Measures of Central Tendency

A measure of central tendency is a single value that is used to represent an entire set of data.

Measure of central tendency is also known as an '**Average**'.
The three most commonly used measures of central tendency or 'averages' are:

- Arithmetic Mean
- Median
- Mode

Objectives and functions of averages

1. **To present huge data in a summarised form:** It is difficult to grasp a large amount of data or numerical figures. Averages summarise such data into a single figure which makes it easier to understand and remember.
2. **To facilitate comparison:** Averages are very helpful for making comparative studies as they reduce the mass of statistical data to a single figure or estimate.
3. **To facilitate further statistical analysis:** Various tools of statistical analysis like standard deviation, correlation etc. are based on averages.
4. **To trace precise relationship:** Averages are helpful and even essential when it comes to establishing relationships between different groups of data or variables.
5. **To help in decision-making:** Averages provide values which act as a guideline for decision makers. Most of the decisions to be taken in research or planning are based on the average value of certain variables.

Essentials of a good average / measure of central tendency

1. **It should be rigidly defined:**
   - An average should be clear and there should be only one form of interpretation.
   - It should have a definite and fixed value irrespective of method of calculations or formulae used.

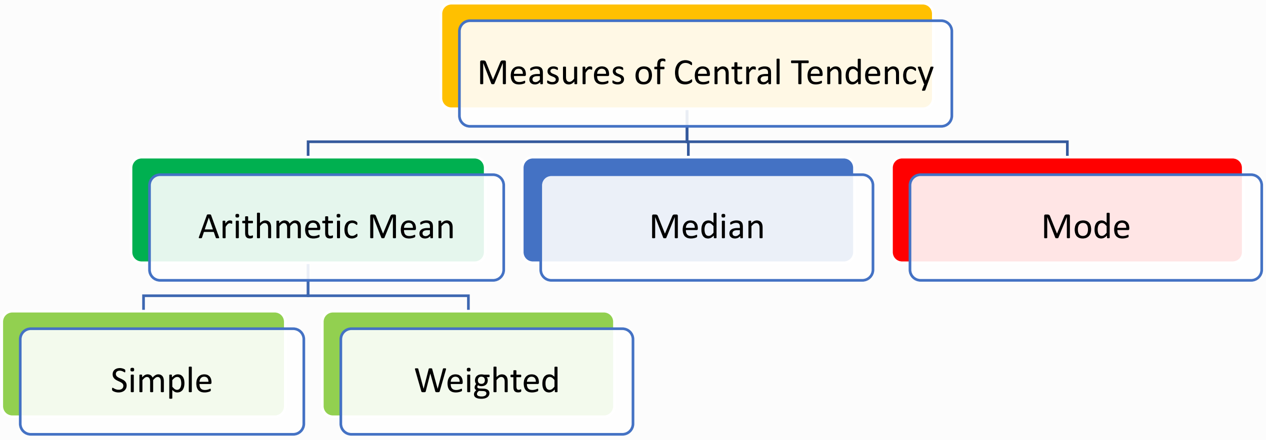
2. **It should be based on all observations:**
   - Average should be calculated by taking into consideration each and every item of the series.
   - If it is not based on all observations, it will not be representative of the whole group.

3. **It should not be affected much by extreme values:**
   - The value of an average should not be affected much by extreme values.
   - One or two very small or very large values should not unduly affect the value of the average significantly.

4. **It should be least affected by fluctuations of sampling:**
   - An average should possess sampling stability i.e. If we take two or more samples from a given population and compute averages for each, then the values thus obtained from different samples should not differ much from each other.

5. **It should be easy to understand and compute:**
   - The value of an average should be computed by using a simple method without reducing its accuracy and other advantages.

6. **It should be capable of further algebraic treatment:**
   - It should be capable of further mathematical and statistical analysis to expand its utility such as to be further used in calculation of measures of dispersion, correlation etc.



Arithmetic Mean

It is defined as the sum of the values of all observations divided by the number of observations. In general, if there are N observations as $X_1, X_2, X_3, \ldots, X_N$, then the Arithmetic Mean is given by:

$$\bar{X} = \frac{X_1 + X_2 + X_2 + \cdots + X_N}{N}$$

For convenience, this will be written in simpler form:

$$\boxed{\bar{X} = \frac{\Sigma X}{N}}$$

where, ΣX = sum of all observations and N = Total number of observations.

Calculation of Mean for different series using different methods

- **INDIVIDUAL SERIES**

1) Direct Method

$$\boxed{\bar{X} = \frac{\Sigma X}{N}}$$

where, ΣX = Sum of all observations and N = Total number of observations.

For example:

The following data shows the weekly income of 10 families. Calculate the arithmetic mean by direct method and interpret the result.

Family	A	B	C	D	E	F	G	H	I	J
Weekly Income (in ₹)	850	700	100	750	5,000	80	420	2,500	400	360

Computation of Arithmetic Mean by Direct Method

Family	A	B	C	D	E	F	G	H	I	J	Total
Weekly Income (X)	850	700	100	750	5,000	80	420	2,500	400	360	**11,160**

$$\bar{X} = \frac{\Sigma X}{N} = \frac{11{,}160}{10} = ₹1{,}116$$

2) Assumed Mean/ Shortcut Method

$$\boxed{\bar{X} = A + \frac{\Sigma d}{N}}$$

where, d = (X – A) ; A = Assumed mean and N = Total number of observations.

For example:

The following table gives the daily income of ten workers in a factory. Find the arithmetic mean by Assumed Mean Method. **(Take A = 250)**

Workers	A	B	C	D	E	F	G	H	I	J
Daily Income (in ₹)	120	150	180	200	250	300	220	350	370	260

Solution: **Computation of Arithmetic Mean by Assumed Mean Method**

Workers	Daily Income (₹) (X)	d = X – 250
A	120	–130
B	150	–100
C	180	–70
D	200	–50
E	250	0
F	300	+50
G	220	–30
H	350	+100
I	370	+120
J	260	+10
N = 10		**– 100**

$$\bar{X} = A + \frac{\Sigma d}{N} = 250 + \left(\frac{-100}{10}\right) = 250 - 10 = ₹240$$

Thus, the average daily income of a worker is ₹240.

3) Step Deviation Method

$$\boxed{\bar{X} = A + \frac{\Sigma d'}{N} \times c}$$

where, $d' = \frac{(X - A)}{C}$; C is the common factor in d ;

A = Assumed mean and N = Total number of observations.

For example:

The following data shows the weekly income of 10 families. Calculate the arithmetic mean by step deviation method and interpret the result. **(Take A = 850)**

Family	A	B	C	D	E	F	G	H	I	J
Weekly Income (in ₹)	850	700	100	750	5,000	80	420	2,500	400	360

Families	Income (X)	d = X – 850	d′ = (X – 850)/10
A	850	0	0
B	700	–150	–15
C	100	–750	–75
D	750	–100	–10
E	5000	+4150	+415
F	80	–770	–77
G	420	–430	–43
H	2500	+1650	+165
I	400	–450	–45
J	360	–490	–49
			+266

$$\bar{X} = A + \frac{\Sigma d'}{N} \times c = 850 + \frac{266}{10} \times 10 = 850 + 266 = ₹1{,}116.$$

Interpretation: The average weekly income of a family is ₹1,116.

- **DISCRETE SERIES**

1) Direct Method

$$\bar{X} = \frac{\Sigma fX}{\Sigma f}$$

where, ΣfX = Sum of all observations multiplied by their respective frequency
and $\Sigma f = N$ = Total number of observations.

For example:

Plots in a housing colony come in only three sizes: 100 sq. metre, 200 sq. meters and 300 sq. metre and the number of plots are respectively 200, 50 and 10. Calculate the mean plot size in the housing colony by direct method.

Solution: Computation of Arithmetic Mean by Direct Method

Plot Size in sq. metre (X)	No. of Plots (f)	f X
100	200	20,000
200	50	10,000
300	10	3,000
	260	**33,000**

$$\bar{X} = \frac{\Sigma fX}{\Sigma f} = \frac{33000}{260} = 126.92$$

Therefore, the mean plot size in the housing colony is 126.92 sq. metre.

2) Assumed Mean / Shortcut Method

$$\bar{X} = A + \frac{\Sigma fd}{\Sigma f}$$

where, $d = (X - A)$; A = Assumed mean and $\Sigma f = N$ = Total number of observations.

For example:

Plots in a housing colony come in only three sizes: 100 sq. metre, 200 sq. meters and 300 sq. metre and the number of plots are respectively 200, 50 and 10. Calculate the mean plot size in the housing colony by assumed mean method. (Take A = 200)

Solution: Computation of Arithmetic Mean by Assumed Mean Method

Plot Size in sq. metre (X)	No. of Plots (f)	d = X–200	fd
100	200	–100	–20,000
200	50	0	0
300	10	+100	+1,000
	260		**–19,000**

$$\text{Mean plot size } \bar{X} = A + \frac{\Sigma fd}{\Sigma f} = 200 + \left(\frac{-19{,}000}{260}\right) = 200 - 73.08 = 126.92 \text{ sq. metre}$$

3) Step Deviation Method

$$\bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c$$

where, $d' = \frac{(X - A)}{C}$; and C is the common factor in d.

A = Assumed mean and $\Sigma f = N$ = Total number of observations.

For example:

Plots in a housing colony come in only three sizes: 100 sq. metre, 200 sq. meters and 300 sq. metre and the number of plots are respectively 200, 50 and 10. Calculate the mean plot size in the housing colony by step-deviation method. (Take A = 200)

Solution: **Computation of Arithmetic Mean by Step Deviation Method**

Plot Size in sq. metre (X)	No. of Plots (f)	d′ = (X–200)/100	fd′
100 200 300	200 50 10	–1 0 +1	–200 0 10
	260		**–190**

$$\text{Mean plot size } \bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c = 200 + \left(\frac{-190}{260}\right) \times 100 = 200 - 73.08 = 126.92 \text{ sq. metre}$$

- **CONTINUOUS SERIES**

1) Direct Method

$$\boxed{\bar{X} = \frac{\Sigma fm}{\Sigma f}}$$

where, Σfm = Sum of midpoints of classes multiplied by their respective class frequency
Σf = N = Total number of observations.

For example:

Calculate average marks of the following students by using direct method:

Marks	0–10	10–20	20–30	0–40	40–50	50–60	60–70
No. of Students	5	12	15	25	8	3	2

Mark (X)	No. of students (f)	Mid value (m)	fm
(1)	(2)	(3)	(4) = (2) × (3)
0–10	5	5	25
10–20	12	15	180
20–30	15	25	375
30–40	25	35	875
40–50	8	45	360
50–60	3	55	165
60–70	2	65	130
	70		**2,110**

$$\bar{X} = \frac{\Sigma fm}{\Sigma f} = \frac{2{,}110}{70} = 30.14$$

Therefore, average marks of 70 students is 30.14.

2) Assumed Mean / Shortcut Method

$$\boxed{\bar{X} = A + \frac{\Sigma fd}{\Sigma f}}$$

where, d = (m – A) and m is the midpoint of the respective class.
A = Assumed mean and Σf = N = Total number of observations.

For example:

Calculate average marks of the following students by using assumed mean method: (Take A = 35)

Marks	0–10	10–20	20–30	0–40	40–50	50–60	60–70
No. of Students	5	12	15	25	8	3	2

Mark (X)	No. of students (f)	Mid value (m)	d= m–35	fd
(1)	(2)	(3)	(4)	(5) = (2)×(4)
0–10	5	5	–30	–150
10–20	12	15	–20	–240
20–30	15	25	–10	–150
30–40	25	35	0	0
40–50	8	45	10	80
50–60	3	55	20	60
60–70	2	65	30	60
	70			**–340**

$$\bar{X} = A + \frac{\Sigma fd}{\Sigma f} = 35 + \left(\frac{-340}{70}\right) = 35 - 4.86 = 30.14$$

Therefore, Average marks of 70 students is 30.14.

3) Step Deviation Method

$$\boxed{\bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c}$$

where, $d' = \frac{(m - A)}{C}$; m is the midpoint of the respective class and C is the common factor in d. A = Assumed mean and $\Sigma f = N$ = Total number of observations.

For example:

Calculate average marks of the following students by using step deviation method: (Take A = 35)

Marks	0–10	10–20	20–30	0–40	40–50	50–60	60–70
No. of Students	5	12	15	25	8	3	2

Solution: Computation of Average Marks for Exclusive Class Interval by Step deviation Method

Mark (X)	No. of students (f)	Mid value (m)	d′ = (m – 35)/10	fd′
(1)	(2)	(3)	(4)	(5) = (2)×(4)
0–10	5	5	–3	–15
10–20	12	15	–2	–24
20–30	15	25	–1	–15
30–40	25	35	0	0
40–50	8	45	1	8
50–60	3	55	2	6
60–70	2	65	3	6
	70			**–34**

Average marks of 70 students, $\bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c = 35 + \frac{(-34)}{70} \times 10 = 30.14$

Properties of Arithmetic Mean

1. **The sum of deviations of observations from their arithmetic mean is always equal to zero.**
   Symbolically $\Sigma(X - \bar{X}) = 0$

Example:

X	$(X - \bar{X})$
10	– 20
20	– 10
30	0
40	+ 10
50	+ 20
ΣX = 150	**$\Sigma(X - \bar{X}) = 0$**

$$\bar{X} = \frac{\Sigma X}{N} = \frac{150}{5} = 30.$$

When we calculate the deviations of all the items from their arithmetic mean ($\bar{X}$ = 30), we find that the sum of the deviations from the arithmetic mean, i.e. $\Sigma(X - \bar{X})$ comes out to be zero.

2. **Arithmetic mean is NOT independent of change of origin**
   If each observation of a series is increased (or decreased) by a constant, then the mean of these observations is also increased (or decreased) by that constant.

3. **Arithmetic mean is NOT independent of change of scale**
   If each observation of a series is multiplied (or divided) by a constant, then the mean of these observations is also multiplied (or divided) by that constant.

4. **The sum of squares of deviations of observations from their arithmetic mean is minimum.**
   $\Sigma(X-\bar{X})^2$ is always minimum.

5. **If arithmetic mean and number of items of two or more related groups are given, then we can compute the combined mean** using the formula given below.

Combined Mean

If we have the arithmetic mean and number of items of two groups, we can compute combined mean of these two groups by applying the following formula:

$$\boxed{\bar{X}_{12} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}}$$

where,

N_1 = Number of items in the first group
$\bar{X}_1$ = Arithmetic mean of the first group
$\bar{X}_{12}$ = Combined mean of the two groups
N_2 = Number of items in the second group
$\bar{X}_2$ = Arithmetic mean of the second group

For example:

The mean marks of 60 students in section A is 40 and mean marks of 40 students in section B is 35. Calculate the combined mean marks of all the students of sections A and B.

Sections	No. of students	Mean marks
A	$N_1 = 60$	$\bar{X}_1 = 40$
B	$N_2 = 40$	$\bar{X}_2 = 35$

Combined mean marks $\bar{X}_{12} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2} = \frac{(60 \times 40) + (40 \times 35)}{60 + 40} = 38$ marks

If we have to find out the combined mean of three groups, the formula will be:

$$\overline{X}_{123} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + N_3\overline{X}_3}{N_1 + N_2 + N_3}$$

Weighted Mean

It refers to the average when different items of a series are given different weights according to their relative importance.

In case of simple arithmetic mean all items of a series are given equal importance. The Weighted Mean is given by:

$$\overline{X}_w = \frac{W_1X_1 + W_2X_2 + ... + W_nX_n}{W_1 + W_2 + ... + W_n} = \frac{\Sigma WX}{\Sigma W}$$

For example:

Find out weighted mean of the following:

Items	Rice	Wheat	Pulses	Cloth	Others
Price	₹100 per kg	₹150 per kg	₹300 per kg	₹15 per metre	₹100 per unit
Weight	6	5	3	2	1

Solution: **Calculation of Weighted Mean**

Items	Price (X)	Weight (W)	WX
Rice	100	6	600
Wheat	150	5	750
Pulses	300	3	900
Cloth	15	2	30
Others	100	1	100
		ΣW = 17	ΣWX = 2380

Weighted mean $\overline{X}w = \frac{\Sigma WX}{\Sigma W} = \frac{2,380}{17} = 140$

Corrected Mean

To find the correct mean when incorrect and correct entries are given:

$$\text{Correct } \overline{X} = \frac{\Sigma X_{(wrong)} + \text{correct values} - \text{incorrect values}}{N}$$

Arithmetic Mean at a Glance

<table>
<tr><th rowspan="2">Ungrouped Data</th><th colspan="2">Grouped Data</th></tr>
<tr><th>Discrete Series</th><th>Continuous Series</th></tr>
<tr>
<td>Direct Method
$\overline{X} = \frac{\Sigma X}{N}$
Assumed Mean Method
$\overline{X} = A + \frac{\Sigma d}{N}$
where, A = assumed mean
$\Sigma d = \Sigma(X - A)$
Step Deviation Method
$\overline{X} = A + \frac{\Sigma d'}{N} \times c$
where, c =common factor
$d' = \frac{(X - A)}{c}$</td>
<td>Direct Method
$\overline{X} = \frac{\Sigma fX}{\Sigma f}$
Assumed Mean Method
$\overline{X} = A + \frac{\Sigma fd}{\Sigma f}$
Step Deviation Method
$\overline{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c$</td>
<td>Direct Method
$\overline{X} = \frac{\Sigma fm}{\Sigma f}$
where, m = mid values
Assumed Mean Method
$\overline{X} = A + \frac{\Sigma fd}{\Sigma f}$
where, d = m – A
Step deviation method
$\overline{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c$
where, $d' = \frac{m - A}{c}$</td>
</tr>
<tr><td colspan="3">Combined mean: $\overline{X}_{12} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}$ Weighted arithmetic mean: $\overline{X}_w = \frac{\Sigma WX}{\Sigma W}$</td></tr>
</table>

Arithmetic mean in special cases

1) Cumulative Series (Less-than or More-than series): Cumulative frequency series is first converted into simple frequency series and then mean is calculated in the usual manner.
2) Mid-Value series: There is no need to convert the mid-value series into classes since only the midpoint is required for calculation of mean.
3) Inclusive series: There is no need to convert inclusive series into exclusive series as the midpoint remains the same in both types of series for calculation of mean.
4) Open-ended series: The missing class limits are assumed according to the pattern of class intervals of other classes and then the mean is calculated in the usual manner.
5) Unequal class series: Mean can be calculated in the usual manner by first calculating the midpoints of each class even if it is of unequal size.

Merits of Arithmetic Mean

- It is based on all observations i.e it takes into consideration all the values in a given series. It is considered to be more representative of the distribution.
- Its value is always definite and it is rigidly defined.
- It is capable of further algebraic treatment. It is widely used in the computation of various statistical measures such as standard deviation, correlation etc.
- Arithmetic mean is the least affected by fluctuations of sampling.

Demerits of Arithmetic Mean

- It is affected by extreme values : Since arithmetic mean is calculated using all the items of a series, it can be unduly affected by extreme values i.e. very small or very large items.
- It may give absurd results: For example, if a teacher says that average number of students in a class is 28.75, it sounds illogical.
- It cannot be obtained graphically like median or mode.
- Arithmetic mean cannot be computed for qualitative data such as honesty, intelligence etc.
- It gives more stress on items of higher value: Arithmetic mean gives more importance to higher items of a series as compared to smaller items or has an upward bias. If out of five, four values are small but one is of a bigger value, the bigger value item will push up the average considerably.

Median

Median is defined as the middle value in the data set when its elements are arranged in a sequential order, that is, in either ascending or descending order.

It is a positional value. Positional average determines the position of variables in the series.

- **INDIVIDUAL SERIES**

  Steps for calculating median

  Step1: First arrange the data in ascending order.
  Step2: Use the given formula to calculate the median.

$$\boxed{\text{Median} = \text{size of} \left(\frac{N+1}{2}\right)^{th} \text{item}}$$

where N = Total number of observations

For example:

From the following data of the wages of 7 workers, compute the median wage and interpret the result:

Wages (in ₹)	1,100	1,150	1,080	1,120	1,200	1,160	1,400

Solution:

Wages (in ₹)	1,080	1,100	1,120	1,150	1,160	1,200	1,400

$$\text{Median} = \text{size of} \left(\frac{N+1}{2}\right)^{th} \text{item} = \text{size of} \left(\frac{7+1}{2}\right)^{th} \text{item} = 4\text{th item} = 1{,}150$$

Therefore, the median wage is ₹1,150.

When the number of observations is an even number

If there are even numbers in the data, there will be two observations which fall in the middle. The median in this case is computed as the arithmetic mean of the two middle values. However, the same formula for calculating median shall apply.

$$\text{Median} = \text{size of} \left(\frac{N+1}{2}\right)^{th} \text{item}$$

<u>For example:</u>

The following data provides marks of 20 students. Calculate median marks and interpret the result:

25, 72, 28, 65, 29, 60, 30, 54, 32, 53, 33, 52, 35, 51, 42, 48, 45, 47, 46, 33

Solution: Arranging the data in an ascending order, we get

25, 28, 29, 30, 32, 33, 33, 35, 42, 45, 46, 47, 48, 51, 52, 53, 54, 60, 65, 72

$$N = 20. \text{ Median} = \text{size of} \left(\frac{N+1}{2}\right)^{th} \text{item} = \text{size of} \left(\frac{20+1}{2}\right)^{th} \text{item} = \text{size of } 10.5\text{th item}$$

$$\text{Size of } 10.5^{th} \text{ item} = \frac{10^{th} \text{ item} + 11^{th} \text{ item}}{2} = \frac{45+46}{2} = 45.5$$

Therefore, Median = 45.5 marks

- **<u>DISCRETE SERIES</u>**

Steps for calculating median

Step 1: Arrange the data in ascending or descending order.

Step 2: Find out the cumulative frequencies.

Step 3: $\text{Median} = \text{size of} \left(\frac{N+1}{2}\right)^{th} \text{items}$, where $N = \Sigma f$

Step 4: The value whose cumulative frequency is equal to $\frac{N+1}{2}$ or next higher to that, is the median value.

For example:

The frequency distribution of the number of persons and their respective incomes (in ₹) are given below. Calculate the median income. Also, interpret the result.

Income (in ₹)	100	200	300	400
Number of persons	2	4	10	4

Solution: In order to calculate the median income, we calculate the cumulative frequencies.

Computation of Median for Discrete Series

Income (in ₹)	No. of persons (f)	Cumulative frequency (c.f.)
100	2	2
200	4	6
300	10	16
400	4	20
	N = Σf = 20	

The median is located in the $\frac{N+1}{2} = \frac{20+1}{2}$ = 10.5th observation.

This can be easily located through cumulative frequency.
The 10.5th observation lies in the c.f. of 16.
The income corresponding to this is ₹300, so the median income is ₹300.

- **CONTINUOUS SERIES**

Steps for calculating median

Step1: Locate the median class where $(N/2)^{th}$ item lies ; N= Σf.

Step2 : Using the formula given below, calculate the median.

$$\text{Median} = L + \frac{\left(\frac{N}{2} - c.f.\right)}{f} \times h$$

where,

L = lower limit of the median class

f = frequency of the median class

c.f. = cumulative frequency of the class preceding the median class

h = magnitude of the median class interval

For example:

Following data relates to daily wages (in ₹) of persons working in a factory. Compute the median daily wage.

Daily wages	550–600	500–550	450–500	400–450	350–400	300–350	250–300	200–250
No. of workers	7	13	15	20	30	33	28	14

Solution: We rearrange the data in ascending order and calculate cumulative frequencies.

Computation of Median for Continuous Series

Daily wages (in ₹)	No. of Workers (f)	Cumulative Frequency
200–250	14	14
250–300	28	42
300–350	33	75
350–400	30	105
400–450	20	125
450–500	15	140
500–550	13	153
550–600	7	160

Median class is the value of $\left(\frac{N}{2}\right)^{th}$ item $= \left(\frac{160}{2}\right)^{th}$ item $= 80^{th}$ item of the series, which lies in 350–400 class interval.

$$\text{Median} = L + \frac{\left(\frac{N}{2} - c.f.\right)}{f} \times h$$

Median = 350 + $\frac{(80-75)}{30}$ X 50 = 350 + 8.33 = ₹358.33

The median daily wage is ₹358.33.

Median Formulae at a Glance

Ungrouped Data	Grouped Data	
	Discrete Series	Continuous Series
Median = size of $\left(\frac{N+1}{2}\right)^{th}$ item where N = number of items	Median = size of $\left(\frac{N+1}{2}\right)^{th}$ item where N = Σf The position of median can be located through cumulative frequency.	Locate median class where $\left(\frac{N}{2}\right)^{th}$ item lies. Median = $L+\frac{\left(\frac{N}{2}-c.f.\right)}{f}\times h$ where, L = lower limit of the median class c.f. = cumulative frequency of preceding class f = frequency of the median class h = magnitude of the median class

Properties of Median

1) The sum of absolute deviations of items from the median (ignoring the signs) is the minimum. i.e $\Sigma|X\text{-Median}|$ is minimum.
2) Median is a positional average so is not affected by change in extreme values.

Median in special cases

1) Cumulative Series (Less-than or More-than series): Cumulative frequency series is first converted into simple frequency series and then median is calculated in the usual manner.
2) Mid-Value series: Mid- values are first converted into classes and then median is calculated in the usual manner
3) Inclusive series: The inclusive series is first converted into exclusive series and then median is calculated in the usual manner.
4) Open-ended series: There is no need to complete the class intervals to calculate the median.
5) Unequal class series: Median can be calculated in the usual manner and there is no need to make the class intervals equal.

Merits of Median

- Its value is always definite and it is rigidly defined.
- It is not affected by extreme values.
- It can be obtained graphically using ogives.
- It is appropriate for qualitative data.
- It can be calculated even in case of open-ended distributions.

Demerits of Median

- It is not based on all observations.
- It is not capable of further algebraic treatment.
- It is affected by fluctuations of sampling.
- It requires arrangement of data in ascending or descending order of magnitude.

MODE

Mode is defined as the value occurring most frequently in a given series and around which other items of the set cluster most densely.
The word mode has been derived from the French word '**la Mode**' which signifies the most fashionable values of a distribution, because it is repeated the highest number of times in the series.

- **INDIVIDUAL SERIES**

  The value which occurs maximum number of times is the **mode**.

For example:
Calculate the mode from the following data of the marks obtained by 10 students:

10, 27, 24, 12, 27, 27, 20, 18, 15, 30

Since the value 27 occurs the maximum number of times (thrice) in the series, hence the modal marks = 27

- **DISCRETE SERIES**

There are two methods of calculating mode using grouped data:
a) Inspection or Observation method b) Grouping method

a) Inspection or observation method : The value of the variable against the highest frequency will give the **mode**.

For example:
Calculate the mode from the following data:

Variable	10	20	30	40	50
Frequency	2	8	20	10	5

Since the maximum frequency is 20 in the given series, hence the value of mode is 30.

b) Grouping method

Under this method, grouping of the data is done by preparing a **Grouping Table** consisting of 6 columns, in addition to a column for the values of the variable.

- In column I, the highest frequency is marked or put in a circle.
- In column II, frequencies are grouped in two's. Find out their total and mark the highest total or put it in a circle.
- In column III, leave the first frequency and then group the remaining in two's. Find out their total and mark the highest total or put it in a circle.
- In column IV, frequencies are grouped in three's. Find out their total and mark the highest total or put it in a circle.
- In column V, leave the first frequency and then group the remaining in three's. Find out their total and mark the highest total or put it in a circle.
- In column VI, leave the first two frequencies and then group the remaining in three's. Find out their total and mark the highest total or put it in a circle.

The highest frequency total in each of the six columns is identified and analysed in the **Analysis Column** to determine mode. The last column will be the analysis column ***and the mode will be the value against the highest tally in the analysis column.***

For example: Calculate the mode from the following data using grouping method.

Variable	10	20	30	40	50
Frequency	2	8	20	10	5

Grouping Table

<table>
<tr><th>Age in yrs.</th><th>Column 1
Frequency</th><th>Column 2</th><th>Column 3</th><th>Column 4</th><th>Column 5</th><th>Column 6</th><th>Analysis Column</th></tr>
<tr><td>10</td><td>2</td><td rowspan="2">10</td><td>--</td><td rowspan="3">30</td><td>--</td><td>--</td><td>I</td></tr>
<tr><td>20</td><td>8</td><td rowspan="2">28</td><td rowspan="3">38</td><td>--</td><td>III</td></tr>
<tr><td>30</td><td>20</td><td rowspan="2">30</td><td rowspan="3">35</td><td>𝍸 I</td></tr>
<tr><td>40</td><td>10</td><td rowspan="2">15</td><td></td><td>III</td></tr>
<tr><td>50</td><td>5</td><td></td><td></td><td></td><td>I</td></tr>
</table>

The value 30 occurs maximum number of times (6 times) in the analysis column. Therefore, the value of **mode is 30**.

- **CONTINUOUS SERIES**

Step1: Find the modal class using either inspection or grouping method.

a) **Inspection/ observation method :** The modal class is the class with highest frequency**.**

For example:

Calculate the mode from the following data and interpret the result:

Class Interval	0-5	5-10	10-15	15-20	20-25
Frequency	7	18	25	30	20

By inspection method, the modal class is 15-20 since it has the highest frequency of 30.

b) **Grouping method** *(Steps same as in discrete series)*

Grouping Table

<table>
<tr><th rowspan="2">Marks</th><th>Column 1</th><th rowspan="2">Column 2</th><th rowspan="2">Column 3</th><th rowspan="2">Column 4</th><th rowspan="2">Column 5</th><th rowspan="2">Column 6</th><th rowspan="2">Analysis Column</th></tr>
<tr><th>Frequency</th></tr>
<tr><td>0 - 5</td><td>7</td><td rowspan="2">25</td><td>--</td><td rowspan="3">50</td><td>--</td><td>--</td><td>I</td></tr>
<tr><td>5 - 10</td><td>18</td><td rowspan="2">43</td><td rowspan="3">73</td><td>--</td><td>II</td></tr>
<tr><td>10 - 15</td><td>25</td><td rowspan="2">55</td><td rowspan="3">75</td><td>IIII</td></tr>
<tr><td>15 - 20</td><td>30</td><td rowspan="2">50</td><td></td><td>~~IIII~~</td></tr>
<tr><td>20 - 25</td><td>20</td><td></td><td></td><td></td><td>II</td></tr>
</table>

By grouping method, the modal class is 15-20 since it has the highest frequency (tally) in the analysis column.

Step2: Using the modal class, mode can be calculated by using the formula:

$$M_o = L + \frac{|f_1 - f_0|}{|f_1 - f_0| + |f_1 - f_2|} \times h$$

Where **L =** Lower limit of the modal class ; **h** = width of the modal class

f_1 = frequency of the modal class

f_0 = frequency of the class preceding modal class.

f_2 = frequency of the class succeeding modal class.

Now, **L** = 15, f_1 = 30 ; f_0 = 25 ; f_2 = 20 ; **h** = 5

$|f_1 - f_0|$ = |30 – 25| = 5; $|f_1 - f_2|$ = |30 – 20| = 10

$$M_o = L + \frac{|f_1 - f_0|}{|f_1 - f_0| + |f_1 - f_2|} \times h$$

$$= 15 + \frac{5}{5+10} \times 5 = 15 + 1.67 = 16.67$$

Thus, the **mode is 16.67 marks**.

Mode in special cases

1) Cumulative Series (Less-than or More-than series) Cumulative frequency series is first converted into simple frequency series and then mode is calculated in the usual manner.
2) Mid-Value series: Mid- values are first converted into class intervals and then mode is calculated in the usual manner
3) Inclusive series: The inclusive series is first converted into exclusive series and then mode is calculated in the usual manner.
4) Open-ended series: There is no need to complete the class intervals to calculate the mode.
5) Unequal class series: The unequal classes need to be first converted into equal width classes and frequencies are adjusted before calculating the mode in the usual manner.

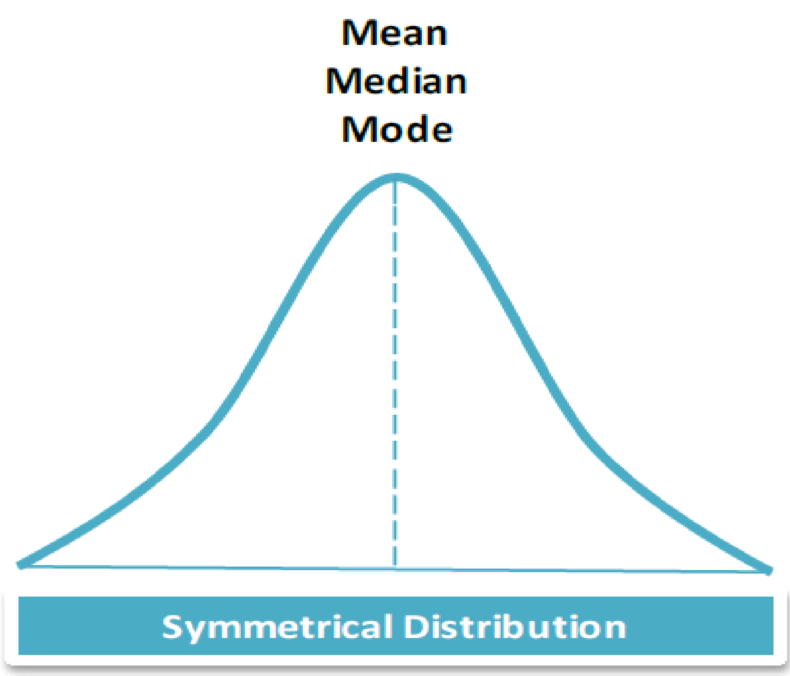
Merits of Mode

- It is not affected by values of extreme items.
- It can be obtained graphically using histogram.
- It can be used to describe quantitative as well as qualitative data.
- It can be calculated even in case of open-ended distributions without finding class limits.

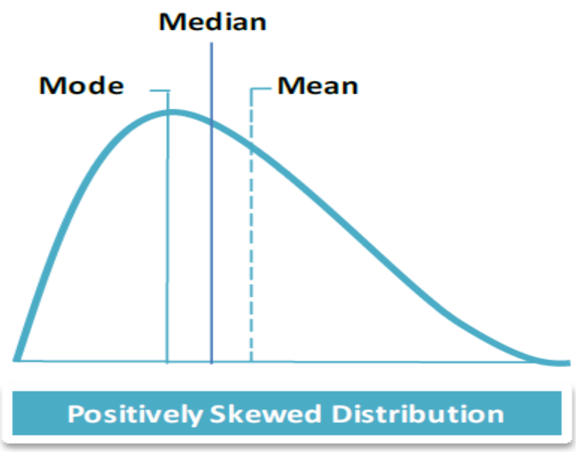
Demerits of Mode

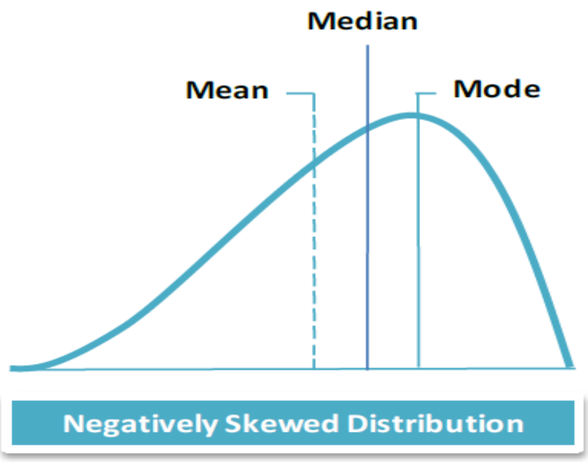
- It is not rigidly defined.
- It is not based on all observations.
- It is not capable of further algebraic treatment.
- It is affected by fluctuations of sampling.

Relationship between Mean, Median and Mode



In a symmetrical distribution: Mean = Median = Mode





In an asymmetrical distribution: Mode = 3 Median – 2 Mean

Symmetrical/normal distribution	Asymmetrical/skewed distribution
In case of symmetrical/normal distribution: Mean = Median = Mode Suppose Arithmetic Mean = M_e, Median = M_i and Mode = M_o $M_e = M_i = M_o$	In case of asymmetrical/skewed distribution: Mode = 3 Median – 2 Mean The median is always between the arithmetic mean and the mode. $M_e > M_i > M_o$ *or* $M_e < M_i < M_o$

Recap

- *The measure of central tendency summarises the data with a single value, which can represent the entire data.*
- *Arithmetic mean is defined as the sum of the values of all observations divided by the number of observations.*
- *The sum of deviations of items from the arithmetic mean is always equal to zero.*
- *Sometimes, it is important to assign weights to various items according to their importance.*
- *Median is the central value of a distribution that divides it into two equal parts with 50% items placed above it and the other 50% items placed below it.*
- *Mode is the value which occurs most frequently in a given series.*

Click on the following links for further explanation of the topics discussed above:

https://www.youtube.com/watch?v=dzX5khpoSaI (Problems Arithmetic Mean)
https://diksha.gov.in/play/content/do_3130390985989652481498 (Mean)
https://diksha.gov.in/play/content/do_3130390986496000001499 (Median)
https://diksha.gov.in/play/content/do_3130390987014062081543 (Mode)