

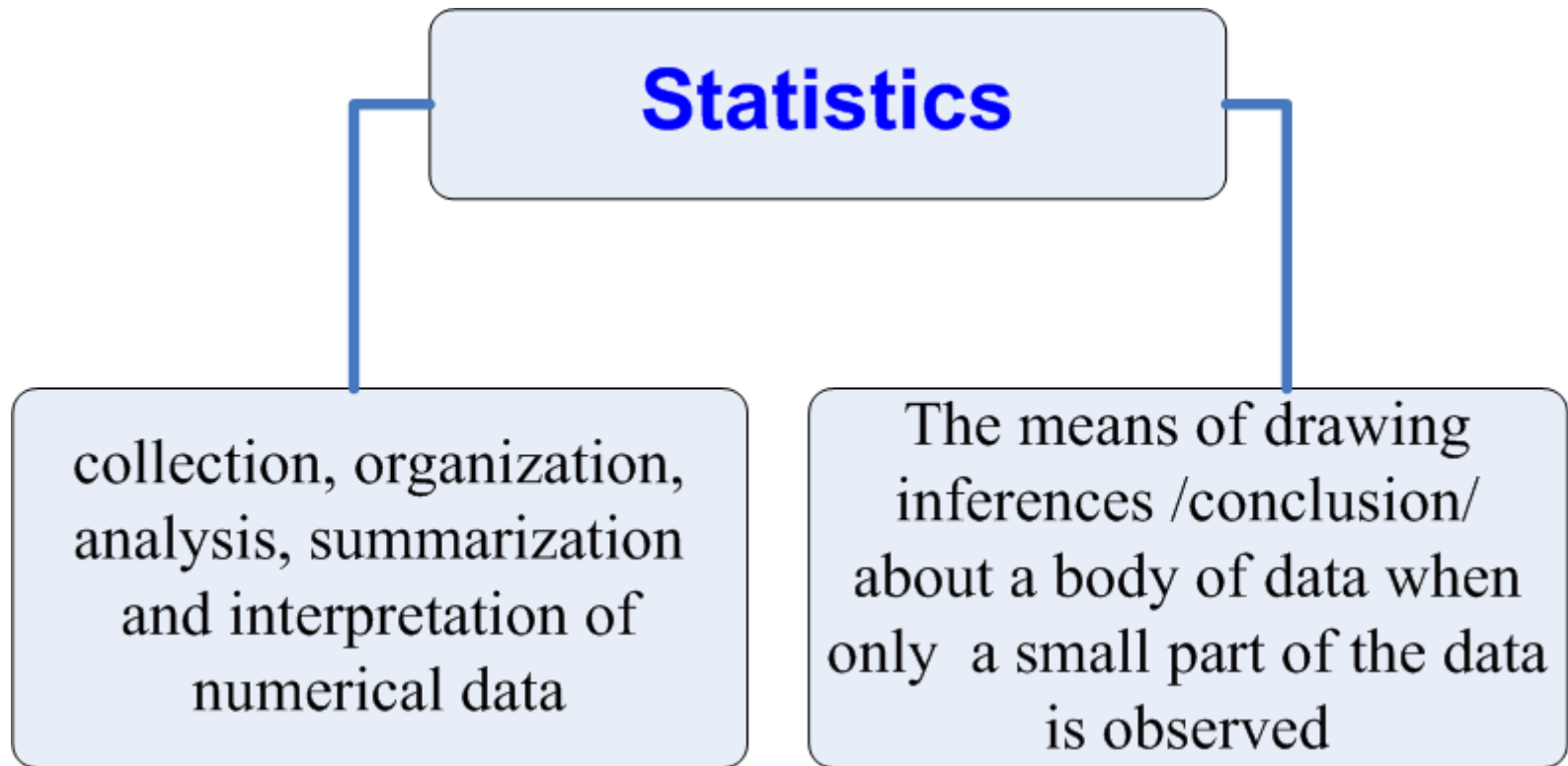
CHAPTER ONE

INTRODUCTION TO BIOSTATISTICS

Discussion points

- What do you think the definition of statistics?
- What is its application in Health science?

What is statistics?



- The tools of statistics are employed in many fields:
 - business, education, psychology, agriculture, and economics, etc.
- **Biostatistics:** is the segment of statistics that deals with data arising from biological processes or medical study.

- Biostatistics:
 - Concerned with *management* of biological data & the *communication* of information derived from these data
- Has central role in medical investigations
- To get *valid interpretations* for row data.

Uses of biostatistics

- Provide methods of **organizing information**
- Assessment of health status
- Health program evaluation
- Resource allocation
- Magnitude of association
 - Strong vs weak association between exposure and outcome

Uses of biostatistics

- Assessing risk factors
 - Cause & effect relationship
- Evaluation of a new vaccine or drug
 - How effective is the vaccine (drug)?
 - Is the effect due to chance or some bias?
- Drawing of inferences
 - Information from sample to population

Types of Statistics

1. Descriptive statistics:

- Ways of **organizing** and **summarizing** data
- Helps to identify the general **features** and **trends** in a set of data and extracting useful information
- Also very important in presenting the final results of a study
- **Example:** tables, graphs, numerical summary measures

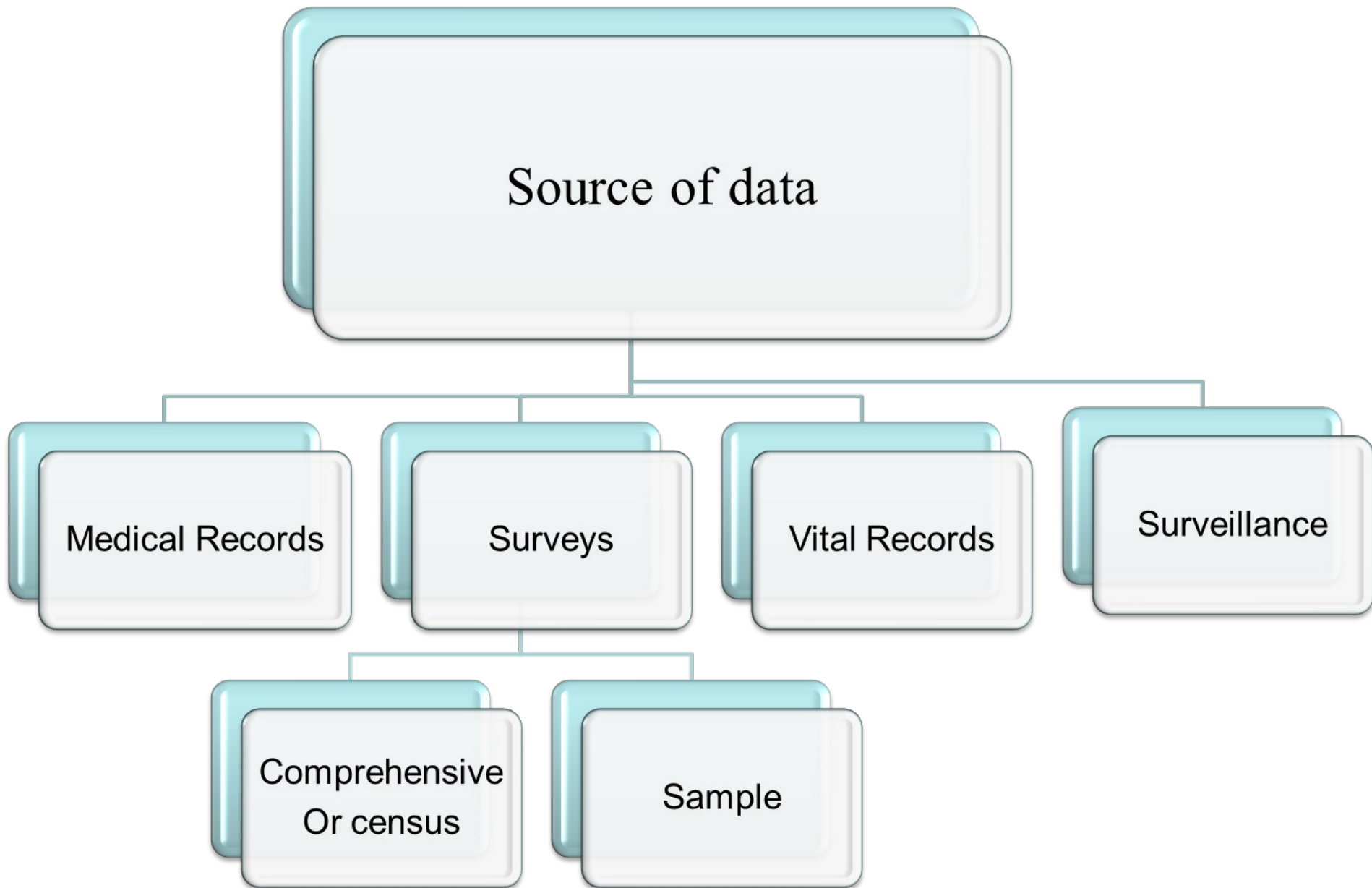
Types of Statistics

2. Inferential statistics:

- Methods used for drawing conclusions about a population based on the information obtained from a sample of observations drawn from that population
- **Example:** Principles of probability, estimation, confidence interval, comparison of two or more means or proportions, hypothesis testing, etc.

Data

- **Data** is the input in the application of statistics
- Is the raw material for statistics
- **Data** are numbers which can be measurements or can be obtained by counting



Types of Data

- 1. Primary data:** collected from the items or individual respondents **directly** by the researcher for the purpose of a study.
- 2. Secondary data:** which had been collected by certain people or organization, & statistically treated and the information contained in it is used for other purpose by other people

Population and Sample

- **Population:**

- Refers to any collection of objects
- is the largest collection of values of a random variable for which we have an interest at a particular time.
 - E.g. Children involved in EPI campaign.

- **Target population:**

- A collection of items that have something **in common** for which we wish to draw conclusions at a particular time.
 - E.g., All hospitals in Ethiopia
- The whole group of interest

Population and Sample

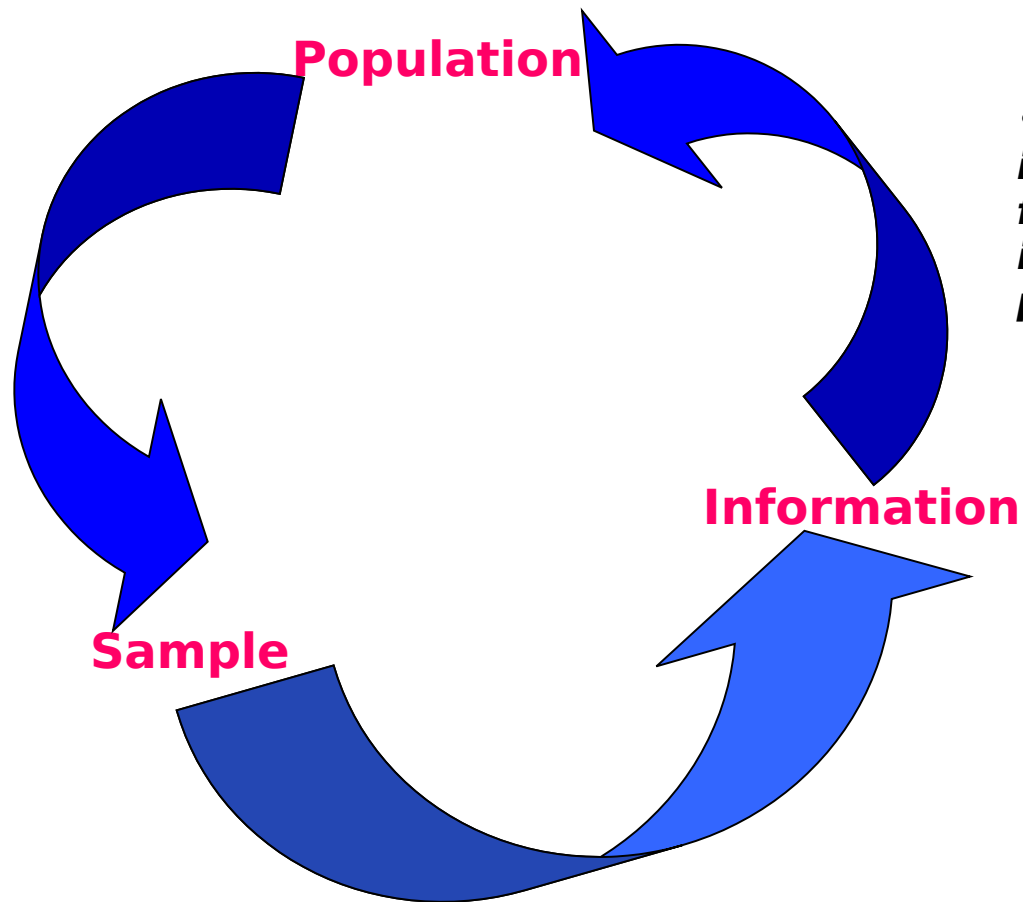
Sampled Population:

- The subset of the target population that has at **least some chance of being sampled**
- The specific population group from which samples are drawn and data are collected

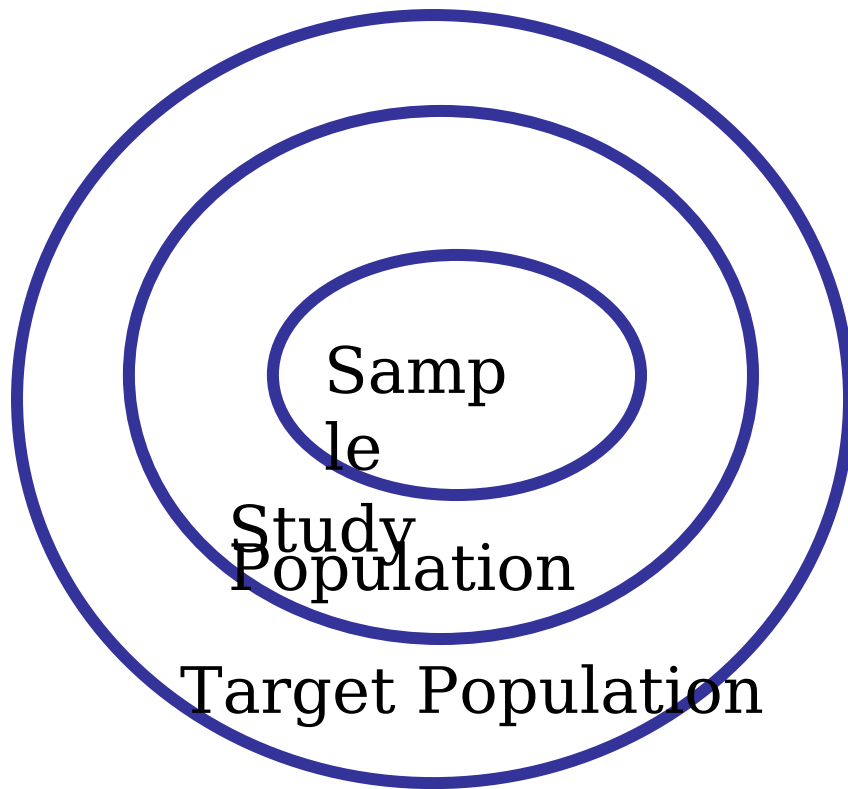
Population and Sample

Sample:

- . A subset of a study population, about which information is **actually** obtained.
- . The individuals who are actually measured and comprise the actual data.



- *Role of statistics in using information from a sample to make inferences about the population*



E.g.: In a study of the prevalence of HIV among adolescents in Ethiopia, a random sample of adolescents in Lideta Kifle Ketema of AA were included.

Target Population: All adolescents in Ethiopia

Study/sampled population: All adolescents in Addis Ababa

Sample: Adolescents in Lideta Kifle Ketema who were included in the

Generalizability

- Is a two-stage procedure:
- We need to be able to generalize from:
 - the sample to the study population, &
 - then from the study population to the target population
- If the sample is **not representative** of the population, the conclusions are restricted to the sample & don't have general applicability

Parameter and Statistic

- **Parameter:** A descriptive measure computed from the data of a population.
 - E.g., the mean (μ) age of the target population
- **Statistic:** A descriptive measure computed from the data of a sample.
 - E.g., sample mean age ()