

Section 1: Collecting data

Notes and examples

These notes contain subsections on

- [Populations and samples](#)
- [Types of sampling](#)
- [Bias in sampling](#)

Populations and samples

A population in statistics means all the individuals you are interested in for a particular investigation, for example: all 6 year-old girls in the UK, or all the trees in a public park.

Often it is not possible to investigate an entire population. If you wanted to gather data about the trees in a public park, it would probably be possible to look at every one, if you had enough time and there weren't too many trees. However it would be very difficult to gather data about every 6 year-old girl in the UK. It would be necessary to select a sample.

Suppose, for example, you wanted to find out the mean height of 6 year-old girls in the UK. There are a number of ways you could choose your sample (see the list of sampling methods below). However, it is important to remember that different samples will give you different results. If your sample is large, and you have used a good sampling method that is free from bias, the mean height from your sample will probably be reasonably close to the mean height of the entire population. However, you must always be cautious about assuming that a sample statistic such as the mean height is representative of the population as a whole.

Types of sampling

Simple random sampling

In this sampling method, the items in the sample are chosen by a random procedure such as drawing tickets out of a box or using a random number generator. Every possible sample of the required size has the same probability of being selected. One way to achieve this is for a finite population is for every member of the population to have an equal chance of being selected as long as sampling is without replacement and selections are independent of each other.

Simple random sampling is not always possible, because you may not have a list of every member of the population.

Opportunity sampling

In opportunity sampling, individuals are chosen to be part of a sample as opportunity arises. Interviewing passers by on the street is one example.

Edexcel AS Maths Data 1 Notes and examples

Stratified sampling

If the parent population can be divided into subgroups, or strata, such as by age or gender, then stratified sampling ensures that all strata are sampled. The subgroups are not expected to be representative of the population. If the numbers sampled from each of the strata are proportional to the sizes of the strata, then this is called proportional stratified sampling. Otherwise, weighting must be used.

Quota sampling

A quota sample is similar to a stratified sample but it is specified in terms of the number of data items required in each stratum, for example, a certain number of males and a certain number of females might be required. This method is often used by interviewers, and the actual selection of the sample members is up to the interviewer, whereas with stratified sampling it will often be done at random.

Systematic sampling

A method of choosing individuals to form a sample. For example, if the parent population was all the Year 11 students in a school, you might obtain an alphabetical list of the students and select every 10th student on the list.

Bias in sampling

Sampling techniques sometimes run the risk of introducing bias into the sample.

Opportunity sampling is prone to bias: for example if you survey passers by in the middle of the morning on a working day, your sample may include a disproportionate number of people of retired age.