

Introduction to sampling

- Video: Welcome to week 3
2 min
- Video: Cliff: Value everyone's contributions
2 min
- Video: Introduction to sampling
4 min
- Reading: The relationship between sample and population
20 min
- Video: The sampling process
5 min
- Reading: The stages of the sampling process
20 min
- Video: Compare sampling methods
6 min
- Reading: Probability sampling methods
20 min
- Video: The impact of bias in sampling
5 min
- Reading: Non-probability sampling methods
20 min
- Ungraded Plugin: Identify: Sampling methods
10 min
- Practice Quiz: Test your knowledge: Introduction to sampling
3 questions

Sampling distributions
Work with sampling distributions in Python
Review: Sampling

Non-probability sampling methods

Recently, you learned that probability sampling methods use random selection, which helps avoid sampling bias. A randomly chosen sample means that all members of the population have an equal chance of being included. In contrast, non-probability sampling methods do not use random selection, so they do not typically generate representative samples. In fact, non-probability methods often result in biased samples. **Sampling bias** occurs when some members of the population are more likely to be selected than other members.

In this reading, you'll learn more about four methods of non-probability sampling, and learn how sampling bias can affect each method. We'll also discuss why non-probability sampling may be useful in certain situations.

Non-probability sampling methods

Non-probability samples use non-random methods of selection, so not all members of a population have an equal chance of being selected. This is why non-probability methods have a high risk of sampling bias. However, non-probability methods are often less expensive and more convenient for researchers to conduct. Sometimes, due to limited time, money, or other resources, it's not possible to use probability sampling. Plus, non-probability methods can be useful for exploratory studies, which seek to develop an initial understanding of a population, rather than make inferences about the population as a whole.

We'll go over four methods of non-probability sampling:

- Convenience sampling
- Voluntary response sampling
- Snowball sampling
- Purposive sampling

Let's explore each method in more detail.

Convenience sampling

For convenience sampling, you choose members of a population that are easy to contact or reach. As the name suggests, conducting a convenience sample involves collecting a sample from somewhere convenient to you, such as your workplace, a local school, or a public park.

For example, to conduct an opinion poll, a researcher might stand at the entrance of a shopping mall during the day and poll people that happen to walk by.



Because these samples are based on convenience to the researcher, and not a broader sample of the population, convenience samples often suffer from undercoverage bias. Undercoverage bias occurs when some members of a population are inadequately represented in the sample. For example, the above sample will underrepresent people who don't like to shop at malls, or prefer to shop at a different mall, or don't visit the mall because they lack transportation.

Convenience sampling is often quick and inexpensive, but it's not a reliable way to get a representative sample.

Voluntary response sampling

A voluntary response sample consists of members of a population who volunteer to participate in a study. Like a convenience sample, a voluntary response sample is often based on convenient access to a population. However, instead of the researcher selecting participants, participants volunteer on their own.

For example, let's say college administrators want to know how students feel about the food served on campus. They email students a link to an online survey about the quality of the food, and ask students to fill out the survey if they have time.

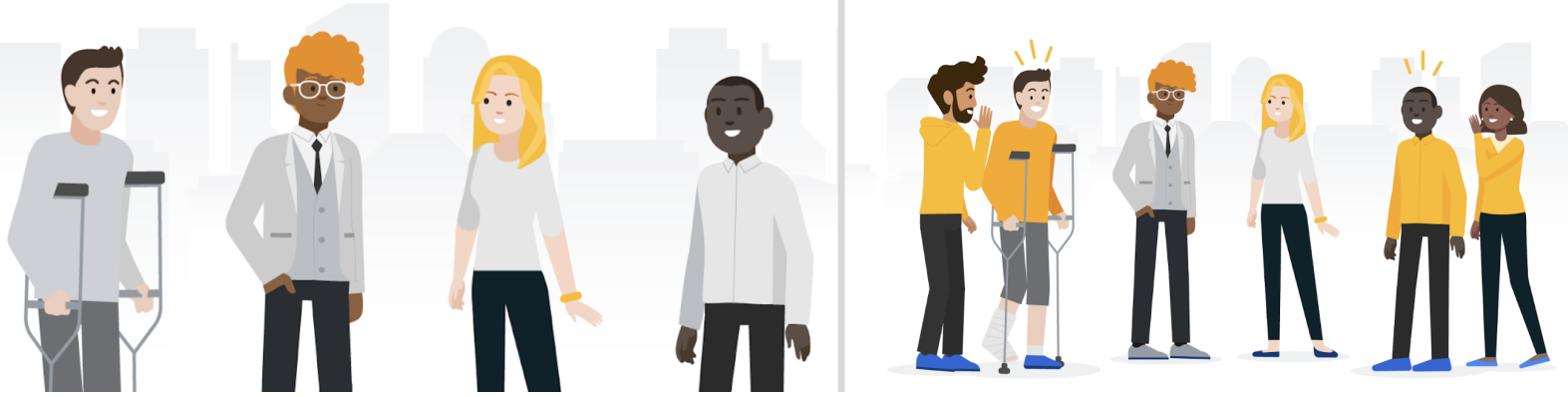


Voluntary response samples tend to suffer from nonresponse bias, which occurs when certain groups of people are less likely to provide responses. People who voluntarily respond will likely have stronger opinions, either positive or negative, than the rest of the population. In this case, only students who really like or really dislike the food may be motivated to fill out the survey. The survey may omit many students who have more mild opinions about the food, or are neutral. This makes the volunteer students unrepresentative of the overall student population.

Snowball sampling

In a snowball sample, researchers recruit initial participants to be in a study and then ask them to recruit other people to participate in the study. Like a snowball, the sample size gets bigger and bigger as more participants join in. Researchers often use snowball sampling when the population they want to study is difficult to access.

For example, imagine a researcher is studying people with a rare medical condition. Due to reasons of confidentiality, it may be difficult for the researcher to obtain contact information for members of this population from hospitals or other official sources. However, if the researcher can find a couple of people willing to participate, these two people may know others with the same condition. The initial participants could then recruit others by sharing the potential benefits of the study.



Snowball sampling can take a lot of time, and researchers must rely on participants to successfully continue the recruiting process and build up the "snowball." This type of recruiting can also lead to sampling bias. Because initial participants recruit additional participants on their own, it's likely that most of them will share similar characteristics, and these characteristics might be unrepresentative of the total population under study.

Purposive sampling

In purposive sampling, researchers select participants based on the purpose of their study. Because participants are selected for the sample according to the needs of the study, applicants who do not fit the profile are rejected.

For example, imagine a game development company wants to conduct market research on a new video game before its public release. The research team only wants to include gaming experts in their sample. So, they survey a group of professional gamers to provide feedback on potential improvements.



In purposive sampling, researchers often intentionally exclude certain groups from the sample to focus on a specific group they think is most relevant to their study. In this case, the researcher excludes amateur gamers. Amateur gamers may purchase the new game for different reasons than professional gamers, and enjoy game features that don't appeal to professionals. This could lead to biased results, because the professionals in the sample are not likely to be representative of the overall gamer population.

Purposive sampling is often used when a researcher wants to gain detailed knowledge about a specific part of a population, or where the population is very small and its members all have similar characteristics. Purposive sampling is not effective for making inferences about a large and diverse population.

Key takeaways

Non-probability sampling is useful for collecting data in situations where you have limited time, budget, and other resources. Non-probability sampling is also useful for exploratory research, when you want to get an initial understanding of a population, rather than make inferences about the population as a whole. However, it's important to remember that non-probability sampling methods have a high risk of sampling bias.

As a data professional, you have to think about bias and fairness from the moment you start collecting sample data to the time you present your conclusions. Once you become aware of some common forms of bias, you can remain on the alert for bias in any form.

Mark as completed