BUSINESS STATISTICS




For Contemporary Decision Making

BLACK • CHAKRAPANI • CASTILLO

Second Canadian Edition

WILEY

WileyPLUS

WileyPLUS is a research-based online environment for effective teaching and learning.

WileyPLUS **builds students' confidence because it takes the guesswork out of studying by providing students with a clear roadmap:**

- what to do
- how to do it
- if they did it right

It offers interactive resources along with a complete digital textbook that help students learn more. With *WileyPLUS*, students take more initiative so you'll have greater impact on their achievement in the classroom and beyond.

Now available for Bb Blackboard

For more information, visit www.wileyplus.com

WileyPLUS

ALL THE HELP, RESOURCES, AND PERSONAL SUPPORT YOU AND YOUR STUDENTS NEED!

www.wileyplus.com/resources

1st DAY OF CLASS ... AND BEYOND!

2-Minute Tutorials and all of the resources you and your students need to get started

WileyPLUS Student Partner Program

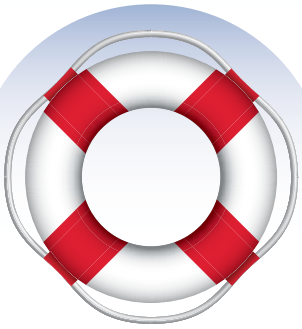
Student support from an experienced student user

Wiley Faculty Network

Collaborate with your colleagues, find a mentor, attend virtual and live events, and view resources
www.WhereFacultyConnect.com

WileyPLUS Quick Start

Pre-loaded, ready-to-use assignments and presentations created by subject matter experts



Technical Support 24/7
FAQs, online chat,
and phone support
www.wileyplus.com/support

Your *WileyPLUS* Account Manager, providing personal training and support

BUSINESS STATISTICS

For Contemporary Decision Making

SECOND CANADIAN EDITION

KEN BLACK

University of Houston–Clear Lake

CHUCK CHAKRAPANI

Ryerson University

IGNACIO CASTILLO

Wilfrid Laurier University

WILEY

For Carolyn, Caycee, and Wendi
Ken

For Murtaza
For his personal and professional friendship
Chuck

For Christian E., Dave G., and Stevie Z.
Thank you for your friendship
Ignacio

Library and Archives Canada Cataloguing in Publication

Black, Ken (Kenneth Urban)
Business statistics: for contemporary decision making / Ken Black, Chuck Chakrapani, Ignacio Castillo. — Second Canadian edition.
Includes index.
ISBN 978-1-118-47695-6 (bound)
1. Commercial statistics—Textbooks. I. Chakrapani, Chuck II. Castillo, Ignacio III. Title.
HF1017.B5722 2014 519.502′465 C2013-905137-6

Production Credits
Acquisitions Editor: Darren Lalonde
Vice President and Publisher: Veronica Visentin
Marketing Manager: Anita Osborne
Editorial Manager: Karen Staudinger
Production Manager: Tegan Wallace
Developmental Editor: Daleara Jamasji Hirjikaka
Media Editor: Channade Fenandoe-Alli
Editorial Assistant: Luisa Begani
Cover and Interior Design: Joanna Vieira
Production and Typesetting: Aptara
Cover Photographs: Shelly Perry, Getty Images (background image),
L-R: ©istockphoto.com/EduardHarkonen; ©istockphoto.com/a-wrangler;
©istockphoto.com/tumpikuja; ©istockphoto.com/tumpikuja
Printing and Binding: Courier
Printed and bound in the United States of America

1 2 3 4 5 CC 18 17 16 15 14

John Wiley & Sons Canada, Ltd.
5353 Dundas Street West, Suite 400
Toronto, ON, M9B 6H8
Canada

Visit our website at: www.wiley.ca

ABOUT THE AUTHORS

Chuck Chakrapani is Distinguished Visiting Professor at the Ted Rogers School of Management, at Ryerson University, President of Leger Analytics, and Chief Knowledge Officer of the Blackstone Group in Chicago.

He was formerly the Chief Executive Officer of Millward Brown Canada, past President and Executive Director of the PMRS (now Marketing Research and Intelligence Association), and Chairman of the Investors Association of Canada. Dr. Chakrapani has held academic appointments at the London Business School in England, and at the universities of Liverpool, Guelph, and McMaster. He was also Editor-in-Chief of the *Canadian Journal of Marketing Research* and Editor of *Marketing Research*, a quarterly publication of the American Marketing Association. Dr. Chakrapani is a Fellow of the Royal Statistical Society and is elected Fellow of the Marketing Research and Intelligence Association for his outstanding contributions to marketing research. He has been an invited speaker in England, the United States, Australia, Canada, India, and the Middle East.

He is the author of more than 10 books and 500 articles on marketing research and investments. His recent books include *Statistics in Market Research* (2004) and *Marketing Research: State of the Art Perspectives* (2000). Dr. Chakrapani is an active academic, business consultant, and seminar leader. He works internationally and his clients include several multinational companies, legal firms, marketing boards, and government bodies. He can be reached at chuck.chakrapani@gmail.com or through his website ChuckChakrapani.com.

Ignacio Castillo is professor of operations and decision sciences at the School of Business & Economics, Wilfrid Laurier University. He received a Ph.D. in industrial engineering from Texas A&M University in 2001, an M.S.E. in industrial engineering from Arizona State University in 1998, and a B.S. (magna cum laude) in applied sciences from Universidad San Francisco de Quito, Ecuador, in 1995.

His research and teaching interests include business statistics, facility location, facility layout and material handling systems, manufacturing and service operations and logistics, and sustainable and closed-loop supply chain management. Dr. Castillo has consulted for several public and private organizations, has supervised over 15 graduate theses or projects, and has produced over 20 articles that have appeared in prestigious academic journals. His research has been mainly supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Dr. Castillo has served as adjunct faculty at the University of Alberta and the University of Waterloo and is a member of the Alpha Pi Mu Industrial Engineering Honor Society, The Honor Society of Phi Kappa Phi, Pinnacle Honor Society, and a Glenn Carroll Teaching Fellow.

Ken Black is currently professor of decision sciences in the School of Business at the University of Houston–Clear Lake. Born in Cambridge, Massachusetts, and raised in Missouri, he earned a bachelor's degree in mathematics from Graceland University, a master's degree in math education from the University of Texas at El Paso, a Ph.D. in business administration in management science, and a Ph.D. in educational research from the University of North Texas.

Since joining the faculty of UHCL in 1979, Professor Black has taught all levels of statistics courses, forecasting, management science, market research, and production/operations management. In 2005, he was awarded the President's Distinguished Teaching Award for the university. He has published over 20 journal articles and 20 professional papers, as well as two textbooks: *Business Statistics: An Introductory Course* and *Business Statistics for Contemporary Decision Making.* Professor Black has consulted for many different companies, including Aetna, the city of Houston, NYLCare, AT&T, Johnson Space Center, Southwest Information Resources, Connect Corporation, and Eagle Engineering.

Ken Black and his wife, Carolyn, have two daughters, Caycee and Wendi. His hobbies include playing the guitar, reading, travelling, and running.

PREFACE

The second Canadian edition of *Business Statistics for Contemporary Decision Making* continues the tradition of using clear, complete, and student-friendly pedagogy to present and explain business statistics topics. The vast ancillary resources available through *WileyPLUS* complement the text in helping instructors effectively deliver this subject matter and assisting students in their learning.

We wrote the chapters in the book to help students construct their knowledge of the big picture of statistics, provide assistance as needed, and afford more opportunities to practise statistical skills. The 19 chapters in this book are organized into five units to facilitate student understanding of the bigger view of statistics. There are 21 high-quality video tutorials with Ken Black that explain difficult key topics and demonstrate how to work problems from challenging sections of the text.

This book is written and designed for a two-semester introductory undergraduate business statistics course or an MBA-level introductory course. In addition, with 19 chapters, this text lends itself nicely to adaptation for a one-semester introductory business statistics course. The text is written with the assumption that the student has a university algebra mathematical background. No calculus is used in the presentation of material in the text.

An underlying philosophical approach to the text is that every statistical tool presented in the book has some business application. While the text contains statistical rigour, it is written so that the student can readily see that the proper application of statistics in the business world goes hand in hand with good decision making. In this edition, statistics are presented as a means for converting data into useful information that can be used to assist the business decision maker in making more thoughtful, information-based decisions. Thus, the text presents business statistics as "value-added" tools in the process of converting data into useful information.

CHANGES FOR THE SECOND EDITION

UNITS AND CHAPTERS

In the second edition, the unit and chapter organization remains the same as in the first edition—there are 19 chapters organized into five units. The purpose of the unit organization is to locate chapters with similar topics together, thereby increasing the likelihood that students will grasp the bigger picture of statistics.

Unit I, Introduction, contains the first four chapters of the text. In this unit, students learn some important foundational tools for understanding topics presented in the rest of the course. In Unit II, Distributions and Sampling, consisting of Chapters 5 through 7, students are presented with six population distributions and two sampling distributions. In Unit III, Making Inferences about Population Parameters, which includes Chapters 8 through 11, students learn about estimating and testing population parameters. Unit IV, Chapters 12 through 15, is called Regression Analysis and Forecasting. In this unit, students explore relationships between variables, including developing models to predict a variable by other variables and developing models to forecast. Unit V, Special Topics, includes nonparametric statistics and quality, which are covered in Chapters 16 through 18. In these chapters, students are presented with a series of well-known nonparametric techniques along with a number of quality-improvement concepts and techniques. This unit also includes Chapter 19 on decision analysis.

Topical Changes

Sections and topics from the first Canadian edition remain virtually unchanged in the second edition, with some exceptions. In Chapter 2, there is a new section on Charts and Graphs for Two Variables. Chapter 13 includes a new section titled Using Regression Analysis: Some Caveats, which helps the student to understand the limitations of multiple regression analysis. And finally, Chapter 14 includes a new section on Logistic Regression.

Decision Dilemma and Decision Dilemma Solved

The Decision Dilemmas are real business vignettes that open each chapter. They set the tone for the chapter by presenting a business dilemma and asking a number of managerial or statistical questions, the solutions to which require the use of techniques presented in the chapter. The Decision Dilemma Solved feature discusses and answers the managerial and statistical questions posed in the Decision Dilemma using techniques from the chapter, thus bringing closure to the chapter. In the second edition, all decision dilemmas have been revised and updated. Solutions given in the Decision Dilemma Solved features have been revised for new data and for new versions of computer output.

In addition, a new Decision Dilemma has been introduced in the second edition. In Chapter 10, we present a Decision Dilemma on a company, L. L. Bean, that is primarily a catalogue and on-line retailer. L. L. Bean of Freeport, Maine, started out as a one-man operation in 1912 selling hunting boots and has grown into an operation that ships catalogues to over 160 countries, with more than 11 million customer contacts per year. In this Decision Dilemma, facts and figures are given on customer contacts, orders placed on-line, the average order size, and the time that L. L. Bean's m-commerce site home page takes to load. In the managerial and statistical questions, students are asked such questions as, "Is the average order size for women greater than the average order size for men?"

Statistics in Business Today

This book includes one or two Statistics in Business Today features in every chapter. This feature presents a real-life example of how the statistics presented in that chapter apply in the business world today. Several of these features have been revised and in some cases replaced with more relevant issues. Some of these include "Plastic Bags vs. Bringing Your Own," "Ethical Differences between Men and Women," and "Risk Taking by Ad Agencies."

Cases

Every chapter in this text contains a unique business **c**ase. All cases in the second edition have been updated for today's market. These business cases are more than just long problems, and in the discussion that follows the business scenario, several issues and questions are posed that can be addressed using techniques presented in the chapter. Some new cases have been added to the second edition, including ones on the Container Store, Coca-Cola in the African market, Whole Foods, and Ceapro.

Problems

Throughout each chapter are problems and demonstration problems for students to apply their new-found knowledge before moving on to other sections. We examined the problems

included in the text for timeliness, appropriateness, and logic before we included them in the second edition. Those that fell short were replaced or rewritten. In the second edition, there are 49 new problems. While the total number of problems in the text is 964, a concerted effort has been made to include only problems that make a significant contribution to the learning process.

All demonstration problems and example problems were thoroughly reviewed and edited for effectiveness. A demonstration problem is an extra example containing both a problem and its solution and is used as an additional pedagogical tool to supplement explanations and examples in the chapters. Virtually all example and demonstration problems in this edition are business oriented and contain current data available to us.

Problems are located at the end of most sections in the chapters. A significant number of additional problems are provided at the end of each chapter in the Supplementary Problems. The Supplementary Problems are "scrambled"—problems using the various techniques in the chapter are mixed—so that students can test themselves on their ability to discriminate and differentiate ideas and concepts.

Databases

Available with the second edition are six databases that provide additional opportunities for students to apply the statistics presented in the text. These six databases represent a variety of business areas, such as the stock market, international labour, finance, energy, agri-business, and new to this edition, the registered retirement savings plan (RRSP) database. The data are gathered from such reliable sources as Statistics Canada, the Toronto Stock Exchange, the Fraser Institute, and the Global Environment Outlook (GEO) Data Portal.

STUDENT VIDEO SUPPORT

With the advent of on-line business statistics courses, increasingly large class sizes, and the number of commuter students who have very limited access to educational resources on business statistics, it is often difficult for students to get the learning assistance that they need to bridge the gap between theory and application on their own. Two innovative features of the package to address this issue are:

VIDEO TUTORIALS BY KEN BLACK

An exciting feature of the package that will enhance the effectiveness of student learning in business statistics and significantly enhance the presentation of course material is the series of video tutorials by Ken Black.

There are 21 video tutorial sessions on key difficult topics in business statistics delivered by Ken Black. The tutorials are:

1. Chapter 1: Levels of Data Measurement
2. Chapter 2: Stem-and-Leaf Plot
3. Chapter 3: Computing Variance and Standard Deviation
4. Chapter 3: Understanding and Using the Empirical Rule
5. Chapter 4: Constructing and Solving Probability Matrices
6. Chapter 4: Solving Probability Word Problems
7. Chapter 5: Solving Binomial Distribution Problems, Part I

8. Chapter 5: Solving Binomial Distribution Problems, Part II
9. Chapter 6: Solving Problems Using the Normal Curve
10. Chapter 7: Solving for Probabilities of Sample Means Using the z Statistic
11. Chapter 8: Confidence Intervals
12. Chapter 8: Determining which Inferential Technique to Use: Confidence Intervals
13. Chapter 9: Hypothesis Testing Using the z Statistic
14. Chapter 9: Establishing Hypotheses
15. Chapter 9: Understanding p Values
16. Chapter 9: Type I and Type II Errors
17. Chapter 9: Two-Tailed Tests
18. Chapter 10: Hypothesis Tests of the Difference in Means of Two Independent Populations Using the t Statistic
19. Chapter 11: Computing and Interpreting a One-Way ANOVA
20. Chapter 12: Testing the Regression Model I—Predicted Values, Residuals, and Sum of Squares of Error
21. Chapter 12: Testing the Regression Model II—Standard Error of the Estimate and r^2

OFFICE HOURS VIDEOS BY IGNACIO CASTILLO

New in the second edition are Office Hours Videos, problem-solving video tutorials based on chapter problems prepared by Canadian author Ignacio Castillo. Icons in the text indicate problems that are accompanied by on-line Office Hours Videos.

The videos by Ken Black and Ignacio Castillo are available for adopters on *WileyPLUS* and can easily be uploaded for classroom usage to augment lectures and enrich classroom presentations. Each video is around 10 minutes in length.

FEATURES AND BENEFITS

CHAPTER FEATURES

Each chapter of the book contains sections called Learning Objectives, a Decision Dilemma, Demonstration Problems, Section Problems, Concept Checks, Points of Interest, Statistics in Business Today, Decision Dilemma Solved, Key Considerations, Why Statistics Is Relevant, Chapter Summary, Key Terms, Formulas, Supplementary Problems, Analyzing the Databases, Case, and Using the Computer.

- **Learning Objectives.** Each chapter begins with a statement concerning the chapter's main learning objectives. This statement gives the reader a list of key topics that will be discussed and the goals to be achieved from studying the chapter.
- **Decision Dilemma.** At the beginning of each chapter, a short case describes a real company or business situation in which managerial and statistical questions are raised. In most Decision Dilemmas, actual data are given and the student is asked to consider how the data can be analyzed to answer the questions.
- **Demonstration Problems.** Virtually every section of every chapter in the book edition contains demonstration problems. A demonstration problem contains both an example problem and its solution, and is used as an additional pedagogical tool to supplement explanations and examples.
- **Section Problems.** Problems for practice are found at the end of almost every section of the text. Most problems use real data gathered from a plethora of sources. Included here are two brief excerpts from some of the real-life problems in the text: "The Canadian

Beef Export Federation reports that the top six destinations for Canadian beef in a recent year were the U.S. with $1,697 million, Mexico with $269 million, Japan with $171 million, South Korea with $28 million, Taiwan with $16 million, and China with $4 million." "Data accumulated by Environment Canada show that the average wind speed in kilometres per hour for Victoria International Airport, located on the Saanich Peninsula in British Columbia, is 9.3."

- **Concept Checks.** Concept checks are conceptual questions aimed to reinforce the conceptual understanding of the material presented. They are presented following each main section of the chapter. These questions do not require students to use calculations.
- **Points of Interest.** Points of interest appear throughout the chapters to provide helpful tips and further explanation regarding key concepts.
- **Statistics in Business Today.** Every chapter in the book contains at least one Statistics in Business Today feature. These focus boxes contain an interesting application of how techniques from that particular chapter are used in the business world today. They are usually based on real companies, surveys, or published research.
- **Decision Dilemma Solved.** Situated at the end of the chapter, the Decision Dilemma Solved feature addresses the managerial and statistical questions raised in the Decision Dilemma. Data given in the Decision Dilemma are analyzed computationally and by computer using techniques presented in the chapter. Answers to the managerial and statistical questions raised in the Decision Dilemma are arrived at by applying chapter concepts, thus bringing closure to the chapter.
- **Key Considerations.** Each chapter contains a Key Considerations feature that is very timely, given that many users of statistics are unaware of the traps that await the unwary. With the abundance of statistical data and analysis, there is also considerable potential for the misuse of statistics in business dealings. The important Key Considerations feature underscores this potential misuse by discussing such topics as lying with statistics, failing to meet statistical assumptions, and failing to include pertinent information for decision makers. Through this feature, instructors can begin to integrate the topic of ethics with applications of business statistics. Here are two excerpts from Key Considerations features: "Invalid or spurious results can be obtained by using the parameters from one population to analyze another population. For example, a market study in Nova Scotia may result in the conclusion that the amount of fish eaten per month by adults is normally distributed with an average of 1 kg of fish per month. A market researcher in Manitoba should not assume that these figures apply to her population. People in Manitoba probably have quite different fish-eating habits than people in Nova Scotia," and "In describing a body of data to an audience, it is best to use whatever measures it takes to present a full picture of the data. By limiting the descriptive measures used, the business researcher may give the audience only part of the picture and can skew the way the receiver understands the data."
- **Why Statistics Is Relevant.** This section in each chapter discusses how the material discussed in that chapter is related to real-life decision making.
- **Chapter Summary.** Each chapter concludes with a summary of the important concepts, ideas, and techniques of the chapter. This feature can serve as a preview of the chapter as well as a chapter review.
- **Key Terms.** Important terms appear in colour and their definitions, where possible, are italicized throughout the text as they are discussed. At the end of the chapter, a list of the key terms from the chapter is presented. In addition, these terms appear with their definitions in the end-of-book Glossary.
- **Formulas.** Important formulas in the text are highlighted to make it easy for students to locate them. In several chapters, formulas are numbered for ease of reference and clarity if they are referred to more than once. At the end of each chapter, most of the chapter's formulas are listed together as a handy reference.
- **Supplementary Problems.** At the end of each chapter is an extensive set of additional problems. The Supplementary Problems are divided into three groups: *Calculating the Statistics*, which are strictly computational problems; *Testing Your Understanding*, which are problems for application and understanding; and *Interpreting the Output*, which are problems that require the interpretation and analysis of software output.

- **Analyzing the Databases.** There are six major databases located on the student companion website that accompanies the book. The end-of-chapter Analyzing the Databases section contains several questions/problems that require the application of techniques from the chapter to data in the variables of the databases. It is assumed that most of these questions/problems will be solved using a computer.
- **Case.** Each end-of-chapter case is based on a real company, many featuring Canadian businesses. These cases give the student an opportunity to use statistical concepts and techniques presented in the chapter to solve a business dilemma. Some cases feature very large companies, such as Starbucks, McCain Foods Limited, or Shell Oil. Others pertain to smaller businesses, such as Delta Wire or Ceapro, that have overcome obstacles to survive and thrive. Most cases include raw data for analysis and questions that encourage the student to use several of the techniques presented in the chapter. In many cases, the student must analyze software output in order to reach conclusions or make decisions.
- **Using the Computer.** The Using the Computer section contains directions for producing the Excel 2010 software output presented in the chapter. It is assumed that students have a general understanding of a Microsoft environment. Directions include specifics about menu bars, drop-down menus, and dialogue boxes. Each dialogue box is not discussed in detail; the intent is to provide enough information for students to produce the same statistical output analyzed and discussed in the chapter.

TREE DIAGRAM OF INFERENTIAL TECHNIQUES

To assist the student in sorting out the plethora of confidence intervals and hypothesis testing techniques presented in the text, tree diagrams are presented at the beginning of Unit III and Chapters 8, 9, 10, and 17. The tree diagram at the beginning of Unit III displays virtually all of the inferential techniques presented in Chapters 8–10 so that the student can construct a view of the "forest for the trees" and determine how each technique plugs into the whole. Then at the beginning of each of these three chapters, an additional tree diagram is presented to display the branch of the tree that applies to techniques in that particular chapter. Chapter 17 includes a tree diagram for just the nonparametric statistics presented in that chapter.

In determining which technique to use, there are several key questions that a student should consider. Listed here are some of the key questions (displayed in a list in the Unit III introduction) that delineate what students should ask themselves in determining the appropriate inferential technique for a particular analysis: Does the problem call for estimation (using a confidence interval) or testing (using a hypothesis test)? How many samples are being analyzed? Are you analyzing means, proportions, or variances? If means are being analyzed, is (are) the variance(s) known or not? If means from two samples are being analyzed, are the samples independent or related? If three or more samples are being analyzed, are there one or two independent variables and is there a blocking variable?

WileyPLUS

WileyPLUS is an innovative, research-based on-line environment for effective teaching and learning.

WileyPLUS builds students' confidence because it takes the guesswork out of studying by providing students with a clear roadmap: **what to do, how to do it, if they did it right**. Students will take more initiative so you'll have greater impact on their achievement in the classroom and beyond.

Among its many features, this on-line learning interface allows students to study and practise using the digital textbook, quizzes, and algorithmic exercises. The immediate feedback helps students understand where they need to focus their study efforts.

In *WileyPLUS*, students will find various helpful tools, such as an ebook, the student study manual, videos with tutorials by the authors, applets, Decision Dilemma and Decision Dilemma Solved animations, learning activities, flash cards for key terms, demonstration problems, databases in Excel, case data in Excel, and problem data in Excel.

- **Ebook.** The complete text is available on *WileyPLUS* with learning links to various features and tools to assist students in their learning.
- **Videos.** There are 21 videos of one of the authors explaining concepts and demonstrating how to work problems for some of the more difficult topics.
- **Office Hours Videos.** These are worked video examples of selected problems from the text prepared by Canadian author Ignacio Castillo.
- **Applets.** Statistical applets are available, affording students the opportunity to learn concepts by iteratively experimenting with various values of statistics and parameters and observing the outcomes.
- **Learning Activities.** There are numerous learning activities to help students better understand concepts and key terms. These activities have been developed to make learning fun, enjoyable, and challenging.
- **Data Sets.** Virtually all problems in the text along with the case problems and the databases are available to students in Excel format.
- **Animations.** To aid students in understanding complex interactions, selected figures from the text that involve dynamic activity have been animated using Flash technology. Students can download these animated figures and run them to improve their understanding of dynamic processes.
- **Kaddstat.** Kaddstat is an easy-to-use add-in to Excel that makes it easier to run complex statistical tests on Excel.
- **Flash Cards.** Key terms will be available to students in a flash card format along with their definition.
- **Student Study Guide.** The study guide contains complete answers to all odd-numbered questions.
- **Demonstration Problems.** Step-by-step solved problems are given for each chapter.

ANCILLARY TEACHING AND LEARNING MATERIALS

www.wiley.com/go/blackcanada **WileyPLUS**

FOR INSTRUCTORS

Several useful supplements and resources are offered on the book's companion website and in *WileyPLUS*. On these sites, instructors will find the Solutions Manual, PowerPoint presentations, Test Bank, Instructor's Manual, Computerized Test Bank, and other valuable teaching resources. The supplements are prepared by subject matter experts and contributors who are often users of the text.

FOR STUDENTS

On the book's companion website and *WileyPLUS*, students will find support materials that will help them develop their conceptual understanding of class material and increase their ability to solve problems. In addition to other resources, students will find the databases accompanying the text, problem files in Excel, formulas, student solutions manual, and other useful supplements.

ACKNOWLEDGEMENTS

John Wiley & Sons Canada, Ltd. and the authors would like to thank the reviewers who cared enough and took the time to provide us with their excellent insights and advice, which was used to shape and mould the second Canadian edition. Our appreciation goes to:

Fouzia Baki, McMaster University
Sylvie Bequet, Bishop's University
Andrew Flostrand, Simon Fraser University
Paramjit Gill, University of British Columbia, Okanagan
Wayne Horn, Carleton University
Muhammad Hossain, MacEwan University
Tracy Jenkin, Queen's University
Denis Keroack, The Northern Alberta Institute of Technology
Ehsan Latif, Thompson Rivers University
Samie Li Shang Ly, Concordia University
Mohammad Mahbobi, Thompson Rivers University
Stan Miles, Thompson Rivers University
Peggy Ng, York University
Cecilia Rodriguez, The Northern Alberta Institute of Technology
Keith Rogers, Queen's University
Brian Smith, McGill University
Don St. Jean, George Brown College
Augustine Wong, York University
Morty Yalovsky, McGill University

Special thanks to our contributors for preparing the various supplements accompanying the text:

Alexei Gokhman, Humber College
Denis Keroack, The Northern Alberta Institute of Technology
Ioulia Kim
Lisa MacKay, SAIT Polytechnic
Cecilia Rodriguez, The Northern Alberta Institute of Technology

There are several people working at or with John Wiley & Sons Canada, Ltd., in bringing this edition to fruition whom we would like to thank for their invaluable assistance on the project. They include Darren Lalonde, Acquisitions Editor; Anita Osborne, Marketing Manager; Karen Staudinger, Editorial Manager; Daleara Hirjikaka, Developmental Editor; Penny Walker, Production Editor; Deanna Durnford, Supplements Coordinator; Laurel Hyatt, Copy Editor; Zofia Laubitz, Proofreader; and Belle Wong, Indexer.

Chuck Chakrapani
Ignacio Castillo

Revising a book is a lonely job. I am thankful for the assistance I received from the Wiley team and from Tansel Erguden of CSCA at Ryerson.

- Chuck

I would like to thank Dr. Erhan Erkut (Özyeğin University, Turkey) and Dr. Hamid Noori (Wilfrid Laurier University) for their support and mentorship over the years. I also would like to thank my parents Ignacio and Lilián, my brother Andrés, my sister Gabriela, my wife Ivonne, and my sons José and Emilio for their unconditional love.

- Ignacio

BRIEF CONTENTS

CONTENTS

UNIT V · SPECIAL TOPICS

Unit I

INTRODUCTION

The study of business statistics is important, valuable, and interesting. However, because it involves a new language of terms, symbols, logic, and mathematics, it can be daunting. For many students, this text is their first and only introduction to business statistics, which instructors often teach as a "survey course." That is, the student is presented with an overview of the subject, including many techniques, concepts, and formulas. It can be overwhelming! One of the main difficulties in studying business statistics in this way is being able to see the forest for the trees; that is, to sort out the myriad topics so they make sense. With this in mind, we have organized the chapters of this text into four units, with each unit containing chapters that present similar material. At the beginning of each unit, an introduction presents the overarching themes to those chapters.

Unit I is titled "Introduction" because the four chapters (1–4) contained herein introduce the study of business statistics. In Chapter 1, students will learn the many meanings of the word *statistics,* the concepts of descriptive and inferential statistics, and the levels of data measurement. In Chapter 2, students will see how raw data can be organized using various graphical and tabular techniques to facilitate their use in making better business decisions. Chapter 3 introduces some essential and basic statistics that will be used both as a way to summarize data and as tools for techniques introduced later in the text. The chapter also includes a discussion of distribution shapes and measures of association. In Chapter 4, the basic laws of probability are presented. The notion of probability underlies virtually every business statistics topic, distribution, and technique, thereby making it important to appreciate and understand the subject. In Unit I, the first four chapters, we are developing building blocks that will enable you to understand and apply statistical concepts, so you can analyze data to assist present and future business managers in making better decisions.

CHAPTER 1

INTRODUCTION TO STATISTICS

Learning Objectives

The primary objective of Chapter 1 is to introduce you to the world of statistics thereby enabling you to:

1. Define statistics and list example applications of statistics in business.
2. Define important statistical terms, including population, sample, and parameter, as they relate to descriptive and inferential statistics.
3. Explain the difference between variables, measurement, and data.
4. Compare the four different levels of data: nominal, ordinal, interval, and ratio.

STATISTICS DESCRIBE THE STATE OF BUSINESS IN INDIA'S COUNTRYSIDE

India is the second-most populous country in the world, with more than a billion people. Nearly three quarters of the people live in rural areas, scattered about the countryside in 600,000 villages. In fact, it may be said that more than one in every ten people in the world live in rural India. The population in rural India can be described as poor and semiliterate. With an annual per capita income of less than $1 (U.S.) per day, rural India accounts for only about one third of total national product sales. Less than 50% of households in rural India have electricity, and many of the roads are not paved. The annual per capita consumption of toothpaste is only 30 g per person in rural India, compared to 160 g in urban India, 400 g in the United States, and 320 g as the world average.

However, in addition to the impressive size of its population, there are other compelling reasons for companies to market their goods and services to rural India. The market of rural India has been growing at five times the rate of the urban Indian market. There is increasing agricultural productivity, leading to growth in disposable income, and there is a reduction in the gap between the tastes of urban and rural customers. The literacy level is increasing, and people are becoming more conscious of their lifestyles and of opportunities for a better life. Nearly two thirds of all middle-income households in India are in rural areas, with the number of middle- and high-income households in rural India expected to grow from 80 million to 111 million over the next three years. More than one third of all rural households now have a main source of income other than farming. Virtually every home has a radio, almost 20% have a television, and more than 30% have at least one bank account.

In the early 1990s, toothpaste consumption in rural India doubled, and the consumption of shampoo increased fourfold. Recently, other products have done well in rural India, accounting for nearly one half of all of the country's sales of televisions, fans, bicycles, bath soap, and other products. According to MART, a New Delhi–based research organization, rural India buys 46% of all soft drinks and 49% of motorcycles sold in India. In one year alone, the market for Coca-Cola in rural India grew by 37%, accounting for 80% of new Coke drinkers in India. Because of such factors, many global and Indian firms, such as Microsoft, General Electric, Kellogg's, Colgate-Palmolive, Hindustan-Unilever, Godrej, Nirma, and others, have entered the rural Indian market with enthusiasm. Marketing to rural customers often involves building categories by persuading them to try products that they may not have used before. Rural India is a huge, relatively untapped market for businesses. However, entering such a market is not without risks and obstacles. The dilemma facing companies is whether to enter this marketplace and, if so, to what extent and how.[1]

Managerial and Statistical Questions

1. Are the statistics presented in this report exact figures or estimates?
2. How and where could the researchers have gathered such data?
3. In measuring the potential of the rural Indian marketplace, what other statistics could have been gathered?
4. What levels of data measurement are represented by data on rural India?
5. How can managers use these and other statistics to make better decisions about entering this marketplace?

Every minute of the working day, decisions are made by businesses around the world that determine whether companies will be profitable and grow or whether they will stagnate and die. Most of these decisions are made with the assistance of information gathered about the marketplace, the economic and financial environment, the workforce, the competition, and other factors. Such information usually comes in the form of data or is accompanied by data. Business statistics provides the tools by which such data are collected, analyzed, summarized, and presented to facilitate the decision-making process, and business statistics plays an important role in the ongoing saga of decision making within the dynamic world of business.

1.1 STATISTICS IN BUSINESS

Learning Objective 1

Define statistics and list example applications of statistics in business.

There is a wide variety of uses and applications of statistics in business. Several examples follow.

- A survey of 1,465 workers by Hotjobs found that 55% of workers believe that the quality of their work is perceived the same when they work remotely as when they are physically in the office.
- In a survey of 477 executives by the Association of Executive Search Consultants, 48% of men and 67% of women said they were more likely to negotiate for less business travel compared with five years earlier.
- A survey of 1,007 adults by RBC Capital Markets showed that 37% of adults would be willing to drive 8 to 15 km to save 5 cents on a litre of gas.
- A Deloitte Retail "Green" survey of 1,080 adults revealed that 54% agreed that plastic, non-compostable shopping bags should be banned.
- A survey by Statistics Canada determined that in 2011, the average annual household net expenditure in Canada was $73,457 and that households averaged $3,711 annually on recreation.[2] In addition, when measured on the basis of gross domestic product and real gross national income, living standards in Canada grew at a faster pace than in the United States between 1997 and 2011.[3]
- In a 2008 survey of 14 countries conducted by GlobeScan for the National Geographic Society, Canada ranked 13th out of 14 when it came to environmentally friendly consumption patterns. This was due mostly to Canadian preferences for bigger houses and an established culture of using privately owned cars as opposed to transit.[4]

Note that, in most of these examples, business researchers have conducted a study and provided rich and interesting information that can be used in business decision making.

In this text, we will examine several types of graphs for depicting data as we study ways to arrange or structure data into forms that are both meaningful and useful to decision makers. We will learn about techniques for sampling from a population that allow studies of the business world to be conducted in a less expensive and more timely manner. We will explore various ways to forecast future values and we will examine techniques for predicting trends. This text also includes many statistical tools for testing hypotheses and for estimating population values. These and many other exciting statistics and statistical techniques await us on this journey through business statistics. Let us begin.

1.2 BASIC STATISTICAL CONCEPTS

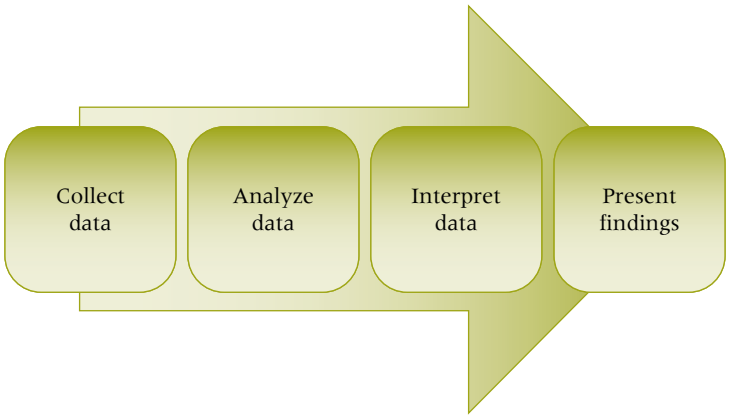
Learning Objective 2

Define important statistical terms, including population, sample, and parameter, as they relate to descriptive and inferential statistics.

Business statistics, like many areas of study, has its own language. It is important to begin our study with an introduction of some basic concepts in order to understand and communicate about the subject. We begin with a discussion of the word *statistics*. This word has many different meanings in our culture. *Webster's Third New International Dictionary* gives a comprehensive definition of **statistics** as *a science dealing with the collection, analysis, interpretation, and presentation of numerical data.* Viewed from this perspective, statistics includes all the topics presented in this text. Figure 1.1 captures the key elements of business statistics.

The study of statistics can be organized in a variety of ways. One of the main ways is to subdivide statistics into two branches: descriptive statistics and inferential statistics. To understand the difference between descriptive and inferential statistics, definitions of *population* and *sample* are helpful. *Webster's Third New International Dictionary* defines **population** as *a collection of persons, objects, or items of interest.* The population can be a widely defined category, such as "all automobiles," or it can be narrowly defined, such as "all Ford Focus cars produced from 2012 to 2013." A population can be a group of people, such as "all workers employed by Microsoft," or it can be a set of objects, such as "all Toyota RAV4s

Figure 1.1
The Key Elements of Statistics



produced in February 2013 by Toyota Canada at the Woodstock, Ontario, plant." The researcher defines the population to be whatever he or she is studying. When researchers *gather data from the whole population for a given measurement of interest,* they call it a **census**. Most people are familiar with the Canadian Census. Every five years, the government attempts to measure all persons living in this country. As another example, if a researcher is interested in ascertaining the grade point average for all students at the University of Toronto, one way to do so is to conduct a census of all students currently enrolled there.

A **sample** is *a portion of the whole* and, if properly taken, is representative of the whole. For various reasons (explained in Chapter 7), researchers often prefer to work with a sample of the population instead of the entire population. For example, in conducting quality control experiments to determine the average life of light bulbs, a light bulb manufacturer might randomly sample only 75 light bulbs during a production run. Because of time and money limitations, a human resources manager might take a random sample of 40 employees instead of using a census to measure company morale.

If a business analyst is *using data gathered on a group to describe or reach conclusions about that same group,* the statistics are called **descriptive statistics**. For example, if an instructor produces statistics to summarize a class's examination results and uses those statistics to reach conclusions about that class only, the statistics are descriptive. The instructor can use these statistics to discuss class average, talk about the range of class scores, or present any other data measurements for the class based on the test.

Most athletic statistics, such as batting average, save percentages, and first downs, are descriptive statistics because they are used to describe an individual or team effort. Many of the statistical data generated by businesses are descriptive. They might include number of employees on vacation during June, average salary at the Edmonton office, corporate sales for 2013, average managerial satisfaction score on a company-wide census of employee attitudes, and average return on investment for the Lofton Company for the years 2004 through 2013.

Another type of statistics is called **inferential statistics**. If a researcher *gathers data from a sample and uses the statistics generated to reach conclusions about the population from which the sample was taken,* the statistics are inferential statistics. The data gathered from the sample are used to infer something about a larger group. Inferential statistics are sometimes referred to as *inductive statistics.* The use and importance of inferential statistics continue to grow.

One application of inferential statistics is in pharmaceutical research. Some new drugs are expensive to produce, and therefore tests must be limited to small samples of patients. Utilizing inferential statistics, researchers can design experiments with small, randomly selected samples of patients and attempt to reach conclusions and make inferences about the population.

Market researchers use inferential statistics to study the impact of advertising on various market segments. Suppose a soft drink company creates an advertisement depicting a dispensing machine that talks to the buyer, and market researchers want to measure the impact of the new advertisement on various age groups. The researcher could stratify the population into age categories ranging from young to old, randomly sample each stratum, and use inferential statistics to determine the effectiveness of the advertisement for the

various age groups in the population. The advantage of using inferential statistics is that they enable the researcher to effectively study a wide range of phenomena without having to conduct a census. Most of the topics discussed in this text pertain to inferential statistics.

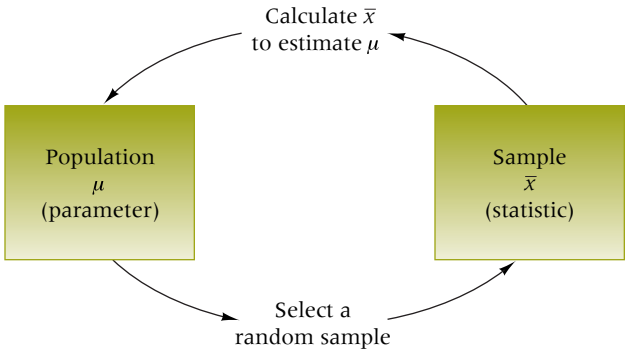
A *descriptive measure of the population* is called a **parameter**. Parameters are usually denoted by Greek letters. Examples of parameters are population mean (μ), population variance (σ^2), and population standard deviation (σ). A *descriptive measure of a sample* is called a **statistic**. Statistics are usually denoted by Roman letters. Examples of statistics are sample mean ($\bar{x}$), sample variance (s^2), and sample standard deviation (s).

Differentiation between the terms *parameter* and *statistic* is important only in the use of inferential statistics. A business researcher often wants to estimate the value of a parameter or conduct tests about the parameter. However, the calculation of parameters is usually either impossible or infeasible because of the amount of time and money required to take a census. In such cases, the business researcher can take a random sample of the population, calculate a statistic on the sample, and infer by estimation the value of the parameter. The basis for inferential statistics, then, is the ability to make decisions about parameters without having to complete a census of the population.

For example, a manufacturer of washing machines would probably want to determine the average number of loads that a new machine can wash before it needs repairs. The parameter is the population mean or average number of washes per machine before repair. A company analyst takes a sample of machines, computes the number of washes before repair for each machine, averages the numbers, and estimates the population value or parameter by using the statistic, which in this case is the sample average. Figure 1.2 demonstrates the inferential process.

Figure 1.2

Process of Inferential Statistics to Estimate a Population Mean (μ)



Inferences about parameters are made under uncertainty. Unless parameters are computed directly from the population, the statistician never knows with certainty whether the estimates or inferences made from samples are true. In an effort to estimate the level of confidence in the result of the process, statisticians use probability statements. For this and other reasons, part of this text is devoted to probability (Chapter 4).

Concept Check

Fill in the blanks.

1. The basic idea of statistics is to extrapolate from a random data _____ you have collected to make general conclusions about the larger _____ from which the random data _____ were derived.
2. Descriptive statistics can be used to _____ the data to describe a data sample either numerically or graphically.
3. Statistical inference is inference about a _____ from a random data _____ drawn from it.

1.3 VARIABLES AND DATA

Learning Objective 3

Explain the difference between variables, measurement, and data.

Business statistics is about measuring phenomena in the business world and organizing, analyzing, and presenting the resulting numerical information in such a way that better, more informed business decisions can be made. Most business statistics studies contain variables, measurements, and data.

In business statistics, a **variable** is *a characteristic of any entity being studied that is capable of taking on different values*. Some examples of variables in business might include return on investment, advertising dollars, labour productivity, stock price, historic cost, total sales, market share, age of worker, earnings per share, kilometres driven to work, time spent in store shopping, and many, many others. In business statistics studies, most variables produce a measurement that can be used for analysis. A **measurement** occurs *when a standard process is used to assign numbers to particular attributes or characteristics of a variable*. Many measurements are obvious, such as the time spent in a store shopping by a customer, the age of the worker, or the number of kilometres driven to work. However, some measurements, such as labour productivity, customer satisfaction, and return on investment, have to be defined by the business researcher or by experts within the field. Once such measurements are recorded and stored, they can be denoted as "data." It can be said that **data** are *recorded measurements*. The processes of measuring and data gathering are basic to all that we do in business statistics. It is data that are analyzed by a business statistician in order to learn more about the variables being studied. Sometimes, sets of data are organized into databases as a way to store data or as a means for more conveniently analyzing data or comparing variables. Valid data are the lifeblood of business statistics, and it is important that the business researcher pay thoughtful attention to the creation of meaningful, valid data before embarking on analysis and reaching conclusions.

1.4 DATA MEASUREMENT

Learning Objective 4

Compare the four different levels of data: nominal, ordinal, interval, and ratio.

Millions of numerical data are gathered in businesses every day, representing myriad items. For example, numbers represent dollar costs of items produced, geographical locations of retail outlets, masses of shipments, and rankings of employees at yearly reviews. Not all such data should be analyzed in the same way statistically because the entities represented by the numbers are different. For this reason, the business researcher needs to know the *level of data measurement* represented by the numbers being analyzed.

The disparate use of numbers can be illustrated by the numbers 40 and 80, which could represent the masses of two objects being shipped, the ratings received on a consumer test by two different products, or the hockey jersey numbers of a centre and a winger. Although 80 kg is twice as much as 40 kg, the winger is probably not twice as big as the centre! Averaging the two masses seems reasonable but averaging the hockey jersey numbers makes no sense. The appropriateness of the data analysis depends on the level of measurement of the data gathered. The phenomenon represented by the numbers determines the level of data measurement. Four common levels of data measurement are:

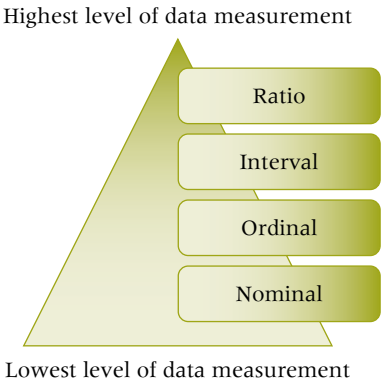
1. Nominal
2. Ordinal
3. Interval
4. Ratio

Nominal is the lowest level of data measurement, followed by ordinal, interval, and ratio. Ratio is the highest level of data, as shown in Figure 1.3.

NOMINAL LEVEL

The *lowest level of data measurement* is the nominal level. Numbers representing **nominal-level data** (the word *level* is often omitted) can be *used only to classify or categorize*. Employee

Figure 1.3
Hierarchy of Levels of Data



identification numbers are an example of nominal data. The numbers are used only to differentiate employees and not to make a value statement about them. Many demographic questions in surveys result in data that are nominal because the questions are used for classification only. The following is an example of a question that would result in nominal data:

Which of the following employment classifications best describes your area of work?

a. Educator
b. Construction worker
c. Manufacturing worker
d. Lawyer
e. Doctor
f. Other

Suppose that, for computing purposes, an educator is assigned a 1, a construction worker is assigned a 2, a manufacturing worker is assigned a 3, and so on. These numbers should be used only to classify respondents. The number 1 does not denote the top classification. It is used only to differentiate an educator (1) from a lawyer (4) or any other occupation.

Other types of variables that often produce nominal-level data are sex, religion, ethnicity, geographic location, and place of birth. Social insurance numbers, telephone numbers, and employee ID numbers are further examples of nominal data. Statistical techniques that are appropriate for analyzing nominal data are limited. However, some of the more widely used statistics, such as the chi-square statistic, can be applied to nominal data, often producing useful information.

ORDINAL LEVEL

Ordinal-level data measurement is higher than nominal level. In addition to having the nominal-level capabilities, ordinal-level measurement can be used to rank or order objects. For example, using ordinal data, a supervisor can evaluate three employees by ranking their productivity with the numbers 1 through 3. The supervisor could identify one employee as the most productive, one as the least productive, and one as somewhere in between by using ordinal data. However, the supervisor could not use ordinal data to establish that the intervals between the employees ranked 1 and 2 and between the employees ranked 2 and 3 are equal; that is, the supervisor could not say that the differences in the amount of productivity between the workers ranked 1, 2, and 3 are necessarily the same. With ordinal data, the distances between consecutive numbers are not always equal.

Some Likert-type scales on questionnaires are considered by many researchers to be ordinal in level. The following is an example of such a scale:

This computer tutorial is	____	____	____	____	____
	not helpful	somewhat helpful	moderately helpful	very helpful	extremely helpful
	1	2	3	4	5

When this survey question is coded for the computer, only the numbers 1 through 5 will remain, not the descriptions. Virtually everyone would agree that a 5 is higher than a 4 on this scale and that it is possible to rank responses. However, most respondents would not consider the differences between not helpful, somewhat helpful, moderately helpful, very helpful, and extremely helpful to be equal.

Mutual funds are sometimes rated in terms of investment risk by using measures of default risk, currency risk, and interest rate risk. These three measures are applied to investments by rating them as high, medium, or low risk. Suppose high risk is assigned a 3, medium risk a 2, and low risk a 1. If a fund is awarded a 3 rather than a 2, it carries more risk, and so on. However, the differences in risk between categories 1, 2, and 3 are not necessarily equal. Thus, these measurements of risk are only ordinal-level measurements. Another example of the use of ordinal numbers in business is the ranking of the 50 best employers in Canada in *Report on Business* magazine. The numbers ranking the companies are only ordinal in measurement. Certain statistical techniques are specifically suited to ordinal data, but many other techniques are not appropriate for use on ordinal data.

Because nominal and ordinal data are often derived from imprecise measurements such as demographic questions, the categorization of people or objects, or the ranking of items, *nominal and ordinal data* are **nonmetric data** and are sometimes referred to as *qualitative data.*

INTERVAL LEVEL

Interval-level data measurement is the *next to the highest level of data, in which the distances between consecutive numbers have meaning and the data are always numerical.* The distances represented by the differences between consecutive numbers are equal; that is, interval data have equal intervals. An example of interval measurement is Celsius temperature. With Celsius temperature numbers, the temperatures can be ranked, and the amounts of heat between consecutive readings, such as 20°, 21°, and 22°, are the same.

In addition, with interval-level data, the zero point is a matter of convention or convenience and not a natural or fixed zero point. Zero is just another point on the scale and does not mean the absence of the phenomenon. For example, zero degrees Celsius is not the lowest possible temperature. Some other examples of interval-level data are the percentage change in employment, the percentage return on a stock, and the dollar change in share price.

RATIO LEVEL

Ratio-level data measurement is *the highest level of data measurement.* Ratio data *have the same properties as interval data,* but ratio data have an *absolute zero* and *the ratio of two numbers is meaningful.* The notion of absolute zero means that zero is fixed, and *the zero value in the data represents the absence of the characteristic being studied.* The value of zero cannot be arbitrarily assigned because it represents a fixed point. This definition enables the statistician to create *ratios* with the data.

Examples of ratio data are height, mass, time, volume, and Kelvin temperature. With ratio data, a researcher can state that 180 kg of mass is twice as much as 90 kg or, in other words, make a ratio of 180:90. Many of the data gathered by machines in industry are ratio data.

Other examples in the business world that are ratio level in measurement are production cycle time, work measurement time, passenger distance, number of trucks sold, complaints per 10,000 flyers, and number of employees. With ratio-level data, no b factor is required in converting units from one measurement to another, that is, $y = ax$. As an example, in converting height from metres to feet, 1 m = 3.28 ft.

Because interval- and ratio-level data are usually gathered by precise instruments often used in production and engineering processes, in standardized testing, or in standardized accounting procedures, they are called **metric data** and are sometimes referred to as *quantitative* data.

COMPARISON OF THE FOUR LEVELS OF DATA

Figure 1.4 shows the relationships of the usage potential among the four levels of data measurement. The concentric squares denote that each higher level of data can be analyzed by