Figure 8.3. PLANNED DATA COLLECTING: Experiments and Sample Surveys

Program 13 in: *Against All Odds: Inside Statistics*

This program has two main purposes. First, we look at some more complex designs for experiments, especially the use of blocking. *Blocks* are groups of experimental units or subjects that are similar in some way important to the response. In a randomized block design, randomization is carried out separately within each block. Forming blocks is another type of control in an experiment. The video shows an agricultural experiment in which the experimental units are small plots of land and the treatments are different *varieties* of strawberries. Because land in different locations varies in fertility, drainage and other ways that affect plant growth, it is helpful to divide the land into blocks containing a few plots in the same location. The different varieties are then assigned at random to plots separately within each block. In addition to blocking, the program also mentions experiments with more than one factor. In such experiments, a treatment combines a particular value of each factor.

The second purpose of this program is to introduce the design of sample surveys. In a sample survey, a *sample* is selected from the *population* of all people or things about which we desire information. Conclusions about the population are based on examination of the sample. The video contrasts a sample with a *census* which attempts to examine *every* member of the population. Even a census is usually not completely accurate, as a discussion of the U.S. population census shows. Some people are missed by the census, and this undercount is concentrated in the minority neighbourhoods of large cities. Only a census can give detailed information about every part of a population, but for overall information a sample is usually faster, less expensive, and just as accurate. The varied uses of sampling are shown by a look at the manufacture of potato chips: sampling is used at every stage, from the arrival of potatoes at the plant to checking the final product on supermarket shelves.

The *design* of a sample refers to the method used to select the sample from the population. The most important statistical principle in sampling design is to select the sample *at random* from the population. The basic probability sample is a *simple random sample*, which gives every possible sample of a given size the same chance to be the sample selected. Simple random samples are chosen by labelling the members of the population and using random digits to select the sample. Random sampling avoids the systematic favoritism or *bias* that often results when a sample is formed by human choice.

Copyright © 1985 by Consortium for Mathematics and Its Applications (COMAP), Inc.
Reprinted with permission of W. H. Freeman and Company.

1 Comment critically on the last two sentences of the video summary given above.

- *Simple random samples are chosen by labelling the members of the population and using random digits to select the sample.*
- *Random sampling avoids the systematic favoritism or* ***bias*** *that often results when a sample is selected by human choice.*

Blank page