

**1. Statistics:** Statistics is the science that deals with the **collection, organization, summarization/presentation, analysis, and interpretation** of data from to assist in making more effective and reasonable decisions. It is an indispensable tool in almost every field, from scientific research to business decision-making.

For example, suppose you want to know the average height of students in your class. **You can measure the height of each student and record the data.** Then you can use descriptive statistics to **calculate the mean, median, mode, range, and standard deviation of the data.** These are numerical summaries that tell you something about the distribution of the data. **You can also use graphs or charts to display the data visually.** For example, you can use a histogram or a box plot to show the shape and spread of the data.

Now suppose you want to know if the average height of students in your class is different from the average height of students in another class. **You can describe the data by using statistical methods and procedures** and **making conclusions based on the analyzed data.**

***Collection:*** Statistics involves the process of gathering data from various sources. This can include surveys, experiments, observations, or existing databases. The goal is to obtain relevant and reliable information that can be analyzed.

***Organization:*** Once the data is collected, it needs to be organized in a systematic manner. This may involve categorizing data into groups or classes, creating tables or charts, or structuring the data in a way that facilitates further analysis.

***Summarization/Presentation:*** After organizing the data, it's important to summarize and present it in a meaningful way. This can be done through statistical measures such as averages, percentages, graphs, or visual representations. Summarizing the data helps in understanding its key characteristics and trends.

***Analysis:*** Statistical analysis involves applying various mathematical and statistical techniques to the collected data. This step aims to explore patterns, relationships, and trends within the data. Statistical analysis helps uncover insights and draw meaningful conclusions from the data.

***Interpretation:*** Once the analysis is done, the results need to be interpreted. This involves drawing conclusions, making inferences, and providing explanations based on the findings of the statistical analysis. It requires understanding the context, limitations, and implications of the data and its analysis.

**2. Types of Statistics:** Statistics can be derived into two areas.

- a) **Descriptive Statistics:** Descriptive statistics are procedures used to summarize, organize, and present a set of data/observations in a meaningful way (e.g., tables, graphs, numerical summaries).

For example, suppose we have the following exam scores (out of 100) for a group of CSE students: 75, 82, 90, 68, 92, 85, 78, 88, 79, 84.

Mean: To find the mean (average) score, we sum up all the scores and divide by the total number of students:

$$\text{Mean} = (75 + 82 + 90 + 68 + 92 + 85 + 78 + 88 + 79 + 84) / 10 = 841 / 10 = 84.1$$

Median: The median is the middle value when the scores are arranged in ascending order. Since we have an even number of observations, the median is the average of the two middle values:

$$\text{Median} = (82 + 84) / 2 = 83$$

Range: The range is the difference between the maximum and minimum scores. In this case:

$$\text{Range} = \text{Maximum score} - \text{Minimum score} = 92 - 68 = 24$$

Standard Deviation: The standard deviation measures the spread or variability of the data around the mean. We can use a formula to calculate it, but for simplicity, let's use a rounded estimate:

$$\text{Standard Deviation} \approx 6.37 \text{ (approximately)}$$

Visualization: We can represent the exam scores graphically using a histogram or a bar chart to visualize the distribution of scores.

- b) **Inferential Statistics:** Inferential statistics is the branch of statistics that uses sample data to make a generalization/conclusion about the population.

For example, suppose we want to investigate whether there is a significant difference in the programming skills between male and female CSE students. We have a random sample of 50 male students and 50 female students from the BRAC university.

**Hypothesis:** Our hypothesis is that there is a difference in programming skills between male and female CSE students.

**Data Collection:** We administer a programming assessment test to both groups of students and record their scores.

### ***Sample Statistics:***

Male Students: Sample mean = 80, Sample standard deviation = 10

Female Students: Sample mean = 85, Sample standard deviation = 8

***Statistical Test:*** We can use a two-sample t-test to determine if the difference in means is statistically significant. The t-test will assess whether the observed difference in means is likely due to chance or if it represents a true difference in programming skills between the two groups.

***Assumptions:*** We assume that the programming scores are normally distributed, and the samples are independent.

**Calculation and Interpretation:** By conducting the t-test, we can calculate the p-value associated with the difference in means. If the p-value is below a predetermined significance level (e.g., 0.05), we can reject the null hypothesis and conclude that there is a significant difference in programming skills between male and female CSE students.

***Conclusion:*** Based on the analysis, if the p-value is less than 0.05, we can infer that there is a statistically significant difference in programming skills between male and female CSE students. This suggests that gender may have an impact on programming abilities within this sample.

In this example, inferential statistics are used to draw conclusions about the entire population of CSE students based on the observed sample. The statistical test helps us make inferences about the population parameters and determine if the observed difference is likely to hold true beyond the sample.

### 3. Difference between Descriptive and Inferential Statistics.

	<b>Descriptive Statistics</b>	<b>Inferential Statistics</b>
<i>Purpose</i>	Describe and summarize the main characteristics, features, or patterns of a dataset or population	Make inferences, predictions, or generalizations about a population based on sample data
<i>Data Analysis</i>	Analyze and interpret the data at hand without making generalizations beyond the observed dataset	Make judgments about the likelihood of observed sample results occurring by chance or representing a true population characteristic
<i>Population vs. Sample</i>	Applicable to both populations and samples	Focuses on drawing conclusions about a larger population based on the analysis of a representative sample
<i>Measures</i>	Uses measures of central tendency, measures of dispersion, and graphical representations to summarize and present the data	Utilizes statistical techniques such as hypothesis testing, confidence intervals, and regression analysis
<i>Examples</i>	Calculating the average income of a population, creating a pie chart to display the distribution of student majors, finding the range of test scores in a class	Conducting a survey to predict election outcomes, testing a new drug to draw conclusions about its effectiveness, estimating the average height of a population based on a sample
<i>Key Goal</i>	Provide a snapshot or summary of the data	Make inferences and draw conclusions about populations based on sample data

#### 4. Example.

Suppose we have a dataset consisting of the programming proficiency scores (out of 100) for 100 CSE students. Average performance level among the male students is 75 and female students is 77. Suppose we want to investigate whether there is a significant difference in programming proficiency between CSE students who attended coding boot camps and those who did not.

\*\*\* Define the term statistics. What are the two basic types of statistics? Describe them and give examples.

\*\*\* How would you classify statistics? / What are the key steps involved in the statistical process? Discuss.

\*\*\* Explain difference between descriptive and inferential statistics with example.

\*\*\* Discuss the importance and uses of statistics.

\*\*\* What is the importance of studying statistics in engineering?

\*\*\* Explain how statistics plays an important role in Computer Science and Engineering field.

**Ans:** Statistics plays a crucial role in the field of Computer Science and Engineering (CSE) in several ways. Here are some key areas where statistics is important in CSE:

- a) **Data Analysis:** CSE involves working with large amounts of data, and statistical techniques are essential for analyzing and making sense of this data. Statistical methods help in extracting meaningful information, identifying patterns, and drawing insights from data sets.
- b) **Machine Learning and Artificial Intelligence:** Statistics forms the foundation of machine learning and artificial intelligence algorithms. Statistical models and techniques are used to train, validate, and optimize machine learning models. These models enable tasks such as image recognition, natural language processing, recommendation systems, and predictive analytics.
- c) **Experiment Design and Evaluation:** In CSE research and development, statistical principles are applied to design experiments and conduct evaluations. Statistical methods help in determining sample sizes, analyzing experimental results, and drawing conclusions about the effectiveness and performance of algorithms, systems, or software solutions.

- d) ***Performance Analysis:*** Statistics is instrumental in analyzing the performance of computer systems and networks. Through statistical techniques, performance metrics and parameters are measured, such as response time, throughput, latency, and reliability. These analyses aid in optimizing system performance, identifying bottlenecks, and improving overall efficiency.
- e) ***Quality Assurance and Testing:*** Statistical methods play a role in quality assurance and software testing. Techniques like hypothesis testing, control charts, and statistical process control are employed to evaluate software reliability, identify defects, and assess the overall quality of software products.
- f) ***Network Analysis and Security:*** Statistical analysis is used in network traffic analysis, intrusion detection, and cybersecurity. Statistical models and algorithms can help identify patterns of abnormal behavior, detect network anomalies, and develop intrusion detection systems.
- g) ***Optimization and Decision Making:*** Statistics is utilized in optimization problems and decision-making processes. Statistical techniques like linear programming, regression analysis, and optimization algorithms help in solving complex problems, optimizing resource allocation, and making data-driven decisions.

In summary, statistics plays a critical role in various aspects of Computer Science and Engineering. It provides the tools and methods necessary to analyze data, develop models, evaluate performance, optimize systems, ensure quality, and make informed decisions in a wide range of CSE applications.

\*\*\* Scope of Statistics in life science and engineering:

Statistics plays a crucial role in both life science and engineering, providing valuable tools and methods for data analysis, decision-making, and research. Here's an overview of the scope of statistics in these fields:

- a) Data Collection and Analysis:* In life science and engineering, vast amounts of data are collected from experiments, observations, and simulations. Statistics helps in organizing, summarizing, and analyzing this data to extract meaningful insights.
- b) Experimental Design:* Statistics guides the planning of experiments, helping researchers design studies that yield reliable and conclusive results while minimizing resource utilization.
- c) Hypothesis Testing:* Statistical tests are used to evaluate hypotheses and determine whether observed differences or effects are statistically significant. This is crucial for drawing valid conclusions from experiments.
- d) Quality Control:* In engineering, statistical process control (SPC) techniques are employed to monitor and improve manufacturing processes, ensuring products meet quality standards.
- e) Reliability Analysis:* Engineers use statistical methods to assess the reliability and durability of products and systems, helping identify and mitigate potential failures.
- f) Risk Assessment:* Statistics plays a role in assessing and quantifying risks in both fields, such as in pharmaceutical development, where it's used to evaluate the safety and efficacy of new drugs.
- g) Biostatistics:* In life sciences, biostatistics is a specialized branch that deals with healthcare data, clinical trials, epidemiology, and medical research, aiding in the development of treatments and healthcare policies.
- h) Machine Learning:* Statistics is a foundational concept in machine learning, which is increasingly applied in life science and engineering for tasks like predictive modeling and pattern recognition.
- i) Big Data Analysis:* With the advent of big data, both fields are adopting advanced statistical techniques and tools to extract insights from massive datasets.
- j) Optimization:* Engineers use statistical optimization methods to improve designs and processes, finding the best solutions given constraints.

- k) Simulation and Modeling:*** Statistics is used to develop and validate mathematical models and simulations, which are essential for understanding complex systems in both fields.
- l) Decision Support:*** Statistics provides a framework for making informed decisions, especially in situations with uncertainty and limited information.
- m) Environmental Studies:*** Statistics is used to analyze environmental data, assess the impact of pollution, and monitor ecological changes.
- n) Pharmaceuticals:*** Statistics is crucial in drug development, from preclinical trials to clinical testing and post-market surveillance.

In summary, statistics is an indispensable tool in life science and engineering, supporting research, quality assurance, decision-making, and problem-solving across a wide range of applications. It helps professionals in these fields make evidence-based decisions, improve processes, and advance scientific knowledge.



**5. Population:** A population is the entire collection of individuals, objects or units whose characteristics are of interest in any particular enquiry.

*Example:* Suppose we are interested in studying about the CGPA distribution of the students in BRACU in a given semester. Here, the collection of all students during the year constitutes the population.

Two types of population:

- a) Finite population: A population consisting of a finite number of units is called finite population.
- b) Infinite population: A population consisting of an infinite number of units is called an infinite population.

**6. Sample:** A small but representative part of the population is called sample.

*Example:* Suppose we are interested in studying about the CGPA distribution of the students in BRACU in a given semester. Here, the collection of all students during the year constitutes the population. Suppose we select students from CSE department only. Then the set of the students constitutes the sample.

\*\*\* Define population and sample with an example.

\*\*\* Distinguish population and sample.

**Definition:** A population is the entire collection of individuals, objects or units whose characteristics are of interest in any particular enquiry.

On the other side, A small but representative part of the population is called sample.

**Size:** The population size can be relatively small or very large, depending on the scope of the study. In some cases, it can be finite, while in others, it may be considered infinite.

On the other side, the sample size is smaller than the population size, but it should be sufficiently large enough to provide reliable and accurate results.

**Inference:** Statistics performed on the entire population is known as descriptive statistics. However, in most cases, analyzing the entire population is impractical, costly, or time-consuming. On the other side, Statistics performed on the sample is known as inferential statistics. The goal of inferential statistics is to generalize the findings from the sample to the entire population

\*\*\* In a health survey, a researcher wants to find the hemoglobin level of the students of BRACU. For this purpose, 400 students were randomly selected and their hemoglobin levels were recorded.

- i. What is the population?
  - ii. What is the sample?
  - iii. What is the variable being measured?
  - iv. What is the nature of variable?
- i. Population: The population in this scenario is the entire group of students at BRACU (BRAC University).
- ii. Sample: The sample is the subset of students from BRACU who were randomly selected and had their hemoglobin levels recorded. In this case, the sample size is 400 students.
- iii. Variable being measured: The variable being measured is the "hemoglobin level" of the students.
- iv. Nature of variable: The hemoglobin level is a continuous numerical variable. It can take any value within a certain range and can be measured with precision, making it a quantitative variable.

\*\*\* In a customer satisfaction survey, a company wants to measure the waiting time experienced by customers in their service centers. They randomly select 200 customers and record their waiting times.

- i. What is the population?
- ii. What is the sample?
- iii. What is the variable being measured?
- iv. What is the nature of variable?

\*\*\* A researcher wants to study the effectiveness of a new medication in reducing blood pressure. They conduct a clinical trial and randomly assign 100 patients to two groups – one receiving the medication and the other receiving a placebo. The blood pressure of each patient is recorded before and after the treatment.

- i. What is the population?
- ii. What is the sample?
- iii. What is the variable being measured?
- iv. What is the nature of variable?

**7. Census:** A census is a systematic collection of **demographics**, **social**, and **economic** data from every member of a population within a specified geographical area at a specific point in time, and providing a comprehensive view of the population's characteristics.

**Census vs Survey:**

	<b>Census</b>	<b>Survey</b>
<i>Purpose</i>	To collect data about an entire population or a large group covering various demographic, social, and economic characteristics.	To collect data from a sample of the population to estimate characteristics and make inferences about the larger population.
<i>Coverage</i>	Covers the entire population or a substantial portion of it.	Covers a selected sample from the population, which is a smaller subset of the whole.
<i>Data Collection</i>	Conducted through direct enumeration or administrative records, aiming to collect data from every individual or household in the target population.	Conducted through interviews, questionnaires, or other data collection methods with a selected sample of individuals or households.
<i>Sample Size</i>	The sample size is equivalent to the population size since it covers the entire population or a significant proportion of it.	The sample size is typically smaller, representing a fraction of the population, usually determined based on statistical considerations.
<i>Representativeness</i>	Provides an accurate and comprehensive picture of the entire population, ensuring a high level of representativeness.	Relies on proper sampling techniques to ensure the sample represents the population accurately, with the level of representativeness depending on the quality of sampling methods.
<i>Data Analysis</i>	Involves processing and analyzing a vast amount of data to produce	Requires statistical analysis techniques to estimate population

	detailed and precise population-level statistics.	characteristics and draw inferences from the sample data to the larger population.
<i>Frequency</i>	Conducted periodically, typically every 5 to 10 years, to capture changes in population characteristics over time.	Can be conducted at various frequencies, depending on the research or monitoring needs. It can be one-time or repeated at regular intervals.
<i>Resource Intensity</i>	Requires substantial resources, including time, manpower, and financial investments, due to its comprehensive nature.	Requires relatively fewer resources compared to a census as it deals with a smaller sample size.
<i>Examples</i>	National censuses conducted by countries to gather demographic and socio-economic data about their entire population.	Surveys conducted by research institutions, government agencies, or market research firms to study specific topics, public opinion, consumer behavior, etc.

**8. Parameter:** A parameter is a numerical value that describes a characteristic of a population. It is a fixed, unknown value that represents the entire population of interest.

Example: The average CGPA of all students in BRACU, the proportion of voters supporting a particular candidate in an election, or the standard deviation of salaries for all employees in a company are examples of parameters.

**9. Statistic:** A statistic is a numerical value that describes a characteristic of a sample. It is a computed value based on the data collected from a subset of the population.

Example: The average CGPA of a sample of 100 students of BRACU, the proportion of voters supporting a candidate based on a survey of 500 people, or the standard deviation of salaries for a randomly selected group of employees are examples of statistics.

**\*\*\* Example:**

The population represents all CSE students at BRACU, while the sample is a subset of 200 students randomly chosen from that population. The parameter refers to the unknown characteristic (e.g., average programming proficiency) of the entire CSE student population, while the statistic represents the computed value (e.g., average programming proficiency) based on the data collected from the selected sample.

	<b>Parameter</b>	<b>Statistics</b>
<i>Definition</i>	Numerical value that describes a characteristic of a population.	Numerical value that describes a characteristic of a sample.
<i>Population Scope</i>	Associated with the entire population of interest.	Derived from a subset of the population, i.e., a sample.
<i>Calculation</i>	Typically, unknown and estimated based on statistics calculated from sample data.	Computed directly from the sample data.
<i>Inference</i>	Used to make inferences or draw conclusions about the population.	Used to estimate population parameters and generalize conclusions from the sample.
<i>Observability</i>	Often not directly observable or measurable.	Observable and measurable from the collected sample data.
<i>Examples</i>	Average height of all adults in a country, proportion of voters supporting a candidate in an election.	Average height of a sample of 100 adults, proportion of voters supporting a candidate based on a survey of 500 people.

**10. Sampling techniques:** Sampling techniques refer to the methods and approaches used to select a subset of individuals or observations from a larger population for the purpose of conducting a study or analysis. These techniques aim to ensure that the selected sample represents the characteristics and diversity of the population in a fair and representative manner.

There are two types of sampling techniques.

- a) Probability sampling
- b) Non-probability sampling

**11. Probability sampling:** Probability sampling is a method of selecting a sample from a population where the selection is based on the principle of randomization. This means that the selection process relies on chance or random selection. This type of sampling ensures that every individual or observation in the population has an equal or known probability of being included in the sample.

There are few probability sampling techniques:

- a) Simple random sampling
- b) Systematic sampling
- c) Stratified sampling
- d) Cluster sampling
- e) Multi-stage sampling

**13. Non probability sampling:** Non probability sampling involves selecting units from a population using a subjective, non-random method. Unlike probability sampling, non-probability sampling does not require a complete survey frame. It is a quick, convenient, and cost-effective approach to gather data.

Few non-probability sampling techniques are,

- a) Convenience sampling
- b) Snowball sampling
- c) Purposive sampling
- d) Quota sampling

**14. Difference between PS and NPS:**

	<b>Probability Sampling</b>	<b>Non-Probability Sampling</b>
<i>Definition</i>	Probability sampling is a method of selecting a sample from a population where the selection is based on the principle of randomization.	Non probability sampling involves selecting units from a population using a subjective, non-random method.
<i>Random Selection</i>	Random selection is used to ensure each member of the population has an equal chance of being selected.	Random selection is not utilized, and participants are chosen based on convenience, judgment, or specific criteria.
<i>Representativeness</i>	Provides a higher level of representativeness, as each member of the population has an equal chance of being included in the sample.	May not provide the same level of representativeness, as the sample may not accurately reflect the population due to non-random selection methods.
<i>Statistical Analysis</i>	Allows for the use of statistical techniques to estimate population parameters, measure variability, and make statistical inferences.	Statistical analysis may be limited or less robust due to potential biases and lack of random selection.
<i>Examples</i>	Simple random sampling, stratified sampling, cluster sampling, systematic sampling.	Convenience sampling, purposive sampling, quota sampling, snowball sampling.

**15. Variable:** If the values or categories of a characteristics vary from unit to unit, the characteristics is called a variable.

For example,

- i. Gender is a variable which is composed of two categories, Male and Female.
- ii. Number of attendances during 24 working days is a variable which is composed of the values 0,1,2,3, ...,24
- iii. Height of a student is a variable

Variables can be classified into two types:

- a) ***Qualitative or Categorical variable***: A variable is said to be qualitative if its values cannot be measured inherently on a numerical scale. Each unit only be classified into one of a group of categories. A categorical variable does not have units.

For example, marital status is a qualitative variable. It cannot be measured on a quantitative scale; it can only be categorized as unmarried, married, widowed, divorced etc.

- b) ***Quantitative variable***: A variable is said to be quantitative if its values are measured inherently on a numerical scale. An essential part of a quantitative variable is its unit.

For example, number of family members (1,2,3, ...), age of students etc.

- i. ***Discrete variable***: A variable is said to be discrete if it takes only the isolated or countable values within a given range.
- ii. ***Continuous variable***: A variable is said to be continuous if it can take any value on some interval. For example, CGPA.

**16. Data:** Data are some information that has been collected from any field of investigation and translated into a form that is efficient for processing.

There are two types of data,

- a) ***Qualitative data***: Observed values of a qualitative variable are called qualitative data.
- b) ***Quantitative data***: Observed values of a quantitative variable are called quantitative data.
  - i. Discrete data
  - ii. Continuous data



17. **Sources of data:** There are two sources of getting statistical data,

- a) ***Primary data:*** The primary data are those which are collected a fresh and for the first time, and thus happen to be original in character.

For example, suppose a researcher wants to investigate the factors influencing the choice of programming languages among CSE students. He collects information from 100 students. This is primary data.

- b) ***Secondary data:*** Secondary data is data which has already been collected by individuals. Secondary data refers to existing data that has been collected by someone else or previously for a different purpose. Secondary data can be collected from publications, newspapers, government reports and statistics, industry reports and surveys and so on.

**19. Scales of measurement:** Scales of measurement, also known as levels of measurement or data measurement scales, refer to the different ways in which variables or data can be categorized or measured. The scale of measurement determines the mathematical operations that can be performed on the data and the level of information conveyed by the data. There are four commonly recognized scales of measurement:

a) **Nominal:** The measurement level in which numbers or symbols are assigned to the categories. The categories are distinct, mutually exclusive and exhaustive. The primary characteristics of a nominal level of measurement are that order, addition, subtraction, multiplication, or division is not meaningful.

b) **Ordinal:** The measurement level in which numbers are assigned to the categories as well as ranking called an ordinal level of measurement. The primary characteristics of an ordinal level measurement are that the rank or order is valid, but addition, subtraction, multiplication, or division is not meaningful.

Examples of ordinal variables include ratings or rankings (e.g., movie ratings from 1 to 5 stars, academic grades of A, B, C), or satisfaction levels (e.g., very satisfied, satisfied, dissatisfied).

c) **Interval:** The measurement level in which numbers are assigned to the variable values in such a way that measurement has order and distance properties but not an absolute zero value (it does not have a true zero point) is called an interval scale of measurement. The characteristics of an interval level measurement are that the rank or order as well as addition or subtraction is valid, but multiplication, or division is not meaningful.

For example, a temperature of 0°C doesn't indicate the absence of temperature-there is still a temperature.

d) **Ratio:** The measurement level in which numbers are assigned to the variable values in such a way that measurement has order, distance, and an absolute zero property is called a ratio level of measurement. The characteristics of a ratio measurement are that the rank or order, addition or subtraction as well as multiplication or division is valid.

**20. A comparison of the levels of measurement.**

Levels	Property			Example
	Order	Difference	Ratio	
Nominal	No	No	No	Gender: Male, Female
Ordinal	Yes	No	No	Wealth index: Poor, Middle, Rich
Interval	Yes	Yes	No	Temperature
Ratio	Yes	Yes	Yes	Height

\*\*\* Classify each of the following as different levels of measurement indicating underlying logic: Gender, Marital Status, Eye Color, Blood Type, Country of Origin, Vehicle Type, Hair Color, Educational Level, Animal Species, Political Affiliation, Likert Scale, Socioeconomic Status, Customer Satisfaction, Performance Rating, Pain Level, Rank in a Competition, Temperature in Celsius, IQ Scores, Age, Height, Weight, Distance, Time, Income.

\*\*\* What is the scale of measurement for a variable that classifies people into categories such as "Male," "Female," and "non-binary" based on gender?

\*\*\* Identify the scale of measurement for a variable that measures a person's level of education as "High School," "Bachelor's Degree," "Master's Degree," and "Ph.D".

\*\*\* In a survey, respondents rate their satisfaction with a product using the options "Very Dissatisfied," "Dissatisfied," "Neutral," "Satisfied," and "Very Satisfied." What is the scale of measurement for this variable?

\*\*\* A researcher records the temperature in degrees Celsius on different days. Identify the scale of measurement for this temperature variable.

\*\*\* A company collects data on the number of hours employees worked in a week. What is the scale of measurement for this variable?

\*\*\* Suppose a study is conducted to analyze the performance of CSE students in a programming course.

- i. What are the individuals in the study?
- ii. What are some possible quantitative variables of interest?
- iii. What are some possible qualitative variables of interest?

### ***i. Individuals in the study:***

The individuals in this study would be the CSE students who are enrolled in the programming course. Each student would be considered an individual, and data will be collected from multiple students to analyze their performance.

### ***ii. Possible quantitative variables of interest:***

- Scores achieved in programming assignments (measured as numerical grades).
- Time taken to complete each programming assignment (measured in hours).
- Number of lines of code written for each assignment.
- Scores obtained in programming quizzes and exams (measured as numerical grades).
- Average number of debugging errors in their code.

### ***iii. Possible qualitative variables of interest:***

- Programming proficiency level of students (e.g., beginner, intermediate, or advanced).
- Programming languages known to the students (e.g., Python, Java, or C++).
- Types of programming projects students have worked on previously (e.g., web development, app development, or data analysis).
- Learning preferences (e.g., self-paced learning, group study, or attending lectures).
- Student satisfaction with the course content and teaching methods (e.g., satisfied, neutral, or dissatisfied).