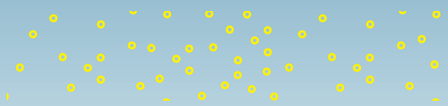
Applied Statistics in Business and Economics

seventh edition

DAVID P. DOANE
LORI E. SEWARD

Mc Graw Hill



Applied Statistics

in Business and Economics

Seventh Edition

The McGraw Hill/Irwin Series in Operations and Decision Sciences

SUPPLY CHAIN MANAGEMENT

Bowersox, Closs, Cooper, and Bowersox
Supply Chain Logistics Management
Fifth Edition

Johnson
Purchasing and Supply Management
Sixteenth Edition

Simchi-Levi, Kaminsky, and Simchi-Levi
Designing and Managing the Supply Chain: Concepts, Strategies, Case Studies
Fourth Edition

PROJECT MANAGEMENT

Larson and Gray
Project Management: The Managerial Process
Eighth Edition

SERVICE OPERATIONS MANAGEMENT

Bordoloi, Fitzsimmons, and Fitzsimmons
Service Management: Operations, Strategy, Information Technology
Ninth Edition

MANAGEMENT SCIENCE

Hillier and Hillier
Introduction to Management Science: A Modeling and Case Studies Approach with Spreadsheets
Sixth Edition

BUSINESS RESEARCH METHODS

Schindler
Business Research Methods
Fourteenth Edition

BUSINESS FORECASTING

Keating and Wilson
Business Forecasting
Seventh Edition

BUSINESS SYSTEMS DYNAMICS

Sterman
Business Dynamics: Systems Thinking and Modeling for a Complex World
First Edition

OPERATIONS MANAGEMENT

Cachon and Terwiesch
Operations Management
Second Edition

Cachon and Terwiesch
Matching Supply with Demand: An Introduction to Operations Management
Fourth Edition

Jacobs and Chase
Operations and Supply Chain Management
Sixteenth Edition

Jacobs and Chase
Operations and Supply Chain Management: The Core
Fifth Edition

Schroeder and Goldstein
Operations Management in the Supply Chain: Decisions and Cases
Eighth Edition

Stevenson
Operations Management
Thirteenth Edition

Swink, Melnyk, and Hartley
Managing Operations Across the Supply Chain
Fourth Edition

BUSINESS MATH

Slater and Wittry
Practical Business Math Procedures
Thirteenth Edition

Slater and Wittry
Math for Business and Finance: An Algebraic Approach
Second Edition

BUSINESS STATISTICS

Bowerman, Drougas, Duckworth, Froelich, Hummel, Moninger, and Schur
Business Statistics in Practice
Ninth Edition

Doane and Seward
Applied Statistics in Business and Economics
Seventh Edition

Doane and Seward
Essential Statistics in Business and Economics
Third Edition

Lind, Marchal, and Wathen
Basic Statistics for Business and Economics
Tenth Edition

Lind, Marchal, and Wathen
Statistical Techniques in Business and Economics
Eighteenth Edition

Jaggia and Kelly
Business Statistics: Communicating with Numbers
Fourth Edition

Jaggia and Kelly
Essentials of Business Statistics: Communicating with Numbers
Second Edition

BUSINESS ANALYTICS

Jaggia, Kelly, Lertwachara, and Chen
Business Analytics: Communicating with Numbers
First Edition

Applied Statistics

in Business and Economics

Seventh Edition

David P. Doane
Oakland University

Lori E. Seward
University of Colorado

Mc Graw Hill

APPLIED STATISTICS IN BUSINESS AND ECONOMICS, SEVENTH EDITION

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 LWI 24 23 22 21

ISBN 978-1-260-71628-3 (bound edition)
MHID 1-260-71628-7 (bound edition)
ISBN 978-1-264-09856-9 (loose-leaf edition)
MHID 1-264-09856-1 (loose-leaf edition)

Portfolio Manager: *Noelle Bathurst*
Product Developers: *Ryan McAndrews*
Marketing Manager: *Harper Christopher*
Content Project Manager: *Amy Gehl /Jamie Koch*
Buyer: *Susan K. Culbertson*
Designer: *Matt Diamond*
Content Licensing Specialists: *Shawntel Schmitt*
Cover Image: *Xuanyu Han/Moment/Getty Images*
Compositor: *SPi Global*

All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

Library of Congress Cataloging-in-Publication Data

Names: Doane, David P., author. | Seward, Lori Welte, 1962- author.
Title: Applied statistics : in business and economics / David P. Doane, Oakland University, Lori E. Seward, University of Colorado.
Description: Seventh Edition. | New York : McGraw Hill LLC, 2021. | Revised edition of the authors' Applied statistics in business and economics, [2019] | Audience: Ages 18+
Identifiers: LCCN 2020036831 (print) | LCCN 2020036832 (ebook) | ISBN 9781260716283 (paperback) | ISBN 9781264098576 (epub)
Subjects: LCSH: Commercial statistics. | Management–Statistical methods. | Economics–Statistical methods. | Statistics.
Classification: LCC HF1017 .D55 2021 (print) | LCC HF1017 (ebook) | DDC 519.5–dc23
LC record available at https://lccn.loc.gov/2020036831
LC ebook record available at https://lccn.loc.gov/2020036832

mheducation.com/highered

About the Authors

Courtesy of David Doane

David P. Doane

David P. Doane is accredited by the American Statistical Association as a Professional Statistician (PStat®) . He is professor emeritus in Oakland University's Department of Decision and Information Sciences. He earned his Bachelor of Arts degree in mathematics and economics at the University of Kansas and his PhD from Purdue University's Krannert Graduate School. His research and teaching interests include applied statistics, forecasting, and statistical education. He is co-recipient of three National Science Foundation grants to develop software to teach statistics and to create a computer classroom. He is a longtime member of the American Statistical Association, serving in 2002 as president of the Detroit ASA. He has consulted with governments, health care organizations, and local firms. He has published articles in many academic journals. He currently belongs to ASA chapters in San Diego and Orange County/Long Beach.

Courtesy of Lori Seward

Lori E. Seward

Lori E. Seward is a teaching professor in The Leeds School of Business at the University of Colorado in Boulder. She earned her Bachelor of Science and Master of Science degrees in Industrial Engineering at Virginia Tech. After several years working as a reliability and quality engineer in the paper and automotive industries, she earned her PhD from Virginia Tech and joined the faculty at The Leeds School in 1998. Professor Seward has served as the faculty director of Leeds' MBA programs since 2017. She currently teaches as well as coordinates the core statistics course for the Leeds full-time, Professional, and Executive MBA programs. She served as the chair of the INFORMS Teachers' Workshop for the annual 2004 meeting. Her teaching interests focus on developing pedagogy that uses technology to create a collaborative learning environment in large undergraduate and MBA statistics courses. Her most recent article, co-authored with David Doane, was published in the *Journal of Statistics Education* (2011).

Dedication

To Robert Hamilton Doane-Solomon

David

To all my students who challenged me to make statistics relevant to their lives.

Lori

From the Authors

> "How often have you heard people/students say about a particular subject, 'I'll never use this in the real world'? I thought statistics was a bit on the 'math-geeky' side at first. Imagine my horror when I saw α, R^2, and correlations on several financial reports at my current job (an intern position at a financial services company). I realized then that I had better try to understand some of this stuff."
>
> *—Jill Odette (an introductory statistics student)*

As recently as a decade ago our students used to ask us, "**How** do I use statistics?" Today we more often hear, "**Why** should I use statistics?" *Applied Statistics in Business and Economics* has attempted to provide real meaning to the use of statistics in our world by using real business situations and real data and appealing to your need to know *why* rather than just *how.*

With over 50 years of teaching statistics between the two of us, we feel we have something to offer. Seeing how students have changed as the new century unfolds has required us to adapt and seek out better ways of instruction. So we wrote *Applied Statistics in Business and Economics* to meet four distinct objectives.

Objective 1: Communicate the Meaning of Variation in a Business Context Variation exists everywhere in the world around us. Successful businesses know how to measure variation. They also know how to tell when variation should be responded to and when it should be left alone. We'll show how businesses do this.

Objective 2: Use Real Data and Real Business Applications Examples, case studies, and problems are taken from published research or real applications whenever possible. Hypothetical data are used when it seems the best way to illustrate a concept.

Objective 3: Incorporate Current Statistical Practices and Offer Practical Advice With the increased reliance on computers, statistics practitioners have changed the way they use statistical tools. We'll show the current practices and explain why they are used the way they are. We also will tell you when each technique should *not* be used.

Objective 4: Provide More In-Depth Explanation of the Why and Let the Software Take Care of the How It is critical to understand the importance of communicating with data. Today's computer capabilities make it much easier to summarize and display data than ever before. We demonstrate easily mastered software techniques using the common software available. We also spend a great deal of time on the idea that there are risks in decision making and those risks should be quantified and directly considered in every business decision.

Our experience tells us that students want to be given credit for the experience they bring to the college classroom. We have tried to honor this by choosing examples and exercises set in situations that will draw on students' already vast knowledge of the world and knowledge gained from other classes. Emphasis is on thinking about data, choosing appropriate analytic tools, using computers effectively, and recognizing limitations of statistics.

What's New in This Seventh Edition?

In this edition we have listened to you and have made many changes that you asked for. We sought advice from students and faculty who are currently using the textbook and reviewers at a variety of colleges and universities. At the end of this preface is a detailed list of chapter-by-chapter improvements, but here are just a few of them:

- New, brighter overall design (format, colors).
- Continued strong focus on Excel and business applications.
- New *Analytics in Action* briefings on careers in data analytics and exciting applications of big data, artificial intelligence, and machine learning (including ethical issues).
- New Appendix J with side-by-side comparison of statistics functions in Excel and R.
- New Appendix K with an easy walk-through to get started with R and RStudio.

- Chapter-end *Software Supplements* showing how to use R for applications in that chapter.
- Updated exercises with emphasis on compatibility with Connect®.
- Updated test bank questions matched with topics and learning objectives.
- New and updated Mini Cases for economics and business.
- New and updated exercise data sets, web links, *Big Data Sets,* and *Related Reading.*
- Many new guided examples on Connect. Students can watch 90 guided examples to aid their learning.
- Connect® supplements including *LearningStat*s demonstrations, illustrations of R calculations for common tasks, and video tutorials (both PC and Mac).

Software

Excel is used throughout this book because it is available everywhere. Some calculations are illustrated using MegaStat and Minitab because they offer more capability than Excel's Data Analysis Tools. In recognition of growing interest in analytics training beyond Excel, our textbook now provides an optional introduction to R with illustrations of topics in each chapter. Our support for R is further enhanced with *LearningStats* modules, tables of R functions, and R-compatible Excel data sets. To further assist students we provide Connect® tutorials or demonstrations on using Excel, Minitab, MegaStat, and R. At the end of each chapter is a list of *LearningStats* demonstrations that illustrate the concepts from the chapter.

Math Level

The assumed level of mathematics is pre-calculus, though there are rare references to calculus where it might help the better-trained reader. All but the simplest proofs and derivations are omitted, though key assumptions are stated clearly. The learner is advised what to do when these assumptions are not fulfilled. Worked examples are included for basic calculations, but the textbook does assume that computers will do the calculations after the statistics class is over, so *interpretation* is paramount. End-of-chapter references and suggested websites are given so that interested readers can deepen their understanding.

Exercises

Simple practice exercises are placed within each section. End-of-chapter exercises tend to be more integrative or to be embedded in more realistic contexts. Attention has been given to revising exercises so that they have clear-cut answers that are matched to specific learning objectives. A few exercises invite short answers rather than just quoting a formula. Answers to odd-numbered exercises are in the back of the book (all of the answers are in the instructor's manual).

LearningStats

Connect users can access *LearningStats,* a collection of Excel spreadsheets, Word documents, and PowerPoints for each chapter. It is intended to let students explore data and concepts at their own pace, ignoring material they already know and focusing on things that interest them. *LearningStats* includes deeper explanations on topics such as how to write effective reports, how to perform calculations, or how to make effective charts. It also includes topics that did not appear prominently in the textbook (e.g., partial F test, Durbin–Watson test, sign test, bootstrap simulation, and logistic regression). Instructors can use *LearningStats* PowerPoint presentations in the classroom, but Connect users also can use them for self-instruction. No instructor can "cover everything," but students can be encouraged to explore *LearningStats* data sets and/or demonstrations, perhaps with an instructor's guidance.

David P. Doane
Lori E. Seward

How Are Chapters Organized

Chapter Contents

Each chapter begins with a short list of section topics that are covered in the chapter.

CHAPTER CONTENTS

Chapter Learning Objectives

Each chapter includes a list of learning objectives students should be able to attain upon reading and studying the chapter material. Learning objectives give students an overview of what is expected and identify the goals for learning. Learning objectives also appear next to chapter topics in the margins.

CHAPTER LEARNING OBJECTIVES

When you finish this chapter, you should be able to

- **LO 1-1** Define statistics and explain some of its uses.
- **LO 1-2** List reasons for a business student to study statistics.
- **LO 1-3** Explain the uses of statistics in business.
- **LO 1-4** State the common challenges facing business professionals using statistics.
- **LO 1-5** List and explain common statistical pitfalls.

Section Exercises

Multiple section exercises are found throughout the chapter so that students can focus on material just learned.

Section Exercises

connect

3.1 (a) Make a stem-and-leaf plot for these 24 observations on the number of customers who used a downtown Citibank ATM during the noon hour on 24 consecutive workdays. (b) Make a dot plot of the ATM data. (c) Describe these two displays. (*Hint:* Refer to center, variability, and shape.) Citibank

39	32	21	26	19	27	32	25
18	26	34	18	31	35	21	33
33	9	16	32	35	42	15	24

3.2 (a) Make a stem-and-leaf plot for the number of defects per 100 vehicles for these 32 brands. (b) Make a dot plot of the defects data. (c) Describe these two displays. (*Hint:* Refer to center, variability, and shape.)

Mini Cases

Every chapter includes two or three mini cases, which are solved applications. They show and illlustrate the analytical application of specific statistical concepts at a deeper level than the examples.

Mini Case 4.8

What Is the DJIA? DJIA

The Dow Jones Industrial Average (commonly called the DJIA) is the oldest U.S. stock market price index, based on the prices of 30 large, widely held, and actively traded "blue chip" public companies in the United States (e.g., Coca-Cola, Microsoft, Walmart, Walt Disney). Actually, only a few of its 30 component companies are "industrial." The DJIA is not a simple average but rather a *weighted average* based on prices of its component stocks. Originally a simple mean of stock prices, the DJIA now is calculated as the sum of the 30 stock prices divided by a "divisor" that compensates for stock splits and other changes over time. The divisor is revised as often as necessary (see *The Wall Street Journal* for the latest divisor value). Because high-priced stocks comprise a larger proportion of the sum, the DJIA is more strongly affected by changes in high-priced stocks. That is, a 10 percent price increase in a $10 stock would have less effect than a 10 percent price increase in a $50 stock, even if both companies have the same total market capitalization (the total number of shares times the price per share; often referred to as "market cap"). Broad-based market price indexes (e.g., NSDQ, AMEX, NYSE, S&P 500, Russ 2K) are widely used by fund managers, but the venerable "Dow" is still the one you see first on CNN or MSNBC.

Analytics in Action

These NEW features bring in real-world examples to illustrate data analytics in action.

Analytics in Action

Walmart, Big Data, and Retail Analytics

Walmart processes over a million customer transactions each hour, which translates into two to three petabytes of data each hour. (A petabyte is a million gigabytes!) What to do with all this data? Walmart is practicing *data democratization*. This term means making large amounts of data available to everyone in the organization so that employees can quickly react to changes in their customers' behaviors.

Walmart operates a Data Café that can be accessed by everyone in the company. When store managers notice changes in sales for particular products, they can go to the Data Café and look at data across all their stores to figure out why the changes

to Promote Student Learning?

Figures and Tables

Throughout the text, there are hundreds of charts, graphs, tables, and spreadsheets to illustrate statistical concepts being applied. These visuals help stimulate student interest and clarify the text explanations.

Figure 3.1

Dot Plot of Business Cycle Duration ($n = 33$)

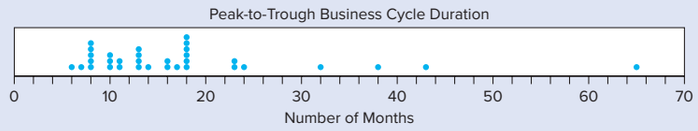


Table 2.5

Random Sampling Methods

Simple Random Sample	Use random numbers to select items from a list (e.g., Visa cardholders).
Systematic Sample	Select every *k*th item from a list or sequence (e.g., restaurant customers).
Stratified Sample	Select randomly within defined strata (e.g., by age, occupation, gender).
Cluster Sample	Select random geographical regions (e.g., zip codes) that represent the population.

Examples

Examples of interest to students are taken from published research or real applications to illustrate the statistics concept. For the most part, examples are focused on business, but there are also some that are more general and don't require any prerequisite knowledge. And there are some that are based on student projects.

Example 3.3

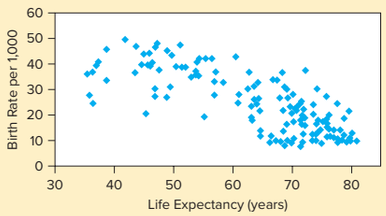
Birth Rates and Life Expectancy

Source: *The World Factbook 2003.* Central Intelligence Agency, 2003. www.cia.gov.

Figure 3.16 shows a scatter plot with life expectancy on the *X*-axis and birth rates on the *Y*-axis. In this illustration, there seems to be an association between *X* and *Y*. That is, nations with higher birth rates tend to have lower life expectancy (and vice versa). No cause-and-effect relationship is implied because, in this example, both variables could be influenced by a third variable that is not mentioned (e.g., GDP per capita).

Figure 3.16

Scatter Plot of Birth Rates and Life Expectancy ($n = 153$ nations) BirthLife



Data Set Icon

A data set icon is used throughout the text to identify data sets used in the figures, examples, and exercises that are included in Connect for the text.

USTrade

How Does This Text Reinforce

Chapter Summary

Chapter summaries provide an overview of the material covered in the chapter.

Chapter Summary

The **mean** and **median** describe a sample's **center** and also indicate **skewness.** The **mode** is useful for discrete data with a small range. The **trimmed mean** eliminates extreme values. The **geometric mean** mitigates high extremes but cannot be used when zeros or negative values are present. The **midrange** is easy to calculate but is sensitive to extremes. Variability is typically measured by the **standard deviation,** while relative dispersion is given by the **coefficient of variation** for nonnegative data. **Standardized data** reveal **outliers** or unusual data values, and the **Empirical Rule** offers a comparison with a normal distribution. In measuring dispersion, the **mean absolute deviation** or **MAD** is easy to understand but lacks nice mathematical properties. **Quartiles** are meaningful even for fairly small data sets, while **percentiles** are used only for large data sets. **Box plots** show the quartiles and data range. The **correlation coefficient** measures the degree of linearity between two variables. The **covariance** measures the degree to which two variables move together. We can estimate many common descriptive statistics from **grouped data.** Sample coefficients of **skewness** and **kurtosis** allow more precise inferences about the **shape** of the population being sampled instead of relying on histograms.

Key Terms

Key terms are highlighted and defined within the text. They are also listed at the ends of chapters to aid in reviewing.

Key Terms

Center	*Variability*	*Shape*	*Other*
geometric mean	Chebyshev's Theorem	bimodal distribution	box plot
mean	coefficient of variation	kurtosis	covariance
median	Empirical Rule	kurtosis coefficient	five-number summary
midhinge	mean absolute deviation	leptokurtic	interquartile range
midrange	outliers	mesokurtic	method of medians
mode	population variance	multimodal distribution	quartiles
trimmed mean	range	negatively skewed	sample correlation coefficient
weighted mean	sample variance	Pearson 2 skewness coefficient	
	standard deviation	platykurtic	
	standardized data	positively skewed	
	two-sum formula	Schield's Rule	
	z-score	skewed left	
		skewed right	
		skewness	
		skewness coefficient	
		symmetric data	

Commonly Used Formulas

Some chapters provide a listing of commonly used formulas for the topic under discussion.

Commonly Used Formulas in Descriptive Statistics

Sample mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Geometric mean: $G = \sqrt[n]{x_1 x_2 \ldots x_n}$

Growth rate: $GR = \sqrt[n-1]{\frac{x_n}{x_1}} - 1$

Range: $\text{Range} = x_{\max} - x_{\min}$

Midrange: $\text{Midrange} = \frac{x_{\max} + x_{\min}}{2}$

Sample standard deviation: $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$

Chapter Review

Each chapter has a list of questions for student self-review or for discussion.

Chapter Review

1. What are descriptive statistics? How do they differ from visual displays of data?
2. Explain each concept: (a) center, (b) variability, and (c) shape.
3. (a) Why is sorting usually the first step in data analysis? (b) Why is it useful to begin a data analysis by thinking about how the data were collected?
4. List strengths and weaknesses of each measure of center and write its Excel function: (a) mean, (b) median, and (c) mode.
5. (a) Why must the deviations around the mean sum to zero? (b) What is the position of the median in the data array when n is even? When n is odd? (c) Why is the mode of little use in continuous data? (d) For what type of data is the mode most useful?
6. (a) What is a bimodal distribution? (b) Explain two ways to detect skewness.
7. List strengths and weaknesses of each measure of center and give its Excel function (if any): (a) midrange, (b) geometric mean, and (c) 10 percent trimmed mean.
8. (a) What is variability? (b) Name five measures of variability. List the main characteristics (strengths, weaknesses) of each measure.
9. (a) Which standard deviation formula (population, sample) is used most often? Why? (b) When is the coefficient of variation useful?
10. (a) To what kind of data does Chebyshev's Theorem apply? (b) To what kind of data does the Empirical Rule apply? (c) What is an outlier? An unusual data value?
11. (a) In a normal distribution, approximately what percent of observations are within 1, 2, and 3 standard deviations of the mean? (b) In a sample of 10,000 observations, about how many observations would you expect beyond 3 standard deviations of the mean?

Student Learning?

Chapter Exercises

Exercises give students an opportunity to test their understanding of the chapter material. Exercises are included at the ends of sections and at the ends of chapters. Some exercises contain data sets, identified by data set icons. Data sets can be accessed through Connect and used to solve problems in the text.

Chapter Exercises

4.54 (a) For each data set, calculate the mean, median, and mode. (b) Which, if any, of these three measures is the weakest indicator of a "typical" data value? Why?

a. Number of e-mail accounts (12 students): 1, 1, 1, 1, 2, 2, 2, 3, 3, 3, 3, 3
b. Number of siblings (5 students): 0, 1, 2, 2, 10
c. Asset turnover ratio (8 retail firms): 1.85, 1.87, 2.02, 2.05, 2.11, 2.18, 2.29, 3.01

4.55 If the mean asset turnover for retail firms is 2.02 with a standard deviation of 0.22, without assuming a normal distribution, within what range will at least 75% of retail firms' asset turnover fall?

4.56 For each data set: (a) Find the mean, median, and mode. (b) Which, if any, of these three measures is the weakest indicator of a "typical" data value? Why?

a. 100 m dash times ($n = 6$ top runners): 9.87, 9.98, 10.02, 10.15, 10.36, 10.36
b. Number of children ($n = 13$ families): 0, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 6

Find Bob's standardized z-score. (b) By the Empirical Rule, is Bob's SAT score unusual?

4.61 Find the data value that corresponds to each of the following z-scores.

a. Final exam scores: Allison's z-score = 2.30, $\mu = 74$, $\sigma = 7$
b. Weekly grocery bill: James' z-score = −1.45, $\mu = \$53$, $\sigma = \$12$
c. Daily video game play time: Eric's z-score = −0.79, $\mu = 4.00$ hours, $\sigma = 1.15$ hours

4.62 The average time a Boulder High varsity lacrosse player plays in a game is 30 minutes with a standard deviation of 7 minutes. Nolan's playing time in last week's game against Fairview was 48 minutes. (a) Calculate the z-score for Nolan's playing time against Fairview. (b) By the Empirical Rule, was Nolan's playing time *unusual* when compared to the typical playing time?

4.63 The number of blueberries in a blueberry muffin baked by EarthHarvest Bakeries can range from 18 to 30 blueberries. (a) Use the Empirical Rule to estimate the standard deviation of the number of blueberries in a muffin. (b) What

More Learning Resources

LearningStats provides a means for Connect users to explore data and concepts at their own pace. Applications that relate to the material in the chapter are identified by topic at the end of each chapter.

CHAPTER 4 More Learning Resources

You can access these LearningStats demonstrations through McGraw-Hill's Connect® to help you understand descriptive statistics.

Topic	*LearningStats Demonstrations*
Overview	Describing Data Using MegaStat Using Minitab Using R
Descriptive statistics	Basic Statistic Quartiles Box Plot Simulation Grouped Data Significant Digits
ScreenCam Tutorials	Using MegaStat Excel Descriptive Statistics Excel Scatter Plots

Key: = PowerPoint = Excel = PDF = ScreenCam Tutorials

Exam Review Questions

At the end of a group of chapters, students can review the material they covered in those chapters. This provides them with an opportunity to test themselves on their grasp of the material.

Exam Review Questions for Chapters 1–4

1. Which type of statistic (descriptive, inferential) is each of the following?
   a. Estimating the default rate on all U.S. mortgages from a random sample of 500 loans.
   b. Reporting the percent of students in your statistics class who use Verizon.
   c. Using a sample of 50 iPhones to predict the average battery life in typical usage.
2. Which is *not* an ethical obligation of a statistician? Explain.
   a. To know and follow accepted procedures.
   b. To ensure data integrity and accurate calculations.
   c. To support client wishes in drawing conclusions from the data.
3. "Driving without a seat belt is not risky. I've done it for 25 years without an accident." This *best* illustrates which fallacy?
   a. Unconscious bias.
   b. Conclusion from a small sample.
4. Which data type (categorical, numerical) is each of the following?
   a. Your current credit card balance.
   b. Your college major.
   c. Your car's odometer mileage reading today.
5. Give the type of measurement (nominal, ordinal, interval, ratio) for each variable.
   a. Length of time required for a randomly chosen vehicle to cross a toll bridge.
   b. Student's ranking of five cell phone service providers.
   c. The type of charge card used by a customer (Visa, Mastercard, AmEx, Other).
6. Tell if each variable is continuous or discrete.
   a. Tonnage carried by an oil tanker at sea.
   b. Wind velocity at 7 o'clock this morning.
   c. Number of text messages you received yesterday.
7. To choose a sample of 12 students from a statistics class of 36 students, which type of sample (simple random, systematic,

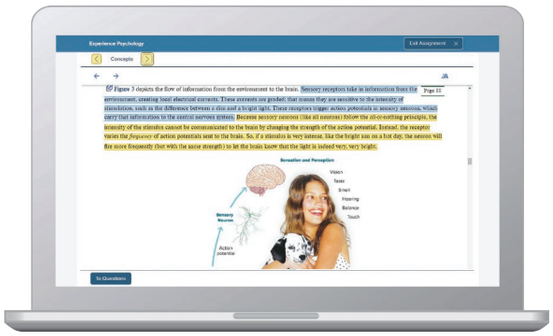
Mc Graw Hill connect®

Instructors: Student Success Starts with You

Tools to enhance your unique voice

Want to build your own course? No problem. Prefer to use our turnkey, prebuilt course? Easy. Want to make changes throughout the semester? Sure. And you'll save time with Connect's auto-grading too.

65%
Less Time Grading



Laptop: McGraw Hill; Woman/dog: George Doyle/Getty Images

Study made personal

Incorporate adaptive study resources like SmartBook® 2.0 into your course and help your students be better prepared in less time. Learn more about the powerful personalized learning experience available in SmartBook 2.0 at **www.mheducation.com/highered/connect/smartbook**

Affordable solutions, added value

Make technology work for you with LMS integration for single sign-on access, mobile access to the digital textbook, and reports to quickly show you how each of your students is doing. And with our Inclusive Access program you can provide all these tools at a discount to your students. Ask your McGraw Hill representative for more information.

Padlock: Jobalou/Getty Images

Solutions for your challenges

A product isn't a solution. Real solutions are affordable, reliable, and come with training and ongoing support when you need it and how you want it. Visit **www.supportateverystep.com** for videos and resources both you and your students can use throughout the semester.

Checkmark: Jobalou/Getty Images

SUPPORT AT every step

Students: Get Learning that Fits You

Effective tools for efficient studying

Connect is designed to make you more productive with simple, flexible, intuitive tools that maximize your study time and meet your individual learning needs. Get learning that works for you with Connect.

Study anytime, anywhere

Download the free ReadAnywhere app and access your online eBook or SmartBook 2.0 assignments when it's convenient, even if you're offline. And since the app automatically syncs with your eBook and SmartBook 2.0 assignments in Connect, all of your work is available every time you open it. Find out more at **www.mheducation.com/readanywhere**

"I really liked this app—it made it easy to study when you don't have your textbook in front of you."

- Jordan Cunningham,
Eastern Washington University

Calendar: owattaphotos/Getty Images

Everything you need in one place

Your Connect course has everything you need—whether reading on your digital eBook or completing assignments for class, Connect makes it easy to get your work done.

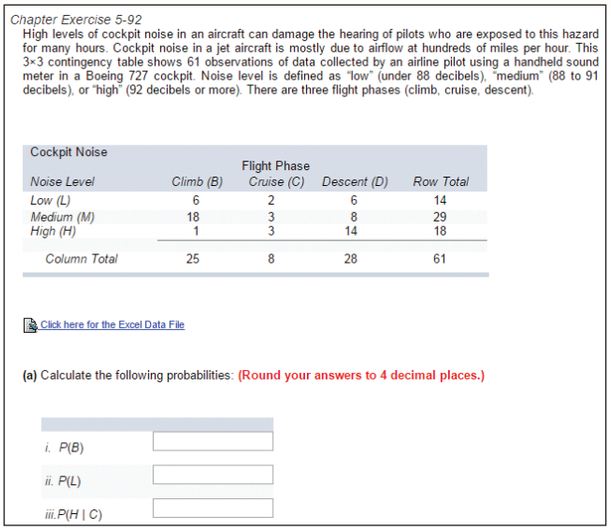
Learning for everyone

McGraw Hill works directly with Accessibility Services Departments and faculty to meet the learning needs of all students. Please contact your Accessibility Services Office and ask them to email accessibility@mheducation.com, or visit **www.mheducation.com/about/accessibility** for more information.

Top: Jenner Images/Getty Images, Left: Hero Images/Getty Images, Right: Hero Images/Getty Images

Additional Connect Features

Excel Data Sets A convenient feature is the inclusion of an Excel data file link in many problems using data files in their calculation. The link allows students to easily launch into Excel, work the problem, and return to Connect to key in the answer.



Chapter Exercise 5-92

High levels of cockpit noise in an aircraft can damage the hearing of pilots who are exposed to this hazard for many hours. Cockpit noise in a jet aircraft is mostly due to airflow at hundreds of miles per hour. This 3×3 contingency table shows 61 observations of data collected by an airline pilot using a handheld sound meter in a Boeing 727 cockpit. Noise level is defined as "low" (under 88 decibels), "medium" (88 to 91 decibels), or "high" (92 decibels or more). There are three flight phases (climb, cruise, descent).

Cockpit Noise	Flight Phase			
Noise Level	*Climb (B)*	*Cruise (C)*	*Descent (D)*	*Row Total*
Low (L)	6	2	6	14
Medium (M)	18	3	8	29
High (H)	1	3	14	18
Column Total	25	8	28	61

Click here for the Excel Data File

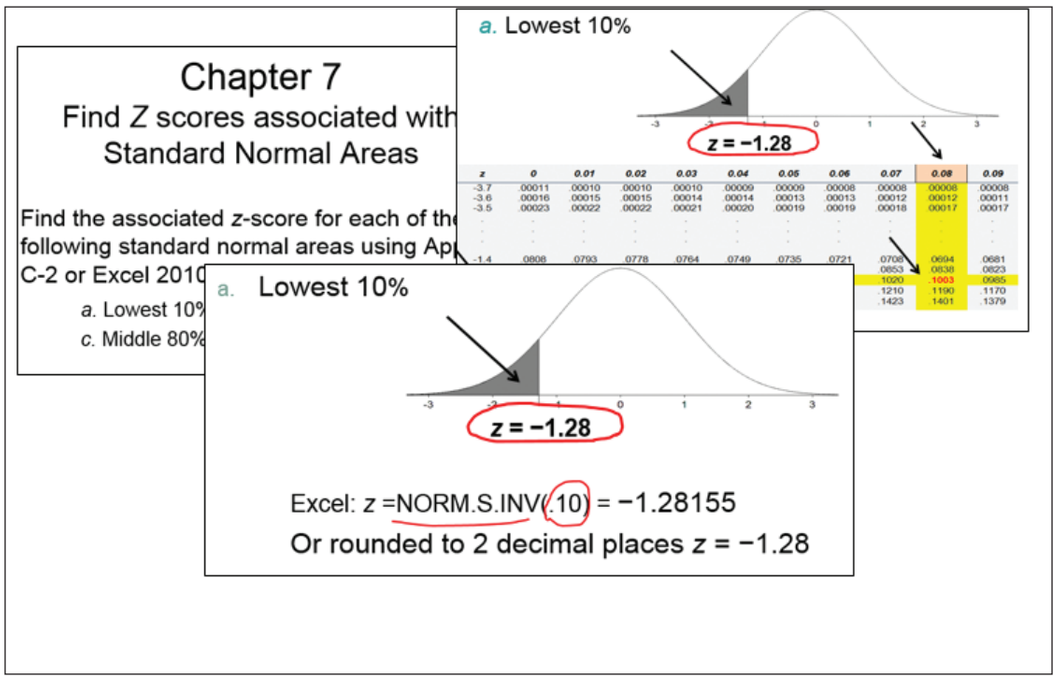
(a) Calculate the following probabilities: **(Round your answers to 4 decimal places.)**

i. $P(B)$

ii. $P(L)$

iii. $P(H \mid C)$

Guided Examples These narrated video walkthroughs provide students with step-by-step guidelines for solving selected exercises similar to those contained in the text. The student is given personalized instruction on how to solve a problem by applying the concepts presented in the chapter. The narrated voiceover shows the steps to take to work through an exercise. Students can go through each example multiple times if needed.

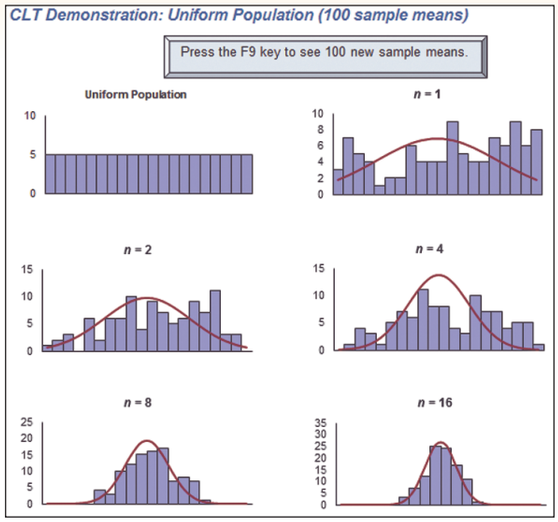


What Resources are Available for Students?

The following software tools are available to assist students in understanding concepts and solving problems.

LearningStats

LearningStats allows students to explore data and concepts at their own pace. It includes demonstrations, simulations, and tutorials that can be downloaded from Connect.



MegaStat® for Excel®

Access Card (ISBN: 0077426274) or online purchase at www.mhhe.com/megastat.
MegaStat is a full-featured Excel add-in that is available with this text. It performs statistical analyses within an Excel workbook. It does basic functions such as descriptive statistics, frequency distributions, and probability calculations as well as hypothesis testing, ANOVA, and regression.

MegaStat output is carefully formatted, and ease-of-use features include Auto Expand for quick data selection and Auto Label detect. Because *MegaStat* is easy to use, students can focus on learning statistics without being distracted by the software. *MegaStat* is always available from Excel's main menu. Selecting a menu item pops up a dialog box. *MegaStat* is updated continuously to work with the latest versions of Excel for Windows and Macintosh users.

Minitab® and Minitab Express®

Free trials and academic versions are available from Minitab at **http://www.minitab.com/en-us/academic/.**

R and RStudio

A sophisticated programming language for statistical computing and graphics plus an integrated development environment. This textbook offers detailed instructions for downloading, installing, and using free versions of R (**https://www.r-project.org/**) and RStudio (**https://rstudio.com/**).

What Resources are Available for Instructors?

Instructor resources are available through the Connect course at connect.mheducation.com. Resources include a complete Instructor's Manual in Word format, the complete Test Bank. Instructor PowerPoint slides, text art files, and more.

Tegrity: Lectures 24/7

Tegrity in Connect is a tool that makes class time available 24/7 by automatically capturing every lecture. With a simple one-click start-and-stop process, you capture all computer screens and corresponding audio in a format that is easy to search, frame by frame. Students can replay any part of any class with easy-to-use, browser-based viewing on a PC, Mac, iPod, or other mobile device.

Educators know that the more students can see, hear, and experience class resources, the better they learn. In fact, studies prove it. Tegrity's unique search feature helps students efficiently find what they need, when they need it, across an entire semester of class recordings. Help turn your students' study time into learning moments immediately supported by your lecture. With Tegrity, you also increase intent listening and class participation by easing students' concerns about note-taking. Using Tegrity in Connect will make it more likely you will see students' faces, not the tops of their heads.

Test Builder in Connect

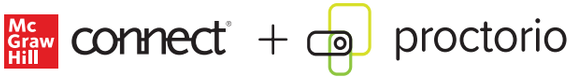
Available within Connect, Test Builder is a cloud-based tool that enables instructors to format tests that can be printed or administered within a LMS. Test Builder offers a modern, streamlined interface for easy content configuration that matches course needs, without requiring a download.

Test Builder allows you to:

- access all test bank content from a particular title.
- easily pinpoint the most relevant content through robust filtering options.
- manipulate the order of questions or scramble questions and/or answers.
- pin questions to a specific location within a test.
- determine your preferred treatment of algorithmic questions.
- choose the layout and spacing.
- add instructions and configure default settings.

Test Builder provides a secure interface for better protection of content and allows for just-in-time updates to flow directly into assessments.

Remote Proctoring & Browser-Locking Capabilities



New remote proctoring and browser-locking capabilities, hosted by Proctorio within Connect, provide control of the assessment environment by enabling security options and verifying the identity of the student.

Seamlessly integrated within Connect, these services allow instructors to control students' assessment experience by restricting browser activity, recording students' activity, and verifying students are doing their own work.

Instant and detailed reporting gives instructors an at-a-glance view of potential academic integrity concerns, thereby avoiding personal bias and supporting evidence-based claims.

McGraw Hill Customer Experience Information

For Customer Support, call **800-331-5094** or visit www.mhhe.com/support. One of our customer experience team members will be able to assist you in a timely fashion.

Acknowledgments

The authors would like to acknowledge some of the many people who have helped with this book. Thomas W. Lauer and Floyd G. Willoughby permitted quotation of a case study. Morgan Elliott, Karl Majeske, Robin McCutcheon, Kevin Murphy, John Sase, T. J. Wharton, and Kenneth M. York permitted questionnaires to be administered in their classes. Mark Isken, Ron Tracy, and Robert Kushler gave generously of their time as expert statistical consultants. Jonathan G. Koomey of E.O. Lawrence Berkeley National Laboratory offered valuable suggestions on visual data presentation.

We are grateful to Farrukh Abbas for his careful scrutiny of the text and for offering ideas on improving the text and exercises. Mark Isken has reliably provided Excel expertise and has suggested health care applications for examples and case studies. John Savio and the Michigan State Employees Credit Union provided ATM data. The Siena Research Institute has made its poll results available. J.D. Power and Associates generously provided permission to use vehicle quality data. The Public Interest Research Group in Michigan (PIRGIM) has generously shared data from its field survey of prescription drug prices.

Phil Rogers has offered numerous suggestions for improvement in both the textbook exercises and Connect. Milo A. Schield shared his research on "quick rules" for measuring skewness from summarized data. We owe special thanks to Aaron Kennedy and Dave Boennighausen of Noodles & Company; to Mark Gasta, Anja Wallace, and Clifton Pacaro of Vail Resorts; to Jim Curtin and Gordon Backman of Ball Corporation; and to Santosh Lakhan from The Verdeo Group for providing suggestions and access to data for Mini Cases and examples. For reviewing the material on quality, we wish to thank Kay Beauregard, administrative director at William Beaumont Hospital, and Ellen Barnes and Karry Roberts of Ford Motor Company. Amy Sheikh provided a new Facebook Friends data set, along with other excellent suggestions and reports from the "front lines" of her classes.

A special debt of gratitude is due to Noelle Bathurst, Harper Christopher, Amy Gehl, and Ryan McAndrews for their direction and support and Harvey Yep and Jamie Koch for managing the text and Connect pieces of the project. Thanks to the many reviewers who provided such valuable feedback including criticism that made the book better, some of whom reviewed several previous editions of the text. Any remaining errors or omissions are the authors' responsibility. Thanks too, to the participants in our focus groups and symposia on teaching business statistics, who have provided teaching ideas and insights from their experiences with students in diverse contexts. We hope you will be able to see in our book and the teaching package consideration of those ideas and insights.

Farrukh Abbas, *Barani Institute of Information Technology (Pakistan)*
Heather Adams, *University of Colorado—Boulder*
Sung Ahn, *Washington State University*
Mostafa Aminzadeh, *Towson University*
Scott Bailey, *Troy University*
Hope Baker, *Kennesaw State University*
Saad Taha Bakir, *Alabama State University*
Steven Bednar, *Elon University*
Adam Bohr, *University of Colorado—Boulder*
Katherine Broneck, *Pima Community College—Downtown*
Alan Cannon, *University of Texas—Arlington*
Deborah Carter, *Coahoma Community College*
Kevin Caskey, *SUNY—New Paltz*
Michael Cervetti, *University of Memphis*
Paven Chennamaneni, *University of Wisconsin—Whitewater*
Alan Chesen, *Wright State University*
Wen-Chyuan Chiang, *University of Tulsa*
Chia-Shin Chung, *Cleveland State University*
Joseph Coleman, *Wright State University—Dayton*
Robert Cutshall, *Texas A&M University—Corpus Christi*
Terry Dalton, *University of Denver*
Douglas Dotterweich, *East Tennessee State University*
Jerry Dunn, *Southwestern Oklahoma State University*
Michael Easley, *University of New Orleans*
Jerry Engeholm, *University of South Carolina—Aiken*
Mark Farber, *University of Miami*
Soheila Kahkashan Fardanesh, *Towson University*
Mark Ferris, *St. Louis University*
Stergios Fotopoulos, *Washington State University*
Vickie Fry, *Westmoreland County Community College*
Joseph Fuhr, *Widener University*
Bob Gillette, *University of Kentucky*
Malcolm Gold, *Avila University*
Don Gren, *Salt Lake City Community College*
Karina Hauser, *University of Colorado—Boulder*
Eric Hernandez, *Miami Dade College*
Clifford Hawley, *West Virginia University*
Yijun He, *Washington State University*
Natalie Hegwood, *Sam Houston State University*

Allen Humbolt, *University of Tulsa*
Patricia Igo, *Northeastern University*
Alam M. Imam, *University of Northern Iowa*
Marc Isaacson, *Augsburg College*
Kishen Iyengar, *University of Colorado—Boulder*
Christopher Johnson, *University of North Florida*
Jennifer Johnson, *San Jose State*
Ronald Johnson, *Central Alabama Community College*
Linda Jones, *Maryville University*
Jerzy Kamburowski, *University of Toledo*
Bob Kitahara, *Troy University*
Drew Koch, *James Madison University*
Mohammad Kazemi, *University of North Carolina–Charlotte*
Agnieszka Kwapisz, *Montana State University*
Mary Beth Camp, *Indiana University—Bloomington*
Kenneth Lawrence, *New Jersey Institute of Technology*
Bob Lynch, *University of Northern Colorado*
Bradley McDonald, *Northern Illinois University*
Richard McGowan, *Boston College*
Kelly McKillop, *University of Massachusetts*
Larry McRae, *Appalachian State University*
Robert Mee, *University of Tennessee—Knoxville*
Mahour Mellat Parast, *University of North Carolina–Pembroke*
John Miller, *Sam Houston State University*
Shelly Moore, *College of Western Idaho*
James E. Moran Jr., *Oregon State University*
Geraldine Moultine, *Northwood University*
Gourab Mukherjee, *University of Southern California*
Adam Munson, *University of Florida*
Joshua Naranjo, *Western Michigan University*
Anthony Narsing, *Macon State College*
Robert Nauss, *University of Missouri–St. Louis*
Pin Ng, *Northern Arizona University*
Thomas Obremski, *University of Denver*
Rahmat Ola Tavallai, *Walsh University*
Grace Onodipe, *Georgia Gwinnett College*
Ceyhun Ozgur, *Valparaiso University*
Ed Pappanastos, *Troy University*
Nitin Paranjpe, *Oakland University*
Eddy Patuwo, *Kent State University*
John Pickett, *University of Arkansas—Little Rock*
James Pokorski, *Virginia Polytechnic Institute & State University*
Stephan Pollard, *California State University—Los Angeles*
Claudia Pragman, *Minnesota State University*
Tammy Prater, *Alabama State University*
Michael Racer, *University of Memphis*
Azar Raiszadeh, *Chattanooga State Community College*
Mohammad Reza Oskoorouchi, *California State University—San Marcos*
Phil Rogers, *University of Southern California*
Milo A. Schield, *Augsburg College*
Sue Schou, *Idaho State University*
Elizabeth Scofidio, *Colorado State University*
Sankara N. Sethuraman, *Augusta State University*
Don Sexton, *Columbia University*
Thomas R. Sexton, *Stony Brook University*
Murali Shanker, *Kent State University*
Gary W. Smith, *Florida State University*
Courtenay Stone, *Ball State University*
Paul Swanson, *Illinois Central College*
Bedassa Tadesse, *University of Minnesota–Duluth*
Rahmat Tavallali, *Walsh University*
Deborah Tesch, *Xavier University*
Dharma S. Thiruvaiyaru, *Augusta State University*
Frank Tian Xie, *University of South Carolina–Aiken*
Bhavneet Walia, *Western Illinois University*
Michael Urizzo, *New Jersey City University*
Jesus M. Valencia, *Slippery Rock University*
Rachel Webb, *Portland State University*
Simone A. Wegge, *City University of New York*
Chao Wen, *Eastern Illinois University*
Alan Wheeler, *University of Missouri—St. Louis*
Blake Whitten, *University of Iowa*
Charles Wilf, *Duquesne University*
Anne Williams, *Gateway Community College*
Janet Wolcutt, *Wichita State University*
Frank Xie, *University of South Carolina—Aiken*
Ye Zhang, *Indiana University–Purdue University Indianapolis*
Mustafa R. Yilmaz, *Northeastern University*

Enhancements for Doane/Seward ASBE 7e

Many changes were motivated by advice from reviewers and users of the textbook. Besides hundreds of small edits and improved topic organization, these changes were common to most chapters:

- New overall design, colors, figures, and exercise layout for a brighter and more efficient look.
- New end-of-chapter *Software Supplements* for R, including two new appendixes (e.g., comparison of R with Excel)
- Updated test bank and updated/expanded *Big Data Sets.*
- Updated *Related Readings* and *Web Sources* for students who want to "dive deeper."
- Revised *LearningStats* demonstrations to illustrate concepts beyond what is possible in a textbook (e.g., simulations).
- Improved illustrations, figures, and tables.

Chapter 1—Overview of Statistics

Three new *Analytics in Action* (New AI Frontiers, Ethical Issues in AI, Algorithms: Friend or Foe).

New *Software Supplement* (Introduction to R and RStudio).

Updated *Related Readings.*

Chapter 2—Data Collection

Updated examples and Super Bowl Mini Case.

New *Analytics in Action* (Too Much Randomness).

Revised discussion of scales, surveys, and sampling.

Chapter 3–Describing Data Visually

Leaner presentation of making charts.

Updated examples (e.g., U.S. foreign trade).

Updated seven exercises and new example (bank earnings).

Updated *Related Readings.*

New *Software Supplement* (Histograms and Scatter Plots Using R).

Chapter 4–Descriptive Statistics

Streamlined discussion of main concepts.

Two new *Analytics in Action* (Future Job Titles, Measuring Volatility).

Enhanced discussion of center and variability.

One new exercise (bank assets) and several updated exercise data sets.

Updated *Related Readings.*

New *Software Supplement* (Descriptive Statistics Using R).

Chapter 5–Probability

Revised Venn diagram presentation.

New *Analytics in Action* (Climate Change).

Substantially revised, leaner treatment of contingency tables.

Two new exercises (clinical trials, 401K retirement contributions).

New *Software Supplement* (Contingency Tables Using R).

Chapter 6–Discrete Probability Distributions

One new *Analytics in Action* (Car Insurance Risk and Driverless Vehicles).

Several revised exercises.

New *Software Supplement* (Discrete Distributions Using R).

Chapter 7–Continuous Probability Distributions

Revised text, exercises, and illustrations.

New *Software Supplement* (Continuous Distributions Using R).

Chapter 8–Sampling Distributions and Estimation

New introduction.

Revised treatment of CLT and deleted section 8.3.

New *Analytics in Action* (Margin of Error in the Era of Big Data).

New discussion of bootstrap method.

New *Software Supplement* (Bootstrap Confidence Intervals in Minitab Express).

New *Software Supplement* (Confidence Intervals Using R).

Chapter 9–One-Sample Hypothesis Tests

New *Analytics in Action* (Walmart, Big Data, and Retail Analytics).

More on choosing α and β, significance and p-values.

Leaner discussion of power, OC curves, and variance tests.

New *Software Supplement* (One-Sample Hypothesis Tests Using R).

Chapter 10—Two-Sample Hypothesis Tests

Leaner discussion of two-tailed tests and confidence intervals.

New *Analytics in Action* (A/B Testing: Old Technique, New Application).

Updates in exercises and *Related Readings.*

New *Software Supplement* (Two-Sample Hypothesis Tests Using R).

Chapter 11—Analysis of Variance

New *Analytics in Action* (Experiments or Big Data?).

Leaner discussion of two-factor ANOVA.

Updated *Related Readings.*

New *Software Supplement* (ANOVA Using R).

Chapter 12—Simple Regression

New *Analytics in Action* (Predictive Maintenance and Machine Learning).

Revised discussion of confidence and prediction intervals.

New MiniCase 12.4 (exports and imports).

Leaner discussion of ill-conditioned data and spurious correlation.

New MiniCase 12.6 (assets and market capitalization).

Revised data set (U.S. price inflation) and updated *Related Readings.*

New *Software Supplement* (Simple Regression Using R).

Chapter 13—Multiple Regression

Simplified introduction and revised treatment of confidence and prediction intervals.

Delete Mini Case 13.4 (cockpit noise).

New *Analytics in Action* (People Analytics at Work).

New correlation matrix illustration (vehicle MPG), updated data set (CPI changes) and a new data set (immunotherapy drug prices).

Updated *Related Readings.*

New *Software Supplement* (Multiple Regression Using R).

Chapter 14—Time Series Analysis

Updated examples (U.S. labor Force, dollar exchange rates).

Updated examples of erratic (hurricanes, snowfall) and consistent patterns (health spending, Amazon revenue).

Leaner trend-fitting presentation, new formula for compound growth rate, and example of decomposition using R.

Two new trend interpretation exercises.

New *Analytics in Action* (Trend? Or Bubble?).

Updated 16 exercise data sets (e.g., bird strikes, renewable energy, PepsiCo, JetBlue, Coca-Cola revenue, revolving debt, plane shipments, federal budget, Boston Marathon, leisure industry, snowboarding, airspace delays).

Updated *Related Readings.*

New *Software Supplement* (Time Trends and Seasonality Using R).

Chapter 15—Chi-Square Tests

Simplified examples of raw data conversion.

New *Analytics in Action* (Confusion Matrix for Machine Learning).

Simplified treatment of GOF tests (Uniform, Normal, ECDF).

New exercise (age and social media preference).

Updated data sets (Derby, NL runs).

New *Software Supplement* (Chi-Square Tests Using R).

Chapter 16—Nonparametric Tests

One new exercise (movie reviews) and updated *Related Readings.*

New *Software Supplement* (Nonparametric Tests Using R).

Chapter 17—Quality Management

New *Analytics in Action* (Big Data Tracks a Virus).

Updated discussion of acceptance sampling.

Updated discussion of software (e.g., R CRAN packages).

Updated *Related Readings.*

Chapter 18—Simulation

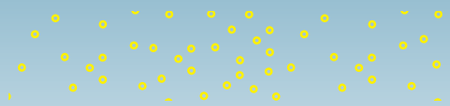
New table for random data in R.

Deleted bootstrap discussion.

Updated *Related Readings.*

Brief Contents

Contents



Applied Statistics

in Business and Economics

Seventh Edition

CHAPTER

1 Overview of Statistics

CHAPTER CONTENTS

CHAPTER LEARNING OBJECTIVES

When you finish this chapter, you should be able to

LO 1-1 Define statistics and explain some of its uses.

LO 1-2 List reasons for a business student to study statistics.

LO 1-3 Explain the uses of statistics in business.

LO 1-4 State the common challenges facing business professionals using statistics.

LO 1-5 List and explain common statistical pitfalls.

Juice Images/Getty Images

When managers are well informed about a company's internal operations (e.g., sales, production, inventory levels, time to market, warranty claims) and competitive position (e.g., market share, customer satisfaction, repeat sales), they can take appropriate actions to improve their business. Managers need reliable, timely information so they can analyze market trends and adjust to changing market conditions. Better data also can help a company decide which types of strategic information it should share with trusted business partners to improve its supply chain. *Statistics* and *statistical analysis* permit *data-based decision making* and reduce managers' need to rely on guesswork.

Statistics is a key component of the field of *business intelligence,* which encompasses all the technologies for collecting, storing, accessing, and analyzing data on the company's operations in order to make better business decisions. Statistics helps convert unstructured "raw" data (e.g., point-of-sale data, customer spending patterns) into *useful information* through online analytical processing (OLAP) and data mining, terms that you may have encountered in your other business classes. Statistical analysis focuses attention on key problems and guides discussion toward issues, not personalities or territorial struggles. While powerful database software and query systems are the key to managing a firm's data warehouse, relatively small Excel spreadsheets are often the focus of discussion among managers when it comes to "bottom line" decisions. That is why Excel is featured prominently in this textbook.

In short, companies increasingly are using **business analytics** to support decision making, to recognize anomalies that require tactical action, or to gain strategic insight to align business processes with business objectives. Answers to questions such as "How likely is this event?" or "What if this trend continues?" will lead to appropriate actions. Businesses that combine managerial judgment with statistical analysis are more successful.

1.1 WHAT IS STATISTICS?

LO 1-1 Define statistics and explain some of its uses.

Statistics is the science of collecting, organizing, analyzing, interpreting, and presenting data. Some experts prefer to call statistics **data science**, a trilogy of tasks involving data modeling, analysis, and decision making. A **statistic** is a single measure, reported as a number, used

to summarize a sample data set. Statistics may be thought of as a collection of methodologies to summarize, draw valid conclusions, and make predictions from empirical measurements. Statistics helps us organize and present information and extract meaning from raw data. Although it is often associated with the sciences and medicine, statistics is now used in every academic field and every area of business.

Plural or Singular?

Statistics The science of collecting, organizing, analyzing, interpreting, and presenting data.

Statistic A single measure, reported as a number, used to summarize a sample data set.

Many different measures can be used to summarize data sets. You will learn throughout this textbook that there can be different measures for different sets of data and different measures for different types of questions about the same data set. Consider, for example, a sample data set that consists of heights of students in a university. There could be many different uses for this data set. Perhaps the manufacturer of graduation gowns wants to know how long to make the gowns; the best *statistic* for this would be the *average* height of the students. But an architect designing a classroom building would want to know how high the doorways should be and would base measurements on the *maximum* height of the students. Both the average and the maximum are examples of a *statistic*.

You may not have a trained statistician in your organization, but any college graduate is expected to know something about statistics, and anyone who creates graphs or interprets data is "doing statistics" without an official title.

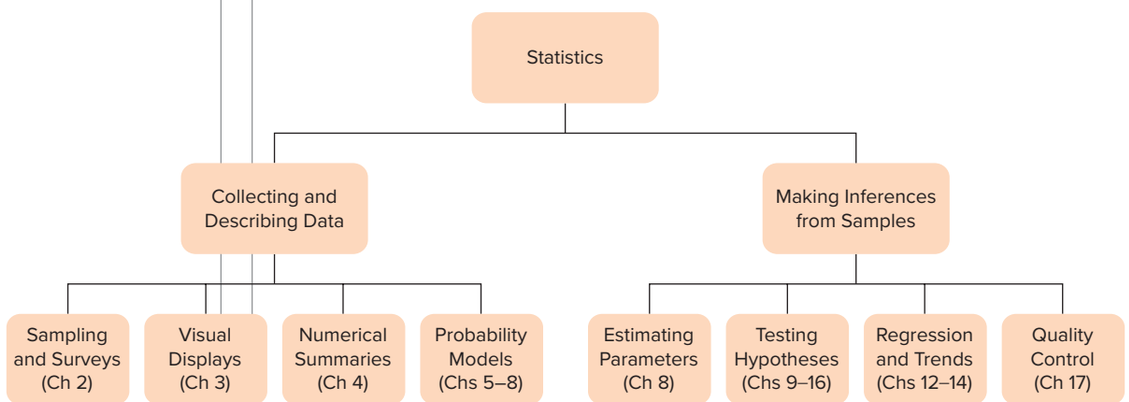
There are two primary kinds of statistics:

- **Descriptive statistics** refers to the collection, organization, presentation, and summary of data (either using charts and graphs or using a numerical summary).
- **Inferential statistics** refers to generalizing from a sample to a population, estimating unknown population parameters, drawing conclusions, and making decisions.

Figure 1.1 identifies the tasks and the text chapters for each.

Figure 1.1

Overview of Statistics



What Is Business Analytics?

Analytics is a broad field that uses statistics, mathematics, and computational tools to extract information from data. Analytics tools fall into three categories: *descriptive, predictive,* and *prescriptive.* This terminology derived from the questions we are trying to answer.

What happened? Businesses use *descriptive analytics* tools to analyze historical data and help them identify trends and patterns.

What is likely to happen next? Businesses use *predictive analytics* tools to predict probabilities of future events and help them forecast consumer behavior.

What actions do we take to achieve our goals? Businesses use *prescriptive analytics* tools to help them make decisions on how to achieve objectives within real-world constraints.

Analytics in Action

Using Analytics to Improve Business

Descriptive Analytics: Google Analytics provides a wide range of metrics for companies that want to better understand how their websites operate. This type of *descriptive analytics* allows companies to see, for example, loading time for their pages in different browsers and on different devices; how much time a viewer spends on the page; and the number of visits per hour, day, or week. In addition, companies can see performance measures such as bounce rate or click-through rate. Bounce rate shows the percentage of viewers who leave the site after viewing only the homepage. Click-through rate is the percentage of viewers who click on the ad when they view it on a webpage. Both of these metrics are important for evaluating the effectiveness of a web design.

Predictive Analytics: In 2015 Oracle acquired Datalogix, a company that analyzes consumer purchasing behavior to help its customers design the most effective digital marketing campaigns. The acquisition of Datalogix strengthened the tools Oracle provides as part of its data cloud offerings. *Predictive analytics* has been the key to Datalogix's success. Dr Pepper is one of its happy customers. When Dr Pepper wanted to increase sales, it turned to Datalogix for help in designing a successful Facebook advertising campaign. Through the use of analytics, Datalogix showed that by targeting Facebook users who purchased carbonated beverages and those who purchased Dr Pepper products in particular, Dr Pepper was able to increase its sales by 1.5 percent. And for every $1 spent on advertising, Dr Pepper saw a $3 increase in revenue. (www.datalogix.com)

Prescriptive Analytics: Hotel rooms and airplane seats are perishable, just like apples and lettuce. When a room or seat is not booked, the company has lost the ability to earn revenue for that night or flight. Setting the right price for the right customer on the right day is complicated. Hotels and airlines feed large data sets into mathematical algorithms that choose the optimal pricing scheme to maximize the revenue for that room or seat. This practice is called revenue management and is an example of *prescriptive analytics.* When large data sets are analyzed along with system constraints, companies can quickly simulate many different outcomes related to a wide variety of decisions. This allows the decision makers to choose a set of actions that will result in the best outcome for the company and its customers.

1.2 WHY STUDY STATISTICS?

LO 1-2

List reasons for a business student to study statistics.

A *BusinessWeek* article called statistics and probability "core skills for business people" in order to know when others are dissembling, to build financial models, or to develop a marketing plan. The 2019 Job Outlook survey conducted by the National Association of Colleges

and Employers (NACE) found that 72 percent of employers ranked "analytical and quantitative skills" as one of the most important attributes for a new hire. When asked to rank the attributes most sought after, employers said that the top seven attributes were (1) communication skills (written), (2) problem-solving skills, (3) ability to work on a team, (4) initiative, (5) analytical/quantitative skills, (6) strong work ethic, and (7) communication skills (verbal). Will these skills earn you a higher salary? According to the Bureau of Labor Statistics, the median salary for a statistician in 2019 was over $88,000 with job growth projected to be 30 percent over the next 10 years. (See www.bls.gov/ooh/math/mathematicians-and-statisticians.htm.)

Data Skills Count

"We look to recruit and groom leaders in our organization who possess strong quantitative skills in addition to a passion for what we do–delivering exceptional experiences at our extraordinary resorts every day. Knowing how to use and interpret data when making important business decisions is one of the keys to our Company's success."

Rob Katz, chairman and chief executive officer of Vail Resorts

Knowing statistics will make you a better consumer of other people's data analyses. You should know enough to handle everyday data problems, to feel confident that others cannot deceive you with spurious arguments, and to know when you've reached the limits of your expertise. Statistical knowledge gives your company a competitive advantage versus those that cannot understand their internal or external market data. Mastery of basic statistics gives you, the individual manager, a competitive advantage as you work your way through the promotion process or when you move to a new employer. For specialized training, many universities now offer master's degrees in business analytics. But here are some reasons for anyone to study statistics.

Communication The language of statistics is widely used in science, social science, education, health care, engineering, and even the humanities. In all areas of business (accounting, finance, human resources, marketing, information systems, operations management), workers use statistical jargon to facilitate communication. In fact, statistical terminology has reached the highest corporate strategic levels (e.g., "Six Sigma" at GE and Motorola). And in the multinational environment, the specialized vocabulary of statistics permeates language barriers to improve problem solving across national boundaries.

Computer Skills Whatever your computer skill level, it can be improved. Every time you create a spreadsheet for data analysis, write a report, or make an oral presentation, you bring together skills you already have and learn new ones. Specialists with advanced training design the databases and decision support systems, but you must handle daily data problems *without* experts. Besides, you can't always find an "expert," and, if you do, the "expert" may not understand your application very well. You need to be able to analyze data, use software with confidence, prepare your own charts, write your own reports, and make electronic presentations on technical topics.

Information Management Statistics can help you handle either too little or too much information. When insufficient data are available, statistical surveys and samples can be used to obtain the necessary market information. But most large organizations are closer to drowning in data than starving for them. Statistics can help summarize large amounts of data and reveal underlying relationships. You've heard of data mining? Statistics is the pick and shovel that you take to the data mine.

Technical Literacy Many of the best career opportunities are in growth industries propelled by advanced technology. Marketing staff may work with engineers, scientists, and manufacturing experts as new products and services are developed. Sales representatives must

understand and explain technical products like pharmaceuticals, medical equipment, and industrial tools to potential customers. Purchasing managers must evaluate suppliers' claims about the quality of raw materials, components, software, or parts.

Process Improvement Large manufacturing firms like Boeing or Toyota have formal systems for continuous quality improvement. The same is true of insurance companies, financial service firms like Vanguard or Fidelity, and the federal government. Statistics helps firms oversee their suppliers, monitor their internal operations, and identify problems. Quality improvement goes far beyond statistics, but every college graduate is expected to know enough statistics to understand its role in quality improvement.

Analytics in Action

Can Big Data Predict Airfares?

When you book an airline ticket online, does it annoy you when the next day you find a cheaper fare on exactly the same flight? Or do you congratulate yourself when you get a "good" fare followed by a price rise? Ticket price variation can be predicted, even though there is some uncertainty in the actual price you end up paying. KAYAK, a subsidiary of Priceline.com, is a travel planning website and mobile app used by millions of people every day to help forecast prices on airfares as well as hotels and rental cars. How does KAYAK predict airfare variation? It uses big data. Each day, KAYAK's analysts process over a billion online queries to forecast whether an airfare will go up or down. The mathematical models they use also provide a statistical confidence in their prediction so you can make decisions on when to purchase your airline ticket. If you travel a lot and take the time to use services such as KAYAK, you could save money. (See www.kayak.com/price-trend-explanation.)

1.3 APPLYING STATISTICS IN BUSINESS

LO 1-3 Explain the uses of statistics in business.

You've seen why statistics is important. Now let's look at some of the ways statistics is used in business.

Auditing A large firm pays over 12,000 invoices to suppliers every month. The firm has learned that some invoices are being paid incorrectly, but it doesn't know how widespread the problem is. The auditors lack the resources to check all the invoices, so they decide to take a sample to estimate the proportion of incorrectly paid invoices. How large should the sample be for the auditors to be confident that the estimate is close enough to the true proportion?

Marketing Many companies use Customer Relationship Management (CRM) to analyze customer data from multiple sources. With statistical and analytics tools such as correlation and data mining, they identify specific needs of different customer groups, and this helps them market their products and services more effectively.

Health Care Health care is a major business (one-sixth of the U.S. GDP). Hospitals, clinics, and their suppliers can save money by finding better ways to manage patient appointments, schedule procedures, or rotate their staff. For example, an outpatient cognitive retraining clinic for victims of closed-head injuries or stroke evaluates 56 incoming patients using a 42-item physical and mental assessment questionnaire. Each patient is evaluated independently by two experienced therapists. Are there statistically significant differences between the two therapists' evaluations of incoming patients' functional status? Are some assessment questions redundant? Do the initial assessment scores accurately predict the patients' lengths of stay in the program?

Quality Improvement A manufacturer of rolled copper tubing for radiators wishes to improve its product quality. It initiates a triple inspection program, sets penalties for workers

who produce poor-quality output, and posts a slogan calling for "zero defects." The approach fails. Why?

Purchasing A food producer purchases plastic containers for packaging its product. Inspection of the most recent shipment of 500 containers found that 3 of the containers were defective. The supplier's historical defect rate is .005. Has the defect rate really risen or is this simply a "bad" batch?

Medicine An experimental drug to treat asthma is given to 75 patients, of whom 24 get better. A placebo is given to a control group of 75 volunteers, of whom 12 get better. Is the new drug better than the placebo, or is the difference within the realm of chance?

Operations Management The Home Depot carries 50,000 different products. To manage this vast inventory, it needs a weekly order forecasting system that can respond to developing patterns in consumer demand. Is there a way to predict weekly demand and place orders from suppliers for every item without an unreasonable commitment of staff time?

Product Warranty A major automaker wants to know the average dollar cost of engine warranty claims on a new hybrid vehicle. It has collected warranty cost data on 4,300 warranty claims during the first six months after the engines are introduced. Using these warranty claims as an estimate of future costs, what is the margin of error associated with this estimate?

Mini Case 1.1

How Do You Sell Noodles with Statistics?

"The best answer starts with a thorough and thoughtful analysis of the data," says Aaron Kennedy, founder of Noodles & Company.

(Visit www.noodles.com to find a Noodles & Company restaurant near you.)
Courtesy of Noodles & Company

Noodles & Company introduced the *quick casual* restaurant concept, redefining the standard for modern casual dining in the United States in the 21st century. Noodles & Company first opened in Colorado in 1995 and has not stopped growing since. It has over 400 restaurants all across the United States from Portland and San Diego to Alexandria and Silver Spring with stops in cities such as Omaha and Naperville.

Noodles & Company has achieved this success with a customer-driven business model and fact-based decision making. Its widespread popularity and high growth rate