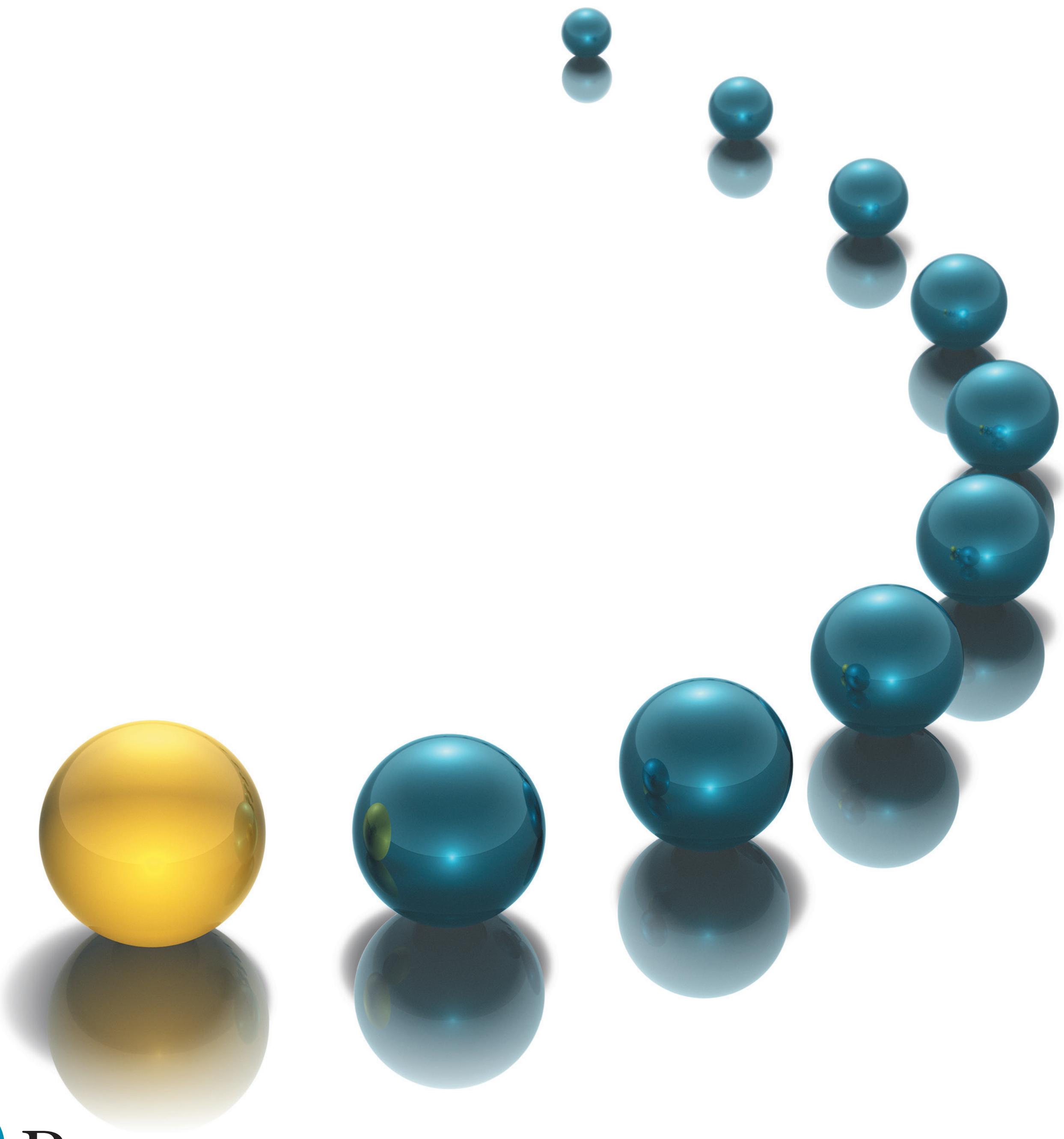
ROBERT STINE | DEAN FOSTER

Statistics for Business

Decision Making and Analysis

Third Edition



Pearson

Statistics for Business

DECISION MAKING AND ANALYSIS

THIRD EDITION

ROBERT STINE

Wharton School of the University of Pennsylvania

DEAN FOSTER

Emeritus, Wharton School of the University of Pennsylvania

Pearson

330 Hudson Street, NY NY 10013

Director, Portfolio Management, Math & Statistics: *Deirdre Lynch*
Portfolio Management Assistant: *Justin Billing*
Managing Producer, Math/Stats: *Karen Wernholm*
Content Producer: *Peggy McMahon*
Producer, Production & Digital Studio, Math & Statistics: *Aimee Thorne*
Manager, Courseware Content Development, Math & Stats: *Robert Carroll*
Manager, Courseware QA: *Mary Durnwald*
Director, Product Marketer, Math & Statistics: *Erin Kelly*
Product Marketing Assistant: *Jennifer Myers*
Senior Author Support/Technology Specialist: *Joe Vetere*
Manager, Rights and Permissions: *Gina Cheselka*
Manufacturing Buyer: *Carol Melville/ LSC Communications*
Production Coordination, Composition, Text Design, Illustrations: *Cenveo® Publisher Services*
Associate Director of Design: *Blair Brown*
Program Design Lead: *Barbara Atkinson*
Cover Design: *Tamara Newnam*
Cover Photo: *Getty Images/BeholdingEye*

The student edition of this title has been catalogued at the Library of Congress as follows:

Library of Congress Cataloging-in-Publication Data

Names: Stine, Robert A., author. | Foster, Dean P., author.
Title: Statistics for business : decision making and analysis / ROBERT STINE, Wharton School of the University of Pennsylvania, DEAN FOSTER, Wharton School of the University of Pennsylvania.
Description: Third Edition. | Boston : Pearson, 2016. | Revised edition of the authors' Statistics for business, 2013. | Includes index.
Identifiers: LCCN 2016016748| ISBN 9780134497167 (hardcover) | ISBN 0134497163 (hardcover)
Subjects: LCSH: Commercial statistics. | Statistics.
Classification: LCC HF1017 .S74 2016 | DDC 519.502/465--dc23
LC record available at https://lccn.loc.gov/2016016748

1 16

Pearson

Student Edition:
ISBN-13: 978-0-13-449716-7
ISBN-10: 0-13-449716-3

ABOUT THE AUTHORS

Robert Stine holds a Ph.D. from Princeton University. He has taught at the Wharton School since 1983, during which time he has regularly taught business statistics. During his tenure, Bob has received a variety of teaching awards, including regularly winning the MBA Core Teaching Award, which is presented to faculty for outstanding teaching of the required curriculum at Wharton. He also received the David W. Hauck Award for Outstanding Teaching, awarded to the most highly rated faculty member teaching in the Wharton undergraduate program. Bob actively consults for industry. His clients include the pharmaceutical firms Merck and Pfizer, and he regularly works with the Federal Reserve Bank of Philadelphia on models for retail credit risk. This collaboration has produced three well-received conferences held at Wharton. His areas of research include computer software, time series analysis and forecasting, and general problems related to model identification and selection. Bob has published numerous articles in research journals, including the *Journal of the American Statistical Association*, *Journal of the Royal Statistical Society*, *Biometrika*, and *The Annals of Statistics*.

Dean Foster holds a Ph.D. from the University of Maryland. He has taught at the Wharton School since 1992 and previously taught at the University of Chicago. Dean has taught courses in introductory business statistics, probability and Markov chains, statistical computing, and advanced statistics for managers. Dean's research areas are statistical inference for stochastic processes, game theory, machine learning, and variable selection. He is published in a wide variety of journals, including *The Annals of Statistics, Operations Research, Games and Economic Behaviour, Journal of Theoretical Population Biology,* and *Econometrica*. Currently Senior Principal Scientist at Amazon; Dean Foster is working on big data there, and his goal is to predict the sales of each and every product that Amazon sells.

Bob Stine and Dean Foster (along with Richard Waterman) have co-authored two casebooks: *Basic Business Statistics* (Springer-Verlag) and *Business Analysis Using Regression* (Springer-Verlag). These casebooks offer a collection of data analysis examples that motivate and illustrate key ideas of statistics, ranging from standard error to regression diagnostics and time series analysis. They also have collaborated on a number of research articles.

CONTENTS

Supplementary Material (online-only)

PREFACE

Knowledge of statistics is a great asset in business, but getting the most value from this asset requires knowing how to ask and answer the right questions. Choosing the right question and solving the problem correctly require an appreciation of business as well as the subtleties of statistics. Unless you understand the business issue from a finance, marketing, management, or accounting perspective, you won't see how statistics can help solve the problem. Performing the statistical analysis must wait until you have grasped the issue facing the business.

Solving Business Problems

This application-directed approach is key to business analytics and shapes our examples. We open each chapter with a business question that motivates the contents of the chapter. For extra practice, worked-out examples within each chapter follow our 4M (Motivation, Method, Mechanics, Message) problem-solving strategy. The *motivation* sets up the problem and explains the relevance of the question at hand. We then identify the appropriate statistical *method* and work through the *mechanics* of its calculation. Finally, the *message* answers the question in language suitable for a business presentation or report. Through the 4Ms, we'll show you how a business context guides the statistical procedure and how the results determine a course of action. Motivation and Message are critical. The Motivation answers the question "Why am I doing this analysis?". If you cannot answer that question, it's hard to get the statistics correct. The Message has to express your answer in language that is used in the business world. Understand the business first, then use statistics to help formulate your conclusion. Notice that we said "help." A statistical analysis by itself is not the final answer. You must frame that analysis in terms that others in the business will understand and find persuasive.

Our emphasis on the substantive use of statistics in business shapes our view that the ideal reader for this text is someone with an interest in learning how statistical thinking improves the ability of a manager to run or contribute to a business. Whether you're an undergraduate with an interest in business, an MBA looking to improve your skills, or a business owner looking for another way to get ahead of the competition, the key is a desire to learn how statistics can produce better decisions and insights from the growing amount of data generated in modern businesses.

We don't assume that readers have mastered the domains of a business education, such as economics, finance, marketing, or accounting. We do assume, though, that you care how ideas from these areas can improve a business. If you're interested in these applications—and we think you will be—then our examples provide the background you will need to appreciate why we want to solve the challenges that we present in each chapter. Readers with more experience will discover that we've simplified the technical details of some applications, such as those in finance or marketing. Even so, we think that the examples offer those with substantive experience a new perspective on familiar problems. We hope that you will agree that the examples are realistic and get to the heart of quantitative applications of statistics in business.

Technology

You cannot do research in modern applied statistics without computing. Data sets have grown in size and complexity, making it impossible to work out the calculations by hand. Rather than dwell on routine calculations, we rely on software (often referred to as a statistics package) to compute the results. Although we emphasize the use of technology, we give the formulas and illustrate the calculations introduced in each chapter so that you will always know what the software is doing. It is essential to appreciate what happens in the calculations: You need to understand how the calculations are done in order to recognize when they are appropriate and when they fail. That does not mean, however, that you need to spend hours doing routine calculations. Your time is precious, and there's only so much of it to go around. We think it makes good economic sense to take advantage of modern technology in order to give us more time to think harder and more thoroughly about the motivating context for an application and to successfully present the business message.

When we present results obtained with a calculator or computer, we typically round them. You don't need to know that the profits from a projected sale are \$123,234.32529. It's usually better to round such a number to \$123 thousand. To let you know when we've rounded a calculation, we say *about* or *approximately*. In expressions, we denote rounding with the symbol $\approx$, as in $1/6 \approx 0.167$.

To help you learn how to use software, each chapter includes hints on using Excel®, MinitabExpress®, and JMP® for calculations. These hints won't replace the help provided by your software, but they will point you in the right direction so that you don't spin your wheels figuring out how to get started with an analysis. Supplemental software study cards are available for specific packages.

Data

Statistical analysis uses data, and we've provided lots of data to give you the opportunity to have some real hands-on experience. As you read through the chapters, you'll discover a variety of data sets that include real estate markets, stocks and bonds, technology, retail sales, human resource management, and fundamental economics. These data come from a range of sources, and each chapter includes a discussion about where we found the data used in examples. We hope you'll use our suggestions and find more.

Prerequisite Knowledge

To appreciate the illustrative calculations and formulas, readers will need to be familiar with basic algebra. Portions of chapters that introduce a statistical method often include some algebra to show where a formula comes from. Usually, we only use basic algebra (up through topics such as exponents and square roots). Several chapters make extensive use of the logarithm function. If you're interested in business and economics, this is a function worth getting to know a lot better. The applications we've provided, such as modeling sales or finding the best price, show why the logarithm is so important. Occasionally, we give credit to calculus for solving a problem, but we don't present derivations using calculus. You'll do fine if you are willing to accept that calculus is a branch of more advanced mathematics that provides, among other things, the ability to derive formulas that have special properties. If you do know calculus, you'll be able to see where these expressions come from.

WHAT'S NEW IN THIS EDITION

This edition adds more of what readers have found really useful:

- **Business analytics** relies on linking data to business decisions. Businesses ranging from traditional banks to the latest game developers are clamoring for employees who can connect data and models to substantive business problems. This edition adds emphasis, examples, and illustrations that stress the importance of these connections. For example, previous editions introduced the 4M paradigm—motivation, method, mechanics, and message—that shows how to combine data and statistics to solve problems in business. This edition carries this metaphor further. By explicitly linking this paradigm to analytics, this edition shows that business analytics requires blending substantive relevance with statistical analysis.
- **Up-to-date applications** explore problems related to "big data" and introduce hot topics such as A/B testing that are popular in today's businesses. Although the methods behind these new topics are familiar within statistics, the names are new. This edition makes sure students know the new names so that they can link what they learn in the classroom to what they read online.
- This edition features more than **90 new and updated data sets**. The changed data range from examples used within chapters to those underlying exercises. Important, highly visible changes include "through the cycle" finance and economic time series that span the 2008 recession.
- **More than 100 enhanced exercises** remove ambiguities and capture nuances in revised data. Many of these changes address issues identified by tracking online student performance in completing related exercises in MyStatLab. Problems that were worded in a way that might confuse students were clarified.
- **Excel** is the workhorse tool of many businesses. This edition adds a section to every chapter that shows step by step how to complete analytic exercises with the latest version of Excel. Excel is the most popular software for introductory statistics, but some prefer the features offered by statistics packages such as Minitab or JMP. We've retained and updated hints in each chapter for these as well.
- It's the little things. **Hundreds of changes** have been made throughout this edition to emphasize and clarify key points. For example, this edition highlights additional tips throughout the text that help readers recognize important points that might be overlooked. Clarified explanations, analogies, and examples in every chapter encourage students to delve deeper and learn for themselves.

COVERAGE AND ORGANIZATION

We have organized the chapters of this book into four parts:

1. Variation
2. Probability
3. Inference
4. Regression Models

Part I. These chapters introduce summary statistics such as the mean and important graphical summaries, including bar charts, histograms, and scatterplots. Even if you are familiar with these methods, we encourage you to skim the examples in these chapters. These examples introduce important terminology that appears in subsequent chapters. A quick review will introduce the notation that we use (which is rather standard) as well as give you a chance to look at some interesting data. If you do skip past these, take advantage of the index of Key Terms in each chapter to find definitions and examples.

Part II. Many courses in mathematics now include topics from probability. Even if you have seen basic probability, you might benefit from reviewing how methods, such as Bayes' Rule, can be used to improve business processes (Chapter 8). If you plan to skip or move briskly through the rest of the chapters in Part 2, be sure that you're familiar with the concept of a random variable (Chapter 9). Statistical models use random variables to present an idealized description of the data in applications. Unless you're familiar with random variables, you won't appreciate the important assumptions that come with their use in practice. Chapter 11 describes special random variables used to model counts, and Chapter 12 defines normal random variables that appear so often in statistical models.

Part III. This part presents the foundations for statistical inference, the process of inferring properties of an entire population from those of a subset known as a sample. Even if you are not interested in quality control, we encourage you to read Chapter 14. Chapter 14 uses quality control to introduce a fundamental concept of inferential statistics, the sampling distribution and standard error. You can get by in statistics with a basic understanding of the concept of a sampling distribution, but the more you know about sampling distributions, the better. Each inferential procedure comes with a checklist of conditions that tell you whether your data and situation match up to the various inferential techniques in these chapters.

Part IV. The chapters in Part 4 describe regression modeling. Regression modeling allows us to associate how differences in data that describe one phenomenon are related to differences in others. Regression models are among the most powerful ways to use statistics in business, providing methods for assessing profitability, setting prices, identifying anomalies, and generating forecasts. We encourage you to slow down and take your time studying these chapters. Even if you don't see yourself doing statistics in your career in business, you can be sure that you will be presented with the results of regression models. Because the examples in these chapters allow us to describe the interconnectedness of several business processes at once, they become even more interesting than those in prior chapters. Be careful if you skip Chapter 20. The material in this chapter shows how to model a richer set of patterns and is less common in business textbooks, but we think these ideas are an essential component of every manager's tool set.

Case Studies

Each of the four main parts of this book includes two supplemental case studies called Statistics in Action. Each case study provides an in-depth look at a business application of statistics. Every case uses real data and takes students through the details of using those data to address a business question. For example, a case study for Part 1 explains details of stock market data, such as how stock returns account for dividends, and elaborates the nuances of financial data beyond the coverage in the surrounding chapters.

We've found that it is easy to have a "chapter-centric" view of any subject; you know how to approach a problem if the question identifies a chapter. Executing the right approach is more difficult without that sort of clue.

Case studies allow us to extend the statistical concepts introduced in the accompanying chapters in the context of a longer, more complex case. For example, the second case in Part 1 carefully explains how to interpret and use logarithms in the context of executive salaries. A case in Part 3 explores the use of many chi-squared tests in an operations management problem that resembles data mining. While logs, chi-squared tests, and issues of multiple testing all appear in the regular flow of the main chapters, case studies provide a means for us to cover these topics in more detail than we thought was appropriate for everyone.

Supplementary Chapters

For this edition, we've added a few supplementary chapters that are available online. These cover topics that are less common in the typical business stats course, but often useful. One chapter covers methods that are needed when the usual approaches don't apply. For example, suppose data are so skewed that one cannot use standard methods for building a confidence interval for the mean. What are you to do? The supplemental chapter **Alternative Approaches to Inference** gives an answer. Two other supplemental chapters go deeper into regression modeling. The chapter **Two-Way Analysis of Variance** goes beyond Chapter 26 and looks at two-way (and higher) analysis of variance, including those with randomized blocking and interactions. The chapter **Regression Modeling with Big Data** goes beyond Chapter 24 and the Statistics in Action cases with coverage of how to build regression models when confronted by "big-data" issues that have become more common in business.

FEATURES

Motivating Examples. Each chapter opens with a business example that frames a question and motivates the contents of the chapter. We return to the example throughout the chapter, as we present the statistical methods that provide answers to the question posed in the opening example.

4M Analytics Examples The 4M (Motivation, Method, Mechanics, Message) problem-solving strategy gives students a clear outline for solving any business problem. Each 4M example first expresses a business question in context, then guides students to determine the best statistical method for working the problem using statistical software, and, finally, frames the analysis in terms that others in the business world will understand.

What Do You Think? Short question sets throughout each chapter give students the opportunity to check their understanding of what they've just read. These questions are intended to be a quick check of key concepts and ideas presented in the chapter; most questions involve very little calculation. Answers are located in a footnote so that students can easily check their answers before moving on in the text.

tip **Tips.** We highlight useful hints for applying statistical methods within the exposition so that students don't miss them.

caution **Caution.** You'll see the caution icon next to material that might be confusing. You should be extra careful to make sure you understand the material being discussed.

✓ **Checklist.** Some statistics presume that the information presented satisfies several conditions or assumptions. For example, certain statistics only detect patterns that resemble lines. You would not want to use these if you were looking for a curve. To help you keep track of the assumptions, the conditions are collected in a checklist.

Best Practices. At the end of each chapter, we include a collection of tips for applying the chapter's concepts successfully and ethically.

Pitfalls. Most of the unintentional mistakes people make when learning statistics are avoidable and usually come from using the wrong method for the situation or misinterpreting the results. This feature at the end of each chapter provides useful tips for avoiding common mistakes.

Data Analytics: The authors analyzed aggregated student usage and performance data from MyStatLab™ for the previous edition of this text. The results of this analysis helped improve the quality and quantity of exercises that matter the most to instructors and students.

Software Hints. Each chapter includes hints on using Excel, Minitab, and JMP for calculations. These hints give students a jumping off point for getting started doing statistical analysis with software. Supplemental study cards for these and other software packages are available from the publisher.

Behind the Math. At the end of most chapters, a Behind the Math section provides interesting technical details that explain important results, such as the justification or interpretation for an underlying formula. If you are so inclined, they will help you appreciate the subtleties and logic behind the mechanics, but they are not necessary for using statistics.

Chapter Summary. These chapter-ending summaries provide a complete review of the content.

- **Key Terms** We provide an index of the chapter's key terms at the end of each chapter to give students a quick and easy way to return to important definitions in the text.
- **Objectives** This new feature provides a list of what students should understand after having read a chapter.
- **Formulas** Important formulas introduced within the chapter are restated.
- **About the Data** This feature provides sources for the data used throughout the chapter with further background.

Exercises. Each chapter contains a variety of exercises at escalating levels of difficulty in order to give students a full complement of practice in problem solving using the skills they've learned in the chapter. Types of exercises include Matching, True/False, Think About It, You Do It, and 4M Exercises. You'll find the data for the 4M and You Do It exercises on Pearson's Math and Statistics Resources Website: http://www.pearsonhighered.com/mathstats

- **Matching and True/False** exercises test students' ability to recognize the basic mathematical symbols and terminology they have learned in the chapter. We avoid unnecessary formulas, but certain symbols and terminology show up so often that students are well served to recognize them.
- **Think About It** exercises ask students to pull together the chapter's concepts in order to solve conceptual problems. You don't need a computer or calculator for most of these.
- **You Do It** exercises give students practice solving problems that reinforce the mechanics they've learned in the chapter. These exercises apply the methods of the chapter to data related to a business

application. Working through the steps of these exercises helps you practice the mechanics. We expect you to use a statistics software package for many of these.

- **4M Analytics** exercises are rich, challenging applications rooted in real business situations. These ask students to apply the statistical knowledge they've developed in the chapter to a set of questions about a particular business problem.

ACKNOWLEDGMENTS

We didn't develop our approach to business statistics in isolation. Our colleagues at Wharton have helped shape our approach to teaching statistics in business. Many of the ideas and examples that you'll find here arose from suggestions made by colleagues, including Andreas Buja, Sasha Rakhlin, Paul Shaman, Richard Waterman, and Adi Wyner. Over the years, members of our department have come to share a common attitude toward the use of statistics in business, and this text reflects that shared perspective. Most of the examples and many exercises from the text have been tried in other classes and improved using that feedback. We owe these friends a debt of gratitude for their willingness to talk about the fundamental use of statistics in business and to explore alternative explanations and examples.

Many thanks to the following reviewers for their comments and suggestions during the revision of this text.

Kunle Adamson, *DeVry University*
Elaine Allen, *Babson College*
Randy Anderson, *California State University—Fresno*
Djeto Assane, *University of Nevada, Las Vegas*
Rajesh K. Barnwal, *Middle Tennessee State University*
Dipankar Basu, *Miami University*
Mark Bloxom, *Alfred State College SUNY College of Technology*
Hannah Bolte, *Indiana University Bloomington*
David Booth, *Kent State University, Main Campus*
John E. Boyer, *Jr., Kansas State University*
Michael Braun, *Southern Methodist University*
Daniel G. Brick, *University of St. Thomas*
Nancy Burnett, *University of Wisconsin—Oshkosh*
Richard Cleary, *Bentley College*
Ismael Dambolena, *Babson College*
Anne Davey, *Northeastern State University*
Dr. Michael Deis, *Clayton University*
Frederick W. Derrick, *Loyola University Maryland*
Neil Desnoyers, *Drexel University*
Joan Donohue, *University of South Carolina*
Steve Erikson, *Babson College*
Nancy Freeman, *Shelton State Community College*
Daniel Friesen, *Midwestern State University*
Deborah J. Gougeon, *University of Scranton*
Christian Grandzol, *Bloomsburg University*
Betsy Greenberg, *University of Texas—Austin*
Ken Griffin, *University of Central Arkansas*
John Grout, *Berry College*
Warren Gulko, *University Of North Carolina, Wilmington*
Marie Halvorsen-Ganepola, *Notre Dame University*
Clifford B. Hawley, *West Virginia University*
Bob Hopfe, *California State University—Sacramento*
Max Houck, *West Virginia University*
David Hudgins, *University of Oklahoma*
Jeffrey Jarrett, *University of Rhode Island*
Chun Jin, *Central Connecticut State University*
Christopher K. Johnson, *Ph.D. University of North Florida*
Morgan Jones, *University of North Carolina*
Ronald K. Klimberg, *Saint Joseph's University*
David Kopcso, *Babson College*
Supriya Lahiri, *University of Massachusetts, Lowell*
Mark T. Leung, *University of Texas—San Antonio*
Tony Lin, *Ph.D., University of Southern California*
John McKenzie, *Babson College*
Kay McKinzie, *University of Central Arkansas*
Mark R. Marino, *Niagara University*
Dennis Mathaisel, *Babson College*
Sherryl May, *University of Pittsburgh*
Bruce McCullough, *Drexel University*
Richard McGowan, *Boston College*
Constance McLaren, *Indiana State University*
Robert Meeks, *Pima Community College*
Jeffrey Michael, *Towson University*

Prakash Mirchandani, *University of Pittsburgh*
Jason Molitierno, *Sacred Heart University*
Carolyn H. Monroe, *Baylor University*
Gourab Mukherjee, *University of Southern California*
Patricia Ann Mullins, *University of Wisconsin—Madison*
Quinton J. Nottingham, *Virginia Polytechnic & State University*
Keith Ord, *Georgetown University*
Michael Parzen, *Emory University*
M. Patterson, *Midwestern State University*
Robert Pred, *Ph.D., Temple University*
Leonard Presby, *William Paterson University*
Darrell Radson, *Drexel University*
Farhad Raiszadeh, *University of Tennessee—Chattanooga*
Ranga Ramasesh, *Texas Christian University*
Deborah Rumsey, *The Ohio State University*
John Saber, *Babson College*
Dr. Subarna Samanta, *The College of New Jersey*
Subarna Samanta, *The College of New Jersey*
Hedayeh Samavati, *Indiana University—Purdue Fort Wayne*
Rose Sebastianelli, *University of Scranton*
Omeed Selbe, *University of Southern California*
Gary Smith, *Florida State University*
Erl Sorensen, *Bentley College*
Bert Steece, *University of Southern California*
J. H. Sullivan, *Mississippi State University*
Dr. Kathryn A. Szabat, *LaSalle University*
Rajesh Tahiliani, *University of Texas—El Paso*
Patrick A. Thompson, *University of Florida*
Denise Sakai Troxell, *Babson College*
Bulent Uyar, *University of Northern Iowa*
John Wang, *Montclair State University*
Dr. William D. Warde, *Oklahoma State University, Main Campus*
Elizabeth Wark, *Worcester State University*
James Weber, *University of Illinois—Chicago*
Fred Wiseman, *Northeastern University*
Roman Wong, *Barry University*
Zhiwei Zhu, *University of Louisiana at Lafayette*

We would also like to thank our accuracy checkers Caroline Swift and Dirk Tempelaar. Thanks also to Lifland et al., and our Pearson Education team for their help and support, especially: Deirdre Lynch, Erin Kelly, Peggy McMahon, Justin Billing, Jennifer Myers, and Aimee Thorne.

Resources for Success

Student Resources

Student's Solutions Manual

Student's Solutions Manual, by Zhiwei Zhu, University of Louisiana, Lafayette

This manual provides detailed, worked-out solutions to all odd-numbered text exercises. (ISBN-10: 0-13-449736-8; ISBN-13: 978-0-13-449736-5)

Study Cards for Business Statistics Software

This series of study cards provides students with easy step-by-step instructions for using statistics software. Available for native Excel® 2016 (0-13-457679-9) Excel 2016 with XLStat™ (0-13-457683-7), Minitab 17 (0-13-457681-0), Minitab Express for PC (0-13-457685-3), Minitab Express for Mac (0-13-457691-8), JMP (0-321-64423-9), and StatCrunch™ (0-13-397513-4)

Instructor Resources

Instructor's Edition

This version of the text contains short answers to all of the exercises within the exercise sets. (ISBN-10: 0-13-449738-4; ISBN-13: 978-0-13-449738-9)

Instructor's Solutions Manual,

Instructor's Solutions Manual, by Zhiwei Zhu, University of Louisiana, Lafayette

This manual provides detailed, worked-out solutions to all of the book's exercises. The Instructor's Solutions Manual is available for download in MyStatLab and from Pearson Education's online catalog (http://www.pearsonhighered.com/irc).

Business Insight Video Assessment Questions

Written to accompany the Business Insight Videos, these video-specific questions and answers can be used for assessment or classroom discussion. These are available for download from MyStatLab or at www.pearsonhighered.com/irc.

Online Test Bank

Online Test Bank, by Paul Lorczak

This test bank contains ready-to-use quizzes and tests that correlate to chapters in the text. The testbank is available for download from Pearson Education's online catalog (http://www.pearsonhighered.com/irc).

TestGen®

TestGen® (www.pearsoned.com/testgen) enables instructors to build, edit, print, and administer tests using a computerized bank of questions developed to cover all the objectives of 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. The software and testbank are available for download from Pearson Education's online catalog (http://www.pearsonhighered.com/irc).

Technology Resources

Data Sets

Data sets formatted for Minitab, Excel, JMP, and text files can be downloaded from MyStatLab or www.pearsonhighered.com/mathstatsresources/.

Business Insight Videos

This series of ten 5- to 7- minute videos, each about a well-known business and the challenges it faces, focuses on statistical concepts as they pertain to the r eal world. The videos can be downloaded from within MyStatLab. Contact your Pearson representative for details.

MyStatLab™ Online Course (access code required)

MyStatLab from Pearson is the world's leading online resource for teaching and learning statistics; integrating interactive homework, assessment, and media in a flexible, easy-to-use format. MyStatLab is a course management system that helps individual students succeed.

- MyStatLab can be implemented successfully in any environment—lab-based, traditional, fully online, or hybrid—and demonstrates the quantifiable difference that integrated usage has on student retention, subsequent success, and overall achievement.
- MyStatLab's comprehensive 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 provides **engaging experiences** that personalize, stimulate, and measure learning for each student. In addition to the resources below, each course includes a full interactive online version of the accompanying textbook.

- **Personalized Learning:** MyStatLab's personalized homework, and adaptive and companion study plan features allow your students to work more efficiently spending time where they really need to.
- **Tutorial Exercises with Multimedia Learning Aids:** The homework and practice exercises in MyStatLab align with the exercises in the textbook, and most regenerate algorithmically to give students unlimited opportunity for practice and mastery. Exercises offer immediate helpful feedback, guided solutions, sample problems, animations, videos, statistical software tutorial videos and eText clips for extra help at point-of-use.
- **Learning Catalytics™:** MyStatLab now provides Learning Catalytics—an interactive student response tool that uses students' smartphones, tablets, or laptops to engage them in more sophisticated tasks and thinking.
- **Videos** tie statistics to the real world.
  - **StatTalk Videos:** Fun-loving statistician Andrew Vickers takes to the streets of Brooklyn, NY, to demonstrate important statistical concepts through interesting stories and real-life events. This series of 24 fun and engaging videos will help students actually understand statistical concepts. Available with an instructor's user guide and assessment questions.
  - **Business Insight Videos** [for Business Statistics only]: 10 engaging videos show managers at top companies using statistics in their everyday work. Assignable question encourage discussion.
- **Additional Question Libraries:** In addition to algorithmically regenerated questions that are aligned with your textbook, MyStatLab courses come with two additional question libraries:
  - **450 exercises** in **Getting Ready for Statistics** cover the developmental math topics students need for the course. These can be assigned as a prerequisite to other assignments, if desired.
  - **1000 exercises** in the **Conceptual Question Library** require students to apply their statistical understanding.
- **StatCrunch™:** MyStatLab integrates the 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 vibrant online community where users can access tens of thousands of shared data sets, create and conduct online surveys, perform complex analyses using the powerful statistical software, and generate compelling reports.
- **Statistical Software, Support and Integration:** We make it easy to copy our data sets, from both the eText and the MyStatLab questions, into software such as StatCrunch, Minitab®, Excel®, and more. Students have access to a variety of support tools—Technology Tutorial Videos, Technology Study Cards, and Technology Manuals for select titles—to learn how to effectively use statistical software.

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.
- Students can also access supplemental animations directly from selected exercises.
- Knowing that students often use external statistical software, we make it easy to copy our data sets, both from the eText 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.

MyStatLab Accessibility

- MyStatLab is compatible with the JAWS screen reader, and enables multiple-choice, fill-in-the-blank and free-response problem-types to be read, and interacted with via keyboard controls and math notation input. MyStatLab also works with screen enlargers, including ZoomText, MAGic®, and SuperNova. And all MyStatLab videos accompanying texts with copyright 2009 and later have closed captioning.
- More information on this functionality is available at http://mystatlab.com/accessibility.

And, MyStatLab comes from an **experienced 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 and our assessment tools work. It means we are committed to making MyStatLab as accessible as possible.
- 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.

PowerPoint Lecture Slides

PowerPoint Lecture Slides provide an outline for use in a lecture setting, presenting definitions, key concepts, and figures from the text. These slides are available within MyStatLab or at www.pearsonhighered.com/irc.

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. The vibrant online community offers tens of thousands of data sets for students to analyze.

- **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. Also, 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 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.

Minitab® 17 and Minitab Express™

Minitab 17 and Minitab Express make learning statistics easy and provide students with a skill-set that's in demand in today's data-driven workforce. Bundling Minitab® software with educational materials ensures students have access to the software they need in the classroom, around campus, and at home. And having 12-month versions of Minitab 17 and Minitab Express available ensures students can use the software for the duration of their course. (ISBN-10: 0-13-445640-8; ISBN-13: 978-0-13-445640-9)

JMP Student Edition

JMP Student Edition is an easy-to-use, streamlined version of JMP desktop statistical discovery software from SAS Institute, Inc., and is available for bundling with the text. Check with your Pearson sales representative for order information. (ISBN-10: 0-13-467979-2; ISBN-13: 978-0-13-467979-2)

XLSTAT™ for Pearson

Used by leading businesses and universities, XLSTAT is an Excel® add-in that offers a wide variety of functions to enhance the analytical capabilities of Microsoft Excel, making it the ideal tool for your everyday data analysis and statistics requirements. XLSTAT is compatible with all Excel versions. Available for bundling with the text. (ISBN-10: 0-321-75940-0; ISBN-13: 978-0-75940-5)

INDEX OF APPLICATIONS

CO = Chapter Opener; **IE** = In-Text Example; **WT** = What Do You Think?; **4M** = Motivation, Method, Mechanics, Message; **P** = Pitfalls; **BP** = Best Practices; **AE** = Analytics in Excel; **AD** = About the Data; **BTM** = Behind the Math; **TAI** = Think About It; **YDI** = You Do It; **SA** = Statistics in Action; **QT** = Questions for Thought

Company Names

Construction

Consumers

Demographics

Distribution and Operations Management

E-Commerce

Economics

Education

Energy

Environment

Finance and Investments

Food/Drink

Games

Government

Human Resource Management/Personnel

Insurance

Labor

Law

Management

Manufacturing

Marketing

Media and Entertainment

Pharmaceuticals, Medicine, and Health

Quality Control

Real Estate

Salary and Benefits

Sales and Retail

Science

Service Industries

Sports

Surveys and Opinion Polls

Technology

Transportation

PART I

Variation

CHAPTER 1 Introduction

1.1 | WHAT IS STATISTICS?

What do you think statistics is all about?

Statistics answers questions using data, or information about the situation. A **statistic** is a property of data, be it a number such as an average or a graph that displays information. Statistics—the discipline—is the science and art of extracting answers from data. Some of these answers do require putting numbers into formulas, but you can also do every statistical analysis with graphs and tables. It's hard to be fooled when the answer stares you in the face.

statistic A property of data.

Think of statistics as art. An artist must choose the right subject, and a good statistician finds the right picture. Rather than learning to paint, in this course you'll learn how to use statistics to interpret data and answer interesting questions. Learning how to use data to solve problems is central to business analytics and data science.

Which questions are interesting? The answer is simple: those you care about. Of course, what interests one person may be of no interest to someone else. In this text we apply statistics to a mix of topics, ranging from finance and marketing to personal choices. Most of the questions in this book concern business, but statistics applies more generally. We'll help you appreciate the generality of statistics by solving problems from health and science, too.

Variation

What kinds of questions can be answered with statistics? Let's start with an example. In November 2011, Barnes and Noble debuted its entry into the market for tablet computers called the Nook Tablet™ and joined the field of challengers to the Apple iPad. A question facing Barnes and Noble was, What's the right price for the Nook Tablet?

That's a hard question. To find an answer, you need to know basic economics, particularly the relationship among price, supply, and demand. A little finance and accounting determine the cost of development and production. Then come questions about the customers. Which customers are interested in a tablet computer? How much are they willing to pay? The Nook Tablet was smaller than an iPad with less technology, making it cheaper to produce. It could be sold at a profit for less than the $499 iPad, but how much less? Should it cost more than the competing $199 Kindle Fire?

Suddenly, the initial pricing question branches into several questions, and the answers depend on whom you ask. There's **variation** among customers; customers react differently. One customer might be willing to pay $300, whereas another would pay only $200. Once you recognize these differences among customers, how are you going to set *one* price? Statistics

variation Differences among individuals or items; also fluctuations over time.

shows how to use your data—what you know about your product and your customers—to set a price that will attract business and earn a profit. That you likely have not heard of the Nook tells you Barnes and Noble didn't find the right combination of price and technology.

Here's another interesting question: Why does a shopper choose a particular box of cereal? Modern grocers have become information-rich retailers, tracking every item purchased by each patron. That's why they give out personalized shopping cards; they're paying customers with discounts in return for tracking purchases. Customers keep retailers off balance because they don't buy the same things every time they shop. Did the customer buy that box of cereal because it was conveniently positioned at the end of an aisle, because he or she had a discount coupon, or simply because a six-year-old just saw a commercial while watching *Sponge Bob*? Variation makes the question harder to answer.

Patterns and Models

pattern A systematic, predictable feature in data.

statistical model A breakdown of variation into a predictable pattern and the remaining variation.

Statistics helps you answer questions by providing methods designed to handle variation. These methods filter out the clutter by revealing patterns. A **pattern** in data is a systematic, predictable feature. If customers who receive coupons typically buy more cereal than customers without coupons, there's a pattern.

Patterns form one part of a **statistical model**. A statistical model describes the variation in data as the combination of a pattern plus a background of remaining, unexplained variation. The pattern in a statistical model describes the variation that we claim to understand. The pattern tells us what we can anticipate in new data and thus goes beyond describing the data we observe. Often, an equation summarizes the pattern in a precise mathematical form. The remaining variation represents the effects of other factors we cannot explain because we lack enough information to do so. For instance, retail sales increase during holiday seasons. Retailers recognize this pattern and prepare by increasing inventories and hiring extra employees. It's impossible, though, for retailers to know exactly which items customers will want and how much they will spend. The pattern does not explain everything.

Good statistical models simplify reality to help us answer questions. Indeed, the word *model* once meant the blueprints, the plans, for a building. Plans answer some questions about the building. How large is the building? Where are the bathrooms? The blueprint isn't the building, but we can learn a lot from this model. A model of an airplane in a wind tunnel provides insights about flight even though it doesn't mimic every detail of flight. Models of data provide answers to questions even though those answers may not be entirely right. A famous statistician, George Box, once said, "All models are wrong, but some are useful."

A simple model that we understand is generally better than a complex model that we do not understand. A challenge in learning statistics is to recognize when a model can be trusted. Models based on physics and engineering often look impressively complex, but don't confuse complexity with being correct. Complex models fail when the science does not mimic reality. For example, NASA used the following elaborate equation to estimate the chance of foam insulation breaking off during take-off and damaging the space shuttle:

$$p = \frac{0.0195(L/d)^{0.45}(d)\rho_F^{0.27}(V - V^*)^{.67}}{S_T^{.25}\rho_T^{.16}}$$

The model represented by this equation failed to anticipate the risk of damage from faulty insulation. Damage from insulation caused the space shuttle *Columbia* to break apart on reentry in 2003.

Models also fail if we mistake random variation for a pattern. People are great at finding patterns. Ancient people looked into the sky and found patterns among the stars. Psychiatrists use the Rorschach ink blot test to probe deep feelings. People even find patterns in clouds, imagining shapes or faces floating in the sky. The phenomenon of perceiving something familiar in a place where it does not occur is so common that it even has a name: pareidolia.[1] A key task in statistics is deciding whether the pattern we have discovered is real or something that we've imagined. Finding a pattern allows us to anticipate what is most likely to happen next, to understand the data in order to plan for the future and make better decisions. But if we imagine a pattern when there is none, we become overconfident and make poor decisions.

1.2 | PREVIEWS

The following two examples preview the use of statistics to answer questions. Movie theaters show previews of coming attractions with lots of action and explosions, and save the character development for later. These examples introduce recurring themes and showcase several methods that are fully developed in later chapters. The point is to advertise the types of analyses you will be able to do after you finish this book.

Each example begins with a question motivated by a story in the news, and then uses a statistical model to formulate an answer to the question. The first example uses a model to predict the future, and the second uses a model to fill in for an absence of data. These are previews, so we emphasize the results and skip the details.

Predicting Employment

In early November 2005, national broadcasts announced surprising and disturbing economic results. The big story was not a recession, but rather the U.S. economy's slower than expected growth. The Labor Department reported that only 56,000 jobs had been created in October 2005, far short of the 100,000 additional jobs expected by Wall Street forecasters.

Financial markets react to surprises. If everyone on Wall Street expects the Labor Department to report large numbers of new jobs, the stock market can tumble if the report announces only modest growth. What should we have expected? What made Wall Street economists expect 100,000 jobs to be created in October? Surely they didn't expect *exactly* 100,000 jobs to be created. Was the modest growth a fluke? These are serious questions. If the shortfall in jobs is the start of a downward trend, it could indicate the start of an economic recession. Businesses need to anticipate where the economy is headed in order to schedule production and supplies.

Was the weather responsible for the modest growth? On August 29, 2005, Hurricane Katrina slammed into Louisiana, devastating the Gulf Coast (see Figure 1.1). Packing sustained winds of 175 miles per hour, Katrina overwhelmed levees in New Orleans, flooded the city, and wrecked the local economy. Estimates of damages reached $130 billion, the highest ever attributed to a hurricane in the United States, with more than 1,000 deaths. Katrina and the hurricanes that followed during the 2005 season devastated the oil industry concentrated along the Gulf of Mexico and disrupted energy supplies around the country. Did Katrina wipe out the missing jobs?

Let's see if we can build our own forecast. Back in September 2005, how could you forecast employment in October?

[1]The *New York Times* article "Is That Jesus in Your Toast" (April 4, 2014) discusses the underlying psychology and offers several more examples.

FIGURE 1.1 Hurricane Katrina on August 29, 2005.

We need two things to get started: relevant data and a method for using these data to address the question at hand. Virtually every statistical analysis proceeds in this way. Let's start with data. At a minimum, we need the number employed before October. For example, if the number of jobs had been steady from January through the summer of 2005, our task would be easy; it's easy to forecast something that doesn't change.

The problem is that employment does change. Table 1.1 shows the number of thousands employed each month since 2003. These are the data behind the story.

TABLE 1.1 Nonfarm employment in the United States, in thousands on payrolls.

	2003	2004	2005
Jan	130,247	130,372	132,573
Feb	130,125	130,466	132,873
Mar	129,907	130,786	132,995
Apr	129,853	131,123	133,287
May	129,827	131,373	133,413
Jun	129,854	131,479	133,588
Jul	129,857	131,562	133,865
Aug	129,859	131,750	134,013
Sep	129,953	131,880	134,005
Oct	130,076	132,162	134,061
Nov	130,172	132,294	
Dec	130,255	132,449	

Each column gives the monthly counts for a year. The first number in the table represents 130,247,000 jobs on payrolls in January 2003. The following number shows that payrolls in February 2003 fell by 122,000. At the bottom of the third column, we can see that employment increased by 56,000 from September to October 2005, as reported by Reuters. This variation complicates the task of forecasting. We've got to figure out how we expect employment to change next month.

We won't replicate the elaborate models used by Wall Street economists, but we can go a long way toward understanding their models by plotting the data. Plots are among the most important tools of statistics. Once we see the plot, we can decide how to make a forecast.

timeplot A chart of values ordered in time, usually with the values along the y-axis and time along the x-axis.

The graph in Figure 1.2 charts employment over time, a common type of display that we'll call a **timeplot**. To keep the vertical axis nicely scaled and avoid showing extraneous digits, we labeled the employment counts in millions

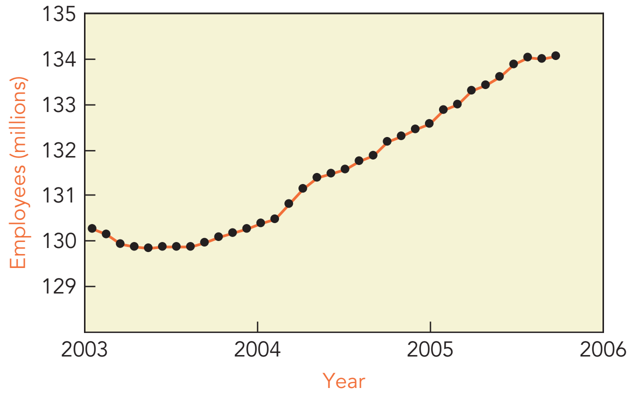


FIGURE 1.2 Timeplot of employment.

rather than thousands. **tip** When displaying numerical information, showing fewer digits often produces a better presentation.

Once you've seen this timeplot, you can probably come up with your own forecast. Employment grew steadily during 2004 into 2005. This steady growth is a pattern. Employment varies from month to month, but the upward trend suggests a way to extrapolate the pattern into the future. The line drawn in the next figure summarizes the pattern we see in the data and suggests a forecast for October.

Models help us simplify complicated phenomena, but that does not make them correct. Good models convey the possibility of an error. The data on employment follow a line after 2004, but that does not mean that this pattern will continue. It's not as though the economy knows about the line in Figure 1.3 and must stick to this trend. This line allows us to anticipate where employment is headed.

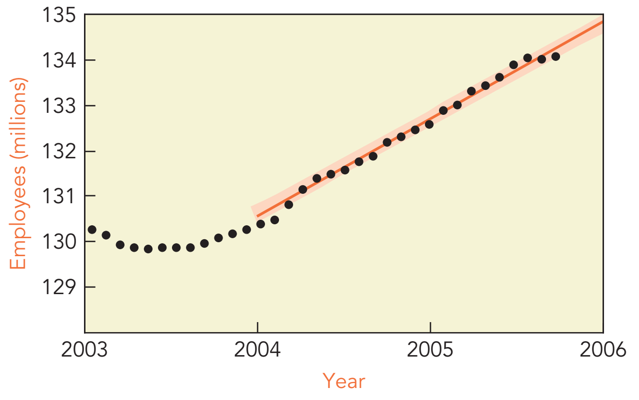


FIGURE 1.3 Linear pattern and region of uncertainty.

Reality may not match our forecast, so we use ranges to convey the uncertainty associated with using a statistical model. Ranges are used in many fields to indicate uncertainty. For example, Figure 1.4 shows the uncertainty surrounding the projected path for Katrina before it made landfall. The wider the cone, the more doubt about the path of the storm.

Similarly, the shaded region around the line in Figure 1.3 indicates how closely the historic data stick to this pattern. In addition, this region suggests the amount of unexplained variation around the pattern. The more closely the data follow the pattern, the narrower this region becomes. The employment data track the line very closely during the period summarized by the line.

FIGURE 1.4 Projected path for Katrina.

To get a forecast, we extend the pattern. It is easy to extend the line beyond the data, as in Figure 1.3. The line passes above the count for October. This model predicts employment to be near 134,150,000 in October 2005, about 90,000 more than the reported count. That's close to the value claimed in the news reports mentioned at the beginning of this subsection. We can also extend the region of uncertainty around the line. We should not expect counts of employment in the future to be closer to the line than those in the past. The line—our pattern—forecasts employment to be near that predicted by Wall Street economists. On the basis of this pattern, we would forecast employment in October to lie between 134,130,000 and 134,533,000 jobs.

Our simple model confirms that the level of employment in October 2005 *is* surprising. Not only is the reported count of 134,061,000 for October less than expected, but it's also outside the anticipated range. Even allowing for variation around the pattern, employment is smaller than expected. Something happened in October that reduced the number of jobs. This is a large break from the pattern and demands our attention. Do we know *why* the employment was less than expected?

Anticipating the impact of weather on employment is an ongoing concern. In the fall of 2008, Hurricanes Ike and Gustav were blamed for putting 50,000 Americans out of work, and the Department of Commerce estimated that Hurricane Sandy put 65,000 out of work in New Jersey and New York in 2012. The only way to arrive at these estimates is to predict what would have happened had these hurricanes not struck. That's a job for a statistical model.

Pricing a Car

For our second preview, let's talk about cars. You may not be interested in cars, but, even so, you have almost certainly heard of the automaker BMW.

Writers for *Car and Driver* magazine love cars made by BMW, particularly the popular sporty sedans that go by numeric names like the 328 or the 335. According to these journalists, competitors have yet to figure out a way to beat the BMW 3-series. Competitors keep trying to knock BMW off the top of the pedestal but keep falling short in comparison. Year after year, *Car and Driver* has included BMW 3-series models in its "top 10" list of best cars of the year.

What's it going to cost to get behind the wheel of a BMW 3-series? The manufacturer's suggested retail price for a basic 2016 BMW 328i is $38,350, excluding options like leather seats that add $1,500 to the bottom line. That's a lot to spend, so let's see what a used BMW costs. For example, a search on the Web turned up a 2013 BMW 328i advertised for $27,000. The car has 40,000 miles, an automatic transmission, and a variety of options. On one hand, it sounds like a lot to spend $27,000 on a used car. On the other hand, its price is $11,350 less than the cost of a stripped-down new car that lacks options like the heated leather seats found on this car.

Companies face similar decisions: Should a company buy new equipment, or should it look for a deal on used substitutes? As for cars, auto dealerships face these questions every day. What should the dealership charge for a three-year lease? The ultimate resale value has a big effect. If, when the car comes back, the dealership can sell it for a good price, then it can offer a better deal on a lease.

Once again, we need to identify data and decide how to use them. To get data on the resale value of used BMWs, students in a class downloaded prices for used cars offered by BMW dealers. They gathered information about 153 cars in the 328-series, from the 2011 through the 2016 model years. Table 1.2 summarizes the prices of these cars, broken out by model year. Eighty-four of these cars are 2013 models, like the car in the online ad. On average, these sell for $29,784. That's $2,784 more than the car in the ad.

TABLE 1.2 Average prices vary by year.

Model Year	Number	Average Price
2011	16	$22,967
2012	3	$25,993
2013	84	$29,784
2014	15	$34,409
2015	34	$38,385
2016	1	$38,999

Before we decide that the used car in the ad is a bargain, we ought to consider other things that affect the price. You can probably name a few. We've taken into account the age (2013 model year) and style (328), but not the mileage. We'd guess you'd be willing to pay more for a car with 10,000 miles than an otherwise similar car with 100,000 miles.

scatterplot A graph that shows pairs of values (x, y) laid out on a two-dimensional grid.

The **scatterplot** in Figure 1.5 shows that mileage is related to price. Each point shows the mileage and price of one of these 84 cars. None of these cars has exactly the same mileage as the advertised car. We can use a statistical model to compensate for our lack of cars with the same mileage. The plot in Figure 1.6 shows a line that relates mileage to price. The plot includes the region of uncertainty around the line. The region of uncertainty is wide because other factors, such as the condition of the car and its options, affect the price. There's much more variation around this pattern than around the line in Figure 1.3.

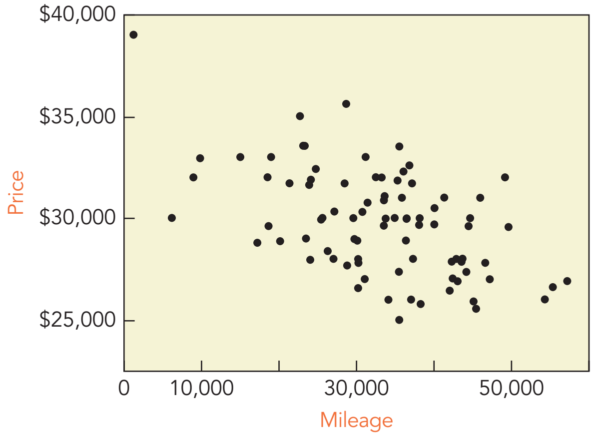


FIGURE 1.5 Scatterplot of price versus mileage.

In our first example, we extrapolated a pattern in historic data. In this example, the pattern serves a different purpose. The line in Figure 1.6 allows us to

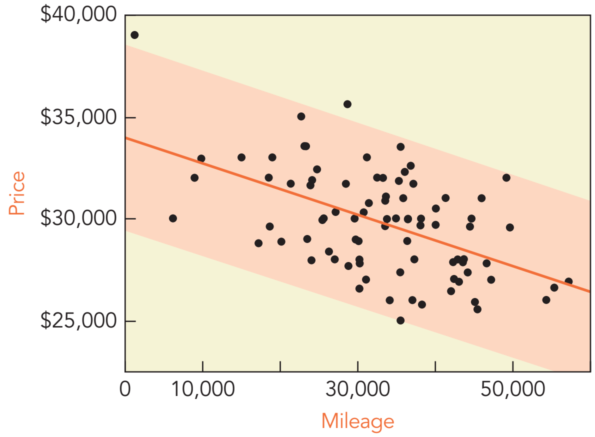


FIGURE 1.6 A line relates mileage to price for these cars.

borrowing strength The use of a model to allow all of the data to answer a specific question.

borrow strength. **Borrowing strength** refers to using a model to glean more from data for answering a question. Rather than search for used cars with exactly 40,000 miles, this model—the line—allows us to estimate the price of a car with 40,000 miles, even though we haven't seen a car with exactly this mileage. Even though the mileage varies among these cars, the pattern allows us to "borrow" some of the information about each car to estimate the price of a car with specific mileage. The estimated value, reading off the height of the line, is about $28,900. As an added bonus, the negative slope of the line in Figure 1.6 shows how higher mileage reduces the value of a car. On average, these cars lose about 13 cents of value per mile.

Having seen this analysis, do you think the car in the classified ad is a bargain? Its $27,000 price is less than the predicted price from this model, but well within the range for cars of this age and mileage. It might be worth a further look, particularly if you like the options that it includes!

The following chapters contain many more examples like these that illustrate the use of statistics in business. The only difference is that you will be doing the analysis, making the choices, and interpreting the results. The first step in learning these methods is to appreciate the importance of data, and that is the subject of the next chapter.