1. Descriptive measures: In “Data Presentations” chapter, we discussed how a raw data set can be organized, and summarized by tables and graphs. Another method of summarizing data set precisely is to compute numbers (a single value). Number that can be describe data sets are called “Descriptive Measures”. There are four types of descriptive measures,

a) Measure of central tendency
b) Measure of location
c) Measure of dispersion
d) Shape of the distribution

2. Measure of Central Tendency: A measure of central tendency is a summary measure that attempts to describe the whole set of data with a single value that represents the center of the distribution.

3. Necessity of measuring the central tendency:

- They give us an idea about the concentration of the values in the central part of the distribution.
- They give precise information, not information of a vague general type.
- It represents all relevant information contained in the data as few numbers as possible.

4. Types of Measures of Central tendency: There are three primary types of measures of central tendency, each serving a distinct purpose in summarizing data.

a) Mean
b) Median
c) Mode

5. Mean: The mean, also known as the average, is calculated by summing up all the data values and dividing by the total number of values. It is sensitive to extreme values and provides a balanced representation of the data distribution. There are different types of mean,

i. Arithmetic mean
ii. Weighted mean
iii. Geometric mean
iv. Harmonic mean → Phy

6. Arithmetic mean: The arithmetic mean of a set of observations is the sum of the observations divided by the number of observations. Suppose, we have n observations $x_1, x_2, \ldots, x_n$. Then the arithmetic mean can be written as,

$$\bar{x} = \frac{Sum\ of\ the\ observations}{Total\ number\ of\ observations} = \frac{\sum_{i=1}^{n} x_i}{n}$$

For population, the arithmetic mean can be written as,

$$\mu = \frac{\sum_{i=1}^{n} X_i}{N}$$

Math: The intelligence quotients (IQ's) of 10 students in a class are given below:

70, 120, 110, 101, 88, 83, 95, 98, 107, 100

Find the mean I.Q.

Solution: Mean I.Q. of 10 students is given by,

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{70 + 120 + 110 + 101 + 88 + 83 + 95 + 98 + 107 + 100}{10} = \frac{972}{10} = 97.2$$

Interpretation: Hence, the average I.Q. score of the students is 97.2

Math: The intelligence quotients (IQ's) of 10 students in a class are given below:

70, 120, 110, 101, P, 83, 95, 98, 107, 100

Here, $\bar{x} = 97.2$, find the value of $\boldsymbol{P}$.

Solution: Here,

$\bar{x} = 97.2$

$\Longrightarrow \frac{\sum x_i}{n} = 97.2$

$\Rightarrow \frac{70 + 120 + 110 + 101 + P + 83 + 95 + 98 + 107 + 100}{10} = 97.2$

$\Rightarrow 884 + P = 972$

$\Rightarrow P = 88$

7. Merits of Arithmetic mean:

a) It is rigidly defined.

b) It is easy to calculate and understand

c) It is based on all the observations.

8. Demerits of Arithmetic mean: The strongest drawback of arithmetic mean is that it is very much affected by extreme observations. Two or three very large values of the variable may unduly affect the value of the arithmetic mean.

Outlier/Extreme observation is data point that vary in large when compared to the other observations in the data set.

For example, imagine placing this data set in ascending order: 3, 6, 7, 10 and 54. 54 might be an outlier because it is significantly higher than the other data points

9. Arithmetic mean for grouped data: The arithmetic mean for grouped data can be written as,

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Here, $f_i = Frequency, and\ x_i = class\ mid\ value$

Math: Construct the frequency distribution table for the follow data set and calculate the average.

17,8,12,19,14,6,10,15,7,18,11,16,9

Solution: We begin by creating a frequency distribution table, following the steps outlined in the 'Data Presentation' chapter.

Table: Frequency distribution table.

Class	*Tally*	*Frequency* (f_i)	*Mid-value* (x_i)	$f_i x_i$
5-9	|||	3	$\frac{5+9}{2} = 7$	21
9-13	||||	4	11	44
13-17	||||	4	15	60
17-21	||	2	19	38
Total		$\sum f_i = 13$		$\sum f_i x_i = 163$

$$\therefore Arithmetic\ mean, \bar{x}_{AM} = \frac{\sum f_i x_i}{\sum f_i} = \frac{163}{13} = 12.54$$

Hence, the average value is 12.54

Math:

No of observation	Frequency
0	46
1	f_1
2	f_2
3	25
4	10
5	5

Find the value of f_1 and f_2, where $\bar{X} = 1.46$ and $\sum f_i = 200$.

Solution:

First, we reconstruct the table,

No of observation (x_i)	***Frequency*** (f_i)	$f_i x_i$
0	46	0
1	f_1	f_1
2	f_2	$2f_2$
3	25	75
4	10	40
5	5	25

Given that,

$$\sum f_i = 200$$

$$\Rightarrow 46 + f_1 + f_2 + 25 + 10 + 5 = 200$$

$$\Rightarrow f_1 + f_2 = 114 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (i)$$

Again,

$$\bar{X} = 1.46$$

$$\Rightarrow \frac{\sum f_i x_i}{\sum f_i} = 1.46$$

$$\Rightarrow \frac{0 + f_1 + 2f_2 + 75 + 40 + 25}{200} = 1.46$$

$$\Longrightarrow \frac{f_1 + 2f_2 + 140}{200} = 1.46$$

$$\Longrightarrow f_1 + 2f_2 = 152 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (ii)$$

Solving equation (i) and (ii), we get,

$f_1 = 76$ and $f_2 = 38$

Weighted arithmetic mean is used when different values in a dataset have varying levels of importance or significance, and you want to calculate a more accurate average that takes these differences into account. Unlike the regular arithmetic mean, where all values contribute equally, the weighted arithmetic mean allows you to assign different weights to different values, reflecting their relative importance.

10. Weighted mean: Let $x_1, x_2, \ldots x_n$ be n values whose relative importance is measures by corresponding positive weights $w_1, w_2, \ldots w_n$. The weighted arithmetic mean is given by,

$$\bar{x}_w = \frac{w_1x_1 + w_2x_2 + \cdots + w_nx_n}{w_1 + w_2 + \cdots + w_n} = \frac{\sum w_ix_i}{\sum w_i}$$

Math: During a one-hour period on a hot Saturday afternoon cabana boy Chris served fifty drinks. He sold five drinks for \$0.50, fifteen for \$0.75, fifteen for \$0.90, and fifteen for \$1.15. Compute the weighted mean of the price of the drinks.

Solution: First we construct a table,

Drink type	**Amount (w_i)**	**Price \$ (x_i)**
Drink 1	5	0.50
Drink 2	15	0.75
Drink 3	15	0.90
Drink 4	15	1.15

Now, the weighted mean is given by,

$$\bar{x}_w = \frac{\sum w_ix_i}{\sum w_i} = \frac{(5 \times 0.50) + (15 \times 0.75) + (15 \times 0.90) + (15 \times 1.15)}{5 + 15 + 15 + 15} = 0.89$$

Hence, the weighted mean of the price of the drinks is 0.89\$.

Math: Below table show the result of a student,

Course	Grade point	Credit hour
1	4	3
2	4	3
3	3.75	2
4	3.5	2
5	3	3

weight associated to each grade.

Calculate the weighted mean.

Solution: Here,

Course	*Grade point* (x_i)	*Credit hour* (w_i)	$w_i \times x_i$
1	4	3	12
2	4	3	12
3	3.75	2	7.5
4	3.5	2	7
5	3	3	9
Total		$\sum w_i = 13$	$\sum w_i x_i = 47.5$

The weighted mean is given by,

$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i} = \frac{47.5}{13} = 3.65$$

Hence, the grade point average is 3.65

Math: Below table show the result of a student,

Course	Grade point	Credit hour
1	4	3
2	4	f_1
3	3.75	2
4	3.5	2
5	3	3

Where, the weighted mean is 3.65. Find the value of f_1.

Solution: Here,

Course	**Grade point (x_i)**	**Credit hour (w_i)**	$\boldsymbol{w_i \times x_i}$
1	4	3	12
2	4	f_1	$4f_1$
3	3.75	2	7.5
4	3.5	2	7
5	3	3	9

Now,

$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i}$$

$$\Rightarrow 3.65 = \frac{12 + 4f_1 + 7.5 + 7 + 9}{3 + f_1 + 2 + 2 + 3}$$

$$\Rightarrow 3.65 = \frac{35.5+4f_1}{10+f_1}$$

$$\Rightarrow 36.5 + 3.65f_1 - 4f_1 = 35.5$$

$$\Rightarrow 0.35f_1 = 1$$

$$\Rightarrow f_1 = 2.86 \approx 3$$

11. Merits (Advantages) of Weighted Mean:

a) ***Reflects Importance:*** Weighted mean allows you to assign different levels of importance to different data points. This is particularly useful when certain data points carry more significance or reliability in the context of your analysis.

b) ***Adjusts for Skewness:*** In datasets with skewed distributions, the weighted mean can help mitigate the impact of extreme values by assigning lower weights to outliers.

c) ***Flexible and Customizable:*** The weighted mean is flexible, as it allows you to tailor the weights to specific criteria. This flexibility is especially valuable when certain groups or factors need to be emphasized in the analysis.

d) ***Better Reflects Real-World Scenarios:*** In scenarios where each data point represents a different sample size or population, the weighted mean can provide a more accurate summary measure that takes these variations into account.

12. Demerits (Disadvantages) of Weighted Mean:

a) ***Complexity:*** Calculating the weighted mean involves extra steps and computations compared to the simple arithmetic mean. This complexity can make it more time-consuming and prone to calculation errors.

b) ***Misuse and Misinterpretation:*** If the weights are not assigned appropriately or justified, the results might not accurately represent the underlying data. Misuse of weighted mean can lead to incorrect conclusions.

c) ***Sensitive to Outliers in Weights:*** Outliers in the weights can significantly influence the weighted mean. If the weights are not well-distributed or contain outliers themselves, the resulting weighted mean may be skewed.

13. Geometric mean: Let a data set contain n observations which all are positive, then geometric mean is the n^{th} positive root of their product. Suppose, we have n positive observations $x_1, x_2, \dots, x_n$. Thus, geometric mean can be written as,

$$\bar{x}_{GM} = (x_1 \times x_2 \times \dots \times x_n)^{\frac{1}{n}}$$

For grouped data, geometric mean can be written as,

$$\bar{x}_{GM} = \left(x_1^{f_1} \times x_2^{f_2} \times \dots \times x_n^{f_n}\right)^{\frac{1}{\sum f_i}}$$

Math: Suppose the population growth rate from 2010 to 2013 were, 1.24, 1.26,1.29 and 1.31 respectively. Find the geometric mean of these growth rates.

Solution: Here, the growth rates are, 1.24, 1.26,1.29, 1.31

The geometric mean can be written as,

$$GM = (1.24 \times 1.26 \times 1.19 \times 1.31)^{\frac{1}{4}} = 1.27$$

Hence, the average growth rate is 1.27

Math: Suppose you have an investment that grew by 10% in the first year, 5% in the second year, 8% in the third year. What was the average growth rate for these three years?

> ***Average growth rate***
> $$= [(\mathbf{1}+\boldsymbol{r_1}) \times (\mathbf{1}+\boldsymbol{r_2}) \times \dots \times (\mathbf{1}+\boldsymbol{r_n})]^{\frac{1}{n}} - \mathbf{1}$$
> **Here, $\boldsymbol{r_1, r_2, \dots, r_n}$ be the growth rate (%).**

Solution: Since, this is a growth rate over the year, the average growth rate can be computed by using geometric mean.

$$\therefore Average\ growth\ rate, \bar{x}_{GM} = \{(1+0.10) \times (1+0.05) \times (1+0.08)\}^{\frac{1}{3}} - 1 = 0.0748$$

Thus, the average growth rate over these three years is approximately 7.48%.

Math: Let's consider an asset that depreciated by 15% in the first year, 8% in the second year, and 12% in the third year. What was the average depreciation rate over these three years?

> ***Average depreciation rate***
> $$= [(\mathbf{1}-\boldsymbol{r_1}) \times (\mathbf{1}-\boldsymbol{r_2}) \times \dots \times (\mathbf{1}-\boldsymbol{r_n})]^{\frac{1}{n}} - \mathbf{1}$$
> **Here, $\boldsymbol{r_1, r_2, \dots, r_n}$ be the diminishing rate (%).**

Solution: Since, this is a depreciation rate over the year, the average depreciation rate can be computed by using geometric mean.

$$\therefore Average\ depreciation\ rate, \bar{x}_{GM} = \{(1-0.15) \times (1-0.08) \times (1-0.12)\}^{\frac{1}{3}} - 1$$
$$= -0.1041$$

Thus, the average depreciation rate over these three years is approximately 10.41%.

Math: Economic survey was conducted each year for a five years period in a country. The economic growth rate at the end of year was 5%, 6%, 7.5%, 9% and 10.5% respectively in the survey. What was the average growth rate for this five-year period? **(Ans: 7.57%)**

Math: A given machine is assumed to depreciated 35% in value in the first year, 25% in the second year, 20% in the third year, and 10% in the fourth year. Percentage being calculated on the diminishing value. What is the average depreciation recorded on the diminishing value for the period of four years? **(Ans: 23.03%)**

Math: Consider a series of data transfer rates (in Mbps) for a computer network over four successive years: 10, 15, 18, and 22. Calculate the geometric mean data transfer rate for these four years.

Math: Consider a company's profits for the years 2001, 2002, and 2003. The profit was \$10,000 in 2000, increased to \$12,000 in 2001, further increased to \$15,000 in 2002, and then decreased to \$13,000 in 2003. Compute the percentage increase, and subsequently, find the average increase rate for this company over the three consecutive years.

Hints:

$$\boldsymbol{Increase\ rate} = \frac{(\boldsymbol{Profit\ in\ current\ year - Profit\ in\ Base\ year})}{\boldsymbol{Profit\ in\ base\ year}} \times \mathbf{100}$$

14. Situation where the geometric mean is employed: The geometric mean is useful when dealing with data that exhibits exponential growth or multiplication. If your rate of change involves variables that follow a multiplicative pattern, such as growth rates or compound interest, the geometric mean might provide a more accurate representation. That is, it's particularly suited for calculating average growth rates, comparing values with multiplicative relationships, and analyzing data that changes exponentially over time.

15. Merits of geometric mean:

a) Geometric mean is rigidly defined.

b) It is based on all the observations.

c) It is not affected much by fluctuations of sampling.

d) As compared with arithmetic mean, GM is affected to a lesser extent by extreme observations.

16. Demerits of geometric mean:

a) Because of its abstract mathematical character, geometric mean is not easy to understand and to calculate for a non-mathematical person.

b) If any one of the observations is zero, geometric mean becomes zero and if any one of the observations is negative, geometric mean becomes imaginary regardless of the magnitude of the other items.

17. Harmonic mean: If a data set contains non-zero observations, then harmonic mean is the reciprocal of arithmetic mean of the reciprocal of the observations. Suppose, we have a set of n non-zero observations, $x_1, \dots, x_n$. Then the harmonic mean can be written as,

$$HM = \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}} = \frac{n}{\sum \frac{1}{x_i}}$$

For grouped data, harmonic mean can be written as,

$$\bar{X}_{HM} = \frac{f_1 + f_2 + \cdots + f_N}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \cdots + \frac{f_N}{x_N}} = \frac{\sum f_i}{\sum \frac{f_i}{x_i}}$$

Math: A car travels 50 miles at 40 mph, 60miles at 50mph and 40 miles at 60mph.What is the average speed of the trip?

<u>**Solution:**</u> Here,

Distance: 50 miles, 60 miles, 40 miles

Speed: 40 mph, 50 mph, 60 mph

The mean (harmonic mean) can be written as,

$$HM = \frac{\sum f_i}{\sum \left(\frac{f_i}{x_i}\right)} = \frac{50 + 60 + 40}{\frac{50}{40} + \frac{60}{50} + \frac{40}{60}} = 48.08\ mph$$

Math: Suppose the distance between two places A and B is 18 km. A car drive for the first 6 km at a speed of 30 km/h, and second 5 km at a speed 40 km/h. Remaining distance at a speed of 20 km/h. What was the average speed in this case?

18. Uses/ Situation where the harmonic mean is employed: Harmonic mean is especially useful in averaging rates and ratios where time factor is variable and the act being performed e.g., distance is constant. That is, harmonic mean is suitable for situations involving rates, ratios, or averages that have an inverse relationship. If your rate of change involves variables that have an inverse relationship, such as speed and time, where higher speeds correspond to shorter times, then the harmonic mean might be more appropriate.

19. Merits of HM:

a) Harmonic mean is rigidly defined.
b) It is based on all the observations.
c) It is suitable for further mathematical treatment.
d) Less affected by extreme values

20. Demerits of HM:

a) It is not easy to understand and calculate.
b) Its value cannot be obtained if any one of the observations is zero

- $AM \geq GM \geq HM$
- $GM = \sqrt{AM \times HM}$
- **Arithmetic mean and geometric mean of two positive numbers are 25 and 15, respectively. Find harmonic mean. (Ans: 9)**

21. Criteria of a good measure of central tendency:

It should be

- Easy to understand
- Easy to compute
- Based upon all observations
- Rigidly defined
- Excessively effected by extreme values
- Suitable for further algebraic treatment
- Less affected by sampling fluctuation

EXTRA

1. There are two units of a garment in two different cities employing 760 and 800 persons, respectively. The arithmetic means of monthly salaries paid to persons in these two units are tk 18750 and tk. 16950 respectively. Find the combined arithmetic mean of salaries of the employees in both the units. [Ans: tk. 17827 (appx.)]
2. An investor allocates 1500$ for purchasing shares in a company each month. In the first five months, he buys shares at prices of 25$, 30$, 20$, 40$, and 50$ per share. After five months, what is the average price paid for the shares by the investor? [Hints: weighted mean, Ans: 29.70$ Approximately]
3. From a certain frequency distribution consisting of 180 observations with population mean 7. But on comparing the original data, it was found that an observation 12 was misreported as 21 in the computation. Compute the correct mean. [Ans: 6.95]
4. The average declared by a group of 10 chemical companies was 18 percent. Later on, it was discovered that one correct figure, 12 was misread as 22. Find the correct average dividend. [Ans: 17%]