



PES
UNIVERSITY

CELEBRATING 50 YEARS

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Unit 1: Population & Sampling

Mamatha.H.R

Department of Computer Science and
Engineering

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Unit 1: Population & Sampling

Mamatha H R

Department of Computer Science and Engineering

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Topics to be covered

- ❖ Statistical Analysis
- ❖ Population
- ❖ Sample
- ❖ Sampling
- ❖ Types of Population

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Problems to be solved



Suppose, you are interested in finding

- Mean height of all male students of all the universities in India.
OR
- Average marks of all female students of PES University. OR
- Relationship between the time a student spends on studying and the grades that he gets. OR
- Impact of rise in number of student assignments on their grades.

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

What is Statistical Analysis?

It's the science of **collecting, exploring and presenting large amounts of data to discover underlying patterns and trends.**

Statistics are applied every day – in research, industry and government – to become more scientific about decisions that need to be made.

The basic idea behind all statistical methods of data analysis is to **make inferences** about a **population** by **studying a relatively small sample** chosen from it.



MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Population

A population is the **entire collection of objects or outcomes** about which information is sought.

As mentioned, statistical methods are based on the idea of analyzing a sample drawn from a population.

For this idea to work, **identifying the population, sample** and **choosing the sample** in an appropriate manner becomes important.

In research, a population doesn't always refer to people. It can mean a group containing elements of anything you want to study, such as objects, events, organizations, countries, species, organisms, etc.



Types of population

1. Tangible or concrete population

2. Conceptual population

Tangible population

Populations where the **members are physical objects**, such as cars, bolts, apples, etc., are called **tangible or concrete populations**.

Such populations are assumed to be **always finite** and therefore **involves counting**.

After an item is sampled, the population size decreases by 1.

In principle, one could in some cases return the sampled item to the population, with a chance to sample it again, but this is rarely done in practice.



Conceptual population

Populations that **do not consist of physical or actual objects** are called **Conceptual populations**.

Conceptual populations are mostly the **result of a measurement**.
It involves **measuring something multiple times**.

Ex: length of a metal rod.

It consists of a **not well-defined group** of which all elements are not available at the time the sample is collected as **the population increases every day**.

The **size** of a conceptual population is **usually large**.

Ex: a measuring scale population can be all the possible outputs it can give. i.e. infinite. The measured values can be thought of as a sample from this infinite population.

Tangible & Conceptual population examples

Define the population, and state whether it is tangible or conceptual.

- A shipment of bolts is received from a vendor. To check whether the shipment is acceptable with regard to shear strength, an engineer reaches into the container and selects 10 bolts, one by one to test.

Ans: All the bolts in the shipment: Tangible population

- The resistance of a certain resistor is measured 5 times with the same ohmmeter.

Ans: All measurements that could be made on that resistor with that ohmmeter : Conceptual population

Tangible & Conceptual population examples

Define the population, and state whether it is tangible or conceptual.

- A geologist weighs a rock several times on a sensitive scale.
Ans: All the readings that the scale could produce: Conceptual population
- A pollster samples 1000 registered voters in a certain state and asks them which candidate they support for governor.
Ans: All registered voters in that state : Tangible population

Tangible & Conceptual population examples

Define the population, and state whether it is tangible or conceptual.

- A quality engineer needs to estimate the percentage of bolts manufactured on a certain day that meet a strength specification. At 3:00 in the afternoon he samples the last 100 bolts to be manufactured.

Ans: All bolts manufactured on that day : Tangible population

- In a clinical trial to test a new drug that is designed to lower cholesterol, 100 people with high cholesterol levels are recruited to try the new drug.

Ans: All people with high cholesterol level: Tangible population

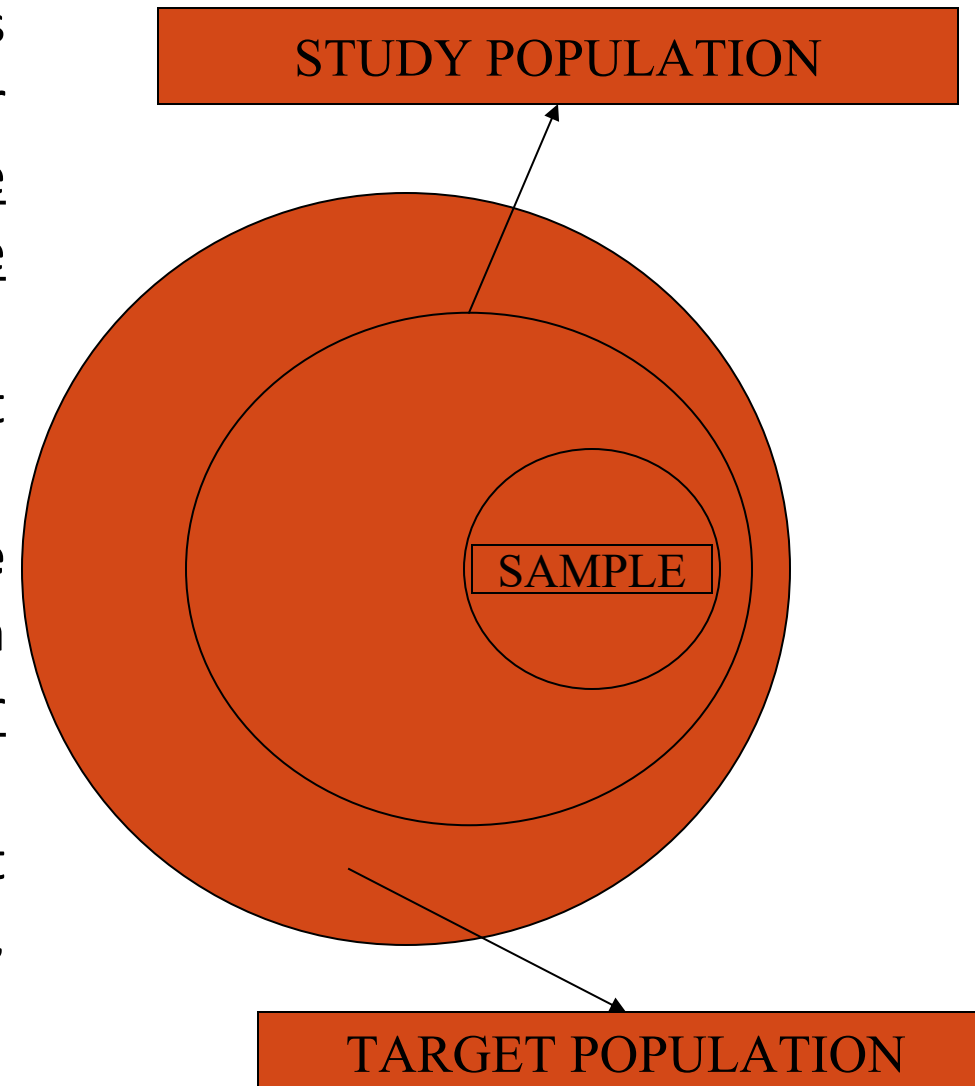
Target and Study population

- **Target or Theoretical population** refers to the entire group of individuals or objects to which researchers are interested in generalizing the conclusions.

It must meet a set of criteria of interest to the researchers.

- **Study population or accessible population** is the population to which the researches can apply their conclusions to.

It is a **subset** of the target population. It may be limited to region, state, city, county, or institution



MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS



Target and Study population examples

Target Population	Study Population
All institutionalized elderly with Alzheimer's	All institutionalized elderly with Alzheimer's in St. Louis county nursing homes
All people with AIDS	All people with AIDS in the metropolitan St. Louis area
All low birth weight infants	All low birth weight infants admitted to the neonatal ICUs in St. Louis city & county
All school-age children with asthma	All school-age children with asthma treated in pediatric asthma clinics in university-affiliated medical centers in the Midwest

Terminologies related to Sampling

- **Target or Theoretical Population:** The population to which the investigator wants to generalize his results.
- **Sampling Frame :** The sampling frame is the list from which the potential respondents are drawn.
Ex: List of Universities, List of Students, List of Airline Companies, Telephone Directory
- **Sampling Unit :** Smallest Unit from which sample can be selected.
- **Sampling Scheme:** Method of selecting sampling units from sampling frame.
- **Sample:** All selected respondents form a sample.

Sampling Breakdown

Study : Find the mean weight of all students of all universities in India.

- Whom do you want to generalize results?
All universities in India
→ **Target or Theoretical population**
- What population can you get access to?
All universities in Karnataka
→ **Study population**
- How can you get access to them?
List of Universities in Karnataka
→ **Sampling frame**
- Who is in your study?
Two Universities from Karnataka
→ **Sample**

1. Which of the following is an example of a tangible population?

- A) All diabetic patients treated at a hospital in 2024
- B) All future patients who will develop diabetes
- C) All possible outcomes of a clinical trial simulation
- D) All hypothetical patients with a rare disease

2. Which of the following best describes a conceptual population in a vaccine study?

- A) Children vaccinated in a specific village last year
- B) All batches of vaccine produced in a factory
- C) All future children who will receive the vaccine
- D) Children currently enrolled in the study

3. Which of the following is not a tangible population?

- A) Students enrolled in Class 10 in 2025
- B) Blood pressure readings from a clinical trial
- C) All cancer patients who will be diagnosed next year
- D) Employees in a specific company today

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Answers

1. A
2. C
3. C



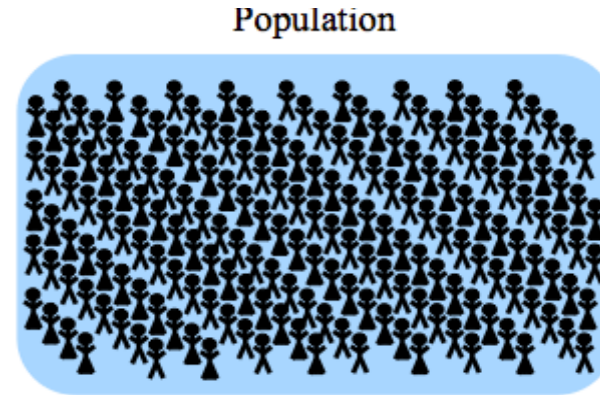
MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Sample

A sample is a **subset of a population**, containing the objects or outcomes that are actually observed.

Sample size: The number of items in a sample is called a sample size. The **size of the sample** is **always less than** the **total size of the population**.

The process of taking a predetermined number of observations from a larger population is called **sampling**.



MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Population vs Sample

Population	Sample
The population is a complete set .	The sample is a subset of the population
Population is hard to define and observe in real life.	A sample is much easier to contact and observe .
It is time consuming and costly to study a population	It is relatively less time consuming and low cost to study a sample.
Population contains all members of a specified group .	Sample is a subset that represents the entire population .
Reports on a population are a true representation of opinion .	Reports on a sample are have a margin of error .

Populations & Samples

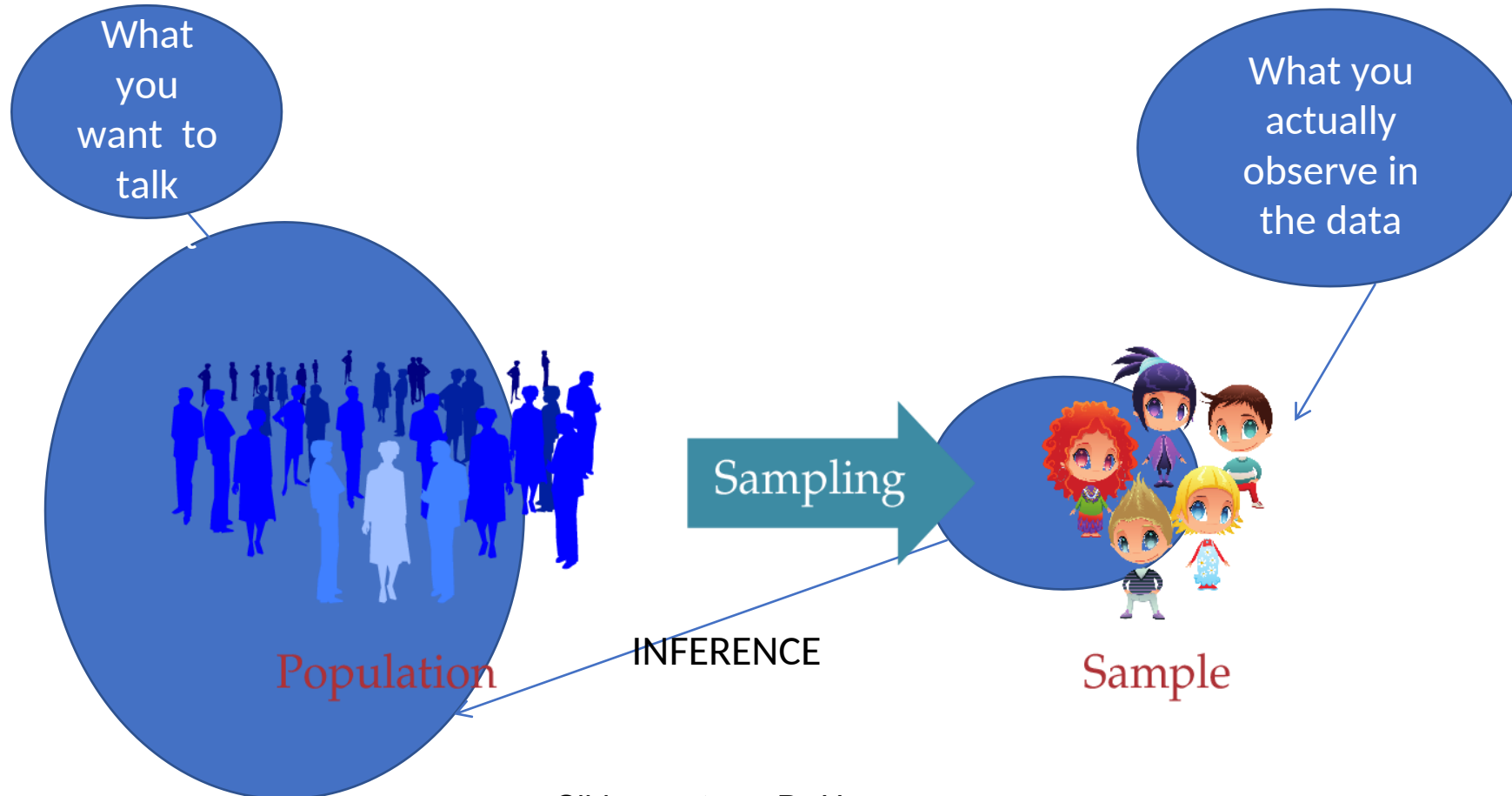
A city council member wanted to know how her constituents felt about a planned rezoning. She randomly selected 75 names from the city phone directory and conducted a phone survey. Identify the population and sample in this setting.

Answer:

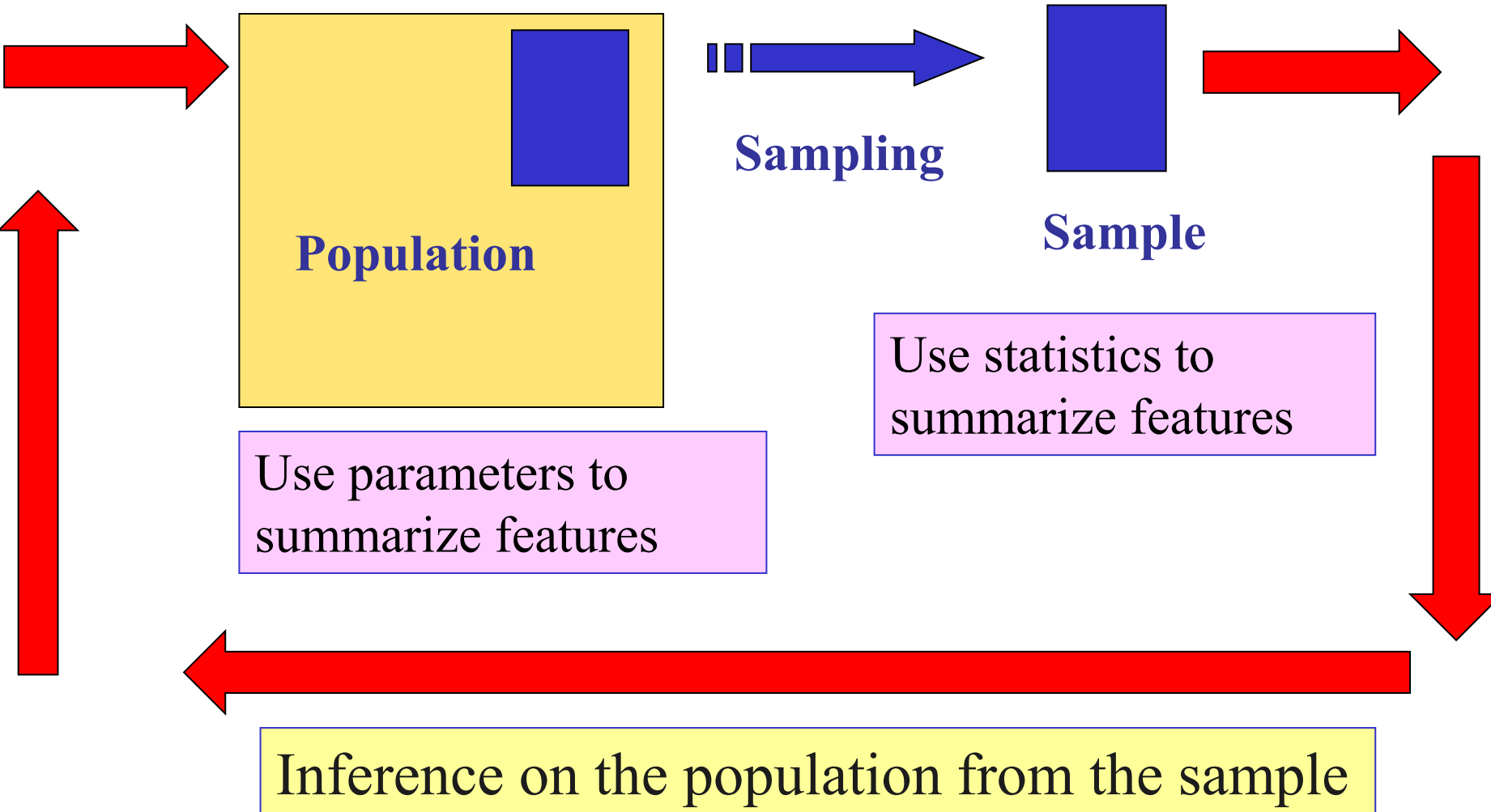
- The population is everyone listed in the city phone directory
- The sample is the 75 people selected to conduct a phone survey.

What is Sampling?

The **process of selecting observations(a sample)** in order to make an inference that can be generalized to the population.



Sampling



Why sampling?

We know that resources such as time, money and people are limited.

When the population is large in size, geographically dispersed, or difficult to contact, it's necessary to use a sample. Thus, most projects aim to gather data from a sample, rather than from the entire population. Some reasons for sampling are:

- **Necessity:** Sometimes it's simply not possible to study the whole population due to its size or inaccessibility.
- **Practicality:** It's easier and more efficient to collect data from a sample.
- **Cost-effectiveness:** There are fewer participant, laboratory, equipment, and researcher costs involved.
- **Manageability:** Storing and running statistical analyses on smaller datasets is easier and reliable.
- **Saves time:** As sample size is relatively less, it increases data-collection speed

Characteristics of a sample

- A sample must be **representative of the population**.
- It must be **appropriately sized**. i.e. it must be sufficiently large to represent the population and provide statistical stability or reliability.
- It must be **unbiased**. It should contain **all types of groups/units** present in the population in **fair proportions**.
- It must be **selected at random**. This means that any item in the group has an equal chance of being and selected and included in the sample.
- It must be **economical**. The objectives of the survey must be achieved in as minimum of cost and effort as possible.
- It must be **goal-oriented**. It must be oriented to the research objectives and fitted to the survey conditions.

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

Is it a good sample?

Study : Survey of the job prospects of the students studying in a university.

Sample: Taking survey from the students who are in Canteen.



Is it a good sample?

This is not an example of a good sample as,

- The students in the canteen are not completely representative of the students studying in the university.
- The size of the sample (i.e. the number of students in the canteen) might not be appropriate or sufficient enough to represent the population (students studying in the university).
- The sample selection is not performed at random as each student studying in the university doesn't have an equal chance of getting selected.

Is it a good sample?

→ **Study** : A test of the effectiveness of a new high school curriculum introduced

Sample : Dividing an area by school district, then choosing a school or set number of schools at random and sampling students from each school.

This type of sampling results in a unbiased sample as it each school district in an area has its representation in the sample. Also, each school has an equal chance of getting chosen.

Is it a good sample?

→ **Study:** Conduct observations to ensure that employees are employing best practices in the company.

Sample: Each employee is assigned a random number using computer software. The same software is used periodically to choose a number of the employees and are observer. This is a good sample as each employee has an equal chance of being selected.

MATHEMATICS FOR COMPUTER SCIENCE ENGINEERS

References

<https://www.questionpro.com/blog/population-vs-sample/>
<https://www.scribbr.com/methodology/population-vs-sample/>

Text Book:

Statistics for Engineers and Scientists, William Navidi.





THANK YOU

Dr.Mamatha H R

Professor, Department of Computer Science

mamathahr@pes.edu

+91 80 2672 1983 Extn 834