2g – A Guide to
Sampling Techniques

Royal Geographical Society with IBG
Advancing geography and geographical learning

A **Sample** is a selection of data chosen from all of that possibly available. Sampling is needed in almost all forms of data collection as in most research processes it is simply not possible to gain data from every available source. For example, if one wished to conduct some interviews within a town it would not be possible to interview absolutely every resident of that town. Instead, a selection of the population would be used to try to get a representation of the town's answers. The method you use to select this sample is known as your **Sampling Technique**.

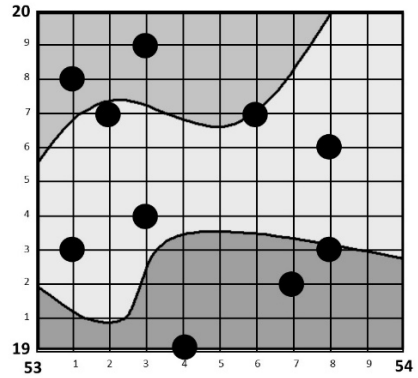
Sampling not only makes conducting your data collection possible, it can also make more efficient use of the time you have to collect your data. It may not even be necessary to collect all points of data for your research, and in many cases your overall population size (or your **Sampling Frame** – the pool of data from which you are drawing a sample) may be an unknown quantity; as a researcher you would therefore not know when to stop collecting data.

Researchers frequently put great effort into deciding on the size of their data sample. The larger the **sample size**, the more representative it is likely to be of your overall sampling frame, and as a result, the more justifiable your conclusions will be. However, the size of any sample is also dependent on the time and resources you have available and how manageable the data collection, and data analysis will be as a result.

While there are many different sampling techniques, there are three common methods used frequently by researchers: **Random**, **Systematic** and **Stratified Sampling**. In the following explanations, two examples will be used. In the first, Researcher A is trying to find a sample of ten data collection locations on a map that contains three different geographical zones. In the second, Researcher B is trying to select a sample of ten interviewees from a population of twenty people of different ages.

Random Sampling

Random sampling is where sources of data are chosen in a completely haphazard way. Once the size of the sample has been decided (maybe as a percentage of the overall sampling frame), researchers use random number generators, which can be found online, to give completely random sets of numbers. These can then be used to create grid references for data collection sites on a map or tell researchers which house numbers to survey within a street.



Nuffield Foundation
This project was funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily those of the Foundation.

Alternatively, if the researcher is looking for a random transect line on a map, a random number generator can give the grid references for the start and end points of that line on a map.

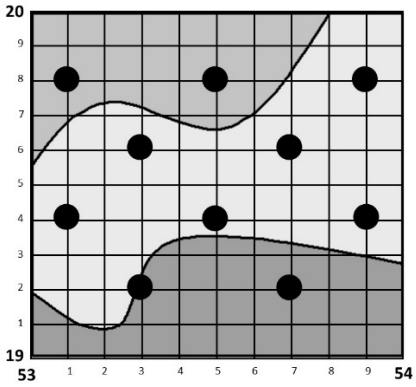
For surveys of natural terrain, where the researcher plans to use a quadrat, a common method for choosing random survey sites is to stand in the centre of the area and throw the quadrat with one's eyes closed. Surveying where the quadrat lands, and then repeating the method from that spot can create a random selection of sites.

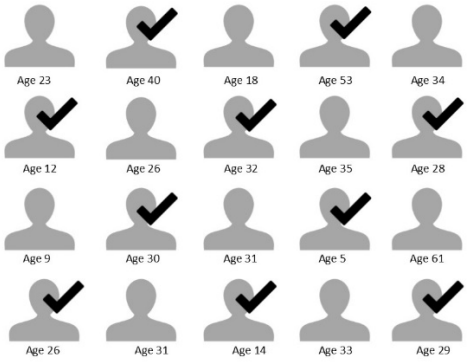
Rolling dice, choosing unseen playing cards, and picking bingo numbers out of a bag can be other ways of making random number selections.

Though randomly generated numbers take a human choice element out of the sampling process and so reduce the chance of human bias in the results, random sampling in general is not always suitable for small sampling frames as there are limited choices to be had.

Systematic Sampling

Systematic sampling is where sources of data are chosen in a completely non-random way. Here the size of the sample may not necessarily be decided before the sampling begins as the chosen system itself may create the sample size on its own. The interval size between sampling points (distance on a map, or every n^{th} person in a survey) is chosen by the researcher and stuck to without compromise.

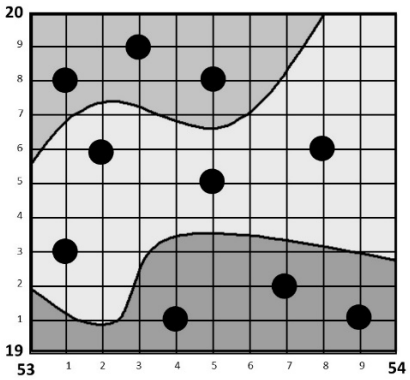


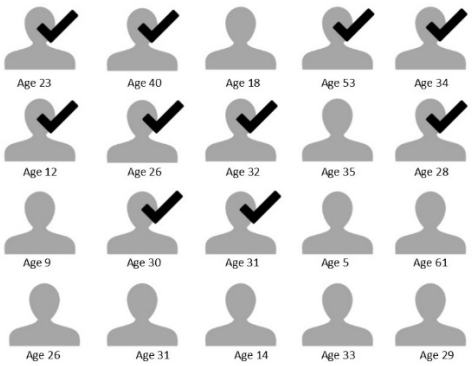


The benefits of systematic sampling are that the researcher is largely removed from the selection process and therefore bias can be avoided. However, in order for the sample to be truly representative of the study area, the researcher must also ensure that the sampling frame itself does not inadvertently create bias. For example, if the sampling frame for a survey were to be taken from a pre-selected list of people, such as the electoral role, it would automatically exclude people who were not eligible to vote, such as those aged under eighteen and some prisoners.

Stratified Sampling

Stratified sampling involves splitting the sample frame into smaller groups or **Strata** and using these strata to 'weight' the sample chosen accordingly to represent the original sampling frame. So if it is known that thirty percent of the sample frame came from a particular location were of a particular age group or belonged to a particular religion, thirty percent of the sample would also represent these strata.





Researcher A has placed more location points in the middle area as it represents a larger percentage of the overall grid square than the other two.

Researcher B has worked out the percentage of people that fall into different age categories in the original sampling frame. They have then used these percentages to determine how many of each age group should appear in the actual sample.

Many researchers believe that stratified sampling represents the most unbiased of the three techniques mentioned here as the sample becomes a true representation of the original sampling frame. However, an extensive knowledge of the sampling frame may be needed before a researcher can choose the strata and a pilot study may be needed in order to decide on the 'weighting' for each strata that is used.

It may be possible, depending on the exact nature of your research, to combine different sampling techniques together. For example, you may choose to draw a random transect line across a data collection site on a map, but then use a systematic or stratified sampling technique to choose the exact points of data collection thereafter.