318en08

8

MEASURES OF CENTRAL TENDENCY

In the previous lesson, we have studied how to collect the raw data, its classification and tabulation in a useful form. Yet, this is not sufficient, for practical purposes; there is a need for further condensation of the data, particularly when we want to compare two or more different distributions data set. We may reduce the entire distribution to one number which represents the distribution using the measures of central tendency.

OBJECTIVES

After completing this lesson, you will be able to:

- explain the concept of measures of central tendency or averages;
- calculate arithmetic mean, combined arithmetic mean and weighted arithmetic mean;
- calculate median and quartiles;
- calculate mode;
- compare the various measures of central tendency; and
- apply these measures for solving various business problems.

8.1 MEANING OF CENTRAL TENDENCY

The measure of central tendency is defined as the statistical measure that identifies a single value as the representative of an entire distribution. It aims to provide an accurate description of the entire data. It is the single value that is most typical/ representative of the data. Since such typical values tend to lie centrally within a set of observations when arranged according to magnitudes, averages are called measures of central tendency. In other words, the measure of central tendency summarizes the data in a single value in such a way that this single value can represent the entire data. The word average is commonly used in day-to-day conversations. For example, we may talk of average income of an Indian, average rainfall, average production, average price etc.

8.2 TYPES OF AVERAGES (OR) MEASURES OF CENTRAL TENDENCY

The following are the important types of averages:

- Arithmetic mean.
  - Simple arithmetic mean.
  - Weighted arithmetic mean.
- Median
- Quartiles.
- Mode.

The first one is called **'mathematical average'** where as other three are called **'measures of location'** or **'measures of position'** or **'positional averages'.**

8.2.1 Arithmetic Mean:

Arithmetic mean is the most commonly used measure of central tendency. Arithmetic mean is computed by adding all the values in the data set divided by the number of observations in it.

8.2.1.1 Computation of Arithmetic Mean in case of Individual Series

The arithmetic mean in case of individual series can be computed using following methods:

- **Direct Method**
- **Assumed Mean Method**
- **Direct Method**:

If there are N observations as $X_1, X_2 X_3 ... X_N$ then the Arithmetic Mean (usually denoted by $\overline{X}$, which is read as X bar) in case of individual series using direct method is given by:

$$\overline{X} = \frac{X_1 + X_2 + X_3 + ... X_N}{N}$$

$$\overline{X} = \frac{\Sigma X}{N}$$

Where, ΣX = sum of all N number of observations and N = total number of observations.

Example 1: Calculate Arithmetic Mean from the data showing the marks obtained by 7 students of class XIth in certain examination 5,11,16,17,19,24,30.

The Arithmetic Mean of marks is given by:

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

$$= \frac{5+11+16+17+19+24+30}{7}$$

$$= \frac{122}{7} = 17.43 \text{ marks}$$

The average marks are 17.43.

- **Assumed Mean Method:**

Assumed mean method also called shortcut method is useful if the number of observations in the data is large and/or figures are in fraction. It helps to simplify the calculations. In this method a particular value is assumed as arithmetic mean on the basis of some logic or experience. The deviation from the said assumed mean from each of the observation is computed. The summation of these deviations is taken and then it is divided by the number of observations in the data. The actual arithmetic mean is calculated using the following formula:

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

Where $\bar{X}$ = Arithemetic mean, A = Assumed mean, $\Sigma d = \Sigma(X - A)$

Σd = sum of deviations, N = Number of Individual observations

Note: Any value whether existing in the data or not can be taken as the assumed mean but the final answer would be the same.

Example 2: Data of exports of a certain firms for the year 2013-2014 are mentioned in the following table:

Firms	1	2	3	4	5	6	7	8	9
Value of Exports (₹ Cr)	10	20	30	40	50	60	70	80	90

Compute average value of exports for these firms using Assumed Mean Method.

Solution:

Table 8.1: Computation of Arithmetic Mean by Assumed Mean Method

Firms	Exports(X) (in ₹Cr)	Deviation from assumed mean (A=60); d=(X-60)
1	10	-50
2	20	-40
3	30	-30
4	40	-20
5	50	-10
6	60	0
7	70	10
8	80	20
9	90	30
N = 9	ΣX = 450	Σd = –90

$$\bar{X} = A + \frac{\Sigma d}{N} = 60 + \frac{(-90)}{9} = \text{Rupees 50 crores}$$

INTEXT QUESTIONS 8.1

1. A researcher has collected the following sample individual data.

   5 12 6 8 5 6 7 5 12 4

   The mean of the data is

   (a) 5 (b) 6 (c) 7 (d) 8

2. Find the mean of the set of numbers below:

   3, 4, –1, 22, 14, 0, 9, 18, 7, 0, 1

8.2.1.2 Computation of Arithmetic Mean for Grouped data

- **Discrete Series:**

In case of discrete series where the variable X takes the values $X_1, X_2 ... X_N$ with respective frequencies $f_1, f_2 ... f_N$ the arithmetic mean can be calculated by applying:

(i) Direct Method

(ii) Assumed Mean Method.

(iii) Step Deviation Method

(i) Direct Method: According to this method the arithmetic mean is given by:

$$\bar{X} = \frac{f_1X_1 + f_2X_2 + ...f_NX_N}{f_1 + f_2 + ...f_N} = \frac{\sum fX}{\sum f}$$

where Σf = total frequency

Example 3: The following data gives the weekly wages (in ₹) **of 20 workers in a factory:**

Weekly Earnings (in ₹)	100	140	170	200	250
No. of workers	5	2	6	4	3

Calculate the average weekly earnings of the workers.

Solution:

Table 8.2: Computation of Arithmetic Mean

Weekly Wages(X)	No. of workers(f)	fX	(A=170); d=X-170	fd
100	5	500	-70	-350
140	2	280	-30	-60
170	6	1020	0	0
200	4	800	30	120
250	3	750	80	240
	Σf=20	**ΣfX=3350**		**Σfd=-50**

Arithmetic mean using direct method, the average weekly wages are:

$$\frac{\sum fX}{\sum f} = \frac{3350}{20} = \text{Rs } 167.50$$

(ii) Assumed Mean Method: Since in the discrete series frequency (f) of each item is given, we multiply each deviation (d) from the assumed mean (A) with the respective frequency (f) to get fd. According to this method, arithmetic mean is given by:

$$\bar{X} = A + \frac{\sum fd}{N}$$

where A = assumed mean, d = X – A and N = Σf

Arithmetic mean using Assumed Mean Method, the average weekly earnings are:

$$\bar{X} = A + \frac{\sum fd}{N} = 170 + \frac{-50}{20} = \text{Rs.}167.50$$

(iii) Step Deviation Method: In this case the deviations taken from assumed mean are divided by the common factor 'c' which simplifies the calculation.

Here we estimate $d' = \frac{d}{c} = \frac{X - A}{c}$ in order to reduce the size of numerical figures for easier calculation. The arithmetic mean is given by:

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

Example 4: From the following data of the marks obtained by 60 students of a class.

Marks	20	30	40	50	60	70
No of students	8	12	20	10	6	4

Calculate the arithmetic mean by

(i) Direct Method

(ii) Assumed Mean Method

(iii) Step-Deviation Method

Solution:

Table 8.3: Calculation of Arithmetic mean

Marks (X)	No. of students(f)	fX	d = (X – 40)	d′ = d/10	fd	fd′
20	8	160	–20	–2	–160	–16
30	12	360	–10	–1	–120	–12
40	20	800	0	0	0	0
50	10	500	10	1	100	10
60	6	360	20	2	120	12
70	4	280	30	3	120	12
	Σf = 60	ΣfX = 2,460			Σfd = 60	Σfd′ = 6

Notes

(i) Direct method:

$$\bar{X} = \frac{\Sigma fX}{\Sigma f} = \frac{2460}{60} = 41$$

Hence the average marks = 41.

(ii) Assumed Mean Method:

$$\bar{X} = A + \frac{\Sigma fd}{N} = 40 + \frac{60}{60} = 40 + 1 = 41$$

Hence the average marks = 41.

(iii) Step-Deviation Method

$$\bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c = 40 + \frac{6}{60} \times 10 = 41$$

Hence the average marks = 41.

Example 5: From the following data, find the missing item if the Mean wage is **₹ 115.86**

Wages in (₹):	110	112	113	117	?	125	128	130
No. of workers :	25	17	13	15	14	8	6	2

Solution: Let x be the missing item.

Table 8.4: Calculation of missing Item

Wages in ₹ (X_i)	Number of workers f_i	f_iX_i
110	25	2750
112	17	1904
113	13	1469
117	15	1755
x	14	14x
125	8	1000
128	6	768
130	2	260
Total	$\Sigma f_i = 100$	$\Sigma f_iX_i = 9906 + 14x$

Now the arithmetic mean $(\bar{X}) = \frac{\Sigma f_i X_i}{\Sigma f_i}$

Therefore, $115.86 = \frac{9906 + 14x}{100}$

or $115.86 \times 100 = 9906 + 14x$

$11586 = 9906 + 14x$

$11586 - 9906 = 14x$

$1680 = 14x$

or $x = \frac{1680}{14}$

$x = 120$

Therefore the missing item is ₹ 120

INTEXT QUESTIONS 8.2

1. Find the mean of the set of ages in the table below:

Age (years)	Frequency
10	0
11	8
12	3
13	2
14	7

2. Find the mean average weekly earnings for the data given in example 3, by using step deviation method.

8.2.1.3 Computation of Arithmetic Mean in case of Continuous series

In case of continuous series, class intervals and frequencies are given. In this case the mid-points of various class intervals are taken for calculating arithmetic mean. It may be noted that class intervals may be exclusive or inclusive or of unequal size. In case of continuous series also the computation of arithmetic mean is done by applying:

(i) Direct Method;

(ii) Assumed Mean Method;

(iii) Step Deviation Method.

(i) Direct Method: The steps involved in the calculation of arithmetic mean are as follows:

1. Calculate the mid-point of each class and denote these mid-points as m as follows:

$$\text{Mid-point } (m) = \frac{\text{Lower Limit} + \text{Upper Limit}}{2}$$

2. Multiply the mid-point with respective frequency and denote these product as fm

3. The arithmetic mean is obtained as follows:

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

(ii) Assumed Mean Method: Under this method the formula for calculating mean is

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

Where A = assumed mean d = m – A

f = frequency of n number of observation.

(iii) Step Deviation Method: To make the calculations simpler, we first find a common figure by which all the values of d can be divided. It will reduce the values of d and make further calculations easier. This common factor by which values of d are divided is termed as c i.e. $\left(\frac{d}{c} = d'\right)$. At a later stage the value of d′ is again multiplied by this common factor so that the final result of arithmetic mean is not affected. The steps involved are as follows:

Step 1: Obtain

$$d' = \frac{m - A}{c}$$

where m = mid-point, A = Assumed mean

c = common factor which is the difference between upper limit and lower limit of a class

Step 2: Apply the following formula to calculate arithmetic mean:

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

Example 6: Calculate average Land- size of the farmers of a village from the following data using (a) Direct method (b) Step deviation method.

Size of Land (in hectares)	0-10	10-20	20-30	30-40	40-50	50-60
No. of Farmers	42	44	58	35	26	15

Solution:

Table 8.5: Computation of Land Size for Exclusive Class Interval

Land Size (in hectares)	No. of Farmers(f)	Mid Points (m)	fm	$d' = \frac{m-35}{10}$	fd'
0-10	42	5	210	–3	–126
10-20	44	15	660	–2	–88
20-30	58	25	1450	–1	–58
30-40	35	35	1225	0	0
40-50	26	45	1170	1	26
50-60	15	55	825	2	30
	$\Sigma f = 220$		$\Sigma fm = 5540$		$\Sigma fd' = -216$

Applying Direct Method:

$$\bar{X} = \frac{\Sigma fm}{\Sigma f} = \frac{5540}{220} = 25.2 \text{ hectares}$$

Applying the Step Deviation Method:

$$\bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c = 35 + \left(\frac{-216}{220}\right) \times 10 = 25.2 \text{ hectares}$$

INTEXT QUESTIONS 8.3

1. The following distribution gives the pattern of overtime work per month done by 180 employees of a company. Calculate the arithmetic mean.

Notes

Overtime (in hrs)	0-10	10-30	30-40	40-50	50-60
No. of Employees	10	60	50	40	20

Solution: Since the class intervals are unequal, the frequencies have to be adjusted to make the class interval equal on the assumption that these are equally distributed throughout the class.

$$\bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c = 45 + \left(\frac{-220}{180}\right) \times 10 = 32.778 \text{ hours}$$

2. A company is planning to improve plant safety. For this accident data for the last 180 weeks were compiled. These data are grouped into the frequency distribution as shown below:

No. of accidents	1-10	11-20	21-30	31-40	41-50	51-60
No. of Weeks	10	20	30	50	40	30

Calculate the arithmetic mean of the number of accident per day.

Solution: In this case the inclusive series are converted into exclusive series by deducting half the difference between upper limit of a class and lower limit of next class from the lower limit of class and adding the same to upper limit of class.

$$\bar{X} = A + \frac{\Sigma fd'}{\Sigma f} \times c = 45.5 + \left(\frac{-180}{180}\right) \times 10 = 35.5 \text{ accident per week}$$

8.2.3 Properties of Arithmetic Mean

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

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

3. The sum of the squares of the deviations of the items taken from arithmetic mean is minimum.

4. If a constant is added or subtracted to all the variables, mean increases or decreases by that constant.

5. If all the variables are multiplied or divided by a constant, mean also gets multiplied or divided by the constant.

INTEXT QUESTIONS 8.4

Choose the correct answer:

1. The sum of deviations of the individual data elements from their mean is:
   (a) always greater than zero
   (b) always less than zero
   (c) sometimes greater than and sometimes less than zero, depending on the data elements
   (d) always equal to zero
2. In a group of 12 scores, the largest score is increased by 36 points. What effect will this have on the mean of the scores?
   (a) It will be increased by 12 points
   (b) It will remain unchanged
   (c) It will be increased by 3 points
   (d) It will increase by 36 points
   (e) There is no way of knowing exactly how many points the mean will be increased.

8.2.4 Combined Mean

If a series of N observations consists of two components having N_1 and N_2 observations ($N_1+N_2=N$), and means $\bar{X}_1$ and $\bar{X}_2$ respectively then the Combined mean $\bar{X}$ of N observations is given by

$$\text{Combined mean } \bar{X} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$$

Example 7: 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: Let X be the average marks of all 150 students taken together.

Average. marks	$\bar{X}_1 = 50;$	$\bar{X}_2 = 55;$	$\bar{X}_3 = 45$
No. of students	$N_1 = 70;$	$N_2 = 50;$	$N_3 = 30$

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

$$= \frac{70\times50+50\times55+30\times45}{70+50+30} = \frac{7600}{150}$$

$$\bar{X}_{123} = 50.67 \text{ marks}$$

INTEXT QUESTIONS 8.5

1. The mean of a certain number of observations is 40. If two or more items with values 50 and 64 are added to this data, the mean rises to 42. Find the number of items in the original data.

2. Eight coins were tossed together and the number of times they fell on the side of heads was observed. The activity was performed 256 times and the frequency obtained for different values of x, (the number of times it fell on heads) is shown in the following table. Calculate then mean by: i) Direct method ii) Short-cut method

X:	0	1	2	3	4	5	6	7	8
f:	1	9	26	59	72	52	29	7	1

3. Calculate the average age of employees working in a company from the following data:

Age (years) below:	25	30	35	40	45	50	55	60
No. of employees:	8	23	51	81	103	113	117	120

8.2.5 Weighted Arithmetic Mean:

In calculating simple arithmetic mean, it is assumed that all the items in the series carry equal importance. But in practice, there are many cases where relative importance should be given to different items. When the mean is computed by giving each data value a weight that reflects its importance, it is referred to as a weighted mean. When data values vary in importance, the analyst must choose the weight that best reflects the importance of each value. If $w_1, w_2, w_3, \ldots w_N$ are weights of N observations in a series $X_1, X_2, X_3, \ldots X_N$ then the weighted mean is calculated as

$$\bar{X}_w = \frac{\Sigma wX}{\Sigma w}$$

Note: If the weights of all the observations are equal i.e. $w_1 = w_2 = w_3, \ldots w_N = w$

Then the weighted A.M is equal to simple A.M i.e. $\bar{X}_w = \bar{X}$.

Example 8: An examination was held to decide the award of scholarship. The weights of various subjects were different. The marks obtained by 3 candidates (out of 100) in each subject are given below:

Subject	Weight	Students		
		A	**B**	**C**
Mathematics	40	60	57	62
History	30	62	61	67
Chemistry	20	55	53	60
English	10	67	77	49

Calculate the weighted A.M. to award the scholarship.

Solution:

Table 8.6: Calculation of the weighted arithmetic mean

Subject	Weight	Students					
		A		**B**		**C**	
		Marks (X_A)	$X_A w_i$	Marks (X_B)	$X_B w_i$	Marks (X_C)	$X_C w_i$
Mathematics	40	60	2400	57	2280	62	2480
History	30	62	1860	61	1830	67	2010
Chemistry	20	55	1100	53	1060	60	1200
English	10	67	670	77	770	49	490
Total	**100**	**244**	**6030**	**248**	**5940**	**238**	**6180**

Applying the formula for weighted mean, we get

$$\bar{X}_{wA} = \frac{6030}{100} = 60.3 \text{ marks}; \quad \bar{X}_A = \frac{244}{4} = 61 \text{ marks.}$$

$$\bar{X}_{wB} = \frac{5940}{100} = 59.4 \text{ marks}; \quad \bar{X}_B = \frac{248}{4} = 62 \text{ marks.}$$

$$\bar{X}_{wC} = \frac{6180}{100} = 61.8 \text{ marks}; \quad \bar{X}_C = \frac{238}{4} = 59.5 \text{ marks.}$$

From the above calculation, it may be noted that student B should get the scholarship as per simple A.M. values, but according to weighted A.M., student C should get the scholarship because all the subjects of examination are not of equal importance.

Notes

INTEXT QUESTIONS 8.6

1. A big mall wants to know the weighted mean of the sales price of 2,000 units of one product that had its final price adjusted according to the first ten days of sales. The table below summarizes the relation between final price and number of sold units.

Price per unit	No. of sold units	Price per unit	No.of sold units
₹ 24.20	354	₹ 24.14	288
₹ 24.10	258	₹ 24.06	240
₹ 24.00	209	₹ 23.95	186
₹ 23.90	133	₹ 23.84	121
₹ 23.82	110	₹ 23.75	101

Compute both the average price and the weighted average sales price of this product

An Evaluation of Arithmatic Mean

Arithmetic mean is easy to calculate**.** All values in the series are used in the calculation of mean, so it can be regarded as more representative of the entire data set. However, mean is affected by extreme items i.e. very high or a very low value in the data set. Thus the mean may be rather lower or higher than most of the values in the data set and so become unrepresentative of the entire data. Mean cannot be calculated in the open-ended frequency distribution.

8.3 MEDIAN

Median is the positional value that divides the series into two equal parts in such a way that half of the items lie above this value and the remaining half lie below this value. In Connor's words - "The median is that value of the variable which divides the group into two equal parts, one part comprising all values greater and the other all values lesser than the median." Median is called a positional average because it is based on the position of a given observation in a series arranged in an ascending or descending order and position of the median is such that equal number of items lie on either side of it. Median is denoted by Med. or M_d.

8.3.1 Computation of Median in case of Individual Series

The steps involved in the calculation of median are as follows:

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

Step 2: Ascertain $\frac{N+1}{2}$th item. This will give the position of the median i.e. item/items at which the median lies.

It may be noted that the formula $\frac{N+1}{2}$th item gives the position of the median in an ordered series, not the median itself. Median is the size of $\frac{N+1}{2}$th item.

Example 9: The following data relates to the no. of patients examined per hour in the hospital:

No. of Patients Examined	10	12	15	20	13	24	17	18

Calculate the median.

Solution:

Arranging the size of item in ascending order:

No. of Patients Examined	10	12	13	15	17	18	20	24

$$\text{Median} = \text{size of}\left(\frac{N+1}{2}\right)\text{th item} = \left(\frac{8+1}{2}\right)\text{th item}$$

$$= \text{size of 4.5th item}$$

$$\text{We get Median} = \frac{15+17}{2} = 16$$

Thus the median no. of patients examined per hour is 16.

Example 10: The following figures represent the number of books in Statistics issued at the counter of a library on 11 different days. 96, 180, 98, 75, 270, 80, 102, 100, 94, 75 and 200. Calculate the median.

Solution: Arrange the data in the ascending order as 75, 75, 80, 94, 96, 98, 100, 102,180, 200, 270.

Now the total number of items ‘N’= 11

$$\text{Therefore, the median} = \text{size of}\left(\frac{N+1}{2}\right)\text{th item}$$

$$= \text{size of } \left(\frac{11+1}{2}\right) \text{th item}$$

$$= \text{size of } 6^{th} \text{ item}$$

$$= 98 \text{ books per day}$$

INTEXT QUESTION S 8.7

1. If a data set has an even number of observations, the median
   (a) cannot be determined
   (b) is the average value of the two middle items
   (c) must be equal to the mean
   (d) is the average value of the two middle items when all items are arranged in ascending order

2. A distribution of 6 scores has a median of 21. If the highest score increases 3 points, the median will become:
   (a) 21
   (b) 21.5
   (c) 24
   (d) Cannot be determined without additional information.
   (e) none of these

8.3.2 Computation of Median in case of Discrete Series:

In case of discrete series the position of median i.e. $\frac{N+1}{2}$th item can be located through cumulative frequency. The steps involved in the calculation of median are as follows:

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

Step 2: find out the cumulative frequency (c.f.)

Step 3: Median = size of $\frac{N+1}{2}$th item

Step 4: Now look at the cumulative frequency column and find that total which is either equal to $\frac{N+1}{2}$ or next higher to that and determine the value of the variable corresponding to it. That gives the value of median.

Example 11: Calculate Median from the following data:

Marks	45	55	25	35	5	15
No of Students	40	30	30	50	10	20

Solution:

Table 8.7: Calculation of Median Marks

Marks in an Ascending order	No. of students	Cumulative frequencies
5	10	10
15	20	30
25	30	60
35	50	110
45	40	150
55	30	180

Median = size of $\frac{N+1}{2}$ th item = $\frac{180+1}{2}$ item = 90.5th item

Cumulative frequency which includes 90.5^{th} item =110

Median = size of item corresponding to 110 = 35marks.

8.3.3 Computation of Median in case of Continuous Series

The steps involved in the calculation of median are as follows:

Step 1: Calculate Cumulative Frequencies.

Step 2: Ascertain $\left[\frac{N}{2}\right]$ th item.

Step 3: Find out the cumulative frequency which includes $\left[\frac{N}{2}\right]$ th item and corresponding class frequency. The corresponding class of this cumulative frequency is called the median class.

Step 4: Calculate Median as Follows:

$$\text{median} = l_1 + \frac{\frac{N}{2} - \text{c.f.}}{\text{f}} \times i$$

where l_1 = Lower limit of the median class

c.f. = cumulative frequency of the preceeding class

f = frequency of the median class

i = class interval of median class

Example 12: Calculate the median of weekly expenditure from the following data:

Weekly Expenditure (in ₹)	0-10	10-20	20-30	30-40	40-50
No. of families	14	23	27	21	15

Solution:

Table 8.8: Calculation of Median Weekly Expenditure (in ₹)

Weekly Expenditure (in ₹)	**No. of families (f)**	**Cumulative frequency (c.f)**
0-10	14	14
10-20	23	37
20-30	27	64
30-40	21	85
40-50	15	100

Ascertain $\left[\frac{N}{2}\right]$ th item $= \left[\frac{100}{2}\right]$ th item = 50th item lies in class interval as 20-30.

Thus the Median Class is 20-30.

Now applying formula of median

$$\text{Median} = l_1 + \frac{\frac{N}{2} - \text{c.f}}{\text{f}} \times \text{i}$$

here $l_1 = 20$, c.f. = 37, f = 27, i = 10

$$\text{Median} = 20 + \frac{50-37}{27} \times 10$$

= ₹ 24.815

Note that while calculating the median of a series, it must be put in the 'exclusive class-interval' form. If the original series is in inclusive type, first convert it into the exclusive type and then find its median.

INTEXT QUESTIONS 8.8

1. Calculate the median age of the persons from the following data.

Age (years) :	20-25	25-30	30-35	35-40	40-45
No. of person :	70	80	180	150	20

2. Calculate the median marks of the students:

Marks :	40-50	30-40	20-30	10-20	0-10
No. of students :	10	12	40	30	8

8.3.4 Important Mathematical property of median

The sum of the deviations of the items from median, ignoring signs is the least.

$$\Sigma|X - Md| \text{ is least.}$$

An Evaluation of Median

Since Median is the middle term it is not affected by extreme values and can also be calculated in the open ended frequency distribution. It is not based on all the values of the data set.

8.4 QUARTILES

Quartile is that value which divides the total distribution into four equal parts. So there are three quartiles, *i.e.* Q_1, Q_2 and Q_3 which are termed as first quartile, second quartile and third quartile or lower quartile, middle quartile and upper quartile respectively. Q_1 (quartile one) covers the first 25% items of the series. Q_1 divides the series in such a way that 25% of the observations have the value less than Q_1 and 75% have the value more than Q_1. Q_2 (quartile two) is the median or middle value of the series and Q_3 (quartile three) covers 75% items of the series. Q_3 divides the series in such a way that 75% of the observations have the value less than Q_3 and 25% have the value more than Q_3.

- **Calculation of Quartiles:**

The calculation of quartiles is done exactly in the same manner as it is in case of the calculation of median.

8.4.1 In case of Individual and Discrete Series

$$Q_k = \text{Size of } \frac{k(N+1)}{4} \text{ th item of the series}$$

8.4.2 In case of Continuous Series

$$Q_k = \text{Size of } k\left(\frac{N}{4}\right) \text{th item of the series},$$

Q_k is calculated as follows:

$$Q_k = l_1 + \frac{k\left(\frac{N}{4}\right) - cf}{f} \times i$$

Where, l_1 = Lower limit of quartile class

l_2 = upper limit of quartile class

c = Cumulative frequency preceding the quartile class

f = Frequency of kth quartile class.

Example 13: Find the Q_1 and Q_3 of the following:

(a) 4, 5, 6, 7, 8, 9, 12, 13, 15, 10, 20

(b) 100, 500, 1000, 800, 600, 400, 7000 and 1200

Solution:

(a) Values of the variable are in ascending order:

i.e. 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 20, So N = 11 (No. of Values)

$$Q_1 = \text{Size of } \frac{(N+1)}{4} \text{ th item of the series}$$

$$= \left(\frac{11+1}{4}\right) = \text{size of } 3^{rd} \text{ item.} = \text{size of 3rd item} = 6$$

$$Q_3 = \text{Size of } \frac{3(N+1)}{4} \text{th item of the series}$$

$$= 3\left(\frac{11+1}{4}\right) = \text{size of } 9^{th} \text{ item.} = \text{size of 9th item} = 13$$

$\therefore$ Required Q_1 and Q_3 are 6 and 13 respectively,

(b) The values of the variable in ascending order are:

100, 400, 500, 600, 700, 800, 1000, 1200, N = 8

$$Q_1 = \text{Size of } \frac{(N+1)}{4} \text{th item of the series}$$

$$= \text{Size of } \frac{(8+1)}{4} \text{th item of the series}$$

= size of 2.25th item

= size of {Second item + 0.25(Third item - Second item)}

= 400 + 0.25 (500 - 400) = 400 + 25 = 425

$$Q_3 = \text{Size of } \frac{3(N+1)}{4} \text{th item of the series}$$

$$= \text{Size of } \frac{3(8+1)}{4} \text{th item of the series}$$

= size of 6.75th item

= size of [6th item + 0.75(7th item - 6th item)]

= 800 + 0.75 (1000 - 800)

= 800 + 150 = 950

Required Q_1 and Q_3 are 425 and 950 respectively.

Example 14: Find the median and Q_1 from the following data.

Marks :	0-10	10-30	30-50	50-80	80-90	90-100
No of Students :	4	12	20	8	4	2

Solution: To locate median class firstly we have to calculate cumulative frequencies.

Table 8.9: Calculation of Median and Quartile Marks

Marks	0-10	10-30	30-50	50-80	80-90	90-100
No of Students	4	12	20	8	4	2
Cumulative frequency	4	16	36	44	48	50

Calculation of Median is shown as under:

Here N = 50, so N/2 = 25, hence median class is 30-50

$$\text{Median} = l_1 + \frac{\frac{N}{2} - c.f}{f} \times i$$

$$\text{Median} = 30 + \frac{25 - 16}{20} \times 20 = 39 \text{ marks}$$

Calculation of Q_1 :

Here N = 50 so N/4 = 12.5, hence Quartile class (Q_1) is 10-30

$$Q_1 = l_1 + \frac{N/4 - cf}{f} \times i$$

$$Q_1 = 10 + \frac{12.5 - 4}{12} \times 20 = 24.16 \text{ marks}$$

8.5 MODE

Mode (M_0) is the value around which maximum concentration of items occurs. For example, a manufacturer would like to know the size of shoes that has maximum demand or style of the shirt that is more frequently demanded. Here, *Mode* is the most appropriate measure. Mode is the value which is repeated the highest number of times in the series. It is the size of that item which possesses the maximum frequency.

8.5.1 Computation of Mode in case of Ungrouped Data/ Individual series

The mode of this series can be obtained by mere inspection. The number which occurs most often is the mode.

Note that if in any series, two or more numbers have the maximum frequency, then the mode will be difficult to calculate. Such series are called as Bi-modal, Tri-modal or Multi-modal series.

Example 15: Find the mode of 15, 21, 26, 25, 21, 23, 28, 21

Solution: The mode is 21 since it occurs three times and the other values occur only once.

INTEXT QUESTIONS 8.9

1. The most frequently occurring value of a data set is called:

   (a) range (b) mode (c) mean (d) median

2. Find the mode of 12, 15, 18, 26, 15, 9, 12, 27

3. The measure of location which is the most likely to be influenced by extreme values in the data set is the:

   (a) median (b) mode (c) mean (d) Quartile

4. A researcher has collected the following sample individual data

5	12	6	8	5	6	7	5	12	4

   The Median is:

   (a) 5 (b) 6 (c) 7 (d) 8

   And the Mode is:

   (a) 5 (b) 6 (c) 7 (d) 8

5. Which of the following can have more than one value?

   (a) Median; (b) Quartile; (c) mode and (d) mean

8.5.2 Computation of mode in case of discrete series

The mode in case of discrete series is calculated by applying the following methods:

(a) Simple inspection method:

By simple inspection, the modal value is the value of the variable against which the frequency is the largest.

Example 16: Find the modal age of boys studying in XII class from the following data.

Age : (in yrs)	5	7	10	12	15	18
No. of Boys :	4	6	9	7	5	3

Solution:

From the above data we can clearly see that modal age is 10 yrs because 10 has occurred maximum number of times i.e. 9.

b) Grouping and Analysis Table method: This method is generally used when the difference between the maximum frequency and the frequency preceding it or succeeding it is very small.

Process of Computation:

In order to find mode, a grouping table and an analysis table are to be prepared in the following manner:

Grouping Table:

A grouping table consists of 6 columns.

1. Arrange the values in ascending order and write down their corresponding frequencies in the column-1.
2. In column-2 the frequencies are grouped into two's and added.
3. In column-3 the frequencies are grouped into two's, leaving the first frequency and added.
4. In column-4 the frequencies are grouped into three's, and added.
5. In column-5 the frequencies are grouped into three's, leaving the first frequency and added.
6. In column-6 the frequencies are grouped into three's, leaving the first and second frequencies and added.
7. Now in each these columns mark the highest total with a circle.

Analysis Table:

After preparing a grouping table, prepare an analysis table. While preparing this table take the column numbers as rows and the values of the variable as columns. Now for each column number see the highest total in the grouping table (Which is marked with a circle) and mark the corresponding values of the variable to which the frequencies are related by using bars in the relevant boxes. Now the value of the variable (class) which gets highest number of bars is the modal value (modal class).

Applying grouping and Analysis Table Method to the given example for calculating the value of mode for discrete series.

Table 8.10: Grouping Table

Age	Frequency					
	Col I	Col II	Col III	Col IV	Col V	Col VI
5	4	10		**19**		
7	6		**15**		**22**	
10	**9**	**16**				**21**
12	7		12	15		
15	5	8				
18	3					

Analysis Table

Column	5	7	10	12	15	18
Col I			1			
Col II			1	1		
Col III		1	1			
Col IV	1	1	1			
Col V		1	1	1		
Col VI			1	1	1	
Total	1	3	**6**	3	1	0

Thus the Modal age of Boys is 10 years.

8.5.3 Computation of Mode in case of Continuous series

In case of continuous series, for calculating mode, first of all ensure that the given continuous series is the exclusive series with equal class intervals. In order to find out the mode we need one step more than those used for discrete series. As explained in the discrete series, modal class is determined by inspection or by preparing grouping and analysis tables. The steps involved are:

1. Determine the modal class which as the maximum frequency.
2. Value of the mode can be calculated by the formula :

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$$

l_1 = lower limit of the modal class

f_1 = frequency of the modal class

f_0 = frequency of the preceeding the modal class

f_2 = frequency of the succeeding the modal class

i = class interval of the modal class

Note: 1.It may be noted that in case of continuous series, class intervals should be equal and series should be exclusive to calculate the mode. If the given series is inclusive and has unequal class interval then the same has to be converted into exclusive series and series with equal class interval.

2. If mid points are given, class intervals are to be obtained.

Example 17: From the following data calculate mode:

Age (in years) :	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No. of Persons :	50	70	80	150	180	120	70	50

Notes

Solution:

Table 8.11: Computation of mode
Grouping Table:

Age	Frequency					
	(1)	(2)	(3)	(4)	(5)	(6)
20-25	50	120		200		
25-30	70		150		300	
30-35	80	230				**410**
35-40	150		**330**	**450**		
40-45	**180**	**300**			**370**	
45-50	120		190			240
50-55	70	120				
55-60	50					

Analysis Table

Column	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
1					1			
2					1	1		
3				1	1			
4				1	1	1		
5					1	1	1	
6			1	1	1			
Total	0	0	1	3	6	3	1	0

The modal class is 40-45. Mode is given by:

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$$

Here L = 40, f_1 = 180, f_0 = 150, f_2 = 120, i = 5

$$\text{Mode} = 40 + \frac{180 - 150}{(2 \times 180) - 150 - 120} \times 5 = 40 + \left[\frac{30}{90}\right] \times 5 = 41.67 \text{ years.}$$

Example 18: Calculate the modal wages from the following data:

Daily wages (in ₹):	20-25	25-30	30-35	35-40	40-45	45-50
No. of workers:	1	3	8	12	7	5

Solution: Here the maximum frequency is 12 by inspection method, corresponding to the class interval (35 - 40) which is the modal class.

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$$

Here L = 35, f_1 = 12, f_0 = 8, f_2 = 7, i = 5

$$\text{Mode} = 35 + \frac{12-8}{(2\times12)-8-7} \times 5 = 35 + \left[\frac{4}{9}\right] \times 5 = 37.22 \text{ (in rupe}$$

Modal wages is ₹ 37.22

An Evaluation of Mode

Mode is not affected by extreme values and can be calculated in the open ended frequency distribution.

Example 19: The following table shows the daily wages of a random sample of construction workers. Calculate its mean, median and mode.

Daily Wages (₹)	Number of Workers
200 - 399	5
400 - 599	15
600 - 799	25
800 - 999	30
1000 - 1199	18
1200 - 1399	7
Total	100

Solution:

Table 8.12: Calculation of Mean

Daily Wages (₹)	Number of Workers (f)	Class Mark m	fm
200 - 399	5	299.5	1,497.5
400 - 599	15	499.5	7,492.5
600 - 799	25	699.5	17,489.5
800 - 999	30	899.5	26,985.0
1000 - 1199	18	1,099.5	19,791.0
1200 - 1399	7	1,299.5	9,096.5
Total	100		82,352.0

Notes

$$\textbf{Mean } (\bar{X}) = \frac{\Sigma fm}{\Sigma f} = \frac{82,352.0}{100} = 823.52 \text{ (in rupees)}$$

Thus Mean Wages are ₹ 823.52

In order to calculate the mode and median the given series has to be converted from inclusive series into exclusive series.

Table 8.13: Calculation of Median

Daily Wages (₹)	Number of Workers(f)	Cumulative Frequency
199.5 – 399.5	5	5
399.5 – 599.5	15	20
599.5 – 799.5	25	45
799.5 – 999.5	30	75
999.5 – 1199.5	18	93
1199.5 – 1399.5	7	100
Total	100	

$$\text{Median} = l_1 + \frac{\frac{N}{2} - \text{c.f}}{f} \times i$$

Here N = 100 so N/2 = 50, hence median class is 799.5 – 999.5

$$\text{Median} = 799.5 + \frac{50-45}{30} \times 200 = 832.83 \text{ (in Rupees)}$$

So the median daily wage is ₹ 832.8

Table 8.14: Computation of Mode

Grouping Table

Daily Wages (₹)	Frequency					
	(1)	(2)	(3)	(4)	(5)	(6)
199.5-399.5	5	20		45		
399.5-599.5	15		40		**70**	
599.5-799.5	25	**55**				**73**
799.5-999.5	**30**		**48**	**55**		
999.5-1199.5	18	25				
1199.5-1399.5	7					

Analysis Table

Col.	199.5- 399.5	399.5- 599.5	599.5-799.5	799.5-999.5	999.5-1199.5	1199.5-1399.5
1				1		
2			1	1		
3				1	1	
4				1	1	1
5		1	1	1		
6			1	1	1	
Total	0	1	3	6	3	1

The modal class is $799.5 - 999.5$. Mode is given by:

$$\text{Mode} = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$$

Here $l_1 = 799.5$, $f_1 = 30$, $f_0 = 25$, $f_2 = 18$, $i = 200$

$$\text{Mode} = 799.5 + \frac{30 - 25}{(2 \times 30) - 25 - 18} \times 200$$

$$= 799.5 + \left[\frac{5}{17}\right] \times 200 = 858.32 \text{ (in Rupees).}$$

Thus the modal wage is ₹858.32

WHAT YOU HAVE LEARNT

- The measure of central tendency identifies the single value that is most typical/ representative of the entire data-set.
- The following are the important measures of central tendency:
  - Arithmetic mean.
    - Simple arithmetic mean.
    - Weighted arithmetic mean.
  - Median
  - Quartiles.
  - Mode.

- The arithmetic mean in case of individual series can be computed using
  - **Direct Method**

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

  - **Assumed Mean Method**

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

where $\bar{X}$ = Arithemetic mean, A = Assumed mean

Σd = sum of deviations, N = Number of Individual observations

- Arithmetic Mean in case of discrete series is given by:
  - **Direct Method**

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

where Σf = total frequency

  - **Assumed Mean Method**

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

where A = assumed mean, d = X – A and $N = \Sigma f$

  - **Step Deviation Method**

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

- Arithmetic Mean in case of continuous series is given by:
  - **Direct Method**

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

where $\text{Mid-point } (m) = \frac{\text{Lower Limit} + \text{Upper Limit}}{2}$

  - **Assumed Mean Method**

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

where A = assumed mean d = m – A

f = frequency of n number of observation.

Notes

- **Step Deviation Method**

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

where $d' = \frac{m - A}{c}$

m = mid-point, A = Assumed mean

c = common factor which is the difference between upper limit and lower limit of a class

- The combined mean of two series is given by

Combined mean $\bar{X} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$

where N_1 and N_2 are no. of items in the two series and $\bar{X}_1$ and $\bar{X}_2$ are the means of two series.

- The weighted arithmetic mean is given by:

$$\frac{\Sigma wX}{\Sigma w}$$

where $w_i = w_1, w_2, w_3, \ldots w_N$ are weights of N observations in a series and $x_i = x_1, x_2, x_3, \ldots x_N$ are N observations in the series.

- Median is the positional value that divides the series into two equal parts in such a way that half of the items lie above this value and the remaining half lie below this value.
- In individual and discrete series, the formula to calculate median is :

Median = size of $\frac{N+1}{2}$ th item

- If the number of observations is even then median is given by

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

- Median in case of cumulative series is given by:

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

$$\text{Median} = l_1 + \frac{\frac{N}{2} - \text{c.f}}{f} \times i$$

where l_1 = Lower limit of the median class

c.f. = cummulative frequency of the preceeding class

f = frequency of the median class

i = class interval of median class

- Quartile is the value which divides the total distribution into four equal parts. There are three quartiles, i.e. Q_1, Q_2 and Q_3 which are termed as first quartile, second quartile and third quartile or lower quartile, middle quartile and upper quartile respectively.

- In case of Individual and Discrete Series, the quartiles are computed by:

$$Q_k = \text{Size of } \frac{k(N+1)}{4} \text{ th item of the series}$$

- In case of Continuous Series the quartiles are computed by:

$$Q = l_1 + \frac{k\left(\frac{N}{4}\right) - cf}{f} \times i$$

where l_1 = Lower limit of ith quartile class

l_2 = upper limit of ith quartile class

c = Cumulative frequency preceding the ith quartile class

f = Frequency of kth quartile class.

- Mode is the value around which maximum concentration of items occurs.

- Mode in case of ungrouped or Individual series is the number which occurs most often in data.

- The mode in case of grouped data (discrete series and continuous series) is the value of the variable against which the frequency is the largest.

TERMINAL EXERCISE

Mean

1. An average daily wages of all 90 workers in a factory is ₹. 60. An average daily wages of non-technical workers is ₹. 45. Calculate an average daily wages of technical workers if one-third workers are technical.

2. For the two frequency distribution given below, the mean calculated from the first was 25.4 and that from the second was 32.5. Find the values of x and y.

Class Interval	Distribution I	Distribution II
10-20	20	4
20-30	15	8
30-40	10	4
40-50	x	2x
50-60	y	y

3. The mean of 99 items is 55. The value of 100^{th} item is 99 more than the mean of 100 items. What is the value of 100^{th} item

Median

4. The length of time taken by each of 18 workers to complete a specific job was observed to be the following:

Time(in min)	5-9	10-14	15-19	20-24	25-29
No. of workers	3	8	4	2	1

Calculate the median time and Q_1 and Q_3

5. Calculate the median from the following data:

Mid values	115	125	135	145	155	165	175	185	195
frequency	6	25	48	72	116	60	38	22	3

6. If the quartiles for the following distribution are Q_1= 23.125 and Q_3=43.5, find the median:

Daily Wages	0-10	10-20	20-30	30-40	40-50	50-60
No. of workers	5	-	20	30	-	10

7. The mean and median of a group of 25 observations are 143,144, and 147 respectively. A set of 6 observations is added to this data with values 132, 125, 130, 160, 165 and 157. Find mean and median for the combined group of 31 observations.

Mode

8. Locate mode in the data:

   7, 12, 8, 5, 9, 6, 10, 9, 4, 9, 9

9. Determine the modal value in the following series:

Value	10	12	14	16	18	20	22	24	26	28	30	32
frequency	7	15	21	38	34	34	11	19	10	38	5	2

10. The median and mode of the following wage distribution are known to be ₹. 33.5 and ₹. 34 respectively. Three frequency values from the table are however missing. Find the missing values.

Wages in ₹	Frequencies
0-10	10
10-20	10
20-30	?
30-40	?
40-50	?
50-60	6
60-70	4
	230

11. The details of monthly salary of various categories of employees working in a university are given below. From these details , calculate mode of monthly salary.

Category	Monthly Salary (₹)	No. of employees
Principal	10,00,000	1
Vice Principal	2,50,000	1
Senior	75,000	5
Professor Professor	30,000	8
Associate Professor	20,000	13
Assistant Professor	18,000	9

12. The distribution of age of patients turned out in a hospital on a particular day was as under:

Age (in years)	No. of patients
More than 10	148
More than 20	124
More than 30	109
More than 40	71
More than 50	30
More than 60	16
More than 70 and upto 80	1

Find the median age and modal age of the patients.

ANSWERS TO INTEXT QUESTIONS

8.1

1. (c) 2. 7

8.2

1. 12.4 years 2. 41

8.4

1. (d) 2. (c)

8.5

1. 15 2. 3.97 3. 36.83 years

8.6

1. Average price ₹23.98

   Weighted mean for sales in the first ten days ₹24.03

8.7

1. (d)
2. (a) Median will remain the same

8.8

1. 32.78 years
2. 23 marks

8.9

1. (b)
2. The modes are 12 and 15 since both occur twice
3. (c)
4. (b)
5. (c)