Business Statistics
Abridged
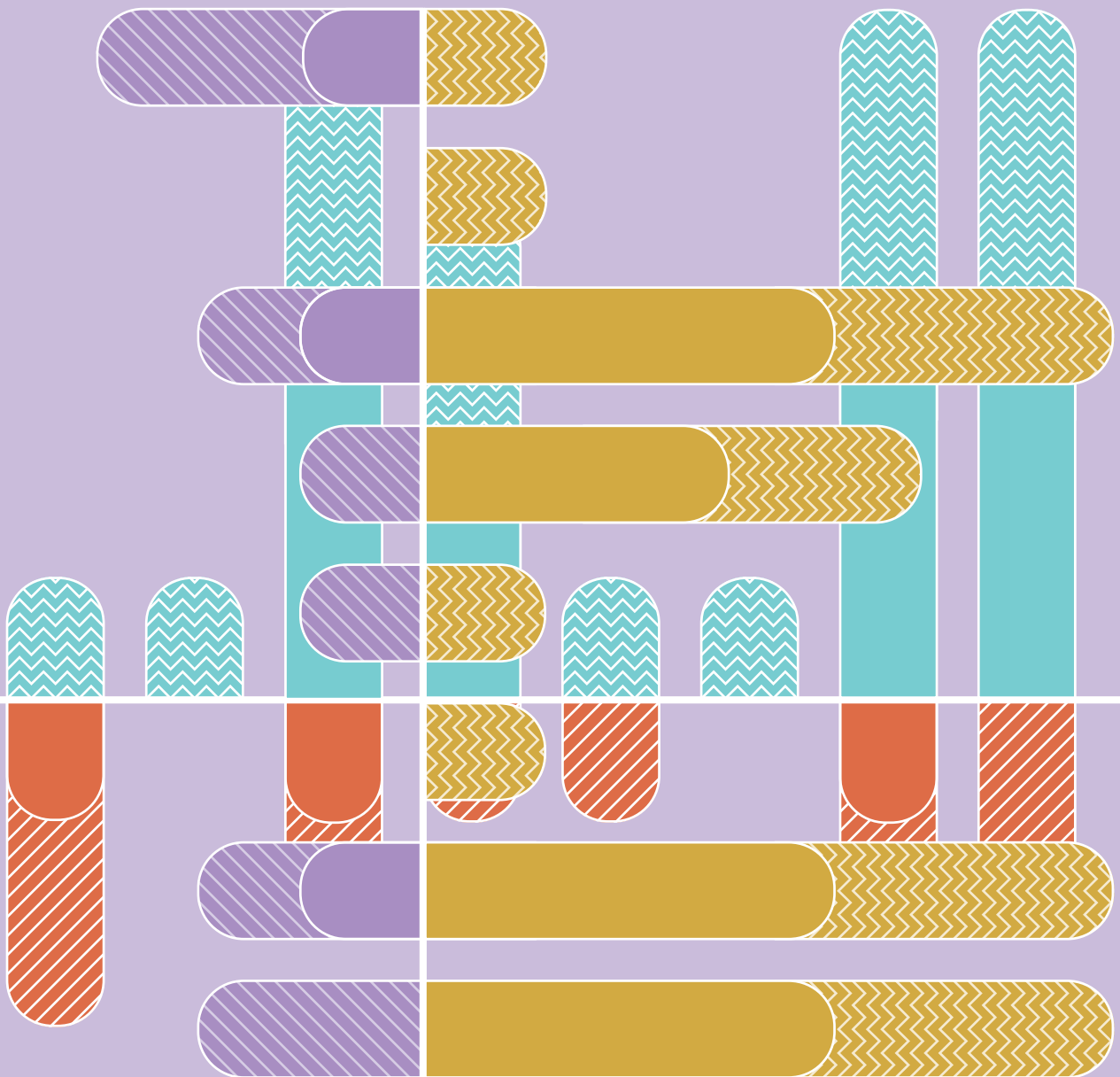
Eliyathamby A. Selvanathan　　Saroja Selvanathan　　Gerald Keller

Business Statistics
Abridged

Australia / New Zealand
8th edition

Eliyathamby A. Selvanathan   Saroja Selvanathan   Gerald Keller

DEDICATION

We would like to dedicate this book to our statistics gurus, the Foundation Lecturers of the Department of Mathematics and Statistics, University of Jaffna, Sri Lanka:
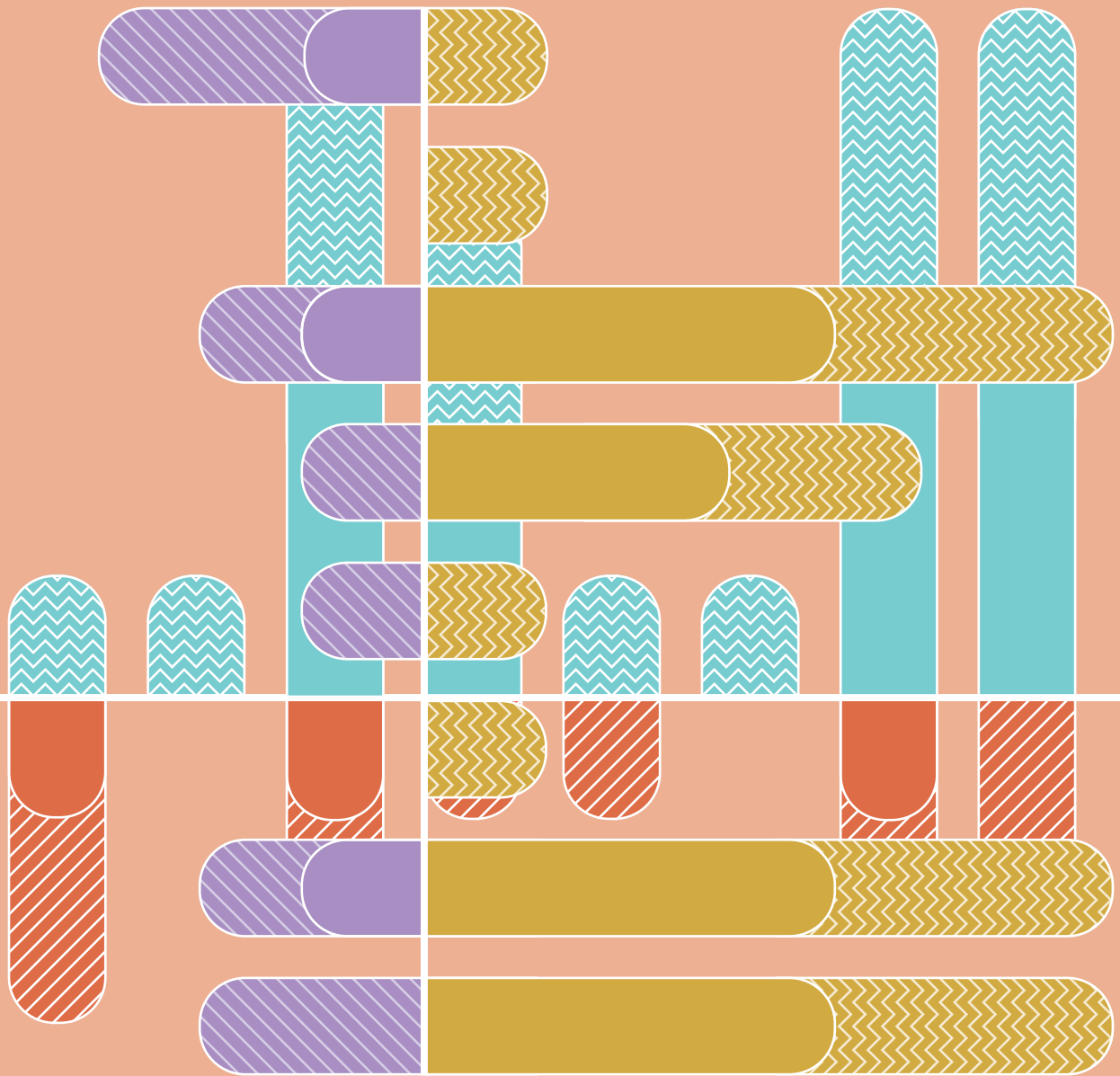
Late Professor Balan Selliah
Professor S Ganesalingam
Late Dr S Varatharajaperumal

Business Statistics
Abridged

Eliyathamby A. Selvanathan   Saroja Selvanathan   Gerald Keller

BRIEF CONTENTS

CONTENTS

PREFACE

Managing a business is a very complex responsibility 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 sets that are potential sources of information for making well-informed objective decisions. Business managers are increasingly using statistical techniques to convert data into meaningful information. However, to achieve this, potential managers need to know which statistical techniques they should use to extract useful information from the available data to 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, Abridged – Australia and New Zealand* is designed to demonstrate that statistical methods are vital tools for today's businesses and business managers to improve their decision-making skills.

This book is a thorough Australasian adaptation of the most popular and best-selling US text, *Statistics for Management and Economics* (11th 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 that 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, additional graphical techniques such as bubble chart and heat maps in the graphical methods chapter, 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.

When solving problems, *Business Statistics, Abridged – Australia and New Zealand* uses its unique 'ICI' approach, which is renowned for its consistent, proven three-step method to solving problems. The ICI approach teaches you how to: (1) *Identify* the appropriate technique, (2) *Compute* the statistics and (3) *Interpret* the results, in the context of the problem at hand. The *compute* stage can be completed in any or all of three ways: manually (with the aid of a calculator) or using Excel or XLSTAT (on the computer). This book contains step-by-step instructions and commands to teach students how to use Microsoft Excel® or XLSTAT to solve statistical problems. For those courses that wish to use the computer extensively, manual calculations can be played down or omitted completely. Conversely, those that wish to emphasise manual calculations may easily do so, and the computer solutions can be selectively introduced or skipped entirely. This approach is designed to provide maximum flexibility, and it leaves to the instructor the decision of when to introduce the computer.

Additionally, most examples, exercises and cases feature raw data. These data sets are available to download from the companion website accessible through https://login.cengagebrain.com/.

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. 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 workbooks that can be used to solve statistical problems beyond Excel's existing capabilities. In this edition we have introduced another Excel Add-in, *XLSTAT*, which is capable of sophisticated statistical analysis to complement Excel's menu of statistical procedures. All statistical techniques introduced in this book can be calculated using either Excel's *Data Analysis (Analysis ToolPak)* or XLSTAT, which is available to download from the companion website accessible through https://login.cengagebrain.com/. The website also contains another Excel add-in *Data Analysis Plus*, which can also be used to perform a number of statistical calculations that Excel *Data Analysis* is unable to perform.

The Excel spreadsheet package is used extensively and presented consistently throughout the book to calculate sample statistics. Most examples in the chapters present manual (*Calculating manually*) and computer (*Using the computer*) 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 (Excel) 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 eighth edition includes hundreds of data files, some consisting of hundreds 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 computer software. 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.
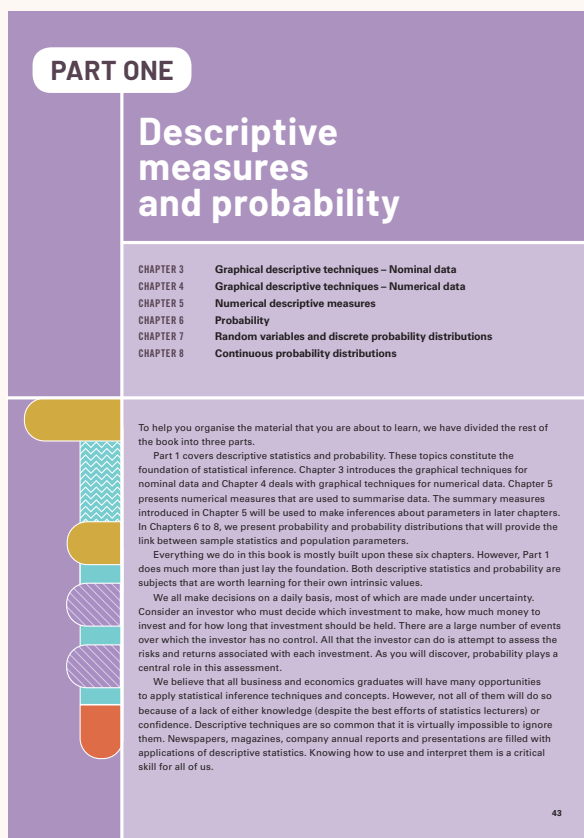
- *In addition to the examples provided in each chapter*, we have included 'Real-life applications' sections which illustrate the fundamental applications of statistics in finance, marketing, human resources management, operations management, accounting and economics.
  - For example, to illustrate graphical techniques, we use an example that compares the histograms of the returns on two different investments. To explain what financial analysts look for in the histograms requires an understanding that risk is measured by the amount of variation in the returns. The example is preceded by a 'Real-Life Applications' box that discusses how return on investment is computed and used.
  - Later when we present the normal distribution, we feature another 'Real-Life Applications' box to show why the standard deviation of the returns measures the risk of that investment.
  - Several 'Real-Life Application' boxes are scattered throughout the book.
- *'In Summary'* 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 is further enhanced by '*Chapter Summary*' at the end of each chapter, which also include individual summary flowcharts in most of the Statistical Inference chapters and an overall summary flowchart in the review chapter.
- *Several new exercises are added to each chapter.*
- *In addition to updating the data in the Examples and Exercises as much as possible, 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.
- *A more sophisticated commercially available Excel Add-in XLSTAT has been incorporated, to enable you to use Excel for almost all statistical procedures introduced in this book.*
- *In addition to Excel Data Analysis and XLSTAT commands,* 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.

Guide to the text

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

PART-OPENING FEATURES

Part outlines show the chapters that are included in each part.

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 that are used to summarise data. The summary measures introduced in Chapter 5 will be used to make inferences about parameters in later chapters. In Chapters 6 to 8, we present probability and probability distributions that will provide the link between sample statistics and population parameters.

Everything we do in this book is mostly built upon these six chapters. However, Part 1 does much more than just lay the foundation. Both descriptive 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.
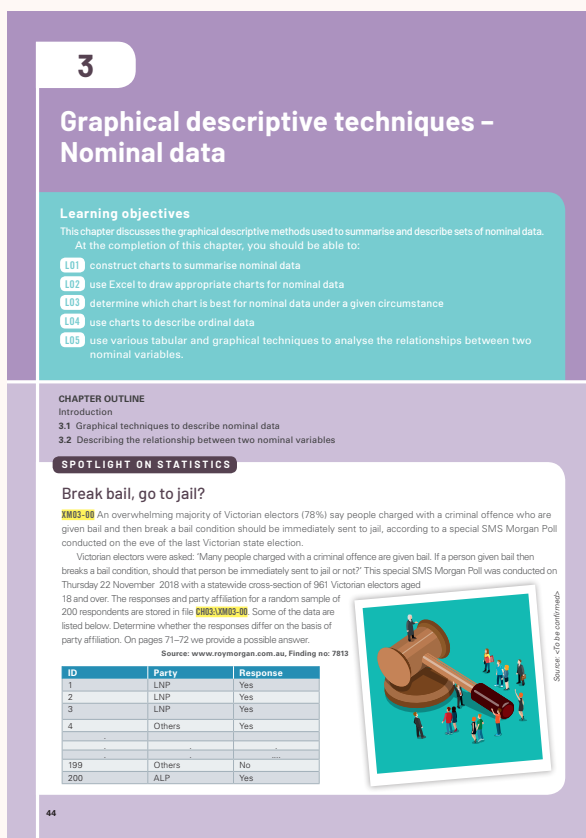
We believe that all business and economics graduates will have many opportunities to apply statistical inference techniques and concepts. However, not all of them will do so because of a lack of either knowledge (despite the best efforts of statistics lecturers) or confidence. Descriptive techniques 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.

43

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

CHAPTER-OPENING FEATURES

Identify the key concepts that the chapter will cover with the **learning objectives** and get a clear sense of what you should be able to do after reading the chapter.

3

Graphical descriptive techniques – Nominal data

Learning objectives

This chapter discusses the graphical descriptive methods used to summarise and describe sets of nominal data. At the completion of this chapter, you should be able to:

- LO1 construct charts to summarise nominal data
- LO2 use Excel to draw appropriate charts for nominal data
- LO3 determine which chart is best for nominal data under a given circumstance
- LO4 use charts to describe ordinal data
- LO5 use various tabular and graphical techniques to analyse the relationships between two nominal variables.

CHAPTER OUTLINE
Introduction
3.1 Graphical techniques to describe nominal data
3.2 Describing the relationship between two nominal variables

SPOTLIGHT ON STATISTICS

Break bail, go to jail?

XM03-00 An overwhelming majority of Victorian electors (78%) say people charged with a criminal offence who are given bail and then break a bail condition should be immediately sent to jail, according to a special SMS Morgan Poll conducted on the eve of the last Victorian state election.

Victorian electors were asked: 'Many people charged with a criminal offence are given bail. If a person given bail then breaks a bail condition, should that person be immediately sent to jail or not?' This special SMS Morgan Poll was conducted on Thursday 22 November 2018 with a statewide cross-section of 961 Victorian electors aged 18 and over. The responses and party affiliation for a random sample of 200 respondents are stored in file CH03\XM03-00. Some of the data are listed below. Determine whether the responses differ on the basis of party affiliation. On pages 71–72 we provide a possible answer.
Source: www.roymorgan.com.au, Finding no: 7813

ID	Party	Response
1	LNP	Yes
2	LNP	Yes
3	LNP	Yes
4	Others	Yes
.	.	.
.	.	.
199	Others	No
200	ALP	Yes

Source: <To be confirmed>

44

Spotlight on Statistics are used at the beginning of every chapter. They highlight a specific problem that can be solved by using the statistical techniques that will be covered in the chapter. The problem is answered later in the chapter.

FEATURES WITHIN CHAPTERS

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.

You will learn to:

- **identify** the right statistical technique to use by focusing on the relationship between the problem and data type and **calculate manually**.

- **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 'Command' boxes, which include step-by-step instructions on how to complete the examples using Microsoft Excel®.

- **interpret** the answer in the context of the problem.

EXAMPLE 11.2 `L07`

Comparing the clothing purchasing habits of men and women

The manager of a major clothing manufacturer wants to compare the annual expenditure on clothing of men and women. She decides to estimate the difference in mean annual expenditure to within $100 with 95% confidence. How large should the two sample sizes be, if we assume that the range of expenditure is $800 for males and $1200 for females and the populations of male and female expenditures on clothing are normal?

Solution

Identifying the technique

The problem objective is to determine the required sample sizes. The data type is numerical. Both populations of interest are normally distributed and the samples are independent.

Calculating manually

The error bound is $100 and the confidence level is 0.95. Hence, $B = \$100$ and $z_{\alpha/2} = z_{0.025} = 1.96$. We approximate σ_1 and σ_2 by using the following formulas:

$$\sigma_1 = \frac{\text{Range}}{4} = \frac{800}{4} = \$200 \text{ (for males)}$$
$$\sigma_2 = \frac{\text{Range}}{4} = \frac{1200}{4} = \$300 \text{ (for females)}$$

Thus,

$$n_1 = n_2 = \left[\frac{1.96\sqrt{200^2 + 300^2}}{100}\right]^2 = 49.94 = 50$$

Interpreting the results

In order to estimate $\mu_1 - \mu_2$ to within $100 with 95% confidence, we should take samples of 50 men and 50 women.

Using the computer

The required sample sizes can be calculated using the **Sample size-2Means** worksheet in the **Estimators** workbook. The output is shown below.

Excel output for Example 11.2

	A	B	C
1	Calculating the sample sizes for a fixed width (2B)		
2		Sample 1	Sample 2
3	Variance	40000	90000
4	Width/2 = B	100	
5	Confidence level	0.95	
6	Sample size	50	50

COMMANDS

To estimate required sample sizes, open the **Sample size-2Means** worksheet in the **Estimators** workbook, then type in the values of the given population variances (or as calculated above), B and the confidence level.

Definition and **Formula** boxes enable you to review your understanding of each new concept.

Geometric mean

The geometric mean of a population of N observations $x_1, x_2, \ldots, x_N$ is defined as:

Population geometric mean: $\mu_g = \sqrt[N]{x_1 x_2 \ldots x_N} = (x_1 x_2 \ldots x_N)^{1/N}$

The geometric mean of a sample of n observations $x_1, x_2, \ldots, x_n$ is defined as:

Sample geometric mean: $\bar{x}_g = \sqrt[n]{x_1 x_2 \ldots x_n} = (x_1 x_2 \ldots x_n)^{1/n}$

In Summary boxes are included after each technique has been introduced, as well as at the end of sections, to allow you to appreciate a technique's essential requirements and to enable you to review your understanding of each technique.

IN SUMMARY

Factors that identify when to calculate the range, variance, standard deviation and coefficient of variation

1 *Objective:* to describe a single set of data
2 *Type of data:* numerical
3 *Descriptive measurement:* variability

FEATURES WITHIN CHAPTERS

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

These include:

- **Learning the techniques** exercises help you to learn the arithmetic involved in a specific procedure
- **Applying the technique** exercises 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®.

EXERCISES

Learning the techniques

3.29 **XR03-29** The following table summarises the data from a survey on the ownership of iPads for families with different levels of income (C1 < C2 < C3). Determine whether the two nominal variables are related.

Ownership	Level of income		
	C_1	C_2	C_3
No	40	32	48
Yes	30	48	52

Applying the techniques

3.30 **XR03-30** **Self-correcting exercise.** The trustee of a company's superannuation scheme has solicited the opinions of a sample of the company's employees regarding a proposed revision of the scheme. A breakdown of the responses is shown in the following table. Use an appropriate graphical presentation to determine whether the responses differ among the three groups of employees.

Responses	Blue-collar workers	White-collar workers	Managers
For	67	32	11
Against	63	18	9

Computer applications

3.31 **XR03-31** The associate dean of a business school was looking for ways to improve the quality of applicants to its MBA program. In particular she wanted to know whether the undergraduate degrees of applicants to her school and the three nearby universities with MBA programs differed. She sampled 100 applicants of her program and an equal number from each of the other universities.

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 student companion website (http://login.cengagebrain.com), the MindTap platform or via your instructor.

EXAMPLE 10.1   L04 L05

Number of hours children spend watching television

XM10-01 The sponsors of television shows targeted at children wanted to know the amount of time children spend watching television, since the types and number of programs and commercials presented are greatly influenced by this information. As a result, a survey was conducted to estimate the average number of hours Australian children spend watching television per week. From past experience, it is known that the population standard deviation σ is 8.0 hours. The following are the data gathered from a sample of 100 children. Find the 95% confidence interval estimate of the average number of hours Australian children spend watching television.

Numb
XM10-01 T
watching
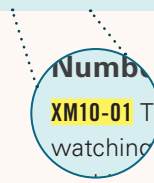
Real-Life Applications are included throughout the text to demonstrate real-world applications of statistics in the areas of finance, marketing, human resource management, accounting and economics.

REAL-LIFE APPLICATIONS

Test marketing

In Chapter 13, we described test marketing, which is often used to assess consumer reaction to changes in one or more elements of the marketing mix. Marketing managers will conduct experiments to determine whether differences in sales exist between different prices for the product, different package designs or different advertising strategies. Some of these experiments are carried out in small communities where it is easy to vary the particular elements that the manager wishes to investigate.

Source: Shutterstock.com/ AlenKadr

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

COMMANDS

1 Open the data file (**XM03-07**). Highlight the data including the column titles (**A1:C355**).
2 Click **INSERT** and **PivotTable** under the **Charts** submenu. Then select **Pivot Chart & Pivot Table**.
3 Make sure that the **Table/Range** is correct and click **OK**.
4 Drag the **Occupation** button from the menu that appears on the right of the screen to the **Drop Row Fields Here** section of the box. Drag the **Newspaper** button to the **Drop Column Fields Here** section. Drag the **Reader** button to the **Drop Value Fields Here** section. Right-click any number in the table, click **Summarize Values By**, and check **Count**. This will produce a contingency table for the counts (frequencies).
5 To convert to row percentages, right-click any number in the table, click **Summarize Values By**, **More options...**, and **Show values as**. Select **% of rows** from the drop-down menu and then click **OK**. Format the data into decimals by highlighting all cells with percentages and right-clicking. Select **Number Format...**, select **Number** under **Category** and select the number of **Decimal places**. Click **OK**. This will produce the contingency table for the row relative frequencies.

Definitions or explanations of important **key terms** are located in the margins for quick reference.

cross-classification table
A first step in graphing the relationship between two nominal variables.

To describe the relationship between two nominal variables, we must remember that we are only permitted to determine the frequency of the values. A variation of the bar chart introduced in Section 3.1 is used to describe the relationship between two nominal (categorical) variables in graphical form. As a first step, we need to produce a cross-classification table (also known as contingency table or cross-tabulation table), which lists the frequency of each combination of the values of the two variables. We will illustrate the use of graphs to describe the relationship between two nominal variables using data from the newspaper readership case in Example 3.7.

END-OF-CHAPTER FEATURES

At the end of each chapter you will find several tools to help you to review, practise and extend your knowledge of the key learning objectives including:

- 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.

Study Tools

CHAPTER SUMMARY

The concept of a random variable permits us to summarise the results of an experiment in terms of numerically valued events. Specifically, a random variable assigns a numerical value to each simple event of an experiment. There are two types of random variables. A discrete random variable is one whose values are countable. A continuous random variable can assume an uncountable number of values. In this chapter we discussed discrete random variables and their probability distributions.

We defined the *expected value*, *variance* and *standard deviation* of a population represented by a discrete probability distribution. We also presented two most important discrete distributions: the *binomial* and the *Poisson*. Finally, in this chapter we also introduced *bivariate discrete distributions* for which an important application in finance was discussed.

SYMBOL

Symbol	Pronounced	Represents
$\sum_{\text{all } x} x$	Sum of x for all values of x	Summation
C_x^n	n-choose-x	Number of combinations
$n!$	n-factorial	$n(n-1)(n-2) \ldots (3)(2)(1)$
e	exponential	2.718...

SUMMARY OF FORMULAS

Expected value (mean)	$E(X) = \mu = \sum_{\text{all } x} x p(x)$
Variance	$V(X) = \sigma^2 = \sum_{\text{all } x} (x-\mu)^2 p(x) = \sum_{\text{all } x} x^2 p(x) - \mu^2$
Standard deviation	$SD(X) = \sigma = \sqrt{V(X)}$
Covariance	$COV(X,Y) = \sigma_{XY} = \sum_{\text{all } x}(x-\mu_x)(y-\mu_y)p(x,y)$ $= \sum_{\text{all } x}\sum_{\text{all } y} xy \, p(x,y) - \mu_x\mu_y$
Coefficient of correlation	$\rho = \dfrac{\sigma_{XY}}{\sigma_x \sigma_y}$

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

SUPPLEMENTARY EXERCISES

9.33 The dean of a business school claims that the average Master of Business Management graduate is offered an annual starting salary of $109500. The standard deviation of the offers is $9900. What is the probability that for a random sample of 38 Master of Business Management graduates, the mean starting annual salary is less than $105000?

9.34 Refer to Exercise 9.33. Suppose that a random sample of 38 Master of Business Management graduates report that their mean starting salary is $105000. What does this tell you about the dean's claim?

9.35 A restaurant in a large commercial building provides coffee for the building's occupants. The restaurateur has determined that the mean number of cups of coffee consumed in one day by all the occupants is 2.0, with a standard deviation of 0.6. A new tenant of the building intends to have a total of 125 new employees. What is the probability that the new employees will consume more than 240 cups of coffee per day?

9.36 The number of pages photocopied each day by the admin staff in a busy office is normally distributed with a mean of 550 and a standard deviation of 150. Determine the probability that in one business week (i.e. 5 days) more than 3000 pages will be copied.

9.37 A university bookstore claims that 50% of its

location they seek, rather than ask for directions. To examine this belief, he took a random sample of 350 male drivers and asked each what they did when lost. If the belief is true, determine the probability that less than 75% said they continue driving.

9.39 The Red Lobster restaurant chain regularly surveys its customers. On the basis of these surveys, management claims that 75% of customers rate the food as excellent. A consumer testing service wants to examine the claim by asking 460 customers to rate the food. What is the probability that less than 70% rate the food as excellent?

9.40 An accounting professor claims that no more than one-quarter of undergraduate business students will major in accounting.
  a What is the probability that in a random sample of 1200 undergraduate business students, 336 or more will major in accounting?
  b A survey of a random sample of 1200 undergraduate business students indicates that there are 336 students who plan to major in accounting. What does this tell you about the professor's claim?

9.41 Statisticians determined that the mortgages of homeowners in a city is normally distributed with a mean of $500000 and a standard deviation of $100000. A random sample of 100 homeowners

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.

Case Studies

CASE 8.1 Average salary of popular business professions in Australia

C08-01 Based on a recent publication (payscale.com/index/au/job), the average salaries of 10 popular business-related jobs in Australia are listed below. It is believed that salaries can be considered as following a normal distribution with a standard deviation of $1500. A high school student would like to choose a degree that could lead to the profession which has a higher probability of gaining a reasonably good salary. What is the likelihood of him receiving an annual salary greater than $60000 for each of the 10 jobs listed in the table?

Average salary in Australia, 10 popular business jobs

Job title	Average salary ($)
Accountant	57139
Business Data Analyst	69823
Finance Manager	93835
Financial Planner	76004
HR Manager	89328
Marketing Manager	78847
Personal Banker	53285
Retail Manager	51413
Supply Chain Manager	103724
Tax Accountant	56316

Source: payscale.com/index/au/job

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

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 $x_1, x_2, \ldots, x_n$. A concise way of representing their sum is:

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

That is:

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

Terminology and notation

1 The symbol Σ is the capital Greek letter sigma, and means 'the sum of'.
2 The letter i is called the *index of summation*. The letter chosen to represent the index of summation is arbitrary.
3 The expression $\sum_{i=1}^{n} x_i$ is read as 'the sum of the terms x_i where i assumes the values from 1 to n inclusive'.

Guide to the online resources

Cengage Learning is pleased to provide you with a selection of resources that will help you prepare your lectures and assessments. These teaching tools are accessible via **cengage.com.au/instructors** for Australia or **cengage.co.nz/instructors** for New Zealand.

MINDTAP

Premium online teaching and learning tools are available on the *MindTap* platform – the personalised eLearning solution.

MindTap is a flexible and easy-to-use platform that helps build student confidence and gives you a clear picture of their progress. We partner with you to ease the transition to digital – we're with you every step of the way.

The *Cengage Mobile App* puts your course directly into students' hands with course materials available on their smartphone or tablet. Students can read on the go, complete practice quizzes or participate in interactive real-time activities.

MindTap for Selvanathan's Business Statistics is full of innovative resources to support critical thinking, and help your students move from memorisation to mastery! Includes:

- Selvanathan's Business Statistics eBook
- Data sets for Examples, Exercises and Cases
- Downloadable Workbooks to perform some of the statistical estimations and tests discussed in the book
- Videos and animations
- Selected solutions for the self-correcting exercises
- Revision quizzes
- Interactive assignment quizzes
- Aplia problem sets
- Interactive flashcards to revise key terms.

MindTap is a premium purchasable eLearning tool. Contact your Cengage learning consultant to find out how *MindTap* can transform your course.

CENGAGE | MINDTAP

SOLUTIONS MANUAL

The **Solutions Manual** includes solutions for all the exercises and supplementary exercises in each chapter.

WORD-BASED TEST BANK

This bank of questions has been developed with the text for the creation of quizzes, tests and exams for your students. Deliver tests from your LMS and your classroom.

POWERPOINT™ PRESENTATIONS

Use the chapter-by-chapter **PowerPoint** presentations to enhance your lecture presentations and handouts to reinforce the key principles of your subject.

ARTWORK FROM THE TEXT

Add the digital files of graphs, pictures and flow charts into your course management system, use them in student handouts, or copy them in your lecture presentations

FOR THE STUDENT

This book is accompanied by a companion website that can be accessed via https://login.cengagebrain.com, which contains the data sets for Examples, Exercises and Cases.

MINDTAP

A new approach to highly personalised online learning, **MindTap** is designed to match your learning style and provides you with an engaging interface that allows you to interact with the course content and multimedia resources as well as with your peers, lecturers and tutors. In the **MindTap** Reader you can make notes, highlight text and even find a definition directly from the page.

To purchase your **MindTap** experience for **Business Statistics**, please contact your instructor.

MindTap also includes:

- Data sets for examples, exercises and cases
- Workbooks for examples, exercises and cases
- Excel estimator workbooks
- Animations plus assistance documentation
- Solutions to selected self-correcting exercises
- Data Analysis Plus with assistance documentation
- Revision quizzes.

CENGAGE | MINDTAP

ACKNOWLEDGEMENTS

ABOUT THE AUTHORS

Professor Eliyathamby A Selvanathan

Eliyathamby 'Selva' Selvanathan is a Professor in Econometrics and Director of the Economics Policy Analysis Program (EPAP) 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, such as *75th Anniversary Distinguished Teaching Award, The University of Western Australia, 1988*, *2006 Minister's Awards for Outstanding Contribution to Improving Literacy and/or Numeracy, Queensland state winner*, *Griffith Awards for Excellence in Teaching 2005 – Innovation Across the Institution*, *Pro-Vice Chancellor's Research Excellence Award 2014 – Research Supervision*, and has received numerous competitive teaching and national 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 Economic Studies, Energy Economics*, *Economic Letters*, *Tourism Analysis*, *Tourism Analysis* and *Marketing Science*, several chapters in books, and several editions of two textbooks.

Professor Saroja Selvanathan

Dr Saroja Selvanathan is Professor in Econometrics at Griffith University and a Visiting Professor at the Madras School of Economics, Anna University, Chennai, India. She has held positions such as such as Deputy Dean (Research and Postgraduate Studies), Acting Dean, Deputy Head of Department, Head of Discipline, Higher Degree Research Convenor and the Director of the Statistics and Research Design Support Unit at Griffith University. Saroja was educated at the University of Jaffna, Murdoch University and The University of Western Australia, and has published several research monographs and a number of research papers in international refereed journals such as the *Review of Economics and Statistics*, *Transportation Research*, *Economics Letters*, *Empirical Economics, Economic Modelling, Tourism Economics and Applied Economics*. She has received several awards, including *Vice Chancellor's Research Excellence Award 2019 – Higher Degree Research Supervision*, *Pro-Vice Chancellor's Research Excellence Award 2018 – Research Supervision*, *Griffith Awards for Excellence in Teaching 2011 – Group Learning and Teaching Citation*, and *Griffith Awards for Excellence in Teaching 2005 – Innovation Across the Institution*.

Professor Gerald Keller

Dr Gerald Keller is Emeritus 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, BSTAT, Essentials of Business Statistics (co-authored), Australian Business Statistics (co-authored),* and *Statistics Laboratory Manual Experiments Using Minitab,* 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:

L01 describe the two major branches of statistics – descriptive statistics and inferential statistics

L02 understand the key statistical concepts – population, sample, parameter, statistic and census

L03 provide examples of practical applications in which statistics have a major role to play and understand how statistics are used by business managers

L04 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 Statistical applications in business

1.3 How managers use statistics

1.4 Statistics and the computer

1.5 Online resources

SPOTLIGHT ON STATISTICS

Census in Australia

Information (or data) gathering on various characteristics of different populations of interest is 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, a census of the whole population is done once in a while in most countries. A census takes place only every five years in Australia. The peak Australian government statistics agency, the Australian Bureau of Statistics (ABS), carries out the census. The two most recent censuses were carried out on the nights of 9 August 2011 and 2016. 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 and the next chapter.

Source: Australian Bureau of Statistics. © Commonwealth of Australia CC BY 4.0. https://creativecommons.org/licenses/by/4.0/

2016 Data in pictures
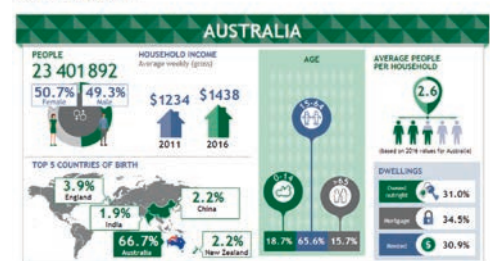
Introduction to statistics

Statistics is a way to get information from data. That's it! Most of this textbook is devoted to describing how, when and why managers and statistics practitioners conduct statistical procedures. You might ask, 'If that's all there is to statistics, why is this book (and most other statistics books) so large?' The answer is that students of applied statistics will be exposed to different kinds of information and data. We demonstrate some of these with case studies and examples throughout this book.

Statistics is a body of principles and methods concerned with extracting useful information from a set of data to help people make decisions. In general, statistics can be subdivided into two basic areas: *descriptive statistics* and *inferential statistics*.

EXAMPLE 1.1

Australia's trade with the world

The total value of imports into Australia continues to exceed the total value of its exports. The Australian government has signed various trade agreements with our major trading partners to relax certain trade restrictions and increase our exports. Over the past decade, successive governments have pledged to reduce the trade deficit (= Imports – Exports) and have only been able to experience the turnaround recently. The picture below shows some useful information on Australian imports and exports in 2017–18 with its major trading partners. For example, the United States (US) is one of our major trading partners and our total imports from the US are about two and a half times our exports to the US. We will discuss various descriptive graphical and numerical techniques to summarise such data in the following chapters.

A$403.2 BILLION
TOTAL VALUE OF EXPORTS 2017–18

TOP 5 EXPORT MARKETS A$M

China	123 274
Japan	51 328
Republic of Korea	23 628
United States	21 424
India	21 145

TOP 5 EXPORTS A$M

Iron ores	61 357
Coal	60 356
Education-related travel	32 434
Natural gas	30 907
Personal travel	21 580

A$395.4 BILLION
TOTAL VALUE OF IMPORTS 2017–18

TOP 5 IMPORT MARKETS A$M

China	71 346
United States	48 752
Republic of Korea	28 674
Japan	26 267
Germany	18 185

TOP 5 IMPORTS A$M

Personal travel	42 496
Passenger motor vehicles	23 299
Refined petroleum	21 655
Ships, boats and floating structures	14 897
Telecom and parts	13 412

AUSTRALIA TRADING WITH THE WORLD 2017–18

Source: Department of Foreign Affairs and Trade CC BY 3.0 AU https://creativecommons.org/licenses/by/3.0/au/legalcode

EXAMPLE 1.2

Summarising the 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 the exam marks for last year's business statistics course. Because, like all statistics lecturers, this one is friendly and helpful, the lecturer obliges and provides a list of the final marks. The overall marks are compiled from 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*.

Descriptive statistics

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 that 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 descriptive statistical technique we use depends on what specific information we would like to extract from a given data set. In Example 1.2, there are at least three important pieces of summary information. The first is the 'typical' mark. We call this a *measure of central location*. The average is one such measure we will introduce in Chapter 5. The student calculated the average mark of the business statistics course, from the list of marks provided by the lecturer, by summing all the marks and dividing the total by the number of students, which was 74. In Chapter 5 we will also introduce another useful measure of central location, the median. In the above example, the median is 81, which is the middle mark of the class when the marks are arranged in ascending or descending order. That is, 50% of the students obtained marks less than the median mark, while 50% received marks greater than the median value.

Now the student knows that the average mark was 74. 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 74 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 mark from the largest. The student noticed that the largest mark is 100 and the smallest is 11, thus the range is (100 − 11) = 89. Unfortunately, this provides little information, as the range doesn't indicate where most of the marks are located. Whether most data are located near 11 or near 100 or somewhere in the middle, the range is still 89. He needs other measures of variability, such as the variance and standard deviation, to reflect the true picture of the spread of the data. These 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 11 and 100. The best way to do this is to use a graphical technique, the histogram, which is introduced in Chapter 4.

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

EXAMPLE 1.3 `LO1` `LO3`

Comparing weekly sales between two outlets

A fast-food franchisee 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
Methods used to draw conclusions about a population based on information provided by a sample of the population.

Inferential statistics

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

EXAMPLE 1.4 `LO1` `LO2` `LO3` `LO4`

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 answer this question, we need to first 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 responses from 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 2 million 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 policy; 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 the new policy.

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 1.4 to determine, for example, the likelihood that less than 10% of the population of its members would purchase the policy, given that 12% of the members sampled said they would purchase. If this likelihood is deemed small enough, the automobile club would 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.

1.1a Population

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 1.4, the population of interest consists of all 2 million or so members.

A descriptive measure of a population is called a parameter. The parameter of interest in Example 1.4 was the proportion of all members who would purchase the new policy.

population
The set of all items of interest.

parameter
A descriptive measure of a population.

1.1b Sample

A sample is a subset of data drawn from the target or studied population. In Example 1.4, 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 1.4, the proportion $(\hat{p})$ 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 (p) who would purchase the life insurance policy. Unlike a parameter, which is a constant, a statistic is a variable whose value varies from sample to sample. In Example 1.4, 12% is a value of the sample statistic based on the selected sample.

sample
A set of data drawn from the studied population.

statistic
A descriptive measure of a sample.

1.1c Statistical inference

Statistical inference is the process of making an estimate, prediction/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 draw conclusions about the population parameters based on sample statistics. In Example 1.4, we make a conclusion about the population proportion, p, based on the sample proportion $\hat{p}$.

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. As decisions involving large amounts of money often hinge on statistical inferences, the reliability of the inference 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 was repeated a very large number of times. For example, a 95% confidence level would mean that for 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 10 and 12. The outcome of the 2019 Australian federal election is an example of how conclusions or predictions can go wrong if the samples selected do not represent the target population.

1.2 Statistical applications in business

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 analyse 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 are also presented) to find out the conclusions and results.

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 4.2 uses graphical techniques to depict the cross-country spread of the global coronavirus pandemic. In Case 5.5 the objective is to compare the central location and variability of the house prices in Sydney and Melbourne. In Case 14.1 it is to compare two populations, the variable of interest being the salary of MBA graduates specialising in marketing and finance. Case 16.1 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 17.3 illustrates another statistical objective. In this case, we need to analyse the relationship between two variables: rate of unemployment and the rate of inflation in New Zealand. By applying the appropriate statistical technique, we will be able to determine whether the two variables are related. As you will discover, the technique also permits statistics practitioners to include other variables to analyse the relationship between the rate of unemployment and inflation.

Case Studies

CASE 3.6 Differing average weekly earnings of men and women in Australia

Although 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 cash 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 full-time employees total) cash earnings (A$), 2018

State/Territory	Males	Females
New South Wales	1807.80	1502.00
Victoria	1696.20	1490.30
Queensland	1777.60	1409.50
South Australia	1585.40	1393.30
Western Australia	2040.00	1495.30
Tasmania	1565.50	1320.00
Northern Territory	1877.90	1552.60
ACT	1977.90	1671.60
Australia	**1823.70**	**1534.10**

Source: Australian Bureau of Statistics, *Average Weekly Earnings,* February 2019, cat. no. 6302.0, ABS, Canberra.

CASE 4.2 Analysing the spread of the Global Coronavirus Pandemic

The coronavirus (COVID-19) pandemic outbreak was first identified in Wuhan, China, in December 2019 and by June 2020 it had spread to more than 200 countries, infected more than 10 million people and resulted in the death of more than 500 000 people globally. Daily data for the number of confirmed cases and deaths are recorded.[1] Depict graphically the number of confirmed cases and number of deaths during the six-month period for the top 10 affected countries and globally.

CASE 5.5 Sydney and Melbourne lead the way in the growth in house prices

A number of recent reports, including those published by the Commonwealth Bank, and two other national housing reports indicated that there is some improvement in house sales activity in Australia combined with a sharp increase in house prices. Some economists attributed this housing market recovery to the low interest rates and first-home owner grants and exemption from stamp duties in some states. The data below show the median (established) house prices in the capital cities of the six states and two territories in Australia for the March quarter of 2002–19. Prepare a summary report using the data below, analysing the movement in house prices in Australia as a whole and in the six states and two territories.

---

1   https://ourworldindata.org/coronavirus-source-data, accessed 29 June 2020

Median house prices ($'000) in eight Australian capital cities, 2002–19, March quarter

Year	Sydney	Melbourne	Brisbane	Adelaide	Perth	Hobart	Darwin	Canberra
2002	365.0	241.0	185.0	166.0	190.0	123.3	190.0	245.0
2003	434.5	270.0	225.0	209.4	216.0	145.0	198.0	300.0
2004	523.0	305.0	302.7	250.0	255.0	200.0	239.5	375.0
2005	486.0	310.0	312.0	270.0	290.0	240.0	275.0	374.7
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
2016	850.0	568.0	483.5	436.0	518.0	360.0	570.0	590.0
2017	930.0	657.5	510.0	451.3	510.0	389.5	530.0	654.5
2018	958.0	730.3	526.0	465.0	510.0	450.0	495.0	699.0
2019	875.0	685.0	533.5	474.0	499.0	471.5	480.0	690.0

Source: Australian Bureau of Statistics, March 2018, *Residential Property Price Indexes: Eight Capital Cities, March Quarter 2018*, cat. no. 6416.0, ABS, Canberra, www.abs.gov.au/AUSSTAT/abs@.nsf/DetailsPage/6416.0 Dec %202018?opendocument.

CASE 14.1 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. For each, she recorded the highest salary offer (including benefits). Can we infer that MBAs with finance majors obtain higher salary offers than do MBAs with marketing majors? Verify the underlying assumptions.

CASE 16.1 Gold lotto

Gold lotto is a national lottery that operates as follows. Players select eight different numbers (six primary and two supplementary numbers) between 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 since Draw 413 (Saturday 6 July 1985). The data recorded in the following table appeared in the *Saturday Gold Lotto* website (https://www.thelott.com/saturday-gold-lotto/results) after the completion of draw 4047 (2 May 2020). 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 (as at 2 May 2020)

Lotto number	Number of times drawn	Lotto number	Number of times drawn	Lotto number	Number of times drawn
1	361	16	324	31	318
2	312	17	301	32	320
.	.	.	.	.	.
.	.	.	.	.	.
14	301	29	320	44	290
15	333	30	301	45	309

CASE 17.3 Does unemployment affect inflation in New Zealand?

A social science research student is interested in investigating the relationship between unemployment and inflation in New Zealand. The student found quarterly data on the rate of unemployment and the quarterly rate of inflation for the period 2015(1) to 2019(1) on the government website, *Statistics New Zealand*. Can you help him to establish the relationship between the two variables using the data he has collected?

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 a manager acquires data, they 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, Coca-Cola Life, 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, Coca-Cola Life and Diet Coke are 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, the Coca-Cola Company 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.