



Sampling Methods

By

Amritansh

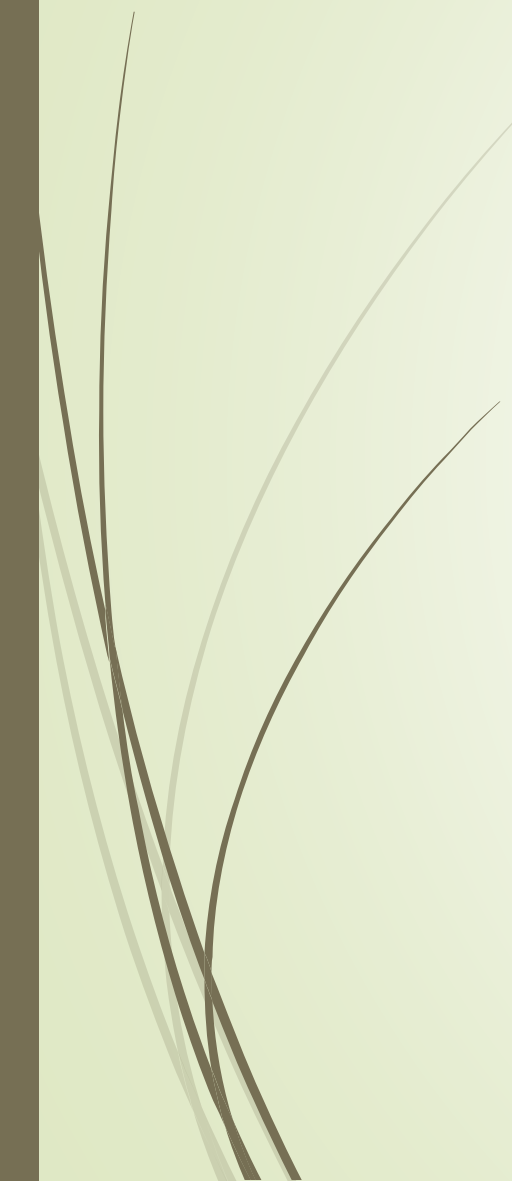


Types of Variables

- Quantitative
 - Measured in amounts e.g. litres, kg, pixels
 - Height , Weight, Picture resolution
- Qualitative
 - Measured in categories
 - Gender, Race, diagnosis, labels
- Discrete
 - Separate categories, Grades: A+, A, A-
- Continuous
 - Continuous values. Infinite values

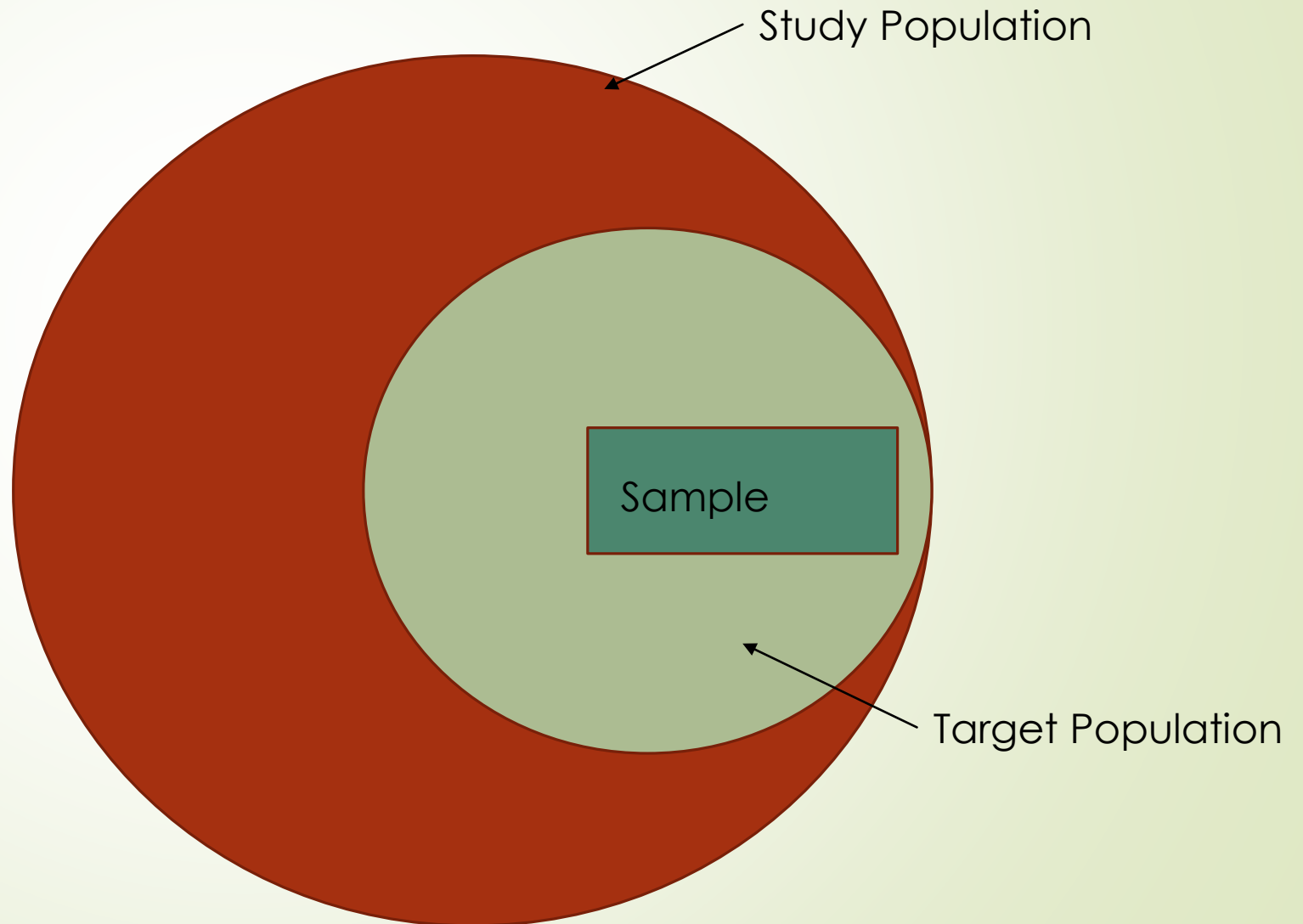


Sampling

- ▶ A sample is a smaller collection of units from a population used to determine truths about the mentioned population
 - ▶ Why bother with samples:
 - ▶ Less Resources (time and computation)
 - ▶ Given accuracy can be trusted mathematically
- 

➤ 3 factors that influence sample representative-ness

- Sampling procedure
- Sample Size
- Participation



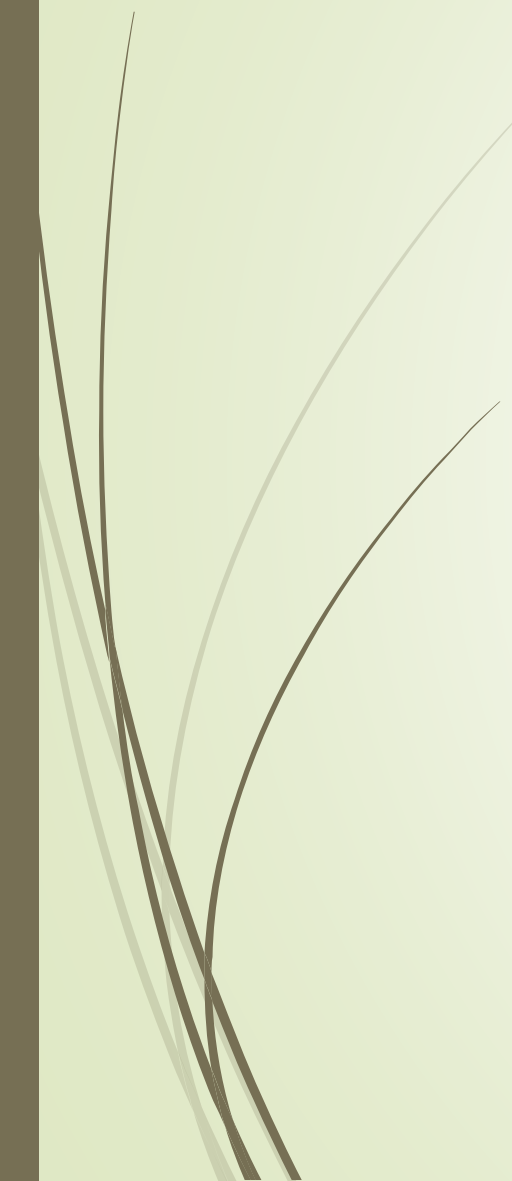


Types of Sampling Methods

- Random Sampling
 - Stratified Random Sampling
 - Cluster Sampling
 - Systematic Sampling
 - Representative Sampling
 - Haphazard Sampling
 - Sampling with Replacement
- 



Random Sampling

- ▶ Applicable when population is small, homogeneous & readily available
 - ▶ All subsets of the frame are given an equal probability. Each element of the frame thus has an equal probability of selection
 - ▶ It provides for greatest number of possible samples. This is done by assigning a number to each unit in the sampling frame.
 - ▶ A table of random number or lottery system is used to determine which units are to be selected.
- 



Pros

Estimates are easy to calculate.

Simple random sampling is always an EPS design, but not all EPS designs are simple random sampling.

Cons

If sampling frame large, this method impracticable

Minority subgroups of interest in population may not be present in sample in sufficient numbers for study.



Systematic Sampling

- Systematic sampling relies on arranging the target population according to some ordering scheme and then selecting elements at regular intervals through that ordered list
- Systematic sampling involves a random start and then proceeds with the selection of every k th element from then onwards. In this case, $k = (\text{population size} / \text{sample size})$
- A simple example would be to select every 10th name from the telephone directory



Pros

Sample easy to select

Suitable sampling frame can be identified easily

Sample evenly spread over entire reference population

Cons

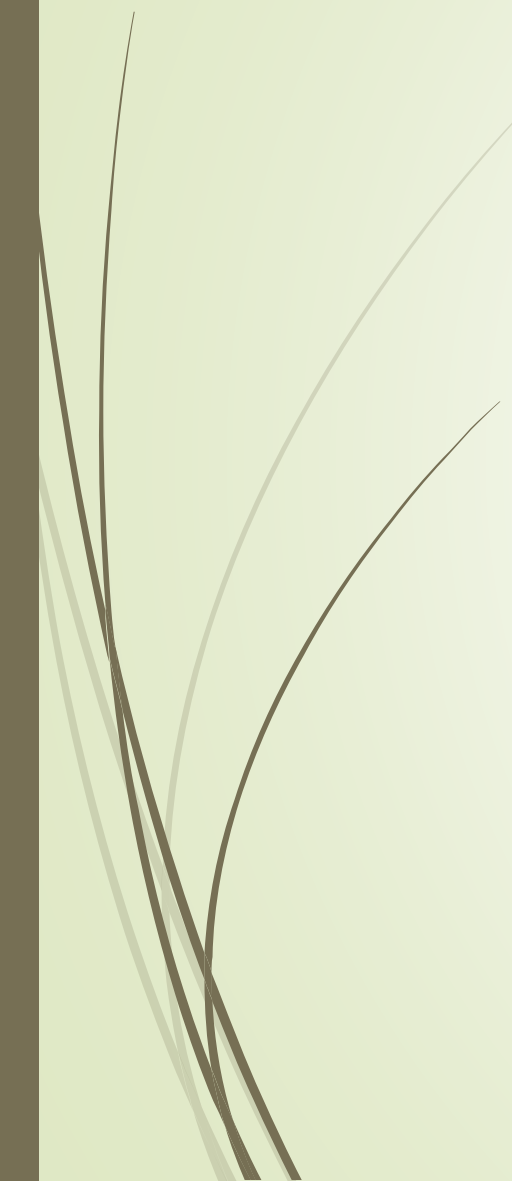
Sample may be biased if hidden periodicity in population coincides with that of selection

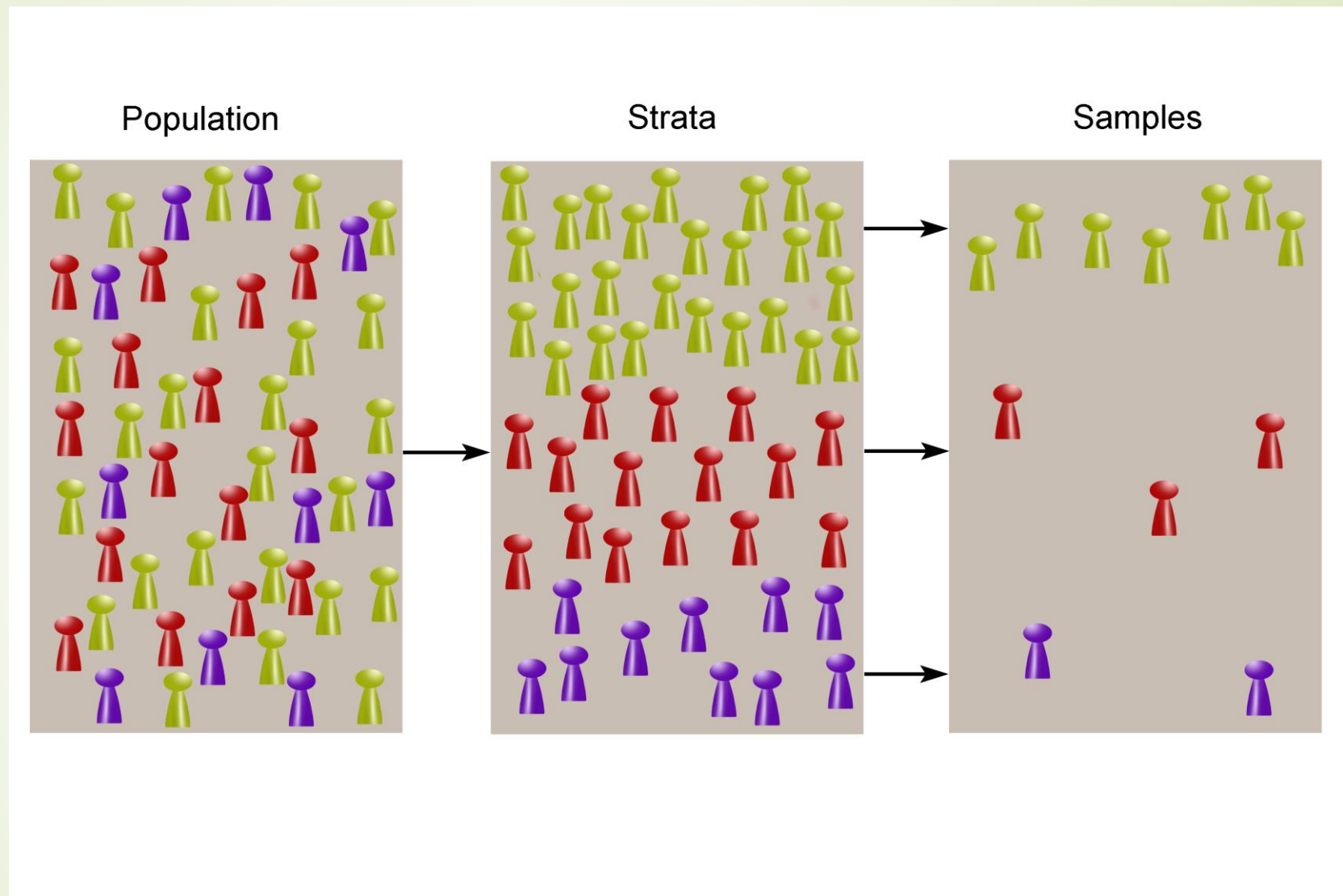
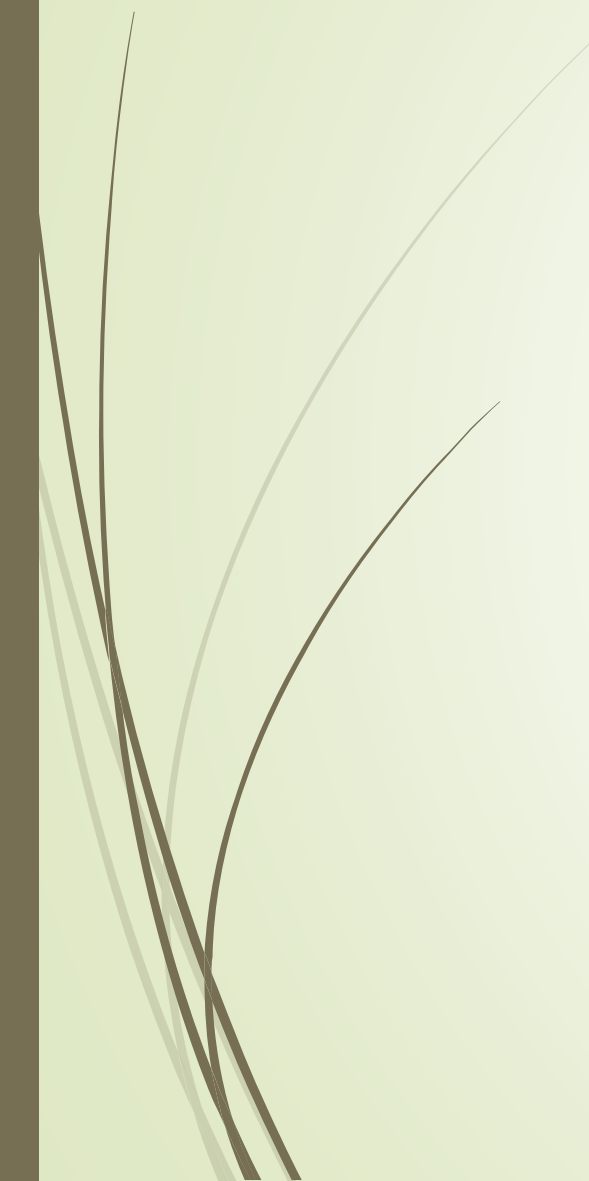
Difficult to assess precision of estimate from one survey





Stratified Sampling

- Where population embraces a number of distinct categories, the frame can be organized into separate "strata." Each stratum is then sampled as an independent sub-population, out of which individual elements can be randomly selected
 - Every unit in a stratum has same chance of being selected
 - Using same sampling fraction for all strata ensures proportionate representation in the sample
 - Adequate representation of minority subgroups of interest can be ensured by stratification & varying sampling fraction between strata as required.
- 





Cons

First, sampling frame of entire population has to be prepared separately for each stratum

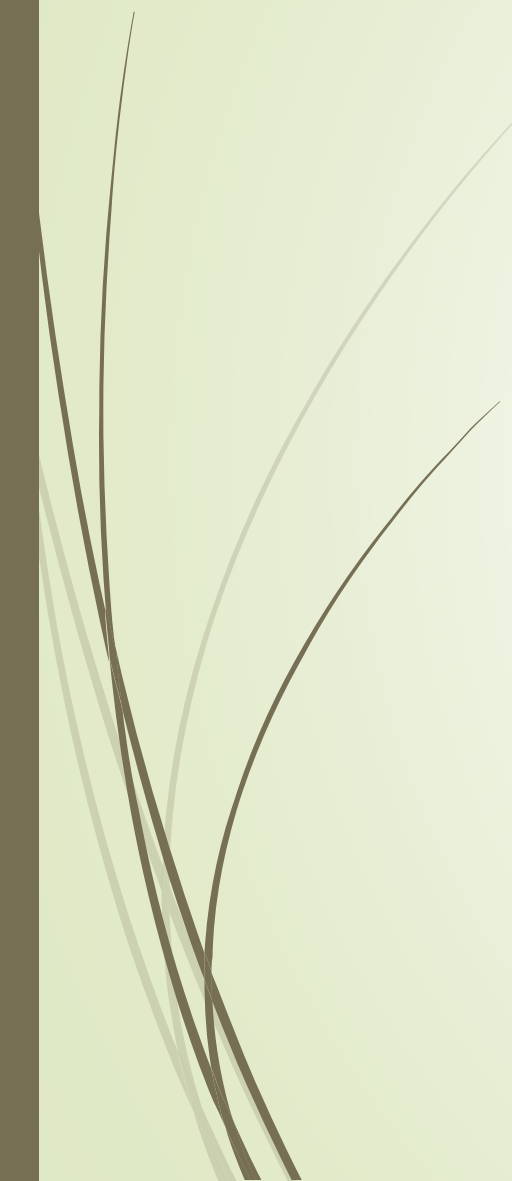
Second, when examining multiple criteria, stratifying variables may be related to some, but not to others, further complicating the design, and potentially reducing the utility of the strata.

Finally, in some cases (such as designs with a large number of strata, or those with a specified minimum sample size per group), stratified sampling can potentially require a larger sample than would other methods

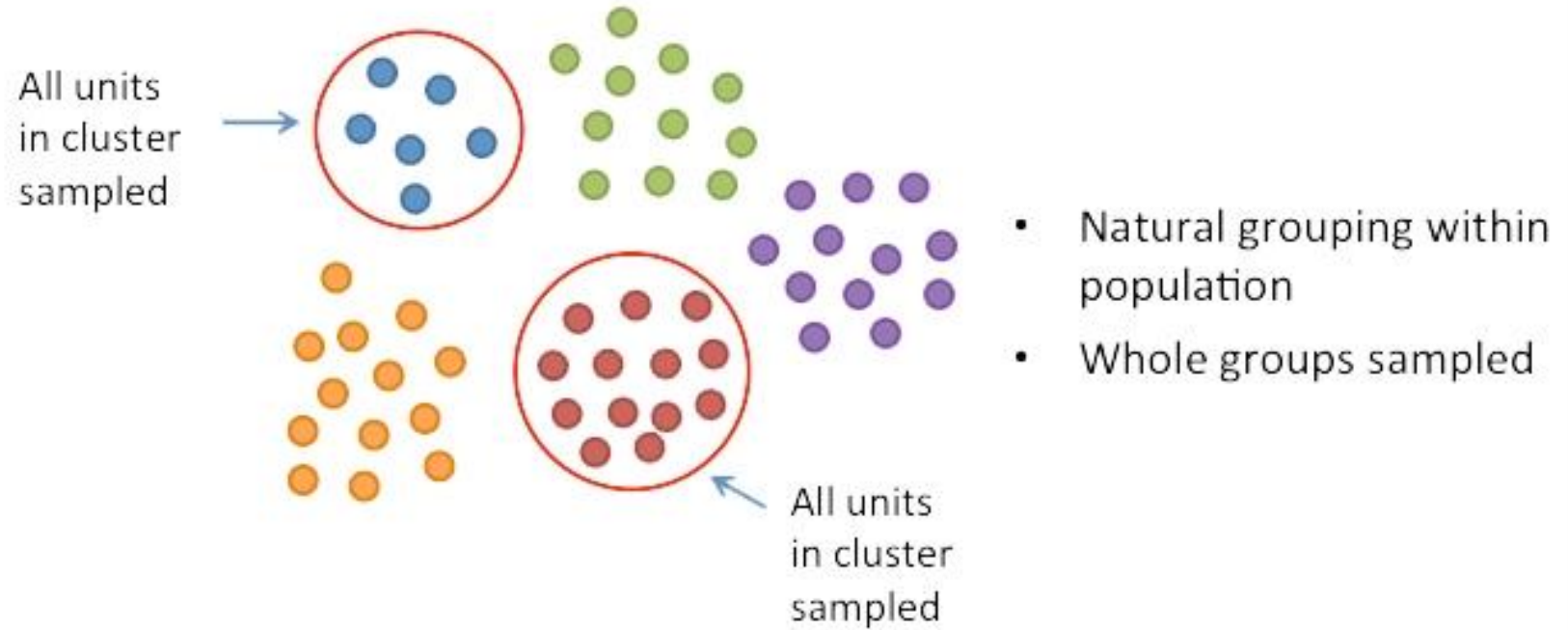




Cluster Sampling

- Cluster sampling is an example of 'two-stage sampling'
 - First stage : a sample of areas is chosen
 - Second stage : a sample of respondents within those areas is selected.
 - Population is divided into clusters of homogeneous units
 - Sampling units are groups rather than individuals
 - A sample of such clusters is then selected.
 - All units from the selected clusters are studied.
- 

Cluster Sampling





Pros

Cuts down on the cost of preparing a sampling frame

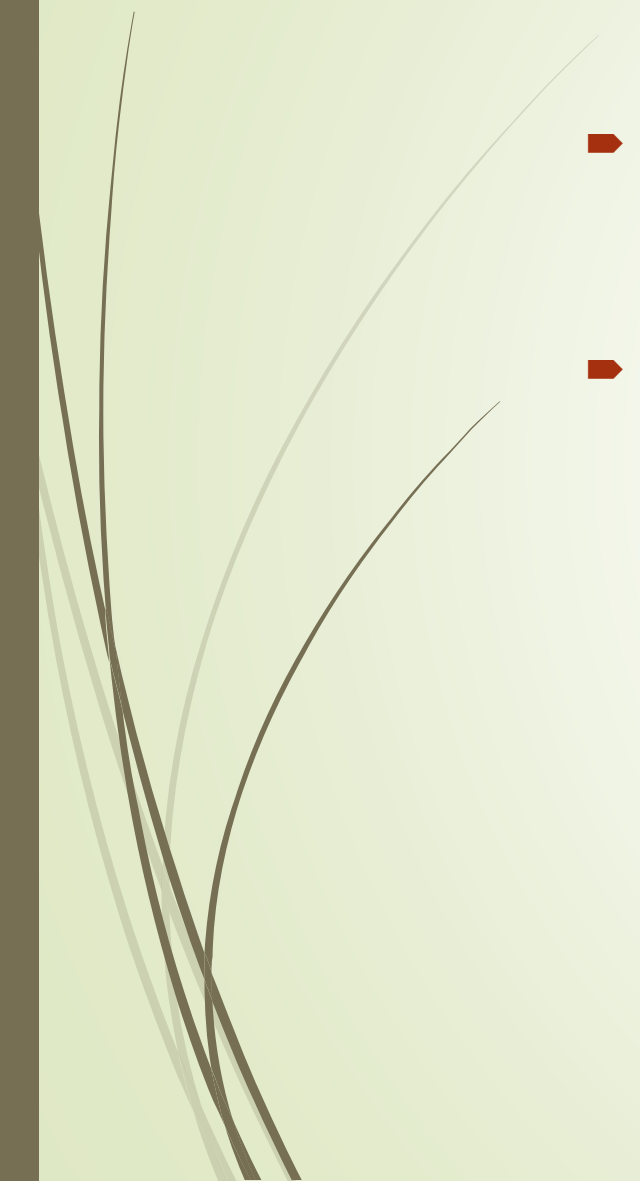
Cons

Sampling error is higher for a simple random sample of same size.



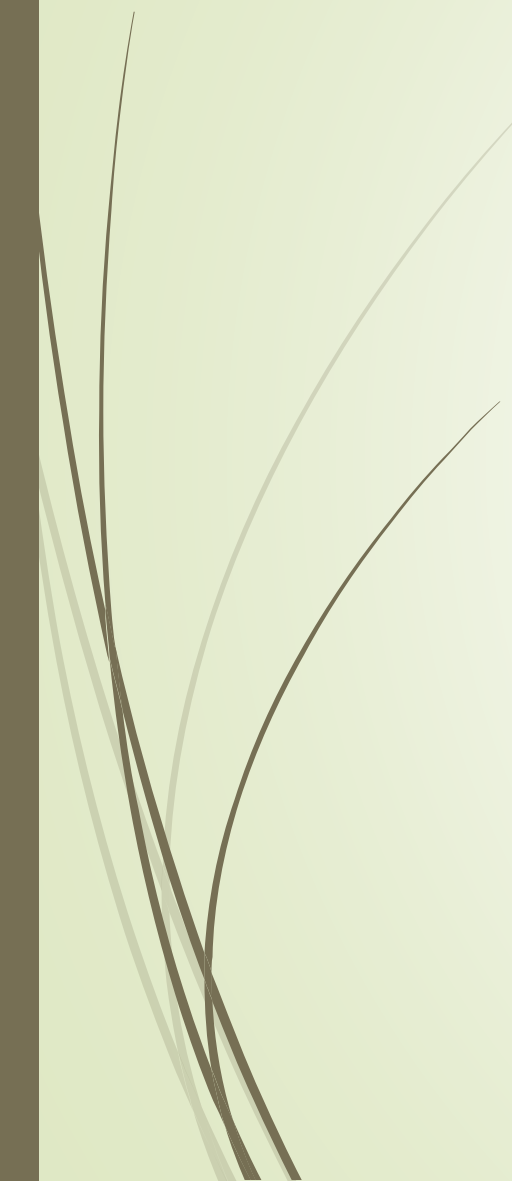


Types of Cluster Sampling

- **One-stage sampling:** All of the elements within selected clusters are included in the sample
 - **Two-stage sampling:** A subset of elements within selected clusters are randomly selected for inclusion in the sample.
- 



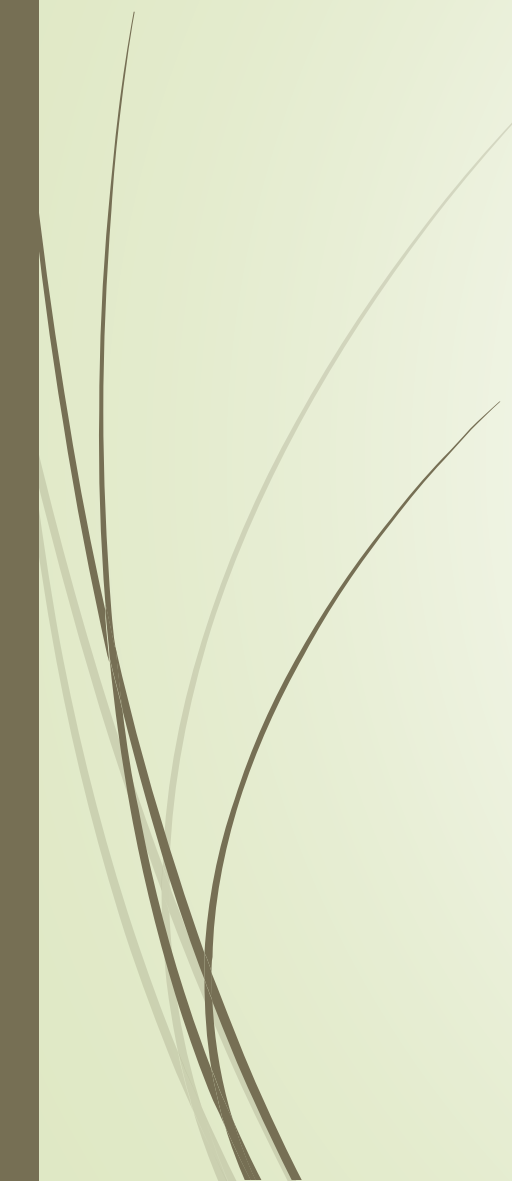
Difference in Strata & Cluster Sampling

- ▶ All strata are represented in the sample; but only a subset of clusters are in the sample
 - ▶ With stratified sampling, the best survey results occur when elements within strata are internally homogeneous. However, with cluster sampling, the best results occur when elements within clusters are internally heterogeneous
- 



CONVENIENCE SAMPLING

Also known as haphazard sampling or accidental sampling

- A type of nonprobability sampling which involves the sample being drawn from that part of the population which is close to hand. That is, readily available and convenient.
 - The researcher using such a sample cannot scientifically make generalizations about the total population from this sample because it would not be representative enough.
 - This type of sampling is most useful for pilot testing
- 

For example, if the interviewer was to conduct a survey at a shopping center early in the morning on a given day, the people that he/she could interview would be limited to those given there at that given time, which would not represent the views of other members of society in such an area, if the survey was to be conducted at different times of day and several times per week.

In social science research, snowball sampling is a similar technique, where existing study subjects are used to recruit more subjects into the sample.

Convenience Sample

select any members of the population who are conveniently and readily available

