



# Business Statistics

Eliyathamby A Selvanathan

Saroja Selvanathan

Gerald Keller

---

**AUSTRALIA / NEW ZEALAND  
EDITION 6**

**Business Statistics: Australia and New Zealand**  
**6th Edition**  
**Eliyathamby A Selvanathan**  
**Saroja Selvanathan**  
**Gerald Keller**

Publishing manager: Dorothy Chiu  
Publishing editor: Geoff Howard  
Developmental editor: Emily Spurr / Kylie McInnes  
Senior project editor: Nathan Katz  
Cover designer: Studio Pounce  
Text designer: Norma Van Rees  
Editor: Sylvia Marson  
Proofreader: Greg Alford  
Indexer: Russell Brooks  
Permissions/Photo researcher: Helen Mammides  
Art direction: Danielle Maccarone  
Typeset by Cenveo Publisher Services

Any URLs contained in this publication were checked for currency during the production process. Note, however, that the publisher cannot vouch for the ongoing currency of URLs.

First published by Nelson Australia in 2000  
Second edition published in 2002  
Third edition published in 2004  
Fourth edition published in 2006  
Fifth edition published in 2011  
This sixth edition published in 2014

© 2014 Cengage Learning Australia Pty Limited

#### Copyright Notice

This Work is copyright. No part of this Work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written permission of the Publisher. Except as permitted under the *Copyright Act 1968*, for example any fair dealing for the purposes of private study, research, criticism or review, subject to certain limitations. These limitations include: Restricting the copying to a maximum of one chapter or 10% of this book, whichever is greater; providing an appropriate notice and warning with the copies of the Work disseminated; taking all reasonable steps to limit access to these copies to people authorised to receive these copies; ensuring you hold the appropriate Licences issued by the Copyright Agency Limited ("CAL"), supply a remuneration notice to CAL and pay any required fees. For details of CAL licences and remuneration notices please contact CAL at Level 15, 233 Castlereagh Street, Sydney NSW 2000, Tel: (02) 9394 7600, Fax: (02) 9394 7601  
Email: [info@copyright.com.au](mailto:info@copyright.com.au)  
Website: [www.copyright.com.au](http://www.copyright.com.au)

For product information and technology assistance,  
in Australia call **1300 790 853**;  
in New Zealand call **0800 449 725**

For permission to use material from this text or product, please email  
**[aust.permissions@cengage.com](mailto:aust.permissions@cengage.com)**

#### National Library of Australia Cataloguing-in-Publication Data

Author: Selvanathan, E. Antony, 1954-author.  
Title: Business statistics : Australia and New Zealand / Eliyathamby A Selvanathan, Saroja Selvanathan, Gerald Keller.  
Edition: 6th edition  
ISBN: 9780170237000 (paperback)  
Notes: Previous ed.: 2011  
Subjects: Commercial statistics. Managerial economics--Australia. Managerial economics--New Zealand. Management--Statistical methods. Economics--Statistical methods.  
Other Authors/Contributors: Selvanathan, Saroja. Keller, Gerald.  
Dewey Number: 519.5

#### Cengage Learning Australia

Level 7, 80 Dorcas Street  
South Melbourne, Victoria Australia 3205

#### Cengage Learning New Zealand

Unit 4B Rosedale Office Park  
331 Rosedale Road, Albany, North Shore 0632, NZ

For learning solutions, visit **[cengage.com.au](http://cengage.com.au)**

Printed in China by China Translation & Printing Services.  
1 2 3 4 5 6 7 17 16 15 14 13

# Brief contents

1	What is statistics?	1
2	Types of data, data collection and sampling	17
<b>PART 1 DESCRIPTIVE MEASURES AND PROBABILITY</b>		<b>41</b>
3	Graphical descriptive techniques – Nominal data	43
4	Graphical descriptive techniques – Numerical data	82
5	Numerical descriptive measures	128
6	Probability	203
7	Random variables and discrete probability distributions	251
8	Continuous probability distributions	302
<b>PART 2 STATISTICAL INFERENCE</b>		<b>339</b>
9	Statistical inference: Introduction	341
10	Sampling distributions	347
11	Estimation: Describing a single population	369
12	Estimation: Comparing two populations	421
13	Hypothesis testing: Describing a single population	455
14	Hypothesis testing: Comparing two populations	517
15	Inference about population variances	567
16	Analysis of variance	597
17	Additional tests for nominal data: Chi-squared tests	665
18	Simple linear regression and correlation	703
19	Multiple regression	767
20	Model building	825
21	Nonparametric techniques	870
22	Statistical inference: Conclusion	915
<b>PART 3 APPLICATIONS</b>		<b>959</b>
23	Time series analysis and forecasting	961
24	Index numbers	1013
25	Decision analysis – online chapter	25-1



# Contents

Preface	xiii
Resources guide	xvii
Acknowledgements	xxiii
About the authors	xxiv
<b>1</b> What is statistics?	<b>1</b>
Introduction to statistics	2
1.1 Key statistical concepts	4
1.2 Practical applications	5
Case 3.6 Differing average weekly earnings of men and women in Australia	6
Case 5.9 Aussies and Kiwis are leading in education	6
Case 15.1 Comparing salary offers for finance and marketing MBA majors - I	6
Case 17.1 Gold lotto	7
Case 18.1 Does unemployment affect the average hourly earnings of New Zealanders?	7
1.3 How managers use statistics	8
1.4 Statistics and the computer	9
1.5 Online resources	10
Appendix 1.A Introduction to Microsoft Excel	13
<b>2</b> Types of data, data collection and sampling	<b>17</b>
Introduction	18
2.1 Types of data	19
2.2 Methods of collecting data	24
2.3 Sampling	28
2.4 Sampling plans	30
2.5 Sampling and non-sampling errors	37
Summary	39
<b>PART 1 DESCRIPTIVE MEASURES AND PROBABILITY</b>	<b>41</b>
<b>3</b> Graphical descriptive techniques – Nominal data	<b>43</b>
Introduction	44
3.1 Graphical techniques to describe nominal data	45
3.2 Selecting a chart: Which chart is best?	53
3.3 Graphical techniques to describe ordinal data	59
3.4 Describing the relationship between two nominal variables	67
Summary	73
Case 3.1 Road fatalities in Australia by state	76
Case 3.2 Corporate tax rates around the world	77
Case 3.3 Revenues of the non-profit sector	78
Case 3.4 Where is the divorce rate heading?	78
Case 3.5 Geographic location of share ownership in Australia	79

Case 3.6	Differing average weekly earnings of men and women in Australia	79
Case 3.7	The demography of Australia	79
Case 3.8	Survey of graduates	80
Case 3.9	Overseas student contributions to Australian educational institutions	81
Case 3.10	Is the Australian overseas student market on the decline?	81
<b>4</b>	<b>Graphical descriptive techniques – Numerical data</b>	<b>82</b>
	Introduction	83
	4.1 Graphical techniques to describe numerical data	83
	4.2 Describing time-series data	104
	4.3 Describing the relationship between numerical two variables	108
	4.4 Graphical excellence and deception	116
	Summary	124
	Case 4.1 The question of global warming	126
	Case 4.2 Economic freedom and prosperity	127
	Case 4.3 An analysis of telephone bills	127
<b>5</b>	<b>Numerical descriptive measures</b>	<b>128</b>
	Introduction	129
	5.1 Measures of central location	130
	5.2 Measures of variability	142
	5.3 Measures of relative standing and box plots	159
	5.4 Approximating descriptive measures for grouped data	167
	5.5 Measures of association	169
	5.6 General guidelines on the exploration of data	185
	Summary	186
	Case 5.1 Return to the global warming question	191
	Case 5.2 Another return to the global warming question	191
	Case 5.3 Performance of major Australian companies	191
	Case 5.4 Consumer income versus expenditure	192
	Case 5.5 The gulf between the rich and the poor – life expectancy and literacy	193
	Case 5.6 Sydney and Melbourne leading the way in the growth in house prices	193
	Case 5.7 Performance of managed funds in Australia – 2-star rated versus 4-star rated funds	194
	Case 5.8 Life in suburbs drives emissions higher	195
	Case 5.9 Aussies and Kiwis are leading in education	196
	Appendix 5.A Summation notation	197
<b>6</b>	<b>Probability</b>	<b>203</b>
	Introduction	204
	6.1 Assigning probabilities to events	204
	6.2 Joint, marginal and conditional probability	215
	6.3 Rules of probability	225
	6.4 Probability trees	231
	6.5 Bayes' law	236
	6.6 Identifying the correct method	243
	Summary	244
	Case 6.1 Let's make a deal	247
	Case 6.2 University admissions in Australia: Does gender matter?	247

Case 6.3	Maternal serum screening test for Down syndrome	248
Case 6.4	Levels of disability among children in Australia	248
Case 6.5	Probability that at least two people in the same room have the same birthday	249
Case 6.6	Home ownership in Australia	249
<b>7</b>	<b>Random variables and discrete probability distributions</b>	<b>251</b>
	Introduction	252
	7.1 Random variables and probability distributions	252
	7.2 Discrete probability distributions	254
	7.3 Expected value and variance	261
	7.4 Bivariate distributions	268
	7.5 Applications in finance: Portfolio diversification and asset allocation	274
	7.6 Binomial distribution	281
	7.7 Poisson distribution	290
	Summary	296
	Case 7.1 Is there a gender disparity in cultural and leisure activities involvement of school children?	299
	Case 7.2 How about a carbon tax on motor vehicle ownership?	300
	Case 7.3 Internet usage by children	300
	Case 7.4 Gains from market timing	301
<b>8</b>	<b>Continuous probability distributions</b>	<b>302</b>
	Introduction	303
	8.1 Probability density functions	303
	8.2 Uniform distribution	306
	8.3 Normal distribution	308
	8.4 Exponential distribution	332
	Summary	337
<b>PART 2</b>	<b>STATISTICAL INFERENCE</b>	<b>339</b>
<b>9</b>	<b>Statistical inference: Introduction</b>	<b>341</b>
	Introduction	342
	9.1 Data type and problem objective	342
	9.2 How, when and why of statistical inference	344
	9.3 Systematic approach to statistical inference: A summary	345
<b>10</b>	<b>Sampling distributions</b>	<b>347</b>
	Introduction	348
	10.1 Sampling distribution of the sample mean $\bar{X}$	348
	10.2 Sampling distribution of the sample proportion $\hat{p}$	362
	10.3 From here to inference	366
	Summary	367
<b>11</b>	<b>Estimation: Describing a single population</b>	<b>369</b>
	Introduction	370
	11.1 Concepts of estimation	371
	11.2 Estimating the population mean $\mu$ when the population variance $\sigma^2$ is known	373

11.3	Estimating the population mean $\mu$ when the population variance $\sigma^2$ is unknown	387
11.4	Estimating the population proportion $p$	400
11.5	Determining the required sample size	406
11.6	Applications in marketing: Market segmentation	412
	Summary	416
	Case 11.1 Estimating average weekly poultry expenditure of New Zealanders	419
	Case 11.2 Cold men and cold women will live longer!	420
	Case 11.3 Super fund managers letting down retirees	420
<b>12</b>	<b>Estimation: Comparing two populations</b>	<b>421</b>
	Introduction	422
12.1	Estimating the difference between two population means $\mu_1 - \mu_2$ when the variances $\sigma_1^2$ and $\sigma_2^2$ are known: Independent samples	423
12.2	Estimating the difference between two population means $\mu_1 - \mu_2$ when the variances $\sigma_1^2$ and $\sigma_2^2$ are unknown: Independent samples	430
12.3	Matched pairs experiments: Dependent samples	440
12.4	Estimating the difference between two population proportions $p_1 - p_2$	444
	Summary	452
	Case 12.1 Australian newspapers are putting up a fight	454
	Case 12.2 Five-star hotels in reach of backpackers	454
<b>13</b>	<b>Hypothesis testing: Describing a single population</b>	<b>455</b>
	Introduction	456
13.1	Concepts of hypothesis testing	457
13.2	Testing the population mean when the variance $\sigma^2$ is known	465
13.3	The $p$ -value of a test of hypothesis	480
13.4	Testing the population mean when the variance $\sigma^2$ is unknown	490
13.5	Calculating the probability of a Type II error	496
13.6	Testing the population proportion	506
	Summary	513
	Case 13.1 Roy Morgan unemployment index	515
	Case 13.2 New Zealand National-led government gains ground	515
	Case 13.3 The republic debate: What Australian women are thinking	515
	Appendix 13.A Excel instructions	516
<b>14</b>	<b>Hypothesis testing: Comparing two populations</b>	<b>517</b>
	Introduction	518
14.1	Testing the difference between two population means: Independent samples	518
14.2	Testing the difference between two population means: Matched pairs experiment	537
14.3	Testing the difference between two population proportions	548
	Summary	558
	Case 14.1 Will the ALP government be re-elected in 2013?	562
	Case 14.2 Direct and broker-purchased mutual funds	562
	Case 14.3 Consumer confidence in New Zealand	563
	Case 14.4 New Zealand Government bond yields: short term versus long term	564

Case 14.5	The price of petrol: city versus country town	564
Case 14.6	Student surrogates in market research	565
Appendix 14.A	Excel instructions	566
<b>15</b>	<b>Inference about population variances</b>	<b>567</b>
	Introduction	568
	15.1 Inference about $\sigma^2$	568
	15.2 Inference about $\sigma_1^2/\sigma_2^2$	581
	Summary	594
Case 15.1	Comparing salary offers for finance and marketing MBA majors – I	595
Case 15.2	Comparing salary offers for finance and marketing MBA majors – II	595
Case 15.3	Risk of an asset: Five-year bonds versus 10-year bonds	596
<b>16</b>	<b>Analysis of variance</b>	<b>597</b>
	Introduction	598
	16.1 Single-factor analysis of variance: Independent samples (one-way ANOVA)	599
	16.2 Multiple comparisons	621
	16.3 Analysis of variance: Experimental designs	629
	16.4 Single-factor analysis of variance: Randomised blocks (two-way ANOVA)	630
	16.5 Two-factor analysis of variance	644
	Summary	658
Case 16.1	Top 300 Australian companies	661
Case 16.2	Effects of financial planning on small businesses	661
Case 16.3	Diversification strategy for multinational companies	662
Case 16.4	Comparing three methods of treating childhood ear infections	663
Case 16.5	Treatment of headaches	663
<b>17</b>	<b>Additional tests for nominal data: Chi-squared tests</b>	<b>665</b>
	Introduction	666
	17.1 Chi-squared test of a multinomial experiment	666
	17.2 Chi-squared test of a contingency table	676
	17.3 Chi-squared test for normality	689
	17.4 Summary of tests on nominal data	695
	Summary	697
Case 17.1	Gold lotto	701
Case 17.2	Which diets work?	701
Case 17.3	How well is the Australian Government managing the asylum seekers affair?	702
<b>18</b>	<b>Simple linear regression and correlation</b>	<b>703</b>
	Introduction	704
	18.1 Model	705
	18.2 Estimating the coefficients	707
	18.3 Error variable: Required conditions	720
	18.4 Assessing the model	722
	18.5 Using the regression equation	736
	18.6 Coefficients of correlation	741
	18.7 Regression diagnostics – I	749



Summary	758
Case 18.1 Does unemployment affect the average hourly earnings of New Zealanders?	764
Case 18.2 Does domestic market capital influence stock prices?	765
Case 18.3 Does unemployment affect inflation in New Zealand?	765
Case 18.4 Cengage Learning Australia	765
Case 18.5 Do higher interest rates reduce the availability of new housing?	765
Case 18.6 Market model of share returns	766
Case 18.7 Life insurance policies	766
<b>19</b> Multiple regression	767
Introduction	768
19.1 Model and required conditions	768
19.2 Estimating the coefficients and assessing the model	770
19.3 Regression diagnostics – II	795
19.4 Regression diagnostics – III (time series)	807
Summary	818
Case 19.1 Are lotteries a tax on the poor and uneducated?	823
Case 19.2 Demand for beer in Australia	823
Case 19.3 Cengage Learning Australia revisited	823
Case 19.4 Average hourly earnings in New Zealand	824
Case 19.5 Testing a more effective device to keep arteries open	824
<b>20</b> Model building	825
Introduction	826
20.1 Polynomial models	826
20.2 Nominal independent variables	837
20.3 Applications in human resources: Pay equity	846
20.4 Logistic regression	851
20.5 Variable selection methods	858
20.6 Stepwise regression	861
20.7 Model building	865
Summary	867
Case 20.1 Does the street number of a house matter?	868
Case 20.2 Track and field performance forecasts	868
<b>21</b> Nonparametric statistics	870
Introduction	871
21.1 Wilcoxon rank sum test	872
21.2 Sign test and Wilcoxon signed rank sum test	884
21.3 Kruskal–Wallis test and Friedman test for independent samples	898
Summary	909
Case 21.1 HR Information Systems Inc.	913
Case 21.2 Bank of Commerce customer survey	913
<b>22</b> Statistical inference: Conclusion	915
Introduction	916
22.1 Identifying the correct technique: Summary of statistical inference	916
22.2 Beyond the statistics subject – the last word	942
Case 22.1 Do banks discriminate against women business owners? – Part I	953
Case 22.2 Do banks discriminate against women business owners? – Part II	954

Case 22.3	Effect of the death of key executives on stock market returns	954
Case 22.4	Northern Territory statehood	955
Case 22.5	Graduate survey report	956
Case 22.6	Evaluation of a new drug	956
Case 22.7	Nutrition education programs	957
Case 22.8	Testing a more effective device to keep arteries open	958
Case 22.9	Type A, B and C personalities and job satisfaction and performance	958
<b>PART 3</b>	<b>APPLICATIONS</b>	<b>959</b>
<b>23</b>	<b>Time series analysis and forecasting</b>	<b>961</b>
	Introduction	962
	23.1 Components of a time series	963
	23.2 Smoothing techniques	966
	23.3 Trend analysis	976
	23.4 Measuring the cyclical effect	982
	23.5 Measuring the seasonal effect	985
	23.6 Introduction to forecasting	994
	23.7 Time series forecasting with exponential smoothing	996
	23.8 Time series forecasting with regression	999
	Summary	1009
	Case 23.1 Part-time employed females	1011
	Case 23.2 Australian tourism: Bed occupancy, New South Wales (NSW)	1012
	Case 23.3 Seasonal and cyclical effects in houses constructed in Victoria	1012
<b>24</b>	<b>Index numbers</b>	<b>1013</b>
	Introduction	1014
	24.1 Constructing unweighted index numbers	1015
	24.2 Constructing weighted index numbers	1021
	24.3 The Australian consumer price index (CPI)	1026
	24.4 Using the CPI to deflate wages and GDP	1029
	24.5 Changing the base period of an index number series	1033
	Summary	1036
	Case 24.1 Soaring petrol prices	1040
	Case 24.2 Australian road deaths in 2011 the lowest in almost 60 years	1040
<b>25</b>	<b>Decision analysis – online chapter</b>	<b>25-1</b>
	Introduction	25-2
	25.1 Decision problem	25-2
	25.2 Using additional information	25-8
	Summary	25-20
	Appendix A: Summary solutions for selected exercises	1041
	Appendix B: Tables	1057
	Glossary	1083
	Index	1089



# Preface

Business is complex and requires effective management to succeed. Managing complexity requires many skills. There are more competitors, more places to sell products and more places to locate workers. As a consequence, effective decision making is more crucial than ever before. On the other hand, nowadays managers have more access to larger and more detailed data that are potential sources of information for making well-informed objective decisions. However, to achieve this, potential managers need to know which statistical techniques they should use to extract useful information from the data available to them and make informed decisions. For students preparing for the business world, it is not enough to focus merely on mastering a diverse set of statistical techniques and calculations. A course and its recommended textbook must provide a complete picture of statistical concepts and their applications to the real world. *Business Statistics – Australia and New Zealand* is designed to demonstrate that statistical methods are vital tools for today's businesses and managers to improve their decision-making skills.

This book is a thorough Australasian adaptation of the most popular and best-selling United States (US) text, *Statistics for Management and Economics* (9th edition) by Gerald Keller. This edition is a further attempt to make the basic business and economics statistics subject a more effective and enjoyable learning experience for both instructors and students at Australasian universities. It uses familiar local terminology, together with examples, exercises and cases which draw upon Australasian data. To enhance flexibility, we have also rearranged a number of chapters from the US edition. For example, we have incorporated the data collection chapter with types of data at the start of the book, introduce estimation and hypothesis testing in separate chapters, present inference about population variance in another chapter and single population and two or more populations in different chapters. Furthermore, we have included a chapter on index numbers which includes some important topics such as the construction of the Australian Consumer Price Index, as well as comparison of Laspeyres and Paasche index numbers, etc.

When solving problems, *Business Statistics – Australia and New Zealand* uses its unique 'ICI' approach which is renowned for its consistent, proven three-step method to solving problems. The 'Identify, Compute and Interpret' approach teaches you how to determine the appropriate technique, how to compute the statistics and how to interpret the results in the context of the problem at hand. The *compute* stage can be completed either manually (with the aid of a calculator) or using Excel (on the computer), or both.

Today, most subjects use the computer and statistical software or spreadsheets. This book contains step-by-step instructions and commands, with screen images, to teach students how to use Microsoft Excel® to solve statistical problems. Additionally, most examples, exercises and cases feature raw data. These data sets are available to download from the **CourseMate for Business Statistics** website, accessible through <http://login.cengage.com> using the access card that comes with each new copy of this book.

## Key features of our approach

### 1. Systematic approach

This edition retains the systematic approach introduced in the US edition, which teaches students how to recognise which statistical technique to use. We believe that this skill is

the most important one to develop, yet it is the one students have the greatest difficulty in mastering. As each technique is introduced, we demonstrate how to recognise when its use is appropriate and when it is not. A feature of this approach is the review chapter (Chapter 22) presented in the book that allows students to hone their technique-selection skills. In the review chapter, a flowchart develops the logical process for choosing the correct technique. Our ICI approach divides the solution of statistical problems into three parts: (1) identify the technique; (2) calculate/compute the required sample statistics; and (3) interpret the results. Our focus has been on the first and third parts, as the sample statistics could be produced relatively easily with a computer.

When demonstrating examples, we start the solutions by reviewing the appropriateness of the method to be used. One of the main benefits of our approach is that it allows instructors to de-emphasise mathematical manipulation. Consequently, students can spend more time properly setting up the procedure and interpreting the statistical results, and less time grinding out the arithmetic.

For students without access to a computer and statistical software, we continue to teach how to calculate statistics manually (with the exception of the most complicated procedures), and most exercises can be solved in this way.

## 2. Cases

Recent academic conferences devoted to improving the teaching of applied statistics have advocated the use of cases to help motivate students. In practice, a statistician often has access only to raw data and the correct procedure to employ is not obvious; our approach allows us to offer more realistic applications. In fact, many of the cases are based on real studies that have been reported in newspapers, magazines, journals, on television and at academic conferences. Several from our own consulting projects have also been included. Such applications can motivate students, who unfortunately often believe that statistics is not very relevant to their future careers. We believe that our approach can change these attitudes. More than 80 cases are included in the book. Students are expected to analyse the cases and draw conclusions in the same way as the original authors did. These cases are neither summaries of what a particular statistician did to solve a problem, nor glorified exercises; rather, they give students the opportunity to see for themselves how statistical problem solving works.

## 3. Review chapter

The review chapter is included in the book to help students practise identifying the correct techniques. This chapter reviews all the statistical methods covered in the book and provides exercises and cases that require the use of several different statistical procedures. It, therefore, provides practice in the technique identification skills that are required for statistics exams and, ultimately, in any real-life application of statistics.

## 4. Use of Excel

Because the use of spreadsheets is so widespread, we believe that Microsoft® Excel is an important addition to this book. However, spreadsheets are not designed for use as statistical software, although they are increasingly capable in data analysis. Because of this limitation, we offer *Data Analysis Plus*® macros and workbooks that can be added to the Excel menu and used to solve sophisticated statistical problems beyond Excel's existing capabilities. Excel macros are created to complement Excel's menu of statistical procedures. All statistical techniques introduced in this book can be calculated using either Excel's *Analysis ToolPak* or

*Data Analysis Plus 9.0*<sup>®</sup> which is compatible with Excel 2010. These macros are available to download from the **CourseMate for Business Statistics** website, accessible through <http://login.cengagebrain.com> using the access card that comes with each new copy of this book. The CD-ROM also contains Version 7.0 of *Data Analysis Plus*<sup>®</sup> which is compatible with previous versions of Excel.

The Excel spreadsheet package is used extensively and presented consistently throughout the book to calculate sample statistics. Most examples in the chapters present manual and using the computer (Excel) solutions, allowing students to see both methods together and to use the preferred method. This feature provides flexibility, allowing the instructor to decide when manual or computer calculations should be emphasised. Detailed instructions and Excel commands provided for the examples make it easy for instructors and students to make use of the computer. They also eliminate the need for instructors to teach how to use the software.

Data files are provided in Excel format for most of the examples, exercises and cases. The sixth edition includes hundreds of data files, some consisting of thousands of observations, which emphasise a central theme in the book – statistical techniques convert data into information. For students who will conduct statistical analyses manually, we have also provided the summary statistics (e.g. means and variances) for exercises, allowing most exercises to be solved manually.

## 5. Exercises

There are over 1500 exercises of varying levels of difficulty. At the end of most sections we supply, under the heading ‘Learning the techniques’, exercises that help students to learn the arithmetic involved in a specific procedure. ‘Applying the techniques’ exercises then stress when and why the technique is used and how the results assist in the decision-making process. ‘Computer applications’ help students gain hands-on experience in applying the techniques to solve problems using real-world data and Microsoft<sup>®</sup> Excel. Supplementary exercises appear at the end of each chapter. As they cover all the topics presented in that chapter, they allow students to practise identifying which of the techniques encountered in that chapter should be employed. They also tend to be more realistic than the other types of exercises.

We are optimistic that the systematic approach used in this book will be successful in helping students to understand how, when and why statistics are used. We hope that the realistic examples, exercises and cases we present, wherever possible with Australasian data, will make the subject more interesting and will persuade students that statistics can play a vital role in managerial decision making.

This text is suitable for a one- or two-semester subject in a business program. Although various sections can be omitted, we strongly urge instructors to attempt to complete most of the statistical inference part of the book. Like a house under construction, the structure of the systematic approach is stronger when most of the various components are in place. Nonetheless, the book has been designed so that chapters can be omitted relatively easily.

## Unique features

- *Chapter opening examples* illustrate the use of techniques introduced in that chapter. These examples are designed to help students learn the concepts in the chapters. These chapter opening examples are revisited at the relevant section of the chapter, where they are solved.
- *‘Seeing statistics’ boxes* refer to applets that are introduced to illustrate statistical concepts visually. They are adapted from *Seeing Statistics* by Gary McClelland and customised to

this text. Nineteen Java applets with a total of 72 applet exercises are available to download from the **CourseMate for Business Statistics** website, accessible through <http://login.cengagebrain.com> using the access card that comes with each new copy of this book.

- *In addition to the examples provided in each chapter*, in this edition we have included 'Real-life applications' sections which illustrate the fundamental applications of statistics in finance, marketing, human resources management, operations management, accounting and economics.
- *'Key concept' boxes* are included after each technique has been introduced. These boxes will allow students to see a technique's essential requirements, in addition to giving them a way to easily review their understanding. This will be further enhanced by the summary flowchart presented in the review chapter.
- *Several new exercises are added to each chapter*. In particular, new applied exercises are added to the chapter on probability to help students visualise and understand concepts.
- *Several new data sets* have been added to the existing computer exercises section of each chapter. For those students who wish to solve the computer exercises containing data sets manually, summary statistics to these data sets are provided within each exercise.
- *Improved and expanded Data Analysis Plus® add-ins for Excel* have been incorporated.
- *In addition to the applets*, we have included several Excel workbooks that feature worksheets for confidence interval estimators and test statistics. By changing one or more inputs, students can learn, for example, the effect of increasing sample sizes on confidence intervals or on test statistics.
- Appendix A: Summary solutions for selected exercises is available in the book.

# Resources guide

As you read this text you will find a number of features in every chapter to enhance your study of **Business Statistics 6e** and help you understand how the theory is applied in the real world.

**Part opening paragraphs** introduce the chapters in each Part to give you an overview of how the chapters relate to each other.

**Learning objectives** give you a clear sense of what each chapter will cover and what you should be able to do after reading the chapter.

**PART ONE**  
Descriptive measures and probability

Chapter 3 Graphical descriptive techniques - Nominal data  
Chapter 4 Graphical descriptive techniques - Numerical data  
Chapter 5 Numerical descriptive measures  
Chapter 6 Probability  
Chapter 7 Random variables and discrete probability distributions  
Chapter 8 Continuous probability distributions

To help you organise the material that you are about to learn, we have divided the rest of the book into three parts.  
Part 1 covers descriptive statistics and probability. These topics constitute the foundation of statistical inference. Chapter 3 introduces the graphical techniques for nominal data and Chapter 4 deals with graphical techniques for numerical data. Chapter 5 presents numerical measures, which are used to summarise data. The summary measures introduced in Chapter 5 will be used to make inferences about parameters in later chapters. Chapters 6 to 8, on present probability and probability distributions that will provide the link between sample statistics and population parameters.  
Everything we do in this book is really built upon these six chapters. However, Part 1 does much more than just lay the foundation. Each chapter on statistics and probability are subjects that are worth learning for their own intrinsic values.  
We all make decisions on a daily basis, most of which are made under uncertainty. Consider an investor who must decide which investment to make, how much money to invest and for how long that investment should be held. There are a large number of events over which the investor has no control. All that the investor can do is attempt to assess the risks and returns associated with each investment. As you will discover, probability plays a central role in this assessment.  
We believe that all business and economics graduates will have many opportunities to apply statistical inference techniques and concepts. However, not all of them will do so because of a lack of other knowledge (despite the best efforts of statistics lecturers) or confidence. Descriptive techniques are so common that it is virtually impossible to ignore them. Newspapers, magazines, company annual reports and presentations are filled with applications of descriptive statistics. Knowing how to use and interpret them is a critical skill for all of us.

**2**  
Types of data, data collection and sampling

**LEARNING OBJECTIVES**  
This chapter discusses the types of data and various methods of collecting data. At the completion of this chapter, you should be able to:

- LO1 describe different types of data
- LO2 understand the primary and secondary sources of statistical data
- LO3 explain the various methods of collecting data
- LO4 explain the basic sampling plans
- LO5 identify the appropriate sampling plan for data collection in a particular experiment
- LO6 understand the main types of errors involved in sampling.

**LEARNING OBJECTIVES**  
This chapter discusses the types of data and various methods of collecting data. At the completion of this chapter, you should be able to:

- LO1 describe different types of data
- LO2 understand the primary and secondary sources of statistical data
- LO3 explain the various methods of collecting data
- LO4 explain the basic sampling plans
- LO5 identify the appropriate sampling plan for data collection in a particular experiment
- LO6 understand the main types of errors involved in sampling.

**Chapter outlines** at the start of each chapter demonstrate the key concepts that will be covered.

**1**  
What is statistics?

**CHAPTER OUTLINE**  
Introduction to statistics  
1.1 Key statistical concepts  
1.2 Practical applications  
1.3 How managers use statistics  
1.4 Statistics and the computer  
1.5 Online resources

**LEARNING OBJECTIVES**  
After reading this chapter, you should be able to:  
LO1 describe the two major branches of statistics – descriptive statistics and inferential statistics.  
LO2 understand the key statistical concepts – population, sample, parameter, statistics, standard deviation.  
LO3 provide examples of practical applications in which statistics have a major role to play.  
LO4 understand how statistics are used by business managers.  
LO5 understand the basics of the computer spreadsheet package Microsoft Excel and its application in doing with statistical data analysis for large amounts of data.

**CHAPTER OUTLINE**  
Introduction to statistics  
1.1 Key statistical concepts  
1.2 Practical applications  
1.3 How managers use statistics  
1.4 Statistics and the computer  
1.5 Online resources

**Census in Australia**  
Introduction to data gathering for various characteristics of different populations of interest are an important part of statistics. When we use all units of the population to record information about the characteristics of interest, this type of data gathering is called a census. One to one and other

**Real World Examples** are used at the beginning of every chapter. They highlight a specific problem that can be solved using the statistical techniques that will be covered in the chapter. The problem is answered later in the chapter.

**Income and its allocation to food**

One of the most important empirical regularities in consumption economics is Engel's law. The law states that the proportion of income allocated to food falls with increasing income. This means that a rich consumer would spend a lesser proportion of his/her income on food than a poor consumer, who would allocate a greater proportion of his/her income to food.

Two statisticians, Working (in 1942) and Leser (in 1963), modelled Engel's law into a linear regression framework. In order to investigate this law, we need to develop the relationship between income and the proportion of income allocated to food. We have collected data from 48 countries, which are stored in the **DATAFILE**. After we have presented the statistical technique, we will return to this problem and solve it (see page 76).

**SSA envelope plan**

An express courier service needs to determine the number of packages to be delivered within 30 days. The list lists an address, and customers are expected to use their own envelopes to return their payments. Currently the mean and standard deviation of the amount of time taken to pay bills are 64 days and 8 days respectively. The ideal financial officer (CFO) believes that including a stamped self-addressed (SSA) envelope would decrease the number of bills. She calculates that the expected cost from a one-day decrease in the payment period would pay for the costs of the envelopes and stamps. Any further decrease in the payment period would generate a profit. To test her belief, she randomly selects 100 customers and includes an SSA envelope with their invoice. The number of days until payment is received are stored in the **DATAFILE/SSA**. Can the CFO conclude that the plan will be profitable?

After we have introduced the topic of tests, we will answer this question (see page 525-526).

**Introduction**

In Chapters 11 and 12, we introduced estimation and showed how it is used. Now we're going to present the second general procedure of making inferences about a population – hypothesis testing. The purpose of this type of inference is to determine whether enough statistical evidence exists to enable us to conclude that a belief or hypothesis about a parameter is supported by the data. You will discover that hypothesis testing has a wide variety of applications in business and economics, as well as many other fields. This chapter will lay the foundation upon which the rest of the book is based. As such it represents a critical contribution to your development as a statistician. In the next section, we will introduce the concepts of hypothesis testing, and in Section 11.2 we will develop the method employed to test a hypothesis about a population mean when the data type is numerical and the population standard deviation is known and extend the analysis to real situations where the population standard deviation is usually unknown in Section 11.4. In Section 11.6, we deal with hypothesis testing when the data type is nominal. The rest of the chapter deals with related topics.

**Hypothesis**  
A proposition or conjecture which the statistician will test by means called hypothesis testing.

**Margin definitions** are key terms, which are defined and included in the margins for easy reference.

**Hypothesis**  
A proposition or conjecture which the statistician will test by means called hypothesis testing.

**Interesting Examples** are used throughout each chapter. These are designed to teach you to use the authors' unique three-step approach to problem solving, and to help you apply statistics to real business problems.

**Identifying the technique**  
The problem objective is to describe the population of Australia. The parameter we wish to estimate is the proportion  $p$  of adults concerned about healthy eating. The confidence interval estimate is

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

**Calculating manually**  
To solve manually, we count the number of 1s in the file. We find

$$\hat{p} = \frac{X}{n} = \frac{269}{1250} = 0.2152$$

$$\hat{q} = 1 - \hat{p} = 1 - 0.2152 = 0.7848$$

**Interpreting the results**  
We estimate that the proportion of Australian adults who are concerned about healthy eating is 0.2152. Because there are 15 834 856 adults in the population, we estimate that the number of adults in the group who are concerned about healthy eating falls between

$$LCL = 15\,834\,856(0.1924) = 3\,046\,626$$

$$UCL = 15\,834\,856(0.2382) = 3\,771\,626$$

**Using the computer**  
The Excel commands are the same as in Example 11.4.

**Excel output for Example 11.7**

	A	B
1	z-estimate of a Proportion	
2		

You will learn to:

- **identify** the right statistical technique to use by focusing on the relationship between the problem and data type
- **compute** the answer either by hand calculation, or by calculating it in Microsoft Excel® when you see a 'Using the computer' heading with instructions and 'commands' boxes, which included step-by-step instructions on how to complete the examples using Microsoft Excel®
- **interpret** the answer in the context of the problem.



**Factors that identify** boxes are included after each technique has been introduced to allow you to appreciate a technique’s essential requirements, and to enable you to review your understanding of each technique.

**Factors that identify the z-interval estimator of  $\mu$**

- 1 Problem objective:** to describe a single population
- 2 Data type:** numerical (quantitative)
- 3 Type of descriptive measurement:** central location
- 4 Population variance:** known

**Interpreting the confidence interval estimate**

In Example 11.1, we found the 90% confidence interval estimate of the average front number of hours that children spend watching television per week to be LCI = 25.623 and UCI = 28.776. Now, people commonly interpret this interval to mean that there is a 90% probability that the population mean lies between 25.623 and 28.776. This interpretation is incorrect because it implies that the population mean is a variable about which we can make probability statements. In fact, the population mean is a fixed but unknown quantity.

**Learning the techniques**

**3.31 XRO3-31** The following table summarises the data from a survey on the ownership of iPads among females with different levels of income. Determine whether the two normal variables are related.

Income	Yes	No
Low	60	32
High	30	48

**Applying the techniques**

**3.32 XRO3-32** Self-correcting exercise. The trustee of a company's superannuation scheme has selected the contents of a sample of the company's employees' responses to a questionnaire. The responses are summarised in the following table. Use an appropriate statistical test to determine whether the responses differ among the three employees.

Employee	Yes	No
John	67	33
Agnes	53	47

**Computer applications**

**3.33 XRO3-33** The associate dean of a business school was looking for ways to improve the quality of responses to the MBA application form. In particular, she wanted to find the undergraduate students of her school and the three years with MBA programs (with 200 applicants of her program).

The text contains over 1500 **exercises**, located at the end of each section in chapters.

These include:

- **Learning the techniques:** exercises that help you to learn the arithmetic involved in a specific procedure.
- **Applying the technique:** exercises that highlight when and why the techniques are used and how the results assist in the decision-making process.
- **Computer applications:** help you gain hands-on experience in applying the techniques to solve problems using real world data and Microsoft Excel®.

**Data files**, highlighted throughout the text, enable you to complete the examples, exercises and case studies in the text without having to spend time inputting raw data. These data files are available on the accompanying **CourseMate for Business Statistics** website (<http://login.cengagebrain.com>).

**96 PART 1 DESCRIPTIVE**

**EXAMPLE 4.3**

**Business statistics marks - manual versus computer calculations**

**OBJECTIVE** The final marks in a statistics course that emphasised mathematical proofs, derivations and manual (hand) calculations, both during the class and on exams, are listed below. The marks obtained by students in the same course after the emphasis was changed to applications with most of the calculations performed using a computer, are also listed below. Draw histograms for both groups and interpret the results.

Manual	Computer
77	87
78	88
79	89
80	90
81	91
82	92
83	93
84	94
85	95
86	96
87	97
88	98
89	99
90	100

Note that Excel puts the upper limit of the class interval at the midpoint of the class. For example, 80 in the histogram refers to the class interval 75 < x < 85.

**Interpreting the results**

As the number of observations is the same for both groups of marks, the absolute frequency histograms can be used for comparison. The histogram of the marks in the 'manual' statistics course is bimodal. The larger modal class consists of the marks in the 75s. The smaller modal class contains the marks in the 80s.

**Real-Life Applications** features are included throughout the text to demonstrate real-world applications of statistics in the areas of finance, marketing, human resource management, accounting and economics.

**REAL-LIFE APPLICATIONS**

**Energy economics**

One variable that has had a large influence on the economies of virtually every country is energy. The 1973 oil crisis in which the price of oil quadrupled over a short period of time is generally considered to be one of the largest financial shocks to the world's economies. In fact, economists often refer to two different economies: before the 1973 oil crisis and after.

Unfortunately, the world will be facing more shocks because of energy for two primary reasons. The first is the depletion of non-renewable sources of energy and the resulting price increases. The second is the possibility that burning fossil fuels and the creation of more carbon dioxide may be the cause of global warming. One economist predicted that the cost of global warming will be calculated in the trillions of dollars. Statistics can play an important role by determining whether Earth's temperature has been increasing and, if so, whether carbon dioxide is the cause (see Case 4-1).

Source	Consumption (TWh)	Percentage (%)
Non-renewable energy sources	100254	94.97
Coal and coal products	40258	4.17
Natural gas	28276	3.43
Oil	0	0
Renewable energy sources	5522	5.23
Hydroelectricity	1522	1.45
Solar	332	0.31
Other (small biomass, geothermal, solar, wind, tide, wave and coast)	548	0.52
Total	105776	100.00

**Seeing statistics** boxes also appear throughout the text. These direct you to the accompanying **CourseMate for Business Statistics** website (<http://login.cengagebrain.com>) where you can view applets, which enhance your comprehension by providing visual insight into statistical concepts.

**Commands** boxes can be found in chapters and include step-by-step instructions on how to complete exercises in Microsoft Excel®.

**COMMANDS**

**Commands for Example 14.1**

- 1 Type or import the data into two columns. Open file **XM14-01**
- 2 Click **Data, Data Analysis** and **z-Test: Two-Sample for Means**. Click **OK**.
- 3 Specify **Variable 1 Range**: **A1:A101**
- 4 Specify **Variable 2 Range**: **B1:B101**
- 5 Specify the **Hypothesised Mean Difference**: **0**
- 6 Specify **Variable 1 Variance (known)**: **22467600**
- 7 Specify **Variable 2 Variance (known)**: **28783225**
- 8 Click **Labels** (if necessary).

Specify a value for  $\alpha$  (Alpha) and click **OK**: **0.05**

Alternatively, if the sample means are already known (given or computed), use the **z-test: 2 Means** worksheet in the **Test Statistics** workbook available from the **CourseMate for Business Statistics** website (accessible through <http://login.cengagebrain.com>) and enter the required sample means, sample sizes and hypothesised difference in means.

**SEEING STATISTICS**

**Normal distribution parameters**

This applet can be used to see the effect of changing the values of the mean and standard deviation of a normal distribution.

Move the top slider left or right to decrease or increase the mean of the distribution. Notice that when you change the value of the mean, the shape stays the same, only the location changes.

Move the second slider to change the standard deviation. The shape of the bell curve is changed when you increase or decrease the standard deviation.

**Applet exercises**

- 4.1 Move the slider bar for the standard deviation so that the standard deviation of the red distribution is greater than 1. What does this do to the spread of the normal distribution? Does it squeeze it or stretch it?
- 4.2 Move the slider bar for the standard deviation so that the standard deviation of the red distribution is less than 1. What does this do to the spread of the normal distribution? Does it squeeze it or stretch it?
- 4.3 Move the mean and the standard deviation sliders so that the red distribution is different from the blue distribution. What would you need to subtract from the red values to slide the red distribution back (forward) so that the centres of the red and blue distributions would overlap? By what would you need to divide the red values to squeeze or stretch the red distribution so that it would have the same spread as the blue distribution?

**SUMMARY**

This chapter dealt with continuous random variables and their distribution. This chapter dealt with continuous random variables and their distribution. This chapter dealt with continuous random variables and their distribution.

**SYMBOLS**

Symbol	Pronounced	Represents
$e$		2.71828 ...
$\pi$	pi	3.14159 ...

**SUMMARY OF FORMULAS**

Standardised normal random variable	$Z = \frac{X - \mu}{\sigma}$
Mean	$\mu = 0$
Variance	$\sigma^2 = 1$
Exponential	$f(x) = \lambda e^{-\lambda x}$
Probability	$P(X < x) = 1 - e^{-\lambda x}$
Mean	$\frac{1}{\lambda}$
Variance	$\frac{1}{\lambda^2}$

**IMPORTANT TERMS**

- continuity correction factor 326
- continuous random variable 303
- exponential distribution 332
- exponential random variable 332
- normal random variable 303
- probability density function 303
- standard normal random variable 303
- uniform distribution 303

At the end of each chapter you will find:

- a **summary** section that consolidates your knowledge of the content of the chapter by reviewing key concepts and drawing out their wider significance,
- a recap of relevant **symbols**,
- a **summary of formulas** from the chapter, and
- a list of **important terms**.

**Supplementary exercises** at the end of each chapter give you the opportunity to further test your understanding of the key concepts covered.

**CHAPTER 7 Random variables and discrete probability distributions**

**739** A study of drivers reveals that, when lost, 45% will stop and ask for directions, 32% will consult a map and 23% will continue driving until the location has been determined. Suppose that a sample of 200 drivers was asked to report what they do when lost. Find the following probabilities.

- At least 100 stop and ask directions.
- At most 50 continue driving.
- Between 50 and 75 (inclusive) consult a map.

**740** The scheduling manager for an electricity supply company knows that there are an average of 12 emergency calls regarding power failures per month. Assume that a month consists of 30 days.

- Find the probability that the company will receive at least 12 emergency calls during a specified month.
- Suppose that the company can handle a maximum of three emergency calls per day. What is the probability that there will be more emergency calls than the company can handle on a given day?

**CASE STUDIES**

**CASE STUDIES**

**Is there a gender disparity in cultural and leisure activities involvement of school children?**

**02701** In addition to involvement in sporting activities, school children these days are also involved in participating in cultural and leisure activities outside school hours. To analyse the extent of their involvement in Australia, data on the participation of children aged 5 to 14 years in selected organised cultural and leisure activities outside of school hours during the 12 months prior to interview in April 2012) was collected. The selected cultural activities include playing a musical instrument, singing, dancing.

No. activities	12051	716.2	1625.4
One (1)	274.5	452.0	897.5
Two (2)	55.0	147.9	107.9
Three (3)	16.8	39.0	39.0
Four or five (4, 5)	5.0	14.0	10.0
Total population aged 5-14 years	442.4	1262.9	2798.5

Source: Department of Education, Children's Participation in School and Leisure Activities, Australia, Table 5, October 2012. Use file: 060110\_002\_Catalogue

**Appendices** throughout the text included step-by-step instructions on how to perform complex statistical calculations.



**Online resources**

Visit <http://login.cengagebrain.com> and login using the code card in the front of this text for a 12-month access to the *Business Statistics 6e CourseMate* website. You'll find interactive self-assessments, quizzes, glossary, flashcards, crosswords, case questions, and more tools to help you excel in your studies.

**CHAPTER 10 STATISTICAL INFERENCE**

**SUPPLEMENTARY EXERCISES**

The following exercises require the use of a computer and software. Use a 5% significance level.

**16.59 XR16-59** The editor of the student newspaper was in the process of making some major changes in the newspaper's layout. He was also contemplating changing the typface of the print used. To help himself make a decision, he set up an experiment in which 20 individuals were asked to read four newspaper pages, with each page printed in a different fontface, at the reading speed

the data listed below. Do these data provide enough evidence to conclude that there are differences in attention span while reading the three products advertised?

Fontface	Product A	Product B	Product C
1	28	40	30
2	36	28	24
3	27	35	—
4	—	—	—
5	29	45	30

**16.60 XR16-60** Upon reexamining the experiment in Exercise 16.59, the participant decides that a child's age may influence the attention span. Consequently, the experiment is rerun in the following way. There are 10 year-olds, three five-year-olds, three eight-year-olds, three seven-year-olds, three six-year-olds, three four-year-olds and three two-year-olds who are randomly assigned to watch one of the commercials, and their attention spans are measured. The data are shown below. Do the results indicate that there are differences in the abilities of the products advertised to hold children's attention?

Age	Commercial A	Commercial B	Commercial C
10	45	43	39
8	44	42	39
7	38	41	37
6	33	41	35
5	33	37	30
4	27	35	30

**16.61** Refer to Exercise 16.60.

If differences exist, which of the following is true?

- Small firms differ from the other two.
- Mid-sized firms differ from the other two.
- Large firms differ from the other two.
- All three firms differ from one another.
- Small firms differ from larger firms.

**Case studies** are included at the end of each chapter to assist you in applying the statistical techniques you are learning to real-world problems

**APPENDIX 5.A**

**Summation notation**

This appendix offers an introduction to the use of summation notation. Because summation notation is used extensively throughout statistics, you should review this appendix even if you have had previous exposure to summation notation. Our coverage of the topic begins with an introduction to the necessary terminology and notation, follows with some examples, and concludes with four rules that are useful in applying summation notation.

Consider  $n$  numbers  $x_1, x_2, \dots, x_n$ . A concise way of representing their sum is

$$\sum_{i=1}^n x_i$$

That is,

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$$

**Terminology and notation**

- The symbol  $\Sigma$  is the capital Greek letter sigma, and means 'the sum of'.
- The letter  $i$  is called the *index of summation*. The letter chosen to represent the index of summation is arbitrary.
- The expression  $\sum_{i=1}^n x_i$  is read 'the sum of the terms  $x_i$  where  $i$  assumes the values from 1 to  $n$  inclusive'.
- The numbers 1 and  $n$  are called the lower and upper limits of summation respectively. Summation notation is best illustrated by means of examples.

**Examples**

1 Suppose that  $x_1 = 5$ ,  $x_2 = 6$ ,  $x_3 = 8$ , and  $x_4 = 10$ . Then:

- $\sum_{i=1}^4 x_i = x_1 + x_2 + x_3 + x_4$   
 $= 5 + 6 + 8 + 10$   
 $= 29$
- $\sum_{i=2}^4 x_i = x_2 + x_3 + x_4$   
 $= 6 + 8 + 10$   
 $= 24$

## For the Instructor

Cengage Learning is pleased to provide you with a selection of resources that will help you prepare for your lectures. These teaching tools are available on the companion website accessible via <http://login.cengage.com>.

### Instructor Solutions Manual

This includes complete solutions to all the end-of chapter conceptual questions and problems in the text. Also provided are full answers to the online student chapter review quizzes.

### PowerPoint™ presentations

Chapter-by-chapter PowerPoint presentations cover the main concepts addressed within the text and can be edited to suit your own requirements. Use these slides to enhance your lecture presentations and to reinforce the key principles of your subject, or for student handouts.

### ExamView® test bank

ExamView helps you to create, customise and deliver tests in minutes for both print and online applications. The Quick Test Wizard and Online test Wizard guide you step by step through the test-creation process. With ExamView's complete word-processing abilities, you can add an unlimited number of new questions to the bank, edit existing questions and build tests of up to 250 questions using up to 12 question types. You can also export the files into Blackboard or WebCT.

**ExamView®**

### Artwork

These digital files of graphs, tables, pictures and flow charts from the text can be used in a variety of media. Add them into your course management system, use them within student handouts or copy them into lecture presentations.

### APLIA

Online homework has never been easier! APLIA is the perfect solution for your homework management needs, with assignable questions and an online gradebook. Once set up by instructors, students can access additional online review questions related to the text.

 **aplia™**



# Acknowledgements

We are grateful to the publishing, editorial and production team at Cengage Learning for their help and support. We are particularly grateful to Cengage Learning's Senior Publishing Editor, Geoff Howard, for his patience, guidance and encouragement; to Sylvia Marson, whose firm hand on the editorial process kept us on task; to Emily Spurr and Kylie McInnes for developmental work; and to Nathan Katz, whose production process converted our manuscript into the polished final product.

We would like to thank the many colleagues, survey participants, reviewers and students for their helpful suggestions, comments and criticisms on various chapters of the book, which have contributed in improving this edition and the previous editions. In particular, we would like to thank Professor Prasada Rao and Dr Mohammed Alauddin, University of Queensland; Professor Ken Clements, University of Western Australia; Professor Bill Griffiths, Melbourne University; Dr Patti Cybinski, Associate Professor Helen Higgs, Dr Suri Rajapakse, Dr Michael Kremmer, Lucille Wong, Dr Tommy Soesanto, Jen Riley, Alexander Gardendal and Danny Williams, Griffith University; Andrew Paultridge, Queensland University of Technology; Dr Christine Williams, Office of Economics and Statistical Research, Queensland Treasury; Professor Eric Sowe, University of New South Wales; Associate Professor Brett Inder, Dr Ann Maharaj, Dr Mindi Nath, John Betts, Bruce Stephens and Lesley Tissera, Monash University; Associate Professor John MacFarlane, Dr Kevin Donegan, Dr Than Pe and Neil Hopkins, University of Western Sydney; Professor Alan Woodland, Tig Ihnatko and John Goodhew, University of Sydney; Dr Geoff Coates, Anne Arnold and Margaret Meyler, University of Adelaide; Dr G.V. Crockett and Dr Fay Rola-Rubzen, Curtin University of Technology; Damon Chernoff, Edith Cowan University; Dr Eddie Oczkowski; Kerrie Cullis, William Collen and Sue Moffatt, Charles Stuart University; Dr Jim Bates, Victoria University; Joy Ross and Lisa Yong, RMIT; Iain Fraser and Suzanne Sommer, La Trobe University; Dr S. Ganesalingham, Massey University, New Zealand; Professor Christine Lim, Waikato University, New Zealand; Professor Brinda Viswanathan, Madras School of Economics, India; Dr N Ravinthirakumaran, University of Colombo; and Dr C. Elankumaran and Dr S. Kalamani, University of Jaffna, Sri Lanka.

The publisher and the authors would like to thank the following reviewers for their incisive and helpful comments:

- Michael Cornish – University of Adelaide
- Ranjith Ihalanayake – Victoria University
- Harmindar Nath – Monash University
- Vivian Piovesan – University of Adelaide
- Michael Pottenger – University of Melbourne
- Trent Smith – University of Otago and several other anonymous reviewers.

Every effort has been made to trace and acknowledge copyright. However, if any infringement has occurred, the publishers tender their apologies and invite the copyright holders to contact them.



## About the authors

### **Professor Eliyathamby A Selvanathan**

Eliyathamby ‘Selva’ Selvanathan is currently a Professor in Econometrics and Applied Statistics in the Griffith Business School at Griffith University, Queensland, Australia. He is also a Visiting Professor at the Madras School of Economics, Anna University, Chennai, India. Selva has also taught previously at the University of Jaffna, Murdoch University, The University of Western Australia and University of Queensland. He has held positions such as the Deputy Dean (Staffing) – Faculty of International Business, Director of Bachelor of International Business Program and Deputy Director of the Statistics and Research Design (STARDS) Unit at Griffith University. Selva was educated at the University of Jaffna, University of Bucharest, Murdoch University and The University of Western Australia. He is the recipient of several individual and group Excellence in Teaching Awards at both the University of Western Australia and Griffith University, and has received numerous competitive teaching and research grants. Selva has published six research monographs and has published widely in international refereed journals such as *Journal of Econometrics*, *Review of Economics and Statistics*, *Journal of Business and Economic Statistics*, *Review of Economics Studies* and *Marketing Science*, and several book chapters.

### **Professor Saroja Selvanathan**

Saroja Selvanathan is a Professor in Econometrics and the Head of the Economics and Business Statistics Discipline in the Department of Accounting, Finance and Economics, Griffith University, Queensland, Australia; a Visiting Professor at the Madras School of Economics, Anna University, Chennai, India; and Research Higher Degrees Convenor, Economics and Business Statistics, at Griffith University. She has also held positions such as Deputy Dean (Research and postgraduate studies), and the Director of the Statistics and Research Design (STARDS) Unit at Griffith University. She has also taught at the University of Jaffna, Murdoch University and The University of Western Australia. Saroja was educated at the University of Jaffna, Murdoch University and The University of Western Australia, and has published several research monographs and research papers in international refereed journals such as the *Review of Economics and Statistics*, *Transportation Research* and *Economics Letters*, and book chapters. Saroja was the joint recipient of *2005 Excellence in Teaching Award* for innovation across the institution for teaching statistics to postgraduate students.

### **Professor Gerald Keller**

Gerald Keller is a Professor of Business at Wilfred Laurier University, where he has taught statistics, management science and operations management since 1974. He has also taught at the University of Toronto, the University of Miami, McMaster University, the University of Windsor and the Beijing Institute of Science and Technology. Gerald has consulted with banks on credit scoring and credit card fraud and has conducted market surveys for the Canadian government on energy conservation. The author of *Applied Statistics with Microsoft Excel*, Gerald has also been published in *Omega*, *IIE Transactions*, *Decision Sciences*, *INFOR*, *Economics Letters* and *Archives of Surgery*.

# 1

## What is statistics?

### LEARNING OBJECTIVES

This chapter provides an introduction to the two general bodies of methods that together constitute the subject called statistics: descriptive statistics and inferential statistics.

At the completion of this chapter, you should be able to:

- LO1** describe the two major branches of statistics – descriptive statistics and inferential statistics
- LO2** understand the key statistical concepts – population, sample, parameter, statistic and census
- LO3** provide examples of practical applications in which statistics have a major role to play
- LO4** understand how statistics are used by business managers
- LO5** understand the basics of the computer spreadsheet package Microsoft Excel and its capabilities in aiding with statistical data analysis for large amounts of data.

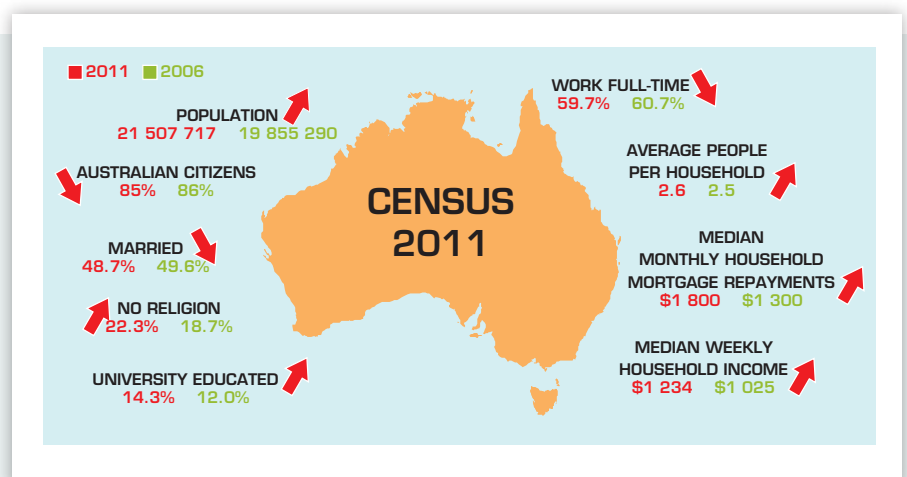
### CHAPTER OUTLINE

Introduction to statistics

- 1.1 Key statistical concepts
- 1.2 Practical applications
- 1.3 How managers use statistics
- 1.4 Statistics and the computer
- 1.5 Online resources

## Census in Australia

Information (or data) gathering on various characteristics of different populations of interest are an important part of statistics. When we use all units of the population to record information about the characteristics of interest, this type of data gathering is called a census. Due to cost and other



resource implications, census of the whole population is done only every five years in Australia. The peak Australian government statistics agency, the *Australian Bureau of Statistics*, carries out the census, the two latest being conducted in 2006 and 2011. A census provides a snapshot of the population characteristics in the year it is held. We will discuss census and other forms of data collection in detail in this chapter.

## Introduction to statistics

Today we have access to more data than ever, through the ever-increasing use of information technology to make informed decisions. Statistics is a body of principles and methods concerned with extracting useful information from a set of data to help people make decisions. The role of this book is to describe how, when and why managers and statisticians conduct statistical procedures. Such description is important as you come across different kinds of information and data to which you need to apply different statistical procedures.

In general, statistics can be subdivided into two basic areas: *descriptive statistics* and *inferential statistics*.

## Descriptive statistics

### descriptive statistics

Methods of organising, summarising and presenting data in ways that are useful, attractive and informative to the reader.

**Descriptive statistics** deals with methods of organising, summarising and presenting data in a convenient and informative form. One form of descriptive statistics uses graphical techniques, which allow statistics practitioners to present data in ways that make it easy for the reader to extract useful information. The main attraction of a graphical presentation is that the message can be easily understood by any layperson. In Chapters 3 and 4 we will present a variety of graphical methods.

Another form of descriptive statistics uses numerical techniques to summarise data. One such method you would have already used frequently is calculating an average or mean. Chapter 5 introduces several numerical statistical measures that describe different features of the data.

The actual technique we use depends on what specific information we would like to extract. Consider the use of descriptive statistics in the following examples.

### EXAMPLE 1.1

LO1

#### Business statistics marks

A student enrolled in a business program is attending his first lecture of the compulsory business statistics course. The student is somewhat apprehensive because he believes the myth that the course is difficult. To alleviate his anxiety, the student asks the lecturer about last year's exam marks of the business statistics course. Because, like all statistics lecturers, this one is friendly and helpful, he obliges and provides a list of the final marks. The marks are composed of all the within-semester assessment items plus the end-of-semester final exam. What information can the student obtain from the list?

This is a typical statistics problem. The student has the data (marks) and needs to apply statistical techniques to get the information he requires. This is a function of *descriptive statistics*.

In this example, we can see at least three important pieces of information. The first is the 'typical' mark. We call this a *measure of central location*. The average is one such measure. In Chapter 5 we will introduce another useful measure of central location, the median. In the above example, the median is the middle mark of the class when the marks are arranged in



ascending or descending order. That is, there are 50% of the students who obtained marks less than the median mark, while 50% received marks greater than the median value.

Suppose the student was told that the average mark last year was 67. Is this enough information to reduce his anxiety? The student would likely respond 'no' because he would like to know whether most of the marks were close to the average mark of 67 or were scattered far below and above the average. He needs a *measure of variability*. The simplest such measure is the *range* (discussed further in Chapter 5), which is calculated by subtracting the smallest number from the largest. Suppose the largest mark is 96 and the smallest is 24, then the range is  $(96 - 24 =) 72$ . Unfortunately, this provides little information as the range doesn't say where most of the marks are located. Whether most data are located near 24 or near 96 or somewhere in the middle, the range is still 72. We need other measures of variability such as the variance and standard deviation, to reflect the true picture of the spread of the data, which will be introduced in Chapter 5. Moreover, the student must determine more about the marks. In particular he needs to know how the marks are distributed between 24 and 96. The best way to do this is to use a graphical technique, the histogram, to be introduced in Chapter 4.

### EXAMPLE 1.2

LO1-3

#### Comparing weekly sales between two outlets

A fast-food franchiser wishes to compare the weekly sales level over the past year at two particular outlets. Descriptive statistical methods could be used to summarise the actual sales levels (perhaps broken down by food item) in terms of a few numerical measures, such as the average weekly sales level and the degree of variation from this average that weekly sales may undergo. Tables and charts could be used to enhance the presentation of the data so that a manager could quickly focus on the essential differences in sales performance at the two outlets.

There is much more to statistics, however, than these descriptive methods. Decision-makers are frequently forced to make decisions based on a set of data that is only a small subgroup (sample) of the total set of relevant data (population).

## Inferential statistics

**Inferential statistics** is a body of methods for drawing conclusions (i.e. making inferences) about characteristics of a population, based on information available in a sample taken from the population. The following example illustrates the basic concepts involved in inferential statistics.

### inferential statistics

Methods used to draw conclusions about a population based on information provided by a sample of the population.

### EXAMPLE 1.3

LO1-4

#### Profitability of a new life insurance policy

An Australia-wide automobile club (consisting of about 2 million members) is contemplating extending its services to its members by introducing a new life insurance policy. After some careful financial analysis, the club has determined that the proposed insurance policy would break even if at least 10% of all current members subscribing to the club also purchase the policy. The question here is how can inferential statistics be used by the automobile club to make a decision about introducing their new life insurance policy?

To obtain additional information before reaching a decision on whether or not to proceed with the new insurance policy, the automobile club has decided to conduct a survey of 500 randomly selected current

members. The collection of all its current 2 million or so members is called the *population*. The 500 members selected from the entire population for the analysis are referred to as a *sample*. Each member in the sample is asked if they would purchase the policy if it were offered at some specified price. Suppose that 60 of the members in this sample reply positively. While a positive response by 60 out of 500 members (12%) is encouraging, it does not assure the automobile club that the proposed insurance policy will be profitable. The challenging question here is how to use the response from these 500 sampled members to conclude that at least 10% of all 2 million or so members would also respond positively. The data are the proportion of positive response among the 500 members in the sample. However, we are not so much interested in the response of the 500 members as we are in knowing what the response would be from all of the club's current members. To accomplish this goal we need another branch of statistics – *inferential statistics*.

If the automobile club concludes, based on the sample information, that at least 10% of all its members in the population would purchase the proposed insurance policy, the club is relying on inferential statistics. The club is drawing a conclusion, or making a statistical inference, about the entire population of its 2 million or so members on the basis of information provided by only a sample of 500 members taken from the population. The available data tell us that 12% of this particular sample of members would purchase; the inference that at least 10% of all its members would purchase the new insurance policy may or may not be correct. It may be that, by chance, the club selected a particularly agreeable sample and that, in fact, no more than 5% of the entire population of members would purchase.

Whenever an inference is made about an entire population on the basis of evidence provided by a sample taken from the population, there is a chance of drawing an incorrect conclusion. Fortunately, other statistical methods allow us to determine the reliability of the statistical inference. They enable us to establish the degree of confidence we can place in the inference, assuming the sample has been properly chosen. These methods would enable the automobile club in Example 1.3 to determine, for example, the likelihood that less than 10% of the population of its members would purchase, given that 12% of the members sampled said they would purchase. If this likelihood is deemed small enough, the automobile club will probably proceed with its new venture.

## 1.1 Key statistical concepts

Statistical inference problems involve three key concepts: the population, the sample and the statistical inference. We now discuss each of these concepts in more detail.

### Population

**population**  
The set of all items of interest.

A **population** is the group of all items of interest to a statistics practitioner. It is frequently very large and may, in fact, be infinitely large. In the language of statistics, the word *population* does not necessarily refer to a group of people. It may, for example, refer to the population of diameters of ball bearings produced at a large plant. In Example 1.3, the population of interest consists of all 2 million or so members.

**parameter**  
A descriptive measure of a population.

A descriptive measure of a population is called a **parameter**. The parameter of interest in Example 1.3 was the proportion of all members who would purchase the new policy.

**sample**  
A set of data drawn from the studied population.

### Sample

A **sample** is a subset of data drawn from the population. In Example 1.3, the sample of interest consists of the 500 selected members.

A descriptive measure of a sample is called a **statistic**. We use sample statistics to make inferences about population parameters. In Example 1.3, the proportion of the 500 members who would purchase the life insurance policy would be a sample statistic that could be used to estimate the corresponding population parameter of interest, the population proportion. Unlike a parameter, which is a constant, a statistic is a variable whose value varies from sample to sample. In Example 1.3, 12% is a value of the sample statistic based on the selected sample.

**statistic**  
A descriptive measure of a sample.

## Statistical inference

Statistical inference is the process of making an estimate, forecast or decision about a population parameter, based on the sample data. Because populations are usually very large, it is impractical and expensive to investigate or survey every member of a population. (Such a survey is called a census.) It is far cheaper and easier to take a sample from the population of interest and to draw conclusions about the population parameters based on information provided by the sample.

For instance, political pollsters predict, on the basis of a sample of about 1500 voters, how the entire 16 million eligible voters from the Australian population will cast their ballots; and quality control supervisors estimate the proportion of defective units being produced in a massive production process from a sample of only several hundred units.

Because a statistical inference is based on a relatively small subgroup of a large population, statistical methods can never decide or estimate with certainty. Since decisions involving large amounts of money often hinge on statistical inferences, the reliability of the inferences is very important. As a result, each statistical technique includes a measure of the reliability of the inference. For example, if a political pollster predicts that a candidate will receive 40% of the vote, the measure of reliability might be that the true proportion (determined on election day) will be within 3% of the estimate on 95% of the occasions when such a prediction is made. For this reason, we build into the statistical inference a measure of reliability. There are two such measures, the **confidence level** and the **significance level**. The *confidence level* is the proportion of times that an estimating procedure would be correct, if the sampling procedure were repeated a very large number of times. For example, a 95% confidence level would mean that, in a very large number of repeated samples, estimates based on this form of statistical inference will be correct 95% of the time. When the purpose of the statistical inference is to draw a conclusion about a population, the *significance level* measures how frequently the conclusion will be wrong in the long run. For example, a 5% significance level means that, in repeated samples, this type of conclusion will be wrong 5% of the time. We will introduce these terms in Chapters 11 and 13.

**confidence level**  
The degree of certainty we have that our interval contains the value of the parameter.

**significance level**  
The relative frequency of a wrong conclusion.

## 1.2 Practical applications

Throughout the text, you will find examples, exercises and cases that describe actual situations from the business world in which statistical procedures have been used to help make decisions. For each example, exercise or case, you will be asked to choose and apply the appropriate statistical technique to the given data and to reach a conclusion. We cover such applications in accounting, economics, finance, management and marketing. Below is a summary of some of the case studies we have analysed in this textbook with partial data, to illustrate additional applications of inferential statistics. But you will have to wait until you work through these cases in the relevant chapters (where some data is also presented) to find out the conclusions and results.

## CASE 3.6

## Differing average weekly earnings of men and women in Australia

While a lot has been achieved in Australia to reduce the difference between men and women in a number of social status indicators, wage differences are still a matter of concern. The following table presents the average weekly earnings of male and female adults for each Australian state and territory and for Australia as a whole. Present the information using appropriate graphical techniques.

### Average weekly (all employees total) earnings (A\$), May 2012

State/Territory	Males	Females
New South Wales	1266.60	843.80
Victoria	1214.60	784.30
Queensland	1270.10	807.40
South Australia	1165.90	762.50
Western Australia	1556.30	834.30
Tasmania	1069.90	742.00
Northern Territory	1419.30	973.20
ACT	1546.50	1172.90
Total (Australia)	1253.10	744.80

Source: Australian Bureau of Statistics, *Average Weekly Earnings*, May 2012, cat. no. 6302.0, ABS, Canberra

## CASE 5.9

## Aussies and Kiwis are leading in education

According to the 2011 published statistics on the human development index (HDI) by the UN, Australians and New Zealanders are leading the world. The HDI is calculated using three indices, namely, education index, GDP index and life expectancy index. The education index data for the top 20, the middle 20 and the bottom 20 of the 176 countries listed in the UN report are recorded. Part of the data are given below. Use suitable numerical summary (central location and variability) measures to analyse the data.

Top 20	Education index	Bottom 20	Education index
Australia	0.993	Eritrea	0.539
Cuba	0.993	...	...
Denmark	0.993	...	...
Finland	0.993	Chad	0.334
New Zealand	0.993	Mali	0.331
...	...	Burkina Faso	0.301
Slovenia	0.969	Niger	0.282

Source: *Human Development Report 2009*, United Nations Development Program (UNDP), New York, 2011

## CASE 15.1

## Comparing salary offers for finance and marketing MBA majors - I

In the last few years, there has been an increase in the number of web-based companies that offer job placement services. The manager of one such company wanted to investigate the job offers recent MBAs were obtaining. In particular, she wanted to know

whether finance majors were being offered higher salaries than marketing majors. In a preliminary study, she randomly sampled 50 recently graduated MBAs, half of whom majored in finance and half in marketing. From each she recorded the highest salary offer (including benefits). Can we infer that finance majors obtain higher salary offers than do marketing majors among MBAs? Verify the underlying assumptions.

## Gold lotto

Gold lotto is a national lottery that operates as follows. Players select eight different numbers (six primary and two supplementary numbers) from 1 and 45. Once a week, the corporation that runs the lottery selects eight numbers (six primary and two supplementary numbers) at random from 1 to 45. Winners are determined by how many numbers on their tickets agree with the numbers drawn. In selecting their numbers, players often look at past patterns as a way to help predict future drawings. A regular feature that appears in the newspaper identifies the number of times each number has occurred in the past. The data recorded in the following table appeared in the 20 January 2013 edition of the *Queensland Sunday Mail* after the completion of draw 3287. What would you recommend to anyone who believes that past patterns of the lottery numbers are useful in predicting future drawings?

### Drawing frequency of lotto numbers since draw 413

Lotto number	Number of times drawn	Lotto number	Number of times drawn	Lotto number	Number of times drawn
1	274	16	246	31	252
2	257	17	249	32	256
.	.	.	.	.	.
.	.	.	.	.	.
14	240	29	248	44	230
15	265	30	231	45	242

## Does unemployment affect the average hourly earnings of New Zealanders?

Wages in the labour market are very much influenced by the demand and supply of labour. In any profession or industry, when there is an over supply of labour, the workers will be at a disadvantage and will not be able to demand high wages and vice versa. During the mining boom in Perth, unusually high wages were paid, due to the shortage of workers to work in the mining fields. This has impacted heavily on the other sectors of the economy (e.g. the house prices in Perth were inflated to astonishingly high levels). In New Zealand, also, in the last few decades the average hourly earnings have fluctuated, depending on the state of the economy, especially on the level of labour supply. When the number of unemployed persons increases, it is expected that the average hourly earnings would fall. The data in the file presents the quarterly data for the average hourly earnings and the total number of unemployed persons in New Zealand during the period March 1994 to Sept 2012. Is there any evidence in New Zealand to support the proposition that the higher (lower) the number of unemployed the lower (higher) the average hourly earnings?

Period	Number unemployed	Average hourly earnings (\$)
Mar. 1994	170 300	14.98
Jun. 1994	147 000	15.10
Sep. 1994	135 800	15.10
Dec. 1994	134 800	15.16
.	.	.
.	.	.
.	.	.
Dec. 2011	150 700	26.59
Mar. 2012	171 200	26.96
Jun. 2012	156 400	27.00
Sep. 2012	170 000	27.31

Source: Income Tables, Statistics New Zealand, December 2012

The objective of the problem described in Case 3.6 is to use the descriptive graphical and numerical techniques to analyse the differences in weekly earnings of men and women in Australia; Case 5.9 is to compare the central location and variability of the education index of the top and bottom 20 countries in the world; Case 15.1 is to compare two populations, the variable of interest being the salary of MBA graduates specialising in marketing and finance. Case 17.1 is a day-to-day real-life application. The objective of the problem is to see how statistical inference can be used to determine whether some numbers in a lotto draw occur more often than others. Case 18.1 illustrates another statistical objective. In this case, we need to analyse the relationship between two variables: average hourly earnings and the total number of unemployed persons in New Zealand. By applying the appropriate statistical technique, we will be able to determine whether the two variables are related, and, if so, whether reducing the number of unemployed leads to higher average salary earnings. As you will discover, the technique also permits statistics practitioners to include other variables to determine whether they affect average salary earnings.

## 1.3 How managers use statistics

As we have already pointed out, statistics is about acquiring and using information. However, the statistical result is not the end product. Managers use statistical techniques to help them make decisions. In general, statistical applications are driven by the managerial problem. The problem creates the need to acquire information. This in turn drives the data-gathering process. When the manager acquires data, he or she must convert the data into information by means of one or more statistical techniques. The information then becomes part of the decision process.

Many business students will take or have already taken a subject in marketing. In the introductory marketing subject, students are taught about market segmentation. Markets are segmented to develop products and services for specific groups of consumers. For example, the Coca-Cola Company produces several different cola products.

There is Coca-Cola Classic, Coca-Cola Vanilla, Coca-Cola Zero, Coke, Diet Coke and Caffeine-Free Diet Coke. Each product is aimed at a different market segment. For example, Coca-Cola Classic is aimed at people who are older than 30, Coca-Cola Vanilla is aimed primarily at women, Coke is aimed at the teen market, Diet Coke is marketed towards individuals concerned about their weight or sugar intake, and Caffeine-Free Diet Coke is for

people who are health-conscious. In order to segment the cola market, Coca-Cola had to determine that all consumers were not identical in their wants and needs. The company then had to determine the different parts of the market and ultimately design products that were profitable for each part of the market. As you might guess, statistics plays a critical but not exclusive role in this process.

Because there is no single way to segment a market, managers must try different segmentation variables. Segmentation variables include geographic (e.g. states, cities, country towns), demographic (e.g. age, gender, occupation, income, religion), psycho-graphic (e.g. social class, lifestyle, personality) and behaviouristic (e.g. brand loyalty, usage, benefits sought). Consumer surveys are generally used by marketing researchers to determine which segmentation variables to use. For example, Coca-Cola used age and lifestyle. The age of consumers generally determines whether they buy Coca-Cola Classic or Coke. Lifestyle determines whether they purchase regular, diet or caffeine-free cola. Surveys and statistical techniques would tell the marketing manager that the 'average' Coca-Cola Classic drinker is older than 30, whereas the 'average' Coke drinker is a teenager. Census data and surveys are used to measure the size of the two segments. Surveys would also inform about the number of cola drinkers who are concerned about kilojoules and/or caffeine. The conversion of the raw data in the survey into statistics is only one part of the process. The marketing manager must then make decisions about which segments to pursue (not all segments are profitable), how to sell and how to advertise.

In this book, we will address the part of the process that collects the raw data and produces the statistical result. By necessity, we must leave the remaining elements of the decision process to the other subjects that constitute business programs. We will demonstrate, however, that all areas of management can and do use statistical techniques as part of the information system.

## 1.4 Statistics and the computer

In almost all practical applications of statistics, the statistics practitioner must deal with large amounts of data. In order to calculate various statistical measures, the statistics practitioner would have to perform various calculations on the data; although the calculations do not require any great mathematical skill, the sheer amount of arithmetic makes this aspect of the statistical method time-consuming and tedious. Fortunately, numerous commercially prepared computer programs are available to perform these calculations. In most of the examples used to illustrate statistical techniques in this book, we will provide two methods for answering the question:

- 1 **Calculating manually.** Except where doing so is prohibitively time-consuming, we will show how to answer the questions using hand calculations (with only the aid of a calculator). It is useful for you to produce some solutions in this way, because by doing so you will gain insights into statistical concepts.
- 2 **Using Microsoft Excel.** Many business students own a spreadsheet package, and university and TAFE subjects incorporate a spreadsheet into their curriculum. We have chosen to use Microsoft Excel 2010 because we believe that it is and will continue to be the most popular spreadsheet package and is the most accessible package. Excel also comes with a limited statistical tool called **Data Analysis**. Consequently, we have included (on the **CourseMate for Business Statistics** website – see Section 1.5) a statistical software add-in, **Data Analysis Plus 9.0** for Excel 2010 (and **Data Analysis Plus 7.0** and **5.1** for earlier versions of Excel), and also created various other macros that can be loaded on to your computer to enable you to use Excel for almost all procedures. Detailed instructions are provided for all techniques. An introduction to the use of Excel is provided in Appendix 1A of this chapter.