AP STATISTICS

2019-2020

Summer Math Assignment

Read "Sampling & Surveys" (Ch. 4.1 that is copied). Then, try the examples and the "check your understanding" problems in the chapter.

Work the "Sampling & Surveys" exercises that are attached and bring your answers with you to class on the first day of school. We will review this section and have a quiz during the first week of school. (Answers keys for all homework exercises will be posted on the HHS website.)

Have a good summer!

Mrs. Warren

Mrs. Catherine Warren
HHS AP Statistics

Designing Studies

Can Magnets Help Reduce Pain?

Early research showed that magnetic fields affected living tissue in humans. Some doctors have begun to use magnets to treat patients with chronic pain. Scientists wondered whether this type of therapy really worked. They designed a study to find out.

Fifty patients with chronic pain were recruited for the study. A doctor identified a painful site on each patient and asked him or her to rate the pain on a scale from 0 (mild pain) to 10 (severe pain). Then, the doctor selected a sealed envelope containing a magnet from a box that contained both active and inactive magnets. That way, neither the doctor nor the patient knew which type of magnet was being used. The chosen magnet was applied to the site of the pain for 45 minutes. After "treatment," each patient was again asked to rate the level of pain from 0 to 10.

In all, 29 patients were given active magnets and 21 patients received inactive magnets. All but one of the patients rated their initial pain as an 8, 9, or 10. So scientists decided to focus on patients' final pain ratings. Here they are, grouped by the type of magnet used:[1]

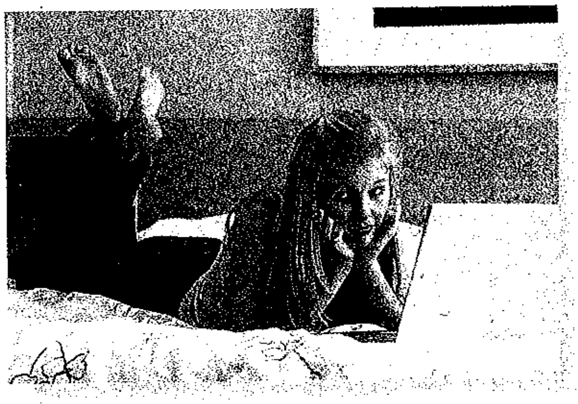
Active:	0, 4, 7, 0, 4, 2, 5, 5, 3, 2, 2, 2, 3, 4, 6, 4, 3, 0, 2, 0, 4, 4, 5, 9, 10, 10, 10, 10, 7
Inactive:	4, 7, 5, 8, 8, 6, 8, 10, 10, 6, 10, 8, 10, 10, 10, 10, 9, 9, 10, 10, 9

What do the data tell us about whether the active magnets helped reduce pain? By the end of the chapter, you'll be ready to interpret the results of this study.

Introduction

You can hardly go a day without hearing the results of a statistical study. Here are some examples:

- The National Highway Traffic Safety Administration (NHTSA) reports that seat belt use in passenger vehicles increased from 83% in 2008 to 84% in 2009.[2]



- According to a recent survey, U.S. teens aged 13 to 18 spend an average of 26.8 hours per week online. Although 59% of the teens said that posting personal information or photos online is unsafe, 62% said they had posted photos of themselves.[3]
- A recent study suggests that lack of sleep increases the risk of catching a cold.[4]
- For their final project, two AP Statistics students showed that listening to music while studying decreased subjects' performance on a memory task.[5]

Can we trust these results? As you'll learn in this chapter, that depends on how the data were produced. Let's take a closer look at where the data came from in each of these studies.

Each year, the NHTSA conducts an *observational study* of seat belt use in vehicles. The NHTSA sends trained observers to record the actual behavior of people in vehicles at randomly selected locations across the country. The idea of an observational study is simple: you can learn a lot just by watching. Or by asking a few questions, as in the survey of teens' online habits. Harris Interactive conducted this survey using a "representative sample" of 655 U.S. 13- to 18-year-olds. Both of these studies use information from a *sample* to draw conclusions about some larger *population*. Section 4.1 examines the issues involved in sampling and surveys.

In the sleep and catching a cold study, 153 volunteers took part. They answered questions about their sleep habits over a two-week period. Then, researchers gave them a virus and waited to see who developed a cold. This was a complicated observational study. Compare this with the *experiment* performed by the AP Statistics students. They recruited 30 students and divided them into two groups of 15 by drawing names from a hat. Students in one group tried to memorize a list of words while listening to music. Students in the other group tried to memorize the same list of words while sitting in silence. Section 4.2 focuses on designing experiments.

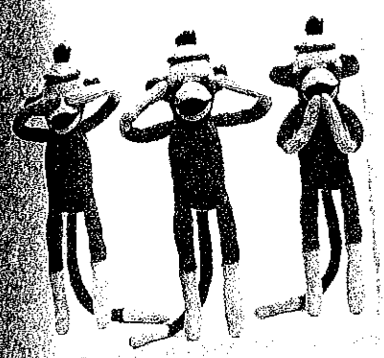
The goal of many statistical studies is to show that changes in one variable *cause* changes in another variable. In Section 4.3, we'll look at why establishing causation is so difficult, especially in observational studies. We'll also consider some of the ethical issues involved in planning and conducting a study.

Here's an Activity that gives you a preview of what lies ahead.

ACTIVITY *See no evil, hear no evil?*

MATERIALS: Two index cards, each with 10 pairs of digits written on it (prepared by your teacher); clock, watch, or stopwatch to measure 30 seconds; and a coin for each pair of students

Confucius said, "I hear and I forget. I see and I remember. I do and I understand." Do people really remember what they see better than what they hear?[6] In this Activity, you will perform an experiment to try to find out.



1. Divide the class into pairs of students by drawing names from a hat.

2. Your teacher will give each pair two index cards with 10 sets of numbers on them. *Do not look at the numbers* until it is time for you to do the experiment.

3. Flip a coin to decide which of you is Student A and which is Student B. Shuffle the index cards and deal one face down to each partner.

4. Student A will attempt a memory task while Student B keeps time.

Directions: Study the pairs of numbers on the index card for 30 seconds. Then turn the card over. Recite the alphabet aloud (A, B, C, etc.). Then tell your partner (Student B) what you think the numbers on the card are. Student B will record how many pairs of numbers you recalled correctly.

5. Now it's Student B's turn to do a memory task while Student A records the data.

Directions: Your partner will read the pairs of numbers on your index card aloud three times slowly. Next, you will recite the alphabet aloud (A, B, C, etc.) and then tell your partner what you think the numbers on the card are. Student A will record how many pairs of numbers you recalled correctly.

6. Your teacher will scale and label axes on the board for parallel dotplots of the results. Plot the number of pairs you remembered correctly on the appropriate graph.

7. Did students in your class remember numbers better when they saw them or when they heard them? Give appropriate evidence to support your answer.

8. Based on the results of this experiment, can we conclude that people in general remember better when they see than when they hear? Why or why not?

4.1

Sampling and Surveys

In Section 4.1, you'll learn about:

- The idea of a sample survey
- How to sample badly
- How to sample well: Random sampling
- Other sampling methods
- Inference for sampling
- Sample surveys: What can go wrong?

Suppose we want to find out what percent of young drivers in the United States text while driving. To answer the question, we will survey 16- to 20-year-olds who live in the United States and drive. Ideally, we would ask them all (take a *census*). But contacting every driver in this age group wouldn't be practical: it would take too much time and cost too much money. Instead, we put the question to a sample chosen to represent the entire **population** of young drivers.

DEFINITION: Population and sample

The **population** in a statistical study is the entire group of individuals about which we want information.

A **sample** is the part of the population from which we actually collect information. We use information from a sample to draw conclusions about the entire population.

The distinction between population and sample is basic to statistics. To make sense of any sample result, you must know what population the sample represents. Here's an example that illustrates this distinction and also introduces some major uses of sampling.

EXAMPLE

Sampling Hardwood and Humans

Populations and samples

PROBLEM: Identify the population and the sample in each of the following settings.

(a) A furniture maker buys hardwood in large batches. The supplier is supposed to dry the wood before shipping (wood that isn't dry won't hold its size and shape). The furniture maker chooses five pieces of wood from each batch and tests their moisture content. If any piece exceeds 12% moisture content, the entire batch is sent back.

(b) Each week, the Gallup Poll questions a sample of about 1500 adult U.S. residents to determine national opinion on a wide variety of issues.

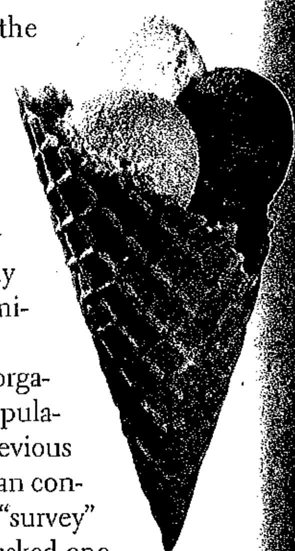
SOLUTION:

(a) The population is all the pieces of hardwood in a batch. The sample is the five pieces of wood that are selected from that batch and tested for moisture content.

(b) Gallup's population is all adult U.S. residents. Their sample is the 1500 adults who actually respond to the survey questions.

For Practice *Try Exercise 1*

The Idea of a Sample Survey



We often draw conclusions about a whole population on the basis of a sample. Have you ever tasted a sample of ice cream and ordered a cone if the sample tastes good? Since ice cream is fairly uniform, the single taste represents the whole. Choosing a representative sample from a large and varied population (like all young U.S. drivers) is not so easy. The first step in planning a **sample survey** is to say exactly *what population* we want to describe. The second step is to say exactly *what we want to measure*, that is, to give exact definitions of our variables.

Sample survey

We reserve the term "sample survey" for studies that use an organized plan to choose a sample that represents some specific population, like the pieces of hardwood and the U.S. adults in the previous example. By our definition, the population in a sample survey can consist of people, animals, or things. Some people use the terms "survey" or "sample survey" to refer only to studies in which people are asked one or more questions, like the Gallup Poll of the last example. We'll avoid this more restrictive terminology.

EXAMPLE

How Does the Current Population Survey Work?

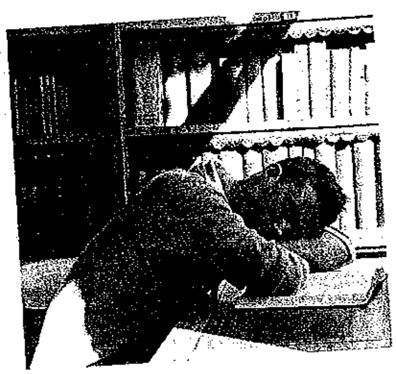
A sample survey

One of the most important government sample surveys in the United States is the monthly Current Population Survey (CPS). The CPS contacts about 60,000 households each month. It produces the monthly unemployment rate and much other economic and social information. To measure unemployment, we must first specify the population we want to describe. The CPS defines its population as all U.S. residents (legal or not) 16 years of age and over who are civilians and are not in an institution such as a prison. The unemployment rate announced in the news refers to this specific population.

What does it mean to be "unemployed"? Someone who is not looking for work—for example, a full-time student—should not be called unemployed just because she is not working for pay. If you are chosen for the CPS sample, the interviewer first asks whether you are available to work and whether you actually looked for work in the past four weeks. If not, you are neither employed nor unemployed—you are not in the labor force.

If you are in the labor force, the interviewer goes on to ask about employment. If you did any work for pay or in your own business during the week of the survey, you are employed. If you worked at least 15 hours in a family business without pay, you are employed. You are also employed if you have a job but didn't work because of vacation, being on strike, or other good reason. An unemployment rate of 9.7% means that 9.7% of the sample was unemployed, using the exact CPS definitions of both "labor force" and "unemployed."

The final step in planning a sample survey is to decide how to choose a sample from the population. Let's take a closer look at some good and not-so-good sampling methods.



How to Sample Badly

How can we choose a sample that we can trust to represent the population? The easiest—but not the best—sampling method just chooses individuals who are close by. Suppose we're interested in finding out how long students at a large high school spent on homework last week, for example. We might go to the school library and ask the first 30 students we see about their homework time. A sample selected by taking the members of the population that are easiest to reach is called a **convenience sample.** Convenience samples often produce unrepresentative data.

DEFINITION: Convenience sample

Choosing individuals who are easiest to reach results in a **convenience sample.**

What's wrong with our convenience sample of students in the library? It's unlikely that this sample represents the homework habits of all students at the school well. In fact, we'd expect the sample to overestimate the average homework time in the population since students who hang out in the library might tend to be more studious. This is **bias:** using a method that will consistently overestimate or underestimate the value you want to know.

DEFINITION: Bias

The design of a statistical study shows **bias** if it systematically favors certain outcomes.

AP EXAM TIP If you're asked to describe how the design of a study leads to bias, you're expected to identify the *direction* of the bias. Suppose you were asked, "Explain how using a convenience sample of students in your statistics class to estimate the proportion of all high school students who own a graphing calculator could result in bias." You might respond, "This sample would probably include a much higher proportion of students with a graphing calculator than in the population at large because a graphing calculator is required for the statistics class. That is, this method would probably lead to an overestimate of the actual population proportion."

Convenience samples are almost guaranteed to show bias. So are **voluntary response samples,** in which people decide whether to join the sample in response to an open invitation. Call-in, write-in, and many Internet polls rely on voluntary response samples. Unfortunately, people who choose to participate are usually not representative of any clearly defined population. Voluntary response samples attract people who feel strongly about the issue in question, often in the same direction. That leads to bias.

The Internet brings voluntary response samples to the computer nearest you. Visit www.misterpoll.com to become part of the sample in any of dozens of online polls. As the site says, "None of these polls are 'scientific,' but do represent the collective opinion of everyone who participates." Unfortunately, such polls don't tell you anything about the views of the population at large.

DEFINITION: Voluntary response sample

A **voluntary response sample** consists of people who choose themselves by responding to a general appeal. Voluntary response samples show bias because people with strong opinions (often in the same direction) are most likely to respond.

Write-in and call-in opinion polls are almost sure to lead to strong bias. In fact, only about 15% of the public has ever responded to a call-in poll, and these tend to be the same people who call radio talk shows. That's not a representative sample of the population as a whole.

EXAMPLE

Illegal Immigrants and Driver's Licenses

Online polls

www

Former CNN commentator Lou Dobbs doesn't like illegal immigration. One of his shows was largely devoted to attacking a proposal to offer driver's licenses to illegal immigrants. During the show, Mr. Dobbs invited his viewers to go to loudobbs.com to vote on the question "Would you be more or less likely to vote for a presidential candidate who supports giving drivers' licenses to illegal aliens? The result: 97% of the 7350 people who voted by the end of the show said "Less likely."

PROBLEM: What type of sample did Mr. Dobbs use in his poll? Explain how this sampling method could lead to bias in the poll results.

SOLUTION: Mr. Dobbs used a voluntary response sample: people chose to go online and respond. Those who voted were viewers of Mr. Dobbs's program, which means that they are likely to support his views. The 97% poll result is probably an extreme overestimate of the percent of people in the population who would be less likely to support a presidential candidate with this position.

For Practice *Try Exercise 5*

CHECK YOUR UNDERSTANDING

For each of the following situations, identify the sampling method used. Then explain how the sampling method could lead to bias.

1. A farmer brings a juice company several crates of oranges each week. A company inspector looks at 10 oranges from the top of each crate before deciding whether to buy all the oranges.

2. The ABC program *Nightline* once asked whether the United Nations should continue to have its headquarters in the United States. Viewers were invited to call one telephone number to respond "Yes" and another for "No." There was a charge for calling either number. More than 186,000 callers responded, and 67% said "No."

How to Sample Well: Random Sampling

Random sampling

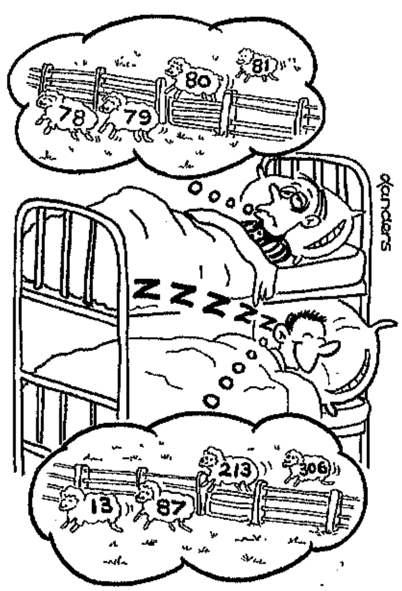
In a voluntary response sample, people choose whether to respond. In a convenience sample, the interviewer makes the choice. In both cases, personal choice produces bias. The statistician's remedy is to allow impersonal chance to choose the sample. A sample chosen by chance rules out both favoritism by the sampler and self-selection by respondents. **Random sampling,** the use of chance to select a sample, is the central principle of statistical sampling.

The simplest way to use chance to select a sample is to place names in a hat (the population) and draw out a handful (the sample). This is the idea of a **simple random sample.**

In everyday life, some people use the word "random" to mean haphazard, as in, "That's so random." In statistics, random means "due to chance." Don't say that a sample was chosen at random if a chance process wasn't used to select the individuals.

DEFINITION: Simple random sample

A **simple random sample (SRS)** of size *n* consists of *n* individuals from the population chosen in such a way that every set of *n* individuals has an equal chance to be the sample actually selected.



Statisticians fall asleep faster by taking a random sample of sheep.

An SRS not only gives each individual an equal chance to be chosen but also gives every possible sample an equal chance to be chosen. There are other random sampling methods that give each individual, but not each sample, an equal chance. Exercise 26 describes one such method. For now, let's think about how to actually select an SRS.

When you think of an SRS, picture drawing names from a hat to remind yourself that an SRS doesn't favor any part of the population. That's why an SRS is a better method of choosing samples than convenience or voluntary response samples. But writing names on slips of paper and drawing them from a hat doesn't work as well if the population is large. Think about the Current Population Survey, for example, which must draw a sample of size 60,000 every month from the population of over 117 million U.S. households.

In practice, people use random numbers generated by a computer or calculator to choose samples. If you don't have technology handy, you can use a **table of random digits.**

DEFINITION: Table of random digits

A **table of random digits** is a long string of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 with these properties:

- Each entry in the table is equally likely to be any of the 10 digits 0 through 9.
- The entries are independent of each other. That is, knowledge of one part of the table gives no information about any other part.

Table D at the back of the book is a table of random digits. It begins with the digits 19223950340575628713. To make the table easier to read, the digits appear in groups of five and in numbered rows. The groups and rows have no meaning—the table is just a long list of randomly chosen digits. There are two steps in using the table to choose a simple random sample.

How to Choose an SRS Using Table D

STEP 1: LABEL. Give each member of the population a numerical label of the *same length.*

STEP 2: TABLE. Read consecutive groups of digits of the appropriate length from Table D.

Your sample contains the individuals whose labels you find.

Always use the shortest labels that will cover your population. For instance, you can label up to 100 individuals with two digits: 01, 02, . . . , 99, 00. As standard practice, we recommend that you begin with label 1 (or 01 or 001, as needed). Reading groups of digits from the table gives all individuals the same chance to be chosen because all labels of the same length have the same chance to be found in the table. For example, any pair of digits in the table is equally likely to be any of the 100 possible labels 01, 02, . . . , 99, 00. Ignore any group of digits that was not used as a label or that duplicates a label already in the sample. You can read digits from Table D in any order—across a row, down a column, and so on—because the table has no order. We recommend reading across rows from left to right.

EXAMPLE

Spring Break!

Choosing an SRS with Table D

The school newspaper is planning an article on family-friendly places to stay over spring break at a nearby beach town. The editors intend to call 4 randomly chosen hotels to ask about their amenities for families with children. They have an alphabetized list of all 28 hotels in the town.

PROBLEM: Use Table D at line 130 to choose an SRS of 4 hotels for the editors to call.

SOLUTION: We'll use the two-step process for selecting an SRS using Table D.

Step 1: Label. Two digits are needed to label the 28 resorts. We have added labels 01 to 28 to the alphabetized list of resorts below.

01 Aloha Kai	08 Captiva	15 Palm Tree	22 Sea Shell
02 Anchor Down	09 Casa del Mar	16 Radisson	23 Silver Beach
03 Banana Bay	10 Coconuts	17 Ramada	24 Sunset Beach
04 Banyan Tree	11 Diplomat	18 Sandpiper	25 Tradewinds
05 Beach Castle	12 Holiday Inn	19 Sea Castle	26 Tropical Breeze
06 Best Western	13 Lime Tree	20 Sea Club	27 Tropical Shores
07 Cabana	14 Outrigger	21 Sea Grape	28 Veranda

Step 2: Table. To use Table D, start at the left-hand side of line 130 and read two-digit groups. Skip any groups that aren't between 01 and 28, as well as any repeated groups. Continue until you have chosen four resorts. Here is the beginning of line 130:

69051 64817 87174 09517 84534 06489 87201 97245

The first 10 two-digit groups are

69	05	16	48	17	87	17	40	95	17
Skip	✓	✓	Skip	✓	Skip	Skip	Skip	Skip	Skip
Too big			Too big		Too big	Repeat	Too big	Too big	Repeat

We skip 5 of these 10 groups because they are too high (over 28) and 2 because they are repeats (both 17s). The hotels labeled 05, 16, and 17 go into the sample. We need one more hotel to complete the sample. The remaining 10 two-digit groups in line 130 are

84	53	40	64	89	87	20	19	72	45
Skip	Skip	Skip	Skip	Skip	Skip	✓			
Too big →									

Our SRS of 4 hotels for the editors to contact is: 05 Beach Castle, 16 Radisson, 17 Ramada, and 20 Sea Club.

For Practice *Try Exercise 11*

As you saw in the previous example, using Table D to select an SRS can be time-consuming. The *Simple Random Sample* applet can help you quickly choose an SRS for populations of up to 500 individuals. Your calculator can do even better.

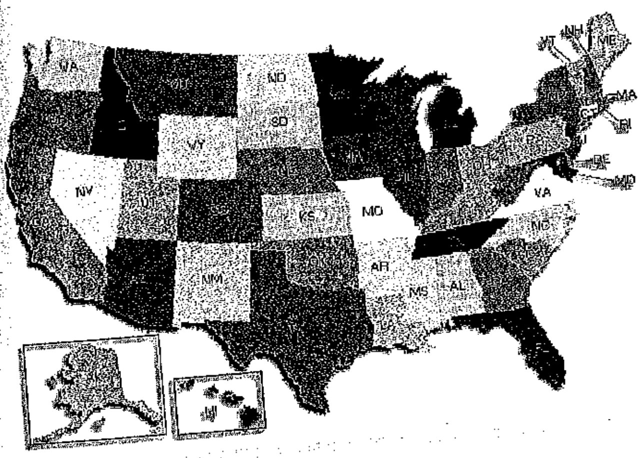
To see how software speeds up choosing an SRS, try the random number generator at www.random.org. The Research Randomizer at www.randomizer.org is another option. Click on "Randomize" and fill in the boxes. You can even ask the Randomizer to arrange your sample in order.

There's an important difference between the samples produced by the *Simple Random Sample* applet and the calculator's `randInt` command in the Technology Corner. The applet sampled *without replacement* from the "Population hopper," but the calculator sampled *with replacement* from the specified population. As a result, the calculator sometimes selects the same number more than once in a given sample. To deal with this problem, you can generate additional random numbers as needed to replace any repeats. Or you can use a method other than `randInt` (like `randSamp` on the TI-Nspire) to sample without replacement. Refer to your device's reference manual or ask your teacher.

We can trust results from an SRS, as well as from other types of random samples that we will meet later, because the use of impersonal chance avoids bias. Online and call-in polls also produce samples. But we can't trust results from these samples because they are chosen in ways that invite bias. *The first question to ask about any sample is whether it was chosen at random.*

ACTIVITY *How large is a typical U.S. state?*

In this Activity, you and your classmates will use two different sampling methods to estimate the average area (in square miles) of the 50 states.



1. Use the map shown to choose a sample of 5 states to estimate the average (mean) land area. You have 15 seconds.
2. Refer to the table of land areas on page N/DS-5. Find the mean land area for your sample.
3. Make a class dotplot of the mean land areas from Step 2.
4. Use a line of Table D assigned by your teacher to choose an SRS of 5 states from the table on page N/DS-5. Find the mean land area for this sample.
5. Make a class dotplot of the mean land areas from Step 4 right above your dotplot from Step 3.
6. How do the class's estimates using the two methods compare? What advantage(s) does random sampling provide?

Other Sampling Methods

The basic idea of sampling is straightforward: take an SRS from the population and use your sample results to gain information about the population. Unfortunately, it's usually very difficult to actually get an SRS from the population of interest. Imagine trying to get a simple random sample of all the batteries produced in one day at a factory. Or an SRS of all U.S. high school students. In either case, it's just not practical to choose an SRS. For starters, it would be difficult to obtain an accurate list of the population from which to draw the sample. It would also be very time-consuming to collect data from each individual that's randomly selected. Sometimes, there are also statistical advantages to using more complex sampling methods.

One of the most common alternatives to an SRS involves sampling important groups (called **strata**) within the population separately. Then these separate "subsamples" are combined to form one **stratified random sample**. This method works best when the individuals within each stratum are similar to one another in a way that affects the variable being measured but there are large differences between strata.

DEFINITION: Stratified random sample and strata

To select a **stratified random sample,** first classify the population into groups of similar individuals, called **strata.** Then choose a separate SRS in each stratum and combine these SRSs to form the full sample.

Choose the strata based on facts known before the sample is taken. For example, in a study of sleep habits on school nights, the population of students in a large high school might be divided into freshman, sophomore, junior, and senior strata. In a preelection poll, a population of election districts might be divided into urban, suburban, and rural strata. *If the individuals in each stratum are less varied than the population as a whole, a stratified random sample can produce better information about the population than an SRS of the same size.* Not convinced? Try the following Activity.

Unfortunately, "stratified random sample" has the same initials as "simple random sample." Only a simple random sample gets abbreviated SRS, however.

ACTIVITY *Sampling sunflowers*

A British farmer grows sunflowers for making sunflower oil. Her field is arranged in a grid pattern, with 10 rows and 10 columns as shown in the figure. Irrigation ditches run along the top and bottom of the field, as shown. The farmer would like to estimate the number of healthy plants in the field so she can project how much money she'll make from selling them. It would take too much time to count the plants in all 100 squares, so she'll accept an estimate based on a sample of 10 squares.

1. Use Table D or technology to take a simple random sample of 10 grid squares. Record the location (for example, B6) of each square you select.

	A	B	C	D	E	F	G	H	I	J
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										

2. This time, you'll take a stratified random sample using the *rows* as strata. Use Table D or technology to randomly select one square from each (horizontal) row. Record the location of each square.

3. Now, take a stratified random sample using the *columns* as strata. Use Table D or technology to randomly select one square from each (vertical) column. Record the location of each square.

4. The table on page N/DS-6 gives the actual number of sunflowers in each grid square. Use the information provided to calculate your estimate of the mean number of sunflowers per square for each of your samples in Steps 1, 2, and 3.

5. Make comparative dotplots showing the mean number of sunflowers obtained using the three different sampling methods for all members of the class. Describe any similarities and differences you see.

6. Your teacher will provide you with the mean number of sunflowers in the population of all 100 grid squares in the field. How did the three sampling methods do?

The following example describes an interesting use of stratified random samples in the music business:

EXAMPLE

Who Wrote That Song?

Using a stratified random sample

A radio station that broadcasts a piece of music owes a royalty to the composer. The organization of composers (called ASCAP) collects these royalties for all its members by charging stations a license fee for the right to play members' songs. ASCAP has four million songs in its catalog and collects $435 million in fees each year. How should ASCAP distribute this income among its members? By sampling: ASCAP tapes about 60,000 hours from the 53 million hours of local radio programs across the country each year.

Radio stations are stratified by type of community (metropolitan, rural), geographic location (New England, Pacific, etc.), and the size of the license fee paid to ASCAP, which reflects the size of the audience. In all, there are 432 strata. Tapes are made at random hours of randomly selected members of each stratum. The tapes are reviewed by experts who can recognize almost every piece of music ever written, and the composers are then paid according to their popularity.[7]

Although a stratified random sample can sometimes give more precise information about a population than an SRS, both sampling methods are hard to use when populations are large and spread out over a wide area. In that situation, we'd prefer a method that selects groups of individuals that are "near" one another. That's the idea of a **cluster sample.**

DEFINITION: Cluster sample and clusters

To take a **cluster sample,** first divide the population into smaller groups. Ideally, these **clusters** should mirror the characteristics of the population. Then choose an SRS of the clusters. All individuals in the chosen clusters are included in the sample.

In a cluster sample, some people take an SRS from each cluster rather than including all members of the cluster.

Imagine a large high school that assigns its students to homerooms alphabetically by last name. The school administration is considering a new schedule and would like student input. Administrators decide to survey 200 randomly selected students. It would be hard to track down an SRS of 200 students, so the administration opts for a cluster sample of homerooms. The principal (who knows some statistics) takes a simple random sample of 8 homerooms and gives the survey to all 25 students in each homeroom.

You might say that strata are ideally "similar within, but different between," while clusters are ideally "different within, but similar between."

Cluster samples are often used for practical reasons, as in the school survey example. They don't offer the statistical advantage of better information about the population that stratified random samples do. That's because clusters are often chosen for ease or convenience, so they may have as much variability as the population itself. Be sure you understand the difference between strata and clusters. We want each stratum to contain similar individuals, and for there to be large differences between strata. For a cluster sample, we'd *like* each cluster to look just like the population, but on a smaller scale.

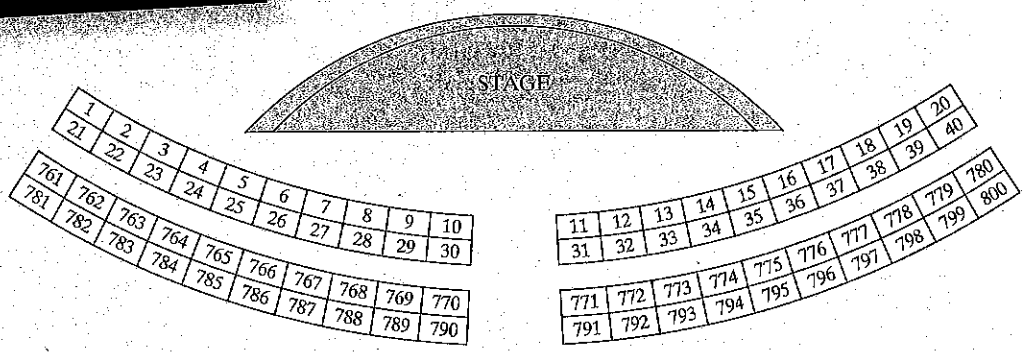
EXAMPLE

Sampling at a School Assembly

Strata or clusters?

The student council wants to conduct a survey during the first five minutes of an all-school assembly in the auditorium about use of the school library. They would like to announce the results of the survey at the end of the assembly. The student council president asks your statistics class to help carry out the survey.

PROBLEM: There are 800 students present at the assembly. A map of the auditorium is shown below. Note that students are seated by grade level and that the seats are numbered from 1 to 800.



9th grade: Seats 601–800	10th grade: Seats 401–600
11th grade: Seats 201–400	12th grade: Seats 1–200

Describe how you would use each of the following sampling methods to select 80 students to complete the survey.

(a) Simple random sample

(b) Stratified random sample

(c) Cluster sample

SOLUTION:

(a) To take an SRS, we need to choose 80 of the seat numbers at random. Use randInt(1,800) on your calculator until 80 different seats are selected. Then give the survey to the students in those seats.

Can you give an advantage and a disadvantage of each sampling method used in the example?

(b) The students in the assembly are seated by grade level. Since students' library use might be similar within grade levels but different across grade levels, we'll use the grade level seating areas as our strata. Within each grade's seating area, we'll select 20 seats at random. For the 9th grade, use randInt(601,800) to select 20 seats. Use randInt(401,600) to pick 20 sophomore seats, randInt(201,400) to get 20 junior seats, and randInt(1,200) to choose 20 senior seats. Give the survey to the students in the selected seats.

(c) With the way students are seated, each column of seats from the stage to the back of the auditorium could be used as a cluster. Note that each cluster contains students from all four grade levels, so each should represent the population well. Since there are 20 clusters, each with 40 seats, we need to choose 2 clusters at random to get 80 students for the survey. Use randInt(1,20) to select two clusters, and then give the survey to all 40 students in each column of seats.

For Practice *Try Exercise* 21

Most large-scale sample surveys use *multistage samples* that combine two or more sampling methods. For example, the Current Population Survey's monthly survey of 60,000 households uses the following design:

Stage 1: Divide the United States into 2007 geographical areas called Primary Sampling Units, or PSUs. Select a sample of 756 PSUs. This sample includes the 428 PSUs with the largest population and a stratified random sample of 328 of the others.

Stage 2: Divide each PSU selected into smaller areas called "neighborhoods." Stratify the neighborhoods using ethnic and other information, and take a stratified random sample of the neighborhoods in each PSU.

Stage 3: Sort the housing units in each neighborhood into clusters of four nearby units. Interview the households in a random sample of these clusters.

Analyzing data from sampling methods more complex than an SRS takes us beyond basic statistics. But the SRS is the building block of more elaborate methods, and the principles of analysis remain much the same for these other methods.

CHECK YOUR UNDERSTANDING

The manager of a sports arena wants to learn more about the financial status of the people who are attending an NBA basketball game. He would like to give a survey to a representative sample of the more than 20,000 fans in attendance. Ticket prices for the game vary a great deal: seats near the court cost over $100 each, while seats in the top rows of the arena cost $25 each. The arena is divided into 30 numbered sections, from 101 to 130. Each section has rows of seats labeled with letters from A (nearest the court) to ZZ (top row of the arena).

1. Explain why it might be difficult to give the survey to an SRS of 200 fans.

2. Which would be a better way to take a stratified random sample of fans: using the lettered rows or the numbered sections as strata? Explain.

3. Which would be a better way to take a cluster sample of fans: using the lettered rows or the numbered sections as clusters? Explain.

Inference for Sampling

Inference

The purpose of a sample is to give us information about a larger population. The process of drawing conclusions about a population on the basis of sample data is called **inference** because we *infer* information about the population from what we *know* about the sample. Inference from convenience samples or voluntary response samples would be misleading because these methods of choosing a sample are biased. We are almost certain that the sample does *not* fairly represent the population. *The first reason to rely on random sampling is to eliminate bias in selecting samples from the list of available individuals.*

Still, it is unlikely that results from a random sample are exactly the same as for the entire population. Sample results, like the unemployment rate obtained from the monthly Current Population Survey, are only estimates of the truth about the population. If we select two samples at random from the same population, we will almost certainly choose different individuals. So the sample results will differ somewhat, just by chance. Properly designed samples avoid systematic bias. But their results are rarely exactly correct, and we expect them to vary from sample to sample.

ACTIVITY *Results may vary . . .*

MATERIALS: At least 200 colored chips with exactly 60% of a particular color (say, red); large bag or other container that students can't see through

In this Activity, you will explore how the results of repeated random samples vary in relation to the population truth.

1. Your teacher will prepare a large population of colored chips with 60% of a particular color.

2. One at a time, each student in the class will take an SRS of 20 chips and record the proportion of chips obtained that are red. The chips should be returned to the container before the next student chooses a sample.

3. Make a class dotplot of the sample proportions for this color of chip. Where is the graph centered? How much do the sample proportions vary around the center? Is there a clear shape?

Why can we trust random samples? As the previous Activity illustrates, the results of random sampling don't change haphazardly from sample to sample. Because we deliberately use chance, the results obey the laws of probability that govern chance behavior. These laws allow us to say how likely it is that sample results are close to the truth about the population. *The second reason to use random sampling is that the laws of probability allow trustworthy inference about the population.* Results from random samples come with a **margin of error** that sets bounds on the size of the likely error. We will discuss the details of inference for sampling later.

Margin of error

One point is worth making now: *larger random samples give better information about the population than smaller samples.* For instance, in the Activity, you would get a better estimate of the proportion of red chips in the population using random samples of 40 chips than 20 chips. By taking a very large sample, you can be confident that the sample result is very close to the truth about the population.

The Current Population Survey contacts about 60,000 households, so we'd expect its estimate of the national unemployment rate to be within about 0.1% of the actual population value. Opinion polls that contact 1000 or 1500 people give less precise results—we expect the sample result to be within about 3% of the actual population percent with a given opinion. Of course, only samples chosen by chance carry this guarantee. Lou Dobbs's online sample tells us little about overall American public opinion even though 7350 people clicked a response.

Sample Surveys: What Can Go Wrong?

Random sampling eliminates bias in choosing a sample. But even a large random sample will give a result that differs from the truth about the population. This "sampling variability" is described by the margin of error that comes with most poll results. So once we see the magic words *randomly selected,* do we know we have trustworthy information? It certainly beats voluntary response, but not always by as much as we might hope. Sampling in the real world is more complex and less reliable than choosing an SRS from a list of names in a textbook exercise.

Most sample surveys are affected by errors in addition to sampling variability. These errors can introduce bias that makes a survey result meaningless. Good sampling technique includes the art of reducing all sources of error. Let's look at the two main sources of errors in sample surveys: sampling errors and nonsampling errors.

Sampling errors The margin of error tells us how much sampling variability to expect, and we can control it by choosing the size of our random sample. It doesn't tell us about sampling errors, mistakes made in the process of taking a sample that could lead to inaccurate information about the population. One source of sampling error is the use of *bad sampling methods,* such as voluntary response. We can avoid bad methods. Other sampling errors are not so easy to handle. Sampling often begins with a list of individuals from which we will draw our sample. This list is called the **sampling frame.** Ideally, the sampling frame should list every individual in the population. Because a list of the entire population is rarely available, most samples suffer from some degree of **undercoverage.**

Sampling frame

DEFINITION: Undercoverage

Undercoverage occurs when some groups in the population are left out of the process of choosing the sample.

A sample survey of households, for example, will miss not only homeless people but prison inmates and students in dormitories. An opinion poll conducted by calling landline telephone numbers will miss households that have only cell phones as well as households without a phone. The results of national sample surveys therefore have some bias due to undercoverage if the people not covered differ systematically from the rest of the population.

Sampling errors in careful sample surveys are usually quite small. The real problems start when someone picks up (or doesn't pick up) the phone. Now nonsampling errors take over.

Nonsampling errors Nonsampling errors are those that can plague even a census. One of the most serious sources of bias in sample surveys is **nonresponse,** which occurs when a selected individual cannot be contacted or refuses to cooperate. Nonresponse to sample surveys often exceeds 50%, even with careful planning and several callbacks. Because nonresponse is higher in urban areas, most sample surveys substitute other people in the same area to avoid favoring rural areas in the final sample. If the people contacted differ from those who are rarely at home or who refuse to answer questions, some bias remains. For example, retired people may be more likely to respond, which would give their opinions more weight. In a poll about Social Security reform, this could give a misleading impression of the population's views.

DEFINITION: Nonresponse

Nonresponse occurs when an individual chosen for the sample can't be contacted or refuses to participate.

CAUTION

Some students misuse the term "voluntary response" to explain why certain individuals don't respond in a sample survey. Their idea is that participation in the survey is optional (voluntary), so anyone can refuse to take part. What the students are actually describing is nonresponse. Think about it this way: nonresponse can occur only after a sample has been selected. In a voluntary response sample, every individual has opted to take part, so there won't be any nonresponse.

EXAMPLE

The ACS, GSS, and Opinion Polls

How bad is nonresponse?

The Census Bureau's American Community Survey (ACS) has the lowest nonresponse rate of any poll we know: only about 1% of the households in the sample refuse to respond. The overall nonresponse rate, including "never at home" and other causes, is just 2.5%.[8] This monthly survey of about 250,000 households replaces the "long form" that in the past was sent to some households in the every-ten-years national census. Participation in the ACS is mandatory, and the Census Bureau follows up by telephone and then in person if a household doesn't return the mail questionnaire.

The University of Chicago's General Social Survey (GSS) is the nation's most important social science survey (see Figure 4.2). The GSS contacts its sample in person, and it is run by a university. Despite these advantages, its most recent survey had a 30% rate of nonresponse.

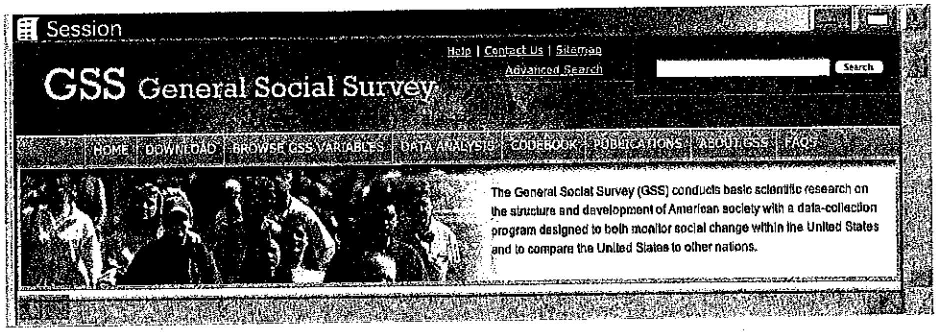
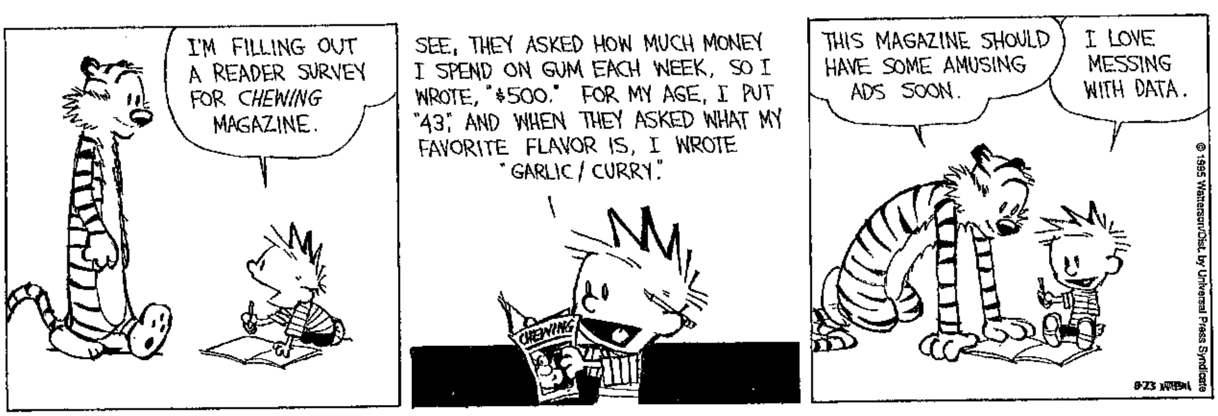


FIGURE 4.2 The home page of the General Social Survey at the University of Chicago's National Opinion Research Center. The GSS has tracked opinions about a wide variety of issues since 1972.

What about opinion polls by news media and opinion-polling firms? We don't know their rates of nonresponse because they won't say. That's a bad sign. The Pew Research Center for the People and the Press imitated a careful random digit dialing survey and published the results: over 5 days, the survey reached 76% of the households in its chosen sample, but "because of busy schedules, skepticism and outright refusals, interviews were completed in just 38% of households that were reached." Combining households that could not be contacted with those who did not complete the interview gave a nonresponse rate of 73%.[9]

Response bias

Another type of nonsampling error occurs when someone gives an incorrect response. A systematic pattern of incorrect responses in a sample survey leads to **response bias.** People know that they should take the trouble to vote, for example, so many who didn't vote in the last election will tell an interviewer that they did. The race or gender of the interviewer can influence responses to questions about race relations or attitudes toward feminism. Answers to questions that ask people to recall past events are often inaccurate because of faulty memory. "Have you visited a dentist in the last 6 months?" will often draw a "Yes" from someone who last visited a dentist 8 months ago.[10] Careful training of interviewers and careful supervision to avoid variation among the interviewers can reduce response bias. Good interviewing technique is another aspect of a well-done sample survey.



Wording of questions

The **wording of questions** is the most important influence on the answers given to a sample survey. Confusing or leading questions can introduce strong bias, and changes in wording can greatly change a survey's outcome. Even the order in which questions are asked matters.

EXAMPLE

How Do Americans Feel about Illegal Immigrants?

Question wording matters

"Should illegal immigrants be prosecuted and deported for being in the U.S. illegally, or shouldn't they?" Asked this question in an opinion poll, 69% favored deportation. But when the very same sample was asked whether illegal immigrants who have worked in the United States for two years "should be given a chance to keep their jobs and eventually apply for legal status," 62% said that they should. Different questions give quite different impressions of attitudes toward illegal immigrants.

CAUTION

Don't trust the results of a sample survey until you have read the exact questions asked. The amount of nonresponse and the date of the survey are also important. Good statistical design is a part, but only a part, of a trustworthy survey.

THINK ABOUT IT

Does the order matter? Ask a sample of college students these two questions:

"How happy are you with your life in general?" (Answers on a scale of 1 to 5)
"How many dates did you have last month?"

There is almost no association between responses to the two questions when asked in this order. It appears that dating has little to do with happiness. Reverse the order of the questions, however, and a much stronger association appears: college students who say they had more dates tend to give higher ratings of happiness about life. Asking a question that brings dating to mind makes dating success a big factor in happiness.

CHECK YOUR UNDERSTANDING

1. Each of the following is a source of error in a sample survey. Label each as *sampling error* or *nonsampling error,* and explain your answers.

(a) The telephone directory is used as a sampling frame.

(b) The person cannot be contacted in five calls.

(c) Interviewers choose people walking by on the sidewalk to interview.

2. A survey paid for by makers of disposable diapers found that 84% of the sample opposed banning disposable diapers. Here is the actual question:

> It is estimated that disposable diapers account for less than 2% of the trash in today's landfills. In contrast, beverage containers, third-class mail and yard wastes are estimated to account for about 21% of the trash in landfills. Given this, in your opinion, would it be fair to ban disposable diapers?[11]

Explain how the wording of the question could result in bias. Be sure to specify the direction of the bias.

SECTION 4.1 Summary

- A **sample survey** selects a **sample** from the **population** of all individuals about which we desire information. We base conclusions about the population on data from the sample. It is important to specify exactly what population you are interested in and what variables you will measure.
- **Random sampling** uses chance to select a sample.
- The basic random sampling method is a **simple random sample** (SRS). An SRS gives every possible sample of a given size the same chance to be chosen. Choose an SRS by labeling the members of the population and using **random digits** to select the sample. Technology can automate this process.
- To choose a **stratified random sample,** divide the population into **strata,** groups of individuals that are similar in some way that might affect their responses. Then choose a separate SRS from each stratum.
- To choose a **cluster sample,** divide the population into groups, or **clusters.** Randomly select some of these clusters. All the individuals in the chosen clusters are included in the sample.
- Failure to use random sampling often results in **bias,** or systematic errors in the way the sample represents the population. **Voluntary response samples,** in which the respondents choose themselves, and **convenience samples,** in which individuals close by are included in the sample, are particularly prone to large bias.
- **Sampling errors** come from the act of choosing a sample. Random sampling error and **undercoverage** are common types of sampling error. Undercoverage occurs when some members of the population are left out of the **sampling frame,** the list from which the sample is actually chosen.
- The most serious errors in most careful surveys, however, are **nonsampling errors.** These have nothing to do with choosing a sample—they are present even in a census. The single biggest problem for sample surveys is **nonresponse:** people can't be contacted or refuse to answer. Incorrect answers by respondents can lead to **response bias.** Finally, the exact **wording of questions** has a big influence on the answers.

4.1 TECHNOLOGY CORNER

Choosing an SRS .... page 214

TI-Nspire instructions in Appendix B

Name: ______________________________ AP STATISTICS

Sampling & Surveys Exercises (Ch. 4.1)

1. A high school's student newspaper plans to survey local businesses about the importance of students as customers. From telephone book listings, the newspaper staff chooses 150 businesses at random. Of these, 73 return the questionnaire mailed by the staff. Identify the population and the sample.

   Population: ____________________ Sample: ______________________

2. A large retailer prepares its customers' monthly credit card bills using an automatic machine that folds the bills, stuffs them into envelopes, and seals the envelopes for mailing. Are the envelopes completely sealed? Inspectors choose 40 envelopes from the 1000 stuffed each hour for visual inspection. Identify the population and the sample.

   Population: ____________________ Sample: ______________________

3. You are on the staff of a member of Congress who is considering a bill that would provide government-sponsored insurance for nursing-home care. You report that 1128 letters have been received on the issue, of which 871 oppose the legislation. "I'm surprised that most of my constituents oppose the bill. I thought it would be quite popular," says the congresswoman. Are you convinced that a majority of the voters oppose the bill? How would you explain the statistical issue to the congresswoman?

4. Many websites include customer reviews of products, restaurants, hotels, and so on. The manager of a hotel was upset to see that 26% of reviewers on a travel website gave the hotel "1 star" - the lowest possible rating. Explain how bias in the sampling method could affect the estimate.

5. How much sleep do high school students get on a typical school night? An interested student designed a survey to find out. To make data collection easier, the student surveyed the first 100 students to arrive at school on a particular morning. These students reported an average of 7.2 hours of sleep on the previous night.

   a) What type of sample did the student obtain? ______________________________

   b) Explain why this sampling method is biased. Is 7.2 hours probably higher or lower than the true average amount of sleep last night for all students at the school? Why?

6. Suppose 1000 iPhones are produced at a factory today. Management would like to ensure that the phones' display screens meet their quality standards before shipping them to retail stores. Since it takes about 10 minutes to inspect an individual phone's display screen, managers decide to inspect a sample of 20 phones from the day's production.

   a) An eager employee suggests that it would be easy to inspect the last 20 iPhones that were produced today. Why isn't this a good idea?

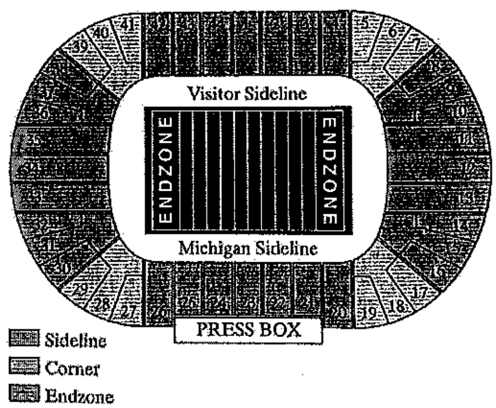
   b) Another employee recommends inspecting every fiftieth iPhone that is produced. Explain why this sampling method is <u>not</u> and SRS.

7. Laying fiber-optic cable is expensive. Cable companies want to make sure that, if they extend their lines out to less dense suburban or rural areas, there will be sufficient demand and the work will be cost-effective. They decide to conduct a survey to determine the proportion of households in a rural subdivision that would buy the service. They select a simple random sample of 5 blocks in the subdivision and survey each family that lives on one of those blocks.

What is the name for this kind of sampling method? ____________________________

8. A corporation employs 2000 male and 500 female engineers. A stratified random sample of 200 male and 50 female engineers gives every individual in the population the same chance to be chosen for the sample. Is it an SRS? Explain.

9. Michigan Stadium, also known as "The Big House," seats over 100,000 fans for a football game. The University of Michigan athletic department plans to conduct a survey about concessions that are sold during games. Tickets are most expensive for seats near the field and on the sideline. The cheapest seats are high up in the end zones. A map of the stadium is shown.



a) The athletic department is considering a stratified random sample. What would you recommend as the strata? Why?

b) Explain why a cluster sample might be easier to obtain. What would you recommend for the clusters? Why?

10. The director of student life at a small college wants to know what percent of students eat regularly in the cafeteria. To find out, the director selects an SRS of 300 students who live in the dorms. Describe how under-coverage might lead to bias in this study. Explain the likely direction of the bias.

11. A total of 300 people participated in a free 12-week weight-loss course at a community health clinic. After one year, administrators emailed each of the 300 participants to see how much weight they had lost since the end of the course. Only 56 participants responded to the survey. The mean weight loss for this sample was 13.6 pounds. Describe how nonresponse might lead to bias in this study. Explain the likely direction of the bias.

12. Two female statistics students asked a random sample of 60 high school boys if they have ever cried during a movie. Thirty of the boys were asked directly and the other 30 were asked anonymously by means of a "secret ballot." When the responses were anonymous, 63% of the boys said "Yes," whereas only 23% of the other group said "Yes." Explain why the two percentages are so different.

13. The following are potential survey questions:
"Some cell phone users have developed brain cancer. Should all cell phones come with a warning label explaining the danger of using cell phones?"

"Do you agree that a national system of health insurance should be favored because it would provide health insurance for everyone and would reduce administrative costs?"

"In view of escalating environmental degradation and incipient resource depletion, would you favor economic incentives for recycling of resource-intensive consumer goods?"

Comment on the errors made in these survey questions.

14. Survey #1:
Ideally, the sampling frame in a survey should list every individual in the population, but in practice, this is often difficult. Suppose that a sample of households in a community is selected at random from the telephone directory.

Survey #2:
Suppose you want to know the average amount of money spent by the fans attending opening day for the Cleveland Indians baseball season. You get permission from the team's management to conduct a survey at the stadium, but they will not allow you to bother the fans in the club seating or box seats (the most expensive seating). Using a computer, you randomly select 500 seats from the rest of the stadium. During the game, you ask the fans in those seats how much they spent that day.

Provide a reason why each of these surveys might yield a biased result.

15. A popular website places opinion poll questions next to many of its news stories. Simply click your response to join the sample. One of the questions was "Do you plan to diet this year?" More than 30,000 people responded, with 68% saying "Yes." Which of the following is true?

   a) About 68% of Americans plan to diet.
   b) The poll used a convenience sample, so the results tell us little about the population of all adults.
   c) The poll uses voluntary response, so the results tell us little about the population of all adults.
   d) The sample is too small to draw any conclusion.
   e) None of these.

16. To gather information about the validity of a new standardized test for 10^{th} grade students in a particular state, a random sample of 15 high schools was selected from the state. The new test was administered to every 10^{th} grade student in the selected high schools. What kind of sample is this?

   a) simple random sample
   b) cluster sample
   c) voluntary response sample
   d) stratified random sample
   e) systematic random sample

17. When we take a census, we attempt to collect data from ...

   a) A stratified random sample.
   b) A convenience sample.
   c) A voluntary response sample.
   d) Every individual in the population.
   e) Every individual selected in an SRS.

18. Suppose that 35% of the voters in a state are registered as Republicans, 40% as Democrats, and 25% as Independents. A newspaper wants to select a sample of 1000 registered voters to predict the outcome of the next election. If it randomly selects 350 Republicans, randomly selects 400 Democrats, and randomly selects 250 Independents. Did this sample procedure result in a simple random sample of registered voters from this state?

   a) Yes, it was a simple random sample.
   b) No, it was a stratified random sample.
   c) No, it was a cluster sample.
   d) No, it was a systematic random sample.
   e) No, it was a convenience sample.

19. A local news agency conducted a survey about unemployment by randomly dialing phone numbers during the work day until it gathered responses from 1000 adults in its state. In the survey, 19% of those who responded said they were not currently employed. In reality, only 6% of the adults in the state were not currently employed at the time of the survey. Which of the following best explains the difference in the two percentages?

a) The difference is due to sampling variability. We shouldn't expect the results of a random sample to match the truth about the population every time.
b) The difference is due to response bias. Adults who are employed are likely to lie and say that they are unemployed.
c) The difference is due to under-coverage bias. The survey included only adults and did not include teenagers who are eligible to work.
d) The difference is due to nonresponse bias. Adults who are employed are less likely to be available for the sample than adults who are unemployed.
e) The difference is due to voluntary response. Adults are able to volunteer as a member of the sample.

20. A simple random sample of 1200 adult Americans is selected, and each person is asked the following question: "In light of the huge national deficit, should the government at this time spend additional money to send humans to Mars?" Only 39% of those responding answered "Yes." This survey ...

a) Is reasonably accurate because it used a large simple random sample.
b) Needs to be larger because only about 24 people were drawn from each state.
c) Probably understates the percent of people who favor sending humans to Mars.
d) Is very inaccurate, but neither understates nor overstates the percent of people who favor sending humans to Mars. Because simple random sampling was used, it is unbiased.
e) Probably overstates the percent of people who favor sending humans to Mars.