SNS COLLEGE OF TECHNOLOGY

An Autonomous Institution

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A' Grade

Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

Department of Computer Science and Engineering

Course Code & Title : 23ADO201 - Data Science Fundamentals

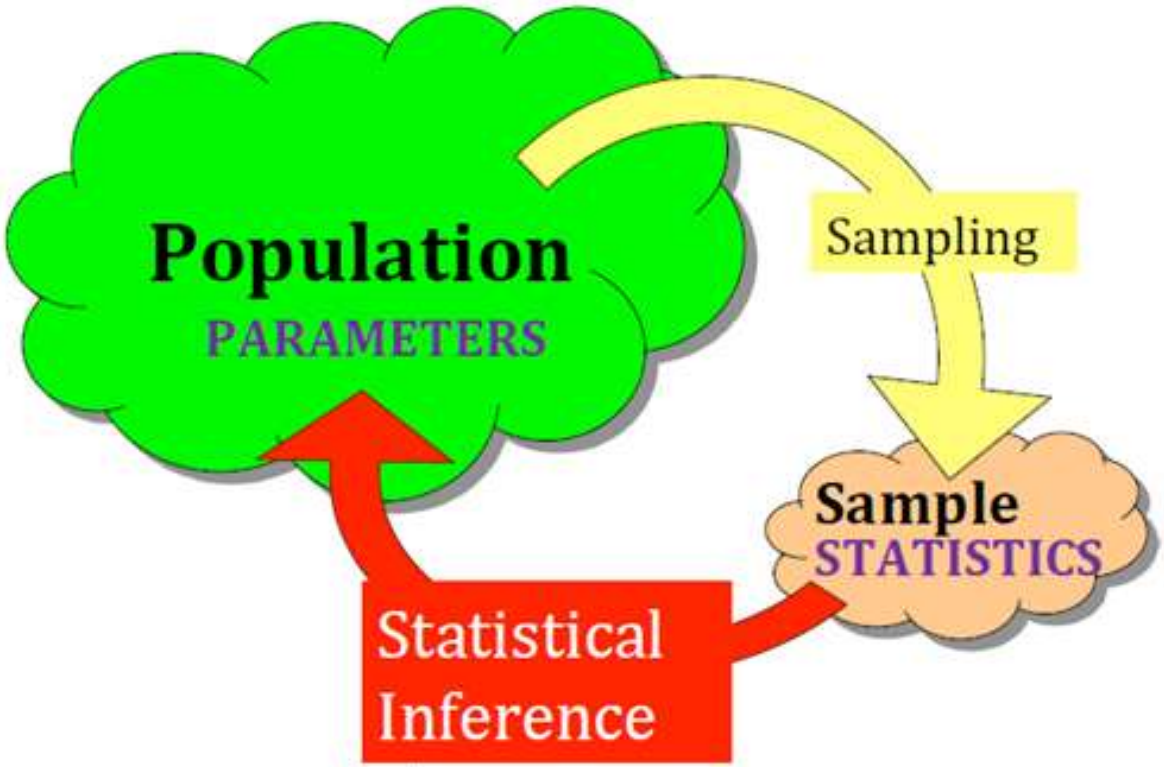
III YEAR / VI SEMESTER - ECE

Unit 3 - INFERENTIAL STATISTICS

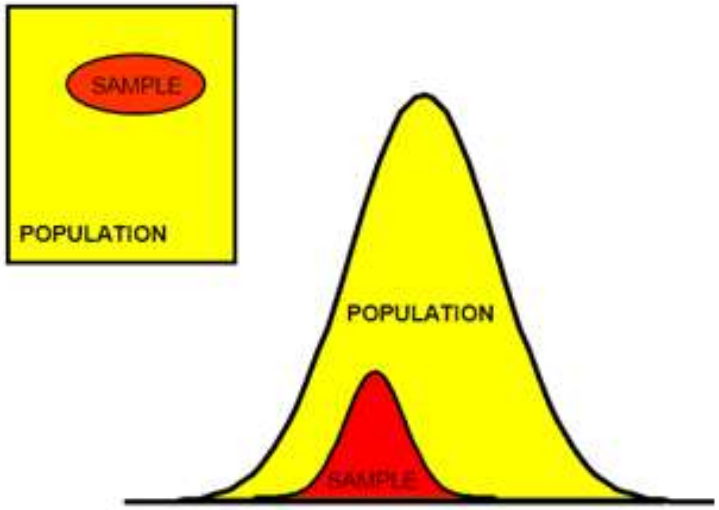
Topic : Populations and samples

K.KARTHIKEYAN AP /CSE

Empathy



- by constructing **confidence intervals** on population parameters
- or by setting up a **hypothesis test** on a population parameter



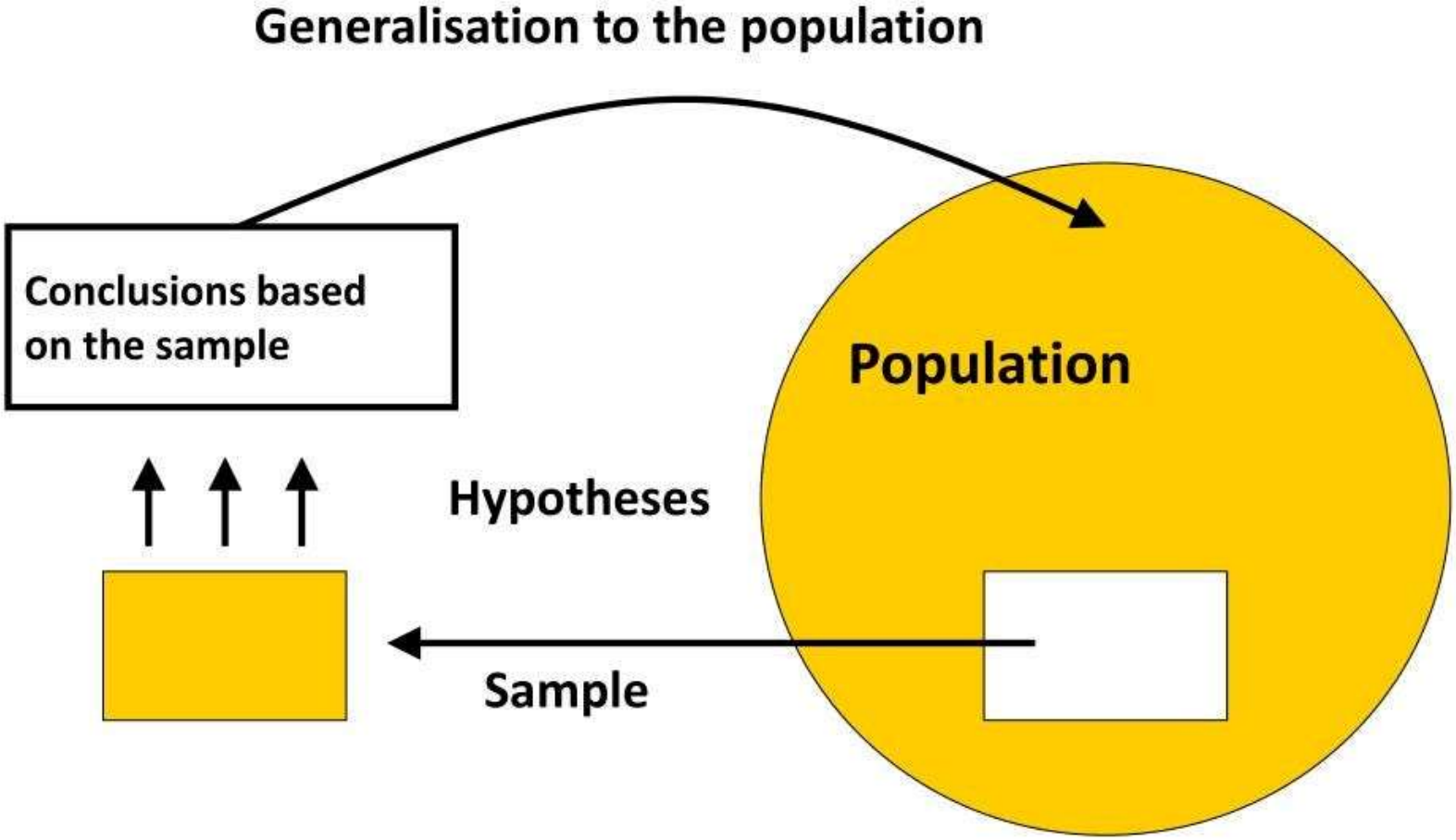
Branches of Statistics

Descriptive Statistics
Involves organizing, summarizing, and displaying data.

e.g. Tables, charts, averages

Inferential Statistics
Involves using ***sample data*** to draw conclusions about a ***population.***

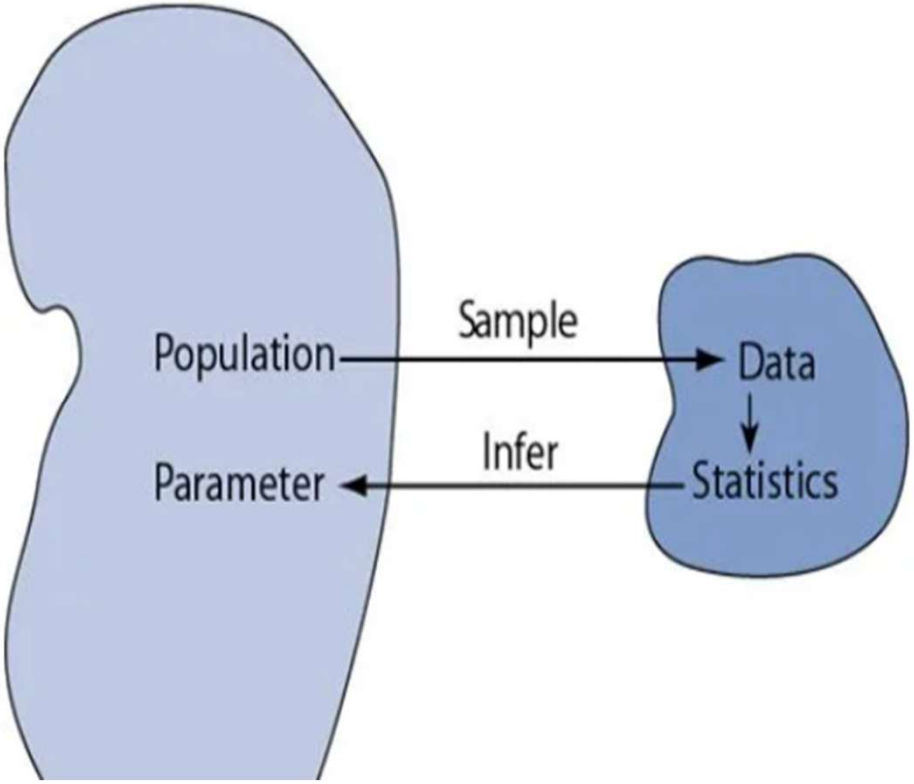
The idea of statistical inference



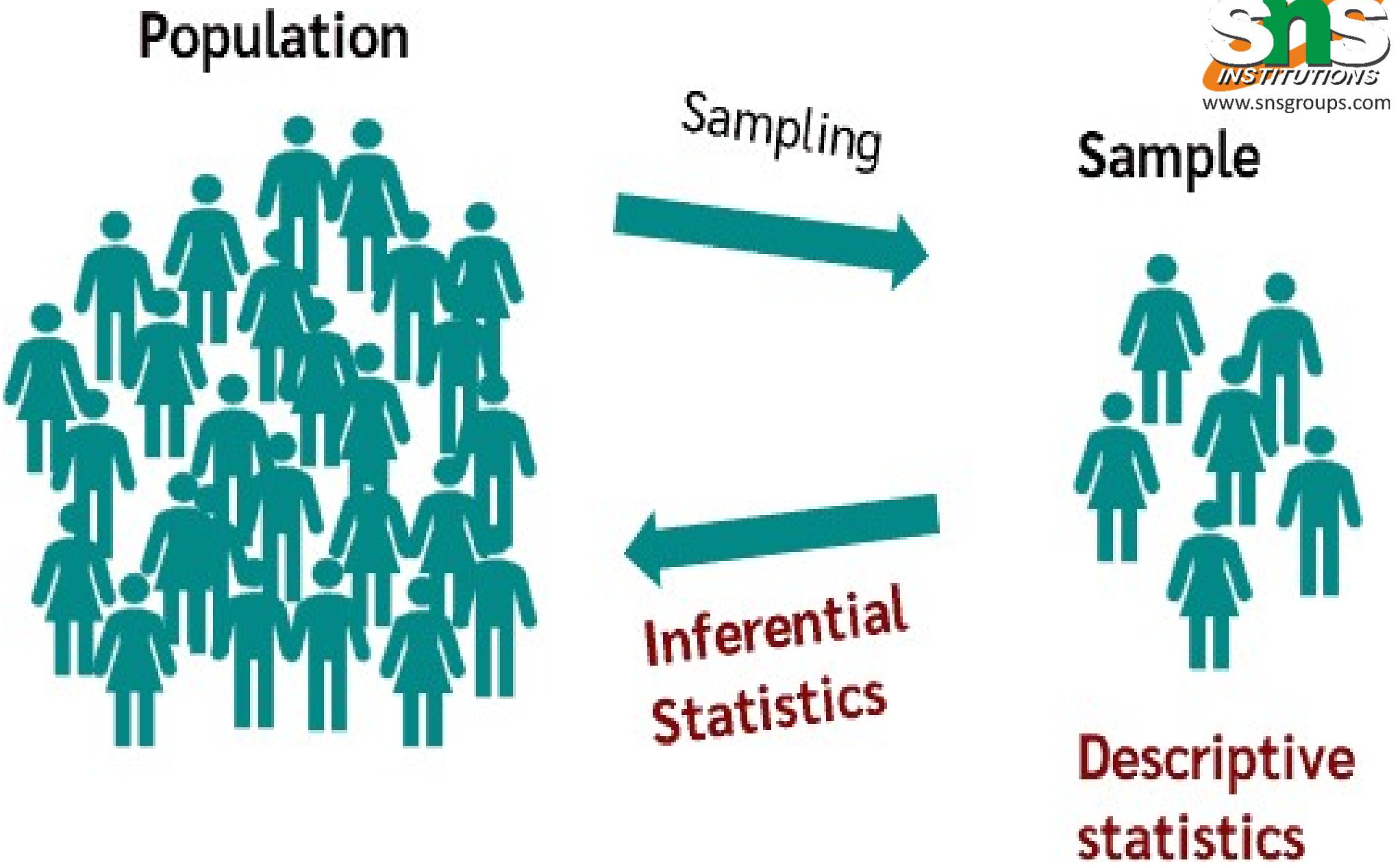
2

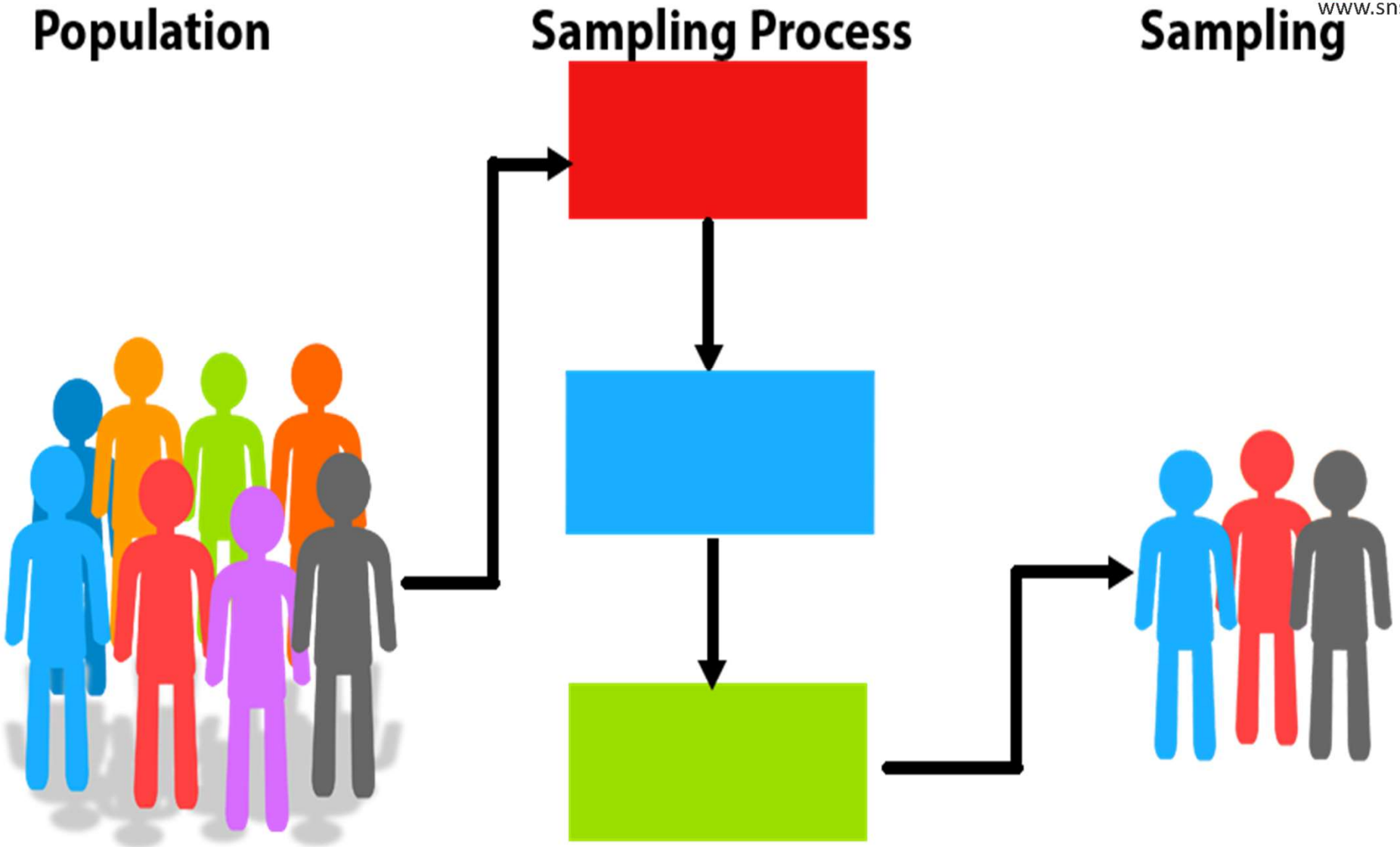
Statistical inference is the act of generalizing from a **sample** to a **population** with calculated degree of certainty.

We want to learn about population *parameters* …

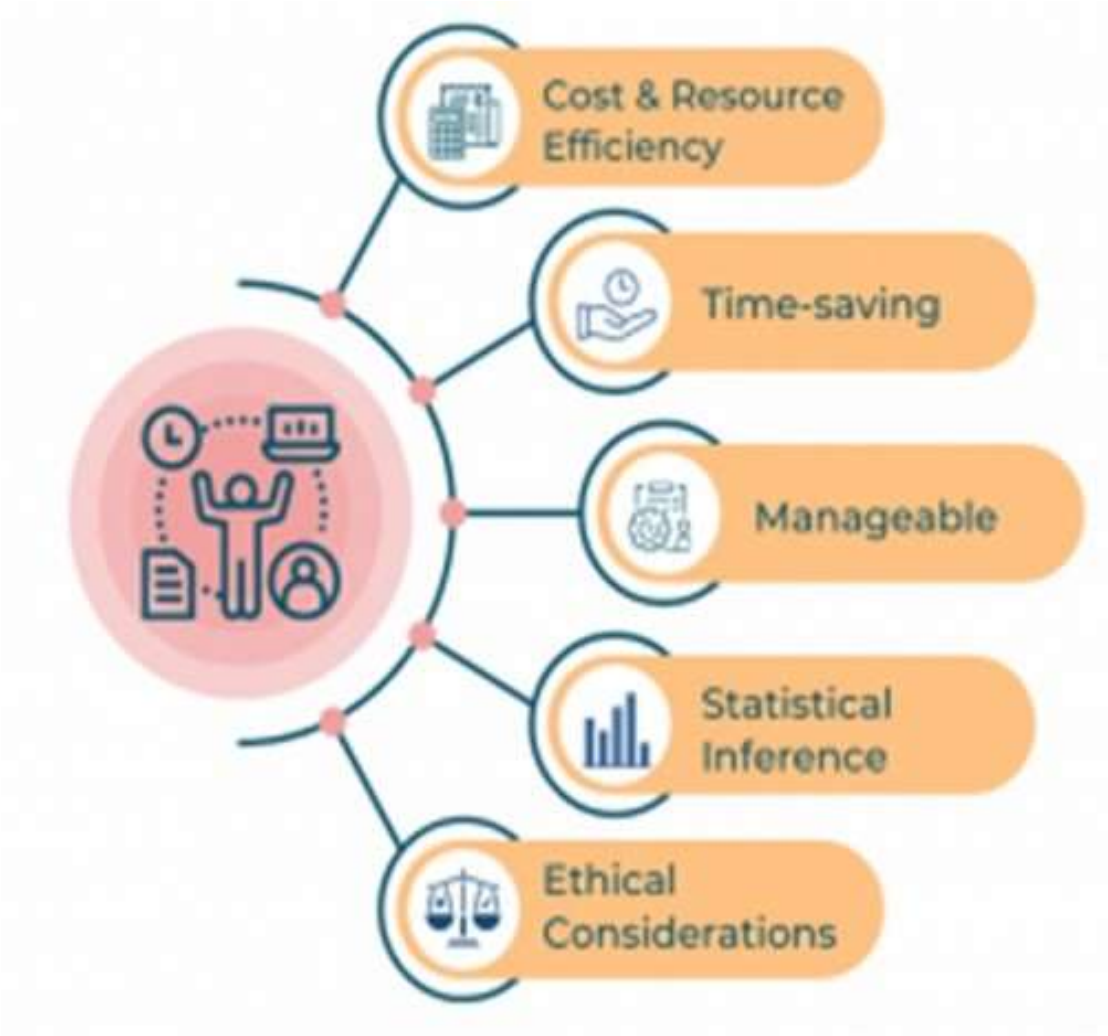


using statistics calculated in the *sample*


Population
Sampling
Sample
Inferential Statistics
Descriptive statistics


Population
Sampling Process
Sampling

Benefits

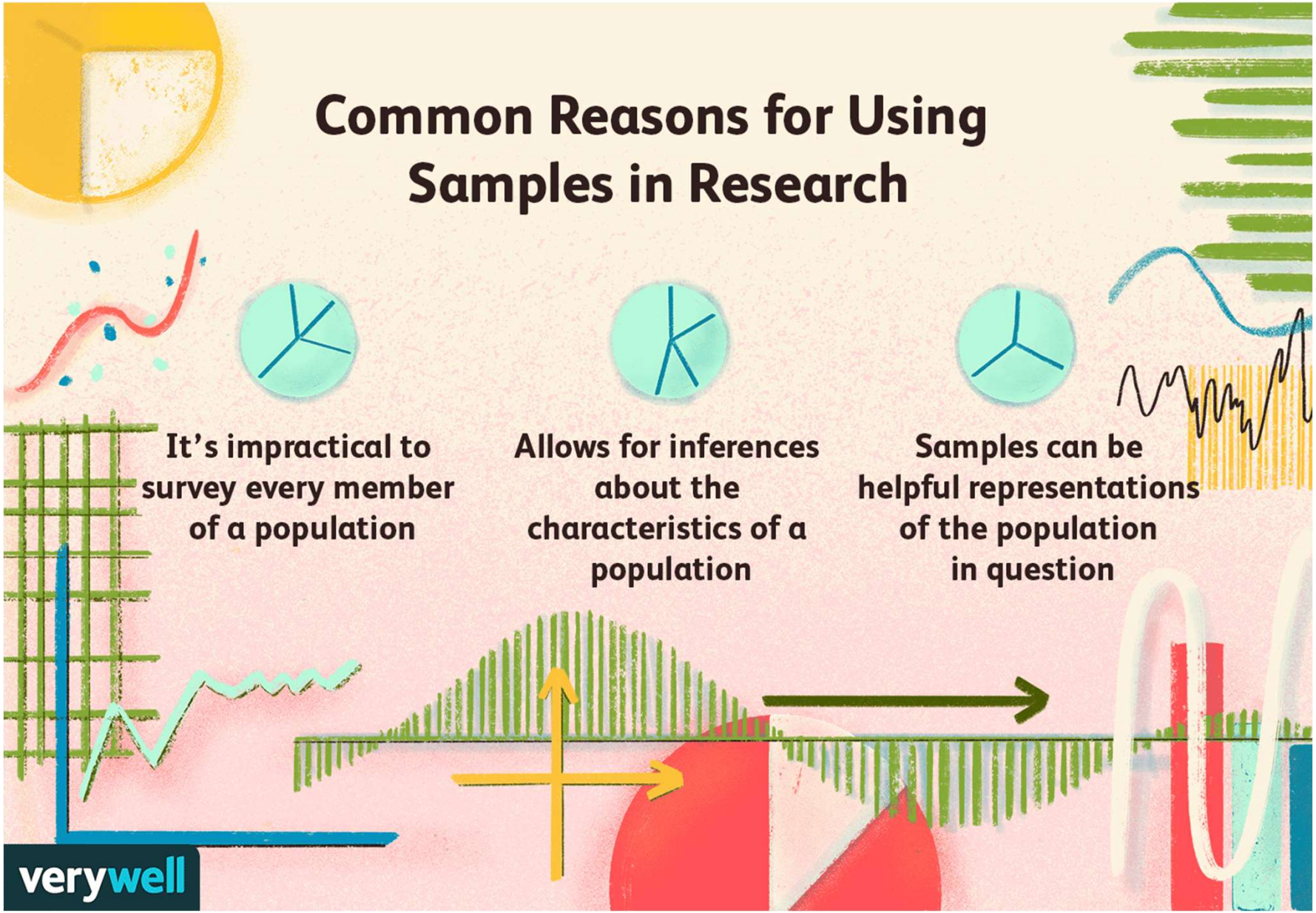

Cost & Resource Efficiency
Time-saving
Manageable
Statistical Inference
Ethical Considerations

Sample

['sam-pəl]

A representative subset of a population.

Investopedia


Common Reasons for Using Samples in Research
It's impractical to survey every member of a population
Allows for inferences about the characteristics of a population
Samples can be helpful representations of the population in question
verywell

Activity

Activity: Identifying Population and Sample

Objective
To help students understand the difference between **population** and **sample**.
To learn how samples are selected from a population for data analysis.
Activity Procedure
Step 1 – Form a Population
The teacher asks all students in the class to stand.
The entire class (for example, **40 students**) is considered the **population**.
Step 2 – Select a Sample
The teacher randomly selects **10 students** from the class.
Step 3 – Collect Data
The selected 10 students are asked a question such as:

How many hours do you study per day?
What is your favorite subject?

Step 4 – Record the Data
Write the responses of the 10 students on the board.
Step 5 – Discussion
Ask students:

What is the **population** in this activity?
What is the **sample**?
Why do we use a sample instead of collecting data from everyone?

Example
Population: All 40 students in the class
Sample: 10 students selected from the class
Conclusion
Students learn that a **population** represents the **entire group**, while a **sample** is a **smaller subset chosen from the population** to analyze and draw conclusions.

MCQ

MCQ questions on Population and Sample:

1. What is meant by population in statistics?

A) A small group selected from a larger group

B) The entire group of individuals or items under study

C) Only numerical data collected in a survey

D) A random method of selecting data

Answer: B) The entire group of individuals or items under study

2. A sample is defined as:

A) The whole group of data

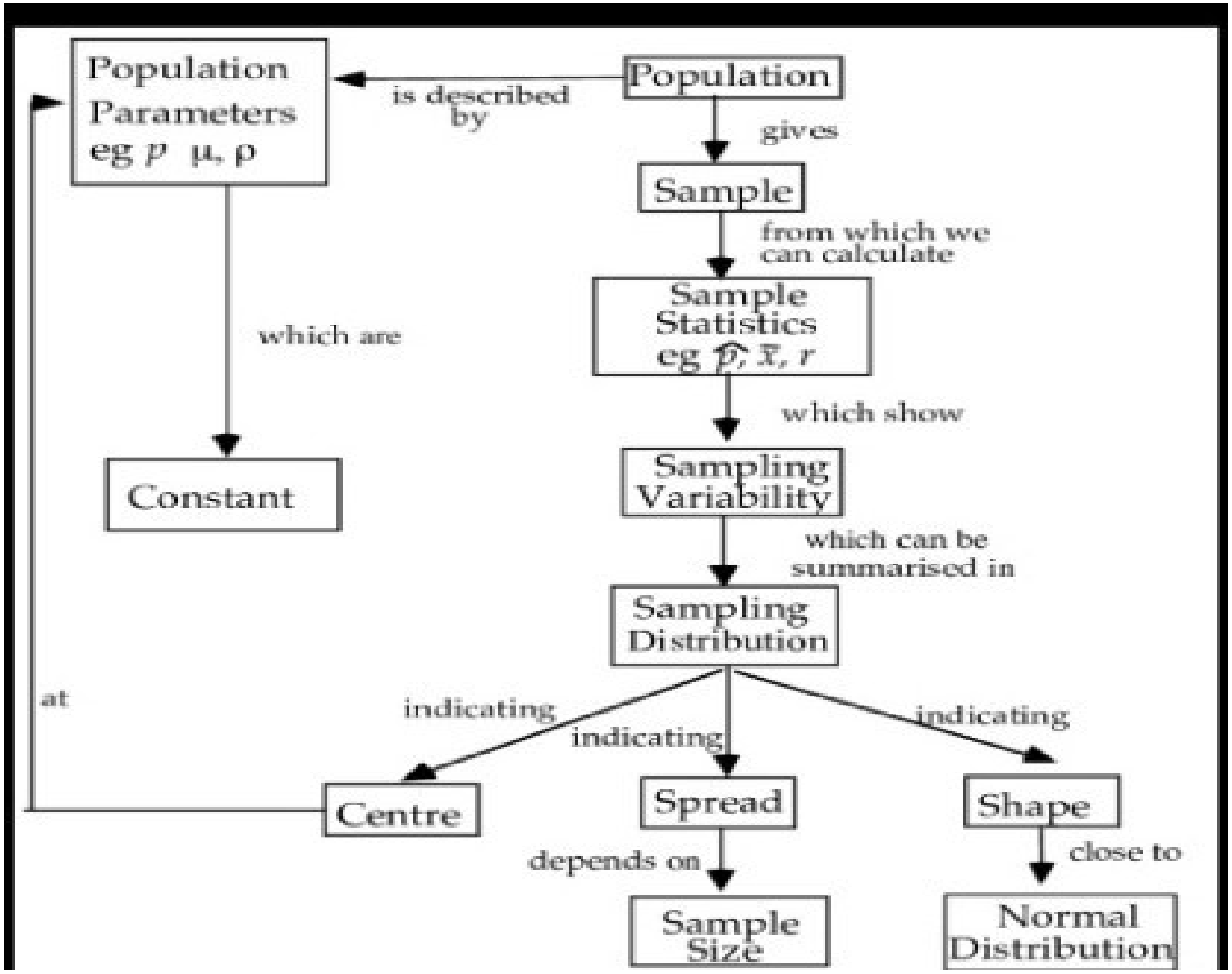
B) A subset of the population

C) Only qualitative data

D) A statistical formula

Answer: B) A subset of the population

MIND MAP


Population Parameters eg p μ, ρ
is described by
Population
gives
Sample
from which we can calculate
Sample Statistics eg $\hat{p}$, $\bar{x}$, r
which are
Constant
which show
Sampling Variability
which can be summarised in
Sampling Distribution
at
indicating
indicating
indicating
Centre
Spread
Shape
depends on
close to
Sample Size
Normal Distribution

REFERENCE BOOKS

1.Allen B. Downey, “Think Stats: Exploratory Data Analysis in Python”, Green Tea Press, 2014.
2.Sanjeev J. Wagh, Manisha S. Bhende, Anuradha D. Thakare, “Fundamentals of Data Science”, CRC Press, 2022.
3.Chirag Shah, “A Hands-On Introduction to Data Science”, Cambridge University Press, 2020.
4.Vineet Raina, Srinath Krishnamurthy, “Building an Effective Data Science Practice: A Framework to Bootstrap and Manage a Successful Data Science Practice”, A press, 2021.


Thank
You