Unit-1
Introduction and presentation of statistical Data

By: Nikunj Mangukiya
Assistant Professor at SDJ International Collage Surat.

1.1
Types of Variable

- **Variable**
- **Definition:** A variable is a property that can take on many values.
- "Age" is a variable. It can take on many different values, such as 18, 49, 72, and so on.
- "Gender" is a variable. It can take on two different values, either male or female.

Types

1. Quantitative Variable:

Definition: A variable which can be measured numerically is called quantitative variable. With measurements of quantitative variables you can do things like add and subtract, and multiply and divide, and get a meaningful result. In the previous example, "Age" was a quantitative variable.

2. Qualitative /Categorical variables

- **Definition:** A variable which can not be measured numerically , it may categorized is called Qualitative variable. These allow for classification based on some characteristic. With measurements of qualitative/categorical variables you cannot do things like add and subtract, and multiply and divide, and get a meaningful result. In the previous example, "Gender" was a qualitative/categorical variable. Gender was categorized as either male or female.
- **There are two further kinds of quantitative variables:**

1. Discrete Variable

- **Definition**: A variable which can take finitely or infinitely countable number of values is said to be **Discrete variable**.
- For example, imagine you rolled an unbiased dice four times and measured how many times you rolled an even number. What are your possible outcomes? {0, 1, 2, 3, 4}

2. Continuous Variable

- **Definition**: A variable which can take infinitely uncountable number of values is said to be **continuous variable.**
- Take temperature for example. Temperature can take on an infinite number of values, such as 80 degrees, or 80.01 degrees, or 80.0050592359 degrees.
- In the previous example we were limited to a finite number of values (you couldn't roll 1.5 even numbers), which is what made it discrete.

1.2
Univariate, bivariate and multivariate data.

- **Definitions.**
- **1.Univariate data:** When we conduct a study that looks at only one variable, we say that we are working with univariate data. Suppose, for example, that we conducted a survey to estimate the average weight of high school students. Since we are only working with one variable (weight), we would be working with univariate data.

- **2. Bivariate data:** When we conduct a study that examines the relationship between two variables, we are working with bivariate data. Suppose we conducted a study to see if there were a relationship between the height and weight of high school students. Since we are working with two variables (height and weight), we would be working with bivariate data.

- 3. **Multivariate Data:** Multivariate data is the data in which analysis are based on more than two variables per observation. Usually multivariate data is used for explanatory purposes.
- Used by financial analysts to estimate profits, cash flow, risk, and even develop new product ideas, based on the types of variables collected. The analysis of these statistics, which consist of a set of random variables that have multiple dimensions is complex and sometimes requires special software systems to aid in the analysis.

1.3 Frequency distribution and its types

- "**Frequency distribution** is statistical table which shows the set of all distinct values of the variable arranged in order of magnitude, either individually or in groups, with their corresponding frequency side by side"
- Thus, frequency distribution possesses two parts. On its left, there are values or sizes or measurements; and on its right the number of items the value is repeated which is called number of items or frequency or number of observation

1.3.1 Discrete Frequency Distribution

Illustration – Discrete Frequency Distribution

Height (in inches)	No. of Students
60	12
62	18
64	10
66	6
68	4

Number of marks	Tally marks	Frequency
1	𝍸 ||	7
2	𝍸	5
3	𝍸 |	6
4	𝍸	5
5	|||	3
Total		**26**

1.3.2 Continuous frequency distribution

Table (a)

Class (Marks)	Frequency
11 - 15	2
16 - 20	3
21 - 25	3
26 - 30	5
31 - 35	6
36 - 40	6
41 - 45	3
46 - 50	2
Total	**30**

Table (b)

Class	Frequency (Marks)
10.5 - 15.5	2
15.5 - 20.5	3
20.5 - 25.5	3
25.5 - 30.5	5
30.5 - 35.5	6
35.5 - 40.5	6
40.5 - 45.5	3
45.5 - 50.5	2
Total	**30**

- Here we can se that the series we are given in table (a) is Inclusive series and in table (b) is Exclusive series.
- In Inclusive series the lower and upper limits both are included in that class itself.
- In Exclusive series the limits of different classes are overlapping. Generally, the lower limits are included in that group or class.

1.3.3 Univariate frequency distribution.

- "constructing frequency distribution by taking only one variable is called Univariate frequency distribution"
- In previous, examples of frequency distribution are of Univariate frequency distribution.

1.3.4 Bivariate frequency distribution

- "constructing frequency distribution by taking two variables is called Bivariate frequency distribution"
- In bivariate frequency distribution we have two study variables.

WEIGHT	HEIGHT					
	60-62	63-65	66-68	69-71	72-74	TOTAL
100_104	5	-	-	-	-	5
105-109	5	10	2	1	-	18
110-114	-	4	6	1	-	11
115_119	-	1	7	6	2	16
TOTAL	10	15	15	8	2	50

- Here in this example we have two variables height and weight and corresponding class frequencies.