CONTENT OVERVIEW

A **census** is an attempt to gather information about every member of some group, called the **population**. This unit introduces the U.S. Census and its problems in collecting data on the entire U.S. population. One of the most serious problems is undercounting certain segments of the population. Unfortunately not all groups of people are undercounted at the same rates. For example, undercount rates for minority groups are higher than for whites and undercounted rates for renters are higher than for homeowners. Moreover, undercount rates for those living in poverty are higher than for the affluent. The U.S. government uses sampling to estimate undercount rates for various groups. However, it never changes the official headcount number based on the results from sampling.

A **sample** allows the researcher to gather information from only a part of the population. Sampling – collecting data from a portion of the population – is the general means of gathering information about a population when it is not possible to get information from each individual in the population. Sampling saves both time and money. In some cases, such as for Frito-Lay potato chips, both the whole potatoes in a sample and the chips in a sample are destroyed as part of the data collection process. In such cases a census would be out of the question or there would be no product left to sell.

In order for a sample to provide good information about a population, the sample needs to be representative of the population. A **simple random sample**, a sample in which each member of the population is equally likely to wind up in the sample, is one means of ensuring that the sample is representative of the population and not biased. A simple random sample can be selected from the population in the same way that a subgroup is randomly selected from a larger group to receive a certain treatment. Hence, you should refer to Unit 15, Designing Experiments, for directions on selecting a random sample.

Sampling bias occurs when a sample is collected in such a way that some members of the population are less likely to be included than others. A voluntary television poll is an example of a biased sample. Since it is voluntary, only those with strong views are likely to call or text in to vote. Furthermore, only those watching the particular station at the time the poll is given will participate. In this case, the entire segment of the population who do not watch that particular station will be left out of the sample.