BUSINESS STATISTICS

For Contemporary Decision Making

Ninth Edition

KEN BLACK

WILEY

WileyPLUS with ORION

A personalized, adaptive learning experience.

WileyPLUS with ORION delivers easy-to-use analytics that help educators and students see strengths and weaknesses to give learners the best chance of succeeding in the course.
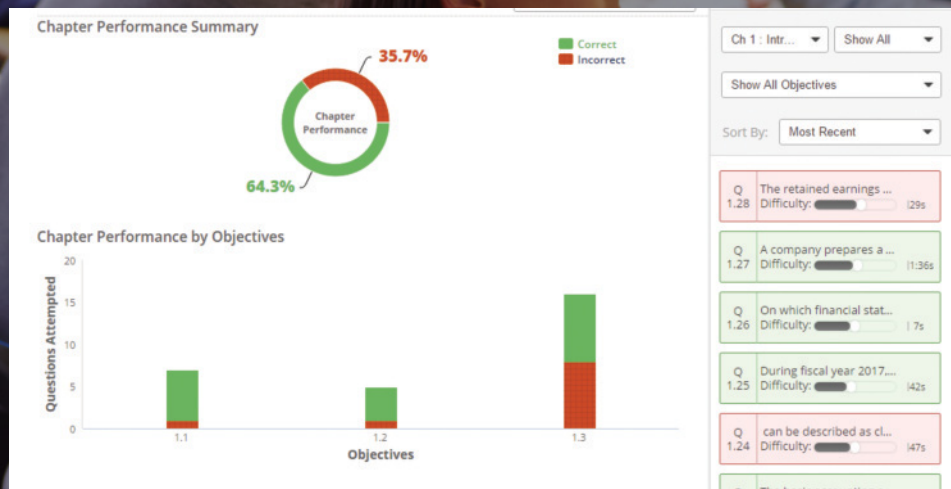

Photo credit: Monkey Business Images/Shutterstock

Identify which students are struggling early in the semester.

Educators assess the real-time engagement and performance of each student to inform teaching decisions. Students always know what they need to work on.

Help students organize their learning and get the practice they need.

With ORION's adaptive practice, students quickly understand what they know and don't know. They can then decide to study or practice based on their proficiency.

Measure outcomes to promote continuous improvement.

With visual reports, it's easy for both students and educators to gauge problem areas and act on what's most important.

www.ORION.wileyplus.com

WILEY

Business Statistics

Ninth Edition

Dedication

For Carolyn, Caycee, and Wendi

Business Statistics

For Contemporary Decision Making

Ninth Edition

KEN BLACK

University of Houston—Clear Lake

WILEY

Founded in 1807, John Wiley & Sons, Inc. has been a valued source of knowledge and understanding for more than 200 years, helping people around the world meet their needs and fulfill their aspirations. Our company is built on a foundation of principles that include responsibility to the communities we serve and where we live and work. In 2008, we launched a Corporate Citizenship Initiative, a global effort to address the environmental, social, economic, and ethical challenges we face in our business. Among the issues we are addressing are carbon impact, paper specifications and procurement, ethical conduct within our business and among our vendors, and community and charitable support. For more information, please visit our website: www.wiley.com/go/citizenship.

The inside back cover will contain printing identification and country of origin if omitted from this page. In addition, if the ISBN on the back cover differs from the ISBN on this page, the one on the back cover is correct.

The ninth edition of *Business Statistics for Contemporary Decision Making* continues the tradition of presenting and explaining the wonders of business statistics through the use of clear, complete, student-friendly pedagogy. The author and Wiley have vast ancillary resources available through WileyPLUS with which to complement the text in helping instructors effectively deliver this subject matter and in assisting students in their learning. With WileyPLUS instructors have far greater latitude in developing and delivering their course than ever before.

In the ninth edition, all the features of the eighth edition have been updated and changed as needed to reflect today's business world. Two new Decision Dilemmas located in Chapters 6 and 7 have been added to the ninth edition. Chapter 6, which introduces the student to continuous distributions including the normal curve, opens with a new Decision Dilemma titled "CSX Corporation," a leading rail transportation company in the U.S. Chapter 7 presents a new Decision Dilemma, "Toro" which discusses the Toro Company's 100 years of providing business solutions for the outdoor environment.

The ninth edition introduces five new cases. Chapter 2 contains the Southwest Airlines case. Chapter 6 presents a case on USAA, the United Services Automobile Association, which offers banking, investment, and insurance to people and families that serve, or served, in the United States military. Chapter 7 contains a new case on the 3M Company, best known for its adhesive products. Chapter 10 has a new case, Five Guys, a family hamburger restaurant chain with U.S. sales of $1.21 billion. Chapter 12 contains a new case on Caterpillar, Inc., the world's leading manufacturer of construction and mining equipment, diesel and natural gas engines, industrial gas turbines and diesel-electric locomotives.

All other cases in the ninth edition have been updated and edited for today's market. Three cases have been significantly revised. These include: "Statistics Describe the State of Business in India's Countryside" in Chapter 1, "Coca Cola Develops the African Market" in Chapter 3, and "Virginia Semiconductor" in Chapter 14.

Every chapter in the ninth edition contains one or two *Thinking Critically About Statistics in Business Today* features that give real-life examples of how the statistics presented in the chapter apply in the business world today. Each of these contains thought-provoking questions called *Things to Ponder* in which the objective is to effect critical thinking on behalf of the student and generate discussion involving critical thinking in the classroom. Some of these include: "Plastic Bags vs. Bringing Your Own in Japan," "Where Are Soft Drinks Sold?," "Recycling Statistics," "Probabilities in the Dry Cleaning Business," "Warehousing," and "Teleworking Facts."

This edition is designed and written for a two-semester introductory undergraduate business statistics course or an MBA-level introductory course. In addition, with 19 chapters, the ninth edition 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 college 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 rigor, 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 Ninth Edition

Chapters

In this edition, several changes have been made in an effort to improve the delivery and completeness of the text. The chapter organization is unchanged; however, the unit organization has been dropped. The text chapter organization allows for both one- and two-semester coverage. The first twelve chapters include the typical chapters covered in a one-semester course. The last seven chapters offer extended coverage allowing students to explore such topics as multiple regression, forecasting, analysis of categorical data, nonparametric statistics, statistical quality control, and decision analysis. For clarity and topic flow, some chapter topics have been combined and reduced in number. In Chapter 1, Statistics in Business has been included in the chapter introduction; and Variables and Data have been combined with Basic Statistical Concepts thereby reducing the number of chapter topics from four to two. In Chapter 4, Methods of Assigning Probabilities has been subsumed with Introduction to Probability.

Topical Changes

Sections and topics from the eighth edition remain virtually unchanged in the ninth edition with the exception that the former end of chapter feature, *Using the Computer*, has been removed from the text. Directions on how to use Excel and/or Minitab to compute statistics featured in the text are now housed exclusively in WileyPLUS.

In addition, in Chapter 14, the section on Logistic Regression has been significantly rewritten to reflect the changes in Minitab moving into version 17. Since much of the material on Logistic Regression is related to computer output, it was necessary to realign section 14.5 with current Minitab output for Logistic Regression. However, the conceptual explanations and the examples remain the same.

Decision Dilemma and the Decision Dilemma Solved

Each chapter of the ninth edition begins with a Decision Dilemma. Decision Dilemmas are real business vignettes that set the tone for each 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. At the end of each chapter, a 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 ninth 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, two new Decision Dilemmas located in Chapters 6 and 7 have been added to the ninth edition. Chapter 6, which introduces the student to continuous distributions including the normal curve, opens with a new Decision Dilemma titled "CSX Corporation" This Decision Dilemma introduces the student to the CSX Corporation which is a leading rail transportation company in the U.S. Based in Jacksonville, Florida, the CSX Transportation network has 21,000 miles of track across the eastern half of the U.S. and two Canadian provinces. Analysis of this vignette focuses on average rail freight line-haul speed, terminal dwell time, and freight train arrival time. Students have to analyze data about these variables using probability distributions presented in Chapter 6.

Chapter 7, Sampling and Sampling Distributions, contains another new Decision Dilemma, "Toro", which discusses the Toro Company's 100 years of providing business solutions for the outdoor environment including turf, snow, and ground-engaging equipment along with irrigation and outdoor lighting solutions. The associated managerial and statistical questions tackle issues related to consumer and household spending on landscape services both in terms of dollars spent and population percentages.

Thinking Critically About Statistics in Business Today

Every chapter in the ninth edition contains at least one or two *Thinking Critically About Statistics in Business Today* features that give real-life examples of how the statistics presented in the chapter apply in the business world today. Each of these contains thought-provoking questions called Things to Ponder in which the objective is to effect critical thinking on behalf of the student and generate discussion involving critical thinking in the classroom. This approach to learning is in sync with various accreditation organizations and their current emphasis on developing critical thinking in our students. Some of these include: "Recycling Statistics," "Warehousing," "Canadian Grocery Shopping Statistics," "Beverage Consumption: America vs. Europe," "Are Facial Characteristics Correlated with CEO Traits?" "Assessing Property Values Using Multiple Regression," "Can Scrap Metal Prices Forecast the Economy?" "City Images of Cruise Destinations in the Taiwan Strait," and "Does an Iranian Auto Parts Manufacturer's Orientation Impact Innovation?" As an example, from "Beverage Consumption: America vs. Europe," Americans drink nearly five times as much soda as do Europeans and almost twice as much beer. On the other hand, Europeans drink more than twice as much tea (hot or cold), more than three times as much wine, and over four times as much tap water as Americans. Statistics show that the average American consumes forty-eight 12 oz. containers of carbonated soda per month compared to only 10 for Europeans. Europeans consume an average of sixteen 4 oz. containers of wine per month compared to an average of only five for Americans. One of the Things to Ponder questions is, "Can you think of some reasons why Americans consume more carbonated soda pop and beer than Europeans, but less wine, hot or iced tea, or tap water? Do you think that these outcomes may change in time?"

Cases

Every chapter in this text contains a unique business case. 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. The ninth edition introduces five new, exciting cases. The Southwest Airlines case in chapter 2 discusses the company's rise from incorporation in 1967 to an operation with more than 3,600 flights a day, serving 94 destinations across the United States and six other countries. In Chapter 6, the reader is introduced to USAA, the United Services Automobile Association, which was founded in 1922 by 25 army officers who came together in San Antonio and decided to insure each other's automobiles. Offering banking, investment, and insurance to people and families that serve, or served, in the United States military, USAA currently has 10.7 million members, 42.6 million products, 26,300 employees and is ranked 57th in net worth ($25 billion) in Fortune 500 companies. Chapter 7 contains a new case on the 3M Company. Born as a small-scale mining company in 1902 and best known for its adhesive products, 3M is a global innovation company with over 100,000 patents, $31 billion in sales, and 90,000 employees. Five Guys, a hamburger restaurant chain whose philosophy is to serve fresh, hand-cooked burgers with "all the toppings you could stuff between fresh-baked buns" along with fresh-cut fries cooked in peanut oil, is the new case in Chapter 10. A family operation which opened its first restaurant in Arlington, Virginia in 1987 on $70,000

has grown to 1,163 U.S. restaurants with a U.S. system wide sales of $1.21 billion and units now existing in other countries. Chapter 12 contains a new case on Caterpillar, Inc. Caterpillar, headquartered in Peoria, Illinois, is an American corporation with a worldwide dealer network which sells machinery, engines, financial products and insurance. Caterpillar is the world's leading manufacturer of construction and mining equipment, diesel and natural gas engines, industrial gas turbines and diesel-electric locomotives.

All other cases in the ninth edition have been updated and edited for today's market. Three cases have been significantly revised. These include: "Statistics Describe the State of Business in India's Countryside" in Chapter 1, "Coca Cola Develops the African Market" in Chapter 3, and "Virginia Semiconductor" in Chapter 14. Here are some excerpts from one case, "The Container Store":

"In the late 1970s, Kip Tindell (chairman and CEO), Garrett Boone (Chairman Emeritus), and John Mullen (architect) drew up plans for a first-of-its-kind retail store specializing in storage solutions for both the home and the office. The vision that they created was realized when on July 1, 1978, the Container Store opened its doors in a small 1,600 square foot retail space in Dallas. The store was stocked with products that were devoted to simplifying people's lives, such as commercial parts bins, wire drawers, mailboxes, milk crates, wire leaf burners, and many others. Some critics even questioned that a store selling "empty boxes" could survive. However, the concept took off, and in the past 33 years, the company has expanded coast to coast in the United States with stores in 49 locations. Now headquartered in Coppell, Texas, the Container Store has 4,000 employees and annual revenues of over $650 million. Besides their innovative product mix, one of the keys to the success of the Container Store is the enthusiasm with which their employees work, the care that employees give to the customer, and employee knowledge of their products."

New Problems

There are 965 practice problems in the text. For ease of effort among adopters in moving from the eighth to the ninth edition, the number of problems and the problem numbers have remained the same. However, 105 of these problems have been updated to include the latest data and information available. Ten of the problems have been replaced with new problems.

Included in virtually every section of every chapter of the text is a demonstration problem which 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 the ninth edition are business oriented and contain the most current data available.

As with the previous edition, 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 ninth edition are nine databases that provide additional opportunities for students to apply the statistics presented in this text. These nine databases represent a wide variety of business areas, such as agribusiness, consumer spending, energy, finance, healthcare, international labor, manufacturing, and the stock market. Altogether, these databases contain 61 variables and 7,722 observations. The data are gathered from such reliable sources as the U.S. government's Bureau of Labor, the U.S. Department of Agriculture, the American Hospital Association, the Energy Information Administration, *Moody's Handbook of Common Stocks,* and the U.S. Census Bureau. Five of the nine databases contain time-series data. The databases are 12-Year Gasoline Database, Consumer Food Database, Manufacturing Database, International Labor Database, Financial Database, Energy Database, U.S. and International Stock Market Database, Hospital Database, and Agribusiness Time-Series Database.

Video Tutorials by Ken Black

An exciting feature of the ninth edition package that will impact the effectiveness of student learning in business statistics and significantly enhance the presentation of course material is the series of videotape tutorials by Ken Black. With the advent of online 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 they need to bridge the gap between theory and application on their own. There are now 21 videotaped tutorial sessions on key difficult topics in business statistics delivered by Ken Black and available for all adopters on WileyPLUS. In addition, these tutorials can easily be uploaded for classroom usage to augment lectures and enrich classroom presentations. Because there is at least one video for each of the first 12 chapters, the instructor has the option to include at least one video in the template of each chapter's plan for most, if not all, of the course. While the video tutorials vary in length, a typical video is about 10 minutes in length. The 21 video 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 Tables
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

Features and Benefits

Each chapter of the ninth edition contains sections called Learning Objectives, a Decision Dilemma, Demonstration Problems, Section Problems, Thinking Critically About Statistics in Business Today, Decision Dilemma Solved, Chapter Summary, Key Terms, Formulas, Ethical Considerations, Supplementary Problems, Analyzing the Databases, and Case.

- **Learning Objectives.** Each chapter begins with a statement of 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 ninth 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.** There are 965 problems in the text. Problems for practice are found at the end of almost every section of the text. Most problems utilize real data gathered from a plethora of sources. Included here are a few brief excerpts from some of the real-life problems in the text: "*The Wall Street Journal* reported that 40% of all workers say they would change jobs for 'slightly higher pay.' In addition, 88% of companies say that there is a shortage of qualified job candidates." "In a study by Peter D. Hart Research Associates for the Nasdaq Stock Market, it was determined that 20% of all stock investors are retired people. In addition, 40% of all U.S. adults have invested in mutual funds." "A survey conducted for the Northwestern National Life Insurance Company revealed that 70% of American workers say job stress caused frequent health problems." "According to Padgett Business Services, 20% of all small-business owners say the most important advice for starting a business is to prepare for long hours and hard work. "According to Nielsen Media Research, approximately 86% of all U.S. households have High-definition television (HDTV). In addition, 49% of all U.S. households own Digital Video Recorders (DVR)."

- **Thinking Critically About Statistics in Business Today.** Every chapter in the ninth edition contains at least one Thinking Critically About Statistics in Business Today feature. These focus boxes contain an interesting application of how techniques of that particular chapter are used in the business world today and ask probing questions of the student. 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.

- **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 are bolded and their definitions 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 a reader to locate them. At the end of the chapter, most of the chapter's formulas are listed together as a handy reference.

- **Ethical Considerations.** Each chapter contains an Ethical Considerations feature that is very timely, given the serious breach of ethics and lack of moral leadership of some business executives in recent years. With the abundance of statistical data and analysis, there is considerable potential for the misuse of statistics in business dealings. The important Ethical 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 a few excerpts from Ethical Considerations features: "It is unprofessional and unethical to draw cause-and-effect conclusions just because two variables are correlated." "The

business researcher needs to conduct the experiment in an environment such that as many concomitant variables are controlled as possible. To the extent that this is not done, the researcher has an ethical responsibility to report that fact in the findings." "The reader is warned that the value lambda is assumed to be constant in a Poisson distribution experiment. Business researchers may produce spurious results if the value of lambda is used throughout a study, but because the study is conducted during different time periods, the value of lambda is actually changing." "In describing a body of data to an audience, it is best to use whatever statistical 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 skew the way the receiver understands the data."

- **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 nine major databases located on the student companion Web site that accompanies the ninth edition and in WileyPLUS. 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 chapter has an end-of-chapter case based on a real company. 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 Shell Oil, Coca-Cola, or Colgate Palmolive. Others pertain to smaller businesses—such as Virginia Semiconductor, The Clarkson Company, or DeBourgh—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.

WileyPLUS

WileyPLUS with ORION 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. This interactive approach focuses on:

Design   Research-based design is based on proven instructional methods. Content is organized into small, more accessible amounts of information, helping students build better time management skills.

Engagement   Students can visually track their progress as they move through the material at a pace that is right for them. Engaging in individualized self-quizzes followed by immediate feedback helps to sustain their motivation to learn.

Outcomes   Self-assessment lets students know the exact outcome of their effort at any time. Advanced reporting allows instructors to easily spot trends in the usage and performance data of their class in order to make more informed decisions.

With *WileyPLUS*, students will always know:

- What to do: Features, such as the course calendar, help students stay on track and manage their time more effectively.
- How to do it: Instant feedback and personalized learning plans are available 24/7.
- If they're doing it right: Self-evaluation tools take the guesswork out of studying and help students focus on the right materials.

WileyPLUS **for** *Business Statistics, Ninth Edition* **includes numerous valuable resources, among them:**

- **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 the author explaining concepts and demonstrating how to work problems for some of the more difficult topics.
- **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 the student 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 both Excel and Minitab format.
- **Animations.**   To aid students in understanding complex interactions, selected figures from the text that involve dynamic activity have been animated. Students can download these animated figures and run them to improve their understanding of dynamic processes.
- **Flash Cards.**   Key terms will be available to students in flash card format along with their definition.
- **Student Study Manual.**   Complete solutions to all odd-numbered questions.
- **Demo Problems.**   Step-by-step solved problems for each chapter.
- **ORION.**   This adaptive, personalized learning experience delivers easy-to-use analytics so you can see exactly where your students excel and where they need help.
  - **Diagnose Early.**   Simply by assigning ORION, you can diagnose the real-time proficiency of each student and see the areas that need reinforcement.

- **Facilitate Engagement.** With ORION's adaptive practice, students can interact with each other as they think more deeply about concepts at hand.
- **Measure Outcomes.** ORION helps you measure students' engagement and proficiency throughout the course so that you can easily assess how things are going at any point in time.

Ancillary Teaching and Learning Materials

www.wiley.com/college/black

Students' Companion Site

The student companion Web site contains:

- All databases in both Excel and Minitab formats for easy access and use.
- Excel and Minitab files of data from all text problems and all cases. Instructors and students now have the option of analyzing any of the data sets using the computer.
- A section on Advanced Exponential Smoothing Techniques (from Chapter 15) that offers the instructor an opportunity

to delve deeper into exponential smoothing if so desired, and derivation of the slope and intercept formulas from Chapter 12.

- A tutorial on summation theory.

Instructor's Resource Kit

All instructor ancillaries are provided on the Instructor Resource Site. Included in this convenient format are:

- **Instructor's Manual.** Prepared by Ken Black, this manual contains the worked-out solutions to virtually all problems in the text. In addition, this manual contains chapter objectives, chapter outlines, chapter teaching strategies, and solutions to the cases.
- **PowerPoint Presentation Slides.** The presentation slides, contain graphics to help instructors create stimulating lectures. The PowerPoint slides may be adapted using PowerPoint software to facilitate classroom use.
- **Test Bank.** The Test Bank includes multiple-choice questions for each chapter. The Test Bank is provided in Microsoft Word format as well as in a Computerized Test Bank.

Acknowledgments

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 (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.* Black has consulted for many different companies, including Aetna, the city of Houston, NYLCare, AT&T, Johnson Space Center, Southwest Information Resources, UTMB, and Doctors Hospital at Renaissance. Black is active in the quality movement and is a certified Master Black Belt in Lean Six Sigma.

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

Brief Contents

CONTENTS

19 Decision Analysis   **705**

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. List quantitative and graphical examples of statistics within a business context.

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.

Decision Dilemma

Statistics Describe the State of Business in India's Countryside

India is the second largest country in the world with more than 1.25 billion people. More than 70% of the people live in rural areas scattered about the countryside in 650,000 villages. In fact, it can be said that one in every 10 people in the world live in rural India. While it has a per capita income of less than $1 (U.S.) per day, rural India, which has been described in the past as poor and semi-illiterate, now contributes to about one-half of the country's gross national product (GNP). However, rural India still has the most households in the world without electricity, over 300,000.

Despite its poverty and economic disadvantages, there are 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 India 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 about their lifestyles and opportunities for a better life.

Around 60% of all middle-income households in India are in rural areas and more than one-third of all rural households in India now have a main source of income other than farming. Virtually every home has a radio, about one-third have a television, and more than one-half of rural households benefit from banking services. Forty-two percent of the people living in India's villages and small towns use toothpastes and that proportion is increasing as rural income rises and as there is greater awareness about oral hygiene.

© Hemis/Alamy Stock Photo

In rural India, consumers are gaining more disposable income due to the movement of manufacturing jobs to rural areas. It is estimated that nearly 75% of the factories that opened in India in the past decade were built in rural areas. Products that are doing well in sales to people in rural India include televisions, fans, bicycles, bath soap, two- or three-wheelers, cars, and many others. According

to MART, a New Delhi-based research organization, rural India buys 46% of all soft drinks and 49% of motorcycles sold in India. Because of such factors, many U.S. and Indian firms such as Microsoft, General Electric, Kellogg's, Colgate-Palmolive, Idea Cellular, Hindustan Lever, Godrej, Nirma Chemical Works, Novartis, Dabur, Tata Motors, and Vodafone India have entered the rural Indian market with enthusiasm. Marketing to rural customers often involves persuading them to try and to adopt 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.

MANAGERIAL AND STATISTICAL QUESTIONS

1. Are the statistics presented in this report exact figures or estimates?
2. How and where could researchers have gathered such data?
3. In measuring the potential of the rural India 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?

Source: Adapted from "Marketing to Rural India: Making the Ends Meet," March 8, 2007, in *India Knowledge@Wharton*, http://knowledge. wharton.upenn.edu/india/article.cfm?articleid=4172; "Rural Segment Quickly Catching Up", September 2015, IBEF (India Brand Equity Foundation) at: www.ibef.org/industry/indian-rural-market.aspx; "Unlocking the Wealth in Rural Markets", June 2014, *Harvard Business Review* at: https://hbr.org/2014/06/unlocking-the-wealth-in-rural-markets; "Much of Rural India Still Waits for Electricity", October 2013, University of Washington, at: artsci.washington.edu/news/2013-10/much-rural-india-still-waits-electricity

Every minute of the working day, decisions are made by businesses around the world that determine whether companies will be profitable and growing 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 tool through 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.

Virtually every area of business uses statistics in decision making. Here are some recent examples:

- According to a national survey of independent business owners conducted by the Institute for Local Self-Reliance in partnership with the Advocates for Independent Business coalition, when asked "Which two public policy changes would most help your business?" (retailers only), 40% said "Pass the Marketplace Fairness Act" and 38% said "Cap Credit Card Swipe Fees".
- A survey of 1465 workers by Hotjobs reports 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.
- A survey of 477 executives by the Association of Executive Search Consultants determined that 48% of men and 67% of women say they are more likely to negotiate for less business travel compared with five years ago.
- A global Family Business Survey of 2,378 respondents sponsored by PwC reported that 65% of family businesses reported growth in the last twelve months and 49% of respondents are apprehensive about their ability to recruit skilled staff in the next twelve months.
- A Deloitte Retail "Green" survey of 1080 adults revealed that 54% agreed that plastic, non-compostable shopping bags should be banned.
- A study of consumer electronics spending by a 2,500 member on-line panel of the NPD group showed that consumers expect to spend $555, on average, per person on new consumer electronics devices this year.

You can see from these few examples that there is a wide variety of uses and applications of statistics in business. Note that in most of these examples, business researchers have conducted a study and provided us rich and interesting information.

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 more inexpensively and in a more timely manner. We will explore various ways to forecast future values and 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.

1.1   Basic Statistical Concepts

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*. The word *statistics* 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** graphically displays the key elements of statistics.

| Collect Data | Analyze Data | Interpret Data | Present Findings |
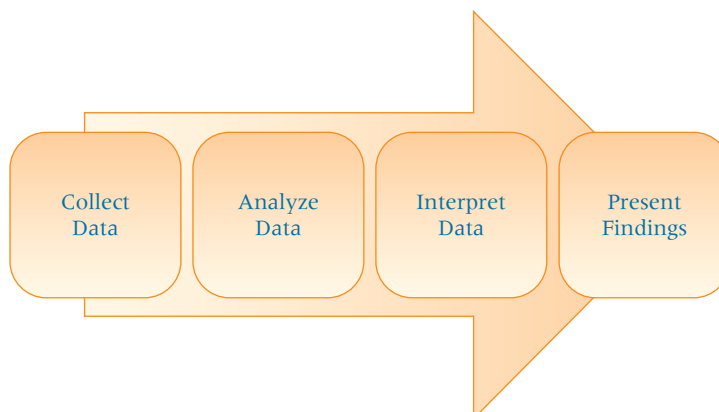
FIGURE 1.1   **The Key Elements of Statistics**

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 Mustang cars produced from 2014 to 2016." A population can be a group of people, such as "all workers presently employed by Microsoft," or it can be a set of objects, such as "all dishwashers produced on February 3, 2016, by the General Electric Company at the Louisville 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 U.S. Census. Every 10 years, the government attempts to measure all persons living in this country.

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 lightbulbs, a lightbulb manufacturer might randomly sample only 75 lightbulbs 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 effort and uses those statistics to reach conclusions about that class only, the statistics are descriptive.

Many of the statistical data generated by businesses are descriptive. They might include number of employees on vacation during June, average salary at the Denver office, corporate sales for 2016, average managerial satisfaction score on a company-wide census of employee attitudes, and average return on investment for the Lofton Company for the years 1996 through 2016.

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 study effectively 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 ($\overline{x}$), 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 researcher 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.

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).

Business statistics is about measuring phenomena in the business world and organizing, analyzing, and presenting the resulting numerical information in such a way such 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, labor productivity, stock price, historical cost, total sales, market share, age of worker, earnings per share, miles 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** is *when a standard process is used to assign numbers to particular attributes or characteristics of a variable*. Many measurements are obvious, such as time spent in a store shopping by a customer, age of the worker, or the number of miles driven to work. However, some measurements, such as labor 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
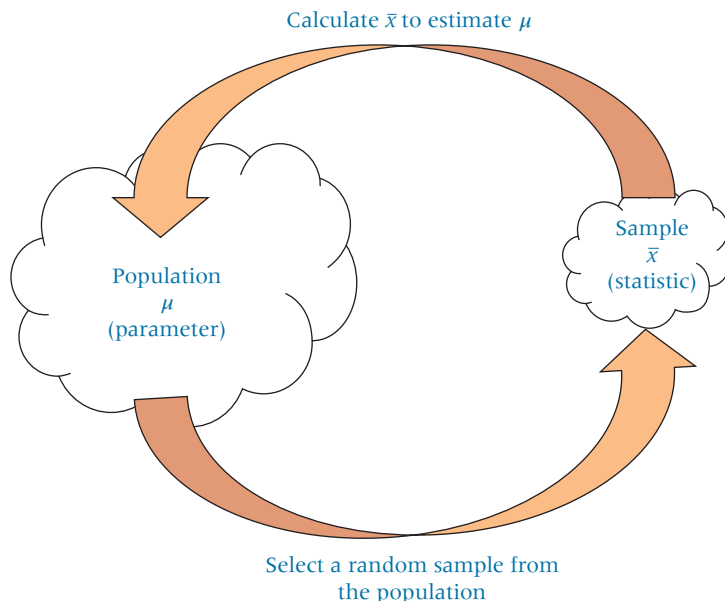
FIGURE 1.2   **The Inferential Process**

are the lifeblood of business statistics, and it is important that the business researcher give thoughtful attention to the creation of meaningful, valid data before embarking on analysis and reaching conclusions.

1.2 Data Measurement

Lecture Video

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, weights of shipments, and rankings of subordinates at yearly reviews. All such data should not be analyzed 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 weights of two objects being shipped, the ratings received on a consumer test by two different products, or football jersey numbers of a fullback and a wide receiver. Although 80 pounds is twice as much as 40 pounds, the wide receiver is probably not twice as big as the fullback! Averaging the two weights seems reasonable, but averaging the football 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 follow.

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 measurement, 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* often is omitted) can be *used only to classify or categorize*. Employee 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 such a question that would result in nominal data:

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

Highest Level of Data Measurement

1. Educator
2. Construction worker
3. Manufacturing worker
4. Lawyer
5. Doctor
6. 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).

Lowest Level of Data Measurement

FIGURE 1.3 **Hierarchy of Levels of Data**

Some other types of variables that often produce nominal-level data are sex, religion, ethnicity, geographic location, and place of birth. Social Security numbers, telephone

numbers, employee ID numbers, and ZIP code 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 the nominal level. In addition to the nominal-level capabilities, ordinal-level measurement can be used to rank or order people or 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 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, she could not say that the differences in the amount of productivity between workers ranked 1, 2, and 3 are necessarily the same. With ordinal data, the distances or spacing represented by consecutive numbers are not always equal.

Some questionnaire Likert-type scales are considered by many researchers to be ordinal in level. The following is an example of one such 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 adjectives. Virtually everyone would agree that a 5 is higher than a 4 on this scale and that ranking responses is possible. 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 as investments are sometimes rated in terms of risk by using measures of default risk, currency risk, and interest rate risk. These three measures are applied to investments by rating them as having high, medium, and 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 top 50 most admired companies in *Fortune* 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. For example, it does not make sense to say that the average of "moderately helpful" and "very helpful" is "moderately helpful and a half."

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 Fahrenheit temperature. With Fahrenheit 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 Fahrenheit 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 stock 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, weight, time, volume, and Kelvin temperature. With ratio data, a researcher can state that 180 pounds of weight is twice as much as 90 pounds or, in other words, make a ratio of $180:90$. Many of the data measured by valves or gauges 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 miles, number of trucks sold, complaints per 10,000 fliers, and number of employees.

Because interval- and ratio-level data are usually gathered by precise instruments often used in production and engineering processes, in national 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 any of the techniques used on lower levels of data but, in addition, can be used in other statistical techniques. Therefore, ratio data can be analyzed by any statistical technique applicable to the other three levels of data plus some others.

Nominal data are the most limited data in terms of the types of statistical analysis that can be used with them. Ordinal data allow the researcher to perform any analysis that can be done with nominal data and some additional analyses. With ratio data, a statistician can make ratio comparisons and appropriately do any analysis that can be performed on nominal, ordinal, or interval data. Some statistical techniques require ratio data and cannot be used to analyze other levels of data.

Statistical techniques can be separated into two categories: parametric statistics and nonparametric statistics. **Parametric statistics** require that data be interval or ratio. If the data are nominal or ordinal, **nonparametric statistics** must be used. Nonparametric statistics can also be used to analyze interval or ratio data. This text focuses largely on parametric statistics, with the exception of Chapter 16 and Chapter 17, which contain nonparametric techniques. Thus much of the material in this text requires that data be interval or ratio data.

Figure 1.5 contains a summary of metric data and nonmetric data.
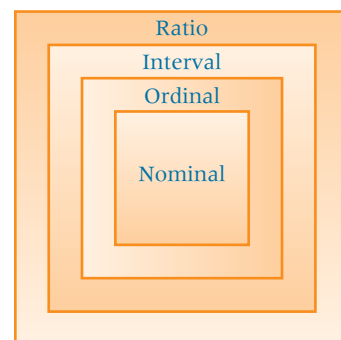
FIGURE 1.4   **Usage Potential of Various Levels of Data**

Ratio
Interval
Ordinal
Nominal

FIGURE 1.5   **Metric vs. Nonmetric Data**

Metric Data
• Higher Level Data
• Interval and Ratio
• Quantitative Data
• Can Use Parametric Statistics

Nonmetric Data
• Lower Level Data
• Nominal and Ordinal
• Qualitative Data
• Must Use Nonparametric Statistics

DEMONSTRATION PROBLEM 1.1

Because of increased competition for patients among providers and the need to determine how providers can better serve their clientele, hospital administrators sometimes administer a quality satisfaction survey to their patients after the patient is released. The following types of questions are sometimes asked on such a survey. These questions will result in what level of data measurement?

1. How long ago were you released from the hospital?

2. Which type of unit were you in for most of your stay?

__Coronary care

__Intensive care

__Maternity care

__Medical unit

__Pediatric/children's unit
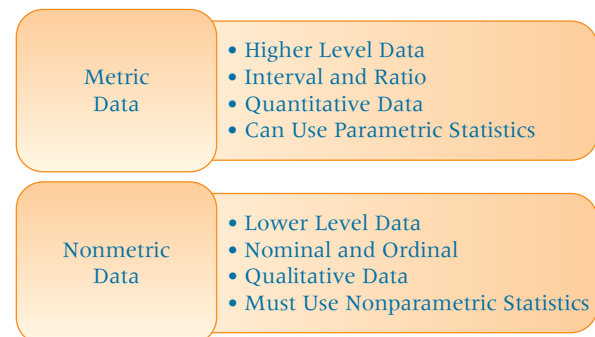
__Surgical unit

3. In choosing a hospital, how important was the hospital's location?

<div align="center">

(circle one)

| Very
Important | Somewhat
Important | Not Very
Important | Not at All
Important |

</div>

4. What was your body temperature when you were admitted to the hospital?

5. Rate the skill of your doctor:

__Excellent   __Very Good   __Good   __Fair   __Poor

Solution: Question 1 is a time measurement with an absolute zero and is therefore ratio-level measurement. A person who has been out of the hospital for two weeks has been out twice as long as someone who has been out of the hospital for one week.

Question 2 yields nominal data because the patient is asked only to categorize the type of unit he or she was in. This question does not require a hierarchy or ranking of the type of unit. Questions 3 and 5 are likely to result in ordinal-level data. Suppose a number is assigned the descriptors in these two questions. For question 3, "very important" might be assigned a 4, "somewhat important" a 3, "not very important" a 2, and "not at all important" a 1. Certainly, the higher the number, the more important is the hospital's location. Thus, these responses can be ranked by selection. However, the increases in importance from 1 to 2 to 3 to 4 are not necessarily equal. This same logic applies to the numeric values assigned in question 5. In question 4, body temperature, if measured on a Fahrenheit or Celsius scale, is interval in measurement.

Statistical Analysis Using the Computer: Excel and Minitab

The advent of the modern computer opened many new opportunities for statistical analysis. The computer allows for storage, retrieval, and transfer of large data sets. Furthermore, computer software has been developed to analyze data by means of sophisticated statistical techniques. Some widely used statistical techniques, such as multiple regression, are so tedious and cumbersome to compute manually that they were of little practical use to researchers before computers were developed.

Business statisticians use many popular statistical software packages, including Minitab, SAS, and SPSS. Many computer spreadsheet software packages also have the capability of analyzing data statistically. In this text, the computer statistical output presented is from both the Minitab and the Microsoft Excel software.