19 : 05

Measures of Central Location

OR Averages (Measures of Central Tendency)

According to Professor Bowley, averages are "statistical constants which enable us to comprehend in a single effort the significance of the whole".

The following are the five measures of central tendency that are in common use:

(i) Arithmetic Mean or Simple Mean
(ii) Median
(iii) Mode
(iv) Geometric Mean (v) Harmonic Mean.

Each of them, in its own way can be called a representative of the characteristic of the whole group and thus the performance of the group as a whole can be described by the single value which each of these measures give. The values of mean, mode and median also help us in comparing two or more groups or their characteristic performance.

Page Date

Arithametic Mean:-

The arithmetic mean of a population is denoted by the symbol μ (mu); and the arithmetic mean of a sample is denoted by the symbol $\bar{X}$ (X-bar)

The arithmetic mean of a sample is defined as the sum of all the observations divided by the number of observations.

In symbols, if n is the no. of observations in a sample and $X_1, X_2, X_3, \ldots X_n$ are the first, second, third, ..., nth etc observations, then their mean is given by $\bar{X}$ where

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$$

For example→ Consider a sample of size 10 on on the no. of florets in a compound flower. Here n is 10, If the observations are:

$X_1 = 25$, $X_2 = 30$, $X_3 = 21$, $X_4 = 55$
$X_5 = 47$, $X_6 = 10$, $X_7 = 15$, $X_8 = 17$
$X_9 = 45$, $X_{10} = 35$

Then $$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9 + X_{10}}{10}$$

$$= \frac{25 + 30 + 21 + 55 + 47 + 10 + 15 + 17 + 45 + 35}{10}$$

$= \frac{300}{10} = 30$

• <u>Note</u> :- In the abbreviated form, the formula can be written as

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

<u>Ques</u>) Calculate the arithmetic mean of a sample of reported cases of mumps in school children as given below:

10, 12, 15, 18, 20, 22, 25, 30

A) ^{h}

<u>~~Q~~</u> : <u>Mean of Grouped Data</u> :-

In ~~the~~ case of frequency distribution $x_i | f_i$, $i = 1, 2, \ldots n$, where f_i is the frequency of variable x_i,

$$\overline{x} = \frac{f_1 x_1 + f_2 x_2 + \cdots + f_n x_n}{f_1 + f_2 + f_3 + \cdots + f_n}$$

or $= \dfrac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$

$= \dfrac{1}{N} \sum_{i=1}^{n} f_i x_i$

where $N = \sum_{i=1}^{n} f_i$

For example,

Ques- Find the arithmetic mean of the following frequency distribution:

X :	1	2	3	4	5	6	7
f :	5	9	12	17	14	10	6

X	f	fX
1	5	5
2	9	18
3	12	36
4	17	68
5	14	70
6	10	60
7	6	42
Total	73	299

$$\overline{X} = \frac{1}{N}\sum fx$$

$$= \frac{299}{73} = 4.09.$$

Ques1- Find the arithmetic mean of the following frequency distribution:

X:	11	12	13	14	15	16	17	18
f:	2	3	5	2	3	2	2	1

Page Date

Mean for Grouped Data With Class - interval More than One :-

When the size of class interval is more than one, i.e., the no. of items in a class interval is more than one, the mid-point of the interval may be used to represent all values falling within the interval.

Ques. Compute the arithmetic mean of the marks from the following table :-

Marks :	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60
No. of students :	12	18	27	20	17	6

Soln

Marks	(f) No. of Students	Mid point (x)	fx
0 - 10	12	5	60
10 - 20	18	15	270
20 - 30	27	25	675
30 - 40	20	35	700
40 - 50	17	45	765
50 - 60	06	55	330
Total	100		2800

Page | Date

Arithmetic Mean , $\overline{X} = \frac{1}{N} \Sigma f x$

$= \frac{1}{100} \times 2800$

$= 28.$

Ques. Calculate the mean of the data on number of florets on sunflowers as given below:

Class-intervals:	0–4	5–9	10–14	15–19	20–24	25–29	30–34	35–39	40–44	45–49
Frequency(f):	2	9	11	26	17	23	36	32	02	01

Calculation of Mean Using Arbitrary Reference Point:-

(Assumed Mean)

We can compute mean from the frequency table by another method in which an arbitrary reference point is taken and then we compute the deviation from this reference point and calculate the mean.

Ques: Calculation of Mean for Group Data Using Assumed Mean Method.

Class-interval	x_i	Frequency (f)	d	fd
25-29	27	1	$\frac{27-52}{4} = -6.25$	-6.25
30-34	32	0	$\frac{32-52}{4} = -5$	0
35-39	37	3	$\frac{37-52}{4} = -3.75$	-11.25
40-44	42	6	$\frac{42-52}{4} = -2.5$	-15
45-49	47	6	$\frac{47-52}{4} = -1.25$	-7.5
$\boxed{50-54}$	52	6	$\frac{52-52}{4} = 0$	0
55-59	57	7	$\frac{57-52}{4} = 1.25$	8.75
60-64	62	4	$\frac{62-52}{4} = 2.5$	10
65-69	67	4	$\frac{67-52}{4} = 3.75$	15
70-74	72	1	$\frac{72-53}{4} = 5$	5
75-79	77	1	$\frac{77-52}{4} = 6.25$	6.25
80-84	82	1	$\frac{82-52}{4} = 7.5$	7.5
Total		N = 40		$\Sigma fd = 12.5$

Page Date

'd' is calculated as

$$d = \frac{X_i - M}{i}$$

Where X_i is the mid point of a particular interval.

i = class interval width

M is the assumed mean.

We have let $M = 52$, $i = 4$

$$\cancel{\bar{X} = \frac{\Sigma f d}{N} \times i + M}$$

$$\bar{X} = M + \frac{\Sigma f d}{N} \times i$$

$$= 52 + \frac{12.5}{\cancel{40}_{10}} \cancel{\times 4}$$

$$= 52 + 1.25$$

$$= 53.25$$

20

Quest Calculate the mean from the following frequency distribution:

Class interval:	0-8	8-16	16-24	24-32	32-40	40-48
Frequency:	8	7	16	24	15	7

Sol^n Here $i = 8$, $M = 28$

Class-interval	Mid-Value (X_i)	Frequency (f)	$d = \frac{X_i - M}{i}$	fd
0-8	4	8	-3	-24
8-16	12	7	-2	-14
16-24	20	16	-1	-16
24-32	$\boxed{28}$	24	0	0
32-40	36	15	1	15
40-48	44	7	2	14
Total		N = 77		-25

$$\overline{X} = M + \frac{\Sigma fd}{N} \times i$$

$$= 28 + \frac{8 \times (-25)}{77}$$

$$= \frac{28 - 200}{77}$$

$$= 25.404$$

Merits of Arithmetic Mean :-

(i) Arithmetic mean is easy to calculate.
(ii) It is easy to understand
(iii) It can be used for further mathematical analysis
(iv) The nature of two or more groups of distribution can be easily compared through arithmetic mean.
(v) It represents the entire group of data in one single value.

Demerits of Arithmetic Mean :-

(i) It can be affected by extreme values.
Eg The mean for 5, 7, 250 and 475 is 184.25 which is not fairly representing any of the above data.
(ii) The qualities like intelligent, efficiency etc. could not be studied in this method.
(iii) It is difficult to calculate mean for open-end frequency distribution.
(iv) It will not represent the data in reasonable manner, if the distribution is

an abnormal one.

Median :-

Median is another frequency used measure of tendency. The median is defined as the point in a distribution with an equal no. of items on each side of it. Half of the observations fall above it and half below it. Median is generally denoted by 'M'.

Eg

X :	2	8	15	19	20	23	27
R :	1	2	3	4	5	6	7

So, Median = 19.

(2) If there is an even no. of items in the data set, the median is the average value of the two middle items, when all the items are arranged in ascending order.

(1) If the total no. of observations is odd then the median is the value of the middle ~~term~~ item when all the items are arranged in an ascending order.

→ * for the calculation of median for the ungrouped data in case of odd no. of items in a data set

$$M = \left(\frac{n+1}{2}\right) \text{th term}$$

where n is the total no. of items and M is median.

* When total no. of items in a data set is even values then median

$$M = \left(\frac{\frac{n}{2}\text{th term} + \left(\frac{n}{2}+1\right)\text{th term}}{2}\right)$$

eg The no. of flowers on 10 red flowers plants are 2, 3, 4, 5, 6, 8, 10, 12, 14, 16. find the median. Here $n = 10$.

This data is arranged in an ascending order.

Soln

$$M = \left(\frac{n}{2}\text{th term} + \left(\frac{n}{2}+1\right)\text{th term}\right)/2$$

$$M = \left(\frac{10}{2}\text{th term} + \left(\frac{10}{2}+1\right)\text{th term}\right)/2$$

$$= \left(5\text{th term} + 6\text{th term}\right)/2$$

$$M = \left(\frac{6+8}{2}\right) = 7$$

In this data, the total no. of items is even.

Ques: Find the median of the following data
6, 2, 8, 10, 9, 7, 4, 12

Solⁿ Let us arrange the data in ascending data or descending order.

2, 4, 6, 7, 8, 9, 10, 12

Here $n = 8$, (Even no.)

$$\text{Median (M)} = \frac{\left(\frac{n}{2}\text{th term} + \left(\frac{n}{2}+1\right)\text{th term}\right)}{2}$$

=

Ques: Find the median
11, 2, 4, 8, 7

Solⁿ Let us arrange the data in ascending order

2, 4, 7, 8, 11 Here $n = 5$ (odd)

$$\text{Median} = \left(\frac{n+1}{2}\right)\text{th term}$$

$$= \left(\frac{5+1}{2}\right)\text{th term.}$$

$= \left(\frac{6}{2}\right)$ th term

$=$ 3 rd term

$= 7.$

Calculation of Median from Frequency Distributions:-

In a frequency distribution the data may be in a discrete series or in a continuous series.

The median is calculated differently in both the cases.

(i) Median from Discrete Series:-

Steps:- (i) Arrange the data in an ascending order

(ii) Find the cumulative frequencies

(iii) Calculate the median by applying the formula

$$M = \text{Value of } \left(\frac{n+1}{2}\right)\text{th term} \quad [\text{if } n \text{ is odd}]$$

where $n = \Sigma f$ = Total no. of observations.

$$M = \text{Value of } \left(\frac{n}{2}\right)\text{th term} \quad [\text{if } n \text{ is even}]$$

CC

34

Page Date

Ques) Calculate the median for the following data on heights of students.

Ans

Height	No. of Students	Cumulative Frequency
40	4	4
42	5	9
45	8	17
50	10	27
60	12	39
65	10	49
68	8	57
Total.	n = 57	

Here $n = 57$

Median = The value of $\left(\frac{n+1}{2}\right)$th item

$= \left(\frac{57+1}{2}\right)$th item

= 29th item.

Median, M = 60.

Ques) Obtain the median for the following frequency distribution:

x_i	1	2	3	4	5	6	7	8	9
f:	8	10	11	16	20	25	15	9	6

Median from Continuous Series:-

- Steps:- (i) Compute the cumulative frequencies
(ii) Determine $n/2$, one half the no. of items
(iii) Find the interval in which the middle value falls and exact limits of this interval
(iv) Calculate median by using formula.

$$\text{Median} = L + \frac{\left(\frac{n}{2} - F\right)}{f_m} \times i$$

where L = lower limit of the interval containing the median

f_m = frequency of the interval containing median.

F = Sum of all frequencies above the interval containing median.

n = Total no. of frequencies $= \sum f$

i = Class interval.

Page Date

Solⁿ Computation of Median →

x	f	Cumulative frequency
1	8	8
2	10	18
3	11	29
4	16	45
5	20	65
6	25	90
7	15	105
8	09	114
9	06	120
Total	$n =$ 120	

Here $n = 120$, (even)

$$\text{Median} = \left(\frac{n}{2}\right)\text{th term}$$

$$= \left(\frac{120}{2}\right)\text{th term} = 60\text{th term}$$

$$= 5.$$

Page: Date:

Quest - Given below is the height of plants grown under restricted light. Calculate the median for this distribution.

Height (Class interval)	No. of plants (frequency)	Cumulative frequency
5-10	7	7
10-15	15	22
15-20	24	46
20-25	31	77
$\boxed{25-30}$ [redacted]	42	119
30-35	30	149
35-40	26	175
40-45	15	190
45-50	10	200
Total	$\Sigma f = 200$	

$n = 200 \Rightarrow \frac{n}{2} = \frac{200}{2} = 100$

$F = 77$, $f_m = 42$, $L = 25$

$i = 30 - 25 = 5$

$$\text{Median} = L + \frac{\frac{n}{2} - F}{f_m} \times i$$

$$= 25 + \frac{(100 - 77)}{42} \times 5$$

२६

40

Page Date

Ques i) - Find the median wage of the following distribution :

Wages (in Rs) :	2,000-3,000	3000 - 4,000	4,000 - 5,000	5,000 - 6,000	6,000 - 7,000
No. of Workers :	3	5	20	10	5

Mode →

Another measure of central tendency is the mode. The mode is the most frequently occurring value in a data set.

Eg ① Consider the observations –
13, 9, 10, 11, 11, 12, 13, 13, 13, 17, 18, 14, 15, 13

Solⁿ Here 13 occurs five times, more frequently than any other value.
So, mode = 13

Eg ② 7, 2, 4, 5, 6, 10, 10, 7, 2, 5, 4, 6

Solⁿ Here all values occur with a frequency of 2.
So, Here we can't calculate the mode.

Eg ③ 9, 9, 10, 10, 13, 13, 14, 14, 14, 15, 15, 15, 17, 18, 19, 20

Solⁿ Here the values 14 & 15 occur with the frequency of 3.

So, Mode $= \frac{14+15}{2} = \frac{29}{2} = 14.5$

Note → When two non-adjacent values of X occur such that the frequencies of both are greater than the frequencies in adjacent intervals, then each value of X may be taken as a mode.

The set of observations is then called **bimodal**.

Eg.

Mode for Continuous Frequency Distribution.

For the group data, the mode can be calculated by identifying the class with the highest frequency as follows:

$$M = L_1 + \frac{(f_m - f_1)}{(2f_m - f_1 - f_2)} \times i$$

Where L_1 = lower limit of mode class
f_m = frequency of mode class
f_1 = frequency of preceding mode class
f_2 = frequency of succeeding mode class
i = width of class interval

Ques) Find the mode for the following distribution:

Class interval	Frequency	Tally Mark
0 - 10	5	𝍸
10 - 20	8	𝍸 |||
20 - 30	7	𝍸 ||
30 - 40	12	𝍸 𝍸 ||
$\boxed{40 - 50}$	$\boxed{28}$	𝍸 𝍸 𝍸 𝍸 𝍸 |||
50 - 60	20	𝍸 𝍸 𝍸 𝍸
60 - 70	10	𝍸 𝍸
70 - 80	10	𝍸 𝍸

$$M = l_1 + \frac{(f_m - f_1)}{(2f_m - f_1 - f_2)} \times i$$

$$= 40 + \frac{(28 - 12)}{(2 \times 28 - 12 - 20)} \times 10$$

$$= 40 + \frac{16}{(56 - 32)} \times 10$$

$$= 40 + \frac{160}{24}$$

$$= 40 + 6.667$$

$$= 46.667$$

Quest- The weight (kg) of 55 students of a class are given below:

42, 74, 40, 60, 82, 115, 41, 61, 75, 83, 63, 53, 76, 84, 50, 67, 65, 78, 77, 56, 95, 68, 69, 104, 80, 79, 54, 73, 59, 81, 100, 66, 44, 77, 90, 84, 76, 42, 54, 70, 80, 72, 50, 79, 52, 103, 96, 51, 86, 78, 94, 71, 20, 70, 79.

Calculate the mode.

Class - interval	Frequency f	Tally Mark
40 - 50	05	𝍸
50 - 60	09	𝍸 ||||
60 - 70	09	𝍸 ||||
70 - 80	15	𝍸 𝍸 𝍸
80 - 90	08	𝍸 |||
90 - 100	04	||||
100 - 110	03	|||
110 - 120	02	||

$f_M = 15, \quad L_1 = 70, \quad f_1 = 09, \quad f_2 = 8$

$i = 80 - 70 = 10$

Median $M = L_1 + \frac{f_M - f_1}{2f_M - f_1 - f_2} \times i$

$= 70 + \frac{(15 - 09)}{(2 \times 15 - 9 - 8)} \times 10$

$= 70 + \frac{6}{(30 - 17)} \times 10$

$= 70 + \frac{60}{13}$

$= 70 + 4.615$

$= 74.615$

Some Other Averages

(i) Geometric Mean →: It is especially useful as a measure of central tendency when the values change exponentially.

If there are only two observations, then the G.M. is the square root of the product of two observations. If there are three observations then it is the cube root of product of three observations.

Thus, if there are 'n' observations, than the G.M will be nth root of the product of the 'n' observations.

$$GM = \sqrt[n]{(X_1)(X_2)\ldots\ldots(X_n)}$$

$$GM = \text{Antilog}\left[\frac{\log X_1 + \log X_2 + \ldots + \log X_n}{n}\right]$$

Ques:- The no. of bacteria ($\times 10^3$) observed in an experiment at hourly intervals are as follows:

10, 38, 50, 151, 296

Calculate the geometric mean of these 5

values.

Sol^n The G.M. of 5 values $= \sqrt[5]{10 \times 38 \times 50 \times 151 \times 996}$

$= (84922400 0)^{1/5}$

$= 61.0667$

Ques:- Calculate G.M for the following data.

Height Class interval	No. of plants frequency	Mid Value (X)	log x	log x × f
0 - 100	20	50	1.6990	33.98
100 - 200	15	150	2.1761	32.6415
200 - 300	12	250	2.3979	28.7748
300 - 400	13	350	2.5441	33.0733
400 - 500	5	450	2.6532	13.266
500 - 600	25	550	2.7404	68.51
600 - 700	17	650	2.8129	47.8193
700 - 800	5	750	2.8751	14.3755
	$\Sigma f = 112$			$\Sigma f \log x = 272.4404$

$$G.M. = \text{Antilog} \frac{\Sigma f \log x}{\Sigma f}$$

$$= \text{Antilog} \frac{272.4404}{112}$$

$$= \text{Antilog } 2.4325$$

$$= 279.7$$

Merits and Demerits of Geometric Mean →

(i) It is rigidly defined
(ii) It can be used for further statistical analysis
(iii) Geometric mean can be used for averaging ratios, rates and percentage
(iv) It is suitable for calculating index number

• Demerits →

(i) It is difficult to calculate and understand
(ii) If any of the value in a series is zero, geometric mean will also be zero. Hence it can't be calculated for such kind of series
(iii) It can't be calculated when the distribution has open-end class.

Harmonic Mean:

It may be defined as the total number of observations divided by the sum of reciprocals of the numbers. It may be expressed as follows:

$$H.M. = \frac{n}{\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} + \cdots + \frac{1}{X_n}} = \frac{n}{\sum \frac{1}{X_n}}$$

where n = Total no. of variables

$\sum \frac{1}{X}$ = Total no. of reciprocals of variables.

Ques:- Compute harmonic mean of the data 2, 97, 150, 5, 285, 427, 550

Soln

X	1/X
2	0.5
97	0.0103
150	0.0067
5	0.2
285	0.0035
427	0.0023
550	0.0018
$n = 7$	$\sum \frac{1}{X} = 0.7246$

$$H.M. = \frac{n}{\Sigma \frac{1}{X}} = \frac{7}{0.7246} = 9.6605$$

Ques)- Calculate the H.M. of the following data,

Age	No. of class (f)	Mid Value	f/X
0-10	7	5	1.4
10-20	8	15	0.533
20-30	20	25	0.8
30-40	11	35	0.314
40-50	9	45	0.2
50-60	3	55	0.055
60-70	4	65	0.062
	n = 62		Σ f/x = 3.364

$$H.M = \frac{n}{\Sigma f/X} = \frac{62}{3.364} = 18.43$$

Merits of H.M. →

(i) It is based on all the observation
(ii) It can be used for further algebraic analysis.
(iii) It is rigidly defined

(iv) It can be used for averaging rates and ratios

: Demerits of H.M. →

(i) It is difficult to understand and to calcu-late

(ii) It can't be calculated if any one or more values of the variables is zero.

(iii) It can't be calculated if there exists both positive and negative values in a series.