Theory Section

* What is the difference between **Descriptive** and **Inferential Statistics**?

**🔹 Descriptive Statistics**

* **Purpose**: Summarizes and describes the **main features** of a dataset.
* **Scope**: Limited to the data you actually have (the sample or population you're working with).
* **Examples**:
  + Mean, median, mode
  + Standard deviation
  + Charts (bar graphs, histograms)
  + Percentages and proportions

✅ **Key Point**: It helps you understand **what the data says**, but **does not** go beyond the data.

**🔹 Inferential Statistics**

* **Purpose**: Makes **predictions or generalizations** about a population based on a sample.
* **Scope**: Goes beyond the observed data to draw **conclusions or inferences**.
* **Examples**:
  + Hypothesis testing
  + Confidence intervals
  + Regression analysis
  + p-values and significance testing

✅ **Key Point**: It helps you make **educated guesses** about a larger group, using sample data.

* + **Define: Population and Sample**

**🔹 Population**

* **Definition**: The **entire group** of individuals, items, or data that you want to study or draw conclusions about.
* **Examples**:
  + All students in a university
  + Every car produced by a company in a year
  + The full list of voters in a country

✅ **Key Point**: A population includes **all** members of a defined group.

**🔹 Sample**

* **Definition**: A **subset** of the population that is selected for analysis.
* **Examples**:
  + 500 students selected from a university for a survey
  + 100 cars tested for quality from a factory
  + A group of 1,000 voters chosen for a poll

✅ **Key Point**: A sample is used to make **inferences** about the population when it's impractical or impossible to study everyone.

* Define **mean**, **median**, and **mode**. How are they different from each other?

**🔹 Mean (Average)**

* **Definition**: The sum of all values divided by the number of values.
* **Formula**:

Mean=Sum of all values / Number of values

**🔹 Median**

* **Definition**: The **middle value** in a sorted list of numbers. If the number of values is even, it's the average of the two middle numbers.
* **Steps**:
  1. Arrange the numbers in order.
  2. Find the middle number.

**🔹 Mode**

* **Definition**: The value(s) that occur **most frequently** in a dataset.

**🔁 Key Differences:**

| **Feature** | **Mean** | **Median** | **Mode** |
| --- | --- | --- | --- |
| **Type** | Mathematical average | Middle value | Most frequent value |
| **Sensitive to Outliers?** | ✅ Yes | 🚫 No | 🚫 No |
| **Uniqueness** | Always one value | One value (or average of two) | Can be none, one, or multiple |
| **Example Use** | Average test score | Income distribution (skewed data) | Most common shoe size in a store |

* What is a **Random Variable**? What are its types?

### 🔹 ****What is a Random Variable?****

A **random variable** is a variable that takes on **numerical values** determined by the outcome of a **random experiment**.

In simple terms:

A random variable assigns numbers to outcomes of a chance process.

There are **two main types**:

### ****Discrete Random Variable****

### Continuous Random Variable