A Slice of Society: The Variety and Significance of Sampling Methods in the Social Sciences

Nicola Righetti

A Slice of Society: The Variety and Significance of Sampling Methods in the Social Sciences

Introduction

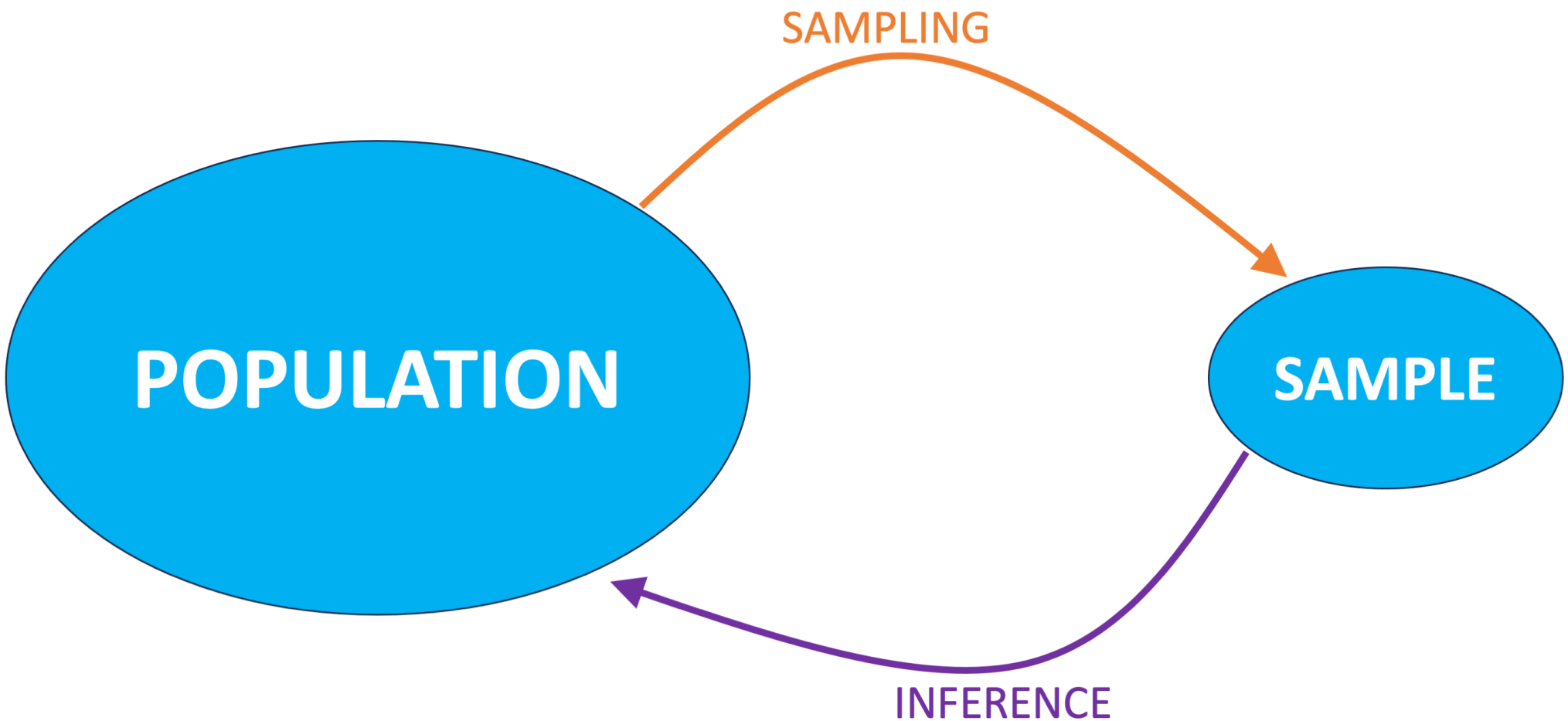
In this lesson we will discuss:

- The importance of sampling in research, and
- The main types of sampling that can be used in research.

As always, you can find the readings indicated in the syllabus in the Moodle folder. The reference of this section is:

> Allen, M. (Ed.). (2017). *The SAGE encyclopedia of communication research methods*. SAGE publications. Pagg. 1523-1552

Sampling, Sample, and Population

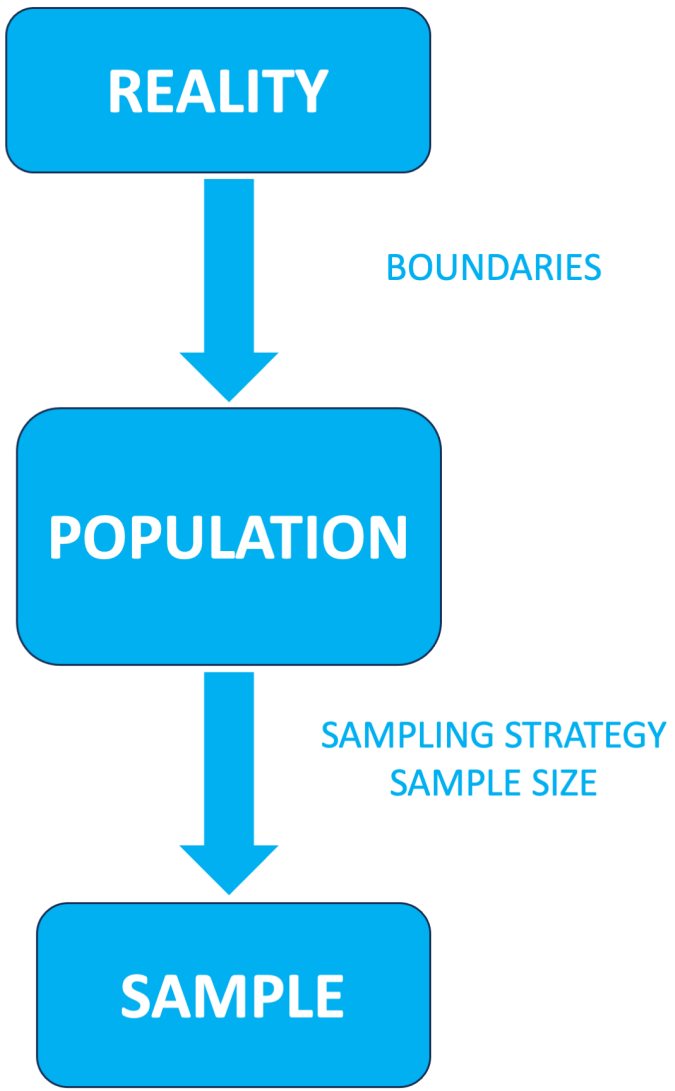


Sampling, Sample, and Population

Sampling is the process of creating *samples*, which are subsets of a *population*.

- **Sampling**: The process of selecting a subset of individuals or elements from a larger group (population) to estimate characteristics of the whole group.
- **Sample**: A subset of individuals or elements selected from a larger population, used to make inferences about the population.
- **Population**: The entire pool of individuals or elements from which a sample is drawn for analysis, representing the total group of interest.

Main Steps of a Sampling Process



Define the Population: Boundaries

- Once researchers have determined what they want to study they must first define their population and determine population boundaries.
- A population boundary separates who or what researchers are interested in studying from those they are not interested in investigating.
- You can think about defining your population in terms of the class of individuals of interest and their attributes.

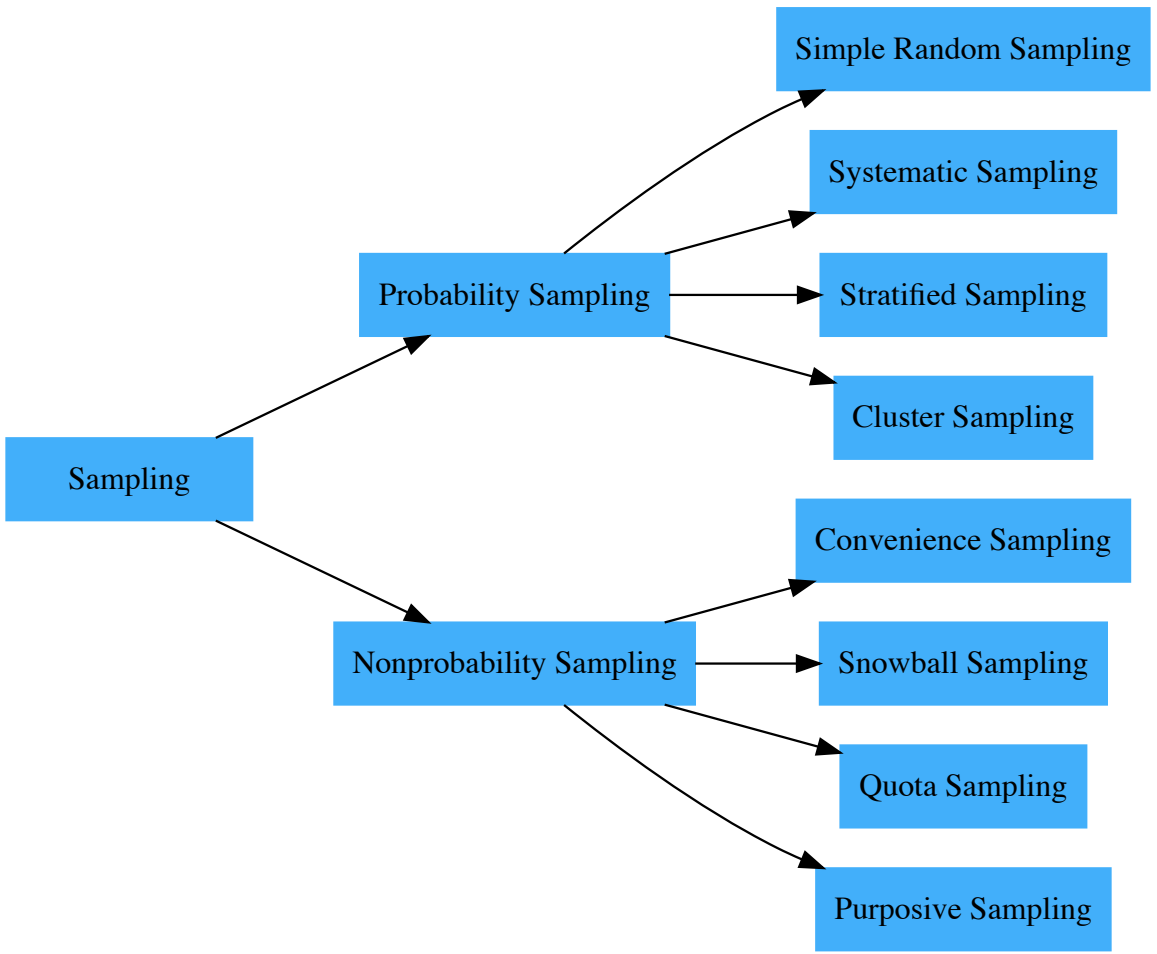
Examples of Populations and their Boundaries

- All the movies (class) that are (attributes) (1) about racism and (2) published on Netflix, and (3) available in Austria.
- All the people (class) that are (attributes) (1) 20-40 y.o., and (2) are students of University of Vienna, and (3) speak German.

Sampling Frame

- Once the population boundaries are set, the researchers should frame the population. This is referred to as a sampling frame.
- A sampling frame is a **list** of individuals who meet the population characteristics for a particular study. From that list it is possible to extract a sample (or multiple samples) to study.
  - For example, when everyone or almost everyone had a landline telephone at home and their number appeared in the telephone registers, that could be used as a list to select a sample of respondents to a telephone survey.
- Unfortunately, having a complete list of individuals in a population is an ideal condition and is only occasionally possible. Sometimes it is possible to have an incomplete list, or no list at all.

Types of Sampling



Probability and Nonprobability Sampling

- **Probability sampling**, also known as random sampling, involves selecting participants from the population that gives all potential participants an equal chance of being selected for participation in the study. It allows the researcher to calculate and report the sampling error, making it easier to generalize results to the larger population.
- **Nonprobability sampling** returns samples that are not randomly selected. Nonprobability samples can be representative of the population but they have a greater chance of biasbecause they are not randomly selected and sampling error cannot be easily calculated.

Sampling Bias and Sampling Errors

- **Sample Bias**: Sample bias occurs when a sample is not representative of the population from which it's drawn, leading to skewed results. It arises when certain members of the population are systematically more likely to be selected in the sample than others, causing the sample to inaccurately reflect the overall population.
- **Sampling Error**: Sampling error refers to the difference between the characteristics of a sample and those of the overall population. It's an inherent part of sampling that results from observing only a subset of the population. Sampling error does not imply a mistake; rather, it's a statistical variation that is expected and can be quantified in probabilistic sampling.

Probability Sampling

- Simple random sampling involves identifying every person in the population and ensuring that each member of the population has an equal chance of being selected.
- Systematic sampling is a probability sampling method in which the researcher will select every nth person from the population list to include in the study.

Probability Sampling

- Stratified sampling involves dividing the population into specific subgroups of interest and then randomly selecting participants from these subgroups.
- Cluster sampling divides the population into groups (or “clusters”), and a random sample of these clusters is selected for study. All members within the selected clusters are then included in the research.

Probability Sampling Examples

- **Simple Random Sampling**: A media researcher wants to understand the average time spent on social media platforms by university students. The researcher compiles a list of all students in a university and randomly selects a group of them to survey, ensuring each student has an equal chance of being chosen.
- **Systematic Sampling**: A communication scientist is studying the frequency of news consumption among residents in a city. Starting from a randomly chosen starting point on a list of all residents, the researcher selects every 50th person to participate in a survey about their news-watching habits.

Probability Sampling Examples

- **Stratified Sampling**: In a study to compare radio listening habits across age groups, a researcher divides the population into different age brackets (e.g., 18-30, 31-45, 46-60, 61+) and then randomly selects participants from each age group to ensure representation across all age demographics.
- **Cluster Sampling**: A researcher studying the impact of a new advertising campaign on consumer behavior realizes it's impractical to survey every consumer in the country. Instead, the researcher divides the country into regions (clusters), randomly selects a few regions, and surveys every consumer within these selected regions to assess the campaign's impact.

Nonprobability Sampling

- Convenience sampling involves selecting participants based on their availability. This method is frequently used in survey research that asks students to participate for course extra credit.
- Snowball sampling involves asking participants who have already completed the study to refer the researcher to others they think might be willing to participate in the study.
- Quota sampling involves identifying population characteristics along with the sample size for each category in the study.
- Purposive sampling involves identifying individuals that are known to possess specific characteristics the researcher is interested in studying.

Nonprobability Sampling Examples

- **Convenience Sampling**: A media studies student conducting a survey on podcast listening preferences sets up a booth outside the university library. The student invites passersby to participate in the survey, selecting individuals based on their willingness and availability at that moment.
- **Snowball Sampling**: A researcher studying the use of underground music forums begins with a few known users and asks them to refer friends who also use these forums. As each participant refers more users, the sample grows, resembling a “snowball” effect.

Nonprobability Sampling Examples

- **Quota Sampling**: In a study examining news consumption habits across different income levels, a researcher decides to include 50 participants from each income bracket (low, middle, and high). The researcher continues to recruit participants until the quota for each income bracket is met.
- **Purposive Sampling**: A communication scientist researching the impact of virtual reality on learning outcomes specifically seeks out educators who have integrated virtual reality technology into their teaching. These educators are known to have firsthand experience with the subject of interest.

Short Light-Hearted Break

The sampling song (by ChatGPT and Nicola Righetti)

(Verse 1: Probabilistic sampling)
🎶
In **Simple Random**, it's fair play, each person has an equal say,
Systematic's neat, you'll see, pick in steps, like one, two, three!
Stratified divides with care, groups by traits they all share,
Cluster's when we can't list all, pick some groups, big or small!

(Chorus)
Sampling, sampling, in our song,
Probabilistic, can't go wrong,
Random, systematic, don't be shy,
Stratified, or cluster, give them a try!

(Verse 2: Nonprobabilistic sampling)
Convenience takes who's near,
Easy to find, have no fear.
Snowball rolls from friend to friend,
Gathering more 'til the end.
Quota fills its preset slots,
Balanced groups, like diverse plots.
Purposive's got a goal in mind,
Specific traits, that's what we find!

(Chorus)
Sampling, sampling, pick your way,
Structured or relaxed, it's okay,
Nonprobabilistic, a bit free-form,
But each has a method, that's the norm!

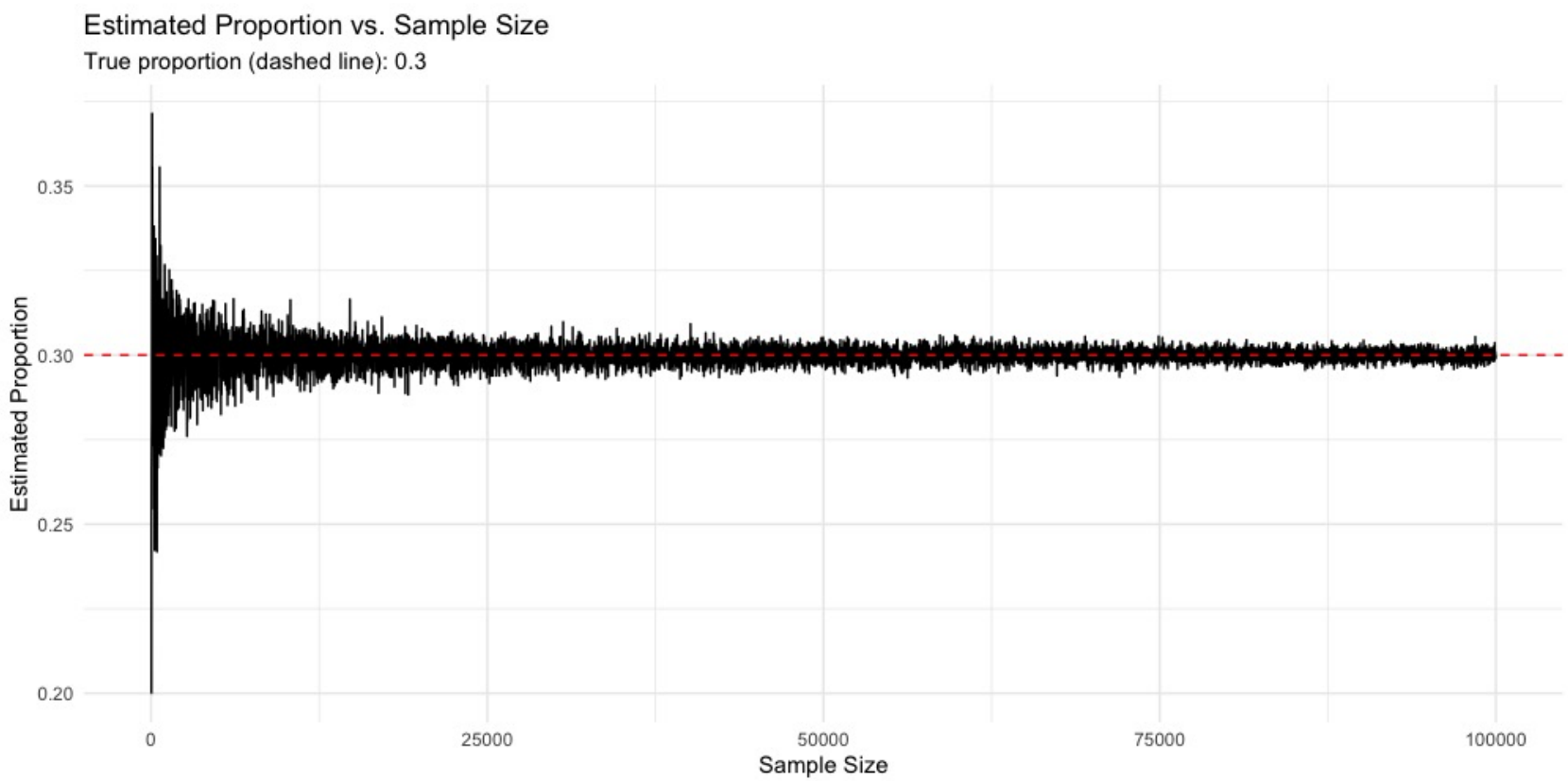
(Outro)
Sampling, sampling, make your choice,
Nonprobabilistic, raise your voice,
Different strategies, in your toolkit,
Pick the right one, and you'll hit it!
🎶

Sample Size

- Once the (1) population of interest has been defined and (2) the ideal type of sampling and the one that can actually be used have been decided (based on time, resources, and general feasibility), the size of the sample must be decided.
- Calculating the sample size can be quite complex. In general term, it is decided based on a number of considerations, including:
    1. the desired precision of the estimates;
    2. the characteristic of the population and the object of study;
    3. the type of sampling;
    4. the overall research design for which the sample is needed.

Sample Size and Estimate Precision

Generally, larger samples are preferable for greater statistical precision. As sample size increases, estimates become more precise, converging to the population's true value ("asymptotically" in infinite populations) or yielding the exact value (in finite populations).



Sample Size and Effect Identification

- An inadequately sized sample can produce inaccurate estimates and may prevent the identification of statistical effects, such as the causal impact of one variable on another.
- Consider the sample size as a magnifying glass: the larger it is, the more capable it is of identifying even small effects. Conversely, a small sample can only detect macroscopic effects.

Sample Size and Power Analysis

- Determining the sample size in a statistical study can be complex, yet for simple estimates like a proportion or an average, there are relatively straightforward formulas.
- There are rough rules of thumb for sample size, such as having at least 10-20 observations per independent variable in multiple regression analysis.
- **Power Analysis** offers a more precise approach. “Power” technically refers to the probability of detecting a true effect when it exists. It requires some assumptions and preliminary information. It can be performed on statistical sofware as R. You can also find Sample Size Calculators online.

Sample Size in Qualitative Research Designs

- Even purely qualitative research makes use of samples. In this case, there are **no formulas** to determine the sample size.
- The criterion used is that of **saturation**, which is the point at which no new information or themes are observed in the data.
- A general rule used in a qualitative design based on interviews, which for example uses a quota sampling strategy, is to have some 10/20 interviews per quota (this paper found that 12 can be enough).