1 Statistics or Sadistics?

It's Up to You

Difficulty Scale ☺☺☺☺☺ (really easy)

WHAT YOU'LL LEARN ABOUT 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," "The math involved is impossible," "I don't know how to use a computer," "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 about at least one of the above, if not actually sharing it with another student, their spouse, a colleague, or a friend.

And all kidding aside, there are some statistics courses that can easily be described as sadistics. That's because the books are repetitiously boring, and the authors have no imagination.

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.

A FIVE-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, and anthropology (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 also 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 cousin of Charles Darwin, by the way), who lived from 1822 to 1911, was very interested in the nature of human intelligence. 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.

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 SPSS®, an IBM company (see Appendix A), 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 $1,000 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 passed on taking this course!) will take any old data they have and think that by running them through some sophisticated SPSS 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 reliable and trustworthy data, what you'll have after your data are analyzed are unreliable and untrustworthy results.

Today, statisticians in all different areas from criminal justice to geophysics to psychology 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) 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 three or four credits.

If you want to learn more about the history of statistics and see a historical time line, a great place to start is a Saint Anselm's College Internet site located at http://www.anselm.edu/homepage/jpitocch/biostatshist.html and http://www.stat.ucla.edu/history (at the University of California at Los Angeles). Tons of good stuff at both places.

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.

In the most general sense, **statistics** describes a set of tools and techniques that is 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 patient complaints when using 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 Sante Fe.

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**.

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 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.

Name	Major	Age	Name	Major	Age
Richard	Education	19	Elizabeth	English	21
Sara	Psychology	18	Bill	Psychology	22
Andrea	Education	19	Hadley	Psychology	23
Steven	Psychology	21	Buffy	Education	21
Jordan	Education	20	Chip	Education	19
Pam	Education	24	Homer	Psychology	18
Michael	Psychology	21	Margaret	English	22
Liz	Psychology	19	Courtney	Psychology	24
Nicole	Chemistry	19	Leonard	Psychology	21
Mike	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. 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 from a smaller group of data (such as our group of 22 students) to a possibly larger one (such as all the undergraduate students in the College of Arts and Sciences).

This 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, NJ, would be a population (it's all the occurrences with certain characteristics—being in fifth grade and living in Newark), whereas a selection of 150 of them would be a sample.

Let's look at another example. Your marketing agency asks you (a newly hired researcher) to determine which of several different 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 that is representative of all potato chip fans and ask them to tell you which one of the three names they like the most. Then, if you did things right, you can easily infer 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 (or a substance that is known not to have any effect) as another alternative, and even nothing as the third alternative to see what happens. Well, you find out that a larger number of patients get better when no action is taken and nature just takes its course! The drug does not have any effect. Then, with that information, you infer to the larger group of patients that suffers from the disease, given the results of your experiment.

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 one you use and when depends on the question you want answered.

WHAT AM I DOING IN A STATISTICS CLASS?

There are probably many reasons why you find yourself using this book. 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 and review for a more advanced class.

In any case, you're a statistics student whether you have to take a final exam at the end of a formal course, you're just in it of your own accord, or you're taking the course online 500 miles from the instructor.

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 my students hear at the beginning of our introductory statistics course.

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. In tough job markets, an edge like this is very important.

2. If this is not a required course, taking basic statistics sets you apart from those students who do not. It shows that you are willing to undertake a course that is above average in difficulty and commitment.

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 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, behavior, and biological pursuits, this course will give you the foundation you need to go further.

6. 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's 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, letting 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 a few 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 like any other new course. Attend the lectures, study the material, and 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. You do have to show up to class and do the homework.

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 work hard 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 life. 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 general idea of what they think. Just don't base it on one spoilsport's experience.

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 (we hope) refer back to this book and use it 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 OK 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 one of the most basic ways to ensure some success in this course. Early in the semester, arrange to study with friends. 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 to be as happy about being there as you are. Studying with others allows you to help them if you know the material better, or to benefit from others who know that 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. How do the exercises do that? An explanation for how each exercise is solved accompanies the problem. If you can answer these end-of-chapter exercises, then you are well on your way to mastering the content of the chapter.

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 in the back of Chapters 1–17 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 look and 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 better and more full 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 indeed 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 THOSE ICONS

An icon is a symbol. Throughout *Statistics for People . . .*, you'll see a variety of different icons. Here's what each one is and what each represents:

This icon represents information that goes beyond the regular text. We might find it necessary to elaborate on a particular point, and we can do that more easily outside of the flow of the usual material.

Here, we select some more technical ideas and tips to discuss and to inform you about what's beyond the scope of this course. You might find these interesting and useful.

Throughout *Statistics for People . . .*, you'll find a small steps icon like the one you see here. This indicates that there is a set of steps coming up that will direct you through a particular process. These steps have been tested and approved by whatever federal agency approves these things.

That finger with the bow is a cute icon, but its primary purpose is to help reinforce important points about the topic that you just read about. Try to emphasize these points in your studying because they are usually central to the topic.

Most of the chapters in *Statistics for People . . .*, provide detailed information about one or more particular statistical procedure and the computation that accompanies them. The computer icon is used to identify the "Using the Computer to . . ." section of the chapter.

Many of these chapters also contain instructions for using Version 19.0 of SPSS to complete the same procedures so that you can have both hands-on experience and experience using one of the most powerful statistical analysis packages available today.

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.

In working the exercises in this book, you will use the data sets in Appendix C, starting on page 366. Also, when you get to each section titled "Using the Computer to . . .," you'll find reference to a data set (such as "Chapter 2 Data Set 1"). Each of these sets is shown in Appendix C, and you will use these data to successfully

complete the "Using the Computer to . . ." sections if you want to follow along. You can enter the data manually or download it from either the website hosted by SAGE at http://www.sagepub.com/salkind4e or the author's website at http://www.onlinefilefolder.com with

User name: *ancillaries*
Password: *files*
The data files are available in either SPSS or Excel format.

KEY TO DIFFICULTY INDEX

1. very hard ☺
2. hard ☺☺
3. not too hard, but not easy either ☺☺☺
4. easy ☺☺☺☺
5. 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 1–17 end with a set of exercises that will help you review the material that was covered in the chapter.

For example, here is the first set of exercises.

1. Interview someone who uses statistics in his or her everyday work. It might be your adviser, an instructor, a researcher who lives on your block, a market analyst for a company, or even a city planner. Ask them what their first statistics

course was like. Find out what they liked and what they didn't. See if they have any suggestions to help you succeed. And most important, ask the person about the ways 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 telephone or online study buddy (or even more than one). 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 and find the results of a survey or interview about any topic. Summarize what the results are 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. It 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 and copy a journal article in your own discipline. Then go through the article with one of those fancy highlighters 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 of these different 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 on the Internet 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 United States Census (at http://www.census.gov/), you'll find a link to Data Tools, which takes you to a page just loaded with links to real live data. Try to find data and information that fits in your own discipline.

6. 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. Be a scientist!