# Section 1.3

**Definition**: In a **census**, measurements or observations from the *entire* population are used.

**Example**: The night classes example was not a census.

**Remark**: A census is impractical for a large population.

**Definition**: In a **sample**, measurements or observations from *part* of the population are used.

**Definition**: In an **observational study**, observations and measurements of individuals are conducted in a way that doesn't change the response or the variable being measured.

**Definition**: In an **experiment**, a *treatment* is deliberately imposed on the individuals in order to observe a possible change in the response or variable being measured.

**Exercise 1.** A four-acre plot of land is used to plant corn. A scientist wishes to test the effect of a new fertilizer on corn yield. Unsure how much fertilizer to apply, the scientist divides the four acres into four homogeneous one-acre plots. Different amounts of fertilizer are used in each of the four plots. Is this an experiment or an observational study?

This is an _____ because we are applying fertilizer to see how it will effect the corn yield.

If it is an experiment, what are the experimental units, response variable, and explanatory variable?

**Exercise 2.** It is desired to ascertain whether studying a foreign language in high school increases mathematical ability. From grade transcripts of high school seniors, we form two groups: those who studied a foreign language and those who did not. A test of mathematical ability is administered to both groups and the results are compared. On average, the students who studied a foreign language performed considerably better than those who did not. Is this an experiment or an observational study?

This is an _____ The test was not administered to change the students' mathematical ability. It was given to discover it.

**Exercise 3.** Classify each of the following as an experiment or an observational study:
(a) Over a 4-month period, among 30 people with bipolar disorder, patients who were given a high dose (10 g/day) of omega-3 fats from fish oil improved more than those given a placebo.

(b) An examination of the medical records of more than 360,000 Swedish men showed that those who were overweight or who had high blood pressure had a higher risk of kidney cancer.

(c) Some gardeners prefer to use nonchemical methods to control insect pests in their gardens. Researchers have designed two kinds of traps, and want to know which design will be more effective. They randomly chose 10 locations in a large garden and place one of each kind of trap at

each location. After a week, they count the number of bugs in each trap.

(d) The May 4, 2000, issue of *Science News* reported that, contrary to popular belief, depressed individuals cry no more often in response to sad situations than nondepressed people. Researchers studied 23 men and 48 women with major depression, and 9 men and 24 women with no depression. They showed the subjects a sad film about a boy whose father has died, noting whether or not the subjects cried. Women cried more often than men, but there were no significant differences between the depressed and the nondepressed groups.

**Class Exercise 1.** Classify each of the following as an experiment or an observational study. If this is an experiment, identify the experimental units.

(a) In a test of roughly 200 men and women, those with moderately high blood pressure (averaging 164/89 mm Hg) did worse on tests of memory and reaction time than those with normal blood pressure.

(b) After menopause many women take supplemental estrogen. There is some concern that if these women also drink alcohol, their estrogen levels will rise too high. Twelve volunteers who were receiving supplemental estrogen were randomly divided into two groups, as were 12 other volunteers not on estrogen. In each case, one group drank an alcoholic beverage, the other a non-alcoholic beverage. An hour later everyone's estrogen level was checked. Only those on supplemental estrogen who drank alcohol showed a marked increase.

(c) In 2002 the journal *Science* reported that a study of women in Finland indicated that having sons shortened the lifespans of mothers by about 34 weeks per son, but that daughters, helped to lengthen the mothers' lives. The data came from church records from the period 1640 to 1870.

(d) Scientists examined the glycogen content of rats' brains at the rats' normal bedtimes and after they had been kept awake for an extra 6, 12, or 24 hours. The scientists found that glycogen was 38% lower among rats that had been sleep-deprived for 12 hours or more, and that the levels recovered during subsequent sleep. These researchers speculate that we may need to sleep in order to restore the brain's energy fuel.

**Exercise 4.** Explain whether an experiment or an observational study would be more appropriate to investigate the following: Whether or not smoking has an effect on ordinary heart disease.

**Class Exercise 2.** Explain whether an experiment or an observational study would be more appropriate to investigate the following:
(a) Whether or not higher SAT scores tend to be positively associated with higher college GPAs.
(b) Whether or not a special coupon attached to the outside of a catalog makes recipients more likely to order products from a mail-order company.

**Class Exercise 3.** A study is to be designed to examine the GPAs of students who take marijuana regularly and those who don't. Which is more appropriate, an observational study or an experiment? Why?

**Exercise 5.** Jim believes that M&M's improves performance on mathematics exams. In order to prove this theory, he gives fifteen of his 30 classmates M&M's and compares the test scores of the M&M's students with the test scores of the non M&M's students. It turns out that the students who received M&M's did do better on their exams. Based on the exam scores, Jim concludes that his theory on M&M's is correct.

Are there any reasons to be suspicious of Jim's conclusion?

First, it is *possible* that Jim sought out the students

Even if Jim didn't                                                                                              , it is possible (by luck) that the

For each of the above possibilities, the better math exam scores would be largely due to the
                                         of the students.

**Definition**: **Randomization** is used to assign individuals to the two treatment groups. This helps prevent bias in selecting members for each group.

**Remark**: Randomization is an important component to any experiment. There is a higher chance of systematic bias without randomizing.
**Remark**: Sometimes, the person conducting an experiment has a hidden agenda. Before looking at the results of an experiment, it is important to see how it was done.

**Exercise 6.** Scientists are interested in seeing which of two weight loss drugs, Drug A and Drug B, is more effective in helping people lose weight. Drug A was given to ten members of Congress and Drug B was given to 10 Hollywood actresses. The drugs were given on October 2 and the weight loss was determined on October 10. The Hollywood actresses (on average) lost more weight. Does this mean that Drug B is more effective than Drug A?

    this                                    that Drug B is more effective than Drug A.

Hollywood actresses are always trying to                         It's very possible that their weight loss has

The sample size is 10 for each group. Generally speaking, this is
                         about the conclusion. The larger the sample size,

**Exercise 7.** Going back to a previous exercise, let's say we have a room full of Hollywood actresses and Congressmen. Our goal again is to see the effect of the weight loss drugs. What is the best strategy to see which drug is more effective? How would you use randomization?

First,                 the actresses from the Congressmen. Next,                 the actresses Drug A and         the actresses Drug B. Then,                 the actresses and then use the Random Number Table to decide which actresses receive Drug A and which actresses receive Drug B. Do the same thing for the Congressmen. Finally,                 the weight loss of the people who received Drug A with the weight loss of the people who received Drug B.

 **Remark**: In practice, scientists use computers (instead of calculators) to make random assignments.

**Remark**: **Replication** is another important component to an experiment. After randomizing, we could still (by chance) have an unevenness between the two groups. The larger the groups, the less likely that there will be a significant difference between the two groups. Suppose that there are 200 Congressmen. It is unlikely that the 100 Congressman who received Drug A would be very different than the 100 Congressmen who received Drug B (on average). On the other hand, suppose there are 4 Congressmen and we give 2 Congressmen Drug A and we give 2 Congressmen

Drug B. It is very possible that there could be a big discrepancy (on average) between the two groups.

**Exercise 8.** Susie is feeling very sick. Her bff Katie says she has a "magic medicine" that will cure her of her illness. Susie takes it and claims she feels better. It turns out that the "magic medicine" was a Tic-Tac. Do Tic-Tac's have a special medical benefit?

**Definition**: A "fake" treatment that looks just like the treatments being tested is called a **placebo**.

**Definition**: The **placebo effect** occurs when a subject reacts favorably to a placebo when in fact, he or she has been given no medicated treatment at all.

**Exercise 9.** "Gastric freezing" is a clever treatment for ulcers in the upper intestine. The patient swallows a deflated balloon with tubes attached, then a refrigerated liquid is pumped through the balloon for an hour. The idea is that cooling the stomach will reduce its production of acid and so relieve ulcers. How would you conduct an experiment to see if "gastric freezing" actually works?

Because the placebo effect can be so strong, it would be foolish to tell subjects in a medical experiment whether they are receiving a new drug or a placebo. Knowing that they are getting "just a placebo" might weaken the placebo effect and bias the experiment in favor of the other treatments. It is also foolish to tell doctors and other medical personnel what treatment each subject is receiving. If they know that a subject is getting "just a placebo," they may expect less than if they know the subject is receiving a promising experimental drug. Doctors' expectations change how they interact with patients and even the way they diagnose a patient's condition. Whenever possible, experiments with human subjects should be *double blind*.

**Definition**: In a **double-blind experiment**, neither the subjects nor the people who work with them know which treatment each subject is receiving.

The last few exercises and remarks illustrate the three major principals of experimental design: randomization, replication, and control.

**Exercise 10.** It is known that the progress of leukemia differs in women and men. A scientist has 48 men and 52 women at her disposable and she wants to test a new drug aimed at curing leukemia. She not only wants to see how the drug works, but how effective the drug is for each gender. How should she conduct the experiment?

**Definition**: A **block** is a group of experimental units or subjects that are known before the experiment to be similar in some way that is expected to affect the response to the treatments.

**Example**: In the previous exercise, the men served as one block and the women served as the other.

**Definition**: An **explanatory variable** explains or causes changes in the response variables.

**Definition**: A **response variable** measures an outcome of a study.

**Class Exercise 4.** In order to determine the effect of a new fertilizer on the growth of tomato plants, researchers applied the fertilizer to tomato plants on the west side of a garden, but did not fertilize the same variety of tomato plants on the east side of the garden. After a specified time period, they measured the weights of the tomatoes and found that the fertilized plants grew larger tomatoes than did the non-fertilized plants. They concluded that the fertilizer works.

(a) Was this an experiment or an observational study? Why?

(b) Describe the explanatory and response variables.

(c) The study is badly flawed. Explain the flaw and how it affects the validity of the researcher's conclusion.

**Class Exercise 5.** It's a common belief that people behave strangely when there's a full moon, and that as a result police and emergency rooms are busier than usual. Design a way you could find out if there is any merit to this belief. Will you use an observational study or an experiment?

**Class Exercise 6.** A medical researcher suspects that giving post-surgical patients large doses of vitamin E will speed their recovery times by helping their incisions heal more quickly. Design an experiment to test this conjecture.

## Sample Survey Issues

For sample surveys, data is obtained through asking people questions. A number of issues can arise if this is how we obtained the data. The next few exercises illustrate some of these issues.

**Exercise 11.** Suppose you ask 100 people the following two questions:

(1) Should teachers be paid more?

(2) Do you favor paying hard-working teachers a little more so that our fine young people can have a decent education?

Which question do you think more people would respond "yes" to?

because it's hard to say no to that question.

**Remark**: The wording of a question can have a huge impact on the response. In the previous exercise, the essence of the questions is the same but the proportion of people answering "yes" will be different.

**Exercise 12.** Let's suppose we wanted to determine the percentage of high school students who smoke marijuana. In order to make an estimate for all high school students, we take a simple random sample of 100 students and then ask them if they have smoked marijuana. What is wrong with this method?

Students are _____ that they have smoked marijuana. If you make an estimate based on the sample, you will likely _____ the proportion of all high school students who smoke marijuana.

**Remark**: Surveys where the respondents are unlikely to be honest are usually given anonymously.

**Class Exercise 7.** Suppose you ask 100 people the following two questions:
(1) Would you support cutting the programs involving aid to the needy?
(2) Would you support cutting public welfare programs?
Which question do you think more people would respond "yes" to? Why?

The following passage was taken from "Exploring Surveys and Information from Samples" by Landwehr, Swift, and Watkins.

People give untruthful answers for several reasons. People, not wanting to appear ignorant, will try to answer a question even if they know nothing about the subject. "In one study, educators were asked how they would rank Princeton's undergraduate business program. In every case, it was rated among the top 10 departments in the country, even though Princeton does not offer an undergraduate business major" (*Los Angeles Times*, November 21, 1982).

People often do not remember numbers they are asked about. For example, one study (*Sociological Methods and Research*, November 1981) asked students to report their grade point averages (GPAs). Researchers then determined the actual GPAs. Over 17% of the students reported a GPA 0.4 or more above their actual average, but about 2% of the students reported a GPA more than 0.4 below their actual GPA!

**Class Exercise 8.** A recent study of deaths in bar fights showed that in 90% of the cases, the person who died started the fight. Do you believe this? Why do you think the number is so high?

**Definition**: A **lurking variable** is one for which no data have been collected but that nevertheless has influence on other variables in the study.

**Definition**: Two variables are **confounded** when the effects of one cannot be distinguished from the effects of the other. Confounding variables may be part of the study, or they may be outside lurking variables.

## Student Feedback

My teaching methods are (I hope) continually subject to improvement. If you have any comments, suggestions, or ideas, please email them to me at Sithparran.Vanniasegaram@evc.edu .

## Homework
### C problems

Section 1.3: 7, 9

### B problems

Section 1.3: 1, 5

### A problems

Section 1.3: 3, 11