between cases in which we have data for a sample, and cases in which we have data for an entire population.

We also need to know the difference between the terms *quantitative data* and *categorical data*. Some numbers, such as those on the shirts of basketball players, are not quantities because they don't measure or count anything, and it would not make sense to perform calculations with such numbers. In this section we describe different types of data. The type of data is one of the key factors that determine the statistical methods we use in our analysis.

Parameter/Statistic

DEFINITIONS

A **parameter** is a numerical measurement describing some characteristic of a *population*.

A **statistic** is a numerical measurement describing some characteristic of a *sample*.

HINT The alliteration in "population parameter" and "sample statistic" helps us remember the meaning of these terms.

Using the foregoing definitions and those given in Section 1-1, we see that the term *statistics* has two possible meanings:

1. Statistics are two or more numerical measurements describing characteristics of samples.
2. Statistics is the science of planning studies and experiments; obtaining data; organizing, summarizing, presenting, analyzing, and interpreting those data; and then drawing conclusions based on them.

We can determine which of these two definitions applies by considering the context in which the term *statistics* is used. The following example uses the first meaning of *statistics* as given above.

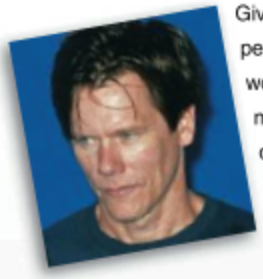
Example 1 Parameter/Statistic

In a Harris Poll, 2320 adults in the United States were surveyed about body piercings, and 5% of the respondents said that they had a body piercing, but not on the face. Based on the latest available data at the time of this writing, there are 241,472,385 adults in the United States. The results from the survey are a sample drawn from the population of all adults.

1. **Parameter:** The population size of 241,472,385 is a *parameter*, because it is based on the entire population of all adults in the United States.
2. **Statistic:** The sample size of 2320 surveyed adults is a *statistic*, because it is based on a sample, not the entire population of all adults in the United States. The value of 5% is another statistic, because it is also based on the sample, not on the entire population.

Six Degrees of Separation

Social psychologists, historians, political scientists, and communications specialists are interested in "The Small World Problem":



Given any two people in the world, how many intermediate links are necessary to connect the two original people? In the 1950s and 1960s, social psychologist Stanley Milgram conducted an experiment in which subjects tried to contact other target people by mailing an information folder to an acquaintance who they thought would be closer to the target. Among 160 such chains that were initiated, only 44 were completed, so the failure rate was 73%. Among the successes, the number of intermediate acquaintances varied from 2 to 10, with a median of 6 (hence "six degrees of separation"). The experiment has been criticized for its high failure rate and its disproportionate inclusion of subjects with above-average incomes. A more recent study conducted by Microsoft researcher Eric Horvitz and Stanford Assistant Professor Jure Leskovec involved 30 billion instant messages and 240 million people. This study found that for instant messages that used Microsoft, the mean length of a path between two individuals is 6.6, suggesting "seven degrees of separation." Work continues in this important and interesting field.