CHAPTER 12

Sample Surveys

In 2007, Pew Research conducted a survey to assess Americans' knowledge of current events. They asked a random sample of 1,502 U.S. adults 23 factual questions about topics currently in the news.[1] Pew also asked respondents where they got their news. Those who frequented major newspaper Web sites or who are regular viewers of the *Daily Show* or *Colbert Report* scored best on knowledge of current events.[2] Even among those viewers, only 54% responded correctly to 15 or more of the questions. Pew claimed that this was close to the true percentage responding correctly that they would have found if they had asked all U.S. adults who got their news from those sources. That step from a small sample to the entire population is impossible without understanding Statistics. To make business decisions, to do science, to choose wise investments, or to understand what voters think they'll do the next election, we need to stretch beyond the data at hand to the world at large.

To make that stretch, we need three ideas. You'll find the first one natural. The second may be more surprising. The third is one of the strange but true facts that often confuse those who don't know Statistics.

Idea 1: Examine a Part of the Whole

A S *Activity:* **Populations and Samples.** Explore the differences between populations and samples.

The first idea is to draw a sample. We'd like to know about an entire **population** of individuals, but examining all of them is usually impractical, if not impossible. So we settle for examining a smaller group of individuals—a **sample**—selected from the population.

You do this every day. For example, suppose you wonder how the vegetable soup you're cooking for dinner tonight is going to go over with your friends. To decide whether it meets your standards, you only need to try a small amount. You might taste just a spoonful or two. You certainly don't have to consume the whole pot. You trust that the taste will *represent* the flavor of the entire pot. The idea behind your tasting is that a small sample, if selected properly, can represent the entire population.

---

[1] For example, two of the questions were "Who is the vice-president of the United States?" and "What party controls Congress?"

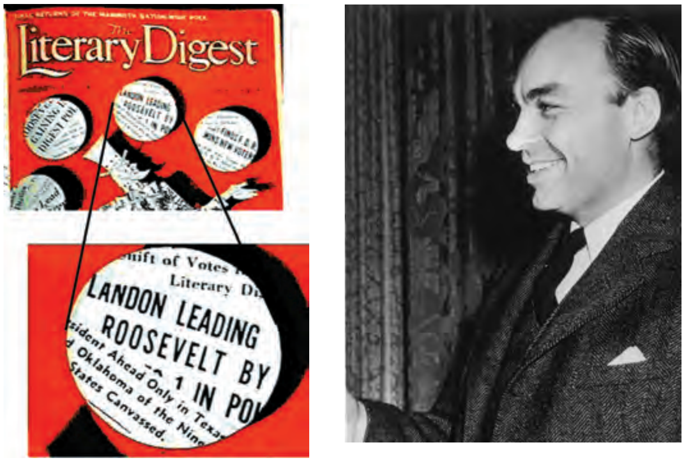
[2] The lowest scores came from those whose main source of news was network morning shows or *Fox News.*

It's hard to go a day without hearing about the latest opinion poll. These polls are examples of **sample surveys,** designed to ask questions of a small group of people in the hope of learning something about the entire population. Most likely, you've never been selected to be part of one of these national opinion polls. That's true of most people. So how can the pollsters claim that a sample is representative of the entire population? The answer is that professional pollsters work quite hard to ensure that the "taste"—the sample that they take—represents the population. If not, the sample can give misleading information about the population.

The W's and Sampling

The population we are interested in is usually determined by the *Why* of our study. The sample we draw will be the *Who*. *When* and *How* we draw the sample may depend on what is practical.

Bias



In 1936, a young pollster named George Gallup used a subsample of only 3000 of the 2.4 million responses that the Literary Digest *received to reproduce the wrong prediction of Landon's victory over Roosevelt. He then used an entirely different sample of 50,000 and predicted that Roosevelt would get 56% of the vote to Landon's 44%. His sample was apparently much more representative of the actual voting populace. The Gallup Organization went on to become one of the leading polling companies.*

Selecting a sample to represent the population fairly is more difficult than it sounds. Polls or surveys most often fail because they use a sampling method that tends to over- or underrepresent parts of the population. The method may overlook subgroups that are harder to find (such as the homeless or those who use only cell phones) or favor others (such as Internet users who like to respond to online surveys). Sampling methods that, by their nature, tend to over- or underemphasize some characteristics of the population are said to be **biased.** Bias is the bane of sampling—the one thing above all to avoid. Conclusions based on samples drawn with biased methods are inherently flawed. There is usually no way to fix bias after the sample is drawn and no way to salvage useful information from it.

Here's a famous example of a really dismal failure. By the beginning of the 20th century, it was common for newspapers to ask readers to return "straw" ballots on a variety of topics. (Today's Internet surveys are the same idea, gone electronic.) The earliest known example of such a straw vote in the United States dates back to 1824.

During the period from 1916 to 1936, the magazine *Literary Digest* regularly surveyed public opinion and forecast election results correctly. During the 1936 presidential campaign between Alf Landon and Franklin Delano Roosevelt, it mailed more than 10 million ballots and got back an astonishing 2.4 million. (Polls were still a relatively novel idea, and many people thought it was important to send back their opinions.) The results were clear: Alf Landon would be the next president by a landslide, 57% to 43%. You remember President Landon? No? In fact, Landon carried only two states. Roosevelt won, 62% to 37%, and, perhaps coincidentally, the *Digest* went bankrupt soon afterward.

What went wrong? One problem was that the *Digest*'s sample wasn't representative. Where would *you* find 10 million names and addresses to sample? The *Digest* used the phone book, as many surveys do.[3] But in 1936, at the height of the Great Depression, telephones were a real luxury, so they sampled more rich than poor voters. The campaign of 1936 focused on the economy, and those who were less well off were more likely to vote for the Democrat. So the *Digest*'s sample was hopelessly biased.

How do modern polls get their samples to *represent* the entire population? You might think that they'd handpick individuals to sample with care and precision.

A S **Video:** ***The Literary Digest*** **Poll and the Election of 1936.** Hear the story of one of the most famous polling failures in history.

---

[3] Today phone numbers are computer-generated to make sure that unlisted numbers are included. But even now, cell phones and VOIP Internet phones are often not included.

But in fact, they do something quite different: They select individuals to sample *at random*. The importance of deliberately using randomness is one of the great insights of Statistics.

Idea 2: Randomize

Think back to the soup sample. Suppose you add some salt to the pot. If you sample it from the top before stirring, you'll get the misleading idea that the whole pot is salty. If you sample from the bottom, you'll get an equally misleading idea that the whole pot is bland. By stirring, you *randomize* the amount of salt throughout the pot, making each taste more typical of the whole pot.

Not only does randomization protect you against factors that you know are in the data, it can also help protect against factors that you didn't even know were there. Suppose, while you weren't looking, a friend added a handful of peas to the soup. If they're down at the bottom of the pot, and you don't randomize the soup by stirring, your test spoonful won't have any peas. By stirring in the salt, you *also* randomize the peas throughout the pot, making your sample taste more typical of the overall pot *even though you didn't know the peas were there.* So randomizing protects us even in this case.

How do we "stir" people in a survey? We select them at random. **Randomizing** protects us from the influences of *all* the features of our population by making sure that, *on average,* the sample looks like the rest of the population.

A S ***Activity:* Sampling from Some Real Populations.** Draw random samples to see how closely they resemble each other and the population.

Why not match the sample to the population? Rather than randomizing, we could try to design our sample so that the people we choose are typical in terms of every characteristic we can think of. We might want the income levels of those we sample to match the population. How about age? Political affiliation? Marital status? Having children? Living in the suburbs? We can't possibly think of all the things that might be important. Even if we could, we wouldn't be able to match our sample to the population for all these characteristics.

FOR EXAMPLE **Is a random sample representative?**

Here are summary statistics comparing two samples of 8000 drawn at random from a company's database of 3.5 million customers:

Age (yr)	White (%)	Female (%)	# of Children	Income Bracket (1–7)	Wealth Bracket (1–9)	Homeowner? (% Yes)
61.4	85.12	56.2	1.54	3.91	5.29	71.36
61.2	84.44	56.4	1.51	3.88	5.33	72.30

Question: Do you think these samples are representative of the population? Explain.

The two samples look very similar with respect to these seven variables. It appears that randomizing has automatically matched them pretty closely. We can reasonably assume that since the two samples don't differ too much from each other, they don't differ much from the rest of the population either.

Idea 3: It's the Sample Size

How large a random sample do we need for the sample to be reasonably representative of the population? Most people think that we need a large percentage, or *fraction,* of the population, but it turns out that what matters is the number of individuals *in the sample*, not the size of the population. A random sample of 100 students in a college represents the student body just about as well as a random sample of 100 voters represents the entire electorate of the United States. This is the *third* idea and probably the most surprising one in designing surveys.

Activity: **Does the Population Size Matter?** Here's the narrated version of this important idea about sampling.

How can it be that only the size of the sample, and not the population, matters? Well, let's return one last time to that pot of soup. If you're cooking for a banquet rather than just for a few people, your pot will be bigger, but do you need a bigger spoon to decide how the soup tastes? Of course not. The same-size spoonful is probably enough to make a decision about the entire pot, no matter how large the pot. The *fraction* of the population that you've sampled doesn't matter.[4] It's the **sample size** itself that's important.

How big a sample do you need? That depends on what you're estimating. To get an idea of what's really in the soup, you'll need a large enough taste to get a *representative* sample from the pot. For a survey that tries to find the proportion of the population falling into a category, you'll usually need several hundred respondents to say anything precise enough to be useful.[5]

A friend who knows that you are taking Statistics asks your advice on her study. What can you possibly say that will be helpful? Just say, "If you could just get a larger sample, it would probably improve your study." Even though a larger sample might not be worth the cost, it will almost always make the results more precise.

TI-nspire

Populations and Samples. How well can a sample reveal the population's shape, center, and spread? Explore what happens as you change the sample size.

What do the pollsters do? How do professional polling agencies do their work? The most common polling method today is to contact respondents by telephone. Computers generate random telephone numbers, so pollsters can even call some people with unlisted phone numbers. The person who answers the phone is invited to respond to the survey—if that person qualifies. (For example, only if it's an adult who lives at that address.) If the person answering doesn't qualify, the caller will ask for an appropriate alternative. In phrasing questions, pollsters often list alternative responses (such as candidates' names) in different orders to avoid biases that might favor the first name on the list.

Do these methods work? The Pew Research Center for the People and the Press, reporting on one survey, says that

> *Across five days of interviewing, surveys today are able to make some kind of contact with the vast majority of households (76%), and there is no decline in this contact rate over the past seven years. But because of busy schedules, skepticism and outright refusals, interviews were completed in just 38% of households that were reached using standard polling procedures.*

Nevertheless, studies indicate that those actually sampled can give a good snapshot of larger populations from which the surveyed households were drawn.

Does a Census Make Sense?

Video: **Frito-Lay Sampling for Quality.** How does a potato chip manufacturer make sure to cook only the best potatoes?

Why bother determining the right sample size? Wouldn't it be better to just include everyone and "sample" the entire population? Such a special sample is called a **census.** Although a census would appear to provide the best possible information about the population, there are a number of reasons why it might not.

First, it can be difficult to complete a census. Some individuals in the population will be hard (and expensive) to locate. Or a census might just be impractical. If you were a taste tester for the Hostess™ Company, you probably wouldn't want to census *all* the Twinkies on the production line. Not only might this be life-endangering, but you wouldn't have any left to sell.

---

[4] Well, that's not exactly true. If the population is small enough and the sample is more than 10% of the whole population, it *can* matter. It doesn't matter whenever, as usual, our sample is a very small fraction of the population.

[5] Chapter 19 gives the details behind this statement and shows how to decide on a sample size for a survey.

Second, populations rarely stand still. In populations of people, babies are born and folks die or leave the country. In opinion surveys, events may cause a shift in opinion during the survey. A census takes longer to complete and the population changes while you work. A sample surveyed in just a few days may give more accurate information.

Third, taking a census can be more complex than sampling. For example, the U.S. Census records too many college students. Many are counted once with their families and are then counted a second time in a report filed by their schools.

The undercount. It's particularly difficult to compile a complete census of a population as large, complex, and spread out as the U.S. population. The U.S. Census is known to miss some residents. On occasion, the undercount has been striking. For example, there have been blocks in inner cities in which the number of residents recorded by the Census was smaller than the number of electric meters for which bills were being paid. What makes the problem particularly important is that some groups have a higher probability of being missed than others—undocumented immigrants, the homeless, the poor. The Census Bureau proposed the use of random sampling to estimate the number of residents missed by the ordinary census. Unfortunately, the resulting debate has become more political than statistical.

Populations and Parameters

Any quantity that we calculate from data could be called a "statistic." But in practice, we usually use a statistic to estimate a population parameter.

A S *Activity:* **Statistics and Parameters.** Explore the difference between statistics and parameters.

Remember: Population model parameters are not just unknown—usually they are *unknowable*. We have to settle for sample statistics.

A study found that teens were less likely to "buckle up." The National Center for Chronic Disease Prevention and Health Promotion reports that 21.7% of U.S. teens never or rarely wear seatbelts. We're sure they didn't take a census, so what *does* the 21.7% mean? We can't know what percentage of teenagers wear seatbelts. Reality is just too complex. But we can simplify the question by building a model.

Models use mathematics to represent reality. Parameters are the key numbers in those models. A parameter used in a model for a population is sometimes called (redundantly) a **population parameter.**

But let's not forget about the data. We use summaries of the data to estimate the population parameters. As we know, any summary found from the data is a **statistic.** Sometimes you'll see the (also redundant) term **sample statistic.**[6]

We've already met two parameters in Chapter 6: the mean, μ, and the standard deviation, σ. We'll try to keep denoting population model parameters with Greek letters and the corresponding statistics with Latin letters. Usually, but not always, the letter used for the statistic and the parameter correspond in a natural way. So the standard deviation of the data is s, and the corresponding parameter is σ (Greek for s). In Chapter 7, we used r to denote the sample correlation. The corresponding correlation in a model for the population would be called ρ (rho). In Chapter 8, b_1 represented the slope of a linear regression estimated from the data. But when we think about a (linear) *model* for the population, we denote the slope parameter β_1 (beta).

Get the pattern? Good. Now it breaks down. We denote the mean of a population model with μ (because μ is the Greek letter for m). It might make sense to denote the sample mean with m, but long-standing convention is to put a bar over anything when we average it, so we write $\bar{y}$. What about proportions? Suppose we want to talk about the proportion of teens who don't wear seatbelts. If we use p to denote the proportion from the data, what is the corresponding model parameter? By all rights it should be π. But statements like $\pi = 0.25$ might be confusing because π has been equal to 3.1415926 . . . for so long, and it's worked so *well*. So, once again we violate the rule. We'll use p for the population model parameter and $\hat{p}$ for the proportion from the data (since, like $\hat{y}$ in regression, it's an estimated value).

---

[6] Where else besides a sample *could* a statistic come from?

Here's a table summarizing the notation:

NOTATION ALERT:

This entire table is a notation alert.

Name	Statistic	Parameter
Mean	$\bar{y}$	μ (mu, pronounced "meeoo," not "moo")
Standard deviation	s	σ (sigma)
Correlation	r	ρ (rho)
Regression coefficient	b	β (beta, pronounced "baytah"[7])
Proportion	$\hat{p}$	p (pronounced "pee"[8])

We draw samples because we can't work with the entire population, but we want the statistics we compute from a sample to reflect the corresponding parameters accurately. A sample that does this is said to be **representative.** A biased sampling methodology tends to over- or underestimate the parameter of interest.

JUST CHECKING

1. Various claims are often made for surveys. Why is each of the following claims not correct?

- **a)** It is always better to take a census than to draw a sample.
- **b)** Stopping students on their way out of the cafeteria is a good way to sample if we want to know about the quality of the food there.
- **c)** We drew a sample of 100 from the 3000 students in a school. To get the same level of precision for a town of 30,000 residents, we'll need a sample of 1000.
- **d)** A poll taken at a statistics support Web site garnered 12,357 responses. The majority said they enjoy doing statistics homework. With a sample size that large, we can be pretty sure that most Statistics students feel this way, too.
- **e)** The true percentage of all Statistics students who enjoy the homework is called a "population statistic."

Simple Random Samples

How would you select a representative sample? Most people would say that every individual in the population should have an equal chance to be selected, and certainly that seems fair. But it's not sufficient. There are many ways to give everyone an equal chance that still wouldn't give a representative sample. Consider, for example, a school that has equal numbers of males and females. We could sample like this: Flip a coin. If it comes up heads, select 100 female students at random. If it comes up tails, select 100 males at random. Everyone has an equal chance of selection, but every sample is of only a single sex—hardly representative.

We need to do better. Suppose we insist that every possible *sample* of the size we plan to draw has an equal chance to be selected. This ensures that situations like the one just described are not likely to occur and still guarantees that each person has an equal chance of being selected. What's different is that with this method, each *combination* of people has an equal chance of being selected as well. A sample drawn in this way is called a **Simple Random Sample,** usually abbreviated **SRS.** An SRS is the standard against which we measure other sampling methods, and the sampling method on which the theory of working with sampled data is based.

To select a sample at random, we first need to define where the sample will come from. The **sampling frame** is a list of individuals from which the sample is drawn.

---

[7] If you're from the United States. If you're British or Canadian, it's "beetah."

[8] Just in case you weren't sure.

For example, to draw a random sample of students at a college, we might obtain a list of all registered full-time students and sample from that list. In defining the sampling frame, we must deal with the details of defining the population. Are part-time students included? How about those who are attending school elsewhere and transferring credits back to the college?

Once we have a sampling frame, the easiest way to choose an SRS is to assign a random number to each individual in the sampling frame. We then select only those whose random numbers satisfy some rule.[9] Let's look at some ways to do this.

FOR EXAMPLE **Using random numbers to get an SRS**

There are 80 students enrolled in an introductory Statistics class; you are to select a sample of 5.

Question: How can you select an SRS of 5 students using these random digits found on the Internet: 05166 29305 77482?

First I'll number the students from 00 to 79. Taking the random numbers two digits at a time gives me 05, 16, 62, 93, 05, 77, and 48. I'll ignore 93 because the students were numbered only up to 79. And, so as not to pick the same person twice, I'll skip the repeated number 05. My simple random sample consists of students with the numbers 05, 16, 62, 77, and 48.

Error Okay, Bias Bad!
Sampling variability is sometimes referred to as *sampling error*, making it sound like it's some kind of mistake. It's not. We understand that samples will vary, so "sampling error" is to be expected. It's *bias* we must strive to avoid. Bias means our sampling method distorts our view of the population, and that will surely lead to mistakes.

- We can be more efficient when we're choosing a larger sample from a sampling frame stored in a data file. First we assign a random number with several digits (say, from 0 to 10,000) to each individual. Then we arrange the random numbers in numerical order, keeping each name with its number. Choosing the first n names from this re-arranged list will give us a random sample of that size.
- Often the sampling frame is so large that it would be too tedious to number everyone consecutively. If our intended sample size is approximately 10% of the sampling frame, we can assign each individual a single random digit 0 to 9. Then we select only those with a specific random digit, say, 5.

Samples drawn at random generally differ one from another. Each draw of random numbers selects *different* people for our sample. These differences lead to different values for the variables we measure. We call these sample-to-sample differences **sampling variability.** Surprisingly, sampling variability isn't a problem; it's an opportunity. In future chapters we'll investigate what the variation in a sample can tell us about its population.

Stratified Sampling

Simple random sampling is not the only fair way to sample. More complicated designs may save time or money or help avoid sampling problems. All statistical sampling designs have in common the idea that chance, rather than human choice, is used to select the sample.

Designs that are used to sample from large populations—especially populations residing across large areas—are often more complicated than simple random samples. Sometimes the population is first sliced into homogeneous groups, called **strata,** before the sample is selected. Then simple random sampling is used within each stratum before the results are combined. This common sampling design is called **stratified random sampling.**

Why would we want to complicate things? Here's an example. Suppose we want to learn how students feel about funding for the football team at a large university. The campus is 60% men and 40% women, and we suspect that men and women have different views on the funding. If we use simple random sampling to select 100 people for the survey, we could end up with 70 men and 30 women or 35 men and 65 women. Our resulting estimates of the level of support for the football funding could vary widely. To help reduce this sampling variability, we can decide to force a representative balance, selecting 60 men at random and 40 women at random. This would guarantee that the proportions of men and women within our sample match the proportions in the population, and that should make such samples more accurate in representing population opinion.

---

[9] Chapter 11 presented ways of finding and working with random numbers.

You can imagine the importance of stratifying by race, income, age, and other characteristics, depending on the questions in the survey. Samples taken within a stratum vary less, so our estimates can be more precise. This reduced sampling variability is the most important benefit of stratifying.

Stratified sampling can also help us notice important differences among groups. As we saw in Chapter 3, if we unthinkingly combine group data, we risk reaching the wrong conclusion, becoming victims of Simpson's paradox.

FOR EXAMPLE **Stratifying the sample**

Recap: You're trying to find out what freshmen think of the food served on campus. Food Services believes that men and women typically have different opinions about the importance of the salad bar.

Question: How should you adjust your sampling strategy to allow for this difference?

I will stratify my sample by drawing an SRS of men and a separate SRS of women—assuming that the data from the registrar include information about each person's sex.

Cluster and Multistage Sampling

Suppose we wanted to assess the reading level of this textbook based on the length of the sentences. Simple random sampling could be awkward; we'd have to number each sentence, then find, for example, the 576th sentence or the 2482nd sentence, and so on. Doesn't sound like much fun, does it?

It would be much easier to pick a few *pages* at random and count the lengths of the sentences on those pages. That works if we believe that each page is representative of the entire book in terms of reading level. Splitting the population into representative **clusters** can make sampling more practical. Then we could simply select one or a few clusters at random and perform a census within each of them. This sampling design is called **cluster sampling.** If each cluster represents the full population fairly, cluster sampling will be unbiased.

FOR EXAMPLE **Cluster sampling**

Recap: In trying to find out what freshmen think about the food served on campus, you've considered both an SRS and a stratified sample. Now you have run into a problem: It's simply too difficult and time consuming to track down the individuals whose names were chosen for your sample. Fortunately, freshmen at your school are all housed in 10 freshman dorms.

Questions: How could you use this fact to draw a cluster sample? How might that alleviate the problem? What concerns do you have?

To draw a cluster sample, I would select one or two dorms at random and then try to contact everyone in each selected dorm. I could save time by simply knocking on doors on a given evening and interviewing people. I'd have to assume that freshmen were assigned to dorms pretty much at random and that the people I'm able to contact are representative of everyone in the dorm.

What's the difference between cluster sampling and stratified sampling? We stratify to ensure that our sample represents different groups in the population, and we sample randomly within each stratum. Strata are internally homogeneous, but differ from one another. By contrast, clusters are internally heterogeneous, each resembling the overall population. We select clusters to make sampling more practical or affordable.

Stratified vs. cluster sampling. Boston cream pie consists of a layer of yellow cake, a layer of pastry creme, another cake layer, and then a chocolate frosting. Suppose you are a professional taster (yes, there really are such people) whose job is to check your company's pies for quality. You'd need to eat small samples of randomly selected pies, tasting all three components: the cake, the creme, and the frosting.

One approach is to cut a thin vertical slice out of the pie. Such a slice will be a lot like the entire pie, so by eating that slice, you'll learn about the whole pie. This vertical slice containing all the different ingredients in the pie would be a *cluster* sample.

Another approach is to sample in *strata:* Select some tastes of the cake at random, some tastes of creme at random, and some bits of frosting at random. You'll end up with a reliable judgment of the pie's quality.

Many populations you might want to learn about are like this Boston cream pie. You can think of the subpopulations of interest as horizontal strata, like the layers of pie. Cluster samples slice vertically across the layers to obtain clusters, each of which is representative of the entire population. Stratified samples represent the population by drawing some from each layer, reducing variability in the results that could arise because of the differences among the layers.

Strata or Clusters?
We may split a population into strata or clusters. What's the difference? We create strata by dividing the population into groups of similar individuals so that each stratum is different from the others. By contrast, since clusters each represent the entire population, they all look pretty much alike.

Sometimes we use a variety of sampling methods together. In trying to assess the reading level of this book, we might worry that it starts out easy and then gets harder as the concepts become more difficult. If so, we'd want to avoid samples that selected heavily from early or from late chapters. To guarantee a fair mix of chapters, we could randomly choose one chapter from each of the seven parts of the book and then randomly select a few pages from each of those chapters. If, altogether, that made too many sentences, we might select a few sentences at random from each of the chosen pages. So, what is our sampling strategy? First we stratify by the part of the book and randomly choose a chapter to represent each stratum. Within each selected chapter, we choose pages as clusters. Finally, we consider an SRS of sentences within each cluster. Sampling schemes that combine several methods are called **multistage samples.** Most surveys conducted by professional polling organizations use some combination of stratified and cluster sampling as well as simple random samples.

FOR EXAMPLE Multistage sampling

Recap: Having learned that freshmen are housed in separate dorms allowed you to sample their attitudes about the campus food by going to dorms chosen at random, but you're still concerned about possible differences in opinions between men and women. It turns out that these freshmen dorms house the sexes on alternate floors.

Question: How can you design a sampling plan that uses this fact to your advantage?

Now I can stratify my sample by sex. I would first choose one or two dorms at random and then select some dorm floors at random from among those that house men and, separately, from among those that house women. I could then treat each floor as a cluster and interview everyone on that floor.

Systematic Samples

Some samples select individuals systematically. For example, you might survey every 10th person on an alphabetical list of students. To make it random, you still must start the systematic selection from a randomly selected individual. When the order of the list is not associated in any way with the responses sought, **systematic sampling** can give a representative sample. Systematic sampling can be much less expensive than true random sampling. When you use a systematic sample, you should justify the assumption that the systematic method is not associated with any of the measured variables.

Think about the reading-level sampling example again. Suppose we have chosen a chapter of the book at random, then three pages at random from that chapter, and now we want to select a sample of 10 sentences from the 73 sentences found on those pages. Instead of numbering each sentence so we can pick a simple random sample, it would be easier to sample systematically. A quick calculation shows $73/10 = 7.3$, so we can get our sample by just picking every seventh sentence on the page. But where should you start? At random, of course. We've accounted for $10 \times 7 = 70$ of the sentences, so we'll throw the extra 3 into the starting group and choose a sentence at random from the first 10. Then we pick every seventh sentence after that and record its length.

JUST CHECKING

2. We need to survey a random sample of the 300 passengers on a flight from San Francisco to Tokyo. Name each sampling method described below.
   a) Pick every 10th passenger as people board the plane.
   b) From the boarding list, randomly choose 5 people flying first class and 25 of the other passengers.
   c) Randomly generate 30 seat numbers and survey the passengers who sit there.
   d) Randomly select a seat position (right window, right center, right aisle, etc.) and survey all the passengers sitting in those seats.

STEP-BY-STEP EXAMPLE Sampling

The assignment says, "Conduct your own sample survey to find out how many hours per week students at your school spend watching TV during the school year." Let's see how we might do this step by step. (Remember, though—actually collecting the data from your sample can be difficult and time consuming.)

Question: How would you design this survey?

THINK

Plan State what you want to know.	I wanted to design a study to find out how many hours of TV students at my school watch.
Population and Parameter Identify the W's of the study. The *Why* determines the population and the associated sampling frame. The *What* identifies the parameter of interest and the variables measured. The *Who* is the sample we actually draw. The *How, When*, and *Where* are given by the sampling plan.	The population studied was students at our school. I obtained a list of all students currently enrolled and used it as the sampling frame. The parameter of interest was the number of TV hours watched per week during the school year, which I attempted to measure by asking students how much TV they watched during the previous week.

Often, thinking about the *Why* will help us see whether the sampling frame and plan are adequate to learn about the population.

Sampling Plan Specify the sampling method and the sample size, n. Specify how the sample was actually drawn. What is the sampling frame? How was the randomization performed?

A good description should be complete enough to allow someone to replicate the procedure, drawing another sample from the same population in the same manner.

I decided against stratifying by class or sex because I didn't think TV watching would differ much between males and females or across classes. I selected a simple random sample of students from the list. I obtained an alphabetical list of students, assigned each a random digit between 0 and 9, and then selected all students who were assigned a "4." This method generated a sample of 212 students from the population of 2133 students.

SHOW

Sampling Practice Specify *When*, *Where*, and *How* the sampling was performed. Specify any other details of your survey, such as how respondents were contacted, what incentives were offered to encourage them to respond, how nonrespondents were treated, and so on.

The survey was taken over the period Oct. 15 to Oct. 25. Surveys were sent to selected students by e-mail, with the request that they respond by e-mail as well. Students who could not be reached by e-mail were handed the survey in person.

TELL

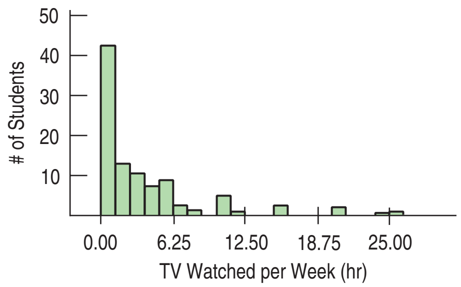
Summary and Conclusion This report should include a discussion of all the elements. In addition, it's good practice to discuss any special circumstances. Professional polling organizations report the *When* of their samples but will also note, for example, any important news that might have changed respondents' opinions during the sampling process. In this survey, perhaps, a major news story or sporting event might change students' TV viewing behavior.

The question you ask also matters. It's better to be specific ("How many hours did you watch TV last week?") than to ask a general question ("How many hours of TV do you usually watch in a week?").

During the period Oct. 15 to Oct. 25, 212 students were randomly selected, using a simple random sample from a list of all students currently enrolled. The survey they received asked the following question: "How many hours did you spend watching television last week?"

Of the 212 students surveyed, 110 responded. It's possible that the nonrespondents differ in the number of TV hours watched from those who responded, but I was unable to follow up on them due to limited time and funds. The 110 respondents reported an average 3.62 hours of TV watching per week. The median was only 2 hours per week. A histogram of the data shows that the distribution is highly right-skewed, indicating that the median might be a more appropriate summary of the typical TV watching of the students.

The report should show a display of the data, provide and interpret the statistics from the sample, and state the conclusions that you reached about the population.



Most of the students (90%) watch between 0 and 10 hours per week, while 30% reported watching less than 1 hour per week. A few watch much more. About 3% reported watching more than 20 hours per week.

Defining the “Who”: You Can’t Always Get What You Want

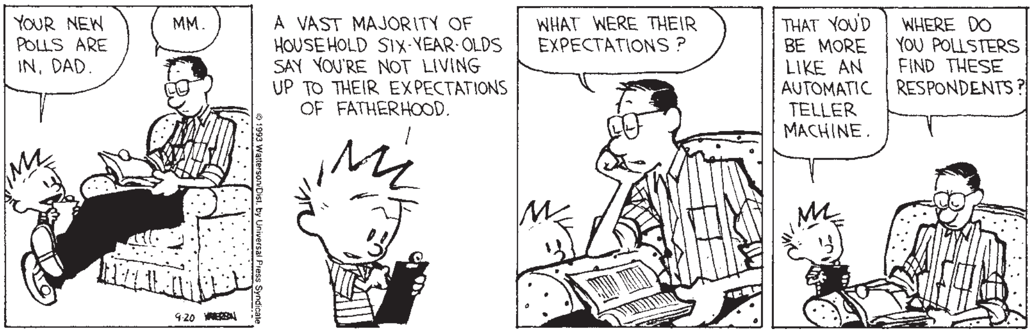
Before you start a survey, think first about the population you want to study. You may find that it’s not the well-defined group you thought it was. Who, exactly, is a student, for example? Even if the population seems well defined, it may not be a practical group from which to draw a sample. For example, election polls want to sample from all those who will vote in the next election—a population that is impossible to identify before Election Day.

The population is determined by the *Why* of the study. Unfortunately, the sample is just those we can reach to obtain responses—the *Who* of the study. This difference could undermine even a well-designed study.

Next, you must specify the sampling frame. (Do you have a list of students to sample from? How about a list of registered voters?) Usually, the sampling frame is not the group you *really* want to know about. (All those registered to vote are not equally likely to show up.) The sampling frame limits what your survey can find out.

Then there’s your target sample. These are the individuals for whom you *intend* to measure responses. You’re not likely to get responses from all of them. (“I know it’s dinnertime, but I’m sure you wouldn’t mind answering a few questions. It’ll only take 20 minutes or so. Oh, you’re busy?”) Nonresponse is a problem in many surveys.

Finally, there’s your sample—the actual respondents. These are the individuals about whom you *do* get data and can draw conclusions. Unfortunately, they might not be representative of the sampling frame or the population.



YOUR NEW POLLS ARE IN, DAD.
MM.
© 1993 Watterson/Dist. by Universal Press Syndicate
9-20 WATTERSON
A VAST MAJORITY OF HOUSEHOLD SIX-YEAR-OLDS SAY YOU'RE NOT LIVING UP TO THEIR EXPECTATIONS OF FATHERHOOD.
WHAT WERE THEIR EXPECTATIONS?
THAT YOU'D BE MORE LIKE AN AUTOMATIC TELLER MACHINE.
WHERE DO YOU POLLSTERS FIND THESE RESPONDENTS?

CALVIN AND HOBBES © 1993 Watterson. Reprinted with permission of Universal Press Syndicate. All rights reserved.

At each step, the group we can study may be constrained further. The *Who* keeps changing, and each constraint can introduce biases. A careful study should address the question of how well each group matches the population of interest. One of the main benefits of simple random sampling is that it never loses its sense of who's *Who*. The *Who* in an SRS is the population of interest from which we've drawn a representative sample. That's not always true for other kinds of samples.

The Valid Survey

It isn't sufficient to just draw a sample and start asking questions. We'll want our survey to be *valid.* A valid survey yields the information we are seeking about the population we are interested in. Before setting out to survey, ask yourself:

- What do I want to know?
- Am I asking the right respondents?
- Am I asking the right questions?
- What would I do with the answers if I had them; would they address the things I want to know?

These questions may sound obvious, but there are a number of pitfalls to avoid.

Know what you want to know. Before considering a survey, understand what you hope to learn and about whom you hope to learn it. Far too often, people decide to perform a survey without any clear idea of what they hope to learn.

Use the right frame. A valid survey obtains responses from the appropriate respondents. Be sure you have a suitable *sampling frame.* Have you identified the population of interest and sampled from it appropriately? A company might survey customers who returned warranty registration cards, a readily available sampling frame. But if the company wants to know how to make their product more attractive, the most important population is the customers who rejected their product in favor of one from a competitor.

Tune your instrument. It is often tempting to ask questions you don't really need, but beware—longer questionnaires yield fewer responses and thus a greater chance of nonresponse bias.

Ask specific rather than general questions. People are not very good at estimating their typical behavior, so it is better to ask "How many hours did you sleep last night?" than "How much do you usually sleep?" Sure, some responses will include some unusual events (My dog was sick; I was up all night.), but overall you'll get better data.

Ask for quantitative results when possible. "How many magazines did you read last week?" is better than "How much do you read: A lot, A moderate amount, A little, or None at all?"

Be careful in phrasing questions. A respondent may not understand the question—or may understand the question differently than the researcher intended it. ("Does anyone in your family belong to a union?" Do you mean just me, my spouse, and my children? Or does "family" include my father, my siblings, and my second cousin once removed? What about my grandfather, who is staying with us? I think he once belonged to the Autoworkers Union.) Respondents are unlikely (or may not have the opportunity) to ask for clarification. A question like "Do you approve of the recent actions of the Secretary of Labor?" is likely not to measure what you want if many respondents don't know who the Secretary of Labor is or what actions he or she recently made.

Respondents may even lie or shade their responses if they feel embarrassed by the question ("Did you have too much to drink last night?"), are intimidated or insulted by the question ("Could you understand our new *Instructions for Dummies* manual, or was it too difficult for you?"), or if they want to avoid offending the interviewer ("Would you hire a man with a tattoo?" asked by a tattooed interviewer). Also, be careful to avoid phrases that have double or regional meanings. "How often do you go to town?" might be interpreted differently by different people and cultures.

Even subtle differences in phrasing can make a difference. In January 2006, the *New York Times* asked half of the 1229 U.S. adults in their sample the following question:

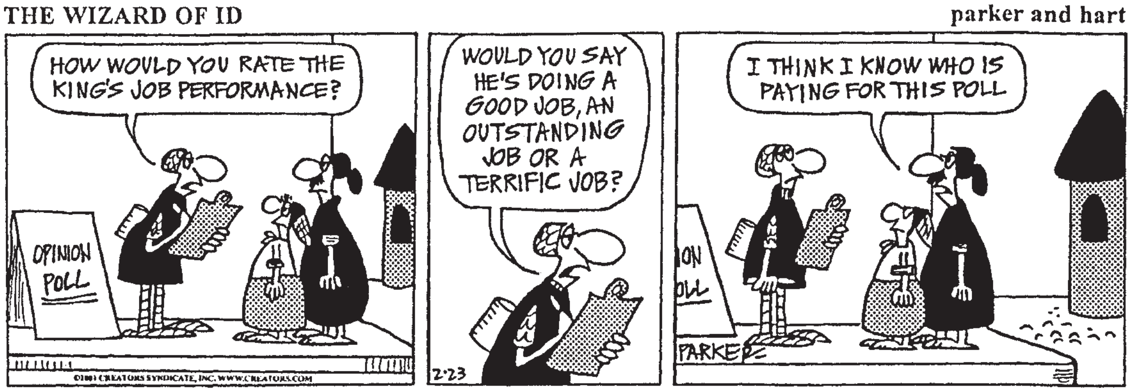
> *After 9/11, President Bush authorized government wiretaps on some phone calls in the U.S. without getting court warrants, saying this was necessary to reduce the threat of terrorism. Do you approve or disapprove of this?*

They found that 53% of respondents approved. But when they asked the other half of their sample a question with only slightly different phrasing,

> *After 9/11, George W. Bush authorized government wiretaps on some phone calls in the U.S. without getting court warrants. Do you approve or disapprove of this?*

only 46% approved.

Be careful in phrasing answers. It's often a good idea to offer choices rather than inviting a free response. Open-ended answers can be difficult to analyze. "How did you like the movie?" may start an interesting debate, but it may be better to give a range of possible responses. Be sure to phrase them in a neutral way. When asking "Do you support higher school taxes?" positive responses could be worded "Yes," "Yes, it is important for our children," or "Yes, our future depends on it." But those are not equivalent answers.


THE WIZARD OF ID
parker and hart
HOW WOULD YOU RATE THE KING'S JOB PERFORMANCE?
OPINION POLL
©2001 CREATORS SYNDICATE, INC. WWW.CREATORS.COM
WOULD YOU SAY HE'S DOING A GOOD JOB, AN OUTSTANDING JOB OR A TERRIFIC JOB?
2·23
I THINK I KNOW WHO IS PAYING FOR THIS POLL
PARKER

The best way to protect a survey from such unanticipated measurement errors is to perform a pilot survey. A **pilot** is a trial run of the survey you eventually plan to give to a larger group, using a draft of your survey questions administered to a small sample drawn from the same sampling frame you intend to use. By analyzing the results from this smaller survey, you can often discover ways to improve your instrument.

WHAT CAN GO WRONG?—OR, HOW TO SAMPLE BADLY

Bad sample designs yield worthless data. Many of the most convenient forms of sampling can be seriously biased. And there is no way to correct for the bias from a bad sample. So it's wise to pay attention to sample design—and to beware of reports based on poor samples.

SAMPLE BADLY WITH VOLUNTEERS

One of the most common dangerous sampling methods is a voluntary response sample. In a **voluntary response sample,** a large group of individuals is invited to respond, and all who do respond are counted. This method is used by call-in shows, 900 numbers, Internet polls, and letters written to members of Congress. Voluntary response samples are almost always biased, and so conclusions drawn from them are almost always wrong.

It's often hard to define the sampling frame of a voluntary response study. Practically, the frames are groups such as Internet users who frequent a particular Web site or those who happen to be watching a particular TV show at the moment. But those sampling frames don't correspond to interesting populations.

Even within the sampling frame, voluntary response samples are often biased toward those with strong opinions or those who are strongly motivated. People with very negative opinions tend to respond more often than those with equally strong positive opinions. The sample is not representative, even though every individual in the population may have been offered the chance to respond. The resulting **voluntary response bias** invalidates the survey.

If you had it to do over again, would you have children? Ann Landers, the advice columnist, asked parents this question. The overwhelming majority—70% of the more than 10,000 people who wrote in—said no, kids weren't worth it. A more carefully designed survey later showed that about 90% of parents actually are happy with their decision to have children. What accounts for the striking difference in these two results? What parents do you think are most likely to respond to the original question?

***Activity:* Sources of Sampling Bias.** Here's a narrated exploration of sampling bias.

FOR EXAMPLE Bias in sampling

Recap: You're trying to find out what freshmen think of the food served on campus, and have thought of a variety of sampling methods, all time consuming. A friend suggests that you set up a "Tell Us What You Think" Web site and invite freshmen to visit the site to complete a questionnaire.

Question: What's wrong with this idea?

Letting each freshman decide whether to participate makes this a voluntary response survey. Students who were dissatisfied might be more likely to go to the Web site to record their complaints, and this could give me a biased view of the opinions of all freshmen.

SAMPLE BADLY, BUT CONVENIENTLY

Another sampling method that doesn't work is convenience sampling. As the name suggests, in **convenience sampling** we simply include the individuals who are convenient for us to sample. Unfortunately, this group may not be representative of the population. A recent survey of 437 potential home buyers in Orange County, California, found, among other things, that

Do you use the Internet?
Click here ○ for yes
Click here ○ for no

> *All but 2 percent of the buyers have at least one computer at home, and 62 percent have two or more. Of those with a computer, 99 percent are connected to the Internet (Jennifer Hieger, "Portrait of Homebuyer Household: 2 Kids and a PC,"* Orange County Register, *27 July 2001).*

Internet convenience surveys are worthless. As voluntary response surveys, they have no well-defined sampling frame (all those who use the Internet and visit their site?) and thus report no useful information. Do not believe them.

Later in the article, we learn that the survey was conducted via the Internet! That was a convenient way to collect data and surely easier than drawing a simple random sample, but perhaps home builders shouldn't conclude from this study that *every* family has a computer and an Internet connection.

Many surveys conducted at shopping malls suffer from the same problem. People in shopping malls are not necessarily representative of the population of interest. Mall shoppers tend to be more affluent and include a larger percentage of teenagers and retirees than the population at large. To make matters worse, survey interviewers tend to select individuals who look "safe," or easy to interview.

FOR EXAMPLE Bias in sampling

Recap: To try to gauge freshman opinion about the food served on campus, Food Services suggests that you just stand outside a school cafeteria at lunchtime and stop people to ask them questions.

Questions: What's wrong with this sampling strategy?

This would be a convenience sample, and it's likely to be biased. I would miss people who use the cafeteria for dinner, but not for lunch, and I'd never hear from anyone who hates the food so much that they have stopped coming to the school cafeterias.

SAMPLE FROM A BAD SAMPLING FRAME

An SRS from an incomplete sampling frame introduces bias because the individuals included may differ from the ones not in the frame. People in prison, homeless people, students, and long-term travelers are all likely to be missed. In telephone surveys, people who have only cell phones or who use VOIP Internet phones are often missing from the sampling frame.

UNDERCOVERAGE

Many survey designs suffer from **undercoverage,** in which some portion of the population is not sampled at all or has a smaller representation in the sample than it has in the population. Undercoverage can arise for a number of reasons, but it's always a potential source of bias.

Telephone surveys are usually conducted when you are likely to be home, interrupting your dinner. If you eat out often, you may be less likely to be surveyed, a possible source of undercoverage.

WHAT Else CAN GO WRONG?

- **Watch out for nonrespondents.** A common and serious potential source of bias for most surveys is **nonresponse bias.** No survey succeeds in getting responses from everyone. The problem is that those who don't respond may differ from those who do. And they may differ on just the variables we care about. The lack of response will

(*continued*)

bias the results. Rather than sending out a large number of surveys for which the response rate will be low, it is often better to design a smaller randomized survey for which you have the resources to ensure a high response rate. One of the problems with nonresponse bias is that it's usually impossible to tell what the nonrespondents might have said.

Remember the *Literary Digest* Survey? It turns out that they were wrong on two counts. First, their list of 10 million people was not representative. There was a selection bias in their sampling frame. There was also a nonresponse bias. We know this because the *Digest* also surveyed a *systematic* sample in Chicago, sending the same question used in the larger survey to every third registered voter. They *still* got a result in favor of Landon, even though Chicago voted overwhelmingly for Roosevelt in the election. This suggests that the Roosevelt supporters were less likely to respond to the *Digest* survey. There's a modern version of this problem: It's been suggested that those who screen their calls with caller ID or an answering machine, and so might not talk to a pollster, may differ in wealth or political views from those who just answer the phone.

A S *Video:* **Biased Question Wording.** Watch a hapless interviewer make every mistake in the book.

A Short Survey

Given the fact that those who understand Statistics are smarter and better looking than those who don't, don't you think it is important to take a course in Statistics?

- **Work hard to avoid influencing responses.** **Response bias**[10] refers to anything in the survey design that influences the responses. Response biases include the tendency of respondents to tailor their responses to try to please the interviewer, the natural unwillingness of respondents to reveal personal facts or admit to illegal or unapproved behavior and the ways in which the wording of the questions can influence responses.

HOW TO THINK ABOUT BIASES

- **Look for biases in any survey you encounter.** If you design one of your own, ask someone else to help look for biases that may not be obvious to you. And do this *before* you collect your data. There's no way to recover from a biased sampling method or a survey that asks biased questions. Sorry, it just can't be done.

  A bigger sample size for a biased study just gives you a bigger useless study. A really big sample gives you a really big useless study. (Think of the 2.4 million *Literary Digest* responses.)

A S *Activity:* **Can a Large Sample Protect Against Bias?** Explore how we can learn about the population from large or repeated samples.

- **Spend your time and resources reducing biases.** No other use of resources is as worthwhile as reducing the biases.
- **If you can, pilot-test your survey.** Administer the survey in the exact form that you intend to use it to a small sample drawn from the population you intend to sample. Look for misunderstandings, misinterpretation, confusion, or other possible biases. Then refine your survey instrument.
- **Always report your sampling methods in detail.** Others may be able to detect biases where you did not expect to find them.

A researcher distributed a survey to an organization before some economizing changes were made. She asked how people felt about a proposed cutback in secretarial and administrative support on a seven-point scale from Very Happy to Very Unhappy.

But virtually all respondents were very unhappy about the cutbacks, so the results weren't particularly useful. If she had pretested the question, she might have chosen a scale that ran from Unhappy to Outraged.

---

[10] Response bias is not the opposite of nonresponse bias. (We don't make these terms up; we just try to explain them.)

CONNECTIONS

With this chapter, we take our first formal steps to relate our sample data to a larger population. Some of these ideas have been lurking in the background as we sought patterns and summaries for data. Even when we only worked with the data at hand, we often thought about implications for a larger population of individuals.

Notice the ongoing central importance of models. We've seen models in several ways in previous chapters. Here we recognize the value of a model for a population. The parameters of such a model are values we will often want to estimate using statistics such as those we've been calculating. The connections to summary statistics for center, spread, correlation, and slope are obvious.

We now have a specific application for random numbers. The idea of applying randomness deliberately showed up in Chapter 11 for simulation. Now we need randomization to get good-quality data from the real world.

WHAT HAVE WE LEARNED?

We've learned that a representative sample can offer us important insights about populations. It's the size of the sample—and not its fraction of the larger population—that determines the precision of the statistics it yields.

We've learned several ways to draw samples, all based on the power of randomness to make them representative of the population of interest:

- A Simple Random Sample (SRS) is our standard. Every possible group of n individuals has an equal chance of being our sample. That's what makes it *simple*.
- Stratified samples can reduce sampling variability by identifying homogeneous subgroups and then randomly sampling within each.
- Cluster samples randomly select among heterogeneous subgroups that each resemble the population at large, making our sampling tasks more manageable.
- Systematic samples can work in some situations and are often the least expensive method of sampling. But we still want to start them randomly.
- Multistage samples combine several random sampling methods.

We've learned that bias can destroy our ability to gain insights from our sample:

- Nonresponse bias can arise when sampled individuals will not or cannot respond.
- Response bias arises when respondents' answers might be affected by external influences, such as question wording or interviewer behavior.

We've learned that bias can also arise from poor sampling methods:

- Voluntary response samples are almost always biased and should be avoided and distrusted.
- Convenience samples are likely to be flawed for similar reasons.
- Even with a reasonable design, sample frames may not be representative. Undercoverage occurs when individuals from a subgroup of the population are selected less often than they should be.

Finally, we've learned to look for biases in any survey we find and to be sure to report our methods whenever we perform a survey so that others can evaluate the fairness and accuracy of our results.

Terms

Population	268. The entire group of individuals or instances about whom we hope to learn.
Sample	268. A (representative) subset of a population, examined in hope of learning about the population.

Sample survey	269. A study that asks questions of a sample drawn from some population in the hope of learning something about the entire population. Polls taken to assess voter preferences are common sample surveys.
Bias	269. Any systematic failure of a sampling method to represent its population is bias. Biased sampling methods tend to over- or underestimate parameters. It is almost impossible to recover from bias, so efforts to avoid it are well spent. Common errors include ▸ relying on voluntary response. ▸ undercoverage of the population. ▸ nonresponse bias. ▸ response bias.
Randomization	270. The best defense against bias is randomization, in which each individual is given a fair, random chance of selection.
Sample size	271. The number of individuals in a sample. The sample size determines how well the sample represents the population, not the fraction of the population sampled.
Census	271. A sample that consists of the entire population is called a census.
Population parameter	272. A numerically valued attribute of a model for a population. We rarely expect to know the true value of a population parameter, but we do hope to estimate it from sampled data. For example, the mean income of all employed people in the country is a population parameter.
Statistic, sample statistic	272. Statistics are values calculated for sampled data. Those that correspond to, and thus estimate, a population parameter, are of particular interest. For example, the mean income of all employed people in a representative sample can provide a good estimate of the corresponding population parameter. The term “sample statistic” is sometimes used, usually to parallel the corresponding term “population parameter.”
Representative	273. A sample is said to be representative if the statistics computed from it accurately reflect the corresponding population parameters.
Simple random sample (SRS)	273. A simple random sample of sample size n is a sample in which each set of n elements in the population has an equal chance of selection.
Sampling frame	273. A list of individuals from whom the sample is drawn is called the sampling frame. Individuals who may be in the population of interest, but who are not in the sampling frame, cannot be included in any sample.
Sampling variability	274. The natural tendency of randomly drawn samples to differ, one from another. Sometimes, unfortunately, called *sampling error*, sampling variability is no error at all, but just the natural result of random sampling.
Stratified random sample	274. A sampling design in which the population is divided into several subpopulations, or **strata,** and random samples are then drawn from each stratum. If the strata are homogeneous, but are different from each other, a stratified sample may yield more consistent results than an SRS.
Cluster sample	275. A sampling design in which entire groups, or **clusters,** are chosen at random. Cluster sampling is usually selected as a matter of convenience, practicality, or cost. Each cluster should be representative of the population, so all the clusters should be heterogeneous and similar to each other.
Multistage sample	276. Sampling schemes that combine several sampling methods are called multistage samples. For example, a national polling service may stratify the country by geographical regions, select a random sample of cities from each region, and then interview a cluster of residents in each city.
Systematic sample	277. A sample drawn by selecting individuals systematically from a sampling frame. When there is no relationship between the order of the sampling frame and the variables of interest, a systematic sample can be representative.
Pilot	281. A small trial run of a survey to check whether questions are clear. A pilot study can reduce errors due to ambiguous questions.
Voluntary response bias	282. Bias introduced to a sample when individuals can choose on their own whether to participate in the sample. Samples based on voluntary response are always invalid and cannot be recovered, no matter how large the sample size.

Convenience sample	282. A convenience sample consists of the individuals who are conveniently available. Convenience samples often fail to be representative because every individual in the population is not equally convenient to sample.
Undercoverage	283. A sampling scheme that biases the sample in a way that gives a part of the population less representation than it has in the population suffers from undercoverage.
Nonresponse bias	283. Bias introduced when a large fraction of those sampled fails to respond. Those who do respond are likely to not represent the entire population. Voluntary response bias is a form of nonresponse bias, but nonresponse may occur for other reasons. For example, those who are at work during the day won't respond to a telephone survey conducted only during working hours.
Response bias	284. Anything in a survey design that influences responses falls under the heading of response bias. One typical response bias arises from the wording of questions, which may suggest a favored response. Voters, for example, are more likely to express support of "the president" than support of the particular person holding that office at the moment.

Skills

THINK

- Know the basic concepts and terminology of sampling (see the preceding list).
- Recognize population parameters in descriptions of populations and samples.
- Understand the value of randomization as a defense against bias.
- Understand the value of sampling to estimate population parameters from statistics calculated on representative samples drawn from the population.
- Understand that the size of the sample (not the fraction of the population) determines the precision of estimates.

SHOW

- Know how to draw a simple random sample from a master list of a population, using a computer or a table of random numbers.

TELL

- Know what to report about a sample as part of your account of a statistical analysis.
- Report possible sources of bias in sampling methods. Recognize voluntary response and nonresponse as sources of bias in a sample survey.

SAMPLING ON THE COMPUTER

Computer-generated pseudorandom numbers are usually good enough for drawing random samples. But there is little reason not to use the truly random values available on the Internet.

Here's a convenient way to draw an SRS of a specified size using a computer-based sampling frame. The sampling frame can be a list of names or of identification numbers arrayed, for example, as a column in a spreadsheet, statistics program, or database:

1. Generate random numbers of enough digits so that each exceeds the size of the sampling frame list by several digits. This makes duplication unlikely.
2. Assign the random numbers arbitrarily to individuals in the sampling frame list. For example, put them in an adjacent column.
3. Sort the list of random numbers, carrying along the sampling frame list.
4. Now the first *n* values in the sorted sampling frame column are an SRS of *n* values from the entire sampling frame.