

## **2. Sampling Methods**

### **[IT2110]**

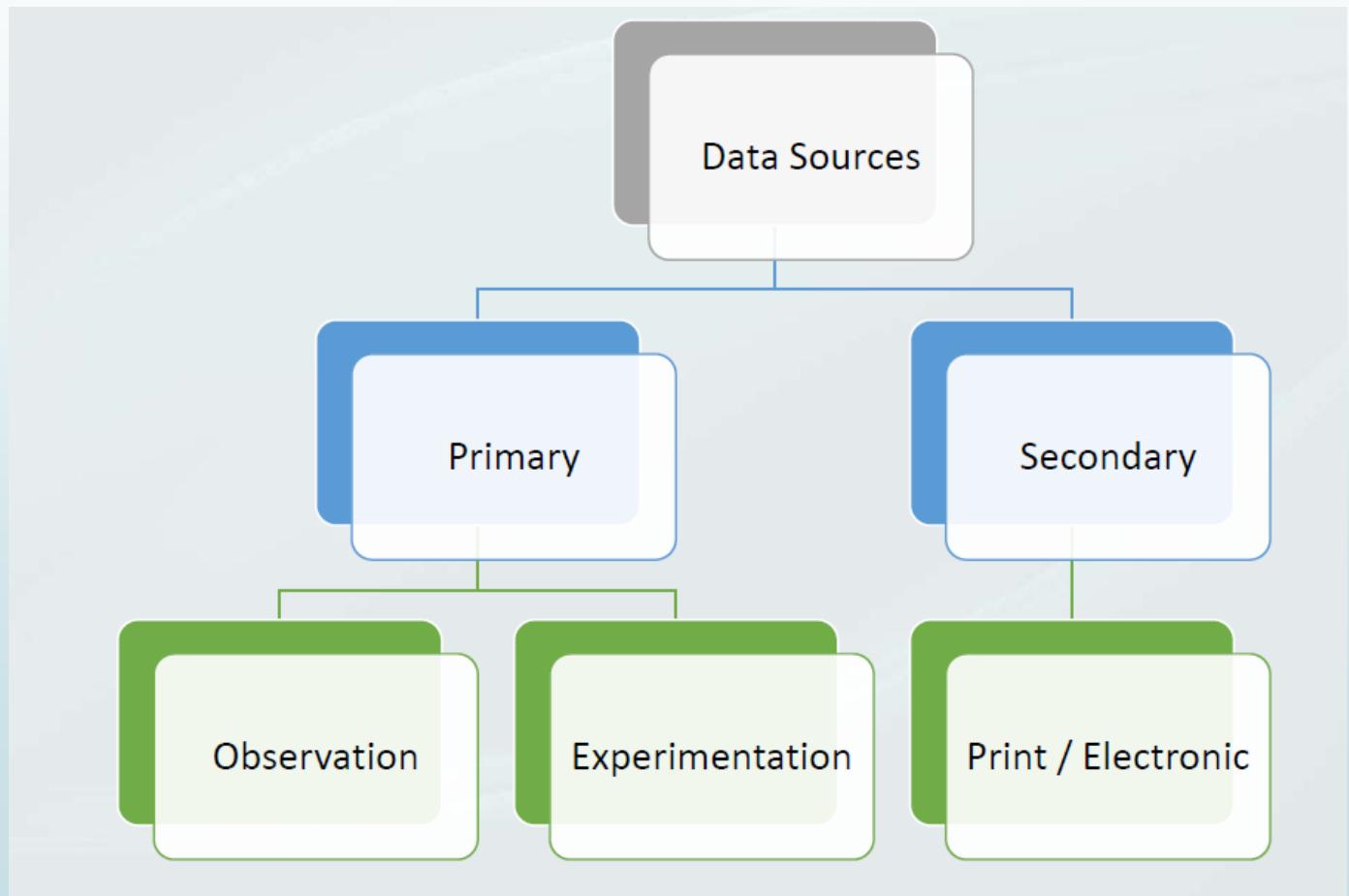


***By Department of Mathematics and Statistics  
Faculty of Humanities and Sciences, SLIIT***



# 1. Introduction to Sampling

# Data Sources

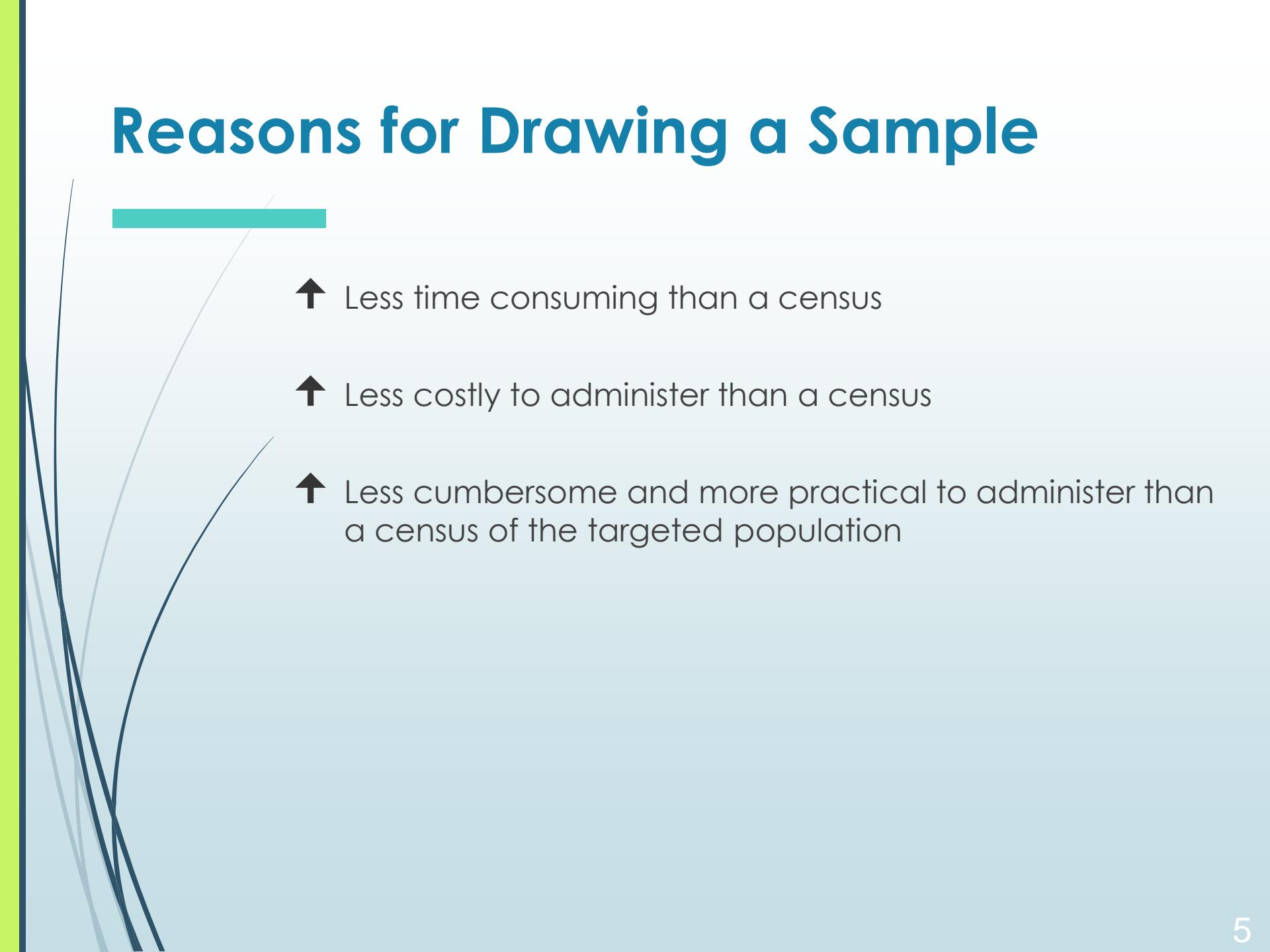


# Introduction to Sampling



- ↑ Data should be collected before describing.
- ↑ If a sampling survey is done, should plan how to select the sample.
- ↑ **Two types of sampling:**
  - ← Probability sampling.
  - ← Non-probability sampling.
- ↑ Why should a proper sample be selected?

# Reasons for Drawing a Sample



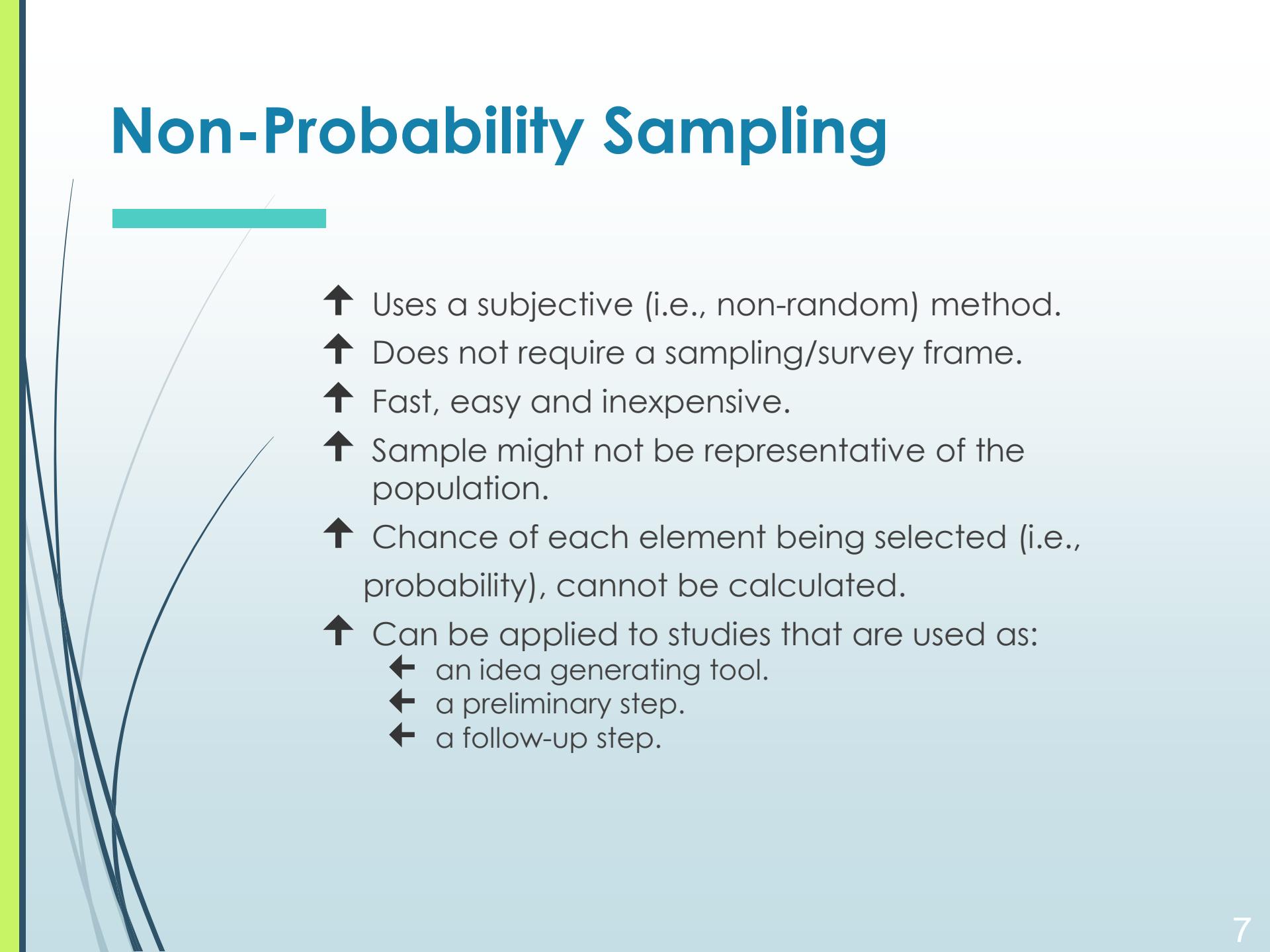
---

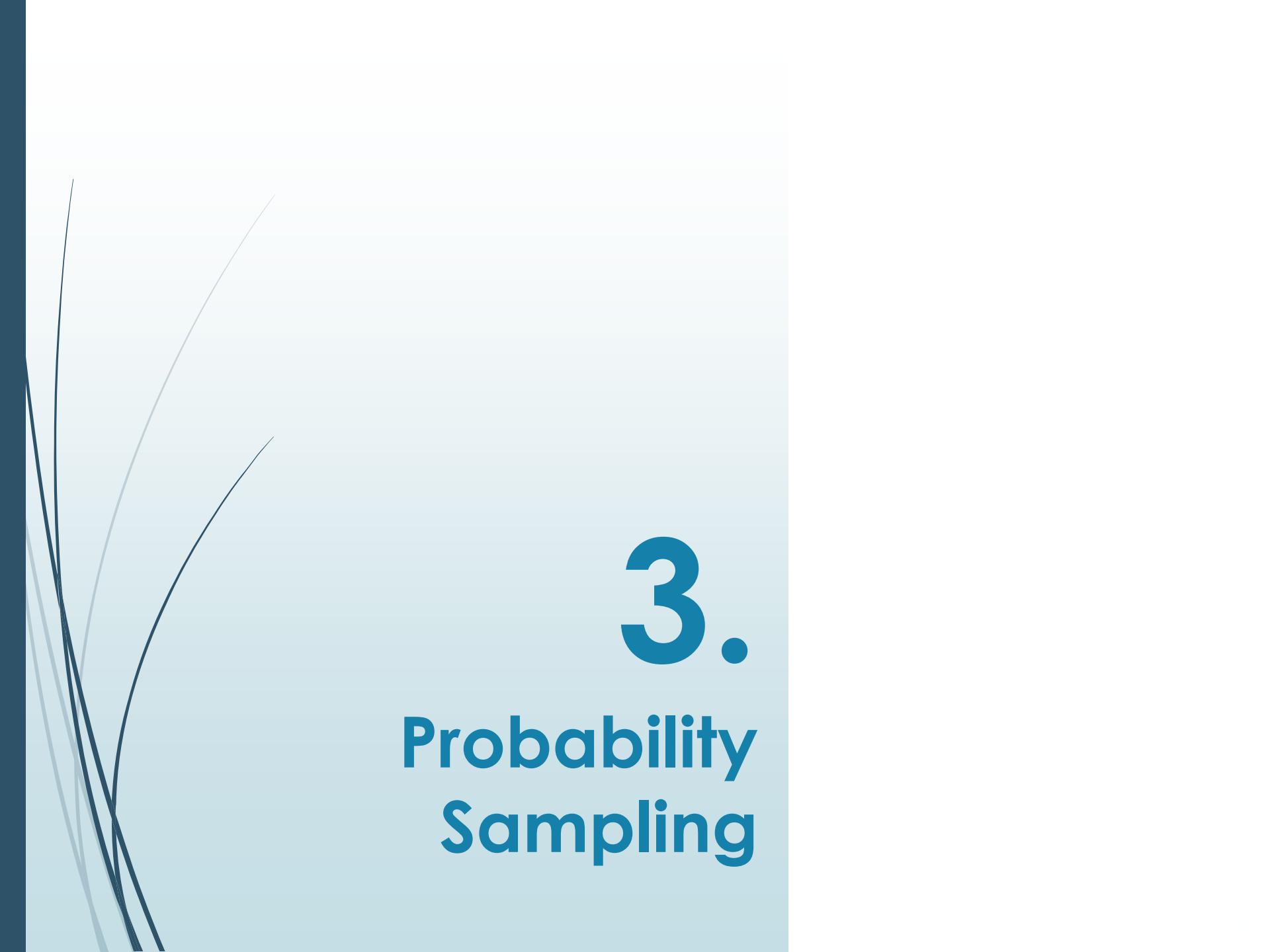
- ↑ Less time consuming than a census
- ↑ Less costly to administer than a census
- ↑ Less cumbersome and more practical to administer than a census of the targeted population

# 2.

## **Non-Probability Sampling**

# Non-Probability Sampling

- 
- ↑ Uses a subjective (i.e., non-random) method.
  - ↑ Does not require a sampling/survey frame.
  - ↑ Fast, easy and inexpensive.
  - ↑ Sample might not be representative of the population.
  - ↑ Chance of each element being selected (i.e., probability), cannot be calculated.
  - ↑ Can be applied to studies that are used as:
    - ← an idea generating tool.
    - ← a preliminary step.
    - ← a follow-up step.



# 3. Probability Sampling

# Probability Sampling



---

- ↑ Based on the principle of randomization or chance.
- ↑ More complex, time consuming and usually more costly.
- ↑ More reliable.
- ↑ Requires a sampling/survey frame.
- ↑ Can use computers or other methods to select elements randomly (e.g.: random number tables).

# Sampling Frame



- ↑ The list of elements from which a sample may be drawn.
- ↑ Also known as: ***working population***.
- ↑ Examples: Telephone directory, List of voters

# Probability Sampling (cont'd.)

- 
- ↑ Commonly used probability sampling methods:
    - ← Simple Random Sampling (SRS).
    - ← Systematic Sampling (SYS).
    - ← Probability-Proportional-to-Size (PPS) Sampling.
    - ← Cluster Sampling.
    - ← Stratified Sampling (STR).
    - ← Multi-Stage Sampling.
    - ← Multi-Phase Sampling.
    - ← Replicated Sampling.

# Simple Random Sampling (SRS)



---

- ↑ Starting point for all probability sampling designs.
- ↑ Each unit in the sample has the same inclusion probability ( $n$  – Sample Size,  $N$  – Population Size).
- ↑ Sampling may be done with or without replacement (SRSWR or SRSWOR).
- ↑ Generally, SRSWOR yields more precise results and is operationally more convenient.

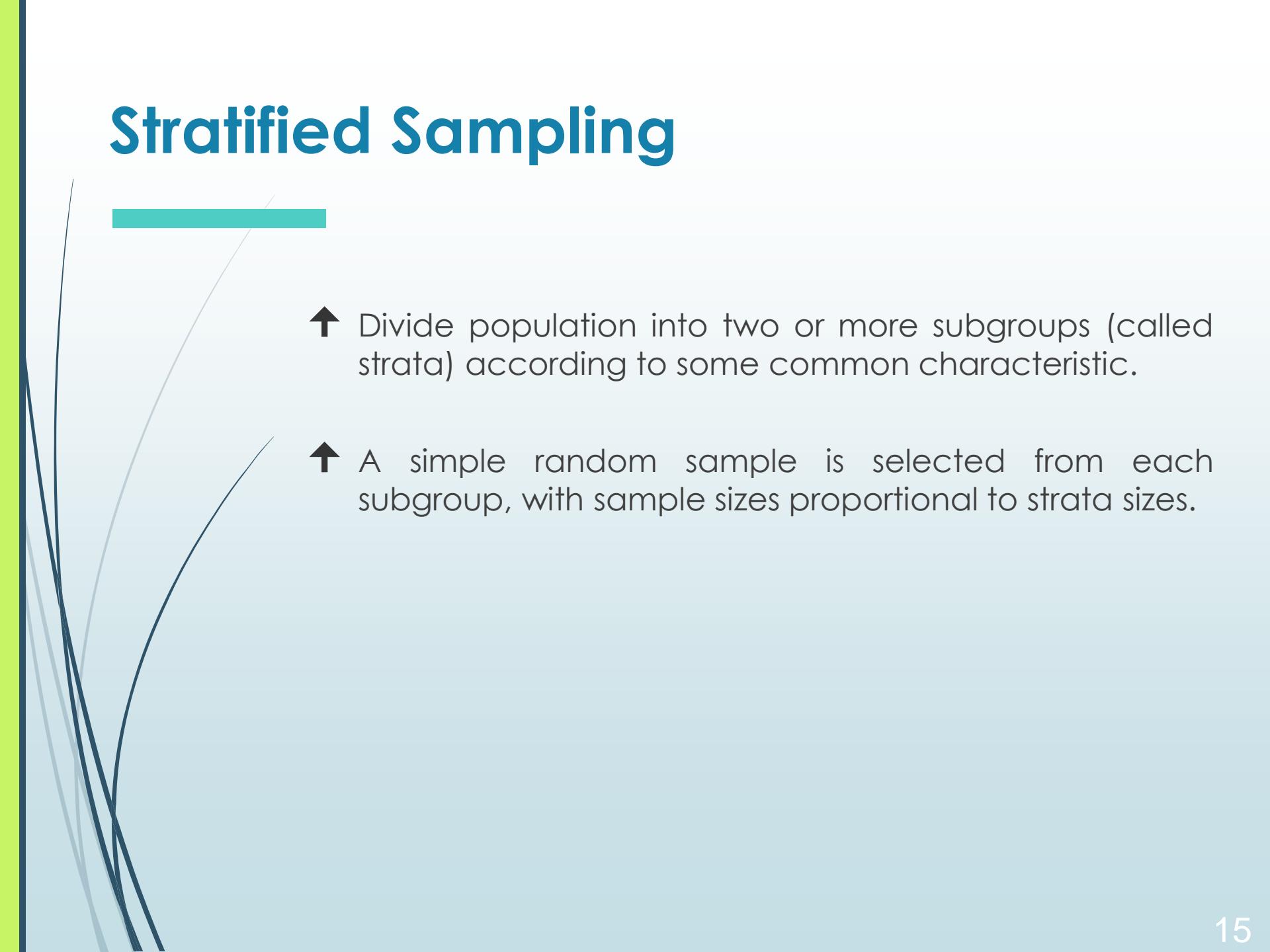
# SRS (cont'd.)

- 
- ↑ Advantages of SRS
    - ← Simplest sampling technique.
    - ← Requires no additional (auxiliary) information on the frame in order to draw the sample.
    - ← Needs no technical development.
  
  - ↑ Disadvantages of SRS
    - ← Makes no use of auxiliary information even if such information exists on the survey frame.
    - ← Can be expensive.
    - ← It is possible to draw a 'bad' SRS sample.

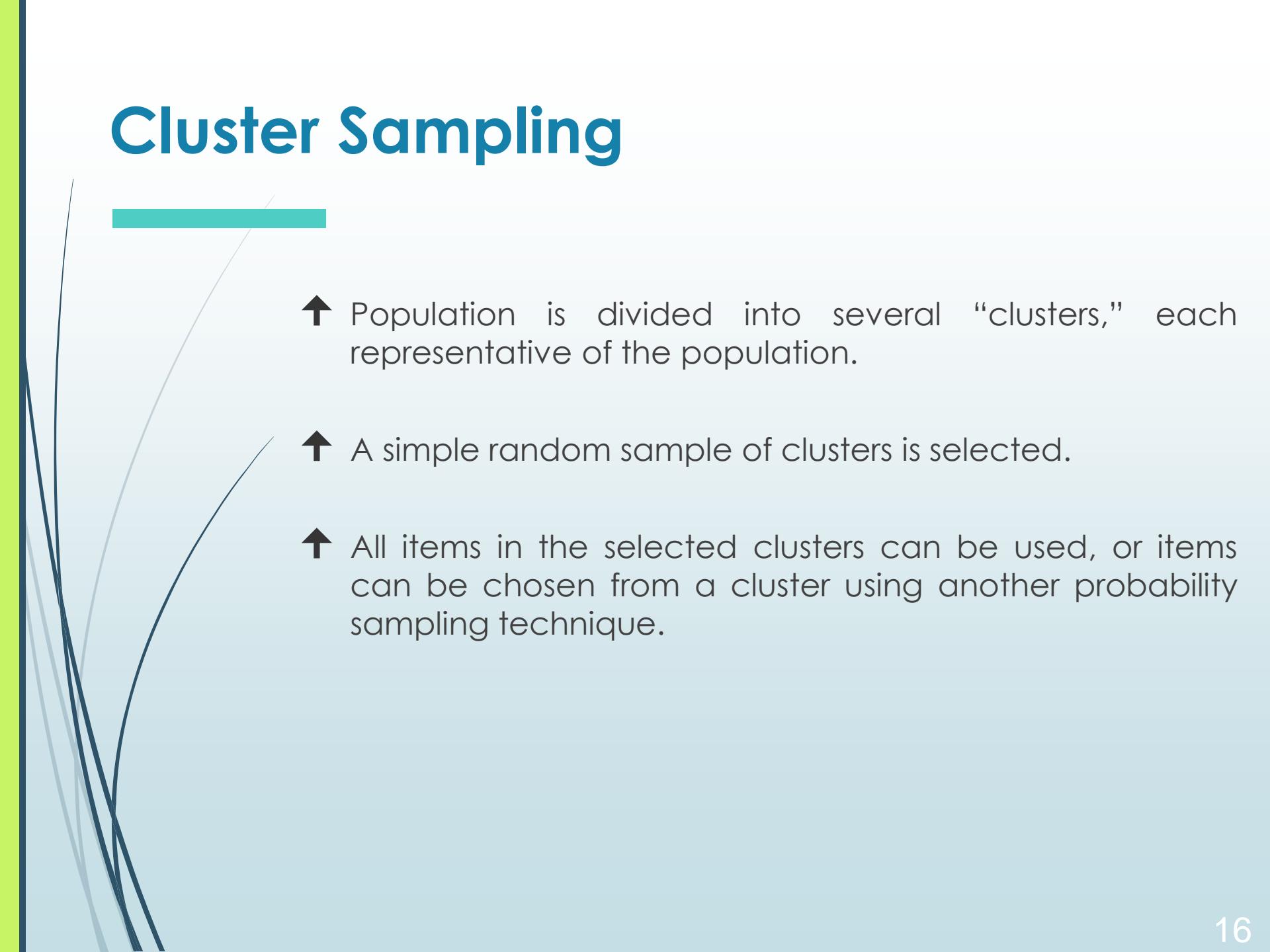
# Systematic Sampling (SYS)

- ↑ Units are selected from the population at regular intervals.
- ↑ A sampling interval ( $k = N/n$ ) and a random start are required.
- ↑ Every  $k^{\text{th}}$  individual thereafter.
- ↑ **Advantages**
  - ← Can result in a sample that is better dispersed than SRS.
  - ← Simpler than SRS.
- ↑ **Disadvantages**
  - ← Can result in a 'bad' sample if the sampling interval matches some periodicity in the population.

# Stratified Sampling

- 
- ↑ Divide population into two or more subgroups (called strata) according to some common characteristic.
  - ↑ A simple random sample is selected from each subgroup, with sample sizes proportional to strata sizes.

# Cluster Sampling



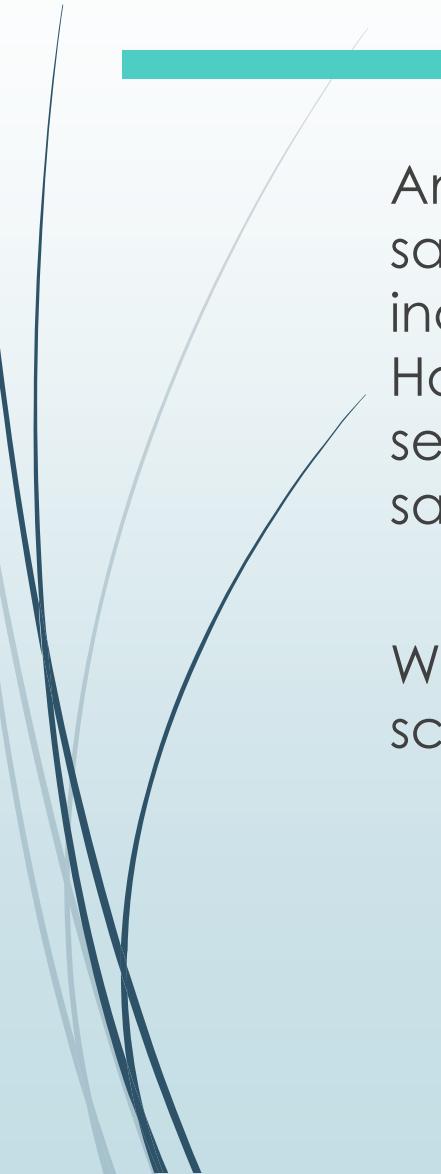
---

- ↑ Population is divided into several “clusters,” each representative of the population.
- ↑ A simple random sample of clusters is selected.
- ↑ All items in the selected clusters can be used, or items can be chosen from a cluster using another probability sampling technique.

# Multistage Sampling

- ↑ With multistage sampling, we select a sample by using combinations of different sampling methods.
- ↑ **Example:-** In Stage 1, we might use cluster sampling to choose clusters from a population. Then, in Stage 2, we might use simple random sampling to select a subset of elements from each chosen cluster for the final sample.

# PROBLEM



An auto analyst is conducting a satisfaction survey, sampling from a list of 10,000 new car buyers. The list includes 2,500 Ford buyers, 2,500 GM buyers, 2,500 Honda buyers, and 2,500 Toyota buyers. The analyst selects a sample of 400 car buyers, by randomly sampling 100 buyers of each brand.

What type of sampling method have used in this scenario?



# Thank You!

## Questins?

---