
Section 2: Ten Tools for Applying Sociology

CHAPTER 2.6: DATA COLLECTION METHODS

QUICK START: In this chapter, you will learn

- The basics of data collection methods.
- To know when to use quantitative and/or qualitative methods in examining social situations.
- To know when to use probability and non-probability sampling.
- To know how to develop a research design to answer a question.

TERMS

Data Collection Methods	Ways to collect data on individual and group characteristics, attitudes, beliefs, values, and behavior.
Quantitative Method	Data collection techniques that convert individual or group characteristics, attitudes, values, beliefs or behaviors into numbers.
Qualitative Method	Data collection techniques that produce a written, audio or video record of people's words or behavior.
Variable	A measure of a concept the value of which can change across people or cases.
Triangulate	Using multiple data collection methods to understand a social issue or problem from multiple angles and perspectives.
Experiments	Exposing one group to a social force and another group to something neutral and then comparing their reactions.

Chapter 2.6: Data Collection Methods

Surveys	Asking standardized questions of a group of people, either in person, or by phone or mail.
Participant Observation	Observing people's behavior in a particular setting on a particular topic as a participant in the setting.
Non-Participant Observation	Observing people's behavior in a particular setting on a particular topic as a non-participant in the setting.
In-Depth Interviews	Asking people questions that require lengthy answers and which they may answer through discussion.
Focus Groups	In-depth interviews with a group of people, in which the interviewees interactively discuss a topic.
Existing or Secondary Data	Using numerical data previously collected by individuals, groups or organizations, such as medical or criminal justice data.
Observation	Observing people's behavior in a particular setting on a particular topic, either as a participant or non-participant in the activity.
Field Notes	A researcher's written notes about observations during or after an interview or observation in a qualitative study.
Historical or Content Analysis	Using written, audio or video documents that were originally recorded for other purposes by individuals, groups or organizations.
Evaluation Research	Collecting data to determine whether an intervention or program works.
Ethnography or Case Study	Studying every aspect of a group, organization, or culture, through participation in and observation of it.
Action Research	Participating in creating social change via a research project. Uses either or both quantitative and qualitative data collection methods.

Chapter 2.6: Data Collection Methods

Sampling	Selecting a subset of a group rather than the whole group to make your observations.
Population	The group of people to whom a study is relevant and to which the study seeks to generalize.
Generalize	To broadly apply the results of research to a larger population.
Sample Size	The number of people in a sample. This may be determined by a formula for probability sampling.
Probability Samples	You can draw inferences from them to the whole group.
Non-Probability Samples	You cannot draw inferences from them to the whole group.
Simple Random Sample	A sample derived by randomly selecting people from off a list.
Systematic Random Sample	A sample derived by randomly selecting every nth person from off a list.
Proportionate Stratified Random	A sample derived by dividing a population into homogeneous groups called strata and then drawing a simple random sample from within each strata in proportion to which the strata exist in the population.
Disproportionate Stratified Random	A sample derived by dividing a population into homogeneous groups called strata and then drawing a simple random sample from within each strata in disproportion to which the strata exist in the population
Convenience Sample	A sample derived by selecting people because they are easily accessible.
Quota Sample	A convenience sample that attempts to select a percentage of people with some characteristic(s), such as sex, race or age, that matches the percentage that these characteristics exist in the population.

Chapter 2.6: Data Collection Methods

Snowball Sample

A sample derived by finding one person in the target population and then asking them to give you access to other people that they know in the population, and then repeated as necessary.

Theoretical Sample

A sample derived as specified by a theory or research question.

TOPICS

Applied sociologists use **data collection methods** that enable them to best answer the question at hand. Sometimes that means using a **quantitative** method, sometimes a **qualitative** method. Applied sociologists often combine or **triangulate** various methods.

Quantitative approaches are the ones you most likely think of when you think of science. A scientist looks at a problem, defining it in terms of concepts. Then the scientist measures the concepts by “reducing” the concept to measurable, observable parts called variables. After measuring the parts, the scientist adds them back up again to describe or understand the original problem. The overall idea is that any problem or topic of study can be broken down into all of its parts, and that the sum of the parts equals the whole problem. For example, let’s say you are an applied sociologist who is hired by a college to determine the professional success of their alumni. The college wants to use this information in their advertising to prospective students. You think about what indicates success and write a survey with questions for each indicator. Some of the questions on your survey might be: whether alumni are employed, how long it took for them to obtain a job post-graduation, job title, income, job satisfaction, and whether alumni attended graduate school or received any graduate degrees. These are **variables**.

You’re familiar with the quantitative approach to studying problems because the physical sciences use this approach. To name a few, biologists, chemists and physicists all apply this approach. Likewise, quantitative social scientists follow this same process too. You’ll see this approach in sociology, psychology, economics and political science. If you plan to research a problem using this approach, you must use a method that breaks a problem down. **Surveys** and **experiments** are examples of the quantitative approach to data collection. We describe these methods in the table below.

Qualitative approaches are different. When you take this approach you start by assuming that the whole problem has a unique quality in its totality. That is, the problem is greater than the sum of its parts. From a qualitative perspective, if you reduce the problem to its parts (as with the quantitative approach), you change the problem’s composition. So if you plan to research a problem using a qualitative approach you must use a method that doesn’t reduce the

Chapter 2.6: Data Collection Methods

phenomenon. **Participant observation**, **non-participant observation** and **in-depth interviews** are examples of qualitative data collection methods. Using these techniques, the researcher observes the problem in its entirety. We describe these methods in the table below.

For example, let's say you are an applied sociologist who is hired by a large real estate company who is merging with another company. The company executives want to know the best way to merge the two groups of employees together, so as to reduce personnel problems, maintain morale, and enhance productivity. To provide answers, you need to understand the organizational culture of both organizations. To truly understand the culture, and why and how people interact as they do in it, you must spend time in the culture and come to know the people in it. Using a method like a survey or an experiment would force you to focus in on certain aspects of the organizational culture, for which you would then derive a handful of measures that represent the whole organizational culture. You may argue that using such an approach would produce a biased picture of the culture. Instead, you opt to observe the employees in their jobs, and to interview them in-depth about their roles in the organization.

Which technique is better? As you can see there are strengths and weaknesses to both quantitative and qualitative methods. Academic sociologists spend a fair amount of time arguing over which of these is best. That's fine; it's their job. But, applied sociologists take a little different approach. Since solving a problem is central, the question is "Which methodological tool(s) provides the best fit to this problem?" Applied problems rarely exist in a perfect scientific vacuum. Hence, an applied sociologist may need to apply multiple methods to get the best understanding of the problem's solution. This might mean both quantitative and qualitative approaches to triangulate the results. For example it would not be unusual for a sociologist doing market research to do a sample survey to provide a general understanding of customer satisfaction, then follow-up the survey with a series of qualitative **focus groups** to enhance the results.

TOOLS

Below is a table that summarizes the main data collection methods that sociologists use.

Chapter 2.6: Data Collection Methods

Quantitative	Qualitative	Triangulation
Experiments: Exposing a group to varying degrees of a social force, then recording and comparing their reactions to a group that was not exposed. The numerically coded reactions serves as the data.	Observation: Observing people's behavior in a particular setting on a particular topic, either as a participant or non-participant observer during the activity. Written field notes about the observations serve as the data.	Evaluation Research: Collecting data to determine whether an intervention or program works. Uses either or both quantitative and qualitative data collection methods.
Surveys: Asking standardized questions of a group of people, either in person, or by phone or mail. Data are the numerically coded survey responses.	In-Depth Interviews: Asking people questions that require lengthy answers and which they may answer through discussion. A focus group is a kind of in-depth interview with multiple people. Traditionally interviews are audiotaped and, later, typed into a written transcript of the interview. They can also be videotaped, and transcribed or catalogued. Field notes and transcriptions of the interviews serve as data.	Ethnography or Case Study: Studying every aspect of a group, organization, or culture, through participation in and observation of it. Can use quantitative or qualitative data collection methods, but usually involves qualitative methods.
Existing or Secondary Data: Using numerical data previously collected by individuals, groups or organizations, such as medical or criminal justice data.	Historical or Content Analysis: Using written, audio or video documents that were originally recorded for other purposes by individuals, groups, and organizations.	Action Research: Participating in creating social change via a research project. Uses either or both quantitative and qualitative data collection methods.

Before we go on, let's talk a little bit more about triangulation. Social behavior is complicated and hard to study. Unlike in the physical sciences, we can't assume that the research cases we are studying will stay in one place. And unlike the physical sciences, we can't contain our research cases, manipulate them and observe how they react, like a scientist studying cells or animals might do. As a result, the research that sociologists do is often more complicated than in the

Chapter 2.6: Data Collection Methods

physical sciences. We must spend a great deal of time designing studies to make sure that the data collection methods we use will provide an accurate window on social behavior. Using one data collection method often does not provide a big enough window. So, applied sociologists will often use more than one data collection method. Let's look at an example.

A music company releases the first album for a new band. They hire you to evaluate their program to promote the band. One data collection method may not be enough. We might do some face-to-face surveys at a mall to see if consumers are aware of the new band. We could also conduct a telephone survey to see if they like the band's music. Liking is not buying, so we might also want to ask how likely they are to actually buy the band's CD. But, likelihood of buying is not really buying either! It's really an intention to buy. So, we might like to look at actual purchases -- how many, when, and by whom. By using several different methods, we start to close in on a reality. Triangulation gives us a better view of reality.

Another data collection tool is **sampling**. A sample is a sub-set of a **population**. With most projects, it is not feasible for sociologists to include every person in a population to study a topic. So we use samples instead. The basic issue of sampling is to get a group of people who accurately represent a population so that you can **generalize** your findings from the sample to the population. There are somewhat complicated formulas that many sociologists use to determine the **sample size** they need in order to make valid generalizations. We do not have time to review all of this science, but we can give you a brief introduction to the types of samples that sociologists use.

There are two categories of samples: **probability samples** and **non-probability samples**. Probability samples use scientific methods to ensure that they accurately represent the larger population. Probability samples are called probability samples because, through the use of scientific methods, every person (or group) in the population has a known probability of being selected for the sample, and often every person has an equal chance of being selected. Generalizations from probability samples are usually valid. Non-probability samples try to represent the larger population too but do not use scientific methods to do so. Hence they often fall short of an accurate representation. Subsequently, generalizations from non-probability samples can be suspect.

Below is a table that summarizes the most common probability and non-probability samples.

Chapter 2.6: Data Collection Methods

Probability Sampling	Non-Probability Sampling
Simple Random: With simple random samples (SRS) you obtain a list of all the people (or groups) in the population and randomly select people from off of it to reach your sample size.	Convenience: With convenience samples, also called availability samples, you select people for a sample because they are easily accessible to you. Many polls done on television shows use convenience samples of TV viewers. Convenience samples rarely represent their populations.
Systematic Random: With systematic random samples you obtain a list of all the people (or groups) in the population and randomly select every n^{th} person from off of it, such as every 10^{th} person. The interval at which you select people depends on the sample size that you need.	Quota: Quota samples are done for the same reasons as convenience samples but you take an additional step to try to represent some of the characteristics of the population. For example, if you were assessing movie viewers' satisfaction with a new movie, you might go to a local movie theater (convenience sample). But instead of just quickly interviewing any movie viewers willing to talk to you, you interview a percentage of men and women in different age groups that matches the percentage to which these groups exist in the population.
Proportionate Stratified Random: With proportionate stratified samples you divide people into homogeneous groups called strata and then draw a SRS from within each strata. You do this when you will not get enough cases with a regular SRS to represent a group. For example, using a national SRS you would likely not obtain enough Asians or Latinos to generalize to, or be inclusive of, those groups. With a proportionate stratified sample, you might divide the population into strata by race and randomly sample, within each strata, a percentage of people that matches the percentage that each group exists in the nation or population.	Snowball: With a snowball sample you find one person in your target population and then ask them to give you access to other people that they know in the population. Then you ask those people to give you access to other people, and so on. You choose a snowball sample when you do not know the members of your target population, or you can not access them. For example, you might use a snowball sample in a study of employment discrimination among gay men and lesbian women.

Chapter 2.6: Data Collection Methods

Probability Sampling	Non-Probability Sampling
Disproportionate Stratified Random: These samples are done for the same reasons as a proportionate stratified sample and follow the same steps except that you draw a disproportionate number of cases from within each strata because a proportionate number would not result in enough cases for analysis. For example, a national sample stratified proportionately by race would likely not generate enough cases of Pacific Islanders with which to conduct generalizable analyses. So, instead we might randomly select a larger percentage within that strata, and a lower percentage within the Caucasian strata.	Theoretical: With theoretical samples, also called purposive samples, you use a theory or information about your topic to choose a sample. For example if you are hired by a state correctional facility to evaluate the effectiveness of domestic violence rehabilitation programs across the state, your topic more or less requires you to sample past participants in these programs, or possibly their families or criminal justice personnel.

Let's walk through the choice of a sampling method and a data collection method to answer a question. Which methods should we use to determine the effect of health information on health behavior among people with cancer? For example, some people that have cancer will try to learn more about their illness and treatment by reading books, articles, Internet sites, and other materials. This information may influence how they eat, or exercise, which in turn may influence the impact of cancer on their body. Knowing how health information influences health behavior could help health care providers better treat illnesses. You might try to answer this question for one of the national organizations for cancer patients.

Your access to cancer patients would drive which data collection method was best to use. It would be difficult to observe cancer patients searching information and then observe the impact the information had on their behavior. Where would you find cancer patients searching for information? Maybe in the library? But how will you know they have cancer, or what information they are seeking? So, observational studies would likely not work.

An experiment is out of the question. It would not be ethical to take a group of cancer patients' and give some of them information about their illness and deny others the same information to see how that information, or lack of it, affects their illness. And, using existing data won't work because data on what cancer patient's read about their illness does not exist. Doctors and hospitals do not collect that data as part of their regular workdays.

Chapter 2.6: Data Collection Methods

Well, we are down to surveys and in-depth interviews. Either may be plausible. With both we need someone to give us access to cancer patients. For example, we could not just do a random sample of households and mail them a survey or call their homes and expect to get enough cancer patients to study our question. So we would need to think about where cancer patients can be found. Doctor's offices would be a good place. Maybe we could get a group of oncologists to grant us access to the patients that they have treated. Oncologists might grant us access if we explained the project to them and the potential benefits of the findings. But we would have no guarantee that some or any of their patients sought information to read about their illness. So we might then also think about where we could find cancer patients who definitely sought information about their illness. We could go to a cancer information website and try to get the webmaster to give us access to their visitors, just as we did with the oncologists. Both the oncology patients and the website visitors would be theoretical samples.

Assuming we get access from both the oncologists and the webmaster, we would then need to decide between in-depth interviews or surveys. In-depth interviews would allow us to dig deeper into what health care information patients read and their subsequent health behaviors. However, in-depth interviews would be more expensive. But an even more critical issue here is privacy. Most people who visit Internet sites prefer to do so anonymously – that is the beauty of the Internet. People who visit a cancer website do not usually identify themselves or how to contact them. So it would be nearly impossible to do in-depth interviews with website visitors. Hence, we would choose a mail survey that can be mailed to the oncologists' patients and that can be posted on the website for visitors to complete.

TASKS

- Your client, the CEO of an information technology firm, wants to know what she can do to reduce the turnover rate among her employees. What data collection methods will you use to help answer her question?
- The director of a program to rehabilitate men convicted of spouse abuse hires you to evaluate the effectiveness of the program. What data collection methods will you use?
- A local PTA hires you to identify what services and programs parents would like the PTA to provide. What data collection methods will you use?