Business Statistics

EDITION 2

ROBERT A. DONNELLY, JR.

BREAKTHROUGH

To improving results

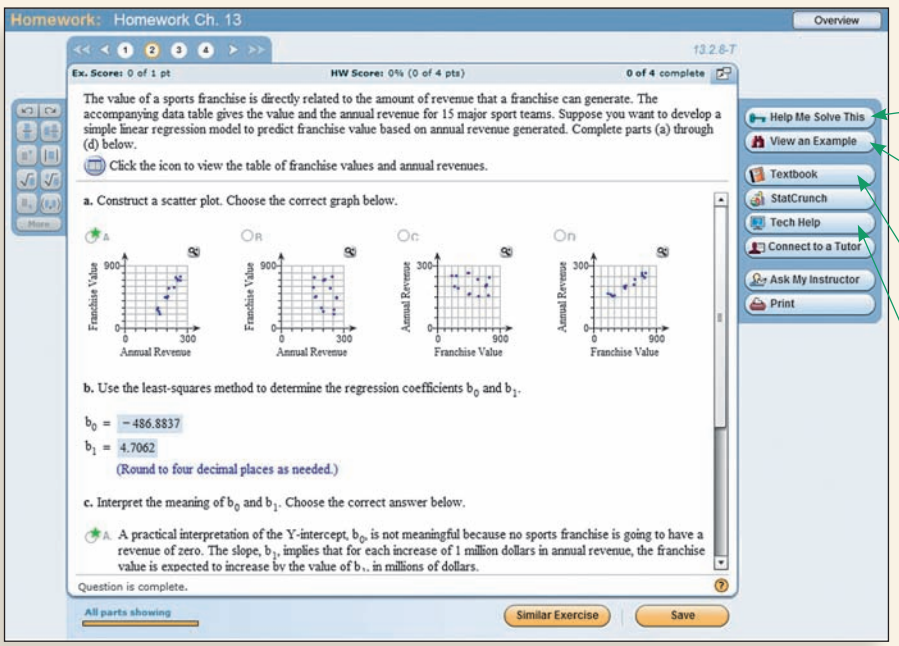
MyStatLab™

for Business Statistics

MyStatLab is a course management system that provides engaging learning experiences and delivers proven results while helping students succeed. Tools are embedded which make it easy to integrate statistical software into the course. And, MyStatLab comes from an experienced partner with educational expertise and an eye on the future.

Tutorial Exercises

MyStatLab homework and practice exercises correlated to the exercises in the textbook are generated algorithmically, giving students unlimited opportunity for practice and mastery. MyStatLab grades homework and provides feedback and guidance.



Help Me Solve This breaks the problem into manageable steps. Students enter answers along the way.

View an Example walks students through a problem similar to the one assigned.

Textbook links to the appropriate section in the etext.

Tech Help is a suite of Technology Tutorial videos that show how to perform statistical calculations using popular software.

Powerful Homework and Test Manager

Create, import, and manage online homework assignments, quizzes, and tests that are automatically graded, allowing you to spend less time grading and more time teaching. Thousands of high-quality and algorithmic exercises of all types and difficulty levels are available to meet the needs of students with diverse mathematical backgrounds.

Ready-to-Go Courses

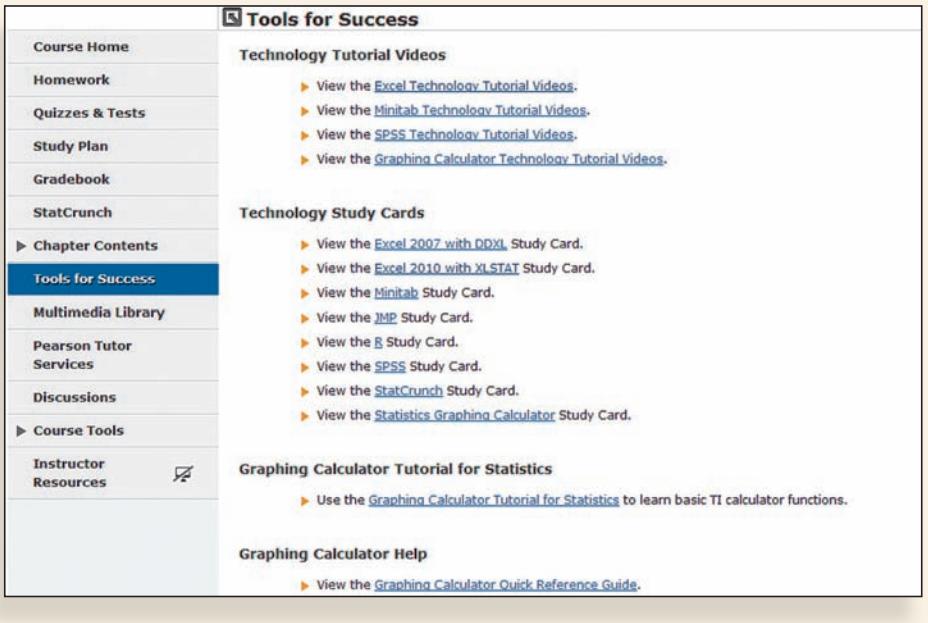
Ready-to-Go Courses make it even easier for first-time users to start using MyStatLab. With the help of experienced instructors, these courses include pre-made assignments that you can alter at any time.

Adaptive Learning

An Adaptive Study Plan serves as a personalized tutor for your students. When enabled, Knewton in MyStatLab monitors student performance and provides personalized recommendations. It gathers information about learning preferences and is continuously adaptive, guiding students though the Study Plan one objective at a time.

Integrated Statistical Software

Copy our data sets, from the eText and the MyStatLab questions, into software such as StatCrunch, Minitab, Excel, and more. Students have access to support tools—videos, Study Cards, and manuals for select titles—to learn how to use statistical software.



StatCrunch

MyStatLab includes web-based statistical software, StatCrunch, within the online assessment platform so that students can analyze data sets from exercises and the text. In addition, MyStatLab includes access to **www.StatCrunch.com,** the full web-based program where users can access thousands of shared data sets, create and conduct online surveys, perform complex analyses using the powerful statistical software, and generate compelling reports.

Engaging Video Resources

- Business Insight Videos are 10 engaging videos showing managers at top companies using statistics in their everyday work. Assignable questions encourage discussion.

- StatTalk Videos, hosted by fun-loving statistician Andrew Vickers, demonstrate important statistical concepts through interesting stories and real-life events. This series of 24 videos includes available assessment questions and an instructor's guide.

2ND EDITION

Business Statistics

ROBERT A. DONNELLY, JR.

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Editor in Chief: Deirdre Lynch
Acquisitions Editor: Marianne Stepanian
Development Editor: Dana Bettez
Assistant Editor: Sonia Ashraf
Senior Managing Editor: Karen Wernholm
Associate Managing Editor: Tamela Ambush
Senior Production Project Manager: Peggy McMahon
Associate Director of Design, USHE EMSS/HSC/EDU: Andrea Nix
Art Director/Cover Designer: Beth Paquin
Interior Design: Nancy Goulet, Studio Wink
Marketing Manager: Erin Lane
Marketing Assistant: Kathleen DeChavez
Senior Author Support/Technology Specialist: Joe Vetere
Project Supervisor, MyStatLab: Robert Carroll
QA Manager, Assessment Content: Marty Wright
Media Producer: Jean Choe
Image Manager: Rachel Youdelman
Image Research: Integra
Procurement Manager: Vincent Scelta
Procurement Specialist: Debbie Rossi
Production Coordination, Technical Illustrations, and Composition: PreMediaGlobal
Printer/Binder: Courier/Kendallville
Cover Printer: Lehigh-Phoenix Color/Hagerstown

Credits and acknowledgments borrowed from other sources and reproduced, with permission, in this textbook appear on page 935 in the back of this book.

Library of Congress Cataloging-in-Publication Data
Donnelly, Robert A.
Business statistics / Robert A. Donnelly, Jr.—2nd ed.
p. cm.
ISBN 978-0-321-92512-1
1. Commercial statistics. 2. Industrial management—Statistical methods. I. Title.
HF1017.S74 2013
519.5024'65—dc23

2012005942

10 9 8 7 6 5 4 3 2 1—CRK—16 15 14 13

PEARSON

ISBN 10: 0-321-92512-2
ISBN 13: 978-0-321-92512-1

To my wife, Debbie,
who supported and
encouraged me every
step of the way.
I could not have done this
without you, Babe.

ABOUT THE Author

ROBERT A. DONNELLY, JR.

Robert (Bob) A. Donnelly, Jr. is a professor at Goldey-Beacom College in Wilmington, Delaware, with more than 25 years of teaching experience. He teaches classes in statistics, operations management, spreadsheet modeling and project management at both the undergraduate and graduate level. Bob earned an undergraduate degree in chemical engineering from the University of Delaware, after which he worked for several years as an engineer with the Diamond Shamrock Corporation in a chlorine plant. Despite success in this field, Bob felt drawn to pursue a career in education. It was his desire to teach that took him back to school to earn his MBA and Ph.D. in Operations Research, also from the University of Delaware. Bob also teaches in the MBA program at the International School of Management in Paris, France. He thoroughly enjoys discussing research methods and business statistics with both his French and American students.

Bob's working experience gathered prior to his teaching career has provided him with many opportunities to incorporate real-life examples into classroom learning. His students appreciate his knowledge of the business world as well as his mastery of the course subject matter. Many former students seek Bob's assistance in work-related issues that deal with his expertise. Typical student comments focus on his genuine concern for their welfare and his desire to help them succeed in reaching their goals.

BRIEF Contents

This page intentionally left blank

TABLE OF Contents

Preface

BUSINESS STATISTICS, ***second edition***, is a one- or two-semester textbook written in a conversational tone designed to reduce the level of anxiety that many business students experience when taking a statistics course.

Many of today's business students are intimidated by their statistics textbook. These students often view their textbook as an obstacle to overcome rather than a tool to help them succeed. To address this issue, I have written *Business Statistics* in a straightforward, conversational tone that helps to reduce the anxiety many students experience with this course. My experience as both a writer and a teacher has taught me that students learn more effectively when they feel a personal connection with their instructor. Many traditional textbooks tend to "talk down" to students in a manner that many find difficult to understand. I prefer a textbook that "talks to" the students as I do in the classroom providing them a sense that I'm on their side, encouraging them every step of the way.

I strongly believe that students learn most effectively when they solve statistics problems as they learn new concepts rather than later (often right before the next exam). To facilitate this philosophy, I provide the student with a parallel problem that I have labeled "Your Turn," which allows them to work alongside the example that I am demonstrating in the chapter. I attempt to motivate them to do these exercises with a little levity but it's not beneath me to downright beg them to give it a try. I show the entire solution at the end of the chapter, so they can quickly check if their answer is correct. I call this my "learn it, do it, check it" cycle, where students learn by reading an example, doing a similar problem on their own, and finally checking their answer to confirm they understand. In effect, the textbook also plays the role of a workbook for the student, keeping them actively engaged in the learning process. Too often, students skim through an example that is completely solved for them in the text and convince themselves they understand the concept—that is, until they are trying to solve a similar problem in an exam for the first time. My approach encourages students to work through examples and confirm they grasp the concept before moving on to the next topic.

I have inserted many author's comments in the margins throughout each chapter, that provide useful insights along the way. This feature is analogous to the side comments you would make to your students during a lecture to help them better understand the material. I have found this to be an effective technique to help keep students focused on material that they may find confusing or overwhelming.

To help students be successful in your course, *Business Statistics,* second edition, has the following attributes:

- **Is written in a straightforward, conversational tone**—to help reduce the anxiety that many business students experience with the topic of statistics.
- **Utilizes a "learn it, do it, check it" cycle**—by providing parallel Your Turn problems throughout each chapter, the textbook essentially serves as a workbook allowing students to convince themselves they really understand a concept before moving on to the next topic.
- **Incorporates author's comments in the margins**—which are analogous to the side comments that an instructor would make during a lecture to help students better understand the material.

NEW TO THIS EDITION

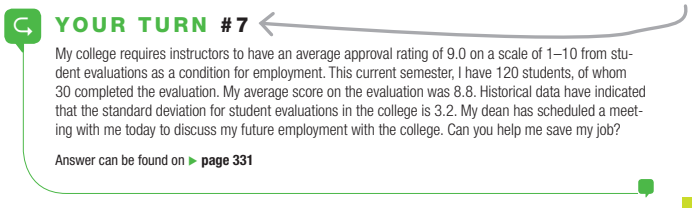
I am very excited to offer several new features to the second edition of *Business Statistics*. I have

- Added two new online chapters: Chapter 17, Decision Analysis, and Chapter 18, Nonparametric Statistics. Chapter 17, Decision Analysis, provides a detailed discussion of decision making under uncertainty and decision making under risk along with a step-by-step description on the construction and analysis of decision trees. Chapter 18, Nonparametric Statistics, provides a detailed description of the following procedures: Sign Test, Wilcoxon Rank-Sum Test, Wilcoxon Signed-Rank Test, Kruskal-Wallis One-Way ANOVA, and Spearman Rank-Order Correlation Coefficient. These chapters can be found on the textbook's Web site www.pearsonhighered.com/donnelly.

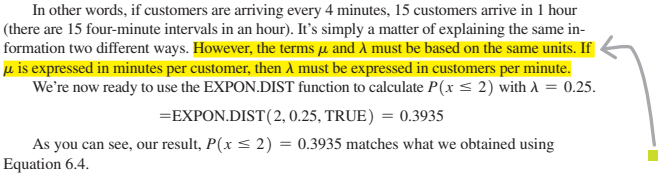
- Updated technology coverage to Microsoft Excel 2013, with instructions for Excel 2011 for Mac and Excel 2010 for Windows provided online as needed. Through my experience in the classroom, I have been aware of the increasing number of Mac users who have been frustrated with software compatibility issues. The version of the Excel Add-in PHStat that is utilized in this edition is compatible with Excel 2011 for Mac. These instructions can be found on the textbook's Web site www.pearsonhighered.com/donnelly.
- Increased the number of problems by 25%, totaling over 1,110 business-related problems. Additionally, 35% of the problems in the text are new or updated.
- Doubled the number of data sets included in problems, examples, and Your Turns, totaling over 340 data sets.
- Introduced topics of covariance and the correlation coefficient at the end of Chapter 3, Calculating Descriptive Statistics. The correlation coefficient is also covered in Chapter 14, Correlation and Simple Linear Regression.
- Used Excel functions to determine *p*-values and critical scores for hypothesis tests that use the normal, student's *t*, *F*, and chi-square distributions. This feature provides students with more options for this type of analysis.
- Removed critical sample mean and critical sample proportion as optional steps to hypothesis testing in Chapters 9 and 10 to streamline these procedures. These two topics are now included in the section describing Type II Errors at the end of Chapter 9.
- Added learning objectives to each chapter opener, which describe the skills that the student is expected to acquire after studying this material.
- Added an Index of Applications that allows faculty and students to conveniently locate specific types of problems and examples.

TEXTBOOK FEATURES

- **Current business examples that keep the students' interest**—Statistical procedures are applied to products and services that students can relate to such as the following:
  - Approximating the probability of an accident similar to the BP oil spill in the Gulf of Mexico occurring again in the near future
  - Liberty Mutual Insurance Company comparing the proportion of auto accident claims for clients with and without good student discounts
  - Comparing satisfaction scores for various smartphone brands
  - The shortage of Internet protocol (IP) addresses using the original IPv4 format



- **Your Turn problems after every major section**—These problems are strategically placed throughout the chapter and are designed to reinforce key concepts. Complete solutions to these problems can be found at the end of the chapter. I feel that students learn more effectively when they actively solve problems rather than skim through examples that are completely solved for them.


In other words, if customers are arriving every 4 minutes, 15 customers arrive in 1 hour (there are 15 four-minute intervals in an hour). It's simply a matter of explaining the same information two different ways. However, the terms μ and λ must be based on the same units. If μ is expressed in minutes per customer, then λ must be expressed in customers per minute.

We're now ready to use the EXPON.DIST function to calculate $P(x \leq 2)$ with $\lambda = 0.25$.

$$=\text{EXPON.DIST}(2, 0.25, \text{TRUE}) = 0.3935$$

As you can see, our result, $P(x \leq 2) = 0.3935$ matches what we obtained using Equation 6.4.

- **Step-by-step approach to complicated statistical procedures**—Many students tend to get "lost in the forest" when facing complicated procedures such as hypothesis testing, analysis of variance (ANOVA), and regression. My approach is to break these procedures down into bite-size, repeatable steps that can be applied to solve a variety of problems. As a result, the student has a consistent road map to follow when deciding how to proceed with a specific technique.
- **Highlighted text**—Throughout the textbook, I have highlighted text to draw the student's attention to a key point in the chapter. This will help to reinforce important concepts that could otherwise be overlooked by the student.

In quality control settings, businesses prefer a smaller standard deviation, which is an indication of more consistency in the process.

- **Author's comments**—In the margins, you will find comments that help clarify specific topics. These comments often point to an appropriate location in the chapter and are analogous to the comments you as an instructor might make in class to provide your students with some additional insight to the material.
- **Statistics in Practice**—Throughout the textbook, examples of how statistics is used in today's business environment are described in specially marked sections within the chapter. Examples include the following:
  - Government reports of unemployment figures using confidence intervals and the interpretation of these results
  - Comparison of customer feedback for different snack food products from Herr's Food Company
  - A statistical technique that health care insurance companies can rely on to investigate unusual billing practices from doctors' offices
  - Comparison of performances of Olympic athletes across different sporting events

Stats in Practice: Using the Range to Measure Volatility

Although the range as a measurement of variability has its limitations, it can be useful when measuring volatility of the stock market. If a volatile day for the stock market is defined as a swing of more than 200 points, we can use the daily range of the stock market index to identify volatile days. Figure 3.9 shows the number of volatile days in the stock market from 2000 until 2012.

The surge in volatile days from 2008 to 2011 reflects the uncertainty in the financial markets as the economy struggled through a significant recession. The quiet period seen in 2012 represents a more stable market as investors showed more confidence in the economy. In this example, we've demonstrated that we can use the range to conclude that the economy is showing signs of recovering from the recent financial turmoil.

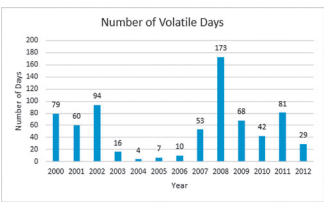
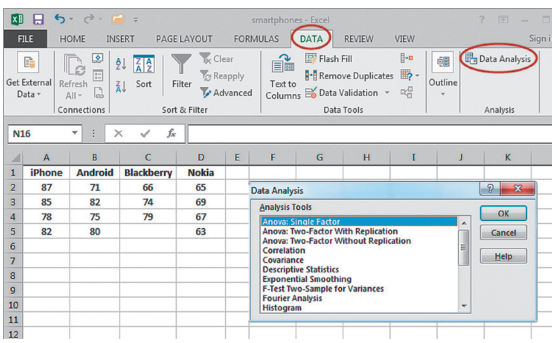


FIGURE 3.9 Number of Volatile Days in the Stock Market

- **The integration of Microsoft Excel® 2013**—I utilize Excel to demonstrate the use of technology in business statistics, but not at the expense of understanding the underlying concepts. I have spoken to students who tell me that they know how to perform a procedure such as ANOVA on the computer, but do not feel comfortable interpreting the results. By letting Excel do all the work they miss the opportunity to understand the underlying concepts of the technique. The philosophy employed in this textbook avoids this unfortunate situation.

FIGURE 11.8A Conducting a One-Way ANOVA Test Using Excel (Steps 1–3)



Supplements

STUDENT RESOURCES

Student Solutions Manual, by Bob Donnelly and Roman Erenshteyn, provides detailed, worked-out solutions to all even-numbered problems (ISBN-10: 0-321-93070-3; ISBN-13: 978-0-321-93070-5).

Study Cards for Business Statistics Software. This series of study cards, available for Excel® with XLSTAT™, Excel® 2013, Minitab®, JMP®, SPSS®, R®, StatCrunch™, and TI-83/84 graphing calculators, provides students with easy, step-by-step guides to the most common business statistics software. Available at www.myPearsonStore.com.

INSTRUCTOR RESOURCES

- **Instructor's Solutions Manual**, by Bob Donnelly and Roman Erenshteyn, provides detailed, worked-out solutions for all problems. Available for download at www.pearsonhighered.com/irc.
- **Test Bank**, by Bob Donnelly, includes true/false, multiple-choice, fill-in, and problem-solving questions based on the definitions, concepts, and ideas developed in each chapter of the text. Available for download at www.pearsonhighered.com/irc.
- **TestGen®** (www.pearsonhighered.com/testgen) enables instructors to build, edit, print, and administer tests, using a computerized bank of questions developed to cover all the objectives in the text. TestGen is algorithmically based, allowing instructors to create multiple but equivalent versions of the same question or test with the click of a button. Instructors can also modify test bank questions or add new questions. Tests can be printed or administered online. The software and test bank are available for download from Pearson Education's online catalog (www.pearsonhighered.com/irc).

TECHNOLOGY RESOURCES

Online Resources can be downloaded from www.pearsonhighered.com/donnelly and include the following:

- Chapters 17, Decision Analysis, and Chapter 18, Nonparametric Statistics.
- **Data sets** in Excel format for chapter problems, examples, and Your Turns.
- **Excel for Mac 2011/Excel 2010 for Windows instructions** for students to refer to when Excel instructions shown in the textbook are not compatible with the Excel for Mac 2011 or Excel 2010 for Windows versions of Excel.
- **Online case files and answers**, by Bob Donnelly, can be used in conjunction with the chapters to further class discussions or to use as class projects or homework assignments.

PHStat, the latest version of PHStat, the Pearson statistical add-in for Windows-based Excel 2003, 2007, and 2010. This version is also compatible with Excel 2011 for Mac. PHStat can be downloaded from www.pearsonhighered.com/phstat.

MyStatLab™ Online Course (access code required) MyStatLab is a course management systems that delivers proven results in helping individual students succeed. MyStatLab can be successfully implemented in any environment—lab-based, hybrid, fully online, traditional—and demonstrates the quantifiable difference that integrated usage has on student retention, subsequent success, and overall achievement. MyStatLab's comprehensive online gradebook automatically tracks students' results on tests, quizzes, homework, and in the study plan. Instructors can use the gradebook to provide positive feedback or intervene if students have trouble. Gradebook data can be easily exported to a variety of spreadsheet programs, such as Microsoft Excel.

MyStatLab™ **MyStatLab** provides engaging experiences that personalize, stimulate, and measure learning for each student.

- **Tutorial Exercises with Multimedia Learning Aids:** The homework and practice exercises in MyStatLab align with the exercises in the textbook, and they regenerate algorithmically to give students unlimited opportunity for practice and mastery. Exercises offer immediate helpful feedback, guided solutions, sample problems, animations, videos, and eText clips for extra help at point-of-use.
- **Adaptive Study Plan:** Pearson now offers an optional focus on adaptive learning in the study plan to allow students to work on just what they need to learn when it makes the most sense to learn it.
- **Additional Statistics Question Libraries:** In addition to questions that are aligned with your textbook, MyStatLab courses come with two additional question libraries. **Four hundred and fifty Getting Ready for Statistics** questions offer the developmental math topics students need for the course. **One thousand Conceptual Question Library** require students to apply their statistical understanding.
- **StatCrunch™:** MyStatLab includes a Web-based statistical software, StatCrunch, within the online assessment platform so that students can easily analyze data sets from exercises and the text. In addition, MyStatLab includes access to **www.StatCrunch.com**, a Web site where users can access shared data sets, conduct online surveys, perform complex analyses, and generate compelling reports.
- **Integration of Statistical Software:** Knowing that students often use external statistical software, we make it easy to copy our data sets, both from the ebook and the MyStatLab questions, into software programs. Students also have access to a variety of support tools—Technology Instruction Videos, Technology Study Cards, and Manuals for select titles.
- **Business Insight Videos:** Ten engaging videos show managers at top companies using statistics in their everyday work. Assignable questions available.

- **StatTalk Videos:** Fun-loving statistician Andrew Vickers takes to the streets of Brooklyn, New York, to demonstrate important statistical concepts through interesting stories and real-life events. A series of 24 videos with accompanying assessment questions and instructor's guide available.
- **Expert Tutoring:** Although many students describe the whole of MyStatLab as "like having your own personal tutor," students also have access to live tutoring with qualified statistics instructors via MyStatLab.

And, MyStatLab comes from a **trusted partner** with educational expertise and an eye on the future.

- Knowing that you are using a Pearson product means knowing that you are using quality content. That means that our eTexts are accurate, that our assessment tools work, and that our questions are error free. And whether you are just getting started with MyStatLab, or have a question along the way, we're here to help you learn about our technologies and how to incorporate them into your course.

To learn more about how MyStatLab combines proven learning applications with powerful assessment, visit **www.mystatlab.com** or contact your Pearson representative.

MyStatLab™ Ready to Go Course (access code required): These new Ready to Go courses provide students with all the same great MyStatLab features that you're used to, but make it easier for instructors to get started. Each course includes preassigned homework and quizzes to make creating your course even simpler. Ask your Pearson representative about the details for this particular course or to see a copy of this course.

MathXL® for Statistics Online Course (access code required): MathXL® is the homework and assessment engine that runs MyStatLab. (MyStatLab is MathXL plus a learning management system.)

With MathXL for Statistics, instructors can

- Create, edit, and assign online homework and tests using algorithmically generated exercises correlated at the objective level to the textbook.
- Create and assign their own online exercises and import TestGen tests for added flexibility.
- Maintain records of all student work, tracked in MathXL's online gradebook.

With MathXL for Statistics, students can

- Take chapter tests in MathXL and receive personalized study plans and/or personalized homework assignments based on their test results.
- Use the study plan and/or the homework to link directly to tutorial exercises for the objectives they need to study.
- Knowing that students often use external statistical software, we make it easy to copy our data sets, both from the ebook and the MyStatLab questions, into software like StatCrunch™, Minitab, Excel, and more.

MathXL for Statistics is available to qualified adopters. For more information, visit our Web site at www.mathxl.com or contact your Pearson representative.

StatCrunch™ StatCrunch is powerful Web-based statistical software that allows users to perform complex analyses, share data sets, and generate compelling reports of their data.

- **Collect.** Users can upload their own data to StatCrunch or search a large library of publicly shared data sets, spanning almost any topic of interest. An online survey tool allows users to quickly collect data via Web-based surveys.
- **Crunch.** A full range of numerical and graphical methods allow users to analyze and gain insights from any data set. Interactive graphics help users understand statistical concepts, and are available for export to enrich reports with visual representations of data.
- **Communicate.** Reporting options help users create a wide variety of visually appealing representations of their data.

Full access to StatCrunch is available with a MyStatLab kit, and StatCrunch is available by itself to qualified adopters. For more information, visit our Web site at www.StatCrunch.com or contact your Pearson representative.

PowerPoint® Lecture Slides. These presentations are available for each chapter. The PowerPoint slides provide an instructor with individual lecture outlines to accompany the text. The slides include many of the figures and tables from the text. Instructors can use these lecture notes as is or can easily modify the notes to reflect specific presentation needs.

The Student Edition of Minitab® is a condensed edition of the professional release of Minitab statistical software. It offers the full range of statistical methods and graphical capabilities, along with worksheets that can include up to 10,000 data points. Individual copies of the software can be bundled with the text (ISBN-10: 0-13-143661-9; ISBN-13: 978-0-13-143661-9).

JMP Student Edition is an easy-to-use, streamlined version of JMP desktop statistical discovery software from AS Institute, Inc., and is available for bundling with this text (ISBN-10: 0-321-89164-3; ISBN-13: 978-0-321-89164-8).

SPECIAL THANKS

I would like especially to thank Deirdre Lynch and Marianne Stepanian for their support and guidance with this second edition, Dana Bettez for keeping the entire project on schedule, Erin Lane for promoting this edition with her marketing skills, Joe Vetere for producing professional quality figures and screenshots in the textbook, Peggy McMahon for coordinating the production process with PreMediaGlobal, Sonia Ashraf for coordinating the reviewers, Amy Ray for improving my writing with her editorial skills, Cathy Zucco-Teveloff and Dirk Tempelaar for their accuracy checking to ensure an error-free textbook, and David Stephan for the development of the Excel Add-in, PHStat.

And most importantly, I would like to thank my wife, Debbie, for her constant love and support throughout the entire project.

Acknowledgments

This textbook was shaped and contributed to by dozens of class testers, consultants, focus group participants, and reviewers who gave generously of their time, expertise, and creativity.

REVIEWERS FOR THE SECOND EDITION

Henry Ander, *Arizona State University*
Kristian Braekkan, *Gustavus Adolphus College*
Chen-Huei Chou, *College of Charleston*
Mark Dahlke, *Colorado State University*
Joan Donohue, *University of South Carolina*
Levent Kaan, *University of Texas Dallas*
Andrew Koch, *James Madison University*
Latika Lagalo, *Emory University*
Rutilio Martinez, *University of Northern Colorado*
Lee McClain, *Western Washington University*
Angela Mitchell, *Wilmington College*
Alex Olbrecht, *Ramapo College of New Jersey*
Robert Patterson, *Penn State-Erie*
Leonard Presby, *William Patterson University*
Carol Puttman, *Penn State-Erie*
Mike Racer, *University of Memphis*
Jena Shafai, *Bellvue University*
Ruben Veliz, *Marymount College*
Candace Sorenson, *Marylhurst University*
Ronald Young, *Stark State College*

Class Testers/Consultant Board/Focus Group Participants/Reviewers for the First Edition

ALABAMA

Scott Bailey, *Troy University*

ARIZONA

Jason Bronowitz, *Arizona State University*
Linda Chattin, *Arizona State University*
Ashley Jacobson, *Chandler Gilbert Community College*
Nicolas Rouse, *Phoenix College*
Susan Sandblom, *Scottsdale Community College*
Yvonne M. Sandoval, *Pima Community College*

ARKANSAS

Tony Hunnicutt, *Ouachita Technical College*

CALIFORNIA

Asatar Bair, *City College of San Francisco*
Michael Brady, *California State University, Dominguez Hills*
Min Li, *California State University, Sacramento*
Khosrow Moshirvaziri, *California State University, Long Beach*
Christopher O'Byrne, *San Diego State University*
Ozgur Ozluk, *San Francisco State University*
Hamid Pourmohammadi, *California State University–Dominquez Hills*
Hindupur Ramakrishna, *University of Redlands*
Sunil Sapra, *California State University, Los Angeles*
Mark Tendall, *Stanford University*

COLORADO

Dale DeBoer, *University of Colorado, Colorado Springs*
Eric Huggins, *Fort Lewis College*
Alexandre Probst, *Colorado Christian University*
Richard Turley, *Colorado State University*

CONNETICUT

Matt Rafferty, *Quinnipiac University*

DELAWARE

Roman Erenshteyn, *Goldey-Beacom College*
James Ford, *University of Delaware*
Angela Mitchell, *Wilmington College*

FLORIDA

Dipankar Basu, *Miami University*
James Mirabella, *Jacksonville University*
Gary Smith, *Florida State University*

GEORGIA

Kim Gilbert, *University of Georgia*
Katherine McClain, *University of Georgia*
Komanduri Murty, *Fort Valley State University*
Hilde Patron-Boenheim, *University of West Georgia*

ILLINOIS

Harold Beck, *Southern Illinois University Edwardsville*
Yanli Cui, *University of Illinois at Chicago*
Jian Li, *Northeastern Illinois University*

INDIANA

Jason Davidson, *Butler University*
Ellen Gundlach, *Purdue University*
Keith Starcher, *Indiana Wesleyan University*

IOWA

Rick Jerz, *St. Ambrose University*
Blake Whitten, *University of Iowa*

KENTUCKY

Bob Gillette, *University of Kentucky*
Kusum Singh, *University of Kentucky*
Meenu Singh, *Murray State University*

LOUSIANA

Brandi Guidry, *University of Lousiana Lafayette*
Begona Perez-Mira, *Northwestern State University of Louisiana*
Zhiwei Zhu, *University of Louisiana at Lafayette*

MARYLAND

John Beyers, *University of Maryland–University College*
Michael Kulansky, *University of Maryland—University College*
Elkanah Faux, *Bowie State University*
Steve South, *Montgomery College, Germantown*

MASSACHUSETTS

Richard McGowan, *Boston College*
Bharatendra Rai, *UMASS–Dartmouth*

MICHIGAN

Fred Gruhl, *Washtenaw Community College*

MISSISSIPPI

Daniel Cameron Montgomery, *Delta State University*
Stephen Trouard, *Mississippi College*

MISSOURI

Eli Snir, *Washington University in St. Louis*

NEBRASKA

Karl Petersen, *Bellevue University*
Jena Shafai, *Bellevue University*

NEW HAMPSHIRE

Irwin Bramson, *New Hampshire College*

NEW JERSEY

Emma Bojinova, *Canisius College*
Ellen Benowitz, *Mercer County Community College*
Raed Dandan, *William Paterson University*
Martin Gritsch, *William Paterson University*
David Letcher, *The College of New Jersey*
Lan Nygren, *Rider University*
Penina Orenstein, *Seton Hall University*
Leonard Presby, *William Paterson University*
Harold Schneider, *Rider University*

NEW MEXICO

Sue Stockly, *Eastern New Mexico University*

NEW YORK

Chen-Huei Chou, *College of Charleston*
Kevin Caskey, *SUNY–New Paltz*
Ann Brandwein, *Baruch College*
Chun-Pin Hsu, *York College, City University of New York*
Jan Pitera, *Broome Community College*
Sean Simpson, *Westchester Community Colleg*e

NORTH CAROLINA

Alexander Deshkovski, *North Carolina Central University*
Casey DiRienzo, *Elon University*
Mahour Parast, *University of North Carolina–Penbroke*

OHIO

Pam Boger, *The Ohio State University*
Susan Emens, *Kent State University at Trumbull*
Eugene Jones, *The Ohio State University*
Joe Nowakowski, *Muskingum University*
Eddy Patuwo, *Kent State University*
Roxana Postolache, *Capital University*
Deborah Rumsey-Johnson, *The Ohio State University*
Anthony Sterns, *Kent State University, University of Maryland—University College*
Robert Stoll, *Cleveland State University*
Michael Welker, *Franciscan University of Steubenville*

OKLAHOMA

Ramona Piearcy, *Connors State College*
Allen White, *Bacone College*
Diana Wolfe, *Oklahoma State University at Oklahoma City*

PENNSLYVANIA

Mohamed Albohali, *Indiana University of Pennsylvania*
Ozgun Caliskan Demirag, *Pennsylvania State-Behrend*
Deborah Gougeon, *University of Scranton*
Ian Langella, *Shippensburg University of Pennsylvania*
Edward Mathis, *Villanova University*
Jerrold H. May, *KGSB/University of Pittsburg*h
Robert O. Neidigh, *Shippensburg University*
Rick Tannery, *Slippery Rock University*
Ray Venkataraman, *Pennsylvania State–Behrend*
Charles Wilf, *Duquesne University*

SOUTH CAROLINA

Joan Donohue, *University of South Carolina*
Lisa Bosman, *Clemson University*
Chen-Huei Chou, *College of Charleston*
Kent Foster, *Winthrop University*
Renu Singh, *South Carolina State University*

TENNESSEE

Mohammad Ahmadi, *The University of Tennessee at Chattanooga*
Michael Racer, *University of Memphis*
Carolyn Rochelle, *East Tennessee State University*
Jeffrey Schultz, *Christian Brothers University*

TEXAS

Jacob Dell, *University of Texas at San Antonio*
Levon R. Hayrapetyan, *Houston Baptist University*
Mark Leung, *University of Texas at San Antonio*
Carolyn Monroe, *Baylor University*
Ranga Ramasesh, *Texas Christian University*
Joseph C. Rhodes Jr., *Lone Star College*
Jesus Tanguma, *The University of Texas–Pan American*
Grace Vaughn, *El Paso Community College*

UTAH

Camille Fairbourn, *Utah State University Brigham City Regional Campus*
Don Gren, *Salt Lake Community College*

VIRGINIA

Kelly Alvey, *Old Dominion University*
Niels-Hugov Blunch, *Washington and Lee University*
Stephen Custer, *Virginia Commonwealth University*
Drew Koch, *James Madison University*
Quinton Nottingham, *Virginia Polytechnic Institute and State University*
Harvey Singer, *George Mason University*
Faye Teer, *James Madison University*

WAHINGTON

Sung Ahn, *Washington State University*
Stergio Fotopoulos, *Washington State University*

WASHINGTON D.C.

Rick Gibson, *American University*
Darius Singpurwalla, *George Washington University*
Ernest Zampelli, *Catholic University of America*

WISCONSIN

Patricia A. Mullins, *University of Wisconsin–Madison*
Cathy Poliak, *University of Wisconsin–Milwaukee*
Lee J. Van Scyoc, *University of Wisconsin–Oshkosh*

Student Reviewers/Focus Group Participants for the First Edition

This textbook was shaped and contributed to by hundreds of students who gave generously of their time, expertise, and creative unique perspective as students. This is who this textbook is for and we appreciate your input. Of particular note, are the following students:

Jay Bhatt, *Goldey-Beacom College*
Rafie K. Boghozian, *East Tennessee State University*
Lisa Bosman, *College of Menominee Nation*
Anne Drougas, *Dominican University*
Ashraf El-Houbi, *Lamar University*
Douglas Evans, *The Ohio State University*
Kevin Foster, *City College of New York*
Fred Howington, *Washtenaw Community College*
Molly Jensen, *University of Arkansas*
Constance McLaren, *Indiana State University*
Lan Ma Nygren, *Rider University*
Eric Ozsen, *San Francisco State Universtiy*
Gopi Parimi, *Goldey Beacom College*
Robert Patterson, *Penn State–Erie, The Behrend College*
Nora Perrone, *Newmann University*
Phillip Pfaff, *Canisius College*
Hamid Pourmohammadi, *California State University–Dominquez Hills*
Alex Riker, *Boston College*
Sanjani Shah, *Goldey-Beacom College*
Tim Sullivan, *Southern Illinois University Edwardsville*
Minghe Sun, *The University of Texas at San Antonio*
Cori Thompson, *Madison Area Technical College*
Linda Williams, *Tidewater Community College*

DEAR STUDENTS

Information overload is a recognized mark of our time, and the daily data tsunami isn't going to recede any time soon. There's no doubt you will be exposed to an unprecedented volume of data and information throughout your career. While many around you are crying "I'm overwhelmed!" or even "man overboard!" you can take a different, much better approach. Developing the skills needed to organize and interpret important information is a key to success, especially in business.

The business statistics course you are beginning will give you invaluable tools that enable you to use data to make good business decisions. For example, if you were a manager for AT&T or Verizon, could you conclude that the dropped-call rates for the two companies are significantly different based on a sample of customers from each company? Would that data alone be enough to affect how you manage your business? Or, would you want to break it down further and analyze it by some additional other factor, the locations of dropped calls perhaps? After you have completed Chapter 10, you'll be able to analyze data such as this and answer these questions with a high degree of reliability.

Throughout your business career you will find people at every level making decisions that have real consequences affecting business profitability and the jobs of real people. A skilled data analysis can give you an amazing window on aspects of a business that are simply not apparent without it.

I have written this textbook with you, the student, as my focus and the overall goal of helping you succeed both in this course and later in your career. I developed my approach over many years of teaching, and on the basis of that experience, I offer the following advice to help you achieve your own goals:

- Make it to class regularly. If you don't, you'll miss the details that help you master the subject. No textbook, no matter how well written, can take the place of your instructor and the classroom interaction. Seriously, go to class!
- Take advantage of the "Your Turn" problems placed throughout the chapters. Solving them will reinforce key concepts and let you know if you fully understand the material. The solutions to the problems at the end of each chapter give you immediate feedback, so I encourage you not to peek at the answers before you solve the problems or you won't really know how you are doing!
- Solve as many of the chapter problems as you can before exams. Business statistics is not a subject that rewards cramming or winging it. The saying "practice makes perfect" holds true in this field, and working through a variety of problems will build your confidence during the semester. I have provided solutions to all the even-numbered problems in Appendix B at the end of the book.

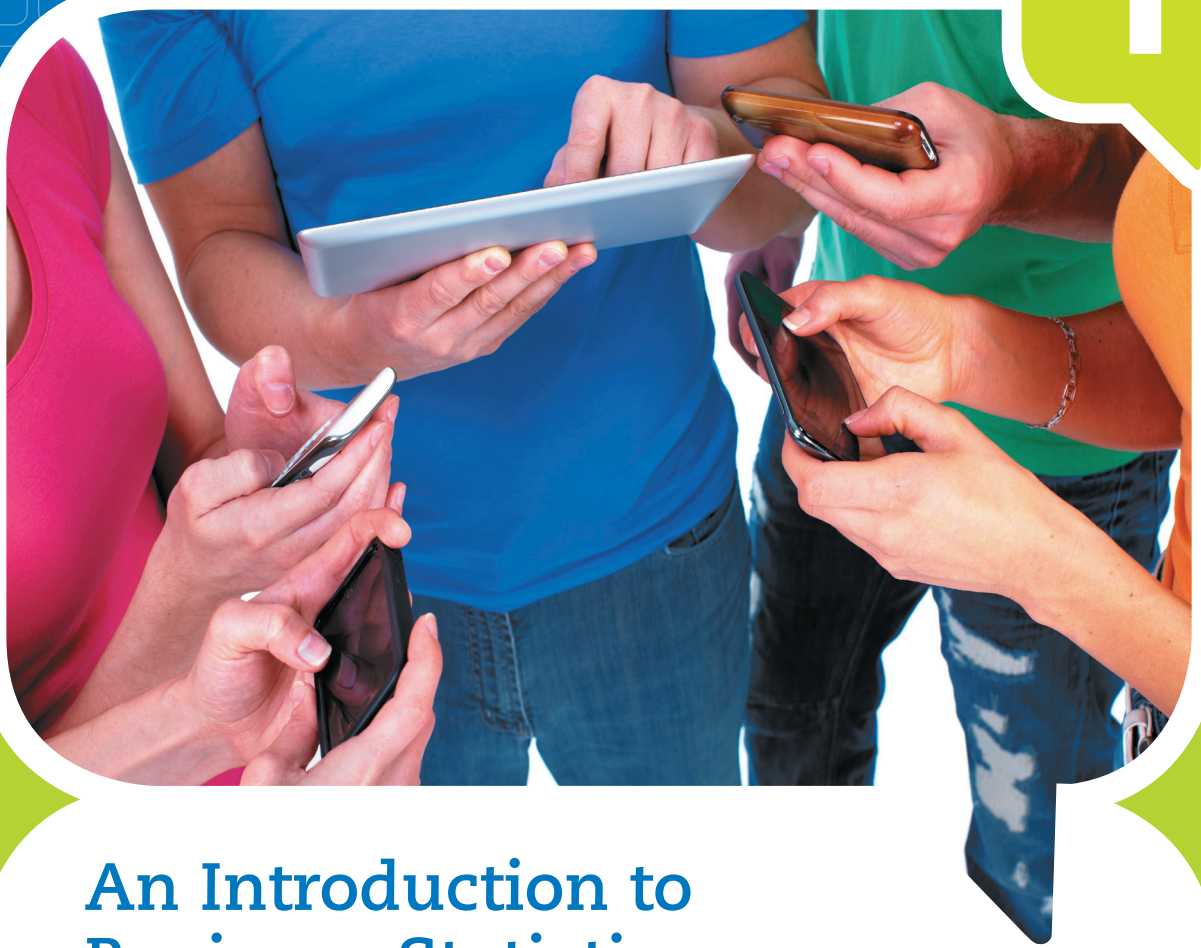
I hope that you come to share my enthusiasm for the value of business statistics this semester and that what you learn in this course contributes to your future success in the business world.

BOB DONNELLY

INDEX OF Applications

This page intentionally left blank

CHAPTER 1



An Introduction to Business Statistics

IN THIS CHAPTER, YOU WILL LEARN:

- How statistics is used in the business world.
- About the sources and data and the methods for collecting it.
- How to classify data by the level of measurement.
- To distinguish between time series and cross-sectional data.
- To distinguish between descriptive and inferential statistics.
- About the ethical implications of misusing statistics.

CHAPTER 1 MAP

Welcome to the world of statistics! Although some of you might be excited about learning statistics, others of you are probably less than thrilled about it. Perhaps you have to take "sadistics" because your major requires it. However, before you write off the value of this learning opportunity, let's discuss the role that statistics can play in your life. In today's world, everyone is a consumer of statistics. By this, I mean that you are continually surrounded by data and statements about those data in an effort to influence you to purchase something, vote for someone, or change your opinion about an issue. Consider the following examples:

- When CBS Sports announces that 109 million viewers tuned in to the 2013 Super Bowl, do you understand the method used to determine this number? (How does CBS know that I, you, or anyone else watched the game?)
- When we hear on TV that President Barack Obama's approval rating is 56% and in small print see $\pm 4\%$, do you understand the significance of this percentage?
- When you read an advertisement claiming that a new product is recommended by four out of five doctors, do you question the validity of the claim? (For instance, were the doctors paid for their endorsements?)
- When an online survey reports that Canon digital cameras are preferred to Nikon, does it concern you that the majority of the respondents could be Canon loyalists who repeatedly voted, skewing the results, or that the survey was conducted on a Canon user's forum?

Never before in the history of humankind have people had more data and information at their fingertips than you do at this moment. Statistics can have a powerful effect on our feelings, our opinions, and the decisions that we make in our personal and professional lives. As a result, it's very important that the statistics we report are both accurate and unbiased to ensure that they are properly utilized.

1.1 Business Statistics and Their Uses

Statistics is the mathematical science that deals with the collection, analysis, and presentation of data, which can then be used as a basis for inference and induction.

Statistics is the mathematical science that deals with the collection, analysis, and presentation of data—data that can then be used as a basis for inference and induction. **Business statistics** are statistics applied to the business world in an effort to improve people's decision making in fields such as marketing, operations, finance, and human resources, to name a few. Let's look at a few examples of how business statistics can help an organization's decision makers.

Marketing Research

Organizations rely heavily on business statistics when they conduct marketing research to determine what consumers want. For example, Kellogg's could perform a taste test to determine if consumers prefer the company's Cheez-It crackers to Nabisco's Cheese Nips. (Being a life-long Cheez-It addict, I know where my vote will go.) Kellogg's could also gather information about each consumer participating in the test in an effort to determine if the people who prefer one brand to the other share similar characteristics. This would be useful information for Kellogg's future marketing efforts.

In the 1980s, Marriott conducted an extensive survey to see how potential customers felt about the company's current hotel offerings. Based on the results, Marriott designed a new hotel chain known as Courtyard by Marriott, which has been a huge success.

Advertising

Television networks set their advertising rates for commercials based on the sizes of viewing audiences. The networks receive the information from Nielsen Media Research, which collects data from approximately 25,000 U.S. households. The sample of households surveyed has been carefully selected so that the results can be used to infer the viewing habits of the entire country. Statistics are used to ensure the sample is properly chosen and to process the data into meaningful information for the networks. Based on this information, CBS was able to charge a record $3.8 million for a 30-second Super Bowl ad in 2013.

Business statistics are statistics applied to the business world to help improve decision making.

Operations

Statistics can also be used to help businesses operate better. Quality control, for example, is a vital concern for all successful organizations. The combination of business statistics and quality control is a marriage made in heaven. If properly implemented, statistics can help manufacturing and service organizations monitor their processes and determine when quality problems begin to occur. For instance, Kellogg's can use statistics to determine if my Cheez-Its are overbaked or too salty. (Based on the box I'm sampling from at this moment, I'd say the company is doing an excellent job with its statistical quality control.)

Finance

When I began writing this textbook, the U.S. economy was in the middle of a deep recession. Part of the economic downturn was due to the poor lending practices of banks, particularly in the mortgage industry. If used properly, business statistics is an excellent tool to help banks identify consumers who are good credit risks and those who are not based on characteristics such as income, education, and home-ownership. For example, Fair Isaac Corporation (FICO) developed the credit scoring system (FICO score) currently used by the industry and which is based on a variety of statistical techniques.

Stats in Practice: Careers in Statistics

It's a great time to be in the job market if you have an interest and aptitude for statistics. This is due to an increase in demand for the skills of individuals who have developed a level of statistical literacy. The biggest driver for this demand is a global society that is both data-rich and data-dependent as technology advances at an explosive rate. In recognition of this development more than 150 professional organizations, including the American Statistical Association, have designated 2013 as the International Year of Statistics.

There are abundant employment opportunities in today's business world that require expertise in statistics. The following jobs are just a small sampling of the type of jobs available to people who have mastered this skill set.

- Transportation statistician—Our daily lives are surrounded by transportation systems that require ongoing attention by people with statistical talents. Examples include:
  - Delivery companies such as FedEx and UPS, which are always seeking opportunities to improve the efficiency of their delivery systems by analyzing data.
  - Airlines such as American Airlines rely on statistics for their yield management system which determines airfares on a continuous basis. You can thank these systems when you experience unexpected swings in airfare as you book your next vacation!
- Financial analyst—Using a variety of statistical tools, financial analysts provide investment advice to businesses and individuals. Banks and investment firms are examples of organizations seeking this type of position.
- Actuary—In order to maintain profitability, insurance companies need actuaries to analyze risk factors for their customers in order to establish appropriate premiums for their service. Statistical techniques play a major role in this analysis. Hospitals, banks, and government agencies are other examples of organizations that rely on these skills.
- Sports statistician—Many professional baseball, basketball, and football teams have hired statisticians in an effort to gain a competitive advantage. Examples include:
  - Assistant General Manager Peter Brand, who was dramatized in the 2011 film *Moneyball*, used statistical analysis to help the Oakland Athletics assemble a competitive Major League Baseball team during the 2002 season. The Boston Red Sox won the World Series in 2004 (their first since 1918!) relying on some of the same statistical modeling used by the Oakland Athletics.
  - The Memphis Grizzlies, an NBA team, recently hired a vice president of basketball operations because of his statistical expertise.
- Political analyst—A great deal of attention has been paid to predicting political outcomes, such as elections, using statistical tools.
  - Nate Silver is an American statistician who has developed an impressive reputation for his accurate predictions in the political arena. I recommend visiting his Web site, fivethirtyeight.blogs.nytimes.com, to gain some insight into the benefits of statistics in this field.

And of course I'd be remiss not mentioning one last type of employment in the field of statistics—education. I can personally testify that teaching statistics to students and writing books to help them learn has been a very rewarding experience.

1.2 Data

Data are values assigned to observations or measurements and are the building blocks of statistical analysis.

Data are the foundation of the field of statistics and can be defined as the values assigned to specific observations or measurements. If I'm collecting data on my wife's snoring behavior, I can do so in different ways. I can measure how many times Deb snores over a 10-minute period. I can measure the length of each snore in seconds. I could also measure how loud each snore is using a descriptive phrase like "That one sounded like a bear just waking up from hibernation" or "Wow! That one sounded like a sea lion calling for its young." (How a sound like that can come from a person who can fit into a pair of size 2 jeans is beyond me.)

In each instance, I would be recording data on the same event but in a different form. In the first instance, I would be measuring a frequency, or number of occurrences. In the second instance, I would be measuring duration, or length of time. And in the final instance, I would be measuring the event by describing its volume using words rather than numbers.

Information is data that are transformed into useful facts that can be used for a specific purpose, such as making a decision.

However, data all by themselves are not particularly useful. By definition, data are just the raw facts and figures that pertain to a measurement of interest. **Information**, on the other hand, is derived from the facts for the purpose of making decisions. One of the major reasons to use statistics is to transform data into information. For example, Table 1.1 shows my golf scores over a two-month period. (For those of you who are non-golfers, lower scores are better.)

TABLE 1.1 | Golf-Score Data

DATE	SCORE
June 13	94
June 20	96
June 27	93
July 10	89
July 16	86
July 24	89

Each individual golf score would be considered a data point. By themselves, the data points have limited value, other than to suggest that I should not quit my day job. (To quit my day job and become a professional golfer, I would need a score in the 65–70 range.) To these data, let's add the fact that in my desperate efforts to buy a better golf game, I quietly purchased a brand new driver on July 1, thinking my wife would not notice a new club in my golf bag (she did). We may be able to conclude, with the help of statistics, that the purchase helped improve my game (imagine a doubtful look on Deb's face). In order to answer this question, we will need to employ a statistical analysis that will be covered later in this text. Stay tuned as I use statistics hopefully to convince my wife that I really needed this new club!

The Sources of Data

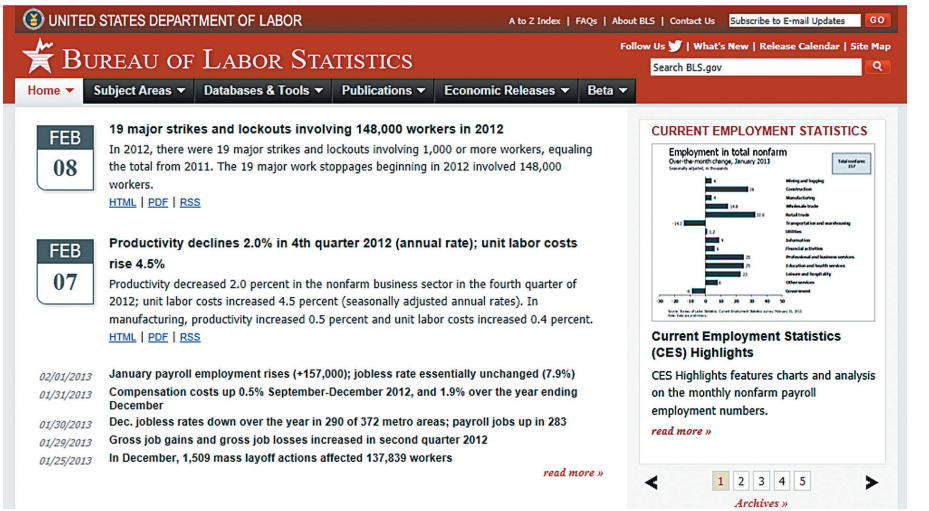
The Internet has become a rich source of data for statistics published by various industries. I once found a Japanese study on the effect of fluoride on toad embryos. Before this discovery, I had no idea toads even had teeth, much less a cavity problem!

Secondary data are data collected by someone else that you are "borrowing."

We classify the sources of data into two broad categories: primary data and secondary data. **Secondary data** are data somebody else has collected and made available for others to use. The U.S. government collects and publishes a variety of data that are readily available online. The U.S. Department of Labor collects mountains of data on topics such as consumer prices, inflation, unemployment, and productivity. The home page for the department's Web site is shown in Figure 1.1.

Every 10 years the U.S. Department of Commerce conducts a nationwide census to gather a wide variety of data related to the country's population. The data are used by Congress to make decisions about the funding for community services throughout the United States. The data are used to adjust the number of representatives each state is allowed to elect to Congress. You can find census data on the Department of Commerce's Web site.

FIGURE 1.1
The Home Page of the Bureau of Labor Statistics



The U.S. Geological Survey (USGS) provides an impressive assortment of scientific information that is used to manage water, energy, biological, and mineral resources across the Earth. For instance, did you know there are 250 species of squirrels in the world? If you don't believe me, you can look the information up at the USGS's Web site and become the local squirrel "expert" in your area.

The main drawback of using secondary data is that you have no control over how the data were collected. People tend to believe anything that's in print, even if it's not true. (You believe me, don't you?) Some of it is wrong, and, as you will learn later in the chapter, some of it is deliberately biased. The advantage to secondary data is that they are cheap (sometimes free) and that they are immediately available. For someone looking for data quickly, secondary data provide instant gratification (assuming the data are accurate, of course).

Primary data are data that you have collected for your own use.

Primary data, on the other hand, are data collected by the person or organization that eventually uses the data. This type of data can be expensive to acquire, but the main advantage is that primary data are your data, and you have nobody else to blame but yourself if you make a mess of it. You can obtain primary data in many ways, such as by direct observation, via experiments, and through surveys.

Direct observation is a method of gathering data while the subjects of interest are in their natural environment, often unaware they are being watched.

Direct observation is a method of gathering data while the subjects of interest are in their natural environment, often unaware they are being watched. Observing wild animals stalking their prey in the forest or teenagers at the mall on Friday night are two examples. (Or are they the same example?) The advantage of this method is that the subjects will unlikely be influenced by the data collection process.

A **focus group** is a direct observational technique whereby individuals are often paid to discuss their attitudes toward products or services in a group setting controlled by a moderator.

A **focus group** is a direct observational technique whereby individuals are often paid to discuss their attitudes toward products or services in a group setting controlled by a moderator. For example, Fisher Price heavily relies on focus groups of both adults and children to obtain valuable feedback on new toy ideas. The participants are aware they are being observed.

In an **experiment**, subjects are exposed to certain treatments and the data of interest are recorded.

In an **experiment**, subjects are exposed to certain treatments and the data of interest are recorded. An experiment that tests the use of a new medical drug is an example. Two different groups would be established: The first group would receive the new drug; the second group would be the control group. People in the control group would be told they are getting the new drug but would in fact get a placebo with no medication. The reactions from each group would be measured and compared to determine whether the new drug is effective.

The benefit of experiments is that they allow statisticians to control factors that could influence the results, such as the gender, age, and education of a participant. One major concern about collecting data through experiments is that the response of the subjects might be influenced by the fact they are participating in a study.

A **survey** involves directly asking people a series of questions and can be administered by e-mail, via the Web, through snail mail, face to face, or over the telephone.

A **survey** involves directly asking people a series of questions. Surveys can be administered by e-mail, via the Web, through snail mail, face to face, or over the telephone. (It's the telephone survey I'm most fond of, especially when I get the call just as I'm sitting down to dinner, getting into the shower, or finally making some progress on the chapter I'm writing.) The questionnaire needs to be carefully designed to avoid any bias that could affect

Research has shown that a question posed in a positive tone will tend to evoke a more positive response. A question posed in a negative tone will tend to evoke a more negative response. A good strategy is to pre-test your questionnaire before releasing it to the actual participants.

participants' responses or confuse them. Bias can occur when a question is stated in a way that encourages or leads a respondent to a particular answer. For example, "Wouldn't you agree that all drivers should wear a seat belt?" is a biased question. The influence the survey itself has on the responses of participants can also affect the quality of the data collected. Some participants will respond in a way they feel the survey would like them to. Figure 1.2 shows a portion of a survey I developed for users of the Claymont Community Center in Delaware. To encourage respondents to participate, an effective survey will state its purpose in the beginning, ask questions in a clear and concise manner, and place the more personal demographic questions last—when the respondents feel more comfortable with the process.

FIGURE 1.2
An Example of a Survey

Claymont Community Center
Health Clinic
Customer Service Survey

Dear Customer,

Our goal is to provide you with the best service possible. You can help us by taking a few minutes to answer the following questions about the Claymont Community Center (CCC). This valuable information will be used to help us to improve our services in the near future. The information gathered with this survey will be kept strictly confidential. Thank you for helping us.

(*Please place this form in the Customer Survey Box in the Health Clinic. If you need more time, you may return the form on your next visit or ask for a stamped envelope to return it by mail.*)

1. How long have you been using the Health Clinic? (Check one)
_____ Less than one month
_____ 1 to 6 months
_____ 6 months to a year
_____ 1 to 3 years

2. How satisfied are you with the quality of service being provided by the medical staff at the Health Clinic?
_____ Very Satisfied
_____ Satisfied
_____ Dissatisfied
_____ Very Dissatisfied
Comment_____

_____

3. Please provide the following personal information. Again, all information will remain confidential.

Zip code of your residence:_____

Gender: _____ Male _____ Female Age: _____

Marital Status: _____ Married _____ Single
_____ Separated _____ Divorced _____ Widowed

Are you the head of your household?
_____ Yes _____ No

What is your annual gross income?
_____ Less than $10,000 _____ $30–40
_____ $10–20 _____ $40+
_____ $20–30

We relied on SurveyMonkey to develop this textbook. Faculty and students tested the book before publication and provided valuable feedback through this Web site. I'm a satisfied customer!

Online surveys are a convenient way to acquire data. Companies such as SurveyMonkey provide people with a low-cost way to design surveys, collect responses, and analyze the data. The SurveyMonkey Web site claims that 80% of the Fortune 100 companies are users of the service. However, there are challenges to Internet surveys, which will be discussed later in this chapter.

To test your understanding of data sources, I encourage you to spend a few minutes answering the questions in the following Your Turn section.

YOUR TURN #1

Identify if the data required for each example are primary or secondary. For primary data, determine the best way in which the data should be collected. In other words, should the data be collected via observation, experiment, or survey?

1. Apple would like to measure the satisfaction levels of customers who purchased its new iPad product.
2. Pepsi would like to determine if consumers prefer the taste of Diet Pepsi to Diet Coke.
3. Cleveland State University needs to determine the current inflation rate to determine the annual salary increases for its staff for the upcoming year.
4. McDonald's would like to determine the average wait time for customers who use its drive-through windows during the lunch hour.

Answers can be found on ▶ **page 19**

The Two Main Types of Data

Another way to classify data is by whether they are quantitative or qualitative:

Quantitative data use numerical values while **qualitative data** relies on descriptive terms to describe something of interest.

- **Quantitative data** use numerical values to describe something of interest either by measuring it (such as its weight, height, or distance) or by counting it (such as the number of customers or repeat customers a business receives).
- **Qualitative data** use descriptive terms to measure or classify something of interest. One example of qualitative data is the name of a respondent in a survey and his or her level of education. Mathematical operations such as addition, subtraction, multiplication, and division cannot be performed on this type of data.

Classifying Data by Their Level of Measurement

Another important way to classify data is by the *way* the data are measured. This distinction is critical because it affects which statistical techniques we can use in our analysis of the data. The four levels of measurement are nominal, ordinal, interval, and ratio.

Nominal data are data that are described as a category or a label. Examples are gender (male or female), marital status (married, single, divorced, widowed), or yes/no responses.

A **nominal level of measurement** deals strictly with qualitative data assigned to predetermined categories. One example is the gender of a survey respondent, with the categories being male and female. This type of data is referred to as nominal data, or categorical data. It does not allow us to perform any mathematical operations on it, such as adding or multiplying. We can only give the data names and categorize them. (The word *nominal* actually means "pertaining to names.") We also cannot rank-order the data in any way from highest to lowest. An example is the state in which the survey respondent resides, such as Delaware or New Jersey (although I would try to rank my home state of Delaware on top.)

Other examples of nominal data are zip codes and telephone numbers, which can't be added, subtracted, or placed in a meaningful order of greater than or less than. Even though the data consist of numbers, they are handled just like qualitative data. Nominal data are considered the lowest level of data, and, as a result, the statistical techniques used to analyze them are the most restrictive.

Ordinal data have all the properties of nominal data, with the added feature that we can rank-order the values from highest to lowest.

An **ordinal level of measurement** can be conducted on data that are on the next level up on the food chain. Ordinal data has all the properties of nominal data, with the added feature that we can rank-order the values from highest to lowest. The following example explains ordinal measurement: Recently, I felt my manhood challenged by two neighbors who claimed their lawnmowers were faster than mine. Naturally, this had to be settled by a lawnmower race down our street, Gaebel Lane. Sadly, I present Figure 1.3, which shows that I lost the race. My neighbor Tom came in first (1), my neighbor Scott came in second (2), and I came in third (3).

FIGURE 1.3
An Example of Ordinal Measurement: Tom, Me, and Scott (from left to right) on Our Mowers

We still can't perform mathematical operations on this type of data, but we can say that Tom's lawnmower was faster than mine and Scott's. However, we cannot say how *much* faster because we didn't record the times of the lawnmowers. We just noted who came in first, second, and third. Ordinal data that have been collected do not allow us to make measurements between the categories or to say, for instance, that Scott's lawnmower is twice as fast as Bob's. For that, we need a different type of data. (In case you were wondering, I have been unsuccessful at convincing Deb that I need more horsepower to restore our family honor at the Second Annual Gaebel Lane Lawnmower Race.)

Education level is another example of ordinal data. A master's degree is ranked higher than a bachelor's degree, which in turn is ranked higher than a high school diploma. However, we are unable to measure the difference between these degrees in a meaningful, mathematical way. For instance, it would not be accurate to claim the difference between a master's and a bachelor's degree is more than the difference between a bachelor's and a high school degree. A property of ordinal data is that the differences between categories are not meaningful and, therefore, cannot be measured.

Ordinal data can also be numerical. One example of numerical ordinal data is when we rate movies with one, two, three, or four stars. Although we can order the movies by their ratings, we can't, for example, claim that a four-star movie is four times as good as a one-star movie.

Interval data, which are strictly quantitative, allow us to measure the differences between the categories with actual numbers in a meaningful way.

The **interval measurement level** is yet a higher level of measurement. It measures interval data, which are strictly quantitative. Temperature measurements in degrees Fahrenheit are an example of interval data. With this level, we can measure the differences between the categories with actual numbers in a meaningful way. For instance, 70°F is 5 degrees warmer than 65°F. However, multiplication and division can't be performed on this level of data. Why not? Simply because we cannot argue that 100°F is twice as warm as 50°F. The logic of this claim becomes more obvious when we convert the temperatures to the Celsius scale. The same two temperatures convert to 38°C and 10°C, respectively, so the twice-as-warm argument does not hold true. To help explain this, try baking a cake at twice the recommended temperature in half the recommended time. Yuck!

Another characteristic of interval data is that they do not have a true zero point. The term *true zero point* means that a zero data value indicates the absence of the object being measured. For instance, 0°F and 0°C do not represent the absence of temperature, even though it may feel like it.

Your grade point average (GPA) is another example of interval data. We can measure the difference between a 4.0 and a 2.0 GPA by simply subtracting the two values. However, it would not be an accurate statement to claim that a 4.0 student is twice as smart as a 2.0 student. Also, GPA has no true zero point because a 0.0 GPA does not indicate the absence of a grade point average.

Ratio data have all the features of interval data, with the added benefit of having a true zero point.

The most versatile of data types is the **ratio level of measurement**. Ratio data are as good as it gets as far as data are concerned. Examples of this type of data are ages, weights, prices, and salaries. Ratio data have all the features of interval data, with the added benefit of a true zero point. For instance, a zero salary indicates the absence of any salary. With a true zero point, we can say that a person who is six feet in height is twice as tall as a three-foot person or that a 20-year-old person is half the age of a 40-year-old.

For example, dollars are considered to be ratio data because $20 is twice as much as $10.

The distinction between interval and ratio data is a fine line. To help identify the proper scale, use the "twice as much" rule. If the phrase "twice as much" accurately describes the relationship between two values that differ by a multiple of two, then the data can be considered to be ratio level.

Table 1.2 summarizes the properties of the four levels of data measurement, and Figure 1.4 shows the relationship between the levels of data measurement and the two main types of data,

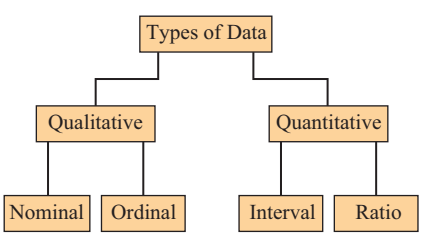


FIGURE 1.4
The Two Main Types of Data and Their Corresponding Levels

TABLE 1.2 | The Four Levels of Data Measurement: A Summary

LEVEL	DESCRIPTION	EXAMPLE
Nominal	Arbitrary labels for data No ranking allowed	Zip Codes (19808, 76137)
Ordinal	Ranking allowed No measurable meaning to the number differences	Education level (master's degree, doctorate degree)
Interval	Meaningful differences No true zero point	Calendar year (2014, 2015)
Ratio	Meaningful differences True zero point	Income ($48,000, $0)

which creates data measurement scales. As we explore different statistical techniques later in this book, we will revisit these different measurement scales. You will discover that specific techniques require certain types of data.

We can summarize classifying data by the level of measurement with the following baseball analogy:

- **Nominal data**. The numbers on the players' uniforms would be nominal data because these values are simply a label and are not used for any type of measurement. A player with the number 30 is not necessarily better than the player wearing number 15.
- **Ordinal data**. The batting order of the players would be considered ordinal data because we can rank-order the players going to the plate to hit. However, there is no useful or meaningful way to measure the differences between a player hitting 5th and a player hitting 6th.
- **Interval data**. The birth years of the players would be considered interval data. We can say that a player born in 1984 is three years older than a player born in 1987, but the "twice-as-much rule" does not hold true for calendar years. A person born in the year 1000 is not twice as old as a person born in the year 2000. This is because there is no true zero point with calendar years. The year 0 does not indicate the absence of age or time; it is merely an arbitrary reference point.
- **Ratio data**. The number of home runs hit by the best hitter on the team is an example of ratio data. A player with 20 home runs has twice as many as a player with 10 home runs. There is a true zero point for these data because zero home runs represent the absence of home-run hits.

Because this is such an important concept, which will be used later in the course, take a few minutes to answer the questions in the following Your Turn section before moving on in the chapter.

YOUR TURN #2

Identify the type of data (qualitative/quantitative) and the level of measurement for each of the following data sources:

a. Your IQ score
b. The price for one gallon of gasoline
c. The letter grade earned in your statistics course
d. The number of boxes of Frosted Flakes on the shelf of a grocery store
e. The types of cars driven by students in your class

Answers can be found on ▶ **page 19**