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* 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 SPSS?

Throughout this book, you'll be shown how to use SPSS, a statistical analysis tool, for the analysis of data. No worries; you'll also be shown how to do the same analysis by hand to assure you of an understanding of both.

Why SPSS? Simple. It's one of the most popular, most powerful analytic tools 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 SPSS as their primary computational tool, and you can look to Appendix A for a refresher on some basic SPSS tasks. Also, the way that technology is advancing, few opportunities to use statistics in research, administration, and everyday work will not require some knowledge of how and when to use tools such as SPSS. That's why we're including it in this book! We will show you how to use it to make your statistics learning experience a better one.

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) are 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 I need to trade for one of those?" and "Uh oh, there are more of them than there are of us!" and "Yes, little one, you may keep one pet saber-toothed tiger, but not more than one"), collecting information became a useful skill.

If counting counted, then one could define the seasons by how often the sun rose and set, 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. This was when the use of *descriptive statistics* began, which we will talk about later. From that point on, scientists (mostly mathematicians at first but, later, 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! (It's probably not.) 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. 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.

PEOPLE WHO LOVED STATISTICS

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 statisti-

cal 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.

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 (no kidding—see the *Wall Street Journal* article at http://www.wsj.com/articles/SB10001424052702304071004579409071015745370), 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.

If you want to learn more about the history of statistics and see a historical time line, great places to start are Saint Anselm's College at www.anselm.edu/homepage/jpitocch/biostatshist.html and the University of California–Los Angeles at www.stat.ucla.edu/history.

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 is a book about basic statistics and how to apply them to a variety of different situations, including the analysis and understanding of information, especially when that information is expressed as numbers and quantities.

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 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 of 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.

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* 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* 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 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.)

Throughout this book, you'll 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 SPSS to do these steps. These steps have been tested and approved by whatever federal agency approves these things.

REAL-WORLD STATS

Real-World Stats will appear at the end of every chapter as appropriate and, it is hoped, will provide you with a demonstration of how a particular method, test, idea, or some aspect of statistics is used in the everyday workplace of scientists, physicians, policy makers, government folks, and others. In this first such exploration, we look at a very short paper where the author recalls and shares the argument that the National Academy of Sciences (first chartered in 1863, by the way!) "shall, whenever called upon by any department of the Government, investigate, examine, experiment, and report upon any subject of science or art." This charter, some 50 years later in 1916, led to the formation of the National Research Council, another federal body that helped provide information that policy makers need to make informed decisions. And often this "information" takes the form of quantitative data—also referred to as statistics—that assist people in evaluating alternative approaches to problems that have a wide-ranging impact on the public. So, this article, as does your book and the class you are taking, points out how important it is to think clearly and use data to support your arguments.

Want to know more? Go online or go to the library, and find . . .

Cicerone, R. (2010). The importance of federal statistics for advancing science and identifying policy options. *The Annals of the American Academy of Political and Social Science, 631,* 25–27.

Appendix A contains an introduction to SPSS. Working through this appendix is all you really need to do to be ready to use SPSS. If you have an earlier version of SPSS (or the Mac version), you will still find this material to be very helpful. In fact, the latest Windows and Mac versions of SPSS are almost identical in appearance and functionality.

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/salkindfrey7e**.

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 SPSS 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)

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.

Time to Practice

Because there's no substitute for the real thing, Chapters 1 through 17 and Chapter 19 will end with a set of exercises that will help you review the material that was covered in the chapter. As noted earlier, the answers to these exercises can be found near the end of the book in Appendix D.

For example, here is the first set of exercises (but don't look for any answers for these because these are kind of "on your own" answers—each question and answer are highly tied to your own experiences and interest).

1. Interview someone who uses statistics in his or her everyday work. It might be your advisor, an instructor, a researcher who lives on your block, a health care analyst, a marketer for a company, a city planner, or . . . Ask the person what his or her first statistics course was like. Find out what the person liked and didn't like. See if this individual has any suggestions to help you succeed. And most important, ask the person about how he or she uses these new-to-you tools at work.
2. We hope that you are part of a study group or, if that is not possible, that you have a phone, email, instant messaging, or webcam study buddy (or even more than one). And, of course, plenty of texting and Facebook friends. Talk to your group or a fellow student in your class about similar likes, dislikes, fears, and so on about the statistics course. What do you have in common? Not in common? Discuss with your fellow student strategies to overcome your fears.
3. Search through your local newspaper (or any other publication) and find the results of a survey or interview about any topic. Summarize the results and do the best job you can describing how the researchers who were involved, or the authors of the survey, came to the conclusions they did. Their methods and reasoning may or may not be apparent. Once you have some idea of what they did, try to speculate as to what other ways the same information might be collected, organized, and summarized.
4. Go to the library (either in person or online) and find a copy of a journal article in your own discipline. Then, go through the article and highlight the section (usually the "Results" section) where statistical procedures were used to organize and analyze the data. You don't know much about the specifics of this yet, but how many different statistical procedures (such as *t* test, mean, and calculation of the standard deviation) can you identify? Can you take the next step and tell your instructor how the results relate to the research question or the primary topic of the research study?

5. Find five websites that contain data on any topic and write a brief description of what type of information is offered and how it is organized. For example, if you go to the mother of all data sites, the U.S. Census (http://www.census.gov), you'll find links to hundreds of databases, tables, and other informative tools. Try to find data and information that fit in your own discipline.
6. And the big extra-credit assignment is to find someone who actually uses SPSS for daily data analysis needs. Ask if there is anything specific about SPSS that makes it stand out as a tool for their type of data analysis. You may very well find these good folks in everything from political science to nursing, so search widely!
7. Finally, as your last in this first set of exercises, come up with five of the most interesting questions you can about your own area of study or interest. Do your best to come up with questions for which you would want real, existing information or data to answer. Be a scientist!

Student Study Site

Get the tools you need to sharpen your study skills! Visit **edge.sagepub.com/salkindfrey7e** to access practice quizzes, eFlashcards, original and curated videos, data sets, and more!