# STATISTICS for the BEHAVIORAL SCIENCES

SECOND EDITION

Gregory J. Privitera

# STATISTICS FOR THE BEHAVIORAL SCIENCES

## SECOND EDITION

Award-winning author Gregory J. Privitera engages students in an ongoing spirit of discovery with **a focus on how statistics apply to modern research problems**. Fully updated with current research, robust pedagogy, and a new four-color design, this new edition includes even more real-world examples.

**Real-world examples** make statistics relevant for students.

Example 2.1 applies the steps for distributing the frequency of scores in a data set.

### Example 2.1

A topic of interest in industrial organizational psychology is studying issues of worker safety (Brondino, Silva, & Pasini, 2012; Loughlin & Frone, 2004). To study this topic, a researcher records the number of complaints about safety filed by employees of 45 local small businesses over the previous 3 years. The results are listed in Table 2.3. In this example, we can construct a frequency distribution of these data.

**TABLE 2.3**

| | | | | |
|---|---|---|---|---|
| 45 | 98 | 83 | 50 | 86 |
| 66 | 66 | 88 | 95 | 73 |
| 88 | 55 | 76 | 115 | 66 |
| 92 | 110 | 79 | 105 | 101 |
| 101 | 85 | 90 | 92 | 81 |
| 55 | 95 | 91 | 92 | |
| 78 | 66 | 73 | 58 | |
| 86 | 92 | 51 | 63 | |
| 91 | 77 | 88 | 86 | |
| 94 | 80 | 102 | 107 | |

The number of safety complaints that employees at 45 local small businesses filed over the previous 3 years.

**Step 1:** Find the real range. The smallest value in Table 2.3 is 45, and the largest value is 115; therefore, 115 − 45 = 70. The real range is 70 + 1 = 71.

**Step 2:** Find the interval width. We can split the data into eight intervals (again, you choose the number of intervals). The interval width is the real range divided by the number of intervals: $\frac{71}{8} = 8.88$. The original data are listed as whole numbers, so we round up to the nearest whole number. The nearest whole number is the degree of accuracy of the data. The interval width is 9.

**Step 3:** Construct the frequency distribution. The frequency distribution table is shown in Table 2.4. The first interval starts with the smallest value (45) and contains nine values. To construct the next interval, add one degree of accuracy, or one whole number in this example, and repeat the steps to construct the remaining intervals.

*"One of the most important things I learned as an undergraduate in psychology was that psychology wasn't just a bunch of interesting things to know but that it is an open field of study where real-world things could be discovered. This book gives that to students."*

—Joshua J. Dobias, *Rutgers University*

### Example 7.3

Participants in a random sample of 100 college students are asked to state the number of hours they spend studying during finals week. Among all college students at this school (the population), the mean study time is equal to 20 hours per week, with a standard deviation equal to 15 hours per week. Construct a sampling distribution of the mean.

To construct the sampling distribution, we must (1) identify the mean of the sampling distribution, (2) compute the standard error of the mean, and (3) distribute the possible sample means 3 *SEM* above and below the mean.

Because the sample mean is an unbiased estimator of the population mean, the mean of the sampling distribution is equal to the population mean. The mean of this sampling distribution is equal to 20. The standard error is the population standard deviation (15) divided by the square root of the sample size (100):

$$\sigma_M = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{100}} = 1.50.$$

### 10.3 The Related-Samples *t* Test: Repeated-Measures Design

Two designs associated with selecting related samples are the repeated-measures design and the matched-pairs design. In Example 10.1, we compute the related-samples *t* test for a study using the repeated-measures design. In Example 10.2 (in Section 10.5), we compute a study using the matched-pairs design.

### Example 10.1

One area of focus in many areas of psychology and in education is on understanding and promoting reading among children and adults (Kent, Wanzek, & Al Otaiba, 2012; White, Chen, & Forsyth, 2010). Suppose we conduct a study with this area of focus by testing if teacher supervision influences the time that elementary school children read. To test this, we stage two 6-minute reading sessions and record the time in seconds that children spend reading in each session. In one session, the children read with a teacher present in the room; in another session, the same group of children read without a teacher present. The difference in time spent reading in the presence versus absence of a teacher is recorded. Table 10.5 lists the results of this hypothetical study with difference scores given. Test whether or not reading times differ using a .05 level of significance.

Step 1: State the hypotheses. Because we are testing whether (=) or not (≠) a difference exists, the null hypothesis states that there is no mean difference, and the alternative hypothesis states that there is a mean difference:

$H_0: \mu_D = 0$ There is no mean difference in time spent reading in the presence versus absence of a teacher.

$H_1: \mu_D \neq 0$ There is a mean difference in time spent reading in the presence versus absence of a teacher.

**FVI**

*The null and alternative hypotheses make statements about a population of mean difference scores.*

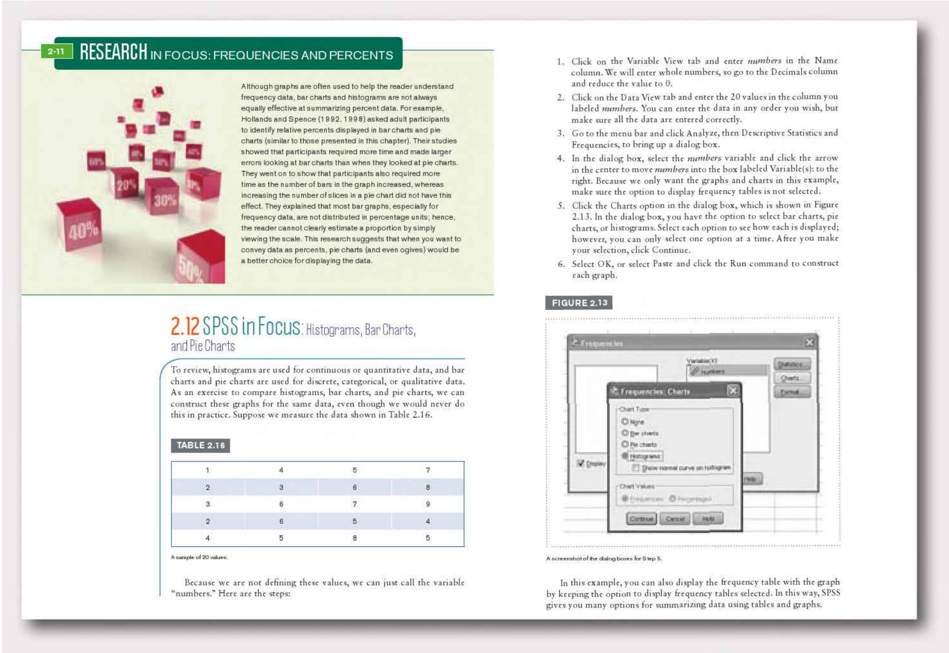
*"I loved the book in the first-edition form and love it even more from the changes I have reviewed. I will continue to use the text and recommend it to colleagues."*

—Jeffrey Kinderdietz, *Arizona State University*

# A FOCUS ON CLARITY

**Research in Focus sections** provide context by reviewing the most current research that illustrates the most important statistical concepts.

**SPSS in Focus sections** (with screenshots) draw from practical research examples to demonstrate how chapter concepts can be applied with SPSS.



2-11 RESEARCH IN FOCUS: FREQUENCIES AND PERCENTS

Although graphs are often used to help the reader understand frequency data, bar charts and histograms are not always equally effective at summarizing percent data. For example, Hollands and Spence (1992, 1998) asked adult participants to identify relative percents displayed in bar charts and pie charts (similar to those presented in this chapter). Their studies showed that participants required more time and made larger errors looking at bar charts than when they looked at pie charts. They went on to show that participants also required more time as the number of bars in the graph increased, whereas increasing the number of slices in a pie chart did not have this effect. They explained that most bar graphs, especially for frequency data, are not distributed in percentage units; hence, the reader cannot clearly estimate a proportion by simply viewing the scale. This research suggests that when you want to convey data as percents, pie charts (and even ogives) would be a better choice for displaying the data.

2.12 SPSS in Focus: Histograms, Bar Charts, and Pie Charts

To review, histograms are used for continuous or quantitative data, and bar charts and pie charts are used for discrete, categorical, or qualitative data. As an exercise to compare histograms, bar charts, and pie charts, we can construct these graphs for the same data, even though we would never do this in practice. Suppose we measure the data shown in Table 2.16.

TABLE 2.16

| | | | |
|---|---|---|---|
| 1 | 4 | 5 | 7 |
| 2 | 3 | 6 | 8 |
| 3 | 6 | 7 | 9 |
| 2 | 6 | 5 | 4 |
| 4 | 5 | 8 | 5 |

A sample of 20 values.

Because we are not defining these values, we can just call the variable "numbers." Here are the steps:

1. Click on the Variable View tab and enter *numbers* in the Name column. We will enter whole numbers, so go to the Decimals column and reduce the value to 0.
2. Click on the Data View tab and enter the 20 values in the column you labeled *numbers*. You can enter the data in any order you wish, but make sure all the data are entered correctly.
3. Go to the menu bar and click Analyze, then Descriptive Statistics and Frequencies, to bring up a dialog box.
4. In the dialog box, select the *numbers* variable and click the arrow in the center to move *numbers* into the box labeled Variable(s): to the right. Because we only want the graphs and charts in this example, make sure the option to display frequency tables is not selected.
5. Click the Charts option in the dialog box, which is shown in Figure 2.13. In the dialog box, you have the option to select bar charts, pie charts, or histograms. Select each option to see how each is displayed; however, you can only select one option at a time. After you make your selection, click Continue.
6. Select OK, or select Paste and click the Run command to construct each graph.

FIGURE 2.13

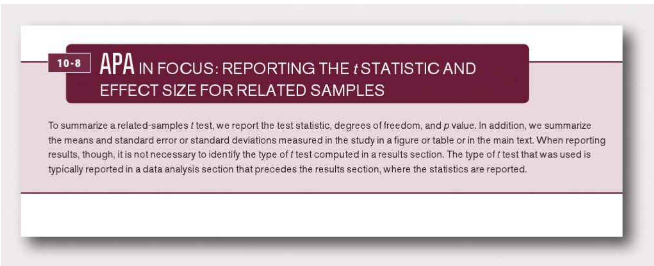
A screenshot of the dialog boxes for Step 5.

In this example, you can also display the frequency table with the graph by keeping the option to display frequency tables selected. In this way, SPSS gives you many options for summarizing data using tables and graphs.

*"The SPSS coverage is exceptional."*

–Walter M. Yamada, *Azusa Pacific University*

● ● ●



10-8 APA IN FOCUS: REPORTING THE *t* STATISTIC AND EFFECT SIZE FOR RELATED SAMPLES

To summarize a related-samples *t* test, we report the test statistic, degrees of freedom, and *p* value. In addition, we summarize the means and standard error or standard deviations measured in the study in a figure or table or in the main text. When reporting results, though, it is not necessary to identify the type of *t* test computed in a results section. The type of *t* test that was used is typically reported in a data analysis section that precedes the results section, where the statistics are reported.

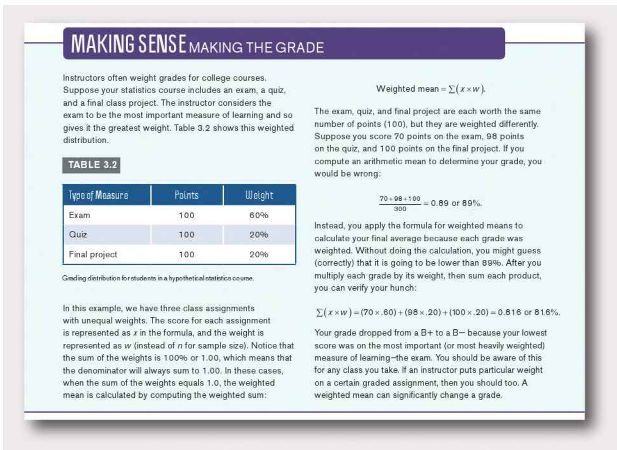
**APA in Focus sections** explain how to read and report statistical results in journals using current APA style.

"APA in Focus *is really useful for introducing students to the reporting standards from their very earliest exposures to these ideas."*

–Kristen T. Begosh, *University of Delaware*

● ● ●

**Making Sense sections** break down the most difficult concepts in statistics.



MAKING SENSE MAKING THE GRADE

Instructors often weight grades for college courses. Suppose your statistics course includes an exam, a quiz, and a final class project. The instructor considers the exam to be the most important measure of learning and so gives it the greatest weight. Table 3.2 shows this weighted distribution.

TABLE 3.2

| Type of Measure | Points | Weight |
|---|---|---|
| Exam | 100 | 60% |
| Quiz | 100 | 20% |
| Final project | 100 | 20% |

Grading distribution for students in a hypothetical statistics course.

In this example, we have three class assignments with unequal weights. The score for each assignment is represented as $x$ in the formula, and the weight is represented as $w$ (instead of $n$ for sample size). Notice that the sum of the weights is 100% or 1.00, which means that the denominator will always sum to 1.00. In these cases, when the sum of the weights equals 1.0, the weighted mean is calculated by computing the weighted sum:

$$\text{Weighted mean} = \sum(x \times w).$$

The exam, quiz, and final project are each worth the same number of points (100), but they are weighted differently. Suppose you score 70 points on the exam, 98 points on the quiz, and 100 points on the final project. If you compute an arithmetic mean to determine your grade, you would be wrong:

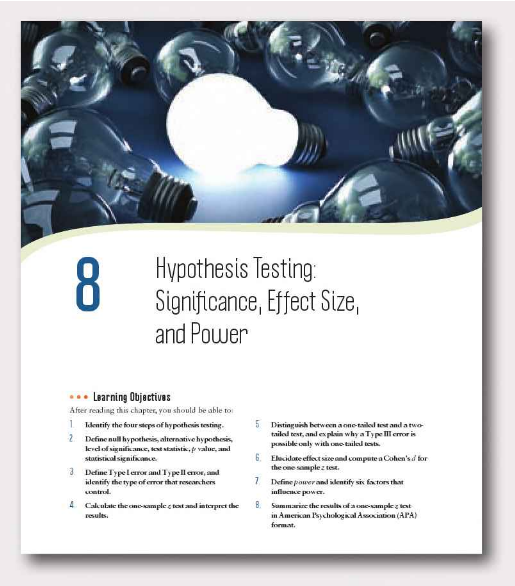
$$\frac{70+98+100}{300} = 0.89 \text{ or } 89\%.$$

Instead, you apply the formula for weighted means to calculate your final average because each grade was weighted. Without doing the calculation, you might guess (correctly) that it is going to be lower than 89%. After you multiply each grade by its weight, then sum each product, you can verify your hunch:

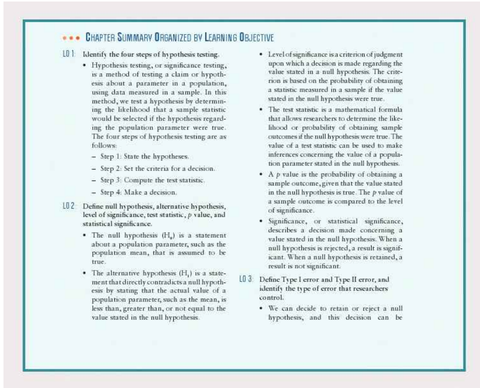
$$\sum(x \times w) = (70 \times .60) + (98 \times .20) + (100 \times .20) = 0.816 \text{ or } 81.6\%.$$

Your grade dropped from a B+ to a B− because your lowest score was on the most important (or most heavily weighted) measure of learning—the exam. You should be aware of this for any class you take. If an instructor puts particular weight on a certain graded assignment, then you should too. A weighted mean can significantly change a grade.

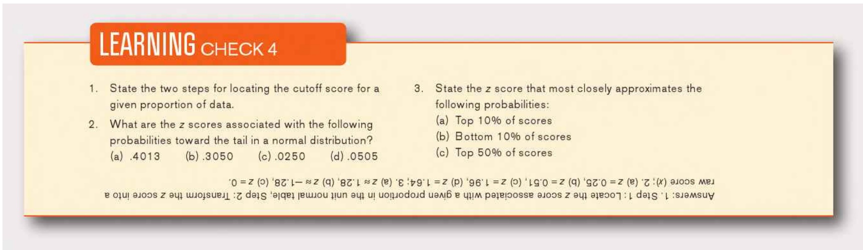
# A FOCUS ON PEDAGOGY AND PRACTICE

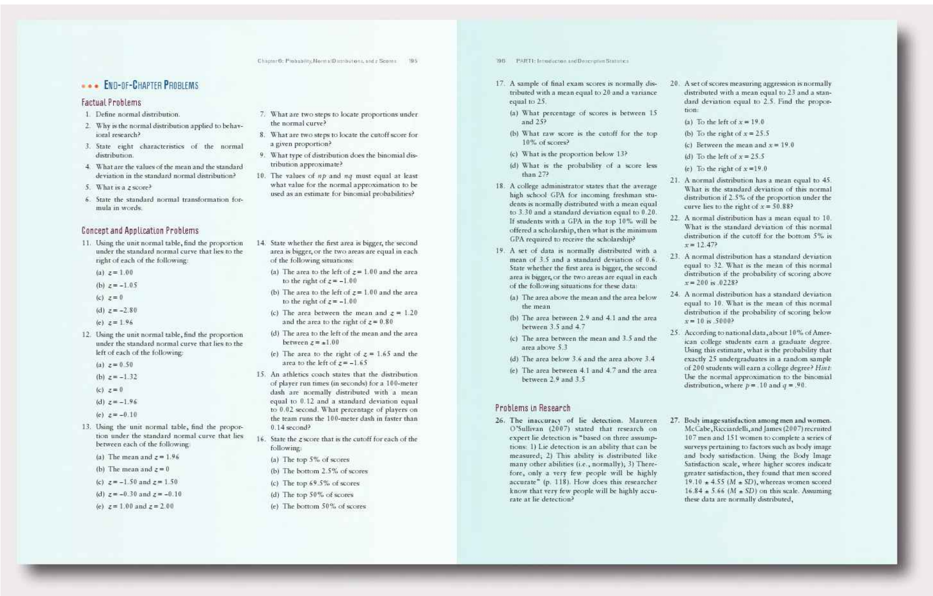



**Chapter Learning Objectives** are revisited and explained in **Chapter Summaries.**




**Learning Checks with answers** appear throughout each chapter, helping students assess their understanding of key concepts.




**More than 30 problems (organized by type) at the end of each chapter** provide a wealth of opportunities for practice.

*"I like the objectives, the readability of the text, the straightforwardness of the presentations of concepts, the problems that are quite appropriate on many levels (computation, theory, etc.), and the emphasis on SPSS."*

–Ted R. Bitner,
*DePauw University*

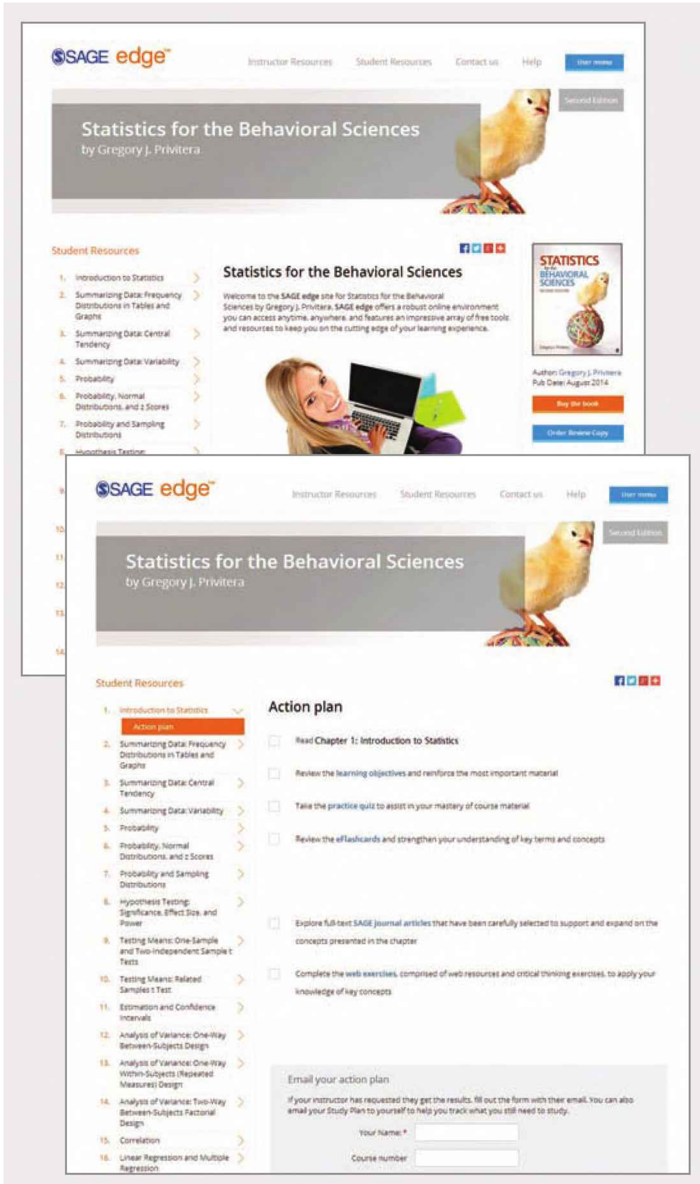
# A FOCUS ON RESOURCES

SAGE edge™ **edge.sagepub.com/priviterastats2e**

## Instructor Edge—

**SAGE edge for Instructors** supports teaching by making it easy to integrate quality content and create a rich learning environment for students. The SAGE edge instructor site includes:

- Extensive author-created test bank
- Sample course syllabi
- Editable, chapter-specific PowerPoint slides
- EXCLUSIVE Access to full-text SAGE journal articles
- Multimedia content
- Lecture notes
- A course cartridge for easy LMS integration




## Student Edge—

**SAGE edge for Students** provides a personalized approach to help students accomplish their coursework goals in an easy-to-use learning environment and features elements such as:

- Mobile-friendly eFlashcards and quizzes
- An online action plan that includes tips and feedback on course progress
- Chapter summaries
- Interactive exercises and meaningful web links
- EXCLUSIVE Access to full-text SAGE journal articles

## Price Edge—

SAGE provides students a great value and is priced **40+% less** than the average competing text price.

## Also Available:

- **WebAssign**, a powerful tool for creating online, auto-graded homework specific to the text
- **Student Study Guide**, with additional activities and exercises—only $5 when bundled with the text!

*"I think the resource package is excellent."*

—Ronald W. Stoffey,
*Kutztown University of Pennsylvania*

# Statistics for the Behavioral Sciences

**Second Edition**

**Gregory J. Privitera**

# Statistics for the Behavioral Sciences

**Second Edition**

**Gregory J. Privitera**
*St. Bonaventure University*

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
33 Pekin Street #02-01
Far East Square
Singapore 048763

Acquisitions Editor: Reid Hester
Associate Editor: Nathan Davidson
Editorial Assistant: Lucy Berbeo
Production Editor: Olivia Weber-Stenis
Copy Editor: Melinda Masson
Typesetter: C&M Digitals (P) Ltd.
Proofreader: Theresa Kay
Indexer: Jean Casalegno
Cover Designer: Bryan Fishman
Marketing Manager: Shari Countryman



Printed in Canada.

Library of Congress Cataloging-in-Publication Data

Privitera, Gregory J.
Statistics for the behavioral sciences/Gregory J. Privitera,
St. Bonaventure University.—Second edition.

pages cm
Includes bibliographical references and index.

ISBN 978-1-4522-8690-7 (hardcover: alk. paper)
1. Social sciences—Statistical methods. 2. Psychology—Statistical methods. I. Title.

HA29.P755 2015
519.5—dc23 2014007660

14 15 16 17 18 10 9 8 7 6 5 4 3 2 1

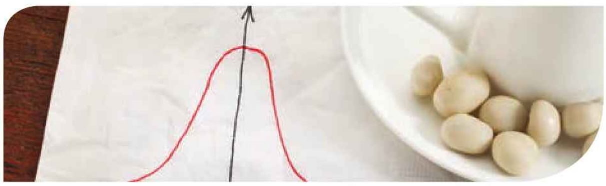
# ... Brief Contents

## PART V. MAKING INFERENCES ABOUT PATTERNS, FREQUENCIES, AND ORDINAL DATA 477

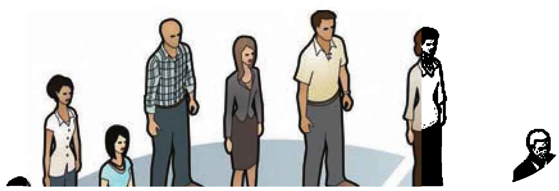
# ... Detailed Contents



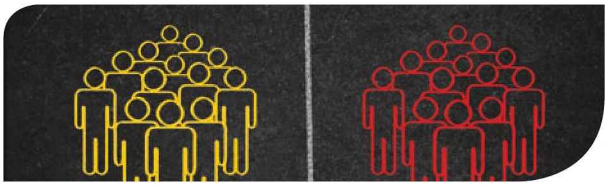

Your odds – 3:1



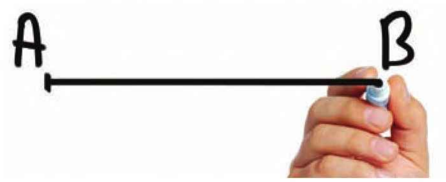




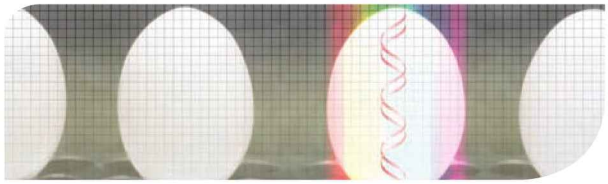
# PART III MAKING INFERENCES ABOUT ONE OR TWO MEANS 229






A
B









# ... About the Author

**Gregory J. Privitera** is an associate professor of psychology at St. Bonaventure University. Dr. Privitera received his PhD in behavioral neuroscience in the field of psychology at the State University of New York at Buffalo. He went on to complete postdoctoral research at Arizona State University before beginning his tenure at St. Bonaventure University. He is an author of multiple books on statistics, research methods, and the psychology of eating, in addition to authoring over two dozen peer-reviewed scientific articles aimed at advancing our understanding of health and promoting the intake of healthier diets. He oversees a variety of undergraduate research projects at St. Bonaventure University where over 20 undergraduate students, many of whom are now earning graduate degrees at various institutions, have coauthored research in his laboratories. For his work with students and fruitful record of teaching and advisement, Dr. Privitera was recognized in 2014 with the Award for Professional Excellence in Teaching, the highest teaching honor at St. Bonaventure University. For his work as an advisor of the Class of 2013, he was also honored as Advisor of the Year. He is the award-winning author of *Research Methods for the Behavioral Sciences,* recipient of the "Most Promising New Textbook" Award from the Text and Academic Authors Association (TAA). In addition to his teaching, research, and advisement, Dr. Privitera is a veteran of the U.S. Marine Corps, and is married with two children: a daughter, Grace, and a son, Aiden.

# ... Acknowledgments

I want to take a moment to thank all those who have been supportive and endearing throughout my career. To my family, friends, acquaintances, and colleagues—thank you for contributing to my perspective in a way that is indubitably recognized and appreciated. In particular to my son, Aiden Andrew, and daughter, Grace Ann—every moment I am with you I am reminded of what is truly important in my life. As a veteran, I also want to thank all those who serve and have served—there is truly no greater honor than to serve something greater than yourself.

To all those at SAGE Publications, know that I am truly grateful to be able to share and work with all of you. It is your vital contributions that have made this book possible and so special to me. Thank you.

I especially want to thank the thousands of statistics students across the country who will use this book. It is your pursuit of education that has inspired this contribution. My hope is that you take away as much from reading this book as I have from writing it.

Last, but certainly not least, I would also like to thank the many reviewers who gave me feedback during the development process.

Jan Andrews, *Vassar College*

Linda Bastone, *Purchase College, State University of New York*

Tara S. Behrend, *George Washington University*

Thomas E. Billimek, *San Antonio College*

Ryan P. Bowles, *Michigan State University*

Kelly Brennan-Jones, *The College at Brockport, State University of New York*

Rebecca Campbell, *Michigan State University*

Mary Jo Carnot, *Chadron State College*

Agatha E. Carroo, *North Carolina Central University*

Maria Czyzewska, *Texas State University*

Scott Drury, *Delta State University*

Armando X. Estrada, *Washington State University Vancouver*

Steve Hallam, *Keuka College*

Deborah J. Hendricks, *West Virginia University*

Michael R. Hulsizer, *Webster University*

Joseph G. Johnson, *Miami University of Ohio*

Angela S. Kelling, *Marshall University*

Lisa Kilanowski-Press, *Niagara University*

E. David Klonsky, *Stony Brook University*
Charlie L. Law, *Penn State Schuylkill*
Jeff Lorentz, *University of Houston*
Shanhong Luo, *University of North Carolina Wilmington*
Justin Marcus, *University of Central Florida*
Douglas McHugh, *Indiana University*
Osvaldo F. Morera, *University of Texas at El Paso*
Paige Muellerleile, *Marshall University*
Bryan Myers, *University of North Carolina Wilmington*
Gayathri Narasimham, *Vanderbilt University*
Fred Oswald, *Rice University*
Craig D. Parks, *Washington State University*
Kelly Peracchi, *University of New Hampshire*
William M. Rogers, *Grand Valley State University*
Bethany Schneider, *Indiana University*
Sonya K. Sedivy, *University of Wisconsin–Milwaukee*
Duane Shuttlesworth, *Delta State University*
Stephanie L. Simon-Dack, *Ball State University*
John M. Spores, *Purdue University North Central*
Cheryl Terrance, *University of North Dakota*
Doris Zahner, *Barnard College*

# ... Preface to the Instructor

## Philosophical Approach

On the basis of years of experience and student feedback, I was inspired to write a book that professors could truly teach from—one that would relate statistics to science using current, practical research examples and one that would be approachable (and dare I say interesting!) to students. I wrote this book in that spirit to give the reader one clear message: Statistics is not something static or antiquated that we used to do in times past; statistics is an ever-evolving discipline with relevance to our daily lives. This book is designed not only to engage students in using statistics to summarize data and make decisions about behavior but also to emphasize the ongoing spirit of discovery that emerges when using today's technologies to understand the application of statistics to modern-day research problems. How does the text achieve this goal? It exposes students to statistical applications in current research, tests their knowledge using current research examples, gives them step-by-step instruction for using SPSS with examples, and makes them aware of how statistics is important for their generation—all through the use of the following key themes, features, and pedagogy.

## Themes, Features, and Pedagogy

### Emphasis on Student Learning

- **Conversational writing style.** I write in a conversational tone that speaks to the reader as if he or she is the researcher. It empowers students to view statistics as something they are capable of understanding and using. It is a positive psychology approach to writing that involves students in the process of statistical analysis and making decisions using statistics. The goal is to motivate and excite students about the topic by making the book easy to read and follow without "dumbing down" the information they need to be successful.
- **Learning objectives.** Clear learning objectives are provided at the start of each chapter to get students focused on and thinking about the material they will be learning. At the close of each chapter, the chapter summaries reiterate these learning objectives and then summarize the key chapter content related to each objective.
- **Learning Checks** are inserted throughout each chapter (for students to review what they learn, as they learn it), and many figures and tables are provided to illustrate statistical concepts and summarize statistical procedures.
- **Making Sense** sections support critical and difficult material. In many years of teaching statistics, I have found certain areas of statistics where students struggle the most. To address this, I include Making Sense sections in each chapter to break down difficult concepts,

review important material, and basically "make sense" of the most difficult material taught in this book. These sections are aimed at easing student stress and making statistics more approachable. Again, this book was written with student learning in mind.

- **Review problems.** At least 32 review problems are included at the end of each chapter. They include *Factual Problems, Concept and Application Problems,* and *Problems in Research.* Unlike the questions in most statistics textbooks, these questions are categorized for you so that you can easily identify and specifically test the type of knowledge you want to assess in the classroom. This format tests student knowledge and application of chapter material while also giving students more exposure to how current research applies to the statistics they learn.
- **Additional features.** Additional features in each chapter are aimed at helping students pull out key concepts and recall important material. For example, key terms are bolded, boxed, and defined as they are introduced to make it easier for students to find these terms when reviewing the material and to grab their attention as they read the chapters. At the end of the book, each key term is summarized in a glossary. Also, margin notes are placed throughout each chapter for students to review important material. They provide simple explanations and summaries based on those given in detail in the text.

## Focus on Current Research

- **Research in Focus.** To introduce the context for using statistics, Chapters 1 to 6 include Research in Focus sections that review pertinent research that makes sense of or illustrates important statistical concepts discussed in the chapter. Giving students current research examples can help them "see" statistical methods as they are applied today, not as they were done 20 years ago.
- **APA in Focus.** As statistical designs are introduced in Chapters 7 to 18, I present APA in Focus sections that explain how to summarize statistical results for each inferential statistic taught. Together, these sections support student learning by putting statistics into context with research and also explaining how to read and report statistical results in research journals that follow American Psychological Association (APA) style.
- **Current research examples.** Many of the statistics computed in this book are based on or use data from published research. This allows students to see the types of questions that behavioral researchers ask while learning about the statistics researchers use to answer research questions. Students do not need a background in research methods to read through the research examples, which is important because most students have not taken a course in research methods prior to taking a statistics course.
- **Problems in Research.** The end-of-chapter review questions include a section of Problems in Research that come straight from the literature.

These classroom-tested problems use the data or conclusions drawn from published research to test knowledge of statistics and are taken from a diverse set of research journals and behavioral disciplines. The problems require students to think critically about published research in a way that reinforces statistical concepts taught in each chapter.

- **Balanced coverage of recent changes in the field of statistics.** I take into account recent developments in the area of statistics. For example, while eta-squared is still the most popular estimate for effect size, there is a great deal of research showing that it overestimates the size of an effect. That being said, a modification to eta-squared, called omega-squared, is considered a better estimate for effect size and is being used more and more in published articles. I teach both, giving students a full appreciation for where statistics currently stands and where it is likely going in the future. Other examples include a full chapter on confidence intervals and detailed reviews of factors that influence power (a key requirement for obtaining grant money and conducting an effective program of research).

## Integration of SPSS

- **Guide to using SPSS with this book.** For professors who teach statistics and SPSS, it can be difficult to teach from a textbook and a separate SPSS manual. The manual often includes different research examples or language that is inconsistent with what appears in the textbook and overall can be difficult for students to follow. This book changes all that by nesting SPSS coverage into the textbook. It begins with the guide at the front of the book, "How to Use SPSS With This Book," which provides students with an easy-to-follow, classroom-tested overview of how SPSS is set up, how to read the Data View and Variable View screens, and how to use the SPSS in Focus sections in the book.
- **SPSS in Focus.** Many statistics textbooks for the behavioral sciences omit SPSS, include it in an appendix separate from the main chapters in the book, include it at the end of chapters with no useful examples or context, or include it in ancillary materials that often are not included with course content. In this edition of *Statistics for the Behavioral Sciences,* SPSS is included in each chapter as statistical concepts are taught. This instruction is given in the SPSS in Focus sections. These sections provide step-by-step, classroom-tested instruction using practical research examples for how the concepts taught in each chapter can be applied using SPSS. Students are supported with screenshot figures and explanations for how to read SPSS outputs.

In addition, there is one more overarching feature that I refer to as *teachability.* While this book is comprehensive and a great reference for any undergraduate student, it is often too difficult for instructors to cover every topic in this book. For this reason, the chapters are organized into sections, each of which can largely stand alone. This gives professors the ability to more easily manage course content by assigning students particular sections in each chapter when they do not want to teach all topics covered in the entire chapter. So this book was not only written with the student in mind; it was also written with the professor in mind. Here are some brief highlights of what you will find in each chapter:

## Chapter Overviews

### Chapter 1. Introduction to Statistics

Students are introduced to scientific thinking and basic research design relevant to the statistical methods discussed in this book. In addition, the types of data that researchers measure and observe are introduced in this chapter. The chapter is to the point and provides an introduction to statistics in the context of research.

### Chapter 2. Summarizing Data: Frequency Distributions in Tables and Graphs

This chapter provides a comprehensive introduction to frequency distributions and graphing using research examples that give students a practical context for when these tables and graphs are used. In addition, students are exposed to summaries for percent data and percentile points. Throughout the chapter, an emphasis is placed on showing students how to decide between the many tables and graphs used to summarize various data sets.

### Chapter 3. Summarizing Data: Central Tendency

This chapter places particular emphasis on what measures of central tendency are, how they are computed, and when they are used. A special emphasis is placed on interpretation and use of the mean, the median, and the mode. Students learn to appropriately use these measures to describe data for many different types of distributions.

### Chapter 4. Summarizing Data: Variability

Variability is often difficult to conceptually understand. So I begin with an illustration for how this chapter will show students what variability is actually measuring. I clarify immediately that variability can never be negative, and I give a simple explanation for why. These are difficult obstacles for students, so I begin with this to support student learning from the very beginning of the chapter. The remainder of the chapter introduces various measures of variability to include variance and standard deviation for data in a sample and population.

### Chapter 5. Probability

This is a true probability chapter with many current research examples. This chapter does not ask about the probability of rolling dice; it looks at how probability problems—from simple probability, to Bayes's theorem, to expected values—are applied to answer questions about behavior. After reading this chapter, students will not feel like they have to gamble in order to apply probability.

## Chapter 6. Probability, Normal Distributions, and *z* Scores

At an introductory level, the normal distribution is center stage. It is at least mentioned in almost every chapter of this book. It is the basis for statistical theory and the precursor to most other distributions students will learn about. For this reason, I dedicate an entire chapter to its introduction. This chapter uses a variety of research examples to help students work through locating probabilities above the mean, below the mean, and between two scores, and even to help them calculate *z* scores.

## Chapter 7. Probability and Sampling Distributions

This is a comprehensive chapter for sampling distributions of both the mean and variance. This chapter introduces the sampling distribution and standard error in a way that helps students to see how the sample mean and sample variance can inform us about the characteristics we want to learn about in some otherwise unknown population. In addition, the chapter is organized in a way that allows professors to easily manage reading assignments for students that are consistent with what they want to discuss in class.

## Chapter 8. Hypothesis Testing: Significance, Effect Size, and Power

In my experience, shifting from descriptive statistics to inferential statistics is particularly difficult for students. For this reason, this chapter provides a comprehensive introduction to hypothesis testing, significance, effect size, power, and more. In addition, students are introduced to power in the context that emphasizes how essential this concept is for research today. Two sections are devoted to this topic, and this chapter uses data from published research to introduce hypothesis testing.

## Chapter 9. Testing Means: One-Sample and Two-Independent-Sample *t* Tests

This chapter introduces students to *t* tests for one sample and two independent samples using current research examples. This allows students to apply these tests in context with the situations in which they are used. In addition, students are introduced to two measures for proportion of variance—one that is most often used (eta-squared) and one that is less biased and becoming more popular (omega-squared). This gives students a real sense of where statistics is and where it is likely going.

## Chapter 10. Testing Means: The Related-Samples *t* Test

Many textbooks teach the related-samples *t* test and spend almost the entire chapter discussing the repeated-measures design. This is misleading because the matched-pairs design is also analyzed using this *t* test. It unnecessarily leads students to believe that this test is limited to a repeated-measures design, and it is not. For this reason, I teach the related-samples *t* test for both designs, explaining that the assumptions, advantages, and disadvantages vary depending on the design used. Students are clearly introduced to the context for using this test and the research situations that require its use.

## Chapter 11. Estimation and Confidence Intervals

Confidence intervals and estimation have become increasingly emphasized among behavioral scientists and statisticians. Although they have a lot in common with significance testing, there are many who believe that someday confidence intervals will replace significance testing. Maybe, maybe not; regardless, this emphasis justifies dedicating a full chapter to reviewing this topic. Particular emphasis is placed on describing the similarities and differences between significance testing and confidence intervals.

## Chapter 12. Analysis of Variance: One-Way Between-Subjects Design

The one-way between-subjects analysis of variance (ANOVA) and its assumptions, hypotheses, and calculations are all reviewed. A particular emphasis is placed on reviewing post hoc designs and what should be done following a significant result. Two post hoc tests are reviewed in order of how powerful they are at detecting an effect. This gives students a decision-focused introduction by showing them how to choose statistics that are associated with the greatest power to detect an effect.

## Chapter 13. Analysis of Variance: One-Way Within-Subjects (Repeated-Measures) Design

The one-way within-subjects ANOVA and its assumptions, hypotheses, and calculations are all reviewed. Students are also introduced to post hoc tests that are most appropriate in situations when samples are related. This is important because many statistics textbooks fail to even recognize that other commonly published post hoc tests are not well adapted for related samples. In addition, a full discussion of consistency and power is included to help students see how this design can increase the power of an analysis to detect an effect.

## Chapter 14. Analysis of Variance: Two-Way Between-Subjects Factorial Design

This chapter provides students with an introduction to the two-way between-subjects factorial design. Students are given illustrations showing exactly how to interpret main effects and interactions, as well as given guidance as to which effects are most informative and how to describe these effects. This is a decision-focused chapter, helping students understand the various effects in a two-way ANOVA design and how they can be analyzed and interpreted to answer a variety of research questions.

### Chapter 15. Correlation

This chapter is unique in that it is organized in a way that introduces the Pearson correlation coefficient, effect size, significance, assumptions, and additional considerations up front before introducing the Spearman, point-biserial, and phi correlation coefficients. This makes it easier for professors who only want to discuss the Pearson correlation (or any other correlation coefficient) to assign students readings that are specific to the concepts they will discuss in lectures. This also minimizes confusion among students and gives professors more control to manage course content and readings.

### Chapter 16. Linear Regression and Multiple Regression

This chapter introduces how a straight line can be used to predict behavioral outcomes. Many figures and tables are included to illustrate and conceptualize regression and how it describes behavior. Also, an analysis of regression is introduced for one (linear regression) and two (multiple regression) predictor variables. Parallels between regression and ANOVA are also drawn to help students see how this analysis relates to other tests taught in previous chapters.

### Chapter 17. Nonparametric Tests: Chi-Square Tests

One of the most difficult parts of teaching chi-square tests can be explaining their interpretation. Much of the interpretation of the results of a chi-square is intuitive or speculative. These issues and the purposes for using these tests are included. In addition, this chapter is linked with the previous chapter by showing students how measures of effect size for the chi-square test are linked with phi correlations. This gives students an appreciation for how these measures are related.

### Chapter 18. Nonparametric Tests: Tests for Ordinal Data

This final chapter is aimed at introducing alternative tests for ordinal data. A key emphasis is to relate each test to those already introduced in previous chapters. The tests taught in this chapter are alternatives for tests taught in Chapters 9, 10, 12, and 13. The tests are introduced in separate sections that make it easier for professors to assign sections of readings for only those tests they want to teach. Again, this can minimize confusion among students and gives the professor more control to manage course content and readings.

## APPENDIXES

Appendix A gives students a basic math review specific to the skills they need for the course. The appendix is specifically written to be unintimidating.

From the beginning, students are reassured that the level of math is basic and that they do not need a strong background in mathematics to be successful in statistics. Learning Checks are included throughout this appendix, and more than 100 end-of-chapter review problems are included to give students all the practice they need to feel comfortable. In addition, Appendix B gives the tables needed to find critical values for the test statistics taught in this book. Appendix C gives the answers for even-numbered problems for the end-of-chapter questions. This allows students to practice additional questions and be able to check their answers in Appendix C.

## Supplements

Ancillaries for this book include the following:

- **Student Study Guide:** Contains chapter learning objectives, chapter outlines, key formulas, tips and cautions, self-tests and quizzes, and exercises designed to test students' understanding of APA style and SPSS.
- **Instructor Teaching Site:** Contains PowerPoint slides, course syllabi, lecture notes, a test bank, solutions for the problems in the Student Study Guide, and solutions for all end-of-chapter problems in the text.
- **Student Study Site:** Contains e-flashcards, "Learning From SAGE Journal Articles" activities with full-text journal articles and discussion questions, web quizzes, and more.

Thank you for choosing *Statistics for the Behavioral Sciences* and best wishes for a successful semester!

Gregory J. Privitera
St. Bonaventure, New York

# … To the Student—How to Use SPSS With This Book

SPSS is an innovative statistical computer program used to compute most statistics taught in this book. This preface provides you with an overview to familiarize you with how to open, view, and understand this software. The screenshots in this book show IBM SPSS Version 21.0 for the PC. Still, even if you use a Mac or different version, the figures and instructions should provide a rather effective guide for helping you use this statistical software (with some minor differences, of course). SPSS is introduced throughout this book, so it will be worthwhile to read this preface before moving into future discussions of SPSS. Included in this preface is a general introduction to familiarize you with this software.

Understanding this software is especially important for those interested in research careers because it is the most widely used statistical program in the social and behavioral sciences. That is not to minimize the importance of understanding how to compute a mean or plot a bar graph by hand—but knowing how to enter, analyze, and interpret statistics using SPSS is equally important for no other reason than you will need it. This is an essential complement to your readings in this book. By knowing how and why you compute certain statistics, you will better understand and interpret the output from SPSS software.

## Overview of SPSS: What Are You Looking At?

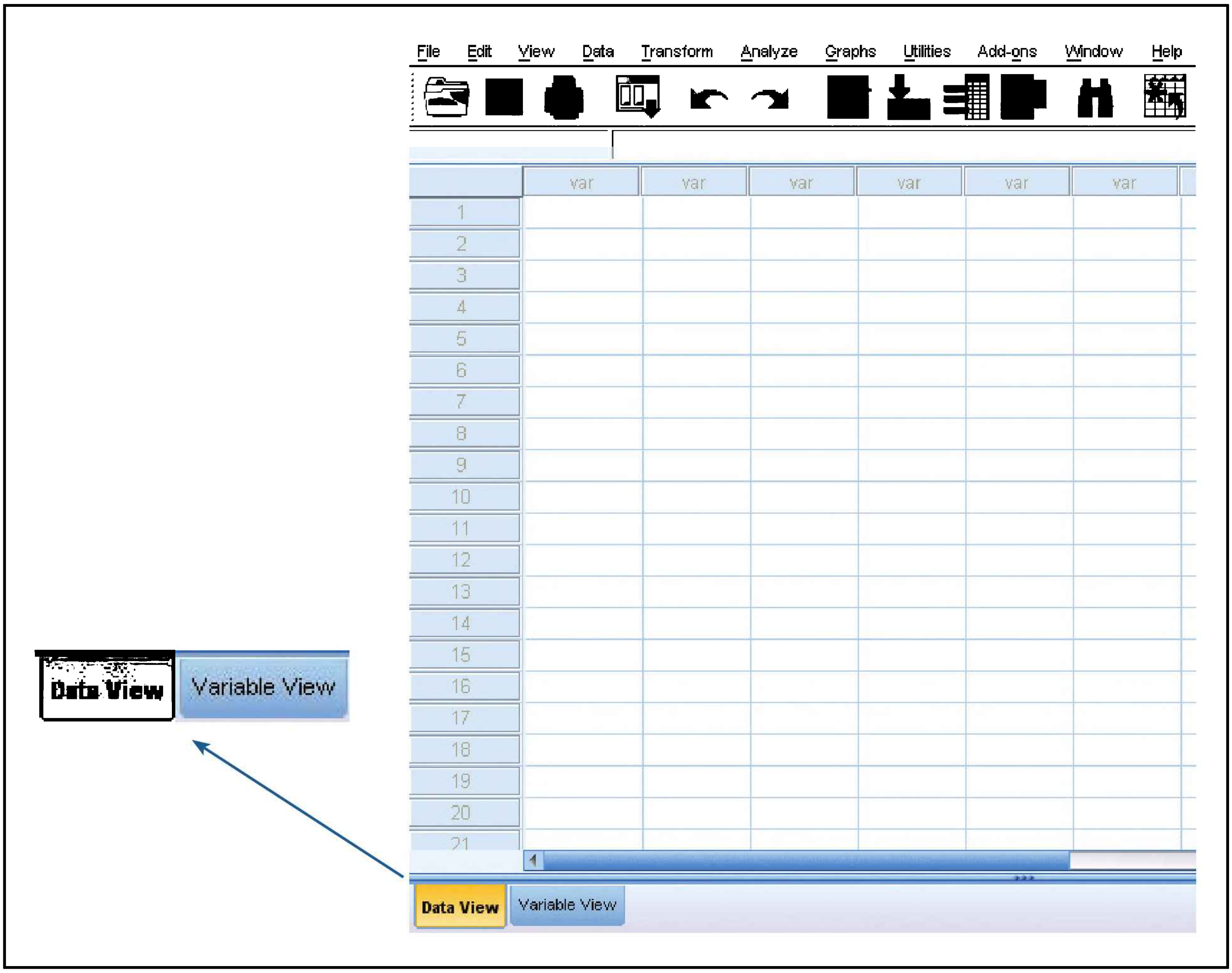
When you open SPSS, you will see a window that looks similar to an Excel spreadsheet. (In many ways, you will enter and view the data like you do in Microsoft Excel.) At the bottom of the window, you will see two tabs as shown in Figure P.1. The Data View tab is open by default. The Variable View tab to the right of it is used to view and define the variables being studied.

### Data View

The Data View screen includes a menu bar (located at the top of the screen), which displays commands that perform most functions that SPSS provides. These commands include File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. Each command is introduced as needed in each chapter in the SPSS in Focus sections, although the command of most use to you will be the Analyze command in the menu bar.

Below the menu bar you will find the toolbar, which includes a row of icons that perform various functions. We use some of these icons, whereas others are beyond the scope of this book. The purpose and function of each icon are introduced as needed in each chapter in the SPSS in Focus sections.

**FIGURE P.1** The Data View Default View in SPSS



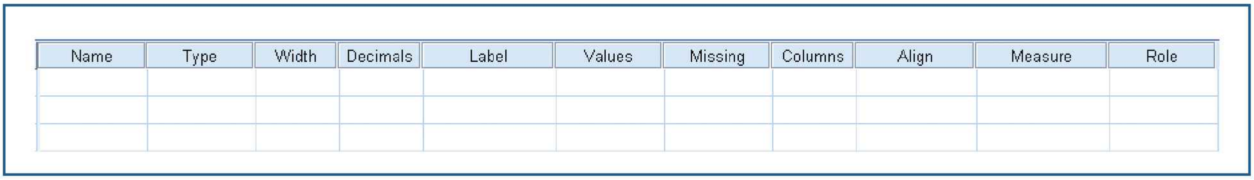

The highlighted tab (pulled out with an arrow in this figure) indicates which view you are looking at. In this figure, the Data View tab is highlighted.

Within the spreadsheet, there are cells organized in columns and rows. The rows are labeled numerically from 1, whereas each column is labeled *var*. Each column will be used to identify your variables, so *var* is short for *variable*. To label your variables with something other than *var,* you need to access the Variable View tab—this is a unique feature to SPSS.

## Variable View

When you click the Variable View tab, a new screen appears. Some features remain the same. For example, the menu bar and toolbar remain at the top of your screen. What changes is the spreadsheet. Notice that the rows are still labeled numerically beginning with 1. What changed are the labels across the columns. There are 11 columns in this view, as shown in Figure P.2: Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure, and Role. We will look at each column.

**FIGURE P.1** The Variable View Page With 11 Columns




Each column allows you to label and characterize variables.

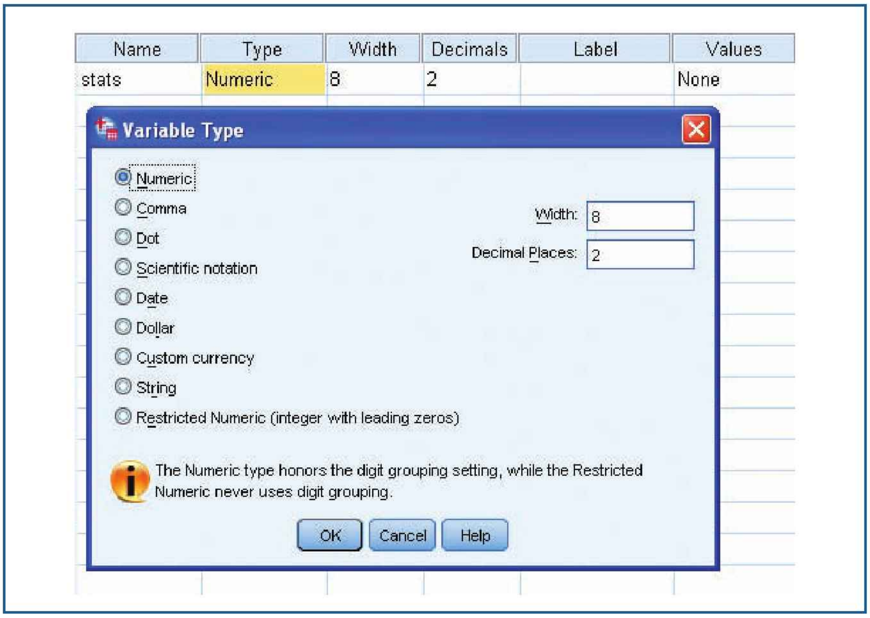
### *Name*

In this column, you enter the names of your variables (but no spaces are allowed). Each row identifies a single variable. Also, once you name your variable, the columns label in Data View will change. For example, while in Variable View, enter the word *stats* in the first cell of this column. Now click on the Data View tab at the bottom left. Notice that the label for Column 1 has now changed from *var* to *stats*. Also notice that once you enter a name for your variable, the row is suddenly filled in with words and numbers. Do not worry; this is supposed to happen.

### *Type*

This cell identifies the type of variable you are defining. When you click in the box, a small gray box with three dots appears. Click on the gray box and a dialog box appears, as shown in Figure P.3. By default, the variable type selected is numeric. This is because your variable will almost always be numeric, so we usually just leave this cell alone.

**FIGURE P.1** Variable Type Dialog Box




The dialog box shown here appears by clicking the small gray box with three dots in the Type column. This allows you to define the type of variable being measured.

### Width

The Width column is used to identify the largest number or longest string of your variable. For example, grade point average, or GPA, would have a width of 4: one digit to the left of the decimal, one space for the decimal, and two digits to the right. The default width is 8. So if none of your variables are longer than eight digits, you can just leave this alone. Otherwise, when you click in the box, you can select the up and down arrows that appear to the right of the cell to change the width.

### Decimals

This cell allows you to identify the number of places beyond the decimal point your variables are. As with the Width cell, when you click in the Decimals box, you can select the up and down arrows that appear to the right of the cell to change the decimals. If you want to enter whole numbers, for example, you can simply set this to 0.

### Label

The Label column allows you to label any variable whose meaning is not clear. For example, we can label the variable name *stats* as *statistics* in the label column, as shown in Figure P.4. This clarifies the meaning of the *stats* variable name.

**FIGURE P.1** Labeling Variables

| Name | Type | Width | Decimals | Label |
|---|---|---|---|---|
| stats | Numeric | 8 | 2 | statistics |

In this example, we labeled the variable name *stats* as *statistics* in the Label column.

### Values

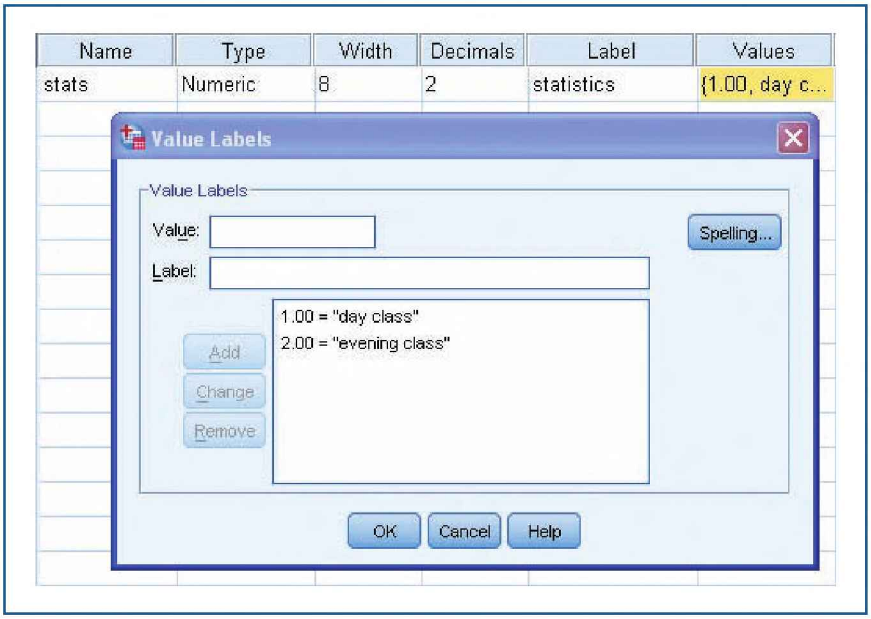
This column allows you to identify the levels of your variable. This is especially useful for coded data. Because SPSS recognizes numeric values, nominal data are often coded numerically in SPSS. For example, *gender* could be coded as *1 = male* and *2 = female; seasons* could be coded as *1 = spring, 2 = summer, 3 = fall,* and *4 = winter.*

Click on the small gray box with three dots to display a dialog box where we can label the variable, as shown in Figure P.5. We can label *day class* as *1* and *evening class* as *2* for our *stats* variable. To do this, enter *1* in the Value box and *day class* in the Label box; then click the Add option. Follow these same instructions for the *evening class* label. When both labels have been entered, click OK to finish.

### Missing

It is at times the case that some data researchers collect are missing. In these cases, you can enter a value that, when entered in the Data View tab, means

**FIGURE P.1** Value Labels Dialog Box




The dialog box shown here appears by clicking the small gray box with three dots in the Values column. This function allows you to code data that are not inherently numeric.

the data are missing. A common value used to represent missing data is *99*. To enter this value, click on the small gray box with three dots that appears to the right of the cell when you click in it. In the dialog box, it is most common to click on the second open circle and enter a *99* in the first cell. When this has been entered, click OK to finish. Now, whenever you enter *99* for that variable in the Data View spreadsheet, SPSS will recognize it as missing data.

### Columns

The Columns column lets you identify how much room to allow for your data and labels. For example, the *stats* label is five letters long. If you go to the Data View spreadsheet, you will see *stats* as the columns label. If you wrote *statisticscourse* in the Name column, then this would be too long—notice that this name continues on to a second line in the Data View columns label, because the columns default value is only 8. You can click the up and down arrows to increase or decrease how much room to allow for your columns label.

### Align

The Align column allows you to choose where to align the data you enter. You can change this by selecting the drop-down menu that appears by clicking in the cell. The alignment options are Left, Right, and Center. By default, numeric values are aligned to the right, and string values are aligned to the left.

### *Measure*

This column allows you to select the scale of measurement for the variable (scales of measurement are introduced in Chapter 1). By default, all variables are considered scale (i.e., an interval or ratio scale of measurement). If your variable is an ordinal or nominal variable, you can make this change by selecting the drop-down menu that appears by clicking in the cell.

### *Role*

The Role column is a column that SPSS has added in recent versions. The drop-down menu in the cell allows you to choose among the following commands: Input, Target, Both (Input and Target), None, Partition, and Split. Each of these options in the drop-down menu generally allows you to organize the entry and appearance of data in the Data View tab. While each option is valuable, these are generally needed for data sets that we will not work with in this book.

## Preview of SPSS in Focus

This book is unique in that you will learn how to use SPSS to perform statistical analyses as they are taught in this book. Most statistics textbooks for behavioral science omit such information, include it in an appendix separate from the main chapters in the book, include it at the end of chapters with no useful examples or context, or include it in ancillary materials that often are not included with course content. Instead, this book provides instructions for using SPSS in each chapter as statistical concepts are taught using practical research examples and screenshots to support student learning. You will find this instruction in the SPSS in Focus sections. These sections provide step-by-step instruction for how the concepts taught in each chapter can be applied to research problems using SPSS.

The reason for this inclusion is simple: Most researchers use some kind of statistical software to analyze statistics, and in behavioral science, the most common statistical software used by researchers is SPSS. This textbook brings statistics in research to the 21st century, giving you both the theoretical and computational instruction needed to understand how, when, and why you perform certain statistical analyses under different conditions and the technical instruction you need to succeed in the modern era of data collection, data entry, data analysis, and statistical interpretation using SPSS statistical software. This preface was written to familiarize you with this software. Subsequent SPSS in Focus sections will show you how to use SPSS to perform the applications and statistics taught in this book.

# Part I

## Introduction and Descriptive Statistics

iStock / thumb

# 1 Introduction to Statistics

## • • • Learning Objectives

After reading this chapter, you should be able to:

1. **Distinguish between descriptive and inferential statistics.**
2. **Explain how samples and populations, as well as a sample statistic and population parameter, differ.**
3. **Describe three research methods commonly used in behavioral science.**
4. **State the four scales of measurement and provide an example for each.**
5. **Distinguish between variables that are qualitative or quantitative.**
6. **Distinguish between variables that are discrete or continuous.**
7. **Enter data into SPSS by placing each group in separate columns and each group in a single column (coding is required).**

## 1.1 The Use of Statistics in Science

Why should you study statistics? The topic can be intimidating, and rarely does anyone tell you, "Oh, that's an easy course . . . take statistics!" **Statistics** is a branch of mathematics used to summarize, analyze, and interpret what we observe—to make sense or meaning of our observations. Really, statistics is used to make sense of the observations we make. For example, we can make sense of how good a soccer player is by observing how many goals he or she scores each season, and we can understand climates by looking at average temperature. We can also understand change by looking at the same statistics over time—such as the number of goals scored by a soccer player in each game, and the average temperature over many decades.

Statistics is commonly applied to evaluate scientific observations. Scientific observations are all around you. Whether you are making decisions about what to eat (based on health statistics) or how much to spend (based on the behavior of global markets), you are making decisions based on the statistical evaluation of scientific observations. Scientists who study human behavior gather information about all sorts of behavior of interest to them, such as information on addiction, happiness, worker productivity, resiliency, faith, child development, love, and more. The information that scientists gather is evaluated in two ways; each way reveals the two types of statistics taught in this book:

- Scientists organize and summarize information such that the information is meaningful to those who read about the observations scientists made in a study. This type of evaluation of information is called *descriptive statistics*.
- Scientists use information to answer a question (e.g., Is diet related to obesity?) or make an actionable decision (e.g., Should we implement a public policy change that can reduce obesity rates?). This type of evaluation of information is called *inferential statistics*.

This book describes how to apply and interpret both types of statistics in science and in practice to make you a more informed interpreter of the statistical information you encounter inside and outside of the classroom. For a review of statistical notation (e.g., summation notation) and a basic math review, please see Appendix A. The chapter organization of this book is such that descriptive statistics are described in Chapters 2–5 and applications for

**Statistics** is a branch of mathematics used to summarize, analyze, and interpret a group of numbers or observations.

### Chapter Outline

probability are further introduced in Chapters 6–7, to transition to a discussion of inferential statistics in the remainder of the book in Chapters 8–18.

The reason it is important to study statistics can be described by the words of Mark Twain: *There are lies, damned lies, and statistics.* He meant that statistics could be deceiving, and so can interpreting them. Statistics are all around you—from your college grade point average (GPA) to a *Newsweek* poll predicting which political candidate is likely to win an election. In each case, statistics are used to inform you. The challenge as you move into your careers is to be able to identify statistics and to interpret what they mean. Statistics are part of your everyday life, and they are subject to interpretation. The interpreter, of course, is *you.*

In many ways, statistics allow a story to be told. For example, your GPA may reflect the story of how well you are doing in school; the *Newsweek* poll may tell the story of which candidate is likely to win an election. In storytelling, there are many ways to tell a story. Similarly, in statistics, there are many ways to evaluate the information gathered in a study. For this reason, you will want to be a critical consumer of the information you come across, even information that is scientific. In this book, you will learn the fundamentals of statistical evaluation, which can help you to critically evaluate any information presented to you.

In this chapter, we begin by introducing the two general types of statistics identified here:

**FYI**

*Two types of statistics are descriptive statistics and inferential statistics.*

- Descriptive statistics: applying statistics to organize and summarize information
- Inferential statistics: applying statistics to interpret the meaning of information

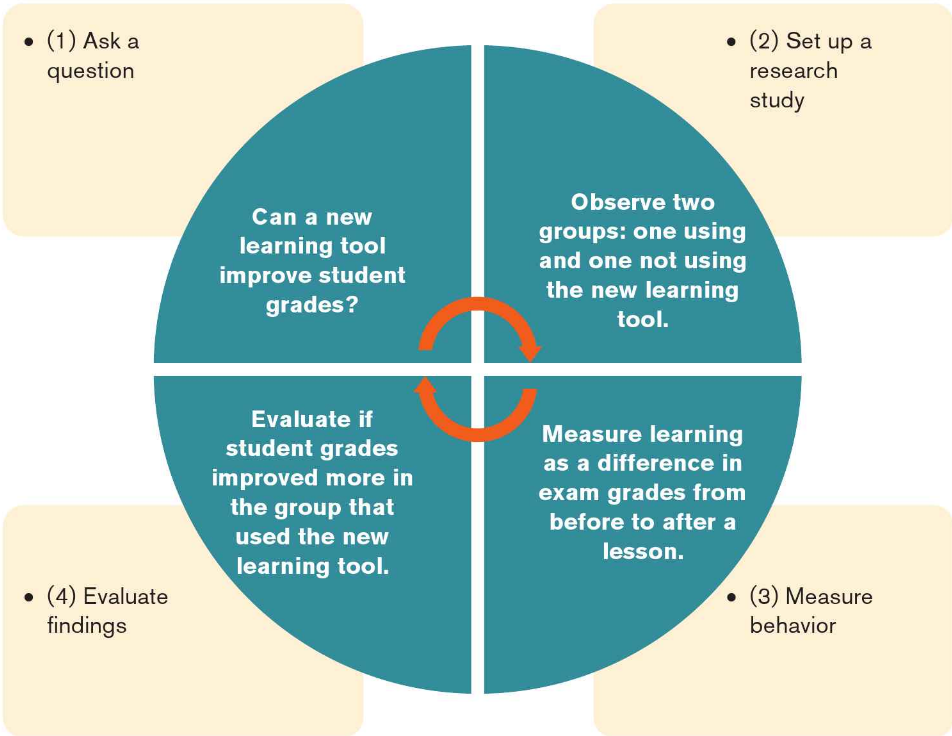
## 1.2 Descriptive and Inferential Statistics

The research process typically begins with a question or statement that can only be answered or addressed by making an observation. The observations researchers make are typically recorded as **data** (i.e., numeric values). To illustrate, Figure 1.1 describes the general structure for making scientific observations, using an example to illustrate. In the example, we suppose a researcher asks if a new learning tool can improve grades. To find the answer, the researcher first sets up a research study. In this example, the researcher creates two groups: one group of students who use the new learning tool, and a second group of students who do not use the new learning tool during a lesson. Now to observe learning, the researcher must measure data for learning. In this example, suppose she decides to measure the difference in exam grades from before and following the lesson. Exam grades, scored from 0 to 100, are the data in this example. If the new learning tool effectively improves grades, then we expect that grades will increase more in the group that uses the new learning tool compared to the group that does not.

**Data** (plural) are measurements or observations that are typically numeric. A **datum** (singular) is a single measurement or observation, usually referred to as a **score** or **raw score**.

In this section, we will introduce how descriptive and inferential statistics allow researchers to assess the data they measure in a research study, using the example given here and in Figure 1.1.

**FIGURE 1.1** General Structure for Making Scientific Observations




The general structure for making scientific observations, using an example for testing the effectiveness of a new learning tool to improve student grades.

## Descriptive Statistics

One way in which researchers can use statistics in research is to use procedures developed to help organize, summarize, and make sense of measurements or data. These procedures, called **descriptive statistics**, are typically used to quantify the behaviors researchers measure. In our example, *learning* could be described as the acquisition of knowledge. This certainly describes learning, but not numerically—in a way that allows us to measure learning. Instead, we stated that learning is the difference in exam scores from before and following a lesson. Here, we define learning as a value from 0 to 100; hence, learning can now be measured. If we observe hundreds of students, then the data in a spreadsheet will be overwhelming. Presenting a spreadsheet with the score for each individual student is not very clear. For this reason, researchers use descriptive statistics to summarize sets of individual measurements so they can be clearly presented and interpreted.

Data are generally presented in summary. Typically, this means that data are presented graphically, in tabular form (in tables), or as summary statistics (e.g., an average). For example, instead of listing each individual score or increase on an exam, we could summarize all scores by stating the

**FYI**

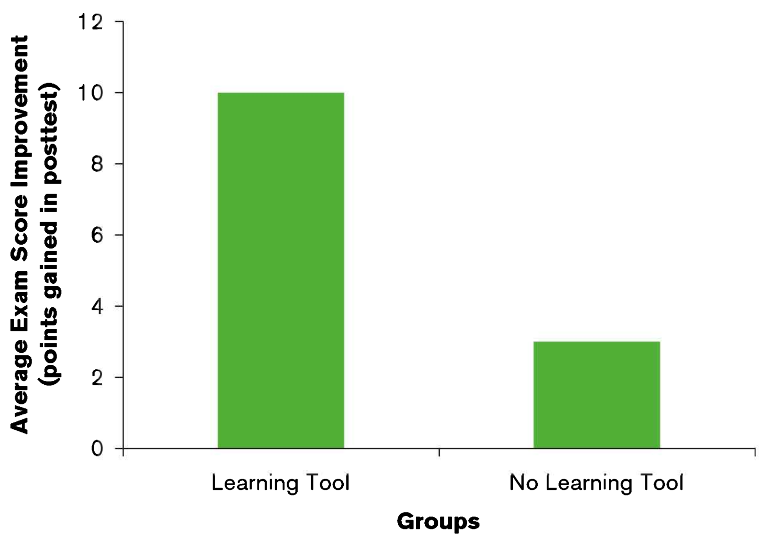
*Descriptive statistics summarize data to make sense or meaning of a list of numeric values.*

**Descriptive statistics** are procedures used to summarize, organize, and make sense of a set of scores or observations. Descriptive statistics are typically presented graphically, in tabular form (in tables), or as summary statistics (single values).

average (mean), middle (median), or most common (mode) score among all individuals, which can be more meaningful.

Tables and graphs serve a similar purpose to summarize large and small sets of data. One particular advantage of tables and graphs is that they can clarify findings in a research study. For example, to evaluate the findings for our study, we expect that average exam grades will improve more in the new learning tool group than in a group that does not use the new learning tool. Figure 1.2 displays these expected findings. Notice how summarizing the average improvement in each group in a figure can clarify research findings.

**FIGURE 1.2** Summary of Expected Findings




A graphical summary of the expected findings if the new learning tool effectively improves student grades.

## Inferential Statistics

Most research studies include only a select group of participants, not all participants who are members of a particular group of interest. In other words, most scientists have limited access to the phenomena they study, especially behavioral phenomena. Hence, researchers select a portion of all members of a group (the *sample*) mostly because they do not have access to all members of a group (the *population*). Imagine, for example, trying to identify every person who has experienced exam anxiety. The same is true for most behaviors—the population of all people who exhibit those behaviors is likely too large. Because it is often not possible to identify all individuals in a population, researchers require statistical procedures, called **inferential statistics**, to infer that observations made with a sample are also likely to be observed in the larger population from which the sample was selected.

**Inferential statistics** are procedures used that allow researchers to infer or generalize observations made with samples to the larger population from which they were selected.

To illustrate, we can continue with the new learning tool study. If we are interested in students in the United States, then all U.S. students would constitute the **population**. Specifically, we want to test if a new learning tool can improve learning in this population; this characteristic (learning) in the population is called a **population parameter**. Learning, then, is the characteristic we will measure, but not in the population. The characteristics of interest are typically descriptive statistics. In the new learning tool study, the characteristic of interest is learning, which was measured as the difference in exam scores (scored from 0 to 100) from before and following a lesson. In practice, however, researchers will not have access to the entire population of U.S. students. They simply do not have the time, money, or other resources to even consider studying all students enrolled in college.

An alternative to selecting all members of a population is to select a portion or **sample** of individuals in the population. Selecting a sample is more practical, and most scientific research is based upon findings in samples, not populations. In our example, we can select any portion of students from the larger population of all students in the United States; the portion of students we select will constitute our sample. A characteristic that describes a sample, such as learning, is called a **sample statistic** and is the value that is measured in the study. A sample statistic is measured to estimate the population parameter. In this way, a sample is selected from a population to learn more about the characteristics in the population of interest.

**FYI**

*Inferential statistics are used to help the researcher infer how well statistics in a sample reflect parameters in a population.*

A **population** is the set of all individuals, items, or data of interest. This is the group about which scientists will generalize.

A characteristic (usually numeric) that describes a population is called a **population parameter.**

A **sample** is a set of individuals, items, or data selected from a population of interest.

A characteristic (usually numeric) that describes a sample is referred to as a **sample statistic.**

## MAKING SENSE POPULATIONS AND SAMPLES

A population is identified as any group of interest, whether that group is all students worldwide or all students in a professor's class. Think of any group you are interested in. Maybe you want to understand why college students join fraternities and sororities. So students who join fraternities and sororities is the group you are interested in. Hence, to you, this group is a population of interest. You identified a population of interest just as researchers identify populations they are interested in.

Remember that researchers select samples only because they do not have access to all individuals in a population. Imagine having to identify every person who has fallen in love, experienced anxiety, been attracted to someone else, suffered with depression, or taken a college exam. It is ridiculous to consider that we can identify all individuals in such populations. So researchers use data gathered from samples (a portion of individuals from the population) to make inferences concerning a population.

To make sense of this, suppose you want to get an idea of how people in general feel about a new pair of shoes you just bought. To find out, you put your new shoes on and ask 20 people at random throughout the day whether or not they like the shoes. Now, do you really care about the opinion of only those 20 people you asked? Not really—you actually care more about the opinion of people in general. In other words, you only asked the 20 people (your sample) to get an idea of the opinions of people in general (the population of interest). Sampling from populations follows a similar logic.

Example 1.1 applies the process of sampling to distinguish between a sample and a population.

## Example 1.1

On the basis of the following example, we will identify the population, sample, population parameter, and sample statistic: Suppose you read an article in the local college newspaper citing that the average college student plays 2 hours of video games per week. To test whether this is true for your school, you randomly approach 20 fellow students and ask them how long (in hours) they play video games per week. You find that the average student, among those you asked, plays video games for 1 hour per week. Distinguish the population from the sample.

In this example, all college students at your school constitute the population of interest, and the 20 students you approached is the sample that was selected from this population of interest. Because it is purported that the average college student plays 2 hours of video games per week, this is the population parameter (2 hours). The average number of hours playing video games in the sample is the sample statistic (1 hour).

## LEARNING CHECK 1

1. ______________ are procedures used to summarize, organize, and make sense of a set of scores or observations.
2. ___________ describe(s) characteristics in a population, whereas ______________ describe(s) characteristics in a sample.
   a. Statistics; parameters
   b. Parameters; statistics
   c. Descriptive; inferential
   d. Inferential; descriptive
3. A psychologist wants to study a small population of 40 students in a local private school. If the researcher was interested in selecting the entire population of students for this study, then how many students must the psychologist include?
   a. None, because it is not possible to study an entire population in this case.
   b. At least half, because this would constitute the majority of the population.
   c. All 40 students, because all students constitute the population.
4. True or false: Inferential statistics are used to help the researcher *infer* the unknown parameters in a given population.

Answers: 1. Descriptive statistics; 2. b; 3. c; 4. True.

## 1.3 Research Methods and Statistics

**Science** is the study of phenomena, such as behavior, through strict observation, evaluation, interpretation, and theoretical explanation.

The **research method**, or **scientific method**, is a set of systematic techniques used to acquire, modify, and integrate knowledge concerning observable and measurable phenomena.

This book will describe many ways of measuring and interpreting data. Yet, simply collecting data does not make you a scientist. To engage in science, you must follow specific procedures for collecting data. Think of this as playing a game. Without the rules and procedures for playing, the game itself would be lost. The same is true in science; without the rules and procedures for collecting data, the ability to draw scientific conclusions would be lost. Ultimately, statistics are often used in the context of **science**. In the behavioral sciences, *science* is specifically applied using the **research method**. To use the research method, we make observations using systematic techniques of scientific inquiry. In this section, we introduce three research methods. Because these methods are often used together with statistical analysis, it is necessary to introduce them.

Can Stock Photo Inc. / iofoto

To illustrate the basic premise of engaging in science, suppose you come across the following problem first noted by the famous psychologist Edward Thorndike in 1898:

> Dogs get lost hundreds of times and no one ever notices it or sends an account of it to a scientific magazine, but let one find his way from Brooklyn to Yonkers and the fact immediately becomes a circulating anecdote. Thousands of cats on thousands of occasions sit helplessly yowling, and no one takes thought of it or writes to his friend, the professor; but let one cat claw at the knob of a door supposedly as a signal to be let out, and straightway this cat becomes the representative of the cat-mind in all books. . . . In short, the anecdotes give really . . . supernormal psychology of animals. (pp. 4–5)

Here the problem was to determine the animal mind. Thorndike posed the question of whether animals were truly smart, based on the many observations he made. This is where the scientific process typically begins: with a question. To answer questions in a scientific manner, researchers need more than just statistics; they need a set of strict procedures for making the observations and measurements. In this section, we introduce three research methods commonly used in behavioral research: experimental, quasi-experimental, and correlational methods. Each method involves examining the relationship between variables, and is introduced here because we will apply these methods throughout the book.

## Experimental Method

Any study that demonstrates cause is called an **experiment**. To demonstrate cause, though, an experiment must follow strict procedures to ensure that all other possible causes are eliminated or highly unlikely. Hence, researchers must control the conditions under which observations are made in order to isolate cause-and-effect relationships between variables. Figure 1.3 shows the general structure of an experiment using an example to illustrate the structure that is described.

The experiment illustrated in Figure 1.3 was designed to determine the effect of distraction on student test scores. A sample of students was selected from a population of all undergraduates. In one group, the professor sat quietly while students took the exam (low-distraction group); in the other group, the professor rattled papers, tapped her foot, and made other sounds during the exam (high-distraction group). Exam scores in both groups were measured and compared.

An **experiment** is the use of methods and procedures to make observations in which the researcher fully controls the conditions and experiences of participants by applying three required elements of control (manipulation, randomization, and comparison/control) to isolate cause-and-effect relationships between variables.

For this study to be called an experiment, researchers must satisfy three requirements. These requirements are regarded as the necessary steps to