STATISTICS for the BEHAVIORAL SCIENCES

THIRD EDITION

GREGORY J. PRIVITERA

Brief Summary of Statistical Notation and Formulas

CENTRAL TENDENCY (CHAPTER 3)

$\mu = \frac{\Sigma x}{N}$ (Population mean)

$M = \frac{\Sigma x}{n}$ (Sample mean)

VARIABILITY (CHAPTER 4)

$\sigma^2 = \frac{SS}{N}$ (Population variance)

$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{SS}{N}}$ (Population standard deviation)

$s^2 = \frac{SS}{n-1} = \frac{SS}{df}$ (Sample variance)

$s = \sqrt{s^2} = \sqrt{\frac{SS}{n-1}} = \sqrt{\frac{SS}{df}}$ (Sample standard deviation)

z TRANSFORMATIONS AND STANDARD ERROR (CHAPTERS 6 AND 7)

$z = \frac{x - \mu}{\sigma}$ (*z* transformation for a population of scores)

$z = \frac{x - M}{SD}$ (*z* transformation for a sample of scores)

$\sigma_M = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$ (Standard error of the mean)

THE *z* TEST (CHAPTER 8)

$z_{obt} = \frac{M-\mu}{\sigma_M}$ (Test statistic for a one-sample *z* test)

$d = \frac{M-\mu}{\sigma}$ (Cohen's *d* effect size measure for the *z* test)

THE *t* TESTS (CHAPTERS 9 AND 10)

One-Sample *t*

$t_{obt} = \frac{M-\mu}{s_M}$ (Test statistic for the one-sample *t* test)

Two-Independent-Sample *t*

$t_{obt} = \frac{(M_1-M_2)-(\mu_1-\mu_2)}{s_{M_1-M_2}}$ (Test statistic for the two-independent-sample *t* test)

Related-Samples *t*

$t_{obt} = \frac{M_D-\mu_D}{s_{MD}}$ (Test statistic for the related-samples *t* test)

Effect Size

$d = \frac{M-\mu}{SD}$ (Estimated Cohen's *d* for one-sample *t* test)

$\frac{M_1-M_2}{\sqrt{s_p^2}}$ (Estimated Cohen's *d* for two-independent-sample *t* test)

$d = \frac{M_D}{s_D}$ (Estimated Cohen's *d* for related-samples *t* test)

ONE-WAY BETWEEN-SUBJECTS ANALYSIS OF VARIANCE (CHAPTER 12)

Between-Subjects Design

$F_{obt} = \frac{MS_{BG}}{MS_E}$ (Test statistic for the one-way between-subjects ANOVA)

Effect Size (Between-Subjects Design)

$R^2 = \eta^2 = \frac{SS_{BG}}{SS_T}$ (Eta-squared estimate for proportion of variance)

$\omega^2 = \frac{SS_{BG} - df_{BG}(MS_E)}{SS_T + MS_E}$ (Omega-squared estimate for proportion of variance)

ONE-WAY WITHIN-SUBJECTS ANALYSIS OF VARIANCE (CHAPTER 13)

Within-Subjects Design

$F_{obt} = \frac{MS_{BG}}{MS_E}$ (Test statistic for the one-way within-subjects ANOVA)

Effect Size (Within-Subjects Design)

$\eta_P^2 = \frac{SS_{BG}}{SS_T - SS_{BP}}$ (Partial eta-squared for proportion of variance)

$\omega_P^2 = \frac{SS_{BG} - df_{BG}(MS_E)}{(SS_T - SS_{BP}) + MS_E}$ (Partial omega-squared for proportion of variance)

TWO FACTOR ANALYSIS OF VARIANCE (CHAPTER 14)

$F_A = \frac{MS_A}{MS_E}$ (Test statistic for the main effect on factor A)

$F_B = \frac{MS_B}{MS_E}$ (Test statistic for the main effect on factor B)

$F_{A \times B} = \frac{MS_{A \times B}}{MS_E}$ (Test statistic for the A × B interaction)

CORRELATION AND REGRESSION (CHAPTERS 15 AND 16)

Correlation Coefficient

$r = \frac{SS_{XY}}{\sqrt{SS_X SS_Y}}$ (Pearson correlation coefficient)

Analysis of Regression

$F_{obt} = \frac{MS_{regression}}{MS_{residual}}$ (Test statistic for analysis of regression and analysis of multiple regression)

CHI-SQUARE TESTS (CHAPTER 17)

One-Way and Two-Way Chi-Square Tests

$\chi^2_{obt} = \Sigma \frac{(f_o - f_e)^2}{f_e}$ (Test statistic for the chi-square goodness-of-fit test and the chi-square test for independence)

Effect Size (Test for Independence)

$V = \sqrt{\frac{\chi^2}{n \times df_{smaller}}}$ (Cramer's *V* effect size estimate)

TESTS FOR ORDINAL DATA (CHAPTER 18)

The Sign Test

$z = \frac{x - np}{\sqrt{np(1-p)}}$ (Test statistic for the normal approximation for the sign test)

Wilcoxon Signed-Ranks *T* Test

$z = \frac{T - \mu_T}{\sigma_T}$ (Test statistic for the normal approximation of the Wilcoxon *T*)

Mann-Whitney *U* Test

$z = \frac{U - \mu_U}{\sigma_U}$ (Test statistic for the normal approximation of the Mann-Whitney *U*)

The Kruskal-Wallis *H* Test

$H = \frac{12}{N(N+1)}\left(\Sigma \frac{R^2}{n}\right) - 3(N+1)$ (Test statistic for the Kruskal-Wallis *H* test)

The Friedman Test

$\chi^2_R = \frac{12}{nk(k+1)} \sum R^2 - 3n(k+1)$ (Test statistic for the Friedman test)

STATISTICS FOR THE BEHAVIORAL SCIENCES

THIRD EDITION

The engaging **Third Edition** of **Statistics for the Behavioral Sciences** shows students that statistics can be understandable, interesting, and relevant to their daily lives. Using a conversational tone, award-winning teacher and author Gregory J. Privitera speaks to the reader as researcher when covering statistical theory, computation, and application. Robust pedagogy allows students to continually check their comprehension and hone their skills while working through carefully developed problems and exercises that include current research and seamless integration of IBM® SPSS®. This edition not only prepares students to be lab-ready but also gives them the confidence to use statistics to summarize data and make decisions about behavior.

Real-world examples make statistics relevant for students.

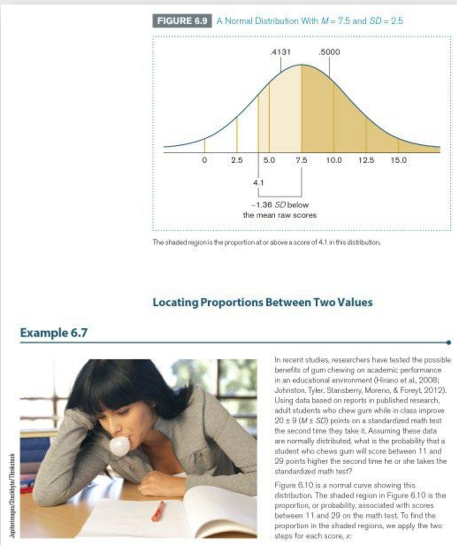


FIGURE 6.9 A Normal Distribution With $M = 7.5$ and $SD = 2.5$

The shaded region is the proportion at or above a score of 4.1 in this distribution.

Locating Proportions Between Two Values

Example 6.7

Jupiterimages/Stockbyte/Thinkstock

In recent studies, researchers have tested the possible benefits of gum chewing on academic performance in an educational environment (Hirano et al., 2008; Johnston, Tyler, Stansberry, Moreno, & Foreyt, 2012). Using data based on reports in published research, adult students who chew gum while in class improve 20 ± 9 ($M \pm SD$) points on a standardized math test the second time they take it. Assuming these data are normally distributed, what is the probability that a student who chews gum will score between 11 and 29 points higher the second time he or she takes the standardized math test?

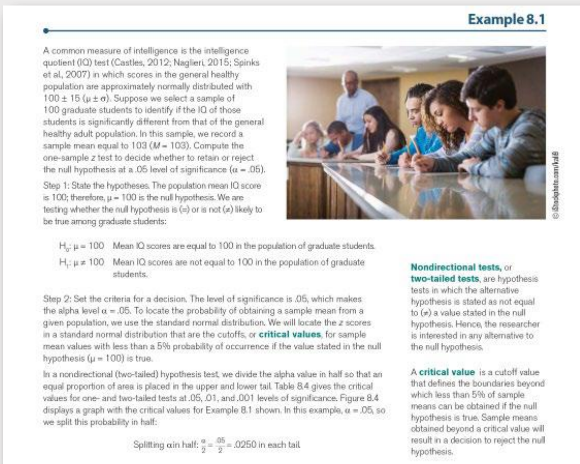
Figure 6.10 is a normal curve showing this distribution. The shaded region in Figure 6.10 is the proportion, or probability, associated with scores between 11 and 29 on the math test. To find the proportion in the shaded regions, we apply the two steps for each score, x:

"Really drives the point home that statistical methods are an integral part of psychology research…This text does a much better job of making this connection than any other text I've seen."

—Scott D. Bradshaw, *Elizabeth City State University*

"Most impressive is [the book's] relevance to the current state of the field (e.g., introduction to multiple regression)."

—Alexander O. Crenshaw, *University of Utah*



Example 8.1

A common measure of intelligence is the intelligence quotient (IQ) test (Castles, 2012; Naglieri, 2015; Spinks et al., 2007) in which scores in the general healthy population are approximately normally distributed with 100 ± 15 ($\mu \pm \sigma$). Suppose we select a sample of 100 graduate students to identify if the IQ of those students is significantly different from that of the general healthy adult population. In this sample, we record a sample mean equal to 103 ($M = 103$). Compute the one-sample z test to decide whether to retain or reject the null hypothesis at a .05 level of significance ($\alpha = .05$).

© iStockphoto.com/kali9

Step 1: State the hypotheses. The population mean IQ score is 100; therefore, $\mu = 100$ is the null hypothesis. We are testing whether the null hypothesis is (=) or is not (≠) likely to be true among graduate students:

H_0: $\mu = 100$ Mean IQ scores are equal to 100 in the population of graduate students.

H_1: $\mu \neq 100$ Mean IQ scores are not equal to 100 in the population of graduate students.

Step 2: Set the criteria for a decision. The level of significance is .05, which makes the alpha level $\alpha = .05$. To locate the probability of obtaining a sample mean from a given population, we use the standard normal distribution. We will locate the z scores in a standard normal distribution that are the cutoffs, or **critical values**, for sample mean values with less than a 5% probability of occurrence if the value stated in the null hypothesis ($\mu = 100$) is true.

In a nondirectional (two-tailed) hypothesis test, we divide the alpha value in half so that an equal proportion of area is placed in the upper and lower tail. Table 8.4 gives the critical values for one- and two-tailed tests at .05, .01, and .001 levels of significance. Figure 8.4 displays a graph with the critical values for Example 8.1 shown. In this example, $\alpha = .05$, so we split this probability in half:

$$\text{Splitting } \alpha \text{ in half: } \frac{\alpha}{2} = \frac{.05}{2} = .0250 \text{ in each tail}$$

Nondirectional tests, or **two-tailed tests**, are hypothesis tests in which the alternative hypothesis is stated as not equal to (≠) a value stated in the null hypothesis. Hence, the researcher is interested in any alternative to the null hypothesis.

A **critical value** is a cutoff value that defines the boundaries beyond which less than 5% of sample means can be obtained if the null hypothesis is true. Sample means obtained beyond a critical value will result in a decision to reject the null hypothesis.

A FOCUS ON CLARITY

SPSS® in Focus sections provide step-by-step, classroom-tested instruction using practical research examples of how chapter concepts can be applied using SPSS®. All SPSS® screenshots were updated for version 24 of SPSS®. A new Appendix B connects each SPSS® in Focus section to the page number where it can be found in the book and provides a general instruction guide for using SPSS®.

4.11 SPSS in Focus: Range, Variance, and Standard Deviation

In the SPSS in Focus section of Chapter 3, we computed the mean, the median, and the mode for creativity, measured as the time (in seconds) it took participants to list 10 uses for a paper clip. The data for this example, originally given in Table 3.11, are reproduced here in Table 4.10. In this section, we will use these same data to compute the range, variance, and standard deviation.

1. Click on the Variable View tab and enter *creativity* in the Name column. We will enter whole numbers, so reduce the value in the Decimals column to 0.
2. Click on the Data View tab and enter the 20 values in the *creativity* column.
3. Go to the menu bar and click Analyze, then Descriptive Statistics and Frequencies, to display a dialog box.
4. When you select the *creativity* variable and click the arrow in the center of the dialog box, *creativity* will move into the box labeled Variable(s): to the right. Make sure the option to display frequency tables is not selected, and then click on Statistics to bring up another dialog box.
5. In this dialog box, select Std. Deviation, Variance, and Range; then select Continue.
6. Select OK, or select Paste and click the Run command. Table 4.11 shows the SPSS output table.

© Can Stock Photo Inc. /csp18310594

TABLE 4.10 The Time (in seconds) That It Took a Group of 20 Participants to Write Down 10 Uses for a Paper Clip on a Creativity Test

41	80
65	80
123	64
46	59
48	51
87	36
38	80
90	143
132	122
115	100

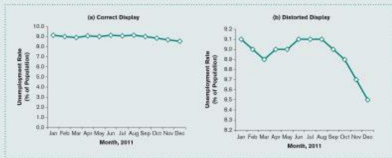
This table originally appeared as Table 3.11 in Chapter 3.

MAKING SENSE ETHICS IN FOCUS: DECEPTION DUE TO THE DISTORTION OF DATA

It was Mark Twain who once said *there are lies, damned lies, and statistics*. His statement identified that statistics can be deceiving—and so can interpreting them. Descriptive statistics are used to inform us. Therefore, being able to identify statistics and correctly interpret what they mean is an important part of the research process. Presenting data can be an ethical concern when the data are distorted in any way, whether on accident or intentionally. The distortion of data can occur for data presented graphically or as summary statistics. Here, we will describe how the presentation of data can be distorted.

When a graph is distorted, it can deceive the reader into thinking differences exist, when in truth differences are negligible (Frankfort-Nachmias & Leon-Guerrero, 2006; Privitera 2016b). Three common distortions to look for in graphs are (1) displays with an unlabeled axis, (2) displays with one axis altered in relation to the other axis, and (3) displays in which the vertical axis (*y*-axis) does not begin with 0. As an example of how a graphical display can be distorted, Figure 2.13 displays a frequency polygon for U.S. unemployment rates in 2011. Figure 2.13a displays the data correctly with the *y*-axis starting at 0%; Figure 2.13b displays the same data with the *y*-axis distorted and beginning at 8.2%. When the graph is distorted in this way, it can make the slope of the line appear steeper as if unemployment rates are significantly declining, although it is clear from Figure 2.13a that this is not the case—in fact, U.S. unemployment rates were rather stable in 2011. To avoid misleading or deceiving readers, pay attention to how data are displayed in graphs to make sure that the data are accurately and appropriately presented.

FIGURE 2.13 Two Graphical Displays for the Same Data



(a) This graph is a correct display and (b) this is a display that is distorted because the *y*-axis does not begin at 0%. Data are of actual unemployment rates in the United States in 2011 (U.S. Bureau of Labor Statistics, 2013).

Source: Privitera, G. J. (2017). *Research methods for the behavioral sciences* (2nd ed.). Thousand Oaks, CA: SAGE. Reprinted with permission from SAGE.

(Continued)

Making Sense sections break down the statistical concepts students typically find most challenging.

1.6 RESEARCH IN FOCUS: EVALUATING DATA AND SCALES OF MEASUREMENT

While qualitative variables are often measured in behavioral research, this book will focus largely on quantitative variables. The reason is twofold: (1) Quantitative measures are more common in behavioral research, and (2) most statistical tests taught in this book are adapted for quantitative measures. Indeed, many researchers who measure qualitative variables will also measure those that are quantitative in the same study.

© iStockphoto.com/PeopleImages

For example, Jones, Blackey, Fitzgibbon, and Chew (2010) explored the costs and benefits of social networking among college students. The researchers used a qualitative method to interview each student in their sample. In the interview, students could respond openly to questions asked during the interview. These researchers then summarized responses into categories related to learning, studying, and social life. For example, the following student response was categorized as an example of independent learning experience for employability: "I think it [social software] can be beneficial . . . in the real working environment" (Jones et al., 2010, p. 780).

The limitation for this analysis is that categories are on a nominal scale (the least informative scale). So many researchers who record qualitative data also use some quantitative measures. For example, researchers in this study also asked students to rate their usage of a variety of social software technologies, such as PowerPoint and personal websites, on a scale from 1 (*never*)

Research in Focus sections in Chapters 1–6 and 8 provide context by reviewing pertinent, current research that clarifies or illustrates important statistical concepts discussed in the chapter.

11.11 APA IN FOCUS: REPORTING CONFIDENCE INTERVALS

To report the results of estimation in scientific journals, we state the level of confidence, point estimate, and interval estimate of each confidence interval. For example, we could report the confidence interval for the one-sample *t* test in Example 11.2 as follows:

> Social functioning was worse among relatives of patients with OCD (M = 62.00) than in the general healthy population (95% confidence interval [CI] 51.58–72.42).

We can also report the point estimate with the interval estimate given in brackets. For example, we could report the confidence interval for the related-samples *t* test in Example 11.4 as follows:

> In a 6-minute reading session, children read longer in the presence versus the absence of a teacher (mean [95% CI] = 15.00 [2.35–27.65]).

These examples provide a concise summary of two of the confidence intervals we computed in this chapter. In addition, we can summarize the sample means, the standard error, and the standard deviations in a figure or table or in the main text.

APA in Focus sections explain how to summarize statistical results for each inferential statistic taught and how to read and report statistical results in research journals that follow APA style.

A FOCUS ON PEDAGOGY AND PRACTICE

2 Summarizing Data
Frequency Distributions in Tables and Graphs

••• **Learning Objectives**

After reading this chapter, you should be able to:

1. **Construct a simple frequency distribution for grouped and ungrouped data.**
2. **Determine whether data should be grouped or ungrouped.**
3. **Identify when it is appropriate to distribute the cumulative frequency, relative frequency, relative percent, cumulative relative frequency, and cumulative percent.**
4. **Identify percentile points and percentile ranks in a cumulative percent distribution.**
5. **Construct and interpret graphs for distributions of continuous data.**
6. **Construct and interpret graphs for distributions of discrete data.**
7. **Construct frequency distributions for quantitative and categorical data using SPSS.**
8. **Construct histograms, bar charts, and pie charts using SPSS.**

Updated Learning Objectives and Chapter Summaries improve chapter organization and help students retain important information.

••• CHAPTER SUMMARY

LO 1–2: **Construct a simple frequency distribution for grouped and ungrouped data; determine whether data should be grouped or ungrouped.**

- A frequency distribution is a summary display for a distribution of data organized or summarized in terms of how often or frequently scores occur.
- A simple frequency distribution for grouped data displays the frequency of data in intervals. Each interval is equidistant, no interval overlaps, and the degree of accuracy for each interval is the same as in the original data. This distribution can be constructed using three steps:

  Step 1: Find the real range.

  Step 2: Find the interval width.

  Step 3: Construct the frequency distribution.
- A simple frequency distribution for ungrouped data displays the frequency of categories or whole units

Learning Checks with answers support a deeper understanding of the material.

LEARNING CHECK 1

1. What is the main limitation of the z test in behavioral research?
2. The ______________ is a normal-like distribution with greater variability in the tails than a normal distribution because the sample variance is substituted for the population variance to estimate the standard error in this distribution.
3. How does increasing sample size affect the estimate of the population variance?
4. What is the calculation for the degrees of freedom of a t distribution?
5. What are the degrees of freedom for each of the following samples?
   (a) $n = 12$ (b) $n = 22$ (c) $n = 5$ (d) $n = 30$
6. Assuming $\alpha = .05$, what is the critical value in each of the following t distributions?
   (a) $df = 10$, two-tailed test
   (b) $df = 30$, one-tailed test (lower tail critical)
   (c) $df = \infty$, two-tailed test
   (d) $df = 15$, one-tailed test (upper tail critical)

Answers: 1. It requires that the population variance is known, and behavioral researchers rarely know the value of the population variance; 2. t distribution; 3. As sample size increases, the sample variance more closely approximates the population variance; 4. $n - 1$; 5. (a) $df = 11$, (b) $df = 21$, (c) $df = 4$, (d) $df = 29$; 6. (a) ±2.228, (b) −1.697, (c) ±1.96, (d) 1.753.

Do you cover R, SAS®, or Stata® in your statistics course?

Statistics for the Behavioral Sciences, Third Edition can be packaged with an alternate, companion guide, **Statistical Analysis: In Focus**, that aligns coverage of R, SAS®, and Stata® with all SPSS® components within the textbook. These alternate versions of Privitera's "In Focus" coverage plug into each chapter of the text in context to offer helpful replacement coverage for students not using SPSS® software. (ISBN: 9781544305608)

ALSO AVAILABLE WITH THE TEXT:

- **WebAssign**, a powerful tool for creating online, autograded homework specific to the text
- **An updated Student Study Guide** that provides additional opportunity for review, practice, and mastery of concepts

SAGE coursepacks

Our Content Tailored to Your LMS

Instructors! **SAGE coursepacks** makes it easy to import our quality instructor and student resource content into your school's learning management system (LMS). Intuitive and simple to use, **SAGE coursepacks** allows you to customize course content to meet your students' needs. Learn more at **sagepub.com/coursepacks**.

STATISTICS for the BEHAVIORAL SCIENCES

THIRD EDITION

Sara Miller McCune founded SAGE Publishing in 1965 to support the dissemination of usable knowledge and educate a global community. SAGE publishes more than 1000 journals and over 800 new books each year, spanning a wide range of subject areas. Our growing selection of library products includes archives, data, case studies and video. SAGE remains majority owned by our founder and after her lifetime will become owned by a charitable trust that secures the company's continued independence.

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne

STATISTICS
for the BEHAVIORAL SCIENCES

THIRD EDITION

GREGORY J. PRIVITERA

St. Bonaventure University

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

SAGE

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.
3 Church Street
#10-04 Samsung Hub
Singapore 049483

Acquisitions Editor: Lara Parra
Development Editor: Lucy Berbeo
Editorial Assistant: Zachary Valladon
eLearning Editor: Morgan Shannon
Production Editor: Kelly DeRosa
Copy Editor: Christina West
Typesetter: C&M Digitals (P) Ltd.
Proofreader: Theresa Kay
Indexer: Jean Casalegno
Cover Designer: Michael Dubowe
Marketing Manager: Katherine Hepburn

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Names: Privitera, Gregory J., author.

Title: Statistics for the behavioral sciences / Gregory J. Privitera, St. Bonaventure University.

Description: Third Edition. | Thousand Oaks : SAGE Publications, [2017] | Revised edition of the author's Statistics for the behavioral sciences, [2015] | Includes bibliographical references and index.

Identifiers: LCCN 2017011485 | ISBN 9781506386256 (hardcover : alk. paper)

Subjects: LCSH: Social sciences—Statistical methods. | Psychology—Statistical methods.

Classification: LCC HA29.P755 2017 | DDC 519.5—dc23
LC record available at https://lccn.loc.gov/2017011485

This book is printed on acid-free paper.

17 18 19 20 21 10 9 8 7 6 5 4 3 2 1

Brief Contents

PART IV. MAKING INFERENCES ABOUT THE VARIABILITY OF TWO OR MORE MEANS 363

PART V. MAKING INFERENCES ABOUT PATTERNS, FREQUENCIES, AND ORDINAL DATA 487

Detailed Contents

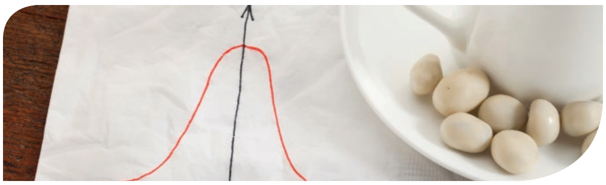
iStock / thumb

iStockphoto.com/Rawpixel

©Can Stock Photo Inc./OG_vision

©iStockphoto.com/Davizro

©iStockphoto.com/Bennewitz



Marek Uliasz/iStock/Thinkstock

iStockphoto.com/JohnnyGreig

cosmin4000/iStock/Thinkstock

©iStockphoto.com/nastinka

Chapter 9. Testing Means: One-Sample and Two-Independent-Sample *t* Tests 274

©iStockphoto.com/Shaiith

Chapter 10. Testing Means: The Related-Samples *t* Test 306

©iStockphoto.com/Ultima_Gaina

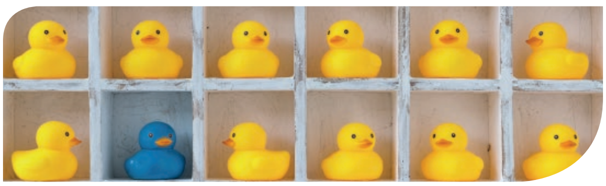
© iStockphoto.com/Borut Trdina

©iStockphoto.com/PeopleImages

©iStockphoto.com/vvmich

©iStockphoto.com/Kalawin

iStockphoto.com/diane39



About the Author

Gregory J. Privitera is a professor and chair of the Department of Psychology at St. Bonaventure University, where he is a recipient of their highest teaching honor, the Award for Professional Excellence in Teaching, and their highest honor for scholarship, the Award for Professional Excellence in Research and Publication. Dr. Privitera received his PhD in behavioral neuroscience in the field of psychology at the State University of New York at Buffalo and continued to complete postdoctoral research at Arizona State University. He is an author of multiple books on statistics, research methods, and health psychology, in addition to authoring more than three dozen peer-reviewed scientific articles aimed at advancing our understanding of health and well-being. He oversees a variety of undergraduate research projects at St. Bonaventure University, where dozens of 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 academic and research advisement, Dr. Privitera was honored as Advisor of the Year by St. Bonaventure University in 2013. In addition, he is the award-winning author of *Research Methods for the Behavioral Sciences*, for which he received the Most Promising New Textbook Award from the Textbook & Academic Authors Association in 2014. 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 Ann, and a son, Aiden Andrew.

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.

Jeff Adams, *Trent University*

Melissa S. Atkins, *Marshall University*

Josh Averbeck, *Western Illinois University*

Kristen T. Begosh, *University of Delaware*

Chloe G. Bland, *College of Saint Elizabeth*

Julia J. C. Blau, *SUNY College at Oneonta*

Scott D. Bradshaw, *Elizabeth City State University*

Shannon Claxton, *Morningside College*

Alexander O. Crenshaw, *University of Utah*

Barry D. Davis, *South Florida Bible College & Theological Seminary*

Diana M. Elliot, *Belmont Abbey College*

Robert Franklin, *Anderson University*

Steven M. Specht, *Utica College*

Sarah Strand, *California State University, Sacramento*

Whitney Alicia Zimmerman, *The Pennsylvania State University*

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 IBM® SPSS® Statistics* 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.

* SPSS is a registered trademark of International Business Machines Corporation.

- **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 and 8 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. Screen-shot figures and explanations provide support for how to read SPSS outputs. In Appendix B, a guide for using SPSS is given for each SPSS in Focus section in the book, with page number references given to make it simple for students to find where those SPSS sections are taught in the book.

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.

Appendix B provides a general instructions guide for using SPSS. Throughout this book, these instructions are provided with an example for how to analyze and interpret data. However, it would be difficult for students to thumb through the book to find each test when needing to refer to these tests later. Therefore, this appendix provides a single place where students can go to get directions for any statistical test taught in this chapter. Each instruction also provides the location within the text where readers can find an example of how to compute each test using SPSS.

Appendix C gives the tables needed to find critical values for the test statistics taught in this book.

Appendix D 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 the appendix.

New to This Edition

The third edition provides substantive changes that have improved clarity of content, linkage to learning objectives, and updated scholarship throughout. The changes allow for a stronger presentation of the material, based on years of feedback from colleagues, instructors, and students, that is more illustrative in nature and meaningful for students. The third edition includes substantive changes that improve clarity of content, linkage to learning objectives, and updated scholarship throughout. Two major overarching themes to the revisions are apparent. A broad summary of changes in the third edition is given briefly here.

One theme in the writing of the third edition was that the figures, tables, and writing were revised to improve clarity throughout. Many revisions were specifically based on feedback from instructors and students, such as revisions to clarify the different distributions of data as well as to expand on parts of the hypothesis testing chapters. Changes included revising figures and tables, in addition to revising and adding new content throughout to build stronger writing around the content being presented, as per feedback from students and instructors.

Another theme that arose was updating scholarship throughout. As disciplines in the behavioral sciences advance, it is important to link hypothesis testing and statistical analysis to current examples to help students realize the value and real-world application of statistics in the behavioral sciences. Scholarship was updated throughout to provide dozens of new references, while also removing references that are now outdated. The scholarship was updated both in the text and in the end-of-chapter problems to bolster student learning.

In addition, Appendix B was added. This new appendix has two key benefits in the book: (1) it gives readers a step-by-step instructional guide for using SPSS throughout the book, and (2) it links each instruction to the specific chapter and page number in the book where students can locate within the text where to find an example of how to compute each test taught using SPSS. Additional changes in the book include learning objectives that were updated throughout and learning objective summaries that were revised with those corresponding changes. Examples were added and revised as needed to further clarify the examples in chapters and make the writing more concise where appropriate. End-of-chapter pedagogy was revised and updated to include new key terms and new content and remove old content no longer in the book. If end-of-chapter materials required an answer key, then the answer key was also updated in the back of the study guide. Overall, the changes allow for a stronger presentation of the material based on years of feedback from colleagues, instructors, and students.

Supplements

edge.sagepub.com/priviterastats3e

SAGE edge for Instructors supports your teaching by making it easy to integrate quality content and create a rich learning environment for students.

The password-protected Instructor Resources site features author-created tools designed to help instructors plan and teach their course. These include an extensive test bank, chapter-specific PowerPoint presentations, lecture notes, sample syllabi for semester and quarter courses, class exercises, Excel® data sets structured by discipline, SPSS in Focus data sets, and more.

SAGE edge for Students provides a personalized approach to help you accomplish your coursework goals in an easy-to-use learning environment.

The open-access Student Resources site provides eFlashcards, web quizzes, access to full-text SAGE journal articles with accompanying assessments, and multimedia resources.

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 24.0 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). Note that an alternative guide has been created that corresponds to this book for all SPSS in Focus sections: *Statistical Analysis "In Focus": Alternate Guides for R, SAS, and Stata*. If you prefer instead to use one of those alternative statistical software packages (R® , SAS® , or Stata®*), the alternative guide will be a valuable resource. Within this book, SPSS is introduced, so it will be worthwhile to read this preface before moving into future discussions of SPSS. This preface includes a general introduction to familiarize you with this software.

Understanding SPSS 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

* RStudio is a trademark of RStudio, Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Stata is a registered trademark of StataCorp LLC.

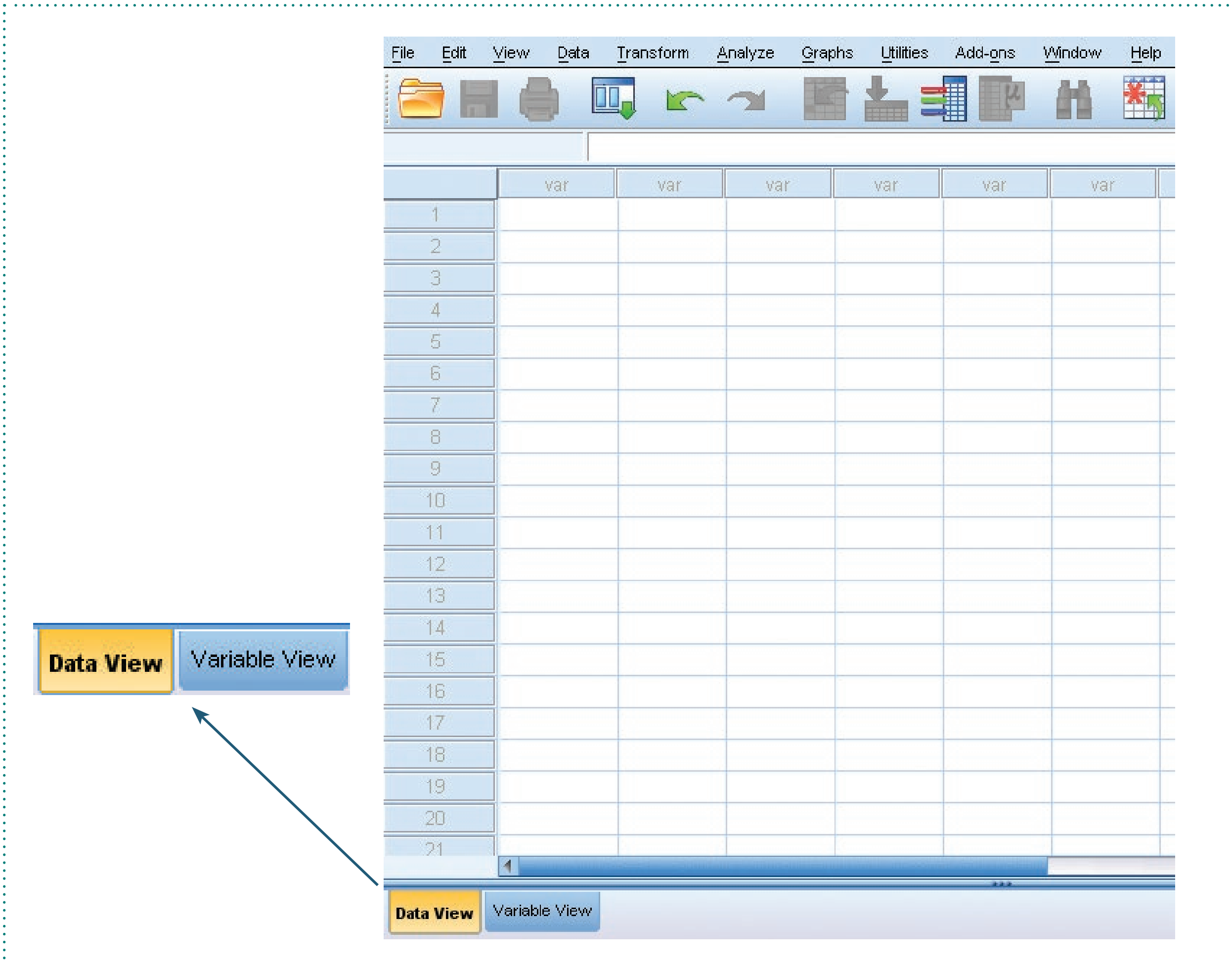
shown in Figure P.1. The Data View tab is open by default. The Variable View tab to the right of it is used to view and define the variables being studied.

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

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.

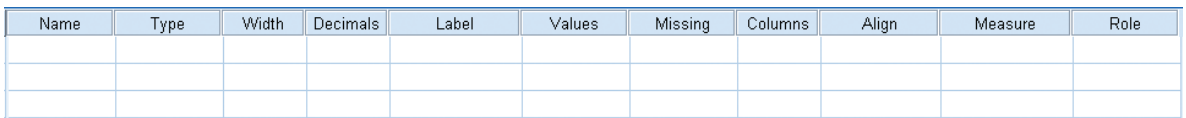
others are beyond the scope of this book. The purpose and function of each icon are introduced as needed in each chapter in the SPSS in Focus sections.

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.2 The Variable View Page With 11 Columns



Each column allows you to label and characterize variables.

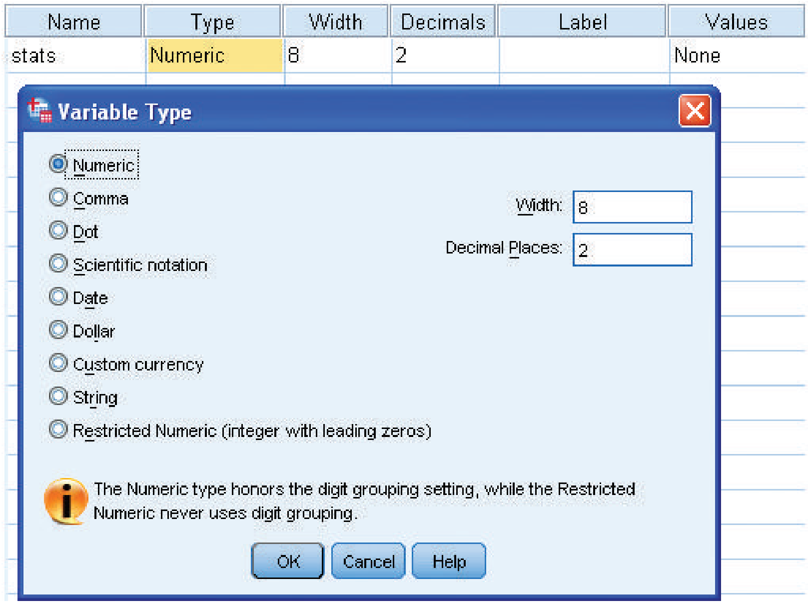
Name

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.3 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.4 Labeling Variables

Name	Type	Width	Decimals	Label
stats	Numeric	8	2	statistics

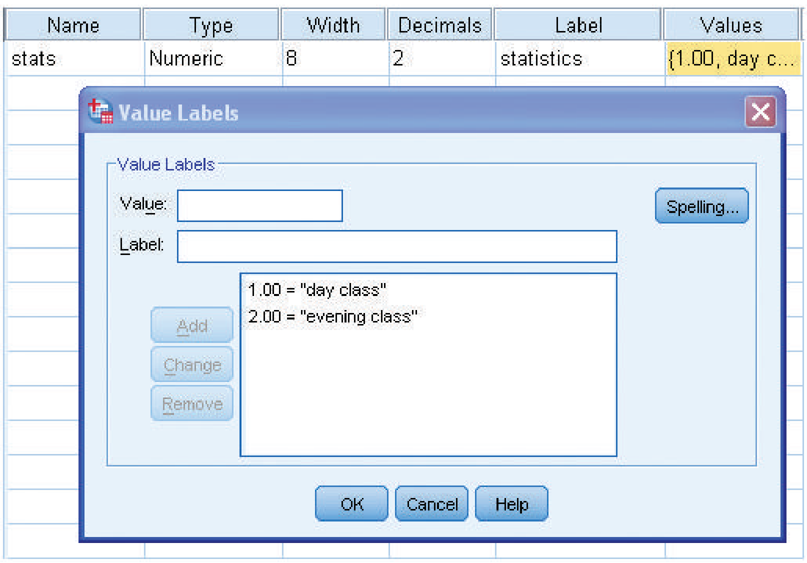
In this example, we labeled the variable name *stats* as *statistics* in the Label column.

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.

FIGURE P.5 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.

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 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.

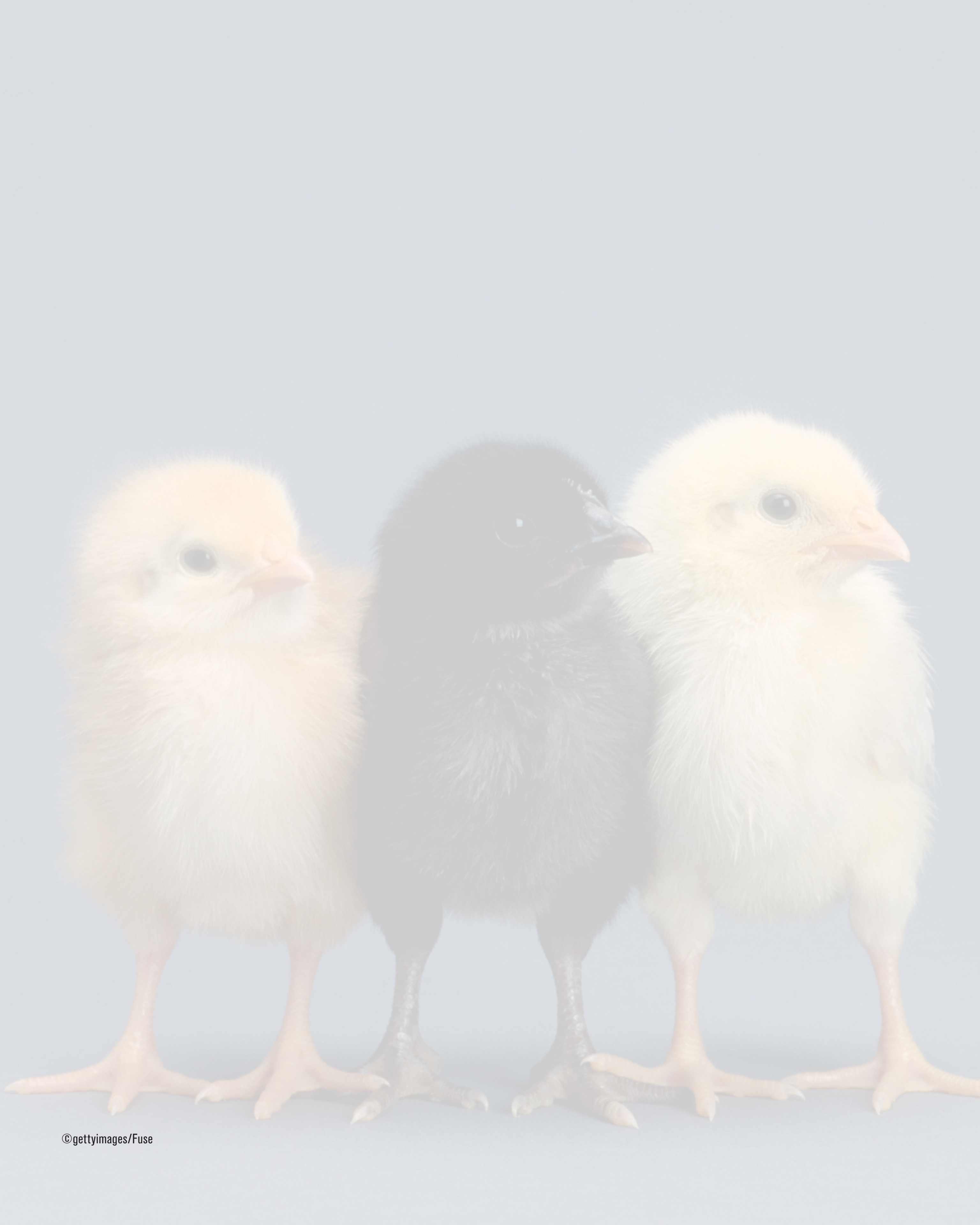
Role

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 inclusion of SPSS is simple: Most researchers use some kind of statistical software to analyze statistics; 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

Master the content.

edge.sagepub.com/priviterastats3e

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

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

Chapter Outline

in Chapters 2–4 and applications for probability are further introduced in Chapters 5–7, 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.

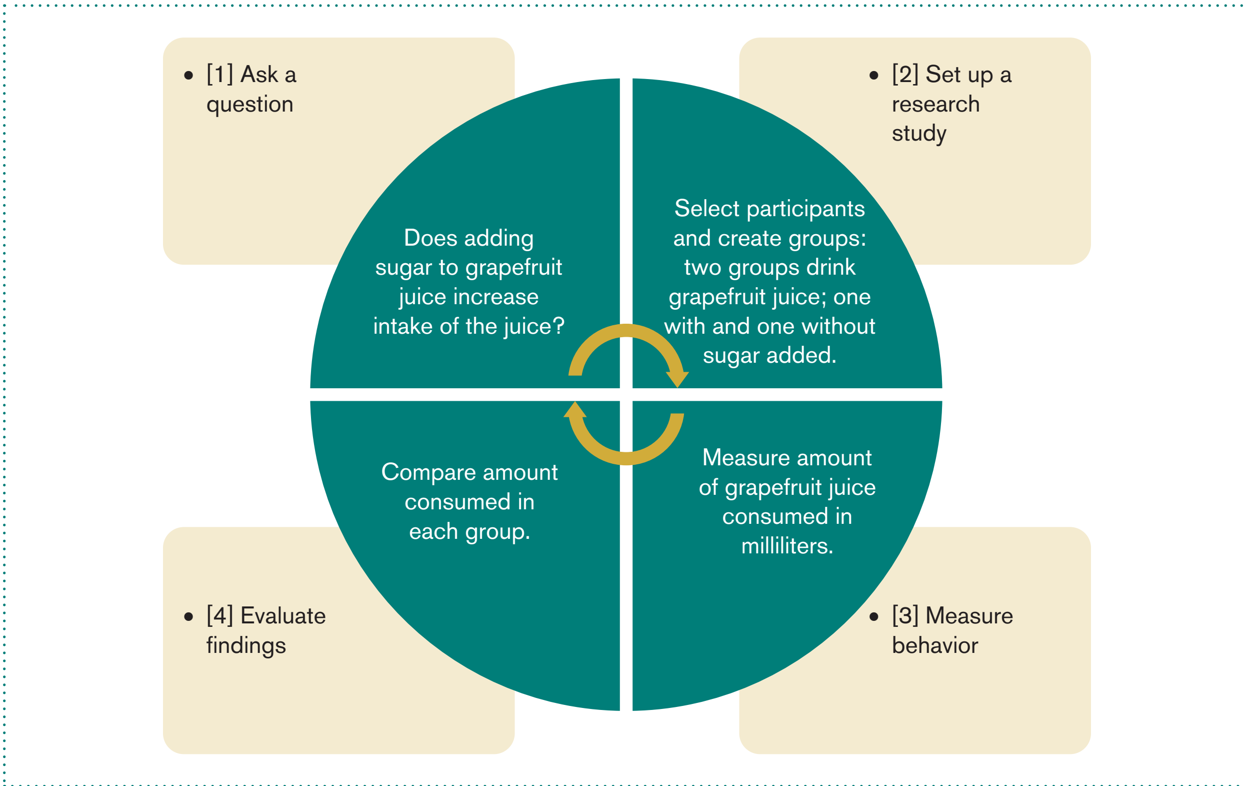
- Descriptive statistics: applying statistics to organize and summarize information
- Inferential statistics: applying statistics to interpret the meaning of information

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. As a basic example adapted from larger-scale studies looking at obesity and healthy food choice (Capaldi & Privitera, 2008; Privitera, 2016a), suppose a researcher asks if adding sugar to a sour-tasting fruit juice (a grapefruit juice) can increase intake of this healthy juice. To test this question, the researcher first identifies a group of participants who dislike plain grapefruit juice and sets up a research study to create two groups: Group No Sugar (this group drinks the grapefruit juice without any added sugar), and Group Sugar (this group drinks the grapefruit juice with sugar added). In this study, the researcher measures intake (i.e., how much juice is consumed). Suppose she decides to measure amount consumed in milliliters (note: 30 milliliters equals about 1 ounce). The data in this example are the volume of drink consumed in milliliters. If adding sugar increases

Data (plural) are a set of scores, 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**.

FIGURE 1.1 General Structure for Making Scientific Observations



The general structure for making scientific observations, using an example for testing if adding sugar increases intake of grapefruit juice.

intake of grapefruit juice, then we expect that participants will consume more of the grapefruit juice when sugar is added (i.e., Group Sugar will consume more milliliters of the juice than Group No Sugar).

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.

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. Thus, we measure or record data (e.g., milliliters consumed), then use descriptive statistics to summarize or make sense of those data, which describe the phenomenon of interest (e.g., intake of a healthy fruit juice). In our example, *intake* could be described simply as amount consumed, which certainly describes intake, but not numerically—or in a way that allows us to record data on intake. Instead, we stated that intake is milliliters consumed of the juice. Here, we define intake as a value that can be measured numerically; hence,

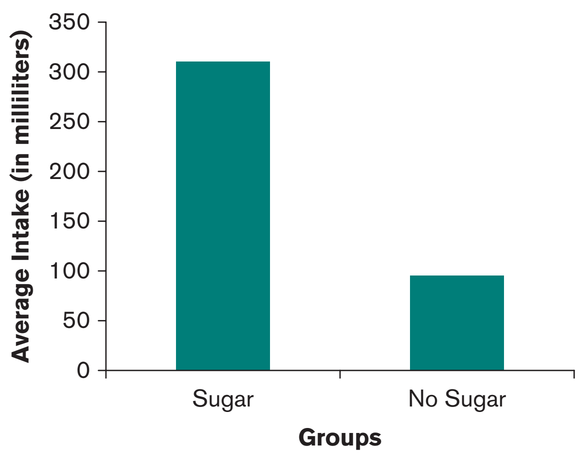
Descriptive statistics are procedures used to summarize, organize, and make sense of a set of scores called *data*. Descriptive statistics are typically presented graphically, in tabular form (in tables), or as summary statistics (single values).

intake can now be measured. If we observe hundreds of participants, then the data in a spreadsheet will be overwhelming. Presenting a spreadsheet with the intake for each individual participant 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 measure of intake, we could summarize the average (mean), middle (median), or most common (mode) amount consumed in milliliters among all participants, 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 participants will consume more grapefruit juice in milliliters if sugar is added to the juice. Figure 1.2 displays these expected findings. Notice how summarizing the average intake 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 adding sugar increases intake of grapefruit juice.

Inferential Statistics

Most research studies include only a select group of participants, and 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.

To illustrate, we can continue with the grapefruit juice study. If we are interested in all those who have a general dislike for sour-tasting grapefruit juice, then this group would constitute the **population** of interest. Specifically, we want to test if adding sugar increases intake of grapefruit juice in this population; this characteristic (intake of grapefruit juice) in the population is called a **population parameter**. Intake, then, is the characteristic we will measure, but not in the population. In practice, researchers will not have access to an entire population. They simply do not have the time, money, or other resources to even consider studying all those who have a general dislike for sour-tasting grapefruit juice.

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 those who have a general dislike for sour-tasting grapefruit juice from the larger population; the portion of those we select will constitute our sample. A characteristic that describes a sample, such as intake, is called a **sample statistic** and is the value that is measured in a 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 a population of interest.

FYI

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

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.

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 called *data*.
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

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