continuous data (for example, height). Histograms are introduced on pages **16–17** of the students' book.

In a histogram, the horizontal axis is a continuous number line. Its start and finish are defined by the data (there is no point starting at 0 if the first data value is, for example, 127). The sides of the bars represent the beginning and end of an "interval". The bars are best labelled using the outer limits of each interval:



As a guide, 10 or 12 is a reasonable maximum for the number of bars. Any more and it would be difficult to see patterns. When deciding on a suitable interval, consider both the greatest and least values and the total number of values in the data set. Use "natural" steps, such as 2, 5, 10, or 20 (not 3, 6, 7, 8, 9, and so on), and keep them the same throughout. As an example, the arm span data in the stem-and-leaf example on page 16 ranges from 118 to 182 centimetres. There are 24 values in the data set. This suggests a histogram with eight intervals of 10 centimetres each:

Arm span (cm)	Frequency
110–	1
120–	2
130–	2
140–	5
150–	6
160–	5
170–	2
180–190	1



Other Statistical Terms

Axes (singular: **axis**) are the two lines, one horizontal and one vertical, that form the framework for most graphs. As a general principle, the vertical axis is used for frequency and the horizontal axis for categories, values, or time. This means that bars are equally spaced and vertical. (Pictographs and bar graphs do not always observe this rule.)

Collate means to collect and combine.

A **correlation** is said to exist between two variables (for example, smoking and heart disease or latitude and temperature) when there appears to be some kind of relationship between them.

To **extrapolate** is to go beyond the available data and make an educated prediction about what will happen "off the edge of the graph". (For example, using population data for the past few years, a reasonable prediction could be made for New Zealand's population next year, or in 5 years.)

To **interpolate** is to estimate a value that lies somewhere between known data values. For the purposes of both interpolation and extrapolation, it is assumed that the observable pattern continues between and will continue beyond the available data. This will not necessarily be true. Extrapolation is generally less reliable than interpolation because there is no guarantee that a previous trend will be maintained.

The number of data in a category or interval is known as its **frequency**. Frequency can be thought of as "number" or "total".

A **frequency table** is a table that organises data by category or interval and gives the frequency for each category or interval. For example:

Weeks between haircuts	2	4	6	8	10	12	14+
Number of classmates	4	6	9	5	2	3	1

"**I wonder**" questions are **investigative questions** – statistical questions or problems to be answered or solved. They consider the entire data set or population and do not involve locating an individual within a data set.

Two types of investigative question are of particular interest at this level:

- **Summary questions**, which usually involve a single variable and require the data to be described in some detail (for example, "I wonder how long it typically takes a year 6 student to run 100 metres?")

- **Comparative questions**, which involve comparing two or more subsets of data, for example, male and female, young and old, in relation to a common variable such as speed (for example, "I wonder whether year 6 girls are typically faster than year 6 boys?").

For a more detailed explanation in a context (It's on the Cards), see pages 26–27.

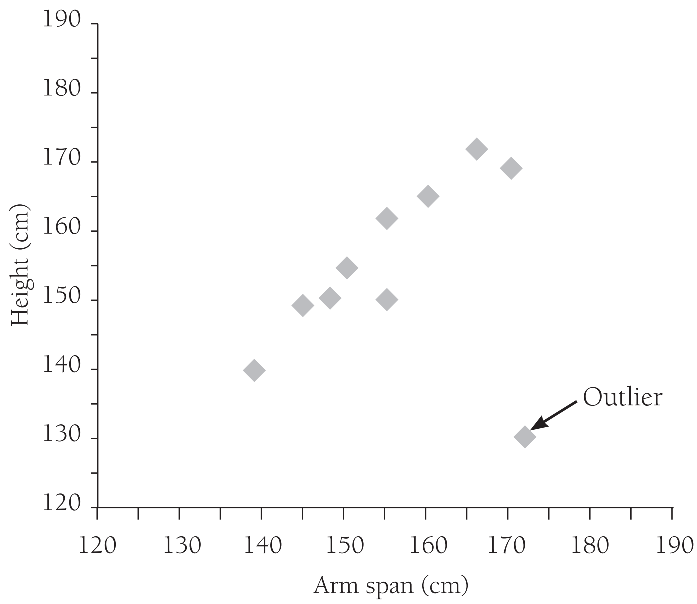
The **mean** of a data set is the sum of all the data values in the set divided by the number of values. It is sometimes called the average.

The **median** is the middle value in a data set when all the values are arranged in order from smallest to biggest or biggest to smallest.

The **mode** is the most commonly occurring value in a data set.

An **outlier** is an outlying value in a data set. It is a term that is often used in statistics. An outlier in a scatter plot is a point that is a long way from the rest. (It may be the result of a counting or measurement error or some other factor.) Outliers can affect the average (or mean) quite considerably.

Relationship between Arm Span and Height



Probability:

- **Probability** and **chance** are the same thing, although one or the other term may be more usual in a particular context.
- **Outcome**: the result of a trial (for example, a match or no match)
- **Trial**: performance of an action or actions where the outcome is uncertain (for example, the toss of a coin)
- **Experiment**: sometimes used interchangeably with trial; otherwise, a series of trials
- **Experimental probability**: the likelihood that something will happen, based on a number of trials
- **Theoretical probability** (**expectation**): the likelihood that something will happen, based on reasoning or calculation.

Tally marks (|) are used when counting or categorising data by hand. Every fifth stroke is drawn across the previous four, facilitating skip-counting by 5s and 10s. For example, ~~||||~~ ~~||||~~ || stands for 12.

In a **tally chart**, information is presented in three columns: category, tally, and frequency. For example:

Footware	**Tally**	**Frequency**
Shoes	𝍸	5
Sandals	𝍸 II	7

Variation is the term used to refer to the differences between data, particularly differences from an expected pattern or trend. To illustrate: if a coin is tossed a very large number of times, we would expect that the numbers of heads and tails would be approximately equal (because the two outcomes are equally likely). But in practice, if we were to toss a coin 100 times and then repeat this experiment 10 times, we would almost certainly get widely differing results, for example:

Experiment	1	2	3	4	5	6	7	8	9	10
Heads	41	44	52	53	42	49	47	50	38	45
Tails	59	56	48	47	58	51	53	50	62	55

Variation can be described and, at later levels, measured, using a variety of measures of spread from the simple to the sophisticated.

Links to *The New Zealand Curriculum*

Achievement Objectives

Achievement objectives in the teachers' notes for existing, revised, or new material in the *Statistics* students' book are from the mathematics and statistics area of *The New Zealand Curriculum*.

Key Competencies

The New Zealand Curriculum identifies key competencies that students will develop over time and in a range of settings. Schools can develop the key competencies within the mathematics and statistics learning area as well as encouraging and modelling values for students to explore.

The five key competencies identified in *The New Zealand Curriculum* are:

- thinking
- using language, symbols, and texts
- managing self
- relating to others
- participating and contributing.

The notes for the student activities in this revised *Statistics* book suggest one or more key competencies that the activities could help to develop. (You may, of course, decide to focus on key competencies other than those suggested.)

Activity Notes

Page 1: Wild about Juice

Achievement Objective

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Wild about Juice can be used to develop these key competencies:

- thinking
- participating and contributing.

Activity

In question **1**, the students must analyse the **category data*** presented in the table before they can agree or disagree with the statements made by the students in Room 2 and make up their own. Encourage statistical argumentation, insisting that it be settled with reference to the evidence in the table. Have the students report their findings to the class. The different statements in the activity can be used to help students learn the difference between acceptable generalisations and sweeping assumptions.

It is noteworthy that Orange was the second-best-selling juice on Monday but achieved zero sales for the rest of the week. This poses a problem of interpretation: did no one actually want Orange on the remaining 4 days or had the canteen run out? If the canteen ran out on Monday, was it during or at the end of lunchtime? Obviously there is no way of knowing, but it is worth exploring the implications of the different scenarios and seeing if it is possible to get agreement on the "balance of probabilities". Amiri's assertion that "Orange is least popular" raises the issue, and it is a complication in question **2**.

In question **2**, the students need to divide their totals by 20 and make statements regarding future canteen purchases. Encourage them to explain their answers as fully as possible. For example: "The canteen needs to purchase 3 boxes of Tropical because in 1 week they sold 27 Tropical. This means that they could sell 54 Tropical in 2 weeks. 54 is close to 60, so they should order 3 boxes of 20, even if some may be left over." It is particularly important that students understand that while **extrapolations** of this kind have to be made all the time (how else do shops maintain stocks?) they are far from certain in terms of **outcomes**. It may be that, next week, Tropical is the best seller and Orange & Mango gathers dust on the shelves!

* For bolded terms, see introductory section.

Pages 2–3: Petty Differences

Achievement Objective

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Petty Differences can be used to develop these key competencies:

- using language, symbols, and texts
- managing self
- relating to others.

Activity One

This activity is based on Wiremu's quest to find out what pets other students in his class have, presumably to help him decide what pet he will choose for his birthday present. It introduces data cards as a statistical tool. The question Wiremu asks is an "open" question, in which his classmates can name any pets.

Data cards are a means of collecting data in such a way that it can be easily sorted and re-sorted. The data on a card relates to a single person or object. These two cards were collected in surveys on travel to school and lunch activities:

Gender: Female Age: 12 Travel to school: walk Time taken: 10–20 minutes Lunch activity: basketball	Gender: Male Age: 12 Travel to school: bus Time taken: 15 minutes Lunch activity: athletics

Data cards can be simple, containing only one or two **variables***, or more complex, as in the examples above.

Use this activity to introduce the students to the concept of data "cleaning". By this, we mean removing irrelevant or difficult-to-interpret data. The cards in this activity contain a number of comments that come into this category. For example, "1 is sick and being treated" is beside the point. The students will also have to decide how they will interpret "a cat in Aussie" or "1 cat, but she's my mum's". Any data that appears "dirty" can be removed, but you should encourage your students to discuss what impact this will have on the overall distribution.

The students also need to understand how to use **tally charts**. If your students have had little experience with them, remind them that the fundamental idea is to group marks in fives to make totalling easy.

* For bolded terms, see introductory section.

Question **4** can be discussed and answered in small groups, with a follow-up teacher-led discussion. Particular attention could be paid to the place of fish in the statistics. Do they warrant the dominant position they have in the first chart? This activity provides a good example of the fact that the same data set can often be used to answer different questions. In this case, there is a world of difference between the questions "How many cats do my classmates have?" and "How many of my classmates have cats?"

Wiremu might go on to ask further "**I wonder**" **questions** and answer them by collecting other information from his classmates. For example, "I wonder why cats and dogs are the most popular pets?" or "I wonder why so few people have turtles (or axolotls or horses) as pets?"

Activity Two

This works best as a teacher-directed activity. Making a **strip graph** from the tally chart data provides an excellent opportunity for the students to visualise the link between the data and its visual representation.

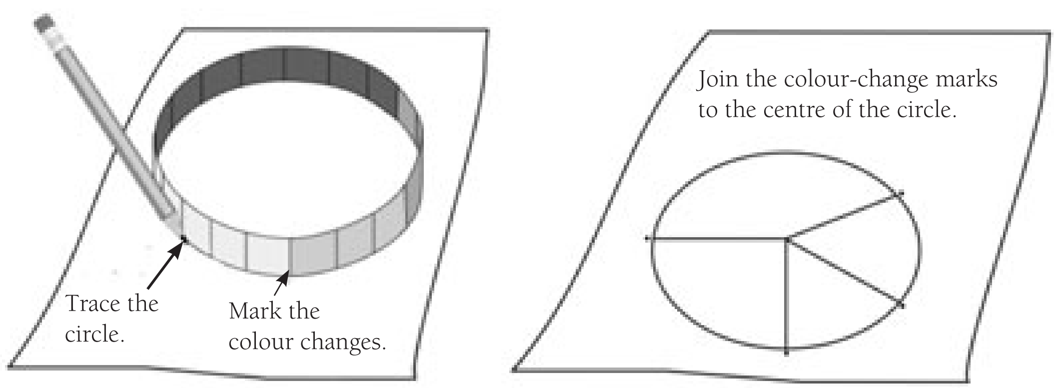
The students need to divide their strip of cardboard into 30 equal parts to match the favourite pet data from the class next door. Before the students start the task of dividing up the cardboard strips, you could try a practical activity with the class. Give eight students red cards to identify them as horses, give four students blue cards to represent guinea pigs, and so on. Then ask them to arrange themselves so that they model:

- a strip graph
- a **pie chart**
- a **bar graph**.

This will help the students to understand that they are dealing with the same data displayed in three different ways, not three distinct sets of data.

Discuss how each type of graphical display represents the data. The strip graph and pie chart show the relative proportions of data in each category. The bar graph shows the actual **frequency** of each data type (for example, the *number* of horses). Reinforce the need for the bars to be separated by a gap to show that the data is in the form of discrete categories.

To make the pie chart in **1d**, the students can put the bent strip graph on a piece of paper and trace its outline, marking where the colour changes. They then mark a centre point and join it to the marks on the circumference of the circle:



The students should label their graphs meaningfully and provide a succinct, accurate title for each.

As an extension to question **1e**, your students could:

- use a graph to estimate what fraction of the class prefers cats. (Which graph did they use and why?)
- use the tally chart to work out what fraction of the class does not prefer cats
- use the tally chart to work out what fraction of the class prefers cats and then compare this answer with the estimate based on the graph
- work out what fraction of the class prefers horses and dogs.

Question **2**, which involves evaluating the effectiveness of the various displays, provides another opportunity to develop statistical literacy.

Pages 4–6: All in the Family

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - gathering, sorting, and displaying multivariate category and whole-number data and simple time-series data to answer questions;
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

All in the Family can be used to develop these key competencies:

- thinking
- using language, symbols, and texts
- participating and contributing.

Activity One

In this activity, the students total the data in Table A and then summarise it in Table B so that deductions can be made. Three graphs are based on the information in Table B. Not all are equally suitable: the students need to realise that some graphs are better for certain types of information.

A **line graph*** is inappropriate for this data because graphs of this type are most suited to showing trends in data over a period of time (with time on the horizontal **axis**). Line graphs should not be used for **discrete** data when there is no meaning that can be attached to non-whole values on the horizontal axis (for example, a family with 2.7 children).

* For bolded terms, see introductory section.

A **bar graph** shows the frequency or number of values in each category and how data is clustered or spread. In this case, the graph makes it clear that there are more 2-sibling families represented in the class than any other kind of family.

Pie charts show proportions visually. The pie chart shown makes it reasonably easy for students to estimate the fraction of the class that has a given number of siblings, but it does not show how many families fall into each category.

To answer question **3**, students have to be able to make sense of the expression "two or more siblings", which will be unfamiliar to some. In terms of the data, it means "three or more children" because the class member is included as part of his or her own family. Those *not* included in Lani's statement are the families with two or fewer children.

Activity Two

This activity involves reorganising (or recategorising) data so that it can be analysed on the basis of family patterns and allow other features to become evident. This is possible because the data is **multivariate** (for each student, the data collected includes both number and gender). As part of their developing statistical literacy, students need to understand that grouping data (to create, for example, a pie chart or bar graph) always involves discarding potentially useful data. They also need to aim to collect multivariate data wherever possible because it can be analysed in more and more interesting ways and it is able to tell more stories. Discuss with your students what other related data Lani could have easily collected at the same time as she collected the age and gender of the children in her classmates' families. The most obvious is age. This would have made it possible to explore "place in family" patterns, for example, and it may have made it possible to better answer the original question relating to family passes.

Activity Three

This is a whole-class activity that will require your involvement in the planning. It could begin at the planning stage of the **PPDAC cycle**, given the preparation done in activities **One** and **Two**. Alternatively, the scope of the investigation could be broadened (perhaps by collecting age data as suggested above) and other questions answered, in which case students will need to begin at the problem stage of the investigation cycle. Data cards could be a useful tool for collecting the class data.

Ideally, students will have access to a computer spreadsheet/graphing program so that they can construct a range of graphs with ease. First check that your students have an adequate understanding of how to use the available spreadsheet programs.

Ensure that this activity does not just become a data-gathering and graphing exercise. The two final stages of the PPDAC cycle (Analysis and Conclusion) are the most important: what stories can this data tell us?

Page 7: It's on the Cards

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - gathering, sorting, and displaying multivariate category and whole-number data and simple time-series data to answer questions;
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

It's on the Cards can be used to develop these key competencies:

- thinking
- using language, symbols, and texts
- participating and contributing.

Activity

This activity focuses on the importance of asking good questions. At this level, students should be able to write and answer **summary*** and **comparative investigative questions**.

One of the most difficult aspects of conducting a survey is making sure that the survey questions you ask give you the data you need. Using the three data card questions in this activity, discuss:

- What is being asked?
- What sorts of answers will be given?
- Who will be interested in the results?

Emphasise that any investigation starts with a purpose and ends with conclusions and/or actions. "What is your favourite spare-time activity?" appears to be a reasonably straightforward question, but the students need to consider how those being surveyed might respond. If there is a clear favourite, there is no problem. But if those surveyed list lots of different activities, it might be difficult to discern a favourite. To avoid this problem, surveyors often offer a list of responses (perhaps with an "other" option) and ask respondents to select one. In such cases, the question is "closed" rather than "open".

You may need to remind the students how data cards are used to summarise information that is collected.

Be aware that questions such as "**I wonder** if there is a relationship between height and foot length?" involve understandings and require tools that are not normally introduced at this level.

* For bolded terms, see introductory section.

A **summary question** for this activity might be: "I wonder what the typical bedtime for a student is?" (The use of "typical" leaves scope for a variety of measures and attributes of the variable ["bedtime"] to be investigated.)

A **comparative question** for this activity might be: "I wonder if girls tend to go to bed later than boys?"

To help the students complete question **2**, you will need to review the kinds of graph that are likely to be useful and appropriate for this situation. **Strip graphs**, **dot plots**, **stem-and-leaf graphs**, **bar graphs**, and **frequency tables** normally show a single variable but can be paired (or in the case of a bar graph, have double bars) so that comparisons can be made for a second variable. While computers make it very easy to create certain standard types of graph (for example, pie chart and bar graph), computer graphs can be difficult to customise. Many graphing programs will not create dot plots.

Question **3** is a whole-class activity that requires your involvement in the planning process. The students need to fully understand the **PPDAC cycle**. The task could be extended to allow the students (possibly in groups) to collect the data and work through each stage of the investigation cycle.

Pages 8–9: Television Times

Achievement Objective

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - gathering, sorting, and displaying multivariate category and whole-number data and simple time-series data to answer questions;
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Television Times can be used to develop these key competencies:

- managing self
- relating to others.

These activities require the students to analyse the data displayed in a graph and to consider reasons for some of the features evident in the data.

Activity One

For this activity, the students need to have prior knowledge of the **stem-and-leaf graph***, understanding how to draw this type of graph and recognising its important features. They should also know the difference between ordered and unordered stem-and-leaf graphs. As an extension, you could introduce them to **back-to-back stem-and-leaf graphs**.

You could also use this activity to introduce **mode** and **median**.

---

* For bolded terms, see introductory section.

The data shown in Ngaio and Maaka's stem-and-leaf graph is well clustered and has no real **outliers**. The mode is 1 hour and 30 minutes. Because there is an even number of data values in this activity, the median will be midway between readings 15 and 16, both of which are 1 hour and 30 minutes. So the median is 1 hour and 30 minutes. This means that half the students in the class watch TV for this amount of time or less and the other half watch TV for this amount of time or more.

Be aware that some of the statements in question **3** pose questions of interpretation. These include: (**i**) and (**iv**) by "most" do we mean a simple majority ("more than half") or "nearly everyone" or somewhere between these two definitions?; (**iii**) by "30 minutes" do we mean "exactly 30 minutes" or "at least 30 minutes"?; and (**vii**) by "200 minutes" do we mean "exactly 200 minutes" or "at least 200 minutes"?

Allow your students to find these issues for themselves and get them to justify the decisions they make.

Activity Two

This is a teacher-directed activity. Once the data is collected, the students need to sort it into a stem-and-leaf graph, working either individually or in groups. They should start with an unordered graph and then create a second graph in which the leaves are ordered.

If your students need a challenge, you could consider collecting a greater range of data (data relating to several **variables**) and then using this database to try and answer a wider range of questions.

When students gather data from their peers for this activity, have them examine their results to look for clusters and **outliers**. What could explain the outliers? How are the mode and median affected by significant outliers? What if three people were to watch 4 hours of TV, as in the following distribution?

Time Spent Watching TV

Hours	Minutes
0	00 30 45
1	00 15 30 30 30 30 45
2	00 15 30 45
3	30
4	00 00 00

The graph above shows that more than half of this class (10 out of 18) watch less than 2 hours of TV each night. Because the number of students represented by the graph is an even number (18), the median viewing time is midway between students 9 and 10 (midway between 1 hour 30 minutes and 1 hour 45 minutes), which is 1 hour $37\frac{1}{2}$ minutes.

Pages 10–11: Well Weathered

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - gathering, sorting, and displaying multivariate category and whole-number data and simple time-series data to answer questions;
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Well Weathered can be used to develop these key competencies:

- using language, symbols, and texts
- managing self
- relating to others.

Activity

In this activity, the students analyse and interpret information found in graphs and tables. They also carry out their own investigation, comparing the weather in their area with that of another location of their choosing.

Before beginning, you will need to explain (or get the students to find out) that NIWA is an abbreviation for the National Institute of Water & Atmospheric Research Limited, the Crown Research Institute that studies the scientific basis for the sustainable management and development of New Zealand's atmospheric, marine, and freshwater systems and associated resources.

The two graphs on which the first part of the activity is based are **time-series graphs***. They show the average daily temperature and average monthly sunshine hours per year. It is very important that the students think about exactly what this means in each case, and how the data in these graphs would have been gathered and collated. Students could do this thinking in pairs or small groups and then present their thoughts to the class.

The data-gathering and analysis process for the first graph would involve:

- Taking the temperature at regular intervals throughout each 24-hour period. (Yes, it's *daily*, not *daytime*!)
- For each month, finding the average (**mean**) of all these observations.
- Doing the above for a number of years and averaging the results for each month (so that the data doesn't just reflect one year).

Depending on where your students are at, you may need to explain or clarify the meaning of "average" and "mean". While they need to have some understanding of the notion, they do not need to know how to calculate a mean to do this activity or the investigation that follows.

* For bolded terms, see introductory section.

The data-gathering and analysis process for the second graph would involve:

- using an electronic device to detect sunshine and keep a running total of the hours for each month
- doing this for a number of years and averaging the results for each month.

Note that, although based on continuous **variables** (time and temperature), the data in the two graphs is actually **discrete** because, in each case, it is summarised as a series of 12 data values. The lines that join each pair of points are there to show the trend or pattern.

As the students do the different parts of the activity, encourage them to look at the detail as well as the overall picture:

- In what part of the year are the average daily temperatures closest together? How do the graphs show this?
- In what months of the year are the temperatures the most different? Why might this be?
- In what months are the average sunshine hours almost exactly the same for the two places? How do the graphs tell us this?

When it comes to reasons for the differences between Kaitaia and Tekapō, the students need to go beyond what they can find in the graphs and look for geophysical explanations. There is excellent scope here for investigating mathematics in a context that relates to another learning area.

Investigation

This investigation involves the students creating their own time-series graphs. Such graphs can be conveniently created in most graphing programs. However, it may not be obvious how to get the horizontal (time) **axis** labelled correctly, so if you are not sure how to do this, it would be best to go through the process yourself first. A likely process is:

- Select the number data only (not the names of the months) and create a line graph using the *Line chart* option.
- From the *Chart* menu (which appears on menu bar when you click on a graph), select *Source data* and then the *Series* tab. Click the cursor in the box that says *Category (X) axis* labels and then select the cells in your spreadsheet that contain the names of the months. (Alternatively, type the names of the months into this box, separated by commas.)
- To get the labels on the horizontal axis aligned with the "tick marks", right-click on the horizontal axis to obtain the *Format axis* options. Select the *Scale* tab and then remove the tick from the check box beside *Value (Y) axis crosses between categories*.

Data on the NIWA website is given to one decimal place. Students will need to be at stage 6 on the Number Framework if they are to understand how to round decimals to the nearest whole number. To round data to the nearest whole number in a typical spreadsheet, go to the *Format* menu, select *Cells*, then the *Number* tab. From the category list, select *Number*. A box labelled *Decimal places* appears to the right. Make this number zero and click OK.

During the course of the investigation, encourage the students to write full sentences in paragraph form when describing their findings. They should also use information from their graphs to support any comments they make.

Pages 12–13: Winning Dots

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - gathering, sorting, and displaying multivariate category and whole-number data and simple time-series data to answer questions;
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Winning Dots can be used to develop these key competencies:

- thinking
- using language, symbols, and texts
- participating and contributing.

These activities involve the students in drawing and interpreting **dot plots***. The students have to be able to find evidence within the dot plot to either support or refute various statements made about **discrete** data. They also carry out an activity that begins with a whole-class data-gathering exercise.

Activity One

This activity provides an ideal opportunity to discuss the difference between discrete and **continuous** data. In this case, the **variable** is distance (or length). Distance is not restricted to whole metres (for example, 14 or 15 metres); it can also be any distance in between (for example, 14.5 or 14.675 metres). For this reason, distance is a continuous variable.

In this activity, however, the data has been recorded as if it were discrete. You could discuss with your students why and how this might have been done. As the event was a fun competition, not the Olympic Games, there was neither the need nor the ability to record throws to the nearest millimetre. The throw may have taken place on the netball court, with chalk lines marking out the metres. If a throw landed between the 14 and 15 metres mark, it would be recorded as 14 metres. In a situation like this, the "rounding up" convention would not normally be applied.

You will need to discuss the idea of clustering. A "cluster" is a group of data found together in a clump.

If your students have had prior experience with dot plots, you may want them to do this activity with minimal further guidance. If they are uncertain about how to interpret them, ask a few simple questions to get them started, for example:

- What was the most common distance thrown?

* For bolded terms, see introductory section.

- What was the shortest distance thrown?
- What is the distance between the longest throw and the shortest throw?

When the students go on to make their own dot plots (first in question **2**), remind them to record one dot for each piece of data.

Activity Two

This whole-class activity will require direction from you. The class need to determine what data they are going to collect, and they need an understanding of the **PPDAC cycle**. The amount of scaffolding required will depend on the ability levels of the students.

This task reinforces the idea of clustering. The range of results may lead to an interesting discussion about the importance of context when looking at patterns in data. Discuss what is meant by **outliers** so that the class are able to decide if values at the extremes are to be retained as legitimate or rejected on the grounds that they are the result of some accident or abnormality in the process.

Comparing their own class data with that of another class is a very valuable exercise. This is an appropriate time to discuss **variation** in data sets. It is a feature of all statistics data gathering that each different sample gives different results. Smaller samples tend to be more variable. Larger samples tend to be more similar to each other.

Make sure that the students are aware that one of the big advantages of dot plots is that the distribution of the data is clearly evident. The "picture" can be lost when data is summarised in a table.

Pages 14–15: Old Enough?

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Old Enough? can be used to develop these key competencies:

- thinking
- managing self
- relating to others.

Activity

This task gives the students the opportunity to display and interpret a given set of data. A whole-class discussion as to what to do with the data, including the responses that are a little unclear, would be useful before the students attempt question **1**.

Discuss with your students the graphing of **discrete data*** and the various **bar graph** options available with your spreadsheet or graphing program and the advantages and disadvantages of each type. The students should make their graphs completely self-explanatory. This means labelling the **axes**, providing a key or legend if necessary, and giving the graph an accurate title.

Question **1b** asks for a bar graph. If creating the graph in a computer graphing program, the tricky bit is to get the horizontal axis labelled correctly because it has two sets of labels (No, Yes, No, Yes, No, Yes and Children, Adults, Total). A likely process is:

- Copy the table in question **1** into a spreadsheet, just as it appears in the book.
- Selecting only the number data, create a bar graph.
- From the *Chart* menu (which appears when you click on a graph), select *Source data* and then the *Series* tab. Click the cursor in the box that says *Category (X) axis labels* and then select all the cells in your spreadsheet that contain the Children, Adults, and Total labels AND the Yes and No labels. Click on OK.

A **pie chart** is asked for in question **1d**. This can easily be created by computer, but if the students are doing the task by hand, make sure they know how to divide a circle into the correct proportions using a calculator and a protractor.

Question **2** involves drawing a second bar graph, this time showing the suggested ages. Along with the first bar graph and the pie chart, this third graph provides the basis for the discussion in question **3**.

The answer to question **4** is yes, but some students may find it difficult to see why. Discussion should highlight the importance of thinking extremely carefully about what data is wanted and exactly what questions are needed to ferret it out.

---

* For bolded terms, see introductory section.

Pages 16–17: Way to Go

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - gathering, sorting, and displaying multivariate category and whole-number data and simple time-series data to answer questions;
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Investigation		Literacy		Probability
P	P	D	A	C

Key Competencies

Way to Go can be used to develop these key competencies:

- using language, symbols, and texts
- relating to others.

These activities give you an opportunity to introduce the students to **histograms***, which they will meet more fully in level 4 of the curriculum.

Activity One

This activity involves **continuous data**. You may need to remind the students what this is. They need to know that all non-rounded measurement data is continuous and that a continuous **variable** can take on any value within an **interval**. The graph that best displays grouped continuous data is the histogram. A **stem-and-leaf graph** could have been used to display this information, but a histogram is more visually informative.

Room 2's histogram comes from the following table:

Distance from school (km)	Number of students
0–1	3
1–2	8
2–3	6
3–4	0
4–5	2
5–6	1
6+	0

The speech bubble in the students' book raises the issue of the boundaries between intervals: if a distance is measured to be 4 kilometres, should it be included in the 3–4 or the 4–5 interval? While there is no mathematical reason for it being one or the other, the usual practice is to

* For bolded terms, see introductory section.

include the boundary case in the next interval. We say, for example, that an interval includes "all distances that are equal to or greater than 3 kilometres but less than 4 kilometres". This is too wordy to fit in a column in a table. (Students will learn numeric short cuts at a later stage.)

By comparing the two histograms, the students should see that the data for Room 2 is very tightly clustered around the shorter distances, while the data for the Otago school is clustered around the further distances. In short, the Auckland students all live within a few kilometres (walking distance?) of school; the great majority of the Otago students live well beyond walking distance. (How do they get to school?)

Question **3** is there to make a simple point: the scale on the horizontal **axis** begins and ends wherever the data requires. It will vary with every graph. In this case, if the same scale were used, the first histogram would have to be very narrow and the second very wide.

Activity Two

Before assigning this task, you will need to obtain a local street or district map and photocopy it for the students to use. You will also need lengths of string or narrow strips of thin card that can be marked out with units that match the scale of the map to be used.

If an appropriate map does not exist, then make it a homework task to find the distance between home and school. A car odometer will do a good job! Encourage the students to predict the results before starting the exercise.

The measuring exercise in question **1** is valuable in its own right, not just as a means of creating data for a statistics task. The students should bend the string or strip of card to conform to the route they take from home to school. If using a strip of marked card, they can then simply read the distance in kilometres off the strip. If using string, they will need to measure the length against a ruler and then translate centimetres into kilometres. Students who are new to the notion of scale will need an introduction. If you think it appropriate, you can simplify the measurement task by pre-preparing measuring strips and photocopying them for student use. In any case, you will need to discuss with your students what level of accuracy is to be aimed for. For most maps, 0.1 kilometre should be suitable. Given that whole kilometres (for example, 5 kilometres) are included in the 5–6 kilometre interval, a distance of (say) 4.95 kilometres should be rounded down to 4.9 (rather than up to 5) to ensure that it is included in the correct interval.

The aim in questions **2** and **3** is to have your students thinking about the realities behind the data portrayed in the graphs. Ask students to replace context-empty statements such as "the 5 to 6 bar is tallest" or "17 to 18 kilometres is an outlier" with interpretive statements such as "more students travel between 5 and 6 kilometres than any other distance" and "one student has to travel between 17 and 18 kilometres – that's 4 kilometres further than anyone else in the class".

In answering question **3**, the students should compare these features and interpret them in terms of the context:

- symmetry (or lack of it)
- clusters (are there obvious clusters, and if so, where are they and how big are they?)
- range (the difference between the largest and smallest values in a data set)
- **outliers**
- **median**
- central tendency (the extent to which data is clustered around some central point) and spread.

Once they have completed questions **2** and **3**, encourage the students to:

- explain their answers to the class
- develop "**I wonder**" **questions** from their conclusions.

Pages 18–19: Scratch 'n' Win

Achievement Objective

Probability

- Investigate simple situations that involve elements of chance by comparing experimental results with expectations from models of all the outcomes, acknowledging that samples vary (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Scratch 'n' Win can be used to develop these key competencies:

- thinking
- using language, symbols, and texts.

Activity

In this activity, the students carry out a simple **probability*** (or **chance**) **experiment** a number of times and make deductions based on the **outcomes** they obtain. Some advance preparation is involved (photocopying) as each pair of students needs a double set of the copymaster cards.

The activity is designed to help students develop their understandings experientially, so all they should need by way of introduction is a clear understanding of what they are to do. There are two initial experiments (questions **1a** and **1b**). These will give you the opportunity to make sure that everyone knows what to do and should demonstrate to students that the outcome of a **trial** is very **variable**.

Following the initial experiment, the students carry out a series of 20 further experiments, recording the outcomes in a table.

For question **1c**, challenge your students to get beyond "pretty difficult" for an answer and use the statistics in their **tally chart** to express the probability of a match as a number or a ratio. This number is the **experimental probability** of obtaining a match. The number they come up with is likely to be near $\frac{1}{4}$ because, whatever first symbol is selected, there are only 4 to choose from for the second symbol. (Because all four symbols have an equal chance of being picked, the **theoretical probability** of a match is said to be $\frac{1}{4}$.) By pooling their results, the students should obtain further confirmation that the probability of a match is indeed 1 in 4.

The answer to question **1e** is, of course, that *no* number of trials will guarantee a match. While probability can predict and explain patterns, it can never guarantee the outcome of a trial or the result of an experiment. In this case, the probability of a match is $\frac{1}{4}$, but it is possible that 100 or even 1000 trials might not produce one – very unlikely, but nevertheless possible.

Question **2** alters a key variable (the number of symbols on a card) and invites the students to predict how this change will affect the probability of a match. This should help them generalise their findings from question **1**.

As an extension to this activity, you could challenge your students to predict what the probability would be of "scratching" two specified symbols (two crosses, for example). (For the 4-symbol

---

* For bolded terms, see introductory section.

card, there is only a 1 in 4 probability of picking the cross in the top row and then only a 1 in 4 probability of picking the cross in the bottom row. So the probability of picking both crosses is "a quarter of a quarter" or $\frac{1}{16}$.)

Note that the 4-symbol experiment could equally well be modelled by a pair of 4-sided dice and 6-symbol cards by a pair of regular 6-sided dice.

Pages 20–21: Superbeans

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - gathering, sorting, and displaying multivariate category and whole-number data and simple time-series data to answer questions;
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Probability

- Investigate simple situations that involve elements of chance by comparing experimental results with expectations from models of all the outcomes, acknowledging that samples vary (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Superbeans can be used to develop these key competencies:

- thinking
- using language, symbols, and texts.

In these activities, students take "beans" from a bag and then compare what they get with what they expected to get.

Activity One

The kind of table found in question **1** is often used as a convenient way of setting out the possible **outcomes*** for a two-**trial probability experiment**. If your students have not filled in one before, they may need help to get started (using the HH already entered).

When completed, it should be clear that there are four possible outcomes: HH, TH, HT, and TT. A head and a tail are equally likely outcomes when a coin is tossed, so each of these four outcomes is also equally likely. This means, for example, that there is a 1 in 4 ($\frac{1}{4}$) probability of getting HH when two coins are tossed. Because two of the outcomes consist of a head and a tail, the probability of a head and a tail (ignoring order) is 2 out of 4, or $\frac{1}{2}$.

* For bolded terms, see introductory section.

These fractions ($\frac{1}{4}$ and $\frac{1}{2}$) are both between 0 and 1, as are all probabilities. 0 means completely impossible; 1 means completely certain. All other outcomes lie somewhere on the continuum between them. This is an important probability concept.

Activity Two

Activity One was designed to help students understand that there may be more than one path to an outcome (both HT and TH qualify as "a head and a tail"). They need to use this knowledge in this second, considerably more complex, activity.

Question **1** explains how the Superbean simulation is to work and asks the students to do 20 trials, recording the outcomes in the shorthand introduced in **Activity One**.

In questions **2** and **3**, the students collate and analyse their results. They may find that no clear pattern emerges from the initial 20 trials. This is likely to change when they combine their outcomes with those of classmates to get a bigger data set (when it comes to number of trials, the bigger the better). What they are likely to find is that most outcomes contain two Superbeans of the same colour; much less common are those that consist of three different colours, and rarest of all are those in which all Superbeans are the same colour. Using the language of question **3b**, two of the same colour is the "most likely" outcome and all the same colour is the "least likely" outcome.

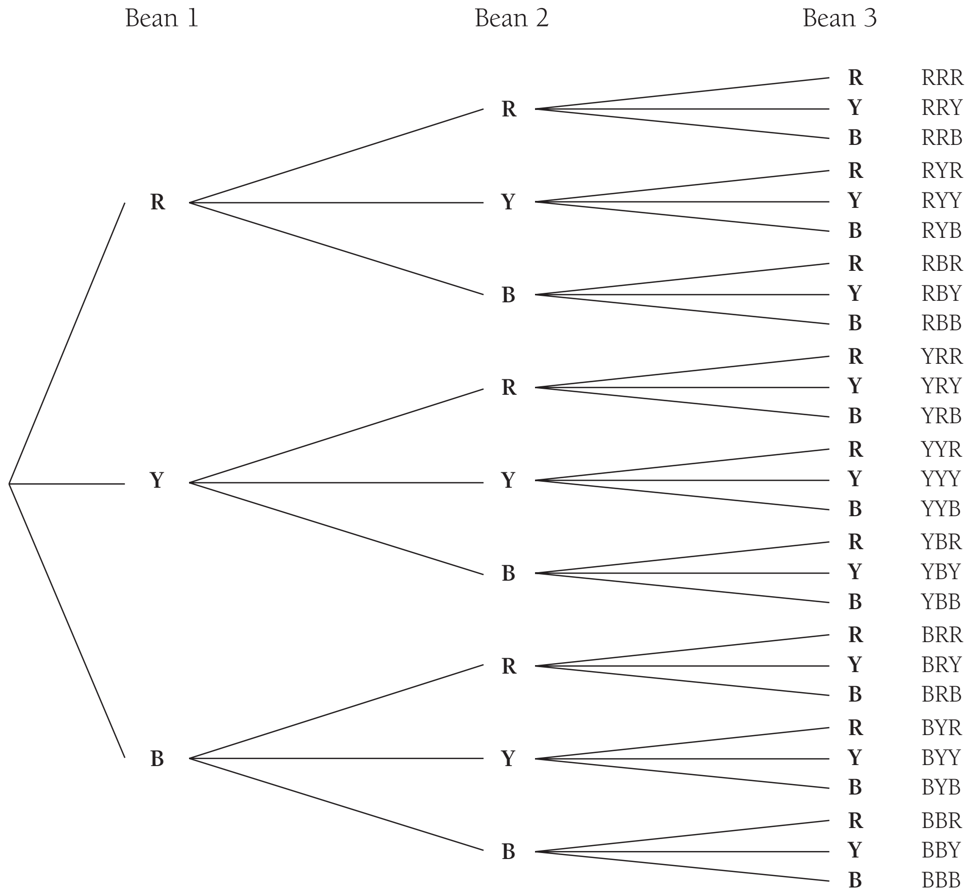
In question **4**, the students list all the possible outcomes for the 3-Superbeans experiment. They will have to do this systematically if they are to find them all. Challenge them to find a way of doing this. Once they have done so, they will be able to see that there are a total of 27 possible outcomes. 3 of these outcomes (RRR, YYY, and BBB) are same-colour, 6 are all-different, and the remaining 18 are two-the-same combinations. This information should go a long way towards explaining the experimental results obtained by the students in the earlier parts of this activity.

Note (for your information only)
The above explanation is correct as far as it goes. But it would not be correct to go on to say that the **theoretical probability** of getting (for example) a same-colour outcome is $\frac{3}{27} = \frac{1}{9}$. The reason for this is that once a Superbean has been removed from the bag, the probabilities change for the remaining Superbeans. We illustrate this with reference to RRR:

To start with, the bag contains 30 Superbeans of which 10 are red, so the probability that the first Superbean is red is $\frac{10}{30}$. If the first Superbean removed from the bag is red, there are now 9 red Superbeans in a bag that contains 29 Superbeans, so the probability that the second Superbean is red is $\frac{9}{29}$. Similarly, if the second Superbean is red, the probability that the third is also red is $\frac{8}{28}$. The probability that all three will be red is therefore $\frac{10}{30} \times \frac{9}{29} \times \frac{8}{28}$, which is rather less than $\frac{10}{30} \times \frac{10}{30} \times \frac{10}{30} = \frac{1}{27}$.

Extension

If you have a student or students in need of a challenge, you could suggest that they create a tree diagram (see following page) to show the possible outcomes for the 3-Superbeans experiment. This can most easily be done using a computer drawing program.



Bean 1
Bean 2
Bean 3
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
R
Y
B
RRR
RRY
RRB
RYR
RYY
RYB
RBR
RBY
RBB
YRR
YRY
YRB
YYR
YYY
YYB
YBR
YBY
YBB
BRR
BRY
BRB
BYR
BYY
BYB
BBR
BBY
BBB

Pages 22–23: Cold Coffee

Achievement Objectives

Statistical investigation

- Conduct investigations using the statistical enquiry cycle:
  - identifying patterns and trends in context, within and between data sets;
  - communicating findings, using data displays (Statistics, level 3).

Statistical literacy

- Evaluate the effectiveness of different displays in representing the findings of a statistical investigation or probability activity undertaken by others (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Cold Coffee can be used to develop these key competencies:

- managing self
- relating to others
- participating and contributing.

Activity One

This scenario involves temperature data gathered over a 48-minute period. The task is to graph and then interpret the data.

The appropriate graph is a **time-series graph***. Such a graph is best created using a computer graphing program, if available. A likely process is:

- Enter the data from the table in the students' book into the computer spreadsheet (keep the same table structure).
- Select the three columns of temperature data and their headings (but not the data from the time column) and create a line graph using the *Line chart* option.
- From the *Chart* menu, select *Source data* and then the *Series* tab. Click the cursor in the box that says *Category (X) axis labels* and then select the cells in your spreadsheet that contain the time intervals (0, 3, 6 …).
- To get the labels on the horizontal axis aligned with the "tick marks", right-click on the horizontal axis to obtain the *Format axis* options. Select the *Scale* tab and then remove the tick from the check box beside *Value (Y) axis crosses between categories*.

Activity Two

This activity involves a simple science experiment. Like all experiments, it needs careful preparation if it is to yield worthwhile data. The idea is to find out how the physical properties of a mug affect the temperature of the contents. For the purposes of a classroom experiment, iced water is a less risky option than boiling water.

* For bolded terms, see introductory section.

Involve the students fully in setting up this experiment and expect them to use all the knowledge they have acquired in their science learning. List all the **variables** that will influence the **outcome** and decide how these can best be controlled so that the experiment is a fair test of the mugs' insulating properties. For example, the volume of water in each mug should be the same, as should the thermometers used to register the temperature. It doesn't matter exactly what mugs are used as long as they are likely to have very different insulating properties. It is this difference that will make the experiment interesting.

In the debrief that follows the experiment, discussion could focus on these areas:

- A comparison of the two time-series graphs: how are they different and how are they similar?
- What factors might be responsible for the initial, rapid temperature change?
- If the experiment was continued for hours or even days, what would happen to the graphs? (They would level out as they reached the ambient [surrounding] air temperature and then rise and fall in response to changes [including night–day fluctuations] in the ambient temperature.)
- What does the slope of the graph tell us? (An upwards slope indicates an increase in temperature and a downwards slope a decrease. The steeper the slope, the faster the increase or decrease.)

Page 24: Dicey Differences

Achievement Objective

Probability

- Investigate simple situations that involve elements of chance by comparing experimental results with expectations from models of all the outcomes, acknowledging that samples vary (Statistics, level 3).

Investigation	Literacy	Probability

P	P	D	A	C

Key Competencies

Dicey Differences can be used to develop these key competencies:

- thinking
- participating and contributing.

Activity

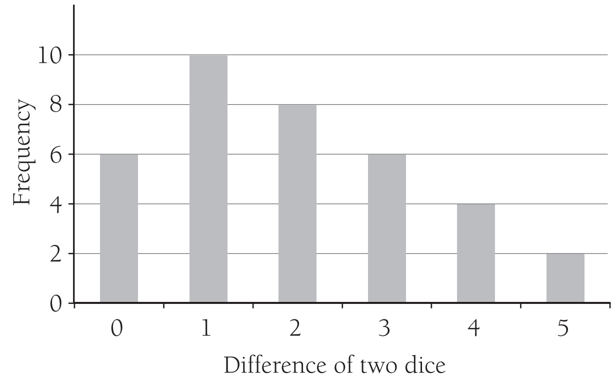
In this activity, the students explore the concept of fairness by playing a game that turns out to be absolutely unfair.

The concept is extremely simple, and students are likely to get most out of the activity if they are allowed to play the game and investigate its fairness with minimal teacher guidance.

It should take no more than 10 rounds, as suggested in question **2**, to establish that the game strongly favours one player. If students combine their results with those of others, as suggested in question **3**, they will have enough data to provide some clear patterns.

These patterns could be confirmed by entering the totals into a computer spreadsheet and graphing the results as a **bar graph***. The resulting graph is likely to have a shape similar to this one:

Dicey Differences: Scores Table



Question **4** asks the students to try and explain the pattern of results. This requires them to move from the **experimental** (playing the game and recording **outcomes**) to the **theoretical** (looking for a mathematical reason).

As a first step, they need to consider all the possible outcomes when they toss two dice. The key to this is the understanding that the two dice must be treated as different and that the throw (2, 3 [a 2 on the first dice and a 3 on the second]) is not the same as (3, 2). Equipped with this understanding, they can create a difference table (also found in the Answers). The numbers in the left-hand column represent the first dice and the numbers across the top the second dice (not that it matters which dice goes where). The numbers in the body of the table are the resulting differences.

Dice 2

Dice 1	**1**	**2**	**3**	**4**	**5**	**6**
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

From this table, it can be seen, for example, that there are only two throws that will give a difference of 5 [(6, 1) and (1, 6)] but 10 that will give a difference of 1. This means that the **probability** of getting a difference of 1 is 5 times greater than the probability of getting a difference of 5.

There are 36 possible outcomes when two dice are thrown, so the probability of scoring a difference of 5 is $\frac{2}{36} = \frac{1}{18}$. It is important that students realise that this does not mean that if they throw a pair of dice 18 times, a difference of 5 will come up exactly once. Probability can predict long-run patterns, but it can never guarantee a particular result.

* For bolded terms, see introductory section.

Copymaster: All in the Family

Table A

Family	Boys	Girls	Total
1	4	1	
2	0	1	
3	2	1	
4	1	1	
5	0	2	
6	2	2	
7	0	4	
8	1	1	
9	1	2	
10	0	1	
11	2	1	
12	0	2	
13	2	1	
14	1	1	
15	3	1	
16	0	2	
17	1	1	
18	1	2	
19	2	0	
20	2	0	
21	1	0	
22	2	1	
23	1	1	
24	2	1	
25	1	2	
26	0	2	
27	1	1	
28	2	2	
29	2	1	
30	1	0	
31	0	2	

Table B

Number of children per family	Number of families
1	
2	
3	
4	
5	

Table A

Family	Boys	Girls	Total
1	4	1	
2	0	1	
3	2	1	
4	1	1	
5	0	2	
6	2	2	
7	0	4	
8	1	1	
9	1	2	
10	0	1	
11	2	1	
12	0	2	
13	2	1	
14	1	1	
15	3	1	
16	0	2	
17	1	1	
18	1	2	
19	2	0	
20	2	0	
21	1	0	
22	2	1	
23	1	1	
24	2	1	
25	1	2	
26	0	2	
27	1	1	
28	2	2	
29	2	1	
30	1	0	
31	0	2	

Table B

Number of children per family	Number of families
1	
2	
3	
4	
5	

Copymaster: It's on the Cards

Girl	Playing sport	Boy	Playing sport	Girl	Playing sport
9 p.m.	A pet cat	8 p.m.	Cellphone	9 p.m.	A computer

Girl	Nothing	Girl	Playing sport	Girl	Watching TV
9 p.m.	Cellphone	8 p.m.	Cellphone	8.30 p.m.	Camera

Girl	Gaming	Girl	Playing sport	Girl	Playing clarinet
11.30 p.m.	Cellphone	8.30 p.m.	Overseas trip	8.30 p.m.	A pet dog

Boy	Reading	Boy	Reading	Boy	Watching TV
10 p.m.	Gaming console	10.30 p.m.	Anything	9.30 p.m.	A book

Copymaster: It's on the Cards

Boy	Playing sport
9.30 p.m.	Everything

Girl	Playing Sport
8.30 p.m.	A hug

Girl	Reading
10 p.m.	Cellphone

Boy	Watching TV
9.30 p.m.	I don't know

Girl	Playing sport
8 p.m.	A pet horse

Boy	Playing guitar
10 p.m.	Fishing gear

Boy	Playing sport
8.30 p.m.	Motorbike

Boy	Gaming
9 p.m.	Gaming console

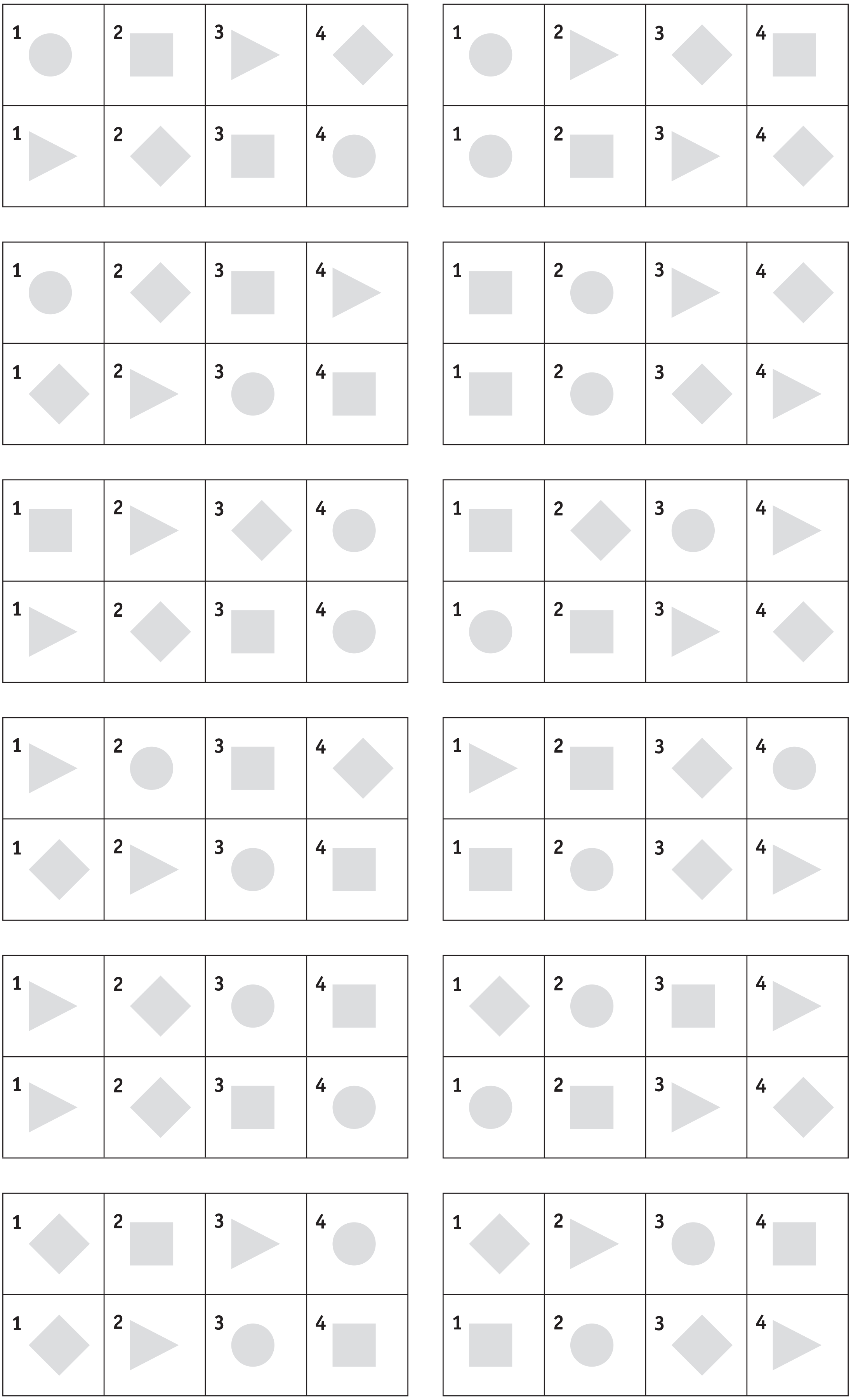
Girl	Watching TV
8 p.m.	Travel

Girl	Watching TV
8.30 p.m.	Money

Boy	Gaming
10.30 p.m.	Gaming console

Girl	Playing sport
9 p.m.	Go Kart

Copymaster: Scratch 'n' Win



Acknowledgments

The Ministry of Education and Learning Media would like to thank Sandra Cathcart (The University of Auckland Faculty of Education) and Karen Gibbs (Taradale High School) for their contribution to the development of these teachers' notes.

On the cover, the photograph is by Mark Coote and the illustrations are by Fraser Williamson. The illustrations on page 23 are by Dawn Barry. All other diagrams and figures, except for the poster noted below, are by Rachel Street.

Editorial team: Susan Slaughter, Ian Reid, June Hannah
Designer: Rachel Street

Published 2008 for the Ministry of Education by
Learning Media Limited, Box 3293, Wellington, New Zealand.
www.learningmedia.co.nz

The *Answers and Teachers' Notes* are available online at
www.tki.org.nz/r/maths/curriculum/figure/level_3_e.php

Dewey number 510
ISBN 978 0 7903 3347 2
Item number 33347
Students' book: item number 33352

Ministry of Education Customer Services: freephone 0800 660 662; freefax 0800 660 663;
email: orders@thechair.minedu.govt.nz; online: www.thechair.minedu.govt.nz
Please quote item number.

This resource has been developed to support
The New Zealand Curriculum