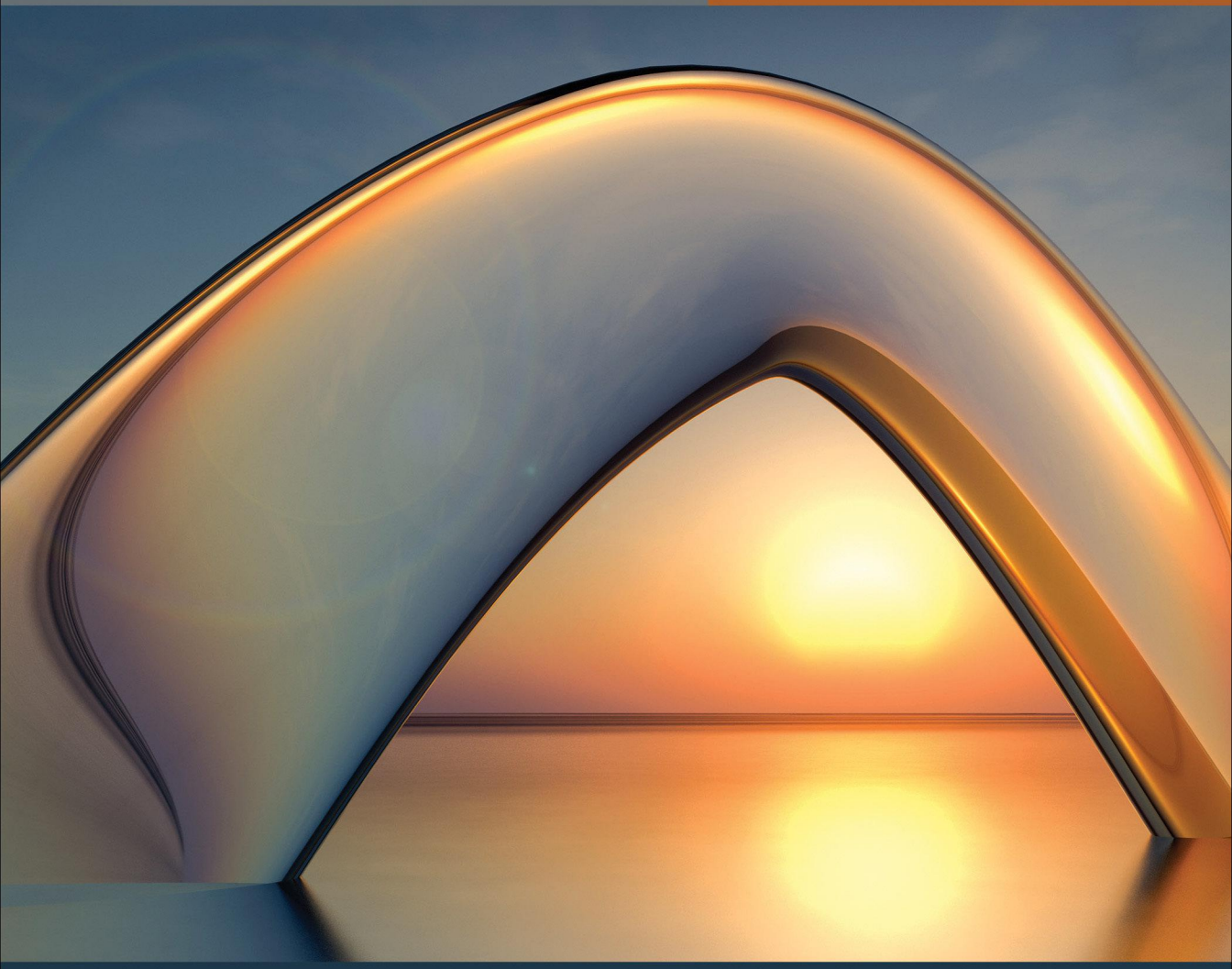
Statistics for People Who (Think They) Hate Statistics

USING MICROSOFT EXCEL©

FIFTH EDITION



Neil J. Salkind | Bruce B. Frey

Statistics for People Who (Think They) Hate Statistics

Fifth Edition

In honor and memory of my friend and mentor, Neil J. Salkind

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 600 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 People Who (Think They) Hate Statistics

Using Microsoft Excel

Fifth Edition

Neil J. Salkind

Bruce B. Frey
University of Kansas

SAGE

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

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.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Acquisitions Editor: Helen Salmon
Editorial Assistant: Elizabeth Cruz
Production Editor: Rebecca Lee
Typesetter: C&M Digitals (P) Ltd.
Proofreader: Jennifer Grubba
Indexer: Integra
Cover Designer: Candice Harman
Marketing Manager: Victoria Velasquez

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Names: Salkind, Neil J., author. | Frey, Bruce, author.

Title: Statistics for people who (think they) hate statistics: using Microsoft Excel / Neil J. Salkind, Bruce B. Frey, University of Kansas.

Description: Fifth edition. | Thousand Oaks, California: SAGE Publications, Inc., [2022] | Includes index.

Identifiers: LCCN 2021006243 | ISBN 978-1-0718-0388-2 (paperback) | ISBN 978-1-0718-0625-8 (epub) | ISBN 978-1-0718-0624-1 (epub) | ISBN 978-1-0718-0623-4 (pdf)

Subjects: LCSH: Statistics.

Classification: LCC HA29 .S2365 2022 | DDC 519.5—dc23

LC record available at https://lccn.loc.gov/2021006243

This book is printed on acid-free paper.

21 22 23 24 25 10 9 8 7 6 5 4 3 2 1

BRIEF CONTENTS

DETAILED CONTENTS

PREFACE

To the student:

This is the fifth edition of *Statistics for People Who (Think They) Hate Statistics, Excel Edition* (and also features the Macintosh version of Excel) and it was much fun (really!) writing this revision. Neil wrote the previous editions and I am honored to be allowed to pick up the mantle and continue with this wonderful series. I hope that the use of the book helps you and that the entire experience is interesting and productive.

What many of the students who study statistics, as well as people in various other roles (researchers, administrators, and professionals in many different fields), have in common (at least at the beginning of the course) is a relatively high level of anxiety, the origin of which is, more often than not, what they've *heard* from their fellow students. Often a small part of what they have heard is true—learning statistics takes an investment of time and effort (and there's the occasional monster for a teacher).

But most of what they've heard (and where most of the anxiety comes from)—that statistics is unbearably difficult and confusing—is just not true. Thousands of fear-struck students have succeeded where they thought they would fail. They did it by taking one thing at a time, pacing themselves, seeing illustrations of basic principles as they are applied to real-life settings, and even having some fun (see the recipe for delicious brownies in Appendix F) along the way. That's what we try to do in writing all the editions of *Statistics for People Who (Think They) Hate Statistics.*

After a great deal of trial and error, and some successful and many unsuccessful attempts, we learned to teach statistics in a way that we (and many of our students) think is unintimidating and informative. We have tried our absolute best to incorporate all of that experience into this book.

What you will learn from this *Statistics for People*...is the information you need to understand what the field and study of basic statistics is all about. You'll learn about the fundamental ideas and the most commonly used techniques to organize and make sense out of data. There's very little theory (but some), and there are a few mathematical proofs or discussion of the rationale for certain mathematical routines. And, for this Excel edition, you'll also learn how to better understand the world of statistics through the use of an easy-to-use, and powerful, tool.

Why isn't theory and other stuff in *Statistics for People Who (Think They) Hate Statistics?* Simple. Right now, you don't need it. It's not that we don't think it is

important. Rather, at this point and time in your studies, we want to offer you material at a level we think you can understand and learn with some reasonable amount of effort, while at the same time not be scared off from taking additional courses in the future. We (and your professor) want you to succeed.

So, if you are looking for a detailed unravelling of the derivation of the analysis of variance *F* ratio, go find another good book from SAGE (we'll be glad to refer you to one). Frankly, neither of us even know how to do that. But if you want to learn why and how statistics can work for you, you're in the right place. This book will help you understand the material you read in journal articles, explain what the results of many statistical analyses mean, and teach you how to perform basic statistical work.

And, if you want to talk about any aspect of teaching or learning statistics, feel free to contact Bruce. You can do this through his email address at school (**bfrey@ku.edu**). You can also keep up on anything new regarding this edition at **http://edge.sagepub.com/salkindexcel5e**, where you can also access the data sets for this book.

Good luck, and let us know how I can improve this book to even better meet the needs of the beginning statistics student. Now, let us talk to your instructor just for a second.

To the instructor:

First, we applaud your efforts at teaching introductory statistics. While the topic may be easier for some students than others, most find the material very challenging. Your patience and hard work is appreciated by all, and if there is anything we can do to help, please send me a note. Each edition of this book has benefitted greatly from the feedback and suggestions from teachers just like you.

Second, the Excel® edition of *Statistics for People Who (Think They) Hate Statistics* is not meant to be a dumbed-down book similar to others you may have seen. Nor is the title meant to convey anything other than the fact that many students new to the subject are very anxious about what's to come. This is not an academic version of a book for "dummies" or anything of its kind. As you know, your students aren't dummies; they are actually pretty bright. (hey, students, quit reading this, it's none of your business!). We have made every effort to address students with the respect they deserve, to not patronize them, and to ensure that the material is approachable. How well we did in these regards is up to you, but we want to convey our very clear intent and feeling that this book contains the information needed in an introductory course and, even though there is some humor involved in our approach, nothing about the intent is anything other than serious. Thank you!

You can sign in at **edge.sagepub.com/salkindexcel5e** to access the test bank, PowerPoint® slides, and data files to accompany this book.

ACKNOWLEDGMENTS

Everybody, and we mean everybody, at SAGE deserves a great deal of thanks for providing us with the support, guidance, and professionalism that takes our idea (way back before the first edition) and makes it into a book like the one you are now reading, and then makes it successful.

We also want to thank the following people for their help in providing feedback on the previous edition as well as this edition. First, thank you to Leslie Shaw, she's the smart writer, who along with Neil, is responsible for *Statistics for People…R Edition*! She contributed greatly to this book! Other folks who provided valuable suggestions for this new edition were:

Lucy Barnard-Brak, University of Alabama

Danielle T. Cooper, University of Newhaven

Jason Jolicoeur, Washburn University

Will Senn, Emporia State University

Candice Vander Weerdt, Cleveland State University

AND NOW, ABOUT THE FIFTH EDITION . . .

What you read up until now about this book reflects our thoughts about why we wrote this book in the first place and how it came to be. But it tells you little about this new edition.

Any book is always a work in progress, and the Excel edition of *Statistics for People Who (Think They) Hate Statistics* is no exception. Over the past decade, many people have found the book useful (and have been nice enough to tell us so), and others have told us how they would like it to change and why. In revising this book, I (Bruce) am trying to meet the needs of all audiences. Some things have remained the same, and some have changed.

There are always new things worth consideration and different ways to present old themes and ideas. Here's a list of what you'll find that's new in *Statistics for People Who (Think They) Hate Statistics: Using Microsoft Excel, Fifth Edition.*

- There continue to be exercises at the end of each chapter, but some have been changed to be more instructive and, I think, improved a bit. These exercises use data sets that are available at **edge.sagepub.com/salkindexcel5e**. The version of Excel that these were developed for is Excel 2016 for Windows and Excel 2016 for the Mac, but the files will work with earlier versions of Excel as well. They will work well with the most recent version, as well (circa 2019).
- Because theoretically there is a newer version of Excel than 2016, we have updated many of the graphics to match what you will really see on your own computers. I describe newer versions as theoretically new because they haven't substantially changed any of the procedures we use in this book.
- A new feature focuses on important actual everyday statisticians with an intentional goal of sharing the histories and contributions of women and people of color who have contributed to the field of research methods.
- At the end of each chapter is a Real-World Stats feature that provides an applied example of the content covered in the chapter. It's an attempt to further show the reader how statistics is, and can be, applied in the everyday world. We have updated many of these examples to deal with topics "right out of today's headlines."

- There is a greater emphasis on more "sophisticated" bells and whistles to help students conduct analyses, talk about statistics, and write about them. This includes calculating and reporting effect sizes (new content added to Chapter 11 and throughout Section IV), talking about hypotheses more precisely and intelligently, thinking more clearly about measurement concepts like validity and reliability, and more closely tying analytical choices to the level of measurement of variables.
- We've reorganized, renumbered, and renamed many of the chapters and appendices. You probably don't care about the specifics here, but rest assured that we did all this to give you more, not less, and to make things easier to navigate.
- The many resources associated with this book remain and have been reviewed and checked for accuracy and useability! This includes:
  - Online Resources for instructors at **edge.sagepub.com/salkind excel5e** include a test bank, PowerPoint slides, and data files.
  - Open-access study tools for students at **edge.sagepub.com/salkindexcel5e** include data files and flashcards of key terms.

As usual, we have been very careful to check for errors or typos or bad jokes and hope that all (well, most) of those have been removed. You can find a list of typos from the previous printings and editions of this book at **edge.sagepub.com/salkindexcel5e**. SAGE and I appreciate all the letters, calls, and emails pointing out these errors and making this fifth edition a better book for it. We have all made every effort in this edition to correct them and hope we did a reasonably good job. Let me hear from you with suggestions, criticisms, nice notes, and so on. Good luck.

Bruce B. Frey
University of Kansas
bfrey@ku.edu

ABOUT THE AUTHORS

Neil J. Salkind received his PhD in human development from the University of Maryland, and after teaching for 35 years at the University of Kansas, he was Professor Emeritus in the Department of Psychology and Research in Education, where he collaborated with colleagues and worked with students. His early interests were in the area of children's cognitive development, and after research in the areas of cognitive style and (what was then known as) hyperactivity, he was a postdoctoral fellow at the University of North Carolina's Bush Center for Child and Family Policy. His work then changed direction to focus on child and family policy, specifically the impact of alternative forms of public support on various child and family outcomes. He delivered more than 150 professional papers and presentations, wrote more than 100 trade and textbooks, and is the author of *Statistics for People Who (Think They) Hate Statistics, Theories of Human Development,* and *Exploring Research*. He has edited several encyclopedias, including the *Encyclopedia of Human Development,* the *Encyclopedia of Measurement and Statistics,* and the *Encyclopedia of Research Design.* He was editor of *Child Development Abstracts and Bibliography* for 13 years. He lived in Lawrence, Kansas, where he liked to read, swim with the River City Sharks, work as the proprietor and sole employee of big boy press, bake brownies (see Appendix I for the recipe), and poke around old Volvos and old houses. He died in 2017 at the age of 70.

Bruce B. Frey, PhD, is an award-winning researcher, teacher, and professor of educational psychology at the University of Kansas. He is the editor of *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation* and author of *There's a Stat for That!, Modern Classroom Assessment*, and *100 Questions (and Answers) About Tests and Measurement*. He is the associate editor of SAGE's *Encyclopedia of Research Design*. He also wrote *Statistics Hacks* for O'Reilly Media. His primary research interests include classroom assessment, instrument development, and program evaluation. In his spare time, Bruce leads a secret life as Professor Bubblegum, host of *Echo Valley*, a podcast that celebrates bubblegum pop music of the late 1960s. The show is wildly popular with the young people.

PART I

YIPPEE! I'M IN STATISTICS!

"Well, if it isn't a significant subset of our study group!"

Not much to shout about, you might say? Let us take a minute and show you how some real-life scientists use this widely used set of tools we call *statistics*.

- Michelle Lampl is a physician, the Samuel Candler Dobbs Professor of Anthropology and the co-director of the Emory-Georgia Tech Predictive Health Initiative. She was having coffee with a friend, who commented on how quickly her young infant was growing. In fact, the new mother spoke as if her son was "growing like a weed." Being a curious scientist (as all scientists should be), Dr. Lampl thought she might actually examine how rapid this child's growth, and that of others, was during infancy. She proceeded to measure a group of children's growth on a daily basis and found, much to her surprise, that some infants grew as much as 1 inch overnight! Some growth spurt.

  Want to know more? Why not read the original work? You can find more about this in Lampl, M., Veldhuis, J. D., & Johnson, M. L. (1992). Saltation and stasis: A model of human growth. *Science, 258,* 801–803.

- Why do some people eat meat and others don't? The traditional explanation has been that meat eaters like the taste of meat. Christopher A. Monteiro at Cornell University wondered if that was so or whether there was more to it than that. Might philosophical beliefs about animals play a role, as well? Dr. Monteiro created a measure to assess the level of two different beliefs: whether it is okay to eat animals and whether it is okay to kill animals, if you are going to eat them. He collected data and used statistics to find relationships among the questions on his measure to identify that there seem to be these different attitudes that work with things like "meat tastes good" to explain the practice of meat eating. Interestingly, those who scored the highest on the "okay to kill animals" scale also tended to score the highest on measures of racism and sexism.

  Want to know more? You can read the whole study! Look this up:

  Monteiro, C. A., Pfeiler, T. M., Patterson, M. D., & Milburn, M. A. (2017). The Carnism Inventory: Measuring the ideology of eating animals. *Appetite, 113,* 51–62.

- Sue Kemper is the Roberts Distinguished Professor of Psychology at the University of Kansas and has worked on the most interesting of projects. She and several other researchers studied a group of nuns and examined how their early experiences, activities, personality characteristics, and other information related to their health during their late adult years. Most notably, this diverse group of scientists (including psychologists, linguists, neurologists, and others) wanted to know how well all this information predicts the occurrence of Alzheimer's disease. Kemper and her colleagues found that the complexity of the nuns' writing during their early 20s was related to the nuns' risk for Alzheimer's 50, 60, and 70 years later.

  Want to know more? Why not read the original work? You can find more about this in Snowdon, D. A., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic ability in early life

and cognitive function and Alzheimer's disease in late life: Findings from the nun study. *Journal of the American Medical Association, 275,* 528–532.

- Do you suffer from migraine headaches? Millions do. Doctors have long been concerned about the risks associated with migraines because there are often serious diseases associated with them. Dr. Tobias Kurth, a German researcher working with Harvard, analyzed data from thousands of women who have migraine headaches and had information about their health across decades. After equalizing the women on a bunch of variables, he discovered that women with migraines were 50% more likely to have heart disease than those who did not have migraine headaches. And they are 62% more likely to have a stroke! The suggestion by Dr. Kurth and his coauthors is that women who suffer from migraines get more frequent checkups that include an evaluation for these other risks.

  Want to know more? Find out for yourself how the study was conducted and how statistics were used to get these estimates. Here are the details: Kurth, T., Winter, A. C., Eliassen, A. H., Dushkes, R., Mukamal, K. J., Rimm, E. B., . . . & Rexrode, K. M. (2016). Migraine and risk of cardiovascular disease in women: Prospective cohort study. *BMJ, 353,* i2610.

All of these researchers had a specific question they found interesting and used their intuition, curiosity, and excellent training to answer it. As part of their investigations, they used this set of tools we call *statistics* to make sense out of all the information they collected. Without these tools, all this information would have been just a collection of unrelated outcomes. The outcomes would be nothing that Lampl could have used to reach a conclusion about children's growth or Kemper could have used to better understand aging and cognition (and perhaps Alzheimer's disease).

Statistics—the science of organizing and analyzing information to make it more easily understood—made these tasks doable. The reason that any of the results from such studies are useful is that we can use statistics to make sense out of them. And that's exactly the goal of this book—to provide you with an understanding of these basic tools and how researchers use them and, of course, how to use them yourself.

In this first part of *Statistics for People Who (Think They) Hate Statistics, Excel 5th Edition,* you will be introduced to what the study of statistics is about and why it's well worth your efforts to master the basics—the important terminology and ideas that are central to the field. This part gives you a solid preparation for the rest of the book.

We'll also be getting right into the Excel material in Chapter 2, with information on formulas and functions and use of the Data Analysis tools in the program.

And this 5th edition works for Excel 2019 (for Windows and the Mac) or Excel 2016 and most any version of Excel you've got. Excel is probably available through the computer center or information technology department at your school or even at your local public library if you don't have access through the class that you might be taking, so check out those possibilities before you buy it.

STATISTICS OR SADISTICS? IT'S UP TO YOU

Difficulty Scale ☺ ☺ ☺ ☺ ☺
(really easy)

WHAT YOU WILL LEARN IN THIS CHAPTER

- What statistics is all about
- Why you should take statistics
- How to succeed in this course

WHY STATISTICS?

You've heard it all before, right?

"Statistics is difficult."

"I'm not a math person."

"I don't know how to use statistics software."

"What do I need this stuff for?"

"What do I do next?"

And the famous cry of the introductory statistics student: "I don't get it!"

Well, relax. Students who study introductory statistics find themselves, at one time or another, thinking at least one of the above and quite possibly sharing the thought with another student, their spouse, a colleague, or a friend.

And all kidding aside, some statistics courses can easily be described as *sadistics*. That's because the books are repetitiously boring, the examples don't seem to apply to real life, and too much math is thrown at you too quickly.

That's not the case for you. The fact that you or your instructor has selected *Statistics for People Who (Think They) Hate Statistics, Using Microsoft Excel, 5th Edition* shows that you're ready to take the right approach—one that is unintimidating, informative, and applied (and even a little fun) and that tries to teach you what you need to know about using statistics as the valuable tool that it is.

If you're using this book in a class, it also means that your instructor is clearly on your side. He or she knows that statistics can be intimidating but has taken steps to see that it is not intimidating for you. As a matter of fact, we'll bet there's a good chance (as hard as it may be to believe) that you'll be enjoying this class in just a few short weeks.

And Why Excel?

Simple. It's the most popular, most powerful spreadsheet tool available today, and it can be an exceedingly important and valuable tool for learning how to use basic and some advanced statistics. In fact, many stats courses taught at the introductory level use Excel as their primary computational tool and ignore other computer programs, such as IBM® SPSS® Statistics (SPSS)* and Minitab. Although we are not going to teach you how to use Excel (see Appendix A for a refresher on some basic tasks), we will show you how to use it to make your statistics learning experience a better one.

But like any program that takes numbers and consolidates and analyzes them, Excel is not a magic bullet or a tool to solve all your problems. It has its limitations. Unless you are an expert programmer and you can program Excel to do just about anything other statistics programs can (the language you would use is Visual Basic Applications or VBA), Excel may not look as pretty as other programs dedicated to statistical analysis or offer as many options. But at the level of introductory statistics, it is a very powerful tool that can do an awful lot of very neat things, and is a great skill for you to have in any workplace.

A bit of terminology about Excel before we move on: The first ever Excel-like computer application was called VisiCalc (thank you, Dan Bricklin and Bob

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

Frankston) and was known as a spreadsheet. Okay, the Excel application is known as a spreadsheet program as well, but each individual sheet is known as a **worksheet**. And worksheets, when combined, constitute what is known as a **workbook**. Fun, huh?

A 5-MINUTE HISTORY OF STATISTICS

Before you read any further, it would be useful to have some historical perspective about this topic called statistics. After all, almost every undergraduate in the social, behavioral, and biological sciences and every graduate student in education, nursing, psychology, social welfare and social services, anthropology, and . . . (you get the picture) is required to take this course. Wouldn't it be nice to have some idea from whence the topic it covers came? Of course it would.

Way, way back, as soon as humans realized that counting was a good idea (as in "How many of these do you need to trade for one of those?"), collecting information became a useful skill.

If counting counted, then one would know how many times the sun would rise in one season, how much food was needed to last the winter, and what amount of resources belonged to whom.

That was just the beginning. Once numbers became part of language, it seemed like the next step was to attach these numbers to outcomes. That started in earnest during the 17th century, when the first set of data pertaining to populations was collected. From that point on, scientists (mostly mathematicians, but then physical and biological scientists) needed to develop specific tools to answer specific questions. For example, Francis Galton (a half-cousin of Charles Darwin, by the way), who lived from 1822 to 1911, was very interested in the nature of human intelligence. He also speculated that hair loss was due to the intense energy that went into thinking. No, really. But back to statistics.

To explore one of his primary questions regarding the similarity of intelligence among family members, he used a specific statistical tool called the correlation coefficient (first developed by mathematicians), and then he popularized its use in the behavioral and social sciences.

PEOPLE WHO LOVED STATISTICS

Anonymous Portrait of Blaise Pascal, 17th century

Inferential statistics, the use of sample observations or data that we can see to make guesses about the likely characteristics of populations that we cannot see, probably started with Blaise Pascal (1623–1662), a French mathematician and religious philosopher. He developed the mathematical formulas that can predict important things like the probability of dice rolls and the likelihood of flipping a coin three times and having it come up heads each time. He even proved that if the coin almost always comes up heads, someone is cheating. The application of these statistical inventions was of immediate practical use to gamblers, and this might be the first time in history that statistics was seen as having practical applications. (Because one could make money by understanding them.) You may notice that Pascal did not live very long. He suffered from various illnesses during the later years of his life, and his cause of death isn't even known for sure, although after his death, he was found to have had stomach cancer and some brain damage. As a deeply religious man, Pascal believed that suffering was necessary for a good life. Thus, he probably would have enjoyed being a stats professor.

You'll learn all about this tool in Chapter 5. In fact, most of the basic statistical procedures that you will learn about were first developed and used in the fields of agriculture, astronomy, and even politics. Their application to human behavior came much later.

The past 100 years have seen great strides in the invention of new ways to use old ideas. The simplest test for examining the differences between the averages of two groups was first advanced during the early 20th century. Techniques that build on this idea were offered decades later and have been greatly refined. And the introduction of personal computers and such programs as Excel has opened up the use of sophisticated techniques to anyone who wants to explore these fascinating topics.

The introduction of these powerful personal computers has been both good and bad. It's good because most statistical analyses no longer require access to a huge and expensive mainframe computer. Instead, a simple personal computer costing

less than $250 or a cloud account can do 95% of what 95% of the people need. On the other hand, less than adequately educated students (such as your fellow students who chose not to take this course!) will take any old data they have and think that by running them through some sophisticated analysis, they will have reliable, trustworthy, and meaningful outcomes—not true. What your professor would say is "Garbage in, garbage out"; if you don't start with data you can trust, what you'll have after your data are analyzed are results you cannot trust.

Today, statisticians in all different areas, from criminal justice to geophysics to psychology to determining whether the "hot" hand really exists in the NBA, find themselves using basically the same techniques to answer different questions. There are, of course, important differences in how data are collected, but for the most part, the analyses (the plural of *analysis*) that are done following the collection of data (the plural of *datum,* which means one piece of information) tend to be very similar, even if called something different. The moral here? This class will provide you with the tools to understand how statistics are used in almost any discipline. Pretty neat, and all for just the cost of a few credit hours.

Okay. Five minutes is up, and you know as much as you need to know for now about the history of statistics. You'll get some more history here and there as we learn about different procedures. Let's move on to what statistics is (and isn't).

STATISTICS: WHAT IT IS (AND ISN'T)

Statistics for People Who (Think They) Hate Statistics, Using Microsoft Excel, 5th Edition is a book about basic statistics and how to apply them to a variety of different situations, including the analysis and understanding of information.

In the most general sense, *statistics* describes a set of tools and techniques that are used for describing, organizing, and interpreting information or data. Those data might be the scores on a test taken by students participating in a special math curriculum, the speed with which problems are solved, the number of side effects when patients use one type of drug rather than another, the number of errors in each inning of a World Series game, or the average price of a dinner in an upscale restaurant in Santa Fe, New Mexico (not cheap).

In all of these examples, and the million more we could think of, data are collected, organized, summarized, and then interpreted. In this book, you'll learn about collecting, organizing, and summarizing data as part of descriptive statistics. And then you'll learn about interpreting data when you learn about the usefulness of inferential statistics.

What Are Descriptive Statistics?

Descriptive statistics are used to organize and describe the characteristics of a collection of data. The collection is sometimes called a **data set** or just **data**.

Scientists would say that descriptive statistics describe a *sample*—a collection of data that you have in front of you.

For example, the following list shows you the names of 22 college students, their major areas of study, and their ages. If you needed to describe what the most popular college major is, you could use a descriptive statistic that summarizes their most frequent choice (called the mode). In this case, the most common major is psychology. And if you wanted to know the average age, you could easily compute another descriptive statistic that identifies this variable (that one's called the mean). Both of these simple descriptive statistics are used to describe data. They do a fine job allowing us to represent the characteristics of a large collection of data such as the 22 cases in our example. Any time there are more than just a few people or things you want to describe, descriptive statistics make that easier. Much easier! That's why descriptive statistics are so popular in the social and natural sciences.

Name	Major	Age	Name	Major	Age
Deja	Education	19	Aliyah	English	21
Sara	Psychology	18	Mateo	Psychology	22
Asma	Education	19	Hadley	Psychology	23
Trevon	Psychology	21	Alejandro	Education	21
Jordan	Education	20	Chip	Education	19
Pam	Education	24	Homer	Psychology	18
Xavier	Psychology	21	Li	English	22
Liz	Psychology	19	Darius	Psychology	24
Nicole	Chemistry	19	Leonard	Psychology	21
Zhang	Nursing	20	Jeffrey	Chemistry	18
Kent	History	18	Emily	Spanish	19

So watch how simple this is. To find the most frequently selected major, just find the one that occurs most often. And to find the average age, just add up all the age values and divide by 22. You're right—the most often occurring major is psychology (9 times), and the average age is 20.3 (actually 20.27). Look, Ma! No hands—you're a statistician.

What Are Inferential Statistics?

Inferential statistics are often (but not always) the next step after you have collected and summarized data. Inferential statistics are used to make inferences based on a smaller group of data (such as our group of 22 students) about a possibly larger one (such as all the undergraduate students in the College of Arts and Sciences).

A smaller group of data is often called a **sample**, which is a portion, or a subset, of a **population**. For example, all the fifth graders in Newark (Neil's fair city of origin), New Jersey, would be a population (the population is all the occurrences with certain characteristics, in this case, being in fifth grade and attending school in Newark), whereas a selection of 150 of these students would be a sample. If we think this sample represents the population well, we can make guesses about the population.

Let's look at another example. Your marketing agency asks you (a newly hired researcher) to determine which of several names is most appealing for a new brand of potato chip. Will it be Chipsters? FunChips? Crunchies? As a statistics pro (we know we're moving a bit ahead of ourselves, but keep the faith), you need to find a small group of potato chip eaters who are representative of all potato chip fans and ask these people to tell you which one of the three names they like the most. Then, if you do things right, you can easily extrapolate the findings to the huge group of potato chip eaters.

Or let's say you're interested in the best treatment for a particular type of disease. Perhaps you'll try a new drug as one alternative, a placebo (a substance that is known not to have any effect) as another alternative, and nothing as the third alternative to see what happens. Well, you find out that more patients get better when no action is taken and nature (and we assume that's the only factor or set of factors that differentiate the groups) just takes its course! The drug does not have any effect. Then, with that information, you can extrapolate to the larger group of patients who suffer from the disease, given the results of your experiment.

Inferring from a sample to a population makes a lot of sense, especially when you are sure the sample represents the population. That's why, as you'll see later, scientists spend a lot of effort getting a representative sample.

In Other Words . . .

Statistics is a tool that helps us understand the world around us. It does so by organizing information we've collected and then letting us make certain statements about how characteristics of those data are applicable to new settings. Descriptive and inferential statistics work hand in hand, and which statistic you use and when depends on the question you want answered and how you happened to measure your variables.

And today, a knowledge of statistics is more important than ever because it provides us with the tools to make decisions that are based on empirical (observed) evidence and not our own biases or beliefs. Want to know whether early intervention programs work? Then test whether they work and provide that evidence to the

court that will make a ruling on the viability of a new school bond issue that could pay for those programs.

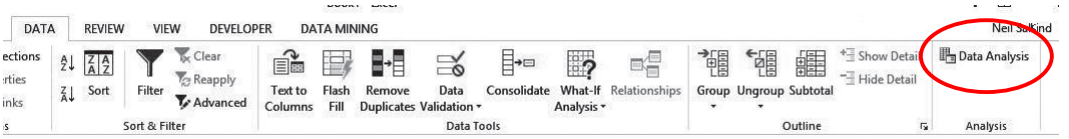
TOOLING AROUND WITH THE DATA ANALYSIS TOOLS

An awful lot of what you need to know about using Excel can be found in Appendix A. However, certain Excel procedures are available only if you have the Data Analysis tools installed (and we use those tools in several chapters throughout the book). Please note that this Excel add-in is sometimes called the Data Analysis Toolpak and sometimes the Data Analysis tool or Data Analysis tools and sometimes the Analysis Toolpak. These are all different names for the same thing. We'll refer to it as Data Analysis tools to cover all the bases.

The Data Analysis tools are a spectacular Excel add-in. Add-ins are special sets of tools that are often not installed when Excel was originally installed.

How do you know whether Data Analysis tools are installed on the computer you are using? If the Data Analysis tools option doesn't appear on your Data tab as Data Analysis tools (as you see it appears in Figure 1.1) in the Windows version, you need to install it. Either ask your instructor to have it installed on the network level where Excel is installed or install it on your own machine by doing the following. In the Mac version, it appears as the Data Analysis option on the Tools menu, so Mac users have no need for any installation steps.

FIGURE 1.1 ⬢ The Data Analysis Tools Option on the Data Tab



1. Click the File tab and then click Options.
2. Click Add-Ins and then in the Add-Ins box, select Analysis Toolpak (see, it's named different things in different places).
3. Click Go under Manage, at the bottom of the screen.
4. In the Add-Ins box, click the Analysis Toolpak check box and then click OK.

You are now ready to make your Excel activities even that much more productive and fun. You can learn how to use the Data Analysis tools in Chapter 2.

WHAT AM I DOING IN A STATISTICS CLASS?

You might find yourself using this book for many reasons. You might be enrolled in an introductory statistics class. Or you might be reviewing for your comprehensive exams. Or you might even be reading this on summer vacation (horrors!) in preparation for a more advanced class.

In any case, you are a statistics student, whether you have to take a final exam at the end of a formal course or you're just in it of your own accord. But there are plenty of good reasons to be studying this material—some fun, some serious, and some both.

Here's the list of some of the things that our students hear at the beginning of our introductory statistics courses:

1. Statistics 101 or Statistics 1 or whatever it's called at your school looks great listed on your transcript. Kidding aside, this may be a required course for you to complete your major. But even if it is not, having these skills is definitely a big plus when it comes time to apply for a job or for further schooling. And with more advanced courses, your résumé will be even more impressive.
2. If this is not a required course, taking basic statistics sets you apart from those who do not. It shows that you are willing to undertake a course that is above average with regard to difficulty and commitment. And, as the political and economic (and sports!) worlds become more "accountable," more emphasis is being placed on analytic skills. Who knows, this course may be your ticket to a job!
3. Basic statistics is an intellectual challenge of a kind that you might not be used to. There's a good deal of thinking that's required, a bit of math, and some integration of ideas and application. The bottom line is that all this activity adds up to what can be an invigorating intellectual experience because you learn about a whole new area or discipline.
4. There's no question that having some background in statistics makes you a better student in the social or behavioral sciences, because you will have a better understanding not only of what you read in journals but also of what your professors and colleagues may be discussing and doing in and out of class. You will be amazed the first time you say to yourself, "Wow, I actually understand what they're talking about." And it will happen over and over again, because you will have the basic tools necessary to understand exactly how scientists reach the conclusions they do.

5. If you plan to pursue a graduate degree in education, anthropology, economics, nursing, sociology, or any one of many other social, behavioral, and biological pursuits, this course will give you the foundation you need to move further.
6. There are many different ways of thinking about, and approaching, different types of problems. The set of tools you learn about in this book (and this course) will help you look at interesting problems from a new perspective. And, while the possibilities may not be apparent now, this new way of thinking can be brought to new situations.
7. Finally, you can brag that you completed a course that everyone thinks is the equivalent of building and running a nuclear reactor.

TEN WAYS TO USE THIS BOOK (AND LEARN STATISTICS AT THE SAME TIME!)

Yep. Just what the world needs—another statistics book. But this one is different. It is directed at the student, is not condescending, is informative, and is as basic as possible in its presentation. It makes no presumptions about what you should know before you start and proceeds in slow, small steps, which lets you pace yourself.

However, there has always been a general aura surrounding the study of statistics that it's a difficult subject to master. And we don't say otherwise, because parts of it are challenging. On the other hand, millions and millions of students have mastered this topic, and you can, too. Here are 10 hints to close this introductory chapter before we move on to our first topic.

1. You're not dumb. That's true. If you were, you would not have gotten this far in school. So, treat statistics as you would any other new course. Attend the lectures, study the material, do the exercises in the book and from class, and you'll do fine. Rocket scientists know statistics, but you don't have to be a rocket scientist to succeed in statistics.

2. How do you know statistics is hard? Is statistics difficult? Yes and no. If you listen to friends who have taken the course and didn't do well, they'll surely volunteer to tell you how hard it was and how much of a disaster it made of their entire semester, if not their lives. And let's not forget—we always tend to hear from complainers. So, we'd suggest that you start this course with the attitude that you'll wait and see how it is and judge the experience for yourself. Better yet, talk to several people who have had the class and get a good idea of what they think. Don't base your expectations on just one spoilsport's experience. Get a bigger sample!

3. Don't skip lessons—work through the chapters in sequence. *Statistics for People Who (Think They) Hate Statistics, Excel 5th Edition* is written so that each chapter provides a foundation for the next one in the book. When you are all done with the course, you will (I hope) continue to use this book as a reference. So if you need a particular value from a table, you might consult Appendix B. Or if you need to remember how to compute the standard deviation, you might turn to Chapter 3. But for now, read each chapter in the sequence that it appears. It's okay to skip around and see what's offered down the road. Just don't study later chapters before you master earlier ones.

4. Form a study group. This is a big hint and one of the most basic ways to ensure some success in this course. Early in the semester, arrange to study with friends or classmates. If you don't have any friends who are in the same class as you, then make some new ones or offer to study with someone who looks as happy to be there as you are. Studying with others allows you to help them if you know the material better or to benefit from those who know some material better than you. Set a specific time each week to get together for an hour and go over the exercises at the end of the chapter or ask questions of one another. Take as much time as you need. Studying with others is an invaluable way to help you understand and master the material in this course.

5. Ask your teacher questions, and then ask a friend. If you do not understand what you are being taught in class, ask your professor to clarify it. Have no doubt—if you don't understand the material, then you can be sure that others do not as well. More often than not, instructors welcome questions. And especially because you've read the material before class, your questions should be well informed and help everyone in class to better understand the material.

6. Do the exercises at the end of a chapter. The exercises are based on the material and the examples in the chapter they follow. They are there to help you apply the concepts that were taught in the chapter and build your confidence at the same time. If you can answer these end-of-chapter exercises, then you are well on your way to mastering the content of the chapter. Correct answers to each exercise are provided in Appendix D.

7. Practice, practice, practice. Yes, it's a very old joke:

Q. How do you get to Carnegie Hall?

A. Practice, practice, practice.

Well, it's no different with basic statistics. You have to use what you learn and use it frequently to master the different ideas and techniques. This means doing the exercises at the end of Chapters 1 through 17 and Chapter 19, as well as taking advantage of any other opportunities you have to understand what you have learned.

8. Look for applications to make it more real. In your other classes, you probably have occasion to read journal articles, talk about the results of research, and generally discuss the importance of the scientific method in your own area of study. These are all opportunities to see how your study of statistics can help you better understand the topics under class discussion as well as the area of beginning statistics. The more you apply these new ideas, the fuller your understanding will be.

9. Browse. Read over the assigned chapter first; then go back and read it with more intention. Take a nice leisurely tour of *Statistics for People Who (Think They) Hate Statistics, Excel 5th Edition* to see what's contained in the various chapters. Don't rush yourself. It's always good to know what topics lie ahead as well as to familiarize yourself with the content that will be covered in your current statistics class.

10. Have fun. This might seem like a strange thing to say, but it all boils down to you mastering this topic rather than letting the course and its demands master you. Set up a study schedule and follow it, ask questions in class, and consider this intellectual exercise to be one of growth. Mastering new material is always exciting and satisfying—it's part of the human spirit. You can experience the same satisfaction here—just keep your eye on the ball and make the necessary commitment to stay current with the assignments and work hard.

ABOUT THE BOOK'S FEATURES

Throughout the book, there are short biographies of **People Who Loved Statistics**. All of the statistical tricks and procedures we will discover in this book were invented by real people, and it's good to realize that they were just like you and me! (Well, a little like you and me.) At other times, in special sections, we will focus on particularly cool aspects of statistics that will help you to start to feel like a real honest-to-goodness statistician—topics like effect sizes (whatever those are), data analysis tricks, decision trees about which statistic to use, and, of course, plenty of applied examples using actual published research articles!

Throughout this book, you'll also find a small-steps icon like the one you see here. This indicates that a set of steps is coming up that will direct you through a particular process. Sometimes you will use Excel to do these steps. These steps have been tested and approved by whatever federal agency approves these things.

Appendix A contains an introduction to Excel. Working through this appendix is all you really need to do to be ready to use Excel. If you have an earlier version of Excel, you will still find this material to be very helpful, as most of the procedures are unchanged from one version to the next.

Appendix B contains important tables you will learn about and need throughout the book.

And, in working through the exercises in this book, you will use the data sets in Appendix C. In the exercises, you'll find references to data sets with names like "Chapter 2 Data Set 1," and each of these sets is shown in Appendix C. You can either enter the data manually or download them from the publisher's site at **edge.sagepub.com/salkindfrey5e**.

Appendix D contains answers to end-of-chapter questions.

Appendix E contains a primer on math for those who could use a refresher.

Appendix F describes some statistical software other than Excel that you might find useful.

Appendix G contains statistics Websites that are fun to play around in.

Appendix H contains the most helpful hints for gathering your own data.

And Appendix I offers the long-sought-after brownie recipe (yes, you finally found it).

KEY TO DIFFICULTY ICONS

To help you along a bit, we placed a difficulty index at the beginning of each chapter. This adds some fun to the start of each chapter, but it's also a useful tip to let you know what's coming and how difficult chapters are in relation to one another. Because the index uses smiley faces, the more smiles the merrier!

☺ (very hard)

☺ ☺ (hard)

☺ ☺ ☺ (not too hard, but not easy either)

☺ ☺ ☺ ☺ (easy)

☺ ☺ ☺ ☺ ☺ (very easy)

KEY TO "HOW MUCH EXCEL" ICONS

To help you along a bit more, we placed a "How Much Excel" index at the beginning of each chapter. This adds even more fun (groan) to the start of each chapter, but it also lets you know how much Excel material is contained in the chapter.

How much Excel?

 (just a mention)

 (some)

 (lots)

 (lots and lots)

 (a ton)

GLOSSARY

Bolded terms in the text are included in the glossary at the back of the book.

SUMMARY

That couldn't have been that bad, right? We want to encourage you to continue reading and not worry about what's difficult or time-consuming or too complex for you to understand and apply. Just take one chapter at a time, as you did this one.

GETTING STARTED IN EXCEL

Difficulty Scale ☺ ☺
(a little tough, but invaluable to stick with)

How much Excel?
(a ton)

WHAT YOU'LL LEARN ABOUT IN THIS CHAPTER

- The difference between formulas and functions
- How to create and use a formula
- The important Excel functions
- How to select and use a function
- What the Data Analysis tools are and what they do

There may be nothing more valuable in your Excel magic tool box than formulas and functions. This is true whether you are using Excel 2019, 2016 or even older versions. Indeed, everything we talk about in this book applies to all the most recent editions! Formulas (or what stats folks like to call *formulae*) and functions both allow you to bypass (very) tedious calculations and get right to the heart of the matter. Both formulas and functions are shortcuts—and both work in different ways and do different things. Let's start with formulas.

WHAT'S A FORMULA?

You probably already know the answer to that question. A **formula** is a set of mathematical operators that performs a particular mathematical task. For example, here's a simple formula:

$$2 + 2 =$$

The operator + tells you to add certain values (a 2 and another 2) together to produce the outcome (4; we did that in our head!). This is a simple one.

Here's one that's a bit more advanced and one with which you will become more familiar in Chapter 17:

$$Y = bX + a \tag{2.1}$$

This is the formula that is used to predict the value of Y' from our knowledge of the values of b, X, and a. We'll worry about what all those symbols mean later. For now, just know this is a formula that contains a bunch of symbols and mathematical operators and helps us compute numbers we need to make decisions.

Excel is a formula engine just ready to help make your learning of statistics easier.

Creating a Formula

A formula is created through these steps:

1. Click on the **cell** in which you want the results of the formula to appear.
2. Enter an equal sign, which looks like this: =. In Excel, all formulas begin with an equal sign, no matter what else they contain.
3. Enter the formula. No spaces in formulas please—Excel does not like them.
4. Press the Enter key, and voilà! The results of the formula will appear in the selected cell.

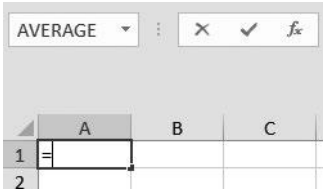
For example, let's enter the formula that was shown earlier— 2 + 2—and see how these steps work.

As you can see in Figure 2.1 we selected Cell A1.

FIGURE 2.1 ⬢ Selecting a Cell Into Which a Formula Will Be Entered

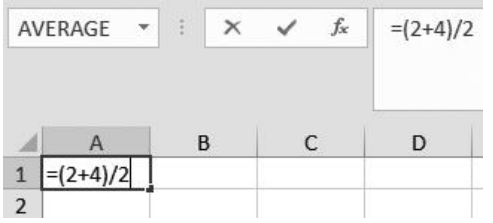
1. The equal sign is entered, as shown in Figure 2.2. And, as you can see, the formula bar at the top of the column becomes active. Everything we enter in Cell A1 will appear in the formula bar. Also, note that Excel automatically enters the name of the last function used (in this example it is AVERAGE—much more about functions soon).

FIGURE 2.2 ⬢ Entering the Equal Sign to Indicate the Beginning of a Formula



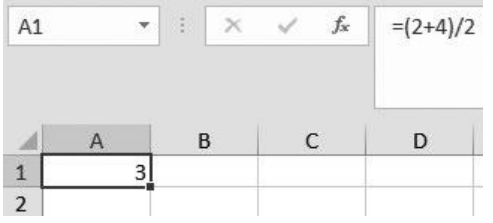
2. Enter the rest of the formula, which in this case is (2+4)/2, as you see in Figure 2.3.

FIGURE 2.3 ⬢ Entering the Formula in Cell A1



3. Press Enter, and the value of the formula is returned to the cell, as you see in Figure 2.4. And, if you click on cell A1 (as we did in Figure 2.4), you can see the formula located in that cell. In all cases, when a formula is entered into a cell, you see the results of that formula in the cell and the formula itself in the formula bar. In this case, the formula (2+4)/2 (shown in the formula bar) returns the result of 3. (If you got 4, instead of 3, you forgot to put in the parentheses! Parentheses tell Excel what to do first.)

FIGURE 2.4 ⬢ The Value of a Formula Being Returned to the Cell



A few notes:

- A formula always begins with an equal sign, which gives Excel a heads up that what follows is the formula.
- The formula itself always appears in the formula bar.
- The results of the formula (and not the formula itself) are returned to the selected cell.

This is the simplest example of how to use a formula. Formulas can become as complex as you need them to be.

Want to see the formula behind the scenes in a worksheet? Just use the Ctrl and ` key combination to toggle between formulas and the results of those formulas. The ` key is to the left of the number 1 key near the top of a standard keyboard.

Operator, Operator—Get Me a Formula!

You have just seen that even the simplest formulas consist of operators. In this case, the only operators are a plus (+) sign and a division (/) sign, which direct Excel to add the two values you see in Figure 2.3, divide by 2, and return the sum to Cell A1.

Addition and division are just two of the operations you can perform. The most important operations and the symbols you use to accomplish them are shown in the following table.

Operator	Symbol	Example	What It Does
Addition	+ (plus)	=2+5	Adds 2 and 5.
Subtraction	– (minus)	=5–3	Subtracts 3 from 5.
Division	/ (slash)	=10/5	Divides 10 by 5.
Multiplication	* (asterisk)	=2*5	Multiplies 2 times 5.
Power of	^ (caret)	=4^2	Takes 4 to the power of 2, or squares 4.

Beware the Parentheses

When you create a formula that goes beyond a very simple one, it is critical for you to consider the order in which operations are carried out and the use of parentheses.

Let's say that we want to find the average score on a weekly test given each Friday for a month and the scores range from 0 to 100.

Here are Willy's scores:

Week 1 78

Week 2 56

Week 3 85

Week 4 92 (Willy finally got it!)

We need to create a formula that will add all of the scores together and divide the sum by 4. We'll name the scores w_1, w_2, w_3, and w_4. Here's one way we might try it:

$$w_1 + w_2 + w_3 + w_4 /4$$

Oops! This will work in the sense that it will produce a number, but it won't give you the outcome you want. What this does is add w_1, w_2, and w_3 together and then adds that sum to the value of (only) w_4 divided by 4. This is not what we want.

Rather, take a look at this formula:

$$(w_1 + w_2 + w_3 + w_4)/4$$

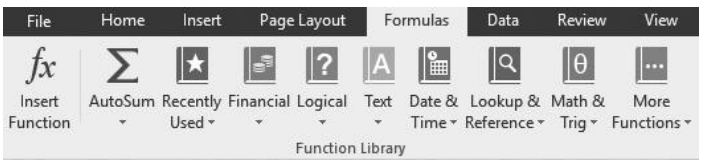
This is more like it. Here, the four values are summed and then that sum is divided by 4. This one works. The lesson? Watch your parentheses! By the way, there is an even easier way to calculate an average, and it is to use a function. Read on.

WHAT'S A FUNCTION?

You know that a formula is a simple set of symbols (such as numbers and operators) that performs some calculation and results in an outcome in the cell where the formula lives.

A **function** is nothing other than a predefined formula. The good people who created Excel developed a whole bunch of functions that can do many different things, but throughout *Statistics for People . . .* , we deal only with those that are relevant to the things we do in these chapters. Functions fall under the general Formulas tab on the ribbon, as you see in Figure 2.5.

FIGURE 2.5 ⬢ The Formulas Tab on the Excel Ribbon



For example, there are groups of financial functions, logical functions, text functions, and others. But we're going to focus (mostly) on the functions that fall in the category of statistical functions and a few database functions. The group of statistical functions are visible on your screen only when you click on the More

Functions drop-down box, which you see in Figure 2.5, and then click Statistical. The group of database functions must be the black sheep of Excel because they don't have their own grouping on the Excel Ribbon. Instead, you have to click the Insert Function button on the left of the ribbon and then specify the group of database functions.

Functions that are relevant to the material covered in this book include AVERAGE (guess what that does) and T.TEST (guess, but you probably don't yet know). Some are too advanced for us to bother with, such as FISHER and GAMALIN (which do not refer to two of Tolkien's elves that accompany Bilbo). We'll leave those for the next course or for you to explore on your own.

Using a Function

Unlike a formula, a function is not created by you. You just tell it which values (located in which cells) you want to work with. Every formula contains two elements—the name of the function and the argument of the function. *Argument* doesn't mean the function is fighting over whose turn it is to wash the dishes. It refers to the values you will use in a function. To understand what an Excel argument is, let's look at an example.

Here's a very simple function that averages a given set of numbers. In this example, this function averages the numbers in Cells A1 through A3:

=AVERAGE(A1:A3)

The *name* of the function is AVERAGE, and the *argument* is A1:A3—the cells on which you want the function to perform its magic. And as you can see, functions (like formulas) always, always, always begin with an equal sign.

Here's another function that produces the sum of a set of cells:

=SUM(A1:A3)

Simple, right? And, you may be thinking, "Well, why not just use a formula in this case?" and you could. But what if you need the sum of a set of 3,267 values like this?

=SUM(A1:A3267)

You really don't want to type in =(A1+A2+A3+A4 . . .) until you get to A3267, right? We thought not. Or what if you need a fancy-schmancy calculation that includes formulas that are very complex? Functions to the rescue!

So, let's get to the way that we use a function, and as an example, we'll use the AVERAGE function.

To use this (or any other function), you follow three steps:

1. Enter the function in the cell where you want the results to appear.
2. Enter the range of cells on which you want the function to operate.
3. Press the Enter key, and voilà! There you have the result located in the cell in which the function was created.

However, there are several ways to accomplish these three steps, and let's deal with those now.

Inserting a Function (When You Know the Function's Name and How It Works)

Here's the old-fashioned way.

1. Enter the function in the cell where you want the results to appear.

   For example, in Figure 2.6, you can see a data set of 10 values. This isn't real data; we just made it up using our super brains! Next, we're going to find the average of those values using the AVERAGE function. And, to make things a bit clearer, we entered a text label in the cell to the left of where we want the sum to appear.

FIGURE 2.6 ⬢ Creating a Data Set and the Location of the AVERAGE Function

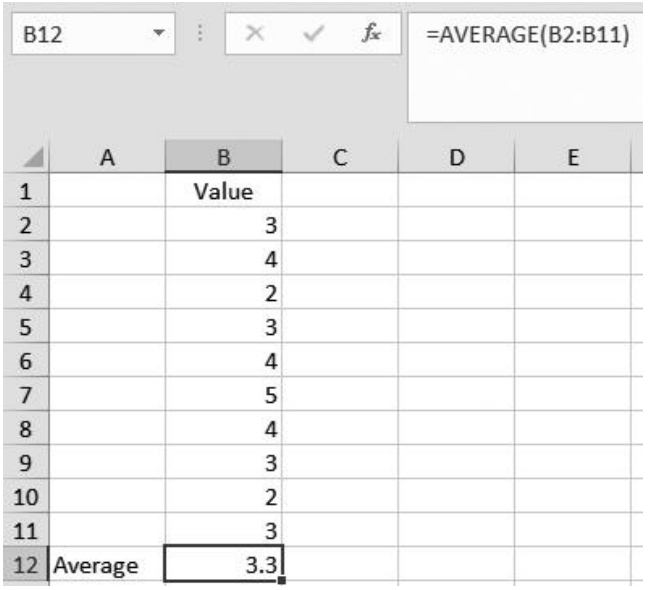
	A	B
1		Value
2		3
3		4
4		2
5		3
6		4
7		5
8		4
9		3
10		2
11		3
12	Average	

2. Type =AVERAGE(B2:B11) in Cell B12. (Using upper case or "capital" letters is the standard way to spell a function, but it will work with lower case letters, too.)
3. Press the Enter key, and presto: As you see in Figure 2.7, the sum shows up in Cell B12, and in the formula bar, you can see the structure of the function.

Notice that the results of the function (3.3) are returned to the same cell (B12) where the function was entered. Pretty cool.

And not very difficult. And very convenient. Remember that you can do this with any function. But how do you know what the structure of the function is? That's where the next step comes in.

FIGURE 2.7 ⬢ The Completed AVERAGE Function



Okay—so how do you know what function to use? Well, certainly one way is through exploring different functions and finding out what they do (which you will do throughout *Statistics for People . . .*). Another is by using Excel Help (press F1 at any time and enter the terms on which you want help). And another way is to look at Table 2.1 on page 33, which gives you a heads-up on which functions we'll be mentioning (some in great detail and others just in passing) throughout the book and what they do.

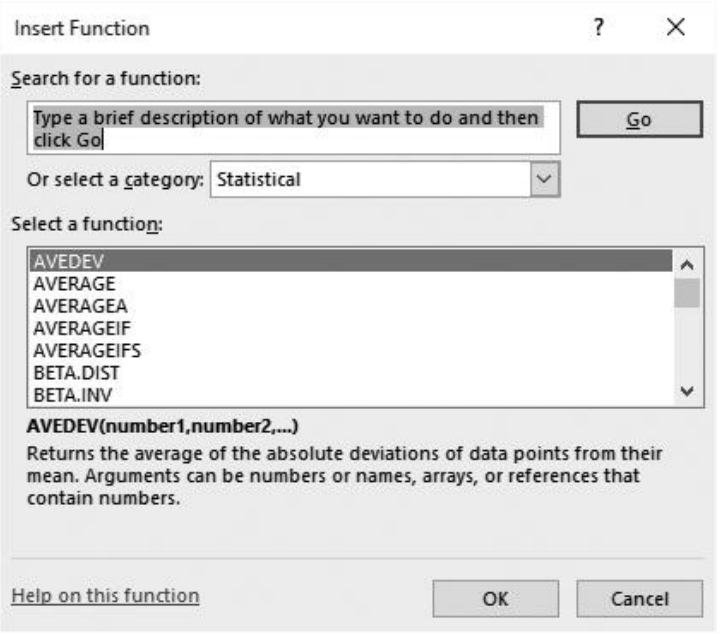
Inserting a Function Using the Insert Function (*fx*) Command

Let's use the same example, the AVERAGE function, and assume you haven't used it before but know this is the one you want to use.

We're using the same data as shown in Figure 2.6. First, erase the results of the function in Cell B12 by selecting the cell and hitting the space bar once and then Enter.

1. Select Cell B12.
2. Click the Formulas tab and the Insert Function command (*fx*). When you do this, you will see the Insert Function dialog box as shown in Figure 2.8. In the Mac version, you see the Formula Builder dialog box, which does the same thing as the Insert Function dialog box in the Windows version.

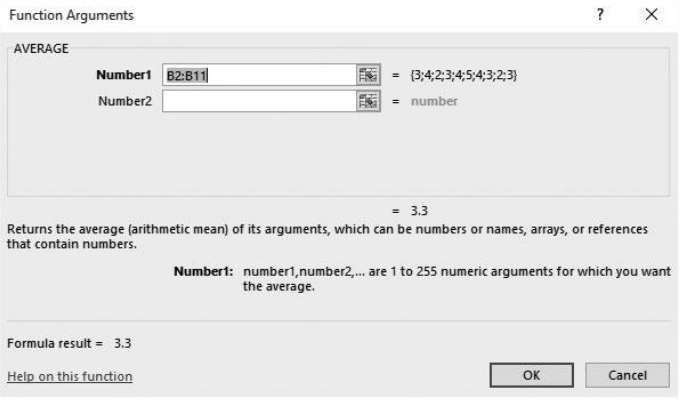
FIGURE 2.8 ⬢ The Insert Function Dialog Box



3. Now you can do one of two things:
   a. Type a brief description of what you want to do, such as *average*, and click the OK button; or
   b. Find the function you want in the list of functions and double-click on it.

We selected option (b), and when we did, the Function arguments dialog box appeared, as shown in Figure 2.9. Notice that Excel automatically assumed that we wanted to average all of the values above the current cell, and it completed the cell range in the Number 1 box.

FIGURE 2.9 ⬢ The Function Arguments Dialog Box



Let's take a look at the different elements in this dialog box.

- There's the name of the function, AVERAGE.
- Then there are text boxes where you enter the range of cells (the argument) on which you want the function to perform its duty. Note that Excel is pretty smart and automatically enters the range of cells it thinks you want to average. Notice also that the actual numbers you want to average are listed to the right of the text box.
- In the middle of the dialog box is a description of what the function does, and above that is the value the function will return to the cell in which it is located (in this case, your average is 3.3). Since we will later learn to distinguish among different averages, notice this specific average is something called the *mean*.
- The syntax (or directions) of how to put the function together is given near the bottom, preceded by **Number 1**.
- The formula result is shown again at the bottom left.
- Finally, there is a link to a place to get help if you need it.

4. Click OK, and you will see the same result as you saw in Figure 2.7. We entered the function using the Insert Function command instead of by typing its name, but we got the same result.

A very nice shortcut to the function command (*fx*) can be found on the Formula bar just to the left of where you see any information that is entered into a cell. Just click that, and you get the Insert Function dialog box.

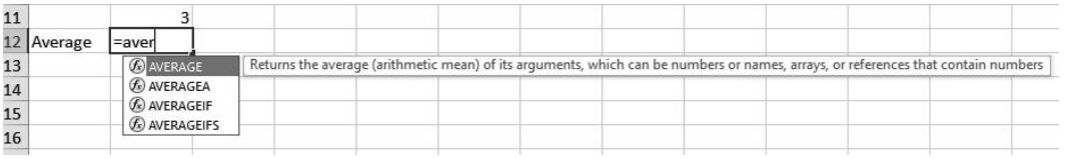
Don't get too excited. A function's argument is not really an argument like a disagreement. An argument in mathematical terms is a set of premises, and that's exactly what you need to provide within the parentheses of any function—a set of premises that the function is to carry out.

Inserting a Function Using Formulas → More Functions → Statistical

That really says it all. Just follow these three mouse clicks, and you will see a list of all the statistical functions that are available. Selecting any of them (such as AVERAGE) provides you with the same dialog box you see in Figure 2.9.

If you enter the following formula . . .	Excel does this . . .
=SUM(3,4)	Adds the values to get 7.
=SUM(A2:A4)	Adds the values located in Cells A2 through A4.
=SUM(A2:A4,6)	Adds the values located in Cells A2 through A4 and also adds the value of 6 to that sum.
=SUM(A6:A8,4)	Adds the values located in Cells A6 to A8 and adds the value of 4 to that sum.

FIGURE 2.10 ⬢ Excel Helps You Choose the Right Function



Most functions can do a lot more than first appears. Excel functions are so useful because they are so flexible. For example, with the simplest of functions such as SUM, you can enter the following variations as arguments and get the following results.

Now you know two ways to insert a function in a worksheet—by typing its name or selecting it through the Insert Function dialog box. And once a specific Function Arguments dialog box (like the one you see in Figure 2.9) is open, you can just enter the cell addresses in the appropriate text box. However, you can also just click in the cell address box and then drag the mouse over the cell addresses you want to include in that box. This is good. But there's another nifty way to go about this. You can click on the Collapse button (which looks like this), which will shrink the entire dialog box and allow you to select the cells you want using the mouse directly on the worksheet. Then click the Expand button, and the dialog box returns to its normal size with the cell addresses included.

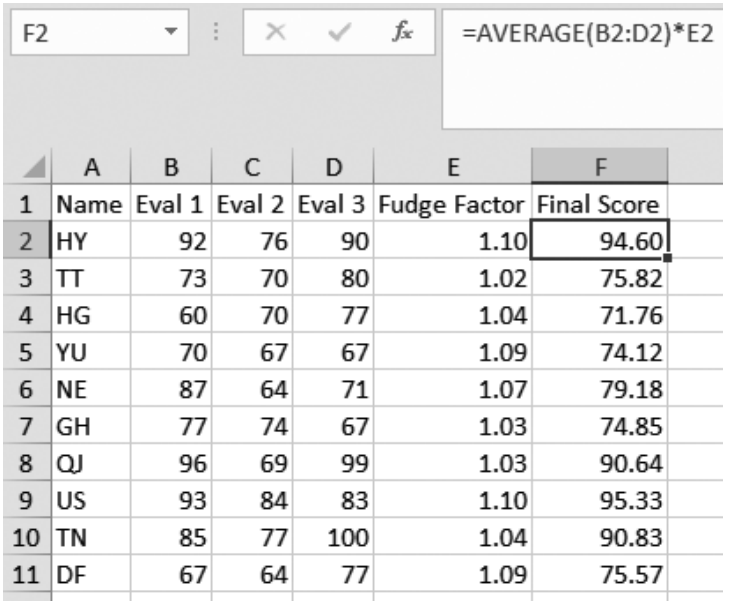
If you insert a function by typing it directly into a cell (or even by typing a partial name), the 2016 version of Excel provides a list of similarly spelled functions as well as a tip about how to use your function, as you see in Figure 2.10. Here we typed in =aver, and as we typed, Excel provided a list of various average functions plus a tip as to what the AVERAGE function does.

Using Functions in Formulas

It's time to get a bit fancy. Now, formulas and functions are basically the same animal—they carry out instructions. So, there's just no reason why you can't include a function in a formula.

For example, let's say that you have three job evaluation scores (Eval 1, Eval 2, and Eval 3) as you see in Figure 2.11. You also have a Fudge Factor (in column E), which is a value you can use to increase or decrease an employee's score at your discretion. For example, you want to increase employee GH's score by 3%, so you multiply the average evaluation score (from Eval 1, Eval 2, and Eval 3) by 1.03. Figure 2.11 shows the formula that includes the AVERAGE function (which you will learn more about in Chapter 2).

FIGURE 2.11 ⬢ Using a Function in a Formula



	A	B	C	D	E	F
1	Name	Eval 1	Eval 2	Eval 3	Fudge Factor	Final Score
2	HY	92	76	90	1.10	94.60
3	TT	73	70	80	1.02	75.82
4	HG	60	70	77	1.04	71.76
5	YU	70	67	67	1.09	74.12
6	NE	87	64	71	1.07	79.18
7	GH	77	74	67	1.03	74.85
8	QJ	96	69	99	1.03	90.64
9	US	93	84	83	1.10	95.33
10	TN	85	77	100	1.04	90.83
11	DF	67	64	77	1.09	75.57

As you can see in the formula bar shown in Figure 2.11, the formula in Cell F2 looks like this:

=AVERAGE(B2:D2)*E2

And it reads like this: The contents of Cells B2 through D2 are averaged, and then that value is multiplied by the contents of Cell E2. We copied the formula from Cell F2 to Cells F3 through F11, and the results are shown in Column F.

We're Taking Names: Naming Ranges

It's certainly easy enough to enter cell addresses such as A1:A3—not much work involved there.

But what if you're dealing with a really large worksheet with hundreds of columns and rows and thousands of cells? Wouldn't it be nice if you could just enter a name that represents a certain range of cells rather than having to remember all those cell addresses? Desire it no more. Excel allows you to name a range, or a collection of cells.

For example, in Figure 2.11, if you want to average the employees' second evaluations, instead of using the cell addresses C2:C11, why not just give the range

of cells a name, such as eval2 or EVAL_2 (no spaces, please!)? Then, the average for that set of scores using the AVERAGE function would look like this—

=AVERAGE(EVAL_2)

rather than like this—

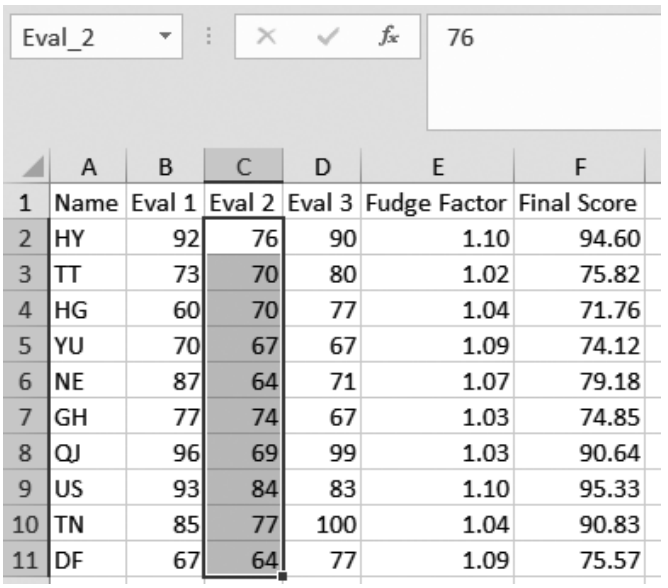
=AVERAGE(C2:C11)

And if you can believe it, this only gets better—you can just paste that name into any formula or function with a few clicks. You don't even have to type anything!

Here's how to assign a name to a range of cells:

1. Highlight the range of cells you want to name.
2. Click the Define Name option in the center of the Formulas bar. You can also get to this option by right clicking and choosing Define Name.
3. Type the name that you want to use to refer to your selection, as shown in Figure 2.12 (we used Eval_2). And again, no spaces please—Excel does not like them in name ranges.

FIGURE 2.12 ⬢ Naming a Selection of Cells as a Range



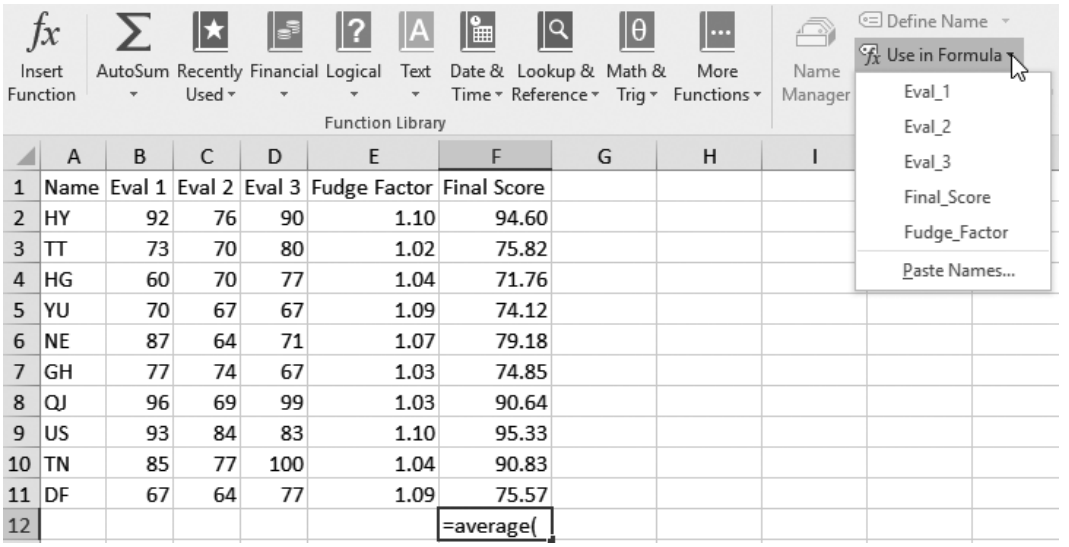
	A	B	C	D	E	F
1	Name	Eval 1	Eval 2	Eval 3	Fudge Factor	Final Score
2	HY	92	76	90	1.10	94.60
3	TT	73	70	80	1.02	75.82
4	HG	60	70	77	1.04	71.76
5	YU	70	67	67	1.09	74.12
6	NE	87	64	71	1.07	79.18
7	GH	77	74	67	1.03	74.85
8	QJ	96	69	99	1.03	90.64
9	US	93	84	83	1.10	95.33
10	TN	85	77	100	1.04	90.83
11	DF	67	64	77	1.09	75.57

4. Press Enter.

Using Ranges

Once a range is defined, you can use the name assigned instead of a cell range. If you remember that the data are named, you can just enter that name by using the drop-down menu now available in the Name Manager command area (on the Formulas tab), as shown in Figure 2.13. Or you'll use the Apply Names option in the Mac version.

FIGURE 2.13 ⬢ Seeing Range Names

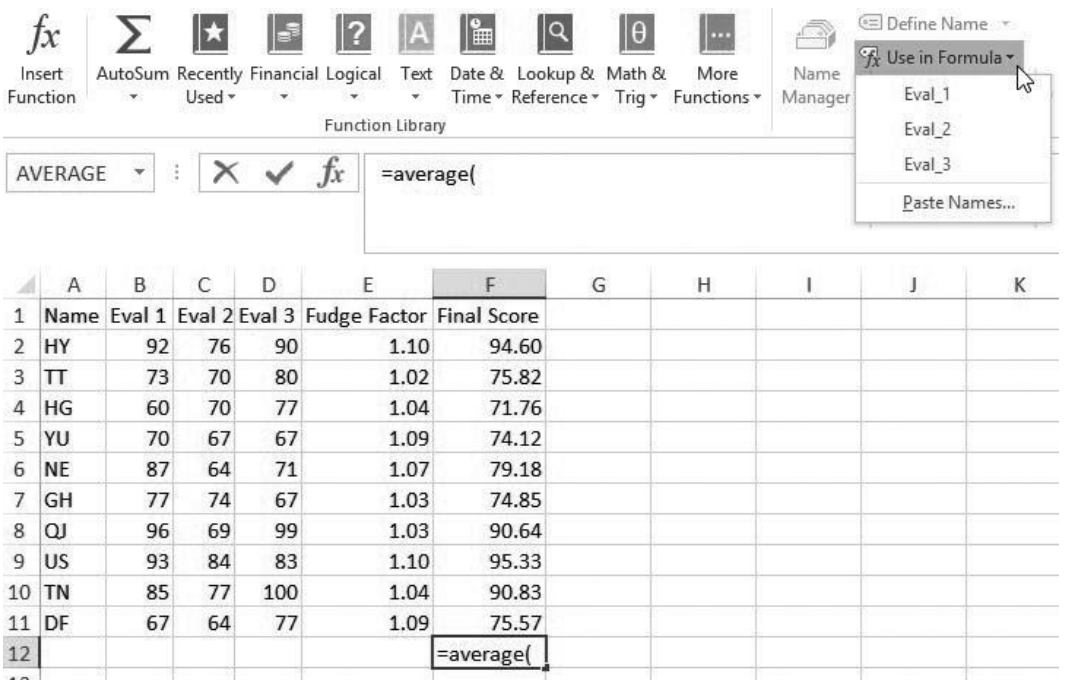


	A	B	C	D	E	F
1	Name	Eval 1	Eval 2	Eval 3	Fudge Factor	Final Score
2	HY	92	76	90	1.10	94.60
3	TT	73	70	80	1.02	75.82
4	HG	60	70	77	1.04	71.76
5	YU	70	67	67	1.09	74.12
6	NE	87	64	71	1.07	79.18
7	GH	77	74	67	1.03	74.85
8	QJ	96	69	99	1.03	90.64
9	US	93	84	83	1.10	95.33
10	TN	85	77	100	1.04	90.83
11	DF	67	64	77	1.09	75.57
12						=average(

Let's use the ranges that were defined and compute the average of all the Fudge Factor scores.

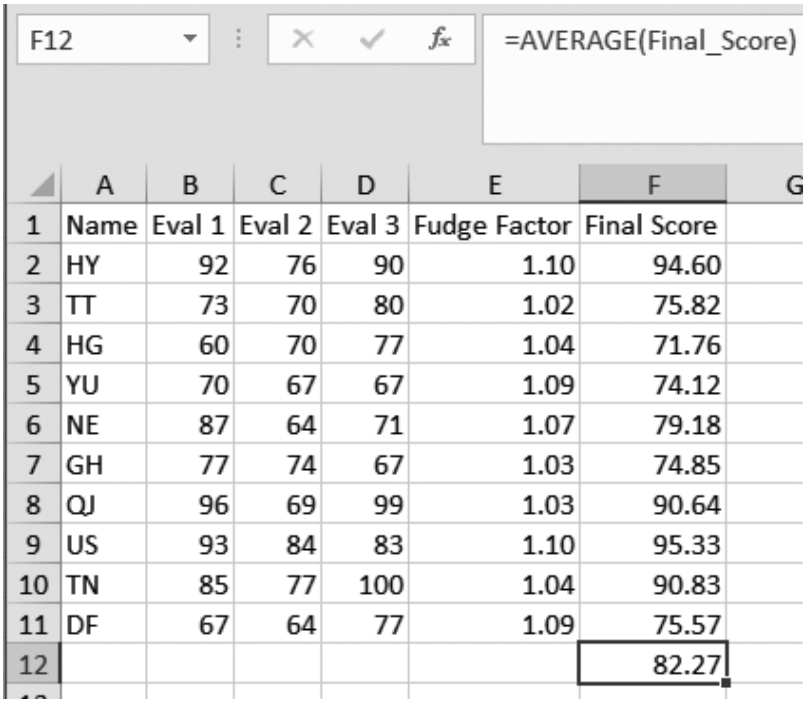
1. Click on Cell F12, where the average will be placed.
2. Type =average(.
3. Click Formulas → Use in Formula, and you will see the Use in Formula drop-down menu, as shown in Figure 2.14.
4. Click on Final_Score.
5. Type).
6. Click Enter and take a look at Figure 2.15, where you can see how the name was used in the function rather than the cell addresses of F2 through F11.

FIGURE 2.14 ⬢ The Menu for Selecting a Range of Cells



	A	B	C	D	E	F	G	H	I	J	K
1	Name	Eval 1	Eval 2	Eval 3	Fudge Factor	Final Score					
2	HY	92	76	90	1.10	94.60					
3	TT	73	70	80	1.02	75.82					
4	HG	60	70	77	1.04	71.76					
5	YU	70	67	67	1.09	74.12					
6	NE	87	64	71	1.07	79.18					
7	GH	77	74	67	1.03	74.85					
8	QJ	96	69	99	1.03	90.64					
9	US	93	84	83	1.10	95.33					
10	TN	85	77	100	1.04	90.83					
11	DF	67	64	77	1.09	75.57					
12						=average(					

FIGURE 2.15 ⬢ Inserting a Cell Range Into a Function



	A	B	C	D	E	F	G
1	Name	Eval 1	Eval 2	Eval 3	Fudge Factor	Final Score	
2	HY	92	76	90	1.10	94.60	
3	TT	73	70	80	1.02	75.82	
4	HG	60	70	77	1.04	71.76	
5	YU	70	67	67	1.09	74.12	
6	NE	87	64	71	1.07	79.18	
7	GH	77	74	67	1.03	74.85	
8	QJ	96	69	99	1.03	90.64	
9	US	93	84	83	1.10	95.33	
10	TN	85	77	100	1.04	90.83	
11	DF	67	64	77	1.09	75.57	
12						82.27	