

Flaw 1: Misleading Graph Figure 1-1 is deceptive. Using a vertical scale that does not start at zero exaggerates the difference between the two numbers of responses. Thus Figure 1-1 makes it appear that the “yes” responses are about three times the number of “no” responses, but examination of the actual response counts shows that the “yes” responses are really about 1.7 times the “no” responses. Deceptive graphs are discussed in more detail in Section 2-4.

Flaw 2: Bad Sampling Method The survey responses are from a recent America OnLine survey of Internet users. The survey question was posted on the America OnLine Web site and Internet users decided whether to respond. This is an example of a *voluntary response sample*—a sample in which respondents decide themselves whether to participate. With a voluntary response sample, it often happens that those with a strong interest in the topic are more likely to participate, so the results are very questionable. The large number of respondents does not overcome this flaw of having a voluntary response sample. When we want to use sample data to learn something about a population, it is *extremely* important to obtain sample data that are representative of the population from which the data are drawn. As we proceed through this chapter and discuss types of data and sampling methods, we should focus on these key concepts:

- **Sample data must be collected in an appropriate way, such as through a process of random selection.**
- **If sample data are not collected in an appropriate way, the data may be so completely useless that no amount of statistical torturing can salvage them.**

It would be easy to accept the preceding survey results and blindly proceed with calculations and statistical analyses, but if we did so, we would miss the critical two flaws described above. We might then develop conclusions that are fundamentally wrong and misleading. Instead, we should develop skills in statistical thinking and critical thinking so we can understand why the survey is so seriously flawed and why we should not rely on it to yield any valid information.

1-1 Review and Preview

The first section of each of Chapter 1 through Chapter 14 begins with a brief review of what preceded the chapter, and a preview of what the chapter includes. This first chapter isn’t preceded by much of anything except the Preface, and we won’t review that (most people don’t even read it). However, we can review and formally define some statistical terms that are commonly used. The Chapter Problem discussed an America OnLine poll that collected sample data. Polls collect data from a small part of a larger group so that we can learn something about the larger group. This is a common and important goal of statistics: Learn about a large group by examining sample data from some of its members. In this context, the terms *sample* and *population* have special meanings. Formal definitions for these and other basic terms are given here.