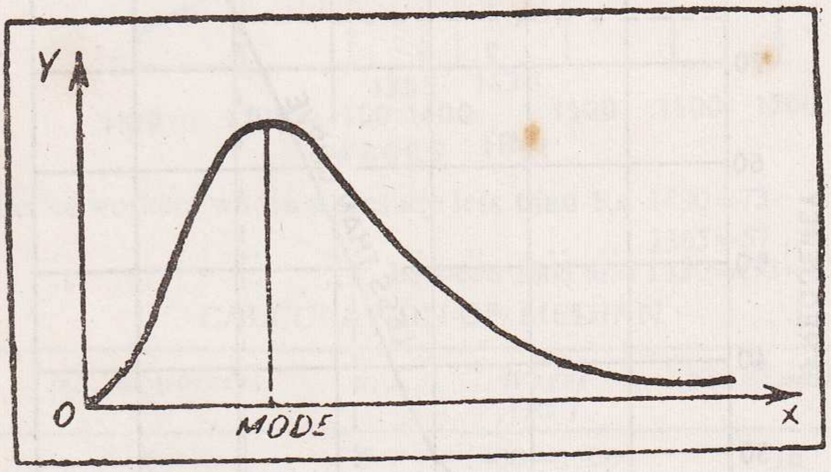
7·6 C. MODE*

The mode or the modal value is that value in a series of observations which occurs with the greatest frequency. For example, the mode of the series 3, 5, 8, 5, 4, 5, 9, 3 would be 5, since this value occurs more frequently than any of the others.

The mode is often said to be that value which occurs most often in the data, that is, with the highest frequency. While this statement is quite helpful in interpreting the mode, it cannot safely be applied to any distribution, because of the vagaries of sampling. Even fairly large samples drawn from a statistical population with a single well defined mode may exhibit very erratic fluctuations in this average if the mode is defined as that exact value in the ungrouped data of each sample which occurs most frequently. Rather it should be thought as the value about which the items are most closely concentrated. It is the value which has the greatest frequency density *in its immediate neighbourhood.*† For this reason mode is also called the most typical or fashionable value of a distribution.

The following diagram shows the modal value :



The value of the variable at which the curve reaches a maximum is called the mode. It is the value around which the items tend to be most heavily concentrated.

Although mode is that value which occurs most frequently, yet it does not follow that its frequency represents a majority out of all the total number of frequencies. For example, in the election of college president the votes obtained by three candidates contesting for presidentship out of a total of 816 votes polled are as follows :

Mr. X 268 ; Mr. Y 278 ; Mr. Z 270 : Total 816.

---

*"The mode of a distribution is the value at the point around which the items tend to be most heavily concentrated. It may be regarded as the most typical of a series of values." *Croxton and Cowden*

†Alva M. Tuttle : *Elementary Business and Economic Statistics*.

A set of data may have a single mode in which case it is said to be *unimodal*, it may have two modes which makes it *bimodal* or it may have several modes and be called multimodal.

Mr. *Y* will be elected as president because he has obtained highest votes. But it will be wrong to say that he represents majority because there are more votes against him (268+270=538) than those for him.

There are many situations in which arithmetic mean and median fail to reveal the true characteristic of data. For example, when we talk of most common wage, most common income, most common height, most common size of shoe or ready-made garments we have in mind mode and not the arithmetic mean or median discussed earlier. The mean does not always provide an accurate reflection of the data due to the presence of extreme items. Median may also prove to be quite unrepresentative of the data owing to an uneven distribution of the series. For example, the values in the lower half of a distribution range from, say, Rs. 10 to Rs. 100, while the same number of items in the upper half of the series range from Rs. 100 to Rs. 6,000 with most of them near the higher limit. In such a distribution the median value of Rs. 100 will provide little indication of the true nature of the data.

Both these shortcomings may be overcome by the use of mode which refers to the value which occurs most frequently in a distribution. Moreover, mode is the easiest to compute since it is the value corresponding to the highest frequency. For example, if the data are :

Size of shoes	5	6	7	8	9	10	11
No. of persons	10	20	25	40	22	15	6

the modal size is '8' since it appears maximum number of times in the series.

Calculation of Mode

Determining the precise value of the mode of a frequency distribution is by no means an elementary calculation. Essentially it involves fitting mathematically some appropriate type of frequency curve to the grouped data and determination of the value on the *X*-axis below the peak of the curve. However, there are several elementary methods of *estimating* the mode. These methods have been discussed for individual observations, discrete series and continuous series.

Calculation of Mode—Individual Observations

For determining mode count the number of times the various values repeat themselves and the value occurring maximum number of times is the modal value. The more often the modal value appears relatively, the more valuable the measure is an average to represent data.

Illustration 20. Calculate the mode from the following data of the marks obtained by 10 students :

Sl. No.	*Marks obtained*	*Sl. N.*	*Marks obtained*
1	10	6	27
2	27	7	20
3	24	8	18
4	12	9	15
5	27	10	30

(B. Com., Mysore Univ , 1986)

Solution. CALCULATION OF MODE

Size of item	*Number of times it occurs*	*Size of item*	*Number of times it occurs*
10	1	20	1
12	1	24	1
15	1	27	3
18	1	30	1
		Total	10

Since the item 27 occurs the maximum number of times, *i.e.*, 3, hence the modal marks are 27.

Note. Thus the process of determining mode in case of individual observations essentially involves grouping of data.

When there are two or more values having the same maximum frequency, one cannot say which is the modal value and hence mode is said to be ill-defined. Such a series is also known as bi-modal or multi-modal. For example, observe the following data :

Income (in Rs.) 110, 120, 130, 120, 110, 140, 130, 120, 130, 140

Size of item	110	120	130	140
No. of times it occurs	2	3	3	2

Since 120 and 130 have the same maximum frequency, *i.e.*, 3, mode is ill defined in this case.

Calculation of Mode—Discrete Series

In discrete series quite often mode can be determined just by inspection, *i.e.*, by looking to that value of the variable around which the items are most heavily concentrated. For example, observe the following data :

Size of garment	28	29	30	31	32	33
No. of persons wearing :	10	20	40	65	50	15

From the above data we can clearly say that the modal size is 31 because the value 31 has occurred the maximum number of times, *i.e.*, 65. However, where the mode is determined just by inspection, an error of judgment is possible in those cases where the difference between the maximum frequency and the frequency preceding it or succeeding it is very small and the items are heavily concentrated on either side. In such cases it is desirable to prepare a grouping table and an analysis table. These tables help us in ascertaining the modal class.

A grouping table has six columns. In column 1 the maximum frequency is marked or put in a circle ; in column 2 frequencies are grouped in two's ; in column 3 leave the first frequency and then group the remaining in two's ; in column 4 group the frequencies in three's ; in column 5 leave the first frequency and group the remaining in three's ; and in column 6 leave the first two frequencies and then group the remaining in three's. In each of these cases take the maximum total and mark it in a circle or by bold type.

After preparing the grouping table, prepare an analysis table. While preparing this table put column number on the left-hand side and the various probable values of mode on the right-hand side. The values against which frequencies are the highest are marked in the grouping table and then entered by means of a bar in the relevant 'box' corresponding to the values they represe n

The procedure of preparing grouping table and analysis table shall be clear from the following example :

Illustration 21. Calculate the value of mode for the following data :

Marks :	10	15	20	25	30	35	40
Numbers :	8	12	36	25	28	18	9

(*B. Com., Kerala Univ., 1988*)

Solution. Since it is difficult to say by inspection as to which is the modal value, we prepare grouping and analysis tables.

CALCULATION OF MODE

x	*f* I	II	III	IV	V	VI
10	8	20		56		
15	12		48		83	
20	36	71				99
25	35		63	81		
30	28	46			55	
35	18		27			
40	9					

PROBABLE MODAL VALUE

Col. No.	20	25	30
I	1		
II	1	1	1
III		1	1
IV		1	
V	1	1	
VI	1	1	1
	4	5	3

Corresponding to the maximum total 5, the value of the variable is 25. Hence modal value is 25.

Calculation of Mode—Continuous Series

Steps. (*i*) By preparing grouping table and analysis table or by inspection ascertain the modal class.

(*ii*) Determine the value of mode by applying the following formula :

$$M_0 = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i \qquad \ldots(i)$$

where, L=lower limit of the modal class ; Δ_1=the difference between the frequency of the modal class and the frequency of the pre-modal class, *i.e.*, preceding class (ignoring signs) ; Δ_2=the difference between the frequency of the modal class and the frequency of the post-modal class, *i.e.*, succeeding class (ignoring signs) ; i=the class interval of the modal class.

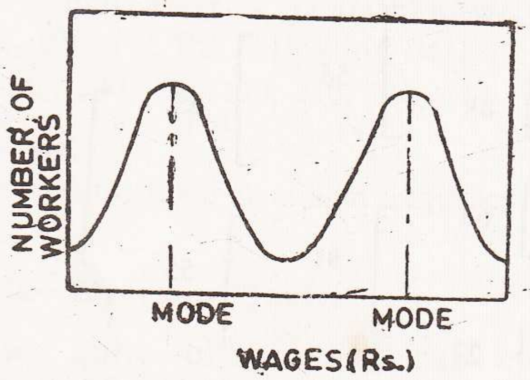
Another form of this formula is :

$$M_0 = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i \qquad \ldots(ii)$$

where, L=lower limit of the modal class ; f_1=frequency of the modal class ; f_0=frequency of the class preceding the modal class ; f_2=frequency of the class succeeding the modal class.

While applying the above formula for calculating mode, it is necessary to see that the class intervals are *uniform* throughout. If they are unequal they should first be made equal on the assumption that the frequencies are equally distributed throughout the class, otherwise we will get misleading results.

There may be two values which occur with equal frequency. The distribution is then called bimodal. The following is a graph of bimodal distribution :



In a bimodal distribution the value of mode cannot be determined with the help of formula given above. If plotted data produce a bimodal distribution, the data themselves should be questioned. Quite often such a condition is caused when the size of the sample is small ; the difficulty can be remedied by increasing the sample size. Another common cause is the use of non-homogeneous data. In instances where a distribution is bimodal and nothing can be done to change it, the mode should not be used as a measure of central tendency.

Where mode is ill-defined, its value may be ascertained by the following formula based upon the relationship between mean, median and mode :

$$\text{Mode} = 3\ \text{Median} - 2\ \text{Mean} \qquad \ldots(iii)$$

This measure is called the *empirical mode.*

Illustration 22. (*a*) Calculate mode from the following data :

Marks	*No. of students*	*Marks*	*No of students*
Above 0	80	Above 60	28
,, 10	77	,, 70	16
,, 20	72	,, 80	10
,, 30	65	,, 90	8
,, 40	55	,, 100	0
,, 50	43		

(*B. Com.. Andhra Univ. 1983 ; Nagarjun Univ., 1986, Kerala Univ , 1986*)

Solution. Since this is cumulative frequency distribution, we first convert it into a simple frequency distribution.

Marks	No. of students	Marks	No. of students
0—10	3	50—60	15
10—20	5	60—70	12
20—30	7	70—80	6
30—40	10	80—90	2
40—50	12	90—100	8

By inspection the modal class is 50—60.

$$M_0=L+\frac{\Delta_1}{\Delta_1+\Delta_2}\times i$$

$$L=50\;;\;\Delta_1=(15-12)=3,\;\Delta_2=(15-12)=3\;;\;i=10$$

$$M_0=50+\frac{3}{3+3}\times 10=50+5=55.$$

Illustration 22. (*b*) Find the value of mode from the data given below :

Weight (kg)	*No. of students*	*Weight (kg)*	*No. of students*
93—97	2	113—117	14
98—102	5	118—122	6
103—107	12	123—127	3
108—112	17	128—132	1

(*B.A. Hons. Econ., Delhi Univ., 1982*)

Solution. By inspection mode lies in the class 108—112. But the real limits of this class are 107·5—112·5.

$$\text{Mode}=L+\frac{\Delta_1}{\Delta_1+\Delta_2}\times i$$

$$L=107{\cdot}5,\;\Delta_1=f_1-f_0=(17-12)=5,\;\Delta_2=f_1-f_2=(17-14)=3,\;i=5$$

$$\text{Mode}=107{\cdot}5+\frac{5}{5+3}\times 5=107{\cdot}5+3{\cdot}125=110{\cdot}625$$

Thus the modal weight is 110·625 kg.

Illustration 23. From the following data of weight of 122 persons determine the modal weight :

Weight (in lbs.)	*No. of persons*	*Weight (in lbs.)*	*No. of persons*
100—110	4	140—150	33
110—120	6	150—160	17
120—130	20	160—170	8
130—140	32	170—180	2

(*B. Com., Sri Venkateshwara Univ., 1986 ; Osmania Univ., 1988*)

Solution. By inspection it is difficult to say which is the modal class. Hence we prepare a grouping table and an analysis table.

GROUPING TABLE

Weight (in lbs.)	*Col. 1*	*Col. 2*	*Col. 3*	*Col. 4*	*Col. 5*	*Col. 6*
	No. of persons					
100—110	4	10		30		
110—120	6		26		58	
120—130	20	52				85
130—140	32		65	82		
140—150	33	50			58	
150—160	17		25			27
160—170	8	10				
170—180	2					

It is clear from the histogram that the modal value is :

Direct Calculation

Mode lies in the class 25—30

$$M_0=L+\frac{\Delta_1}{\Delta_1+\Delta_2}\times i$$

$L=25$, $\Delta_1=(150-110)=40$, $\Delta_2=(150-120)=30$, $i=10$

$$M_0=25+\frac{40}{40+30}\times 5=25+2{\cdot}86=27{\cdot}86$$

The slight difference in the two answers is due to the difficulty of reading very precisely from the graph.

Mode can also be determined from a frequency polygon in which case a perpendicular is drawn on the base from the apex of the polygon and the point where it meets the base gives the modal value.

However, graphic method of determining mode can be used only where there is one class containing the highest frequency. If two or more classes have the same highest frequency, mode cannot be determined graphically. For example, for the data given below mode cannot be graphically ascertained.

Size of shoes	*No. of persons wearing*	*Size of shoes*	*No. of persons wearing*
2—4	10	8—10	8
4—6	15	10—12	2
6—8	15		

Merits and Limitations of Mode

Merits. The main merits of mode are :

1. By definition mode is the most typical or representative value of a distribution. Hence, when we talk of modal wage, modal size of shoe or modal size of family it is this average that we refer to. The mode is a measure which actually does indicate what many people incorrectly believe the arithmetic mean indicates. The mode is the most frequently occurring value. If the modal wage in a factory is Rs. 916 then more workers receive Rs. 916 than any other wage. This is what many believe the "average" wage always indicates, but actually such a meaning is indicated only if the average used is the mode.

2. Like median, the mode is not unduly affected by extreme values. Even if the high values are very high and the low values are very low we choose the most frequent value of the data to the modal value ; for example, the mode of 10, 2, 5, 10, 5, 60, 5, 10, 60 is 10 as this value, *i.e.*, 10 has occurred most often in the data set.

3. Its value can be determined in open-end distributions without ascertaining the class limits.

4. It can be used to describe qualitative phenomenon. For example, if we want to compare the consumer preferences for different types of products, say, soap, toothpaste, etc, or different media of advertising we should complete the modal preferences expressed by different groups of people.

5. The value of mode can also be determined graphically whereas the value of mean cannot be graphically ascertained.

Limitations. The important limitations of this average are :

1. The value of mode cannot always be determined. In some cases we may have a bimodal series.

2. It is not capable of algebraic manipulations. For example, from the modes of two sets of data we cannot calculate the overall mode of the combined data. Similarly, the modal wage times the number of workers will not give the total payroll—except, of course, when the distribution is normal and then the mean, median, and mode are all equal.

3. The value of mode is not based on each and every item of the series.

4. It is not a rigidly defined measure. There are several formulae for calculating the mode, all of which usually give somewhat different answers. In fact, mode is the most unstable average and its value is difficult to determine.

5. While dealing with quantitative data, the disadvantages of the mode outweigh its good features and hence it is seldom used.

Usefulness. The mode is employed when the most typical value of a distribution is desired. It is the most meaningful measure of central tendency in case of highly skewed or non-normal distributions, as it provides the best indication of the point of maximum concentration.

Relationship among Mean, Median and Mode

A distribution in which the values of mean, median and mode