4 Average (Measure of Central Tendency)

When statistical data is properly classified and condensed into a frequency distribution, it is easy for us to study the different characteristics of data. Further, graph and diagrams can also be drawn to convey a better impression to the mind about the data. But these are just visual aids, to know more about the data, we need some descriptive measures. With the results of classification and tabulation, we can only make a qualitative comparison of the data. Often this is not enough and it is necessary to replace a collection of data by a single number. So, for quantitative analysis of the data, we require some statistical measures like central tendency and dispersion.

According to the statistician, Professor Bowley "Measures of Central Tendency (averages) are statistical constants which enable us to comprehend in single effort the significant of the whole".

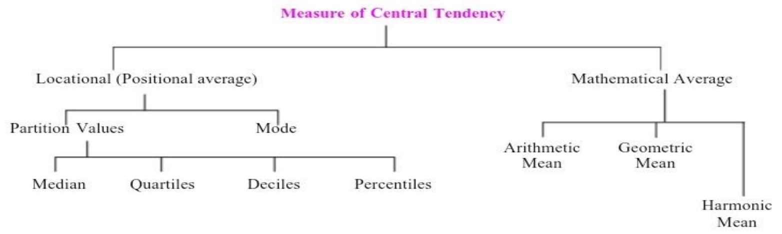
The main objectives of Measure of Central Tendency are :

1. To condense data in a single value.
2. To facilitate comparisons between data.

There are different types of averages, each has its own advantages and disadvantages.

Requisites of a Good Measure of Central Tendency

1. It should be rigidly defined.
2. It should be simple to understand and easy to calculate.
3. It should be based on all the observations of the data.
4. It should be capable of further mathematical treatment.
5. It should be least affected by the fluctuations of the sampling.
6. It should not be unduly affected by the extreme values.
7. It should be easy to interpret.



ARITHMETIC MEAN

Arithmetic mean or mean or average is the most commonly used single descriptive measure of Central Tendency. Mean is simple to compute, easy to understand and interpret. The arithmetic mean is obtained by summing up all elements of the data set and dividing by the number of elements. There are two types of mean namely :

- Simple Arithmetic Mean
- Weighted Arithmetic Mean

1. Simple Arithmetic Mean

This is the most common and useful, single descriptive measure of Central Tendency.

ARITHMETIC MEAN OF RAW DATA

In an ungrouped or raw data, we are given individual items. We also know that the average of n numbers is obtained by finding their sum (by adding) and then dividing it by n. We use the numbers 1, 2, 3, ….. n as subscripts, where the letter 'X' represent an observation. Thus X_1 represent the first observation, X_2 represent the second observation and so on and X_n represent the nth observation. The sample mean is represented by $\overline{X}$ bar and written as

$$\overline{X} = \frac{X_1 + X_2 + X_3 + X_4 + \ldots.\, X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Example : 1

Consider the marks scored by 10 students in Mathematics in a certain examination 25, 35, 17, 15, 45, 30, 55, 21, 47, 10. Find the average performance of the students.

Solution :

Let $\overline{X}$ be the average marks

$\therefore$ sum of all the observations $= 25 + 35 + 17 + 15 + 45 + 30 + 55 + 21 + 47 + 10 = 300$

No. of students = 10

$\therefore$ Arithmetic Mean $= \frac{\Sigma X_i}{n} = \frac{300}{10} = 30$

Example : 2

From the following data of scores of a player in 9 matches, calculate the player's average score, 23, 40, 50, 05, 85, 89, 12, 65, 45

Solution :

No. of matches $n = 9$. Sum of all observations $\Sigma X = 414$.

$$\therefore \quad \overline{X} = \frac{\Sigma X}{n} = \frac{414}{9} = 46.$$

CASE II : WHEN THE DATA IS GROUPED

a) By Direct Method : If a variate X take values X_1 , X_2 , X_3 …… X_n with corresponding frequencies f_1 , f_2 , f_3 ….. f_n respectively, then the arithmetic mean of these values is given by

$$\overline{X} = \frac{f_1 X_1 + f_2 X_2 + f_3 X_3 + \ldots\ldots + f_n X_n}{f_1 + f_2 + f_3 + \ldots\ldots + f_n}$$

or $$\overline{X} = \frac{\sum_{i=1}^{n} f_i X_i}{n}$$

The following algorithm may be used to compute arithmetic mean by direct method.

Example : 3

Find the mean of the following distribution :

X :	04	06	09	10	15
f :	05	10	10	07	08

Solution :

X_i	f_i	$f_i X_i$
04	05	20
06	10	60
09	10	90
10	07	70
15	08	120
	$\Sigma f_i = 40$	$\Sigma f_i X_i = 360$

$$\therefore \text{ Mean } = \overline{X} = \frac{\Sigma f_i X_i}{\Sigma f_i} = \frac{360}{40} = 9$$

Example : 4

If the mean of the following distribution is 6, find the value of P.

X :	2	4	6	10	P + 5
f :	3	2	3	1	2

Solution :

X_i	f_i	$f_i X_i$
2	3	6
4	2	8
6	3	18
10	1	10
P + 5	2	2P + 10
	$N = \Sigma f_i = 11$	$\Sigma f_i X_i = 2P + 52$

$$\text{Mean} = \frac{\Sigma f_i X_i}{\Sigma f_i} \Rightarrow 6 = \frac{2P + 52}{11} \Rightarrow 66 = 2P + 52 \Rightarrow 2P = 14 \Rightarrow P = 7.$$

b) By Short-cut Method : If the values of X or (and f) are large, the calculation of AM by the direct method is quite tedious and time consuming, because calculations involved are lengthy. In such a case to minimize the time involved in calculation, we take deviations from an arbitrary point.

Note : The number 'A' is generally known as the assumed mean and is generally chosen in such a way that the deviations are small.

ALGORITHM

Step I : Prepare the frequency table in such a way that its first column consists of the values of the variable and the second column consists of the corresponding frequencies.

Step II : Choose a number 'A' (preferable among the values in first column) and take deviations $d_i = X_i - A$ of the values X_i of variable 'X' about A. Write these deviations against the corresponding frequencies in the third column.

Step III : Multiply the frequencies in column II with the corresponding deviations d_i in column III to prepare column IV consisting of $f_i\, d_i$.

Step IV : Find the sum of all entries in column IV to obtain $\sum_{i=1}^{n} f_i\, d_i$ and the sum of all frequencies in column II to obtain $\sum_{i=1}^{n} f_i = N$

Step V : Use the formula : $\overline{X} = A + \dfrac{\sum_{i=1}^{n} f_i d_i}{N}$

Example : 6

The following table shows the weight of 12 students :

Weight (in kg) :	67	70	72	73	75
Number of Students :	04	03	02	02	01

Find the mean weight by using short-cut method.

Solution :

Let the assumed mean be A = 72.

Weight in kg X_i	No. of Students f_i	$d_i = X_i - A$ $= X_i - 72$	$f_i d_i$
67	4	-5	-20
70	3	-2	-6
72	2	0	0
73	2	1	2
75	1	3	3
	$N = \Sigma f_i = 12$		$\Sigma f_i d_i = -21$

$$\text{Mean} = A + \frac{\Sigma f_i d_i}{N} = 72 + \left(\frac{-21}{12}\right) = 72 - \frac{7}{4}$$

$$= 72 - 1.75 = 70.25 \text{ kg}$$

$$\therefore \text{ Mean weight} = 70.25 \text{ kg.}$$

Example : 11

The mean of the following frequency table is 50. Find the missing frequencies f_1 and f_2.

Class :	0–20	20–40	40–60	60–80	80–100	Total
Frequency :	17	f_1	32	f_2	19	120

Solution :

Class	Frequency	Class-mark X_i	$f_i X_i$
0–20	17	10	170
20–40	f_1	30	$30\,f_1$
40–60	32	50	1600
60–80	f_1	70	$70\,f_2$
80–100	19	90	1710
	$\Sigma f_i = 68 + f_1 + f_2 = N = 120$		$\Sigma f_i X_i = 3480 + 30\,f_1 + 70\,f_2$

So, we have $\Sigma f_i = 68 + f_1 + f_2$

$$\Sigma f_i X_i = 3480 + 30\,f_1 + 70\,f_2$$

$$\overline{X} = \frac{\Sigma f_i X_i}{N} \Rightarrow 50 = \frac{3480 + 30\,f_1 + 70 f_2}{120}$$

$\Rightarrow 6000 = 3480 + 30\,f_1 + 70\,f_2$

$\Rightarrow 30\,f_1 + 70\,f_2 = 6000 - 3480$

$30\,f_1 + 70\,f_2 = 2520$

$\Rightarrow 3\,f_1 + 7\,f_2 = 252$ ...(A)

Also, $68 + f_1 + f_2 = 120 \Rightarrow f_1 + f_2 = 52$ ...(B)

Solving equation (A) and (B) simultaneously

$f_1 =$ **28** $f_2 =$ **24**

PROPERTIES OF ARITHMETIC MEAN

1. The sum of deviations, of all the values of X, from their arithmetic mean is zero.

   i.e. $\sum_{1}^{n} f_i(X_i - \overline{X} = 0$

2. The product of the arithmetic mean and the number of items gives the total of all items.

   i.e. $\overline{X} = \dfrac{\Sigma f_i X_i}{\Sigma f_i} \Rightarrow \Sigma f_i X_i = \overline{X} \cdot \Sigma f_i$

   or $\overline{X} = \dfrac{\Sigma X_i}{N} \Rightarrow \overline{X} \cdot N = \Sigma X_i$

3. If $\overline{X_1}$ and $\overline{X_2}$ are the arithmetic mean of two samples of size n_1 and n_2 respectively then, the arithmetic mean $\overline{X}$ of the distribution combining the two can be calculated as

   $$\overline{X} = \frac{n_1\overline{X_1} + n_2\overline{X_2}}{n_1 + n_2}$$

   This formula can be extended for still more group or samples.

   $$\overline{X_1} = \frac{\Sigma X_1 i}{n_1} \Rightarrow \Sigma X_1 i = n_1 \overline{X_1}$$

xample : 12

The average marks of three batches of students having 70, 50 and 30 students respectively are 50, 55 and 45. Find the average marks of all the 150 students taken together.

Solution :

$$n_1 = 70, \quad n_2 = 50, \quad n_3 = 30$$

$$X_1 = 50 \quad X_2 = 55 \quad X_3 = 45$$

$$\therefore \quad \overline{X} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2 + n_3\overline{X}_3}{n_1 + n_2 + n_3} = \frac{70 \times 50 + 50 \times 55 + 30 \times 45}{70 + 50 + 30} = \frac{7600}{150}$$

$$\therefore \quad \overline{X} = 50.67 \text{ marks}$$

Example : 13

Example : 13

The average marks of a group of a 100 students in statistics are 60 and for other group of 50 students, the average marks are go. Find the average marks combined group of 150 students.

Solution :

$$n_1 = 100 \qquad n_2 = 50$$

$$\overline{X_1} = 60 \qquad \overline{X_2} = 90$$

$$\overline{X} = \frac{n_1\overline{X_1} + n_2\overline{X_2}}{n_1 + n_2}$$

$$= \frac{100 \times 60 + 50 \times 90}{100 + 50}$$

$$= \frac{10,500}{150}$$

$$= 70$$

Hence the average marks of all 150 students is 70.

Example : 14

Example : 14

An average daily wages of all the 90 workers in a factory is ₹ 60. An average daily wages of female workers is ₹ 45. Calculate an average daily wages of male workers if one third workers are male.

Solution :

n_1 = No. of male workers = 90 × 1/3 = 30

n_2 = No. of female workers = 90 × 2/3 = 60

$\overline{X_1}$ = mean daily wages of female workers = 45

Combined Mean

$$\overline{X} = \frac{n_1\overline{X_1} + n_2\overline{X_2}}{n_1 + n_2}$$

$$60 = \frac{30\,\overline{X_1} + 60 \times 45}{30 + 60}$$

$$60 \times 90 = 30\,\overline{X_1} + 2{,}700$$

$$30\,\overline{X_1} = 5{,}400 - 2{,}700$$

$$\overline{X_1} = \frac{2{,}700}{30}$$

$$= 90$$

Hence, the average daily wages workers is ₹ 90.

6. Mean cannot be calculated when open end class intervals are present in the data.

WEIGHTED ARITHMETIC MEAN

When the different items in a data are assigned weights according to their significance (relatively), the average then calculated is called as weight arithmetic mean.

Symbolically, the weight average of '*n*' items $X_1, X_2, X_3, \ldots\ldots X_n$ whose weights are $W_1, W_2, W_3 \ldots\ldots W_n$ respectively is :

$$\frac{X_1 W_1 + X_2 W_2 + X_3 W_3 + \ldots\ldots + X_n W_n}{W_1 + W_2 + W_3 + \ldots\ldots + W_n}$$

Example : 16

If a person buys five copies of the book at ₹ 12 each, three copies at ₹ 18 each and two copies at ₹ 30 each. Find the average price of a book.

...ution :

$$\overline{X} = \frac{12 \times 5 + 18 \times 3 + 30 \times 2}{5 + 3 + 2} = \frac{60 + 54 + 60}{10} = ₹\,17.40$$

EXERCISE

MEDIAN

The median is the middle value of a distribution i.e. median of a distribution is the value of the variable which divides it into two equal parts. It is the value of the variable such that the number of observations above it is equal to the number of observations below it. Observations are arranged either in ascending order or descending order of their magnitude. Median is a position average whereas the arithmetic mean is a calculated average.

Calculation of Median

When the data is ungrouped

Algorithm

Step I : Arrange the observations X_1 , X_2 , ... X_n is ascending or descending order of magnitude.

Step II : Determine the total no of observations, say n.

Step III : If 'n' is odd, then median is the value of $\left(\frac{n+1}{2}\right)^{\text{th}}$ observation.

If 'n' is even, then median is the AM of the values of $\left(\frac{n}{2}\right)^{\text{th}}$ and $\left(\frac{n}{2}+1\right)^{\text{th}}$ observations.

i.e. M_d or $M = \dfrac{\left(\frac{n}{2}\right)^{\text{th}} \text{ term} + \left(\frac{n}{2}+1\right)^{\text{th}} \text{ term}}{2}$

Example :

Example :

The following are the marks of 9 students in a class. Find the median.

34 , 32 , 48 , 38 , 24 , 30 , 27 , 21 , 35

Solution :

Arranging the data in ascending order of magnitude. We have

21, 24 , 27 , 30 , 32 , 34 , 35 , 38 , 48

Since there are 9 i.e. an odd no. of terms. Therefore, Median is the value of $\left(\frac{9+1}{2}\right)^{th}$ observation i.e. 32.

$\therefore$ **Median = 32**

Example : 17

Find the median of the daily wages of ten workers from the following data :

₹ 20 , 25 , 17 , 18 , 8 , 15 , 22 , 11 , 9 , 14

Solution :

Arranging the wages in ascending order of magnitude, we have

8 , 9 , 11 , 14 , 15 , 17 , 18 , 20 , 22 , 25

Since there are 10 observation, therefore, median is the arithmetic mean of $\left(\frac{10}{2}\right)^{th}$ and $\left(\frac{10}{2}+1\right)^{th}$ observations

$\therefore$ Hence median $= \frac{15+17}{2} = 16$

When the Data is Grouped

CASE I : Discrete Data

Algorithm

Step I : Arrange the value of the variables in ascending or descending order of their magnitude.

Step II : Find the cumulative frequency (C. F)

Step III : Find $\frac{N}{2}$ where $N = \sum_{i=1}^{n} f_i$

Step IV : See the cumulative frequency (C. F) just greater than $\frac{N}{2}$ and determine the corresponding value of the variable.

Step V : The value obtained in step IV is the median.

Example : 19

Obtain the median for the following frequency distribution :

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

Solution :

X	*f*	CF
1	8	8
2	10	18
3	11	29
4	16	45
5	20	65
6	25	90
7	15	105
8	9	114
9	6	120
	$N = \Sigma f = 120$	

Here $N = 120 \Rightarrow \frac{N}{2} = 60$

We find that the cumulative frequency just greater than $\frac{N}{2}$

i.e., 60 is 65 and the value of x, corresponding to 65 is 5.

Therefore, Median = 5.

Example : 20

Example : 20

Find the median of the following frequency distribution.

X :	5	7	9	12	14	17	19	21	Total
f :	6	5	3	6	5	3	2	4	34

Solution :

X :	5	7	9	12	14	17	19	21	Total
f :	6	5	3	6	5	3	2	4	34
C.F	6	11	14	20	25	28	30	34	–

N = 34 i.e. even no. of observations

$$\therefore \text{ Mean} = \text{Average of} \left(\frac{N}{2}\right)^{th} \text{and} \left(\frac{N}{2}+1\right)^{th} \text{observations}$$

$$= \text{Average of 17th and 18th observations}$$

$$= \frac{12+12}{2} = 12$$

The above table shows that all items between 12 to 14 have their values 9.

Since 12th item lies in this interval therefore its value is 9.

Example : 21

The median of the observations 8 , 11 , 13 , 15 , X + 1 , X + 3 , 30 , 35 , 40 , 43 arranged in ascending order is 22. Find the value of X.

Solution :

The observations arranged in ascending order are 8 , 11 , 13 , 15 , X + 1 , X + 3 , 30 , 35 , 40 , 43

Total No. of observations N = 10 (even)

$$\therefore \text{ Median } = \frac{\left(\frac{N}{2}\right)^{th} + \left(\frac{N}{2}+1\right)^{th} \text{ observation}}{2} = \frac{5^{th}\text{ obs} + 6^{th}\text{ obs}}{2}$$

$$22 = \frac{(X+1)+(X+3)}{2} \Rightarrow 44 = 2X + 4 \Rightarrow X = 20$$

CASE II : Continous Data

In order to calculate the median of a grouped or continuous frequency distribution, we use the following algorithm.

Algorithm

Step I : Obtain the frequency distribution.

Step II : Prepare the cumulative frequency column and obtain $N = \Sigma f_i$

Step III : Find N/2

Step IV : See the cumulative frequency just greater than N/2 and determine the corresponding class. This class is known as the median class.

Step V : Use the following formula :

$$\text{Median} = l_1 + \frac{\frac{N}{2} - CF}{F} \times (l_2 - l_1)$$

Where l_1 = lower limit of the median class

l_2 = upper limit of the median class

F = Frequency of the median class

CF = cumulative frequency of the class preceding the median class

Example : 22

Calculate the median from the following distribution :

CI	5–10	10–15	15–20	20–25	25–30	30–35	35–40	40–45	Total
Frequency	5	6	15	10	5	4	2	2	49

Solution :

Let us first prepare a table which gives the frequencies and cumulative frequency to compute the median.

Class Interval	Frequency	C. F.
5–10	5	5
10–15	6	11 $< N/2$
15–20	15	26 $> \frac{N}{2}$
20–25	10	36
25–30	5	41
30–35	4	45
35–40	2	47
40–45	2	49
	N = 49	

We have N = 49. $\therefore \frac{N}{2} = \frac{49}{2} = 24.5$

The cumulative frequency just greater than $\frac{N}{2}$ is 26 and the corresponding class is 15–20 is median class such that $l_1 = 15$, $l_2 = 20$, $f = 15$, CF = 11

$$\therefore \quad \text{Median} = l_1 + \frac{\frac{N}{2} - CF}{F} \times (l_2 - l_1)$$

$$= 15 + \frac{24.5 - 11}{15} \times (20 + 5)$$

$$= 15 + \frac{13.5 \times 15}{15} = 19.5$$

$$\therefore \quad \text{Median} = 19.5$$

Example : 23

Compute the media from the following data :

Mid-Value	115	125	135	145	155	165	175	185	195
Frequency	6	25	48	72	116	60	38	22	3

Solution :

Here, we are given the mid-values. So, should first find the upper and lower limits of the various classes. The difference between two consecutive values is $h = 125 - 115 = 10$.

$\therefore$ Lower limit of a class = Mid value $- h/2$

Upper limit of a class = Mid-value $+ h/2$

Mid Value	Class Groups	Frequency	Cumulative Frequency
115	110–120	6	6
125	120–130	25	31
135	130–140	48	79
145	140–150	72	151 $< N/2$
155	150–160	116	267 $> N/2$
165	160–170	60	327
175	170–180	38	365
185	180–190	22	387
195	190–200	3	390
			$N = \Sigma f_i = 390$

We have, $N = 390$, $\therefore \frac{N}{2} = \frac{390}{2} = 195$

The cumulative frequency just greater than $\frac{N}{2}$ i.e. 195 is 267 and the corresponding class is 150–160. So, 150–160 is the median class.

$\therefore \quad l_1 = 150, \quad l_2 = 160, \; f = 116, \quad CF = 151.$

$$\text{Now, Median} = l_1 + \frac{\frac{N}{2} - CF}{F} \times (l_2 - l_1)$$

$$= 150 + \frac{195 - 151}{116} \times 10$$

$$= 150 + \frac{44 \times 10}{116} = 153.80$$

PARTITION VALUES

Median divides an arrayed series in ascending or descending series into two equal parts. When we are required to divide an arrayed series into more than two equal parts, the dividing places are known as partition values.

Quartiles, Deciles and percentiles are called partition values. Partition values are used to study the scatteredness of the values of the variable in relation to the median. Thus, the special use of partition value is to study dispersion of items in relation to the median i.e. it helps us in understanding the composition of a series.

QUARTILES

Quartiles are those values of the variate which divides an arrayed series into four equal parts.

Each portion contains equal number of items. The first, second and third points are termed as first Quartile (Q_1), second Quartile (commonly called as median) and third Quartile (Q_3). It is important to note that $Q_1 < Q_2 < Q_3$.

Computation of Quartiles for a frequency distribution with class intervals.

Algorithm

Step I : Compute the cumulative frequency of the given distribution.

Step II : Compute $\frac{iN}{4}$. Where $i = 1$ for lower Quartile Q_1, $i = 2$ for middle quartile Q_2 (median), $i = 3$ for upper quartile Q_3.

Step III : Find the cumulative frequency just greater than $\frac{iN}{4}$ and the corresponding class. This class is called the **quartile class**.

Step IV : Use the following formula to calculate Q_1 , Q_2 or Q_3

$$Q_i = l_1 + \frac{\frac{iN}{4} - CF}{F} \times (l_2 - l_1)$$

Algorithm

Step I : Compute the cumulative frequency of the given distribution.

Step II : Compute $\frac{iN}{4}$. Where $i = 1$ for lower Quartile Q_1, $i = 2$ for middle quartile Q_2 (median), $i = 3$ for upper quartile Q_3.

Step III : Find the cumulative frequency just greater than $\frac{iN}{4}$ and the corresponding class. This class is called the **quartile class**.

Step IV : Use the following formula to calculate Q_1, Q_2 or Q_3

$$Q_i = l_1 + \frac{\frac{i\,N}{4} - CF}{F} \times (l_2 - l_1)$$

Where l_1 = lower limit of the quartile class

F = Frequency of the quartile class

CF = cumulative frequency of the class preceding the median class

l_2 = upper limit of the quartile class

Example : 24

Calculate the quartiles i.e. Q_1 and Q_3 from the following data :

Marks	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
No. of Students	08	07	05	12	28	22	08	10

Solution :

Marks	No. of Students	Cumulative Frequency (C.F)
0–10	08	08
10–20	07	15
20–30	05	20 $< n/4$
30–40	12	32 $> n/4$
40–50	28	60 $< 3n/4$
50–60	22	82 $> 3n/4$
60–70	08	90
70–80	10	100

$$Q_1 = \text{Size of} \left(\frac{N}{4}\right)^{th} \text{term} = \frac{100}{4} \text{ i.e. size of 25th term}$$

Q_1 lies in the class interval 30–40

$\therefore \quad l_1 = 30, \quad l_2 = 40, \quad F = 12, \quad CF = 20$

$$\therefore \quad Q_1 = l_1 + \frac{\frac{N}{4} - CF}{F} \times (l_2 - l_1) = 30 + \frac{25 - 20}{12} \times 10 = \mathbf{34.17}$$

Also,

$$Q_3 = \text{Size of} \left(\frac{3N}{4}\right)^{th} \text{term} = \text{75th term}$$

Q_3 lies in the class interval 50–60

$\therefore \quad l_1 = 50, \quad l_2 = 60, \quad F = 22, \quad CF = 60$

$$\therefore \quad Q_3 = l_1 + \frac{\frac{3N}{4} - CF}{F} \times (l_2 - l_1) = 50 + \frac{75 - 60}{22} \times 10 = 56.81$$

Hence $Q_1 = 34.17$, $Q_3 = 56.81$

$\therefore$ Median $= 30.91$

Example : 26

From the following data, calculate median.

Marks More than	0	10	20	30	40	50	60
No. of Students	180	170	150	120	70	30	0

Solution :

Let us first convert more than type data into exclusive type of class interval.

Marks	No. of Students (f)	Cumulative Frequency (C.F)
0 – 10	10	10
10 – 20	20	30
20 – 30	30	60 < N/2
30 – 40	50	110 > N/2
40 – 50	40	150
50 – 60	30	180
	$\Sigma f = N = 180$	

Median is $\frac{N}{2} = \frac{180}{2} = 90^{th}$ term

The C.F. just greater than $\frac{N}{2}$ is 110 and the corresponsing class is 30 – 40 is the median class such that $l_1 = 30$, $l_2 = 40$, $F = 50$, $CF = 60$.

$$\therefore \quad \text{Median} = l_1 + \frac{\frac{N}{2} - CF}{F} \times (l_2 - l_1)$$

$$= 30 + \frac{90 - 60}{50} \times (40 - 30)$$

$$= 30 + \frac{30 \times 10}{50}$$

$$= 36$$

$\therefore$ Median $= 36$

DECILES

The value of the variable which divide the series, when arranged in ascending order in 10 equal parts is called as Decile. Deciles are denoted by $D_1, D_2, D_3 \ldots\ldots D_9$.

Computation of Deciles for a Frequency Distribution

Algorithm

Step I : Compute the cumulative frequency table.

Step II : Compute $\frac{iN}{10}$ to find D_i, where ith Decile $i = 1, 2, 3 \ldots\ldots 9$.

Step III : Find the C.F. just greater than $\frac{iN}{10}$ and the corresponding class. This class is called the decile class.

Step IV : Use the formula

$$D_i = l_1 + \frac{\frac{iN}{10} - CF}{F} \times (l_2 - l_1)$$

Example : 27

Compute D_3 and D_7 for the following frequency distribution.

Marks	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
No. of Students	5	7	8	12	28	22	10	08

Solution :

The cumulative frequency distribution is given below :

Marks	No. of Students	Cumulative Frequency (C.F)
0–10	5	5
10–20	7	12
20–30	8	20 $< 3N/10$
30–40	12	32 $> 3N/10$
40–50	28	60 $< 7N/10$
50–60	22	82 $> 7N/10$
60–70	10	90
70–80	08	100
	$N = \Sigma f = 100$	

Third Decile $= D_3$ $\quad N = 100 \quad = \frac{3N}{10} = \frac{3 \times 100}{10} = 30$

The cumulative frequency just greater than 30 is 32. Thus the third Decile class is 30–40.

$\therefore \quad l_1 = 30, \quad l_2 = 40, \quad F = 12, \quad CF = 20$

$$\therefore \quad D_3 = l_1 + \frac{\frac{3N}{10} - CF}{F} \times (l_2 - l_1)$$

$$= 30 + \frac{30 - 20}{12} \times (40 - 30) = 30 + \frac{10 \times 10}{12} = 38.33$$

$$D_3 = 38.33$$

Seventh Decile $= D_7$. $\quad$ Here $N = 100$ $\quad \frac{7N}{10} = \frac{7 \times 100}{10} = 70$

The cumulative frequency just greater than 70 is 82 and the corresponding class is 50–60.

$\therefore \quad l_1 = 50, \quad l_2 = 60, \quad F = 22, \quad CF = 60$

DECILES

The value of the variable which divide the series, when arranged in ascending order in 10 equal parts is called as Decile. Deciles are denoted by $D_1, D_2, D_3 \ldots\ldots D_9$.

Computation of Deciles for a Frequency Distribution

Algorithm

Step I : Compute the cumulative frequency table.

Step II : Compute $\frac{iN}{10}$ to find D_i, where ith Decile $i = 1, 2, 3 \ldots\ldots 9$.

Step III : Find the C.F. just greater than $\frac{iN}{10}$ and the corresponding class. This class is called the decile class.

Step IV : Use the formula

$$D_i = l_1 + \frac{\frac{iN}{10} - CF}{F} \times (l_2 - l_1)$$

Example : 27

Example : 27

Compute D_3 and D_7 for the following frequency distribution.

Marks	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
No. of Students	5	7	8	12	28	22	10	08

Solution :

The cumulative frequency distribution is given below :

Marks	No. of Students	Cumulative Frequency (C.F)
0–10	5	5
10–20	7	12
20–30	8	20 $< 3N/10$
30–40	12	32 $> 3N/10$
40–50	28	60 $< 7N/10$
50–60	22	82 $> 7N/10$
60–70	10	90
70–80	08	100
	$N = \Sigma f = 100$	

Third Decile = D_3 $\quad N = 100 \quad = \frac{3N}{10} = \frac{3 \times 100}{10} = 30$

The cumulative frequency just greater than 30 is 32. Thus the third Decile class is 30–40.

$\therefore \quad l_1 = 30, \quad l_2 = 40, \quad F = 12, \quad CF = 20$

$$\therefore \quad D_3 = l_1 + \frac{\frac{3N}{10} - CF}{F} \times (l_2 - l_1)$$

$$= 30 + \frac{30 - 20}{12} \times (40 - 30) = 30 + \frac{10 \times 10}{12} = 38.33$$

$$D_3 = 38.33$$

Seventh Decile = D_7. $\quad$ Here $N = 100$ $\quad \frac{7N}{10} = \frac{7 \times 100}{10} = 70$

The cumulative frequency just greater than 70 is 82 and the corresponding class is 50–60.

$\therefore \quad l_1 = 50, \quad l_2 = 60, \quad F = 22, \quad CF = 60$

$$\therefore \quad D_7 = l_1 + \frac{\frac{7N}{10} - CF}{F} \times (l_2 - l_1)$$

$$= 50 + \frac{70 - 60}{22} \times (60 - 50) = 50 + \frac{10 \times 10}{22} = 54.55$$

$$\therefore \quad D_7 = 54.55$$

Hence $D_3 = 38.33$, $D_7 = 54.55$

PERCENTILE

The value of the variable which divide the series, when arranged in ascending order into 100 equal parts are called percentiles. Percentiles are denoted by $P_1, P_2, P_3, \ldots\ldots P_{99}$.

Computation of Percentile for a Frequency Distribution

Algorithm

Step I : Compute the cumulative frequency table.

Step II : Compute $\frac{iN}{10}$ to find P_i, where i^{th} percentile = 1, 2, 3, ……. 99

Step III : Find C.F just greater than $\frac{iN}{10}$ and the corresponding class. This class is called the percentile class.

Step IV : Use the formula.

$$P_i = l_1 + \frac{\frac{iN}{100} - CF}{F} \times (l_2 - l_1)$$

Example : 28

Example : 28

For the following frequency distribution, calculate P_8, P_{23} and P_{90}.

Marks	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
No. of Students	05	07	08	12	28	22	10	08

Solution :

Preparing the cumulative frequency table :

Marks	No. of Students (f)	Cumulative Frequency (C.F)
0–10	05	05
10–20	07	12
20–30	08	20
30–40	12	32
40–50	28	60
50–60	22	82
60–70	10	92
70–80	08	100
	$\Sigma f = N = 100$	

8th percentile = P_8 N = 100 $\quad \frac{8N}{100} \quad = \frac{8 \times 100}{100} = 8$

The cumulative frequency just greater than 8 is 12.

Thus 8th percentile class is 10–20.

$\therefore \quad l_1 = 10, \quad l_2 = 20, \quad F = 07, \quad CF = 05$

$$\therefore \quad P_8 = l_1 + \frac{\frac{8N}{100} - CF}{F} \times (l_2 - l_1)$$

Example : 28

For the following frequency distribution, calculate P_8, P_{23} and P_{90}.

Marks	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
No. of Students	05	07	08	12	28	22	10	08

Solution :

Preparing the cumulative frequency table :

Marks	No. of Students (f)	Cumulative Frequency (C.F)
0–10	05	05
10–20	07	12
20–30	08	20
30–40	12	32
40–50	28	60
50–60	22	82
60–70	10	92
70–80	08	100
	$\Sigma f = N = 100$	

8th percentile = P_8 N = 100 $\quad \frac{8N}{100} \quad = \frac{8 \times 100}{100} = 8$

The cumulative frequency just greater than 8 is 12.

Thus 8th percentile class is 10–20.

$\therefore \quad l_1 = 10, \quad l_2 = 20, \quad F = 07, \quad CF = 05$

$$\therefore \quad P_8 = l_1 + \frac{\frac{8N}{100} - CF}{F} \times (l_2 - l_1)$$

GEOMETRIC MEAN

It is another measure of central tendency based on mathematical footing like arithmetic mean. Geometric mean can be defined in the following terms :

"Geometric mean is the n^{th} positive root of the product of 'n' positive given values".

Hence, geometric mean for a value X containing n values such as $x_1, x_2, x_3, \ldots.. x_n$ is denoted by G.M of X and given as under :

$$\text{G.M of X} = \overline{\text{X}} = \sqrt[n]{x_1\,,\, x_2\,,\, x_3\,,\, \ldots..\,,\, x_n}$$

If we have a series of n positive values with repeated values such as $x_1, x_2, x_3, \ldots. x_k$ are repeated $f_1, f_2, f_3, \ldots. f_k$ times respectively then geometric mean will become

$$\text{GM of X} = \overline{\text{X}} = \sqrt[n]{x_1^{f_1}\,,\, x_2^{f_2}\,,\, x_3^{f_3}\,,\, \ldots.\, x_k^{f_k}}$$

Where $n = f_1 + f_2 + f_3 + \ldots + f_k$

If $X_1, X_2, X_3, \ldots.., X_n$ are n values of a variable X, none of them being zero, then the geometric mean 'G' is defined as :

$$G = (X_1, X_2, X_3, \ldots.., X_n)^{1/n}$$

Taking log of both the sides

$$\log G = \log\ (X_1, X_2, X_3, \ldots.., X_n)^{1/n}$$

$$= \frac{1}{n} \log\,(X_1, X_2, X_3, \ldots.., X_n) = \frac{\log X_1 + \log X_2 + \ldots.. + \log X_n}{n}$$

$$\therefore \quad G = \text{antilog} \frac{(\log X_1 + \log X_2 + \ldots.. + \log X_n)}{n}$$

In a frequency distribution,

Example : 29

Compute the geometric mean (G.M) for the following data :

Marks	0–10	10–20	20–30	30–40	40–50
No. of Students	04	08	10	06	07

Solution :

Marks	Mid-Value X	Frequency	log X	f X log X
0–10	5	4	0.6990	2.7960
10–20	15	8	1.1761	9.4088
20–30	25	10	1.3979	13.9790
30–40	35	06	1.5441	9.2646
40–50	45	07	1.6532	11.5724
		$\Sigma f = n = 35$		$\Sigma f \log X = 47.0208$

$$\log G = \frac{\Sigma f \log X}{N} = \frac{47.0208}{35} = 1.3435$$

$$G = \text{antilog}\,(1.3435) = 22.055 \text{ marks.}$$

Example : 30

Find the geometric mean of the values 10, 5, 15, 8, 12.

Solution :

Here $x_1 = 10, x_2 = 5, x_3 = 15, x_4 = 8, x_5 = 12, n = 5$

$$\therefore \text{GM of X} = \overline{X} = \sqrt[5]{10 \times 5 \times 15 \times 8 \times 12}$$

$$\overline{X} = \sqrt[5]{72000} = (72000)^{1/5} = 9.36$$

Example : 31

Find the geometric mean of the following data

X	13	15	14	16	17
f	2	13	5	7	3

Solution :

$$n = \Sigma f = 2 + 13 + 5 + 7 + 3 = 30$$

$$\text{GM of X} = \overline{X} = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot x_3^{f_3} \cdot x_3^{f_3} \cdot \ldots . x_5^{f_5}}$$

$$= \sqrt[30]{(13)^2 \cdot (15)^{13} \cdot (14)^5 \cdot (16)^7 \cdot (17)^3}$$

$$= \sqrt[30]{2.33292 \times 10^{35}}$$

$$= (2.33292 \times 10^{35})^{1/30}$$

$$\Rightarrow \overline{X} = 15.0984$$

$$\approx 15.10$$

Example : 32

MODE

The mode of a set of numbers is that number, which occurs more number of times than any other number in the set. It is the most frequently occurring value. If two or more values occur with equal or nearly equal number of times, then the distribution is said to have two or more modes. In case, there are three or more modes and the distribution or data set is said to be multimodal.

For finding median, the data is arranged in descending or ascending order, whereas for mode, this is not required.

To illustrate mode, consider the given set of observations :

60 , 29 , 85 , 102 , 29 , 85 , 23 , 85. Though the number 29 is repeated twice but the number 85 is repeated thrice. Hence the modal value is 85. So mode is that value whose frequency is maximum.

MODE FOR RAW DATA

Example : 36

The number of points scored in a series of volleyball games is listed below. What is the mode?

15 , 12 , 13 , 14 , 12 , 14 , 16 , 12 , 14 , 14 , 15

Solution :

Value 14 occurs maximum number of times i.e 4 times.

$\therefore$ Mode Z = 14.

MODE

The mode of a set of numbers is that number, which occurs more number of times than any other number in the set. It is the most frequently occurring value. If two or more values occur with equal or nearly equal number of times, then the distribution is said to have two or more modes. In case, there are three or more modes and the distribution or data set is said to be multimodal.

For finding median, the data is arranged in descending or ascending order, whereas for mode, this is not required.

To illustrate mode, consider the given set of observations :

60 , 29 , 85 , 102 , 29 , 85 , 23 , 85. Though the number 29 is repeated twice but the number 85 is repeated thrice. Hence the modal value is 85. So mode is that value whose frequency is maximum.

MODE FOR RAW DATA

Example : 36

The number of points scored in a series of volleyball games is listed below. What is the mode?

15 , 12 , 13 , 14 , 12 , 14 , 16 , 12 , 14 , 14 , 15

Solution :

Value 14 occurs maximum number of times i.e 4 times.

$\therefore$ Mode Z = 14.

14. Find H.M for the following data :

X :	2	4	8	16
f :	2	3	3	2

15. During one month the total no. of kilometers driven by each track is given below. Find harmonic mean.

Truck No.	1	2	3	4
Km Driven	40	50	60	75

16. Following is the frequency distribution of marks (out of obtained by the students of a certain college in statistics. Calculate harmonic mean.

CI	11–15	16–20	21–25	26–30	31–35
Frequencies	15	20	60	150	15

ANSWERS

(1) 716 (2) 167.13 cm (3) 58 (4) 59.35 (5) 25
(6) 43.18 (8) 146.75
(9) Median = 35.76 yrs, $Q_1 = 30.24$, $Q_3 = 39.55$
$D_7 = 38.79$, $P_{42} = 34.29$
(10) $Q_1 = 77$, $Q_2 = 78$, $Q_3 = 79$
(11) Median = 15 , $D_3 = 11$, $P_{20} = 8.6$
(12) $Q_1 = 47.14$, $Q_2 = 55.58$, $Q_3 = 63.44$, $D_7 = 61.56$, $P_{60} = 58.37$
(13) 2.9 (14) 4.44 (15) 53.33 Km (16) 24.3

MODE

The mode of a set of numbers is that number, which occurs more number of times than any other number in the set. It is the most frequently occurring value. If two or more values occur with equal or nearly equal number of times, then the distribution is said to have two or more modes. In case, there are three or more modes and the distribution or data set is said to be multimodal.

For finding median, the data is arranged in descending or ascending order, whereas for mode, this is not required.

To illustrate mode, consider the given set of observations :

60 , 29 , 85 , 102 , 29 , 85 , 23 , 85. Though the number 29 is repeated twice but the number 85 is repeated thrice. Hence the modal value is 85. So mode is that value whose frequency is maximum.

MODE FOR RAW DATA

Example : 36

The number of points scored in a series of volleyball games is listed below. What is the mode?

15 , 12 , 13 , 14 , 12 , 14 , 16 , 12 , 14 , 14 , 15

Solution :

Value 14 occurs maximum number of times i.e 4 times.

$\therefore$ Mode Z = 14.

Example : 37

The following is the number of problems assigned to a class. What is mode?

10 , 6 , 9 , 9 , 18 , 15 , 9 , 10 , 8 , 14 , 10

GROUPED DATA

Locating the mode in case of discrete data is very simple and involves the following two steps :

- Consider the frequencies and locate the highest frequency in the frequency column (f).
- The 'X' value corresponding to the highest frequency is the mode.

Example : 38

Calculate the mode for the following data :

Marks :	04	05	06	07	08	09
No. of Students :	16	15	18	10	10	07

Maximum frequency in the above example = 18

$\therefore$ Corresponding value of marks = 06

$\therefore$ Mode = Z = 06

Continuous Data

Algorithm

Step I : Obtain the continuous frequency distribution

Step II : Determine the class of maximum frequency. This class is called the modal class.

Step III : Obtain the values of the following from the frequency distribution.

l_1 = Lower limit of the modal class

l_2 = Upper limit of the modal class

F_0 = Frequency of class preceding the modal class

f_1 = frequency of the modal class

f_2 = frequency of the class succeeding the modal class.

i = length of the class interval.

Step IV : Substitute the values obtained in step III in the following formula :

$$\text{Mode (Z)} = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

Following example will illustrate the above algorithm.

Step IV : Substitute the values obtained in step III in the following formula :

$$\text{Mode (Z)} = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

Following example will illustrate the above algorithm.

Example : 39

Compute the mode for the following distribution :

Size of Items	0–4	4–8	8–12	12–16	16–20	20–24	24–28	28–32	32–36	36–40
Frequency	05	07	09	17	12	10	06	03	01	00

Solution :

Here the maximum frequency is 17 and the corresponding class is 12–16. So 12–16 is the modal class.

So, $l_1 = 12$, $l_2 = 16$, $f_0 = 9, f_1 = 17, f_2 = 12$

Substituting the value in the formula

$$Z = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

$$= 12 + \frac{17 - 9}{(17 - 9) + (17 - 12)} \times (16 - 12)$$

$$= 12 + \frac{8}{13} \times 4 = 12 + \frac{32}{13} = 14.46$$

$\therefore$ **Z = 14.46**

Example : 40

Calculate the value of mode for the following frequency distribution :

Class	1–4	5–8	9–12	13–16	17–20	21–24	25–28	29–32	33–36	37–40
Frequency	2	5	8	9	12	14	12	15	11	13

Solution :

Here the classes are not in the inclusive form. So, we first convert them in inclusive form by subtracting $h/2$ from the lower limit and adding $\frac{h}{2}$ to the upper limit of each class. Where h is the difference between the lower limit of a class and upper limit of the preceding class.

Class	Frequency	Class	Frequency
0.5–4.5	2	20.5–24.5	14
4.5–8.5	5	24.5–28.5	12
8.5–12.5	8	28.5–32.5	15
12.5–16.5	9	32.5–36.5	11
16.5–20.5	12	36.5–40.5	13

Clearly, the class 28.5–32.5 has maximum frequency 15. So 28.5 – 32.5 is the modal class.

We have, $l_1 = 28.5$, $l_2 = 32.5$, $f_0 = 12$, $f_1 = 15$, $f_2 = 11$

Substituting the value in the formula

$$Z = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

$$= 28.5 + \frac{15 - 12}{(15 - 12) + (15 - 11)} \times (32.5 - 28.5)$$

$$= 28.5 + \frac{3}{7} \times 4 = 28.5 + \frac{12}{7} = 28.5 + 1.71 = 30.21$$

$\therefore$ **Z = 30.21**

Example : 41

The following table gives the distribution of 100 families according to their expenditures. If the mode of the distribution is 24, find the missing frequency X.

Expenditure in 100 :	0–10	10–20	20–30	30–40	40–50
No. of Families :	14	X	27	21	15

Solution :

Since the modal value is 24. Which lies in class-interval 20–30, therefore 20–30 is the modal class.

Hence, $l_1 = 20,\ l_2 = 30,\ f_0 = x,\ f_1 = 27,\ f_2 = 21,\ Z = 24$

Substituting the value in the formula

$$Z = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

$$24 = 20 + \frac{27 - x}{27 - x + 27 - 21} \times (30 - 20)$$

$$\Rightarrow 24 - 20 = \frac{10\,(27 - x)}{33 - x} \Rightarrow 4\,(33 - x) = 270 - 10x$$

$$\Rightarrow 132 - 4x = 270 - 10x$$

$$\Rightarrow -4x + 10x = 270 - 132$$

$\Rightarrow \; 6x = 138 \qquad \Rightarrow X = 23$

$\therefore$ Value of missing frequency $X = 23$

1. The grades 4, 5, 6, 7, 8, 9 are given to 100 students. Find the average grade of that class.

Grade	4	5	6	7	8	9
No. of Students	10	12	20	24	30	4

Solution :

Let us denote the grades by 'X' and no. of students by f.

X	**F**	**fX**
4	10	40
5	12	60
6	20	120
7	24	168
8	30	240
9	4	36
Total	100	664

$$\overline{X} = \frac{\Sigma fX}{\Sigma f} = \frac{664}{100} = 6.64$$

$\therefore$ average grade of the class is 6.64.

2. The following table gives the yearly profit in 60 textile mills in Mumbai city.

Profits (in lacs)	10–12	12–14	14–16	16–18	18–20	20–22
No. of Mills	5	7	12	22	5	9

Calculate from the above data the average profit in the textile mills.

Solution :

Profits (in lacs)	No. of Mills (f)	Class Mark 'X'	fX
10–12	5	11	55
12–14	7	13	91
14–16	12	15	180
16–18	22	17	374
18–20	5	19	95
20–22	9	21	189
	$\Sigma f = 60$		$\Sigma fX = 984$

$$\text{Mean} = \frac{\Sigma fX}{\Sigma f} = \frac{984}{60} = 16.4$$

3. A factory employing 750 persons reported the following salary bill. Calculate the average salary per employee.

Salary	3000–3500	3500–4000	4000–4500	4500–5000	5000–5500	5500–6000
No. of Employees	30	150	200	185	125	60

Solution :

CI	f	X	$u = \frac{X - A}{i}$	fu
3000–3500	30	3250	– 3	– 90
3500–4000	150	3750	– 2	– 300
4000–4500	200	4250	– 1	– 200
4500–5000	185	4750	0	0
5000–5500	125	5250	1	125
5500–6000	60	5750	2	120
	$\Sigma f = 750$			$\Sigma fu = -345$

$$\text{Mean} = \overline{X} = A + \frac{\Sigma fu}{\Sigma f} \times i$$

Where A = Assumed mean = 4750

i = 500

$$= 4750 + \left(\frac{-345}{750} \times 500\right) = 4520$$

6. The number of students absent in a school was recorded everyday for 147 days and the raw data was presented in the form of the following frequency table.

No. of Students Absent	5	6	7	8	9	10	11	12	13	15	18	20
No. of days	1	5	11	14	16	13	10	70	4	1	1	1

Obtain the median.

Solution :

X_i	f_i	C.F
5	1	1
6	5	6
7	11	17
8	14	31
9	16	47
10	13	60
11	10	70
12	70	140
13	4	144
15	1	145
18	1	146
20	1	147
	$N = \Sigma f_i = 147$	

We have $N = 147 \Rightarrow \frac{N}{2} = \frac{147}{2} = 73.5$

The C.F just greater than $\frac{N}{2}$ is 140 and the corresponding

Value of variable 'X' is 12.

Hence, Median = 12.

8. Compute the median for the following frequency distribution. Also calculate Q_1, Q_3 and P_{72}.

CI	Less than 20	Less than 30	Less than 40	Less than 50	Less than 60	Less than 70	Less than 80	Less than 90	Less than 100
F	0	4	16	30	46	66	82	92	100

Solution :

We are given the cumulative frequency distribution. So, we first construct a frequency table from the given C.F. distribution.

CI	Frequency	C.F
20–30	4	4
30–40	12	16
40–50	14	30
50–60	16	46
60–70	20	66
70–80	16	82
80–90	10	92
90–100	8	100
	$N = \Sigma F = 100$	

Here N = 100 $\therefore \frac{N}{2} = 50$

We observe that the C.F just greater than $\frac{N}{2} = 50$ is 66 and the corresponding class is 60–70.

$\therefore$ Median class is 60–70.

$\therefore$ $l_1 = 60$, $l_2 = 70$, $F = 20$, $CF = 46$

$$\therefore \text{Median} = l_1 + \frac{\frac{N}{2} - CF}{F} \times (l_2 - l_1)$$

$$= 60 + \frac{50 - 46}{20} \times (70 - 60)$$

$$= 60 + \frac{4 \times 10}{20} = 62$$

$\therefore$ **Median = 62**

For lower Quartile Q_1, $\frac{N}{2} = \frac{100}{4} = 25$

Lower Quartile class is 40–50.

$\therefore$ $l_1 = 40$, $l_2 = 50$, $f = 14$, $CF = 16$

$$\therefore Q_1 = l_1 + \frac{\frac{N}{4} - CF}{F} \times (l_2 - l_1)$$

$$= 40 + \frac{25 - 16}{14} \times (50 - 40)$$

$$= 40 + \frac{9 \times 10}{14} = 46.43$$

$\mathbf{Q_1 = 46.43}$

For upper Quartile Q_3, $\frac{3N}{4} = 75$

$\therefore$ upper Quartile class is 70–80.

$\therefore$ $l_1 = 70, \quad l_2 = 80, \; f = 16, \quad CF = 66$

$$\therefore \quad Q_3 = l_1 + \frac{\frac{3N}{4} - CF}{F} \times (l_2 - l_1)$$

$$= 70 + \frac{75 - 66}{16} \times (80 - 70)$$

$$= 70 + \frac{9 \times 10}{16} = 75.625$$

$$\therefore \quad Q_3 = 75.625 = 75.63$$

For 72^{th} percentile, $\frac{72N}{100} = 72$

$\therefore$ P_{72} class is 70–80 $\quad \therefore F = 16 \quad CF = 66$

$$\therefore \quad P_{72} = l_1 + \frac{\frac{72N}{100} - CF}{F} \times (l_2 - l_1)$$

$$= 70 + \frac{72 - 66}{16} \times (80 - 70)$$

$$= 70 + \frac{6 \times 10}{16} = 73.75$$

$\therefore$ Median = 62

$Q_1 = 46.43$

$Q_3 = 75.63$

$P_{72} = 73.75$

$$46-40 = \frac{(115 - 42 - D)}{65} \times 10$$

$$6 \times 65 = 10\,(73 - D)$$

$$390 = 730 - 10D \Rightarrow 10D = 340$$

$$\therefore \boxed{D = 34}$$

Now, $230 = 150 + D + E$

$230 = 150 + 34 + E$

$E = 230 - 184$

$$\boxed{E = 46}$$

10. Calculate the mode for the following frequency distribution.

Class	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
Frequency	5	8	7	12	28	20	10	10

Solution :

As the class 40–50 has maximum frequency, so it is the modal class.

$\therefore f_0 = 12$, $f_1 = 28$, $f_2 = 20$, $l_1 = 40$, $l_2 = 50$

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

$$= 40 + \frac{28 - 12}{(28 - 12) + (28 - 20)} \times (50 - 40)$$

$$= 40 + \frac{2 \times 10}{3} = 40 + \frac{20}{3} = 46.67$$

Hence Mode = 46.67

$$46-40 = \frac{(115 - 42 - D)}{65} \times 10$$

$$6 \times 65 = 10\,(73 - D)$$

$$390 = 730 - 10D \Rightarrow 10D = 340$$

$$\therefore \boxed{D = 34}$$

Now,
$$230 = 150 + D + E$$
$$230 = 150 + 34 + E$$
$$E = 230 - 184$$
$$\boxed{E = 46}$$

10. Calculate the mode for the following frequency distribution.

Class	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
Frequency	5	8	7	12	28	20	10	10

Solution :

As the class 40–50 has maximum frequency, so it is the modal class.

$\therefore$ $f_0 = 12$, $f_1 = 28$, $f_2 = 20$, $l_1 = 40$, $l_2 = 50$

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

$$= 40 + \frac{28 - 12}{(28 - 12) + (28 - 20)} \times (50 - 40)$$

$$= 40 + \frac{2 \times 10}{3} = 40 + \frac{20}{3} = 46.67$$

Hence Mode = 46.67

10. Calculate the mode for the following frequency distribution.

Class	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
Frequency	5	8	7	12	28	20	10	10

Solution :

As the class 40–50 has maximum frequency, so it is the modal class.

$\therefore\ f_0 = 12\ ,\ f_1 = 28\ ,\ f_2 = 20\ ,\ l_1 = 40,\ l_2 = 50$

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times (l_2 - l_1)$$

$$= 40 + \frac{28 - 12}{(28 - 12) + (28 - 20)} \times (50 - 40)$$

$$= 40 + \frac{2 \times 10}{3} = 40 + \frac{20}{3} = 46.67$$

Hence Mode = 46.67

OBJECTIVE QUESTIONS

Mark the Correct Alternative in each of the following :

1. Which of the following is not a measure of central tendency?
   (a) Mean (b) Median
   (c) Mode (d) Standard Deviation
2. For a frequency distribution, mean, median and mode are connected by the relation.
   (a) Mode = 3 Mean – 2 Median (b) Mode = 2 Median – 3 Mean
   (c) Mode = 3 Median – 2 Mean (d) Mode = 3 Median + 2 Mean
3. Which of the following cannot be determined graphically?
   (a) Mean (b) Median
   (c) Mode (d) None of these
4. The median of a given frequency distribution is found graphically with the help of
   (a) Histogram (b) Frequency curve
   (c) Frequency polygon (d) ogive
5. Mode is
   (a) Least frequent value (b) Middle most value
   (c) Most frequent value (d) None of these
6. If the arithmetic mean of X, X + 3, X + 6, X + 9 and X + 12 is 10, then X = ______.
   (a) 1 (b) 2
   (c) 6 (d) 4
7. If the median of the data 24, 25, 26, X + 2, X + 3, 30, 31, 34 arranged in ascending order is 27.5, then X = ______.
   (a) 27 (b) 25
   (c) 28 (d) 30
8. If the mean of 6, 7, x, 8, y, 14 is 9 then
   (a) $X + y = 21$ (b) $X + y = 19$
   (c) $X - y = 19$ (d) $X - y = 21$

9. If the arithmetic mean of 7, 8, X, 11, 14 is X, then X = ______.
   (a) 9 (b) 9.5
   (c) 10 (d) 10.5
10. The arithmetic mean and mode of a data are 24 and 12 respectively then the median is ______.
   (a) 25 (b) 18
   (c) 20 (d) 22
11. In a set of 30 observations, the value 45 is repeated maximum no. of times, so the following measure can be calculated as 45 is
   (a) Arithmetic mean (b) Median
   (c) Mode (d) None of these
12. If there are extreme value present in the data, the following measure is most suitable.
   (a) Arithmetic mean (b) Geometric mean
   (c) Mode (d) None of these
13. If the values of mean and median are 34.5 and 34.1 respectively, the value of mode can be
   (a) 33.3 (b) 35.7
   (c) 40.2 (d) None of these
14. If there are two groups with 100 observations each and 35 and 45 as values of their mean. Then the value of combined mean of 200 observations will be
   (a) 35 (b) 40
   (c) 45 (d) None of these
15. The construction of cumulative frequency table is useful in determining the
   (a) Mean (b) Median
   (c) Mode (d) All the three measures
16. If the mode of the following data is 7, then the value of K in 2, 4, 6, 7, 5, 6, 10, 6, 7, 2K + 1, 9, 7, 13 is
   (a) 3 (b) 7
   (c) 4 (d) 2
17. A data has 25 observations (arranged in descending order). Which observation represents the median?
   (a) 12th (b) 13th
   (c) 14th (d) 15th
18. The algebraic sum of the deviations of a frequency distribution from its mean is
   (a) always positive (b) always negative
   (c) zero (d) a non-zero number
19. If mode of a data is 45, mean is 27, then median is
   (a) 30 (b) 27
   (c) 33 (d) None
20. The class mark of a class interval is
   (a) Lower limit + upper limit (b) Upper limit – lower limit
   (c) $\frac{1}{2}$ (lower limit + upper limit) (d) $\frac{1}{4}$ (upper limit + upper limit)

[**Ans. :** (1 – d), (2 – c), (3 – a), (4 – d), (5 – c), (6 – d), (7 – b), (8 – b), (9 – c), (10 – c), (11 – c), (12 – d), (13 – a), (14 – b), (15 – b), (16 – a), (17 – b), (18 – c), (19 – c), (20 – c)]