

TRANSCRIPT:

This presentation is on the basic principles of sampling as they apply to research design.

SAMPLING BASICS

- What is a sample?
 - ❖ A sample is a subset of the population from which data are collected.
- Why use a sample?
 - ❖ It sometimes is not feasible to collect data from an entire population.
 - ❖ It sometimes is not possible to collect data from an entire population.

SAMPLING BASICS



TRANSCRIPT:

Before we talk about sampling, we should first define what a sample is. Put simply, a sample is a subset of the population from which data are collected. In other words, it is the group of people from whom we attempt to collect data for a research study.

It would always be better to collect data from an entire population, but this is often not feasible because resources may be limited. For example, it might be possible to identify all the students who graduate from public high schools in the U.S. during a certain year, but it is not likely that a researcher would have enough time or financial resources to make that sort of data collection possible.

Another reason a sample might be preferable over a population is that it is not possible to collect data from an entire population. For example, suppose we are interested in studying veterans between the ages of 60 and 75 who suffer from post traumatic stress disorder. It would not be possible to collect data from all individuals who fit into this group because many of them remain undiagnosed.

Let's consider the use of sampling in an education-related research example.

EXAMPLE SAMPLE (GA CCC)

- The GA DoE adopts the Common Core Curriculum for use in its schools. Two years later, they wish to know if teachers have adopted the new curriculum, and whether they are satisfied with it.
- Budget & time limitations make it impossible to collect data from all elementary teachers in the state.
- A subset of elementary teachers is selected to participate in the study ← a sample is needed

SAMPLING BASICS



TRANSCRIPT:

Suppose the Georgia Department of Education were to adopt the Common Core Curriculum for use in its schools, and require that all schools be teaching to Common Core standards in the fall of 2014. Two years later, the state DOE decides to do a study to find out whether or not all teachers have actually adopted the new curriculum within their classrooms, and how they feel about the transition and student progress.

As a starting point, the DoE might decide to start with a study of elementary-level teachers. As you can imagine, though, it would require a lot of time and budget resources to collect data from all the elementary school teachers in the state. Therefore, the state decides to conduct this study by collecting data from only some of the elementary teachers – that smaller group of teachers is the sample, and the total group of elementary teachers is the population from which it is drawn.

This is the basic idea behind sampling. We select a smaller from a large group or population in order to collect data for research studies. There are a lot of different types of sampling, but let's take a minute to talk about some key sampling terms before we get too far into this discussion.

KEY SAMPLING TERMS

- *Target Population*
 - ❖ Complete group of interest
- *Sampling Frame*
 - ❖ List of the sampling units present within the target population
- *Sample*
 - ❖ Subset of the (target) population
- *Participant Group*
 - ❖ Individuals who are selected as part of the sample and who choose to participate

SAMPLING BASICS



TRANSCRIPT:

The key sampling-related terms you need to be familiar with are target population, sample, sampling unit, and sampling frame.

The population is the complete group of interest, or the group about which you want to learn something. In the common core example I discussed a minute ago, the population of interest is all elementary-level school teachers in the state of Georgia.

A sampling frame is a list of the units that exist within the target population. Ideally, a sampling frame should include all the individuals who exist in the population, and each of those individuals should be listed just one time. In the common core curriculum example, the sampling frame is a list of all elementary school teachers at public schools in the state of Georgia.

The sample is a subset of the target population that is selected for participation in the research study because it is impossible or just not feasible to collect data from the entire population. The sample is a smaller group than the population, and includes only a portion of the individuals listed on the sampling frame.

The participant group, then, is those members of the sample who are invited to participate in the study, and actually do choose to participate. This is a very important distinction between the sample and the participant groups. The sample includes everyone you invite to participate, or all those you hope will participate when you begin the study. The participant group is just those people who agree to participate.

Let's look at a visual depiction of how these four concepts relate.

KEY SAMPLING TERMS

Target Population

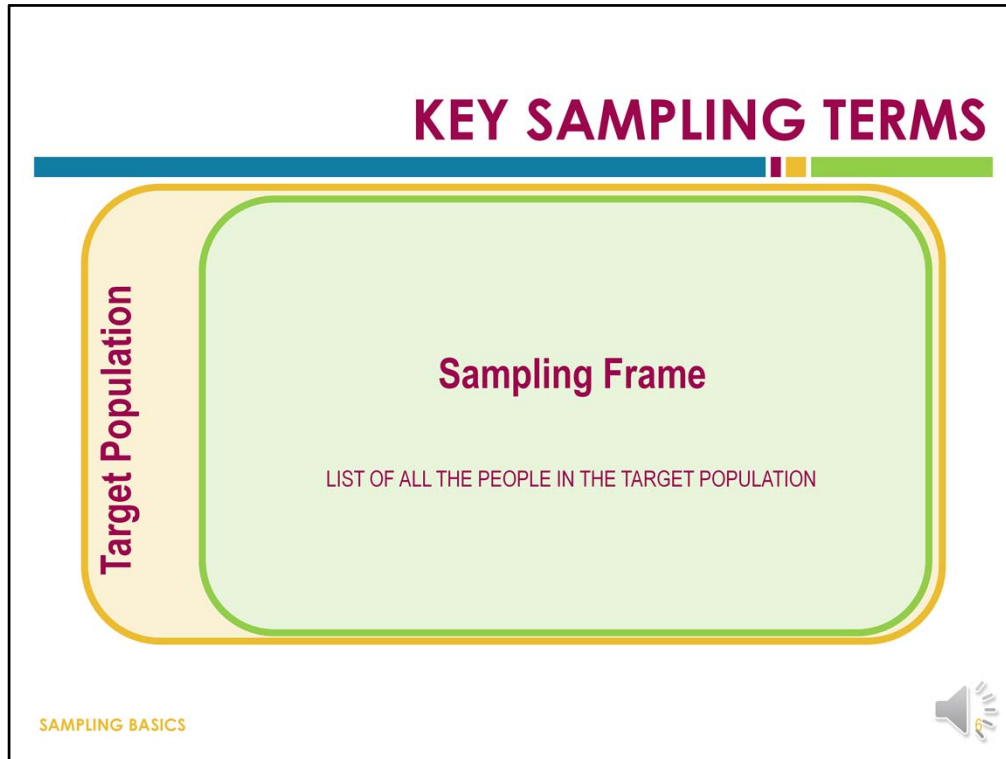
ALL THE PEOPLE YOU WANT
TO KNOW SOMETHING ABOUT

SAMPLING BASICS



TRANSCRIPT:

We start with the target population – all the people we want to know something about. From the target population, we create the sampling frame.

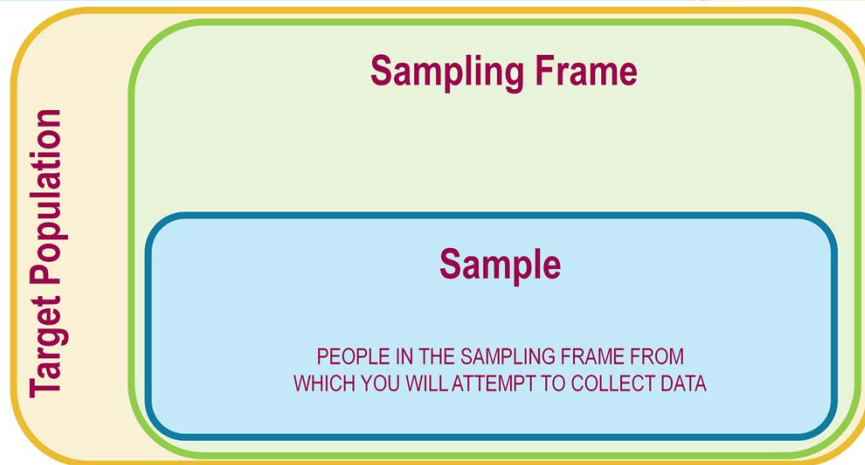


TRANSCRIPT:

In a perfect world, our sampling frame is exactly like the target population. In the real world, this is rarely the case. Consider the previous example of veterans who suffer from PTSD. In that example, there are likely some individuals who have not been diagnosed, and it is possible that some of those who have been diagnosed with PTSD were misdiagnosed.

Even in our common core example, it is possible that the sampling frame will not perfectly match the population because new teachers are hired and others leave, and it can take time for all those changes to be logged in all the relevant systems.

KEY SAMPLING TERMS



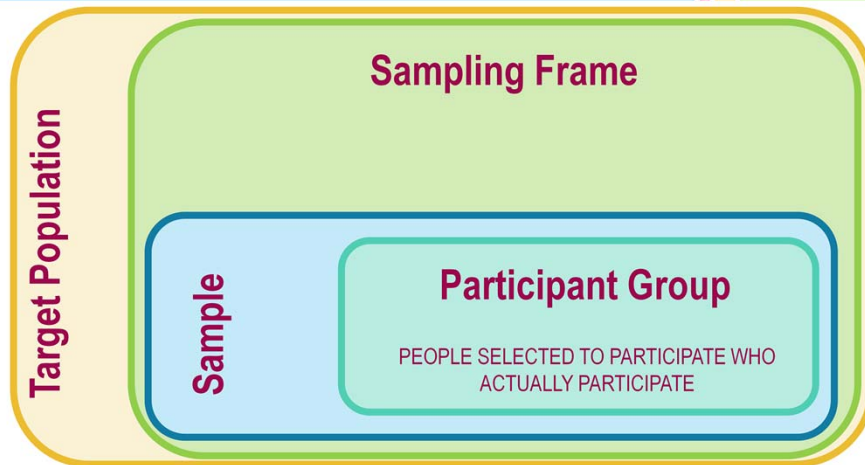
SAMPLING BASICS



TRANSCRIPT:

From the sampling frame, we select our sample. The sample will always include individuals who exist on the sampling frame, but will rarely include all the individuals on the frame. The sample is the group of people who we invite to participate in our research study.

KEY SAMPLING TERMS



SAMPLING BASICS



TRANSCRIPT:

The participant group, then, are the people we invite to be part of the study who actually agree to participate and then do participate. The participant group is sometimes referred to as respondents.

THE PERFECT SAMPLE

➤ *Magic Town* (1947)

- ❖ The people who make up a single town (Grandview) is an exact microcosm of the country
 - Demographics & Employment
 - Political Opinions & General Attitudes
- ❖ The people of Grandview are a perfect representation (representative sample) of the U.S. population

SAMPLING BASICS



TRANSCRIPT:

A perfect sample is one that is an exact representation of the target population. There is an old movie called *Magic Town* from the 1940s in which Jimmy Stewart played a pollster who identified a single town that represented the entire U.S. population perfectly. In this single town, called Grandview, every characteristic of the country's population was presented accurately. The residents' ages, education, race, type of employment, political attitudes, number of children, number of pets, EVERYTHING was exactly, perfectly representative of the U.S. population. This is the perfect sample. It is often unrealistic, but this is the type of relationship between sample and population that researchers want to have.

TYPES OF SAMPLES

- Probability (Representative)
 - ❖ A sample that is selected in such a way as to be representative of the population
- Nonprobability (Non-Representative)
 - ❖ A sample that is not selected in such a way as to be representative of the population

SAMPLING BASICS



TRANSCRIPT:

Many, many different sampling procedures have been developed, but I am only going to discuss those that are most commonly used in education research.

The first distinction between types of samples that should be made is the difference between a probability and non-probability sample. A probability sample is one that is selected in such a way as to be representative of the population.

A nonprobability sample is one that is not selected in a way that increases the extent to which the sample represents the population.

REPRESENTATIVE SAMPLES

➤ *Random*

- ❖ Sample in which each person in the population has an equal chance of being selected
- ❖ Example: A researcher would like to collect data from a random sample of K-12 teachers employed in traditional, public schools in GA
 - List of all K-12 teachers employed in traditional, public schools in the state of GA is obtained from the state DoE (sampling frame)
 - 100 teachers are randomly selected from the list

SAMPLING BASICS



TRANSCRIPT:

One of the most common types of probability or representative samples is the random sample. A random sample is a sampling procedure by which each individual in the population has an equal chance of being selected. For example, assume you are a researcher who would like to collect data from a random sample of K-12 teachers employed in traditional, public schools in GA. You would first develop a sampling frame to include all the names of K-12 teachers in the state. Then, you would randomly select a certain number of teachers from that list.

REPRESENTATIVE SAMPLES

➤ *Stratified*

- ❖ Sample that reflects the population based on specific researcher-selected criteria
- ❖ Example: A school counselor would like to know if it is sufficient to send English-only notices home with students
 - Researcher divides frame into two groups: English as home language, a non-English language as home language. 70% of students' parents report English is primary home language; 70% of sample selected from list of home English speakers.

SAMPLING BASICS



TRANSCRIPT:

Another type of probability or representative sample is a stratified sample. A stratified sample divides the sampling frame up by some characteristic or group of characteristics, and then samples from those divisions – often using random sampling at that point.

For example, suppose you are a school counselor who wants to know if it ok to send school-related information home with students in only English, or if it is necessary to have those messages translated into other languages. To make this decision, you might not care what the primary language spoken in the home is – but whether English can be understood well enough by all households to justify only sending out one version of each notification. To choose your sample, you might look at the sampling frame and discover that 70% of your students' parents report English is the primary language spoken in the home, and 30% report that English is not their primary home language. To create a stratified sample based on this characteristic of home language, you would decide how big you want your total sample to be, and then select 70% of your sample from the English as primary language group, and 30% of your sample from the primary language-other-than-English group.

REPRESENTATIVE SAMPLES

➤ *Census*

- ❖ Sample consisting of the entire population
- ❖ Example: A high school math teacher decides to use a new method of communication with parents to increase parent involvement. At the end of the year, the teacher would like to know if parents thought the new approach was effective.
 - The teacher attempts (and succeeds) to obtain parent responses from the parents of all students in her classes.

SAMPLING BASICS



TRANSCRIPT:

Another type of probability or representative sample is a census sample. A census sample is what happens when we attempt to collect data from every individual listed on the sampling frame, when we assume that the sampling frame perfectly represents the population.

For example, suppose you are a high school math teacher who decides to try out a new method of communicating with your students' parents in an attempt to get parents more involved in their students' education. At the end of the year, you feel pretty good about the new communication method, but you want to know if the parents found it useful as well. In this instance, you could conduct a census sample by simply getting every student to have their parent or parents complete a short survey. This is a census sample because you will have collected data from all the parents.

NON-REPRESENTATIVE SAMPLES

➤ Convenience

- ❖ Sample consisting of units within the population that require minimal effort to reach
- ❖ Example: A researcher would like to investigate the prevalence of suicidal tendencies among college students.
 - The researcher recruits students enrolled in introductory social science (e.g., psychology, sociology) courses.

SAMPLING BASICS



TRANSCRIPT:

There are also many different types of nonprobability or non-representative samples. The most common of these is the convenience sample. This is probably a concept you are familiar with. A research faculty at a university is genuinely interested in the prevalence of suicidal tendencies among college students. Instead of drawing a random sample of students, the researcher gets the faculty in the psychology department to offer extra credit to their students who participate in the study.

There is no way to know exactly how representative the sample is of the population in this case, but it is an easy sample to get. If the study is on a topic that you can assume is pretty well represented throughout the population, then probability sampling may not be a particular concern.

NON-REPRESENTATIVE SAMPLES

➤ *Purposive*

- ❖ Sample selected based on some knowledge of the individuals & the purpose of the research
- ❖ Example: A researcher is investigating attitudes toward online classes among persons who have completed at least one online course.
 - The researcher identifies a group of currently enrolled in online courses, and recruits all individuals in that group who have completed at least one online course.

SAMPLING BASICS



TRANSCRIPT:

Another type of nonprobability or non-representative samples is the purposive or purposeful sample. A purposeful sample is one selected based on some knowledge of the individuals as well as the purpose of the research study. For example, if you are a researcher interested in the attitudes of individuals who have completed online classes toward online education, you might choose to only recruit participants who are currently enrolled in online courses.

NON-REPRESENTATIVE SAMPLES

➤ Quota

- ❖ Sample in which potential respondents are divided into subgroups; data is collected from a minimum number of individuals in each subgroup
- ❖ Example: A district's SPED coordinator would like to know if parents are satisfied with services.
 - The researcher decides to interview the parents of 5 autistic children regardless of the prevalence of autism within the population of SPED students in the district.

SAMPLING BASICS



TRANSCRIPT:

The last type of nonprobability or non-representative sample we are going to discuss is the quota sample. A quota sample exists when potential respondents are divided into subgroups and then individuals are invited to participate until a certain number has been reached for each group.

For example, if you are a school district's Special Education coordinator and you are interested in whether parents are satisfied with the school services their children receive, it might make sense to do a quota sample. You might identify the primary disability groups that your district serves, such as children with down's syndrome, autistic children, dyslexic children, etc., and then collect data until you have responses from the parents of five children for each diagnosis.

The primary disadvantage to this method is that the sample may not represent the population. For example, if you have two children in the district who have turret's syndrome and 20 students with autism, but you collect data from the parents of two children with turret's and two children with autism, you have collected data from 100% of the turret's group, but only a small percentage of the autism group. Therefore, the quota sample is not likely to represent the overall group very well at all.

BOTTOM LINE OF SAMPLING

- *Each method has its place*
- *Representative/probability is more generalizable*
- *Generalizability is not always a priority*
- *Sampling method should fit the study*

SAMPLING BASICS



TRANSCRIPT:

To wrap up, the take away message is that there are many different sampling approaches and each one has at least one situation in which it is most appropriate than any other approach.

The primary advantage of selecting a representative or probability sample over a non-probability sample is that the probability sample will typically allow results to be generalized from the sample to the population more easily. However, we are not always concerned with being able to generalize our results. Sometimes, we are just interested in our sample.

Ultimately, the sampling method that is chosen should be whatever method is most appropriate for the research question and the individual study.