1.1 Why Study Statistics?

There is an old saying that "without data, you are just another person with an opinion." While anecdotes and coincidences may make for interesting stories, you wouldn't want to make important decisions on the basis of anecdotes alone. For example, just because a friend of a friend ate 16 apricots and then experienced relief from joint pain doesn't mean that this is all you need to know to help one of your parents choose a treatment for arthritis! Before recommending the apricot treatment, you would definitely want to consider relevant data—that is, data that would allow you to investigate the effectiveness of this treatment.

It is difficult to function in today's world without a basic understanding of statistics. For example, here are a few headlines from articles that draw conclusions based on data that all appeared in a single issue of *USA Today* (June 29, 2009):

- "Infant Colic May Be Linked to Dads" is the headline of an article reporting on a study of the relationship between excessive crying and parents' depression. The study of more than 7600 babies and their parents concluded that excessive newborn crying is more likely to occur if the father reported symptoms of depression prior to the birth of the baby.
- The article "Many Adults Can't Name a Scientist" summarized the results of a survey of 1000 adults. Of those surveyed, 23% were unable to name a single famous scientist. Of those who did come up with a name, Albert Einstein was the scientist of choice, named by 47% of those surveyed.
- "Few See Themselves as 'Old' No Matter What Their Age" is the title of an article that described results from a large survey of 2969 adults. Those surveyed were asked at what age a person would be considered old. The resulting data revealed that there were notable differences in the answer to the question depending on the age of the responder. The average age identified as old by young adults (age 18–29) was 60, while the average was 69 for those who were age 30 to 49, 72 for those age 50 to 64, and 74 for those age 65 and older.
- The article "Poll Finds Generation Gap Biggest Since Vietnam War" summarized a study that explored opinions regarding social values and political views. Not surprisingly, large behavioral differences between young and old were noted in the use of the Internet, cell phones, and text messaging.
- The graph titled "If you were given $1000, what would you do?" reported on one aspect of a study of consumer purchasing and saving behavior. Something was definitely amiss in this report, however—the percentages for the response categories (save it, pay off credit card debt, use it for a vacation, etc.) added up to 107%!

To be an informed consumer of reports such as those described above, you must be able to do the following:

1. Extract information from tables, charts, and graphs.
2. Follow numerical arguments.
3. Understand the basics of how data should be gathered, summarized, and analyzed to draw statistical conclusions.

Your statistics course will help prepare you to perform these tasks.

Studying statistics will also enable you to collect data in a sensible way and then use the data to answer questions of interest. In addition, studying statistics will allow you to critically evaluate the work of others by providing you with the tools you need to make informed judgments. Throughout your personal and professional life, you will need to understand and use data to make decisions. To do this, you must be able to

1. Decide whether existing data is adequate or whether additional information is required.
2. If necessary, collect more information in a reasonable and thoughtful way.
3. Summarize the available data in a useful and informative manner.
4. Analyze the available data.
5. Draw conclusions, make decisions, and assess the risk of an incorrect decision.

People informally use these steps when making everyday decisions. Should you go out for a sport that involves the risk of injury? Will your college club do better by trying to raise funds with a benefit concert or with a direct appeal for donations? If you choose a particular major, what are your chances of finding a job when you graduate? How should you select a graduate program based on guidebook ratings that include information on percentage of applicants accepted, time to obtain a degree, and so on? The study of statistics formalizes the process of making decisions based on data and provides the tools for accomplishing the steps listed.

We hope that this textbook will help you to understand the logic behind statistical reasoning, prepare you to apply statistical methods appropriately, and enable you to recognize when statistical arguments are faulty.

1.2 The Nature and Role of Variability

Statistical methods allow us to collect, describe, analyze and draw conclusions from data. If we lived in a world where all measurements were identical for every individual, these tasks would be simple. Imagine a population consisting of all students at a particular university. Suppose that *every* student was enrolled in the same number of courses, spent exactly the same amount of money on textbooks this semester, and favored increasing student fees to support expanding library services. For this population, there is *no* variability in number of courses, amount spent on books, or student opinion on the fee increase. A researcher studying students from this population to draw conclusions about these three variables would have a particularly easy task. It would not matter how many students the researcher studied or how the students were selected. In fact, the researcher could collect information on number of courses, amount spent on books, and opinion on the fee increase by just stopping the next student who happened to walk by the library. Because there is no variability in the population, this one individual would provide complete and accurate information about the population, and the researcher could draw conclusions with no risk of error.

The situation just described is obviously unrealistic. Populations with no variability are exceedingly rare, and they are of little statistical interest because they present no challenge! In fact, variability is almost universal. It is variability that makes life (and the life of a statistician, in particular) interesting. We need to understand variability to be able to collect, describe, analyze, and draw conclusions from data in a sensible way.

Examples 1.1 and 1.2 illustrate how describing and understanding variability are the keys to learning from data.

EXAMPLE 1.1 If the Shoe Fits

The graphs in Figure 1.1 are examples of a type of graph called a histogram. (The construction and interpretation of such graphs is discussed in Chapter 3.) Figure 1.1(a) shows the distribution of the heights of female basketball players who played at a particular university between 2000 and 2008. The height of each bar in the graph indicates how many players' heights were in the corresponding interval. For example, 40 basketball players had heights between 72 inches and 74 inches, whereas only 2 players had heights between 66 inches and 68 inches Figure 1.1(b) shows the distribution of heights for members of the women's gymnastics team. Both histograms are based on the heights of 100 women.

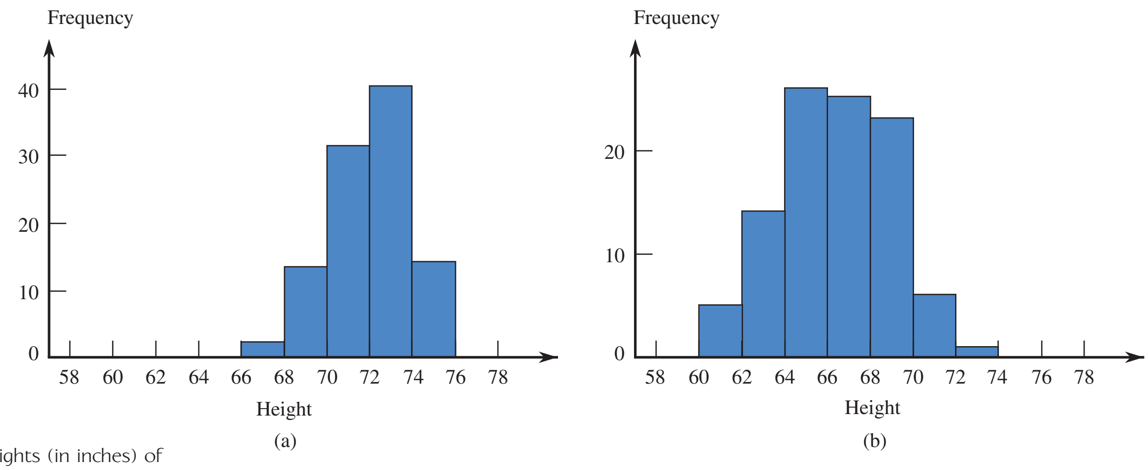

FIGURE 1.1
Histograms of heights (in inches) of female athletes: (a) basketball players; (b) gymnasts.

The first histogram shows that the heights of female basketball players varied, with most heights falling between 68 inches and 76 inches. In the second histogram we see that the heights of female gymnasts also varied, with most heights in the range of 60 inches to 72 inches It is also clear that there is more variation in the heights of the gymnasts than in the heights of the basketball players, because the gymnast histogram spreads out more about its center than does the basketball histogram.

Now suppose that a tall woman (5 feet 11 inches) tells you she is looking for her sister who is practicing with her team at the gym. Would you direct her to where the basketball team is practicing or to where the gymnastics team is practicing? What reasoning would you use to decide? If you found a pair of size 6 shoes left in the locker room, would you first try to return them by checking with members of the basketball team or the gymnastics team?

You probably answered that you would send the woman looking for her sister to the basketball practice and that you would try to return the shoes to a gymnastics team member. To reach these conclusions, you informally used statistical reasoning that combined your own knowledge of the relationship between heights of siblings and between shoe size and height with the information about the distributions of heights presented in Figure 1.1. You might have reasoned that heights of siblings tend to be similar and that a height as great as 5 feet 11 inches, although not impossible, would be unusual for a gymnast. On the other hand, a height as tall as 5 feet 11 inches would be a common occurrence for a basketball player. Similarly, you might have reasoned that tall people tend to have bigger feet and that short people tend to have smaller feet. The shoes found were a small size, so it is more likely that they belong to a gymnast than to a basketball player, because small heights and so small feet are usual for gymnasts and unusual for basketball players.

EXAMPLE 1.2 Monitoring Water Quality

© David Chasey/Photodisc/Getty Images

As part of its regular water quality monitoring efforts, an environmental control board selects five water specimens from a particular well each day. The concentration of contaminants in parts per million (ppm) is measured for each of the five specimens, and then the average of the five measurements is calculated. The histogram in Figure 1.2 summarizes the average contamination values for 200 days.

Now suppose that a chemical spill has occurred at a manufacturing plant 1 mile from the well. It is not known whether a spill of this nature would contaminate groundwater in the area of the spill and, if so, whether a spill this distance from the well would affect the quality of well water.

One month after the spill, five water specimens are collected from the well, and the average contamination is 15.5 ppm. Considering the variation before the spill, would you interpret this as convincing evidence that the well water was affected by the spill? What if the calculated average was 17.4 ppm? 22.0 ppm? How is your reasoning related to the histogram in Figure 1.2?

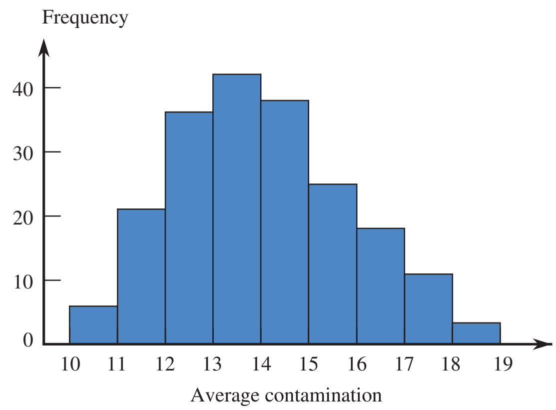

FIGURE 1.2
Frequency of average contamination concentration (in parts per million) in well water.

Before the spill, the average contaminant concentration varied from day to day. An average of 15.5 ppm would not have been an unusual value, so seeing an average of 15.5 ppm after the spill isn't necessarily an indication that contamination has increased. On the other hand, an average as large as 17.4 ppm is less common, and an average as large as 22.0 ppm is not at all typical of the pre-spill values. In this case, we would probably conclude that the well contamination level has increased.

In these two examples, reaching a conclusion required an understanding of variability. Understanding variability allows us to distinguish between usual and unusual values. The ability to recognize unusual values in the presence of variability is an important aspect of most statistical procedures and is also what enables us to quantify the chance of being incorrect when a conclusion is based on data. These concepts will be developed further in subsequent chapters.

1.3 Statistics and the Data Analysis Process

Statistics involves collecting, summarizing, and analyzing data. All three tasks are critical. Without summarization and analysis, raw data are of little value, and even sophisticated analyses can't produce meaningful information from data that were not collected in a sensible way.

Statistical studies are undertaken to answer questions about our world. Is a new flu vaccine effective in preventing illness? Is the use of bicycle helmets on the rise? Are injuries that result from bicycle accidents less severe for riders who wear helmets than for those who do not? How many credit cards do college students have? Do engineering students pay more for textbooks than do psychology students? Data collection and analysis allow researchers to answer such questions.

The data analysis process can be viewed as a sequence of steps that lead from planning to data collection to making informed conclusions based on the resulting data. The process can be organized into the following six steps:

1. **Understanding the nature of the problem.** Effective data analysis requires an understanding of the research problem. We must know the goal of the research and what questions we hope to answer. It is important to have a clear direction before gathering data to ensure that we will be able to answer the questions of interest using the data collected.
2. **Deciding what to measure and how to measure it.** The next step in the process is deciding what information is needed to answer the questions of interest. In some cases, the choice is obvious (for example, in a study of the relationship between the weight of a Division I football player and position played, you would need to collect data on player weight and position), but in other cases the choice of information is not as straightforward (for example, in a study of the relationship between preferred learning style and intelligence, how would you define learning style and measure it and what measure of intelligence would you use?). It is important to carefully define the variables to be studied and to develop appropriate methods for determining their values.
3. **Data collection.** The data collection step is crucial. The researcher must first decide whether an existing data source is adequate or whether new data must be collected. Even if a decision is made to use existing data, it is important to understand how the data were collected and for what purpose, so that any resulting limitations are also fully understood and judged to be acceptable. If new data are to be collected, a careful plan must be developed, because the type of analysis that is appropriate and the subsequent conclusions that can be drawn depend on how the data are collected.
4. **Data summarization and preliminary analysis.** After the data are collected, the next step usually involves a preliminary analysis that includes summarizing the data graphically and numerically. This initial analysis provides insight into important characteristics of the data and can provide guidance in selecting appropriate methods for further analysis.
5. **Formal data analysis.** The data analysis step requires the researcher to select and apply statistical methods. Much of this textbook is devoted to methods that can be used to carry out this step.
6. **Interpretation of results.** Several questions should be addressed in this final step. Some examples are: What can we learn from the data? What conclusions can be drawn from the analysis? and How can our results guide future research? The interpretation step often leads to the formulation of new research questions, which, in turn, leads back to the first step. In this way, good data analysis is often an iterative process.

For example, the admissions director at a large university might be interested in learning why some applicants who were accepted for the fall 2010 term failed to enroll at the university. The population of interest to the director consists of all accepted applicants who did not enroll in the fall 2010 term. Because this population is large and it may be difficult to contact all the individuals, the director might decide to collect data from only 300 selected students. These 300 students constitute a sample.

DEFINITION

The entire collection of individuals or objects about which information is desired is called the **population** of interest. A **sample** is a subset of the population, selected for study.

Deciding how to select the 300 students and what data should be collected from each student are steps 2 and 3 in the data analysis process. The next step in the process involves organizing and summarizing data. Methods for organizing and summarizing data, such as the use of tables, graphs, or numerical summaries, make up the branch of statistics called **descriptive statistics**. The second major branch of statistics, **inferential statistics**, involves generalizing from a sample to the population from which it was selected. When we generalize in this way, we run the risk of an incorrect conclusion, because a conclusion about the population is based on incomplete information. An important aspect in the development of inferential techniques involves quantifying the chance of an incorrect conclusion.

DEFINITION

Descriptive statistics is the branch of statistics that includes methods for organizing and summarizing data. **Inferential statistics** is the branch of statistics that involves generalizing from a sample to the population from which the sample was selected and assessing the reliability of such generalizations.

Example 1.3 illustrates the steps in the data analysis process.

EXAMPLE 1.3 The Benefits of Acting Out

A number of studies have reached the conclusion that stimulating mental activities can lead to improved memory and psychological wellness in older adults. The article "A Short-Term Intervention to Enhance Cognitive and Affective Functioning in Older Adults" (*Journal of Aging and Health* [2004]: 562–585) describes a study to investigate whether training in acting has similar benefits. Acting requires a person to consider the goals of the characters in the story, to remember lines of dialogue, to move on stage as scripted, and to do all of this at the same time. The researchers conducting the study wanted to see if participation in this type of complex multitasking would show an improvement in the ability to function independently in daily life. Participants in the study were assigned to one of three groups. One group took part in an acting class for 4 weeks, one group spent a similar amount of time in a class on visual arts, and the third group was a comparison group (called the "no-treatment group") that did not take either class. A total of 124 adults age 60 to 86 participated in the study. At the beginning of the 4-week study period and again at the end of the 4-week study period, each participant took several tests designed to measure problem solving, memory span, self-esteem, and psychological well-being. After analyzing the data from this study, the researchers concluded that those in the acting group showed greater gains than both the visual arts group and the no-treatment group in both problem solving and psychological well-being. Several new areas of research were suggested in the discussion that followed the analysis. The researchers wondered whether the effect of studying writing or music would be similar to what was observed for acting and described plans to investigate this further. They also noted that the participants in this study were generally well educated and recommended study of a more diverse group before generalizing conclusions about the benefits of studying acting to the larger population of all older adults.

This study illustrates the nature of the data analysis process. A clearly defined research question and an appropriate choice of how to measure the variables of interest (the tests used to measure problem solving, memory span, self-esteem, and psychological well-being) preceded the data collection. Assuming that a reasonable method was used to collect the data (we will see how this can be evaluated in Chapter 2) and that appropriate methods of analysis were employed, the investigators reached the conclusion that the study of acting showed promise. However, they recognized the limitations of the study, which in turn led to plans for further research. As is often the case, the data analysis cycle led to new research questions, and the process began again.

Evaluating a Research Study The six data analysis steps can also be used as a guide for evaluating published research studies. The following questions should be addressed as part of a study evaluation:

- What were the researchers trying to learn? What questions motivated their research?
- Was relevant information collected? Were the right things measured?
- Were the data collected in a sensible way?
- Were the data summarized in an appropriate way?
- Was an appropriate method of analysis used, given the type of data and how the data were collected?
- Are the conclusions drawn by the researchers supported by the data analysis?

Example 1.4 illustrates how these questions can guide an evaluation of a research study.

EXAMPLE 1.4 Afraid of Spiders? You Are Not Alone!

Spider phobia is a common anxiety-producing disorder. In fact, the American Psychiatric Association estimates that between 7% and 15.1% of the population experiences spider phobia. An effective treatment for this condition involves participating in a therapist-led session in which the patient is exposed to live spiders. While this type of treatment has been shown to work for a large proportion of patients, it requires one-on-one time with a therapist trained in this technique. The article "Internet-Based Self-Help versus One-Session Exposure in the Treatment of Spider Phobia" (*Cognitive Behaviour Therapy* [2009]: 114–120), presented results from a study that compared the effectiveness of online self-help modules to in-person treatment. The article states

> A total of 30 patients were included following screening on the Internet and a structured clinical interview. The Internet treatment consisted of five weekly text modules, which were presented on a web page, a video in which exposure was modeled, and support provided via Internet. The live-exposure treatment was delivered in a 3-hour session following a brief orientation session. The main outcome measure was the behavioral approach test (BAT), and the authors used questionnaires measuring anxiety symptoms and depression as secondary measures. Results showed that the groups did not differ at post-treatment or follow-up, with the exception of the proportion showing clinically significant change on the BAT. At post-treatment, 46.2% of the Internet group and 85.7% of the live-exposure group achieved this change. At follow-up, the corresponding figures were 66.7% for the Internet group and 72.7% for the live treatment.

The researchers concluded that online treatment is a promising new approach for the treatment of spider phobia.

The researchers here had a well-defined research question—they wanted to know if online treatment is as effective as in-person exposure treatment. They were interested in this question because online treatment does not require individual time with a therapist, and so, if it works, it might be able to help a larger group of people at a much lower cost. The researchers noted which treatment was received and also recorded results of the BAT and several other measures of anxiety and depression. Participants in the study took these tests prior to beginning treatment, at the end of treatment, and 1 year after the end of treatment. This allowed the researchers to evaluate the immediate and long-term effects of the two treatments and to address the research question.

To assess whether the data were collected in a sensible way, it would be useful to know how the participants were selected and how it was determined which of the two treatments a particular participant received. The article indicates that participants were recruited through advertisements and articles in local newspapers and that most were female university students. We will see in Chapter 2 that this may limit our ability to generalize the results of this study. The participants were assigned to one of the two treatments at random, which is a good strategy for ensuring that one treatment does not tend to be favored over the other. The advantages of random assignment in a study of this type are also discussed in Chapter 2.

We will also have to delay discussion of the data analysis and the appropriateness of the conclusions because we do not yet have the necessary tools to evaluate these aspects of the study.

Many other interesting examples of statistical studies can be found in *Statistics: A Guide to the Unknown* and in *Forty Studies That Changed Psychology: Exploration into the History of Psychological Research* (the complete references for these two books can be found in the back of the book).

EXERCISES 1.1 - 1.11

1.1 Give a brief definition of the terms *descriptive statistics* and *inferential statistics.*

1.2 Give a brief definition of the terms *population* and *sample.*

1.3 Data from a poll conducted by Travelocity led to the following estimates: Approximately 40% of travelers check work e-mail while on vacation, about 33% take cell phones on vacation in order to stay connected with work, and about 25% bring laptop computers on vacation (*San Luis Obispo Tribune,* December 1, 2005). Are the given percentages population values or were they computed from a sample?

1.4 Based on a study of 2121 children between the ages of 1 and 4, researchers at the Medical College of Wisconsin concluded that there was an association between iron deficiency and the length of time that a child is bottle-fed (*Milwaukee Journal Sentinel,* November 26, 2005). Describe the sample and the population of interest for this study.

1.5 The student senate at a university with 15,000 students is interested in the proportion of students who favor a change in the grading system to allow for plus and minus grades (e.g., B+, B, B−, rather than just B). Two hundred students are interviewed to determine their attitude toward this proposed change. What is the population of interest? What group of students constitutes the sample in this problem?

1.6 The increasing popularity of online shopping has many consumers using Internet access at work to browse and shop online. In fact, the Monday after Thanksgiving has been nicknamed "Cyber Monday" because of the large increase in online purchases that occurs on that day. Data from a large-scale survey by a market research firm (*Detroit Free Press,* November 26, 2005) was used to compute estimates of the percent of men and women who shop online while at work. The resulting estimates probably won't make most employers happy—42% of the men and 32% of the women in the sample were shopping online at work! Are the estimates given computed using data from a sample or for the entire population?