STATISTICS

CONTENTS

- Class Mark
- Cummulative Frequency
- Mean
- Median
- Mode
- Ogive Curve

IMPORTANT POINTS

❖ The word data means information (its exact dictionary meaning is: given facts). Statistical data are of two types :

(i) Primary data (ii) Secondary data

❖ When an investigator collects data himself with a definite plan or design in his (her) mind, it is called **Primary data**.

❖ Data which are not originally collected rather obtained from published or unpublished sources are known as **Secondary data**.

❖ After collection of data, the investigator has to find ways to condense then in tabular form in order to study their silent features. Such an arrangement is called **Presentation of data**.

❖ Raw data when put in ascending or descending order of magnitude is called an array or arranged data.

❖ The number of times an observation occurs in the given data is called frequency of the observation.

❖ Classes/class intervals are the groups in which all the observations are divided.

❖ Suppose class-interval is 10-20, then 10 is called lower limit and 20 is called upper limit of the class

❖ Mid-value of class-interval is called **Class-mark**

$$\text{Class-mark} = \frac{\text{lower limit} + \text{upper limit}}{2}$$

$$\text{Class-mark} = \text{lower limit} + \frac{1}{2}$$

(difference between the upper and lower limits)

❖ If the frequency of first class interval is added to the frequency of second class and this sum is added to third class and so on then frequencies so obtained are known as **Cumulative Frequency (c.f.).**

❖ There are two types of cumulative frequencies (a) less than, (b) greater than

❖ EXAMPLES ❖

Ex.1 Given below are the ages of 25 students of class IX in a school. Prepare a discrete frequency distribution.

15, 16, 16, 14, 17, 17, 16, 15, 15, 16, 16, 17, 15, 16, 16, 14, 16, 15, 14, 15, 16, 16, 15, 14, 15.

Sol. Frequency distribution of ages of 25 students

Age	Tally marks	Frequency
14	||||	4
15	~~||||~~ |||	8
16	~~||||~~ ~~||||~~	10
17	|||	3
Total		25

Ex.2 Form a discrete frequency distribution from the following scores:-

Sol. 15, 18, 16, 20, 25, 24, 25, 20, 16, 15, 18, 18, 16, 24, 15, 20, 28, 30, 27, 16, 24, 25, 20, 18, 28, 27, 25, 24, 24, 18, 18, 25, 20, 16, 15, 20, 27, 28, 29, 16.

Frequency Distribution of Scores		
Variate	Tally marks	Frequency
15	||||	4
16	卌 |	6
18	卌 |	6
20	卌 |	6
24	卌	5
25	卌	5
27	|||	3
28	|||	3
29	|	1
30	|	1
Total		40

Ex.3 The water tax bills (in rupees) of 30 houses in a locality are given below. Construct a grouped frequency distribution with class size of 10.

30, 32, 45, 54, 74, 78, 108, 112, 66, 76, 88, 40, 14, 20, 15, 35, 44, 66, 75, 84, 95, 96, 102, 110, 88, 74, 112, 14, 34, 44.

Sol. Here the maximum and minimum values of the variate are 112 and 14 respectively.

$\therefore$ Range = 112 – 14 = 98.

It is given that the class size is 10, and

$$\frac{\text{Range}}{\text{Class size}} = \frac{98}{10} = 9.8$$

So, we shoule have 10 classes each of size 10.

The minimum and maximum values of the variate are 14 and 112 respectively. So we have to make the classes in such a way that first class includes the minimum value and the last class includes the maximum value. If we take the first class as 14-24 it includes the minimum value 14. If the last class is taken as 104-114, then it includes the maximum value 112. Here, we form classes by exclusive method. In the class 14-24, 14 is included but 24 is excluded. Similarly, in other classes, the lower limit is included and the upper limit is excluded.

In the view of above discussion, we construct the frequency distribution table as follows:

Bill (in rupees)	Tally marks	Frequency
14-24	||||	4
24-34	||	2
34-44	|||	3
44-54	|||	3
54-64	|	1
64-74	||	2
74-84	卌	5
84-94	|||	3
94-104	|||	3
104-114	||||	4
Total		30

Ex.4 The marks obtained by 40 students of class IX in an examination are given below :

18, 8, 12, 6, 8, 16, 12, 5, 23, 2,16, 23, 2, 10, 20, 12, 9, 7, 6, 5, 3, 5, 13, 21, 13, 15, 20, 24, 1, 7, 21, 16, 13, 18, 23, 7, 3, 18, 17, 16.

Present the data in the form of a frequency distribution using the same class size, one such class being 15-20 (where 20 is not included)

Sol. The minimum and maximum marks in the given raw data are 0 and 24 respectively. It is given that 15-20 is one of the class intervals and the class size is same. So, the classes of equal size are

0-5, 5-10, 10-15, 15-20 and 20-25

Thus, the frequency distribution is as given under :

Frequency Distribution of Marks

Marks	Tally marks	Frequency
0-5	卌 |	6
5-10	卌 卌	10
10-15	卌 |||	8
15-20	卌 |||	8
20-25	卌 |||	8
	Total	40

Ex.5 The class marks of a distribution are :

47, 52, 57, 62, 67, 72, 77, 82, 87, 92, 97, 102
Determine the class size, the class limits and the true class limits.

Sol. Here the class marks are uniformly spaced. So, the class size is the difference between any two consecutive class marks

$\therefore$ Class size = 52 – 47 = 5

We know that, if a is the class mark of a class interval and h is its class size, then the lower and upper limits of the class interval are $a - \frac{h}{2}$ and $a + \frac{h}{2}$ respectively.

$\therefore$ Lower limit of first class interval

$= 47 - \frac{5}{2} = 44.5$

And, upper limit of first class interval

$= 47 + \frac{5}{2} = 49.5$

So, first class interval is 44.5 – 49.5

Similarly, we obtain the other class limits as given under :

Class marks	Class limits
47	44.5 - 49.5
52	49.5 - 54.5
57	54.5 - 59.5
62	59.5 - 64.5
67	64.5 - 69.5
72	69.5 - 74.5
77	74.5 - 79.5
82	79.5 - 84.5
87	84.5 - 89.5
92	89.5 - 94.5
97	94.5 - 99.5
102	99.5 - 104.5

Since the classes are exclusive (continuous) so the true class limits are same as the class limits.

Ex.6 The class marks of a distribution are 26, 31, 36, 41, 46, 51, 56, 61, 66, 71. Find the true class limits.

Sol. Here the class marks are uniformly spaced. So, the class size is the difference between any two consecutive class marks.

$\therefore$ Class size = 31 – 26 = 5.

If a is the class mark of a class interval of size h, then the lower and upper limits of the class interval are $a - \frac{h}{2}$ and $a + \frac{h}{2}$ respectively.

Here h = 5

$\therefore$ Lower limit of first class interval

$= 26 - \frac{5}{2} = 23.5$

And, upper limit of first class interval

$= 26 + \frac{5}{2} = 28.5$

$\therefore$ First class interval is 23.5 – 28.5.

Thus, the class intervals are**:**

23.5 – 28.5, 28.5 – 33.5, 33.5 – 38.5,38.5 – 43.5, 43.5 – 48.5, 48.5 – 53.5

Since the classs are formed by exclusive method. Therefore, these limits are true class limits.

➤ CUMULATIVE FREQUENCY

A table which displays the manner in which cumulative frequencies are distributed over various classes is called a cumulative frequency distribution or cumulative frequency table.

There are two types of cumulative frequency.

(1) Less than type

(2) Greater than type

❖ EXAMPLES ❖

Ex.7 Write down less than type cumulative frequency and greater than type cumulative frequency.

Height (in cm)	Frequency
140 – 145	10
145 – 150	12
150 – 155	18
155 – 160	35
160 – 165	45
165 – 170	38
170 – 175	22
175 – 180	20

Sol. We have

Height (in cm)	140–145	145–150	150–155	155–160	160–165	165–170	170–175	175–180
Frequency	10	12	18	35	45	38	22	20
Height Less than type	145	150	155	160	165	170	175	180
Cumulative frequency	10	22	40	75	120	158	180	200
Height Greater than type	140	145	150	155	160	165	170	175
Cumulative frequency	200	190	178	160	125	80	42	20

Ex.8 The distances (in km) covered by 24 cars in 2 hours are given below :

125, 140, 128, 108, 96, 149, 136, 112, 84, 123, 130, 120, 103, 89, 65, 103, 145, 97, 102, 87, 67, 78, 98, 126

Represent them as a cumulative frequency table using 60 as the lower limit of the first group and all the classes having the class size of 15.

Sol. We have, Class size = 15

Maximum distance covered = 149 km.

Minimum distance covered = 65 km.

$\therefore$ Range = (149 – 65) km = 84 km.

So, number of classes = 6 $\left[\because \frac{84}{15} = 5.6\right]$

Thus, the class intervals are 60-75, 75-90, 90-105, 105-120, 120-135, 135-50.

The cumulative frequency distribution is as given below **:**

Class interval	Tally marks	Frequency	Cumulative frequency
60-75	||	2	2
75-90	||||	4	6
90-105	~~||||~~ |	6	12
105-120	||	2	14
120-135	~~||||~~ |	6	20
135-150	||||	4	24

Ex.9 The following table gives the marks scored by 378 students in an entrance examination :

Mark	No. of students
0 - 10	3
10 - 20	12
20 - 30	36
30 - 40	76
40 - 50	97
50 - 60	85
60 - 70	39
70 - 80	12
80 - 90	12
90 - 100	6

From this table form (i) the less than series, and (ii) the more than series.

Sol. (i) Less than cumulative frequency table

Marks obtained	Number of students (Cumulative frequency)
Less than 10	3
Less than 20	15
Less than 30	51
Less than 40	127
Less than 50	224
Less than 60	309
Less than 70	348
Less than 80	360
Less than 90	372
Less than 100	378

(ii) More than cumulative frequency table

Marks obtained	Number of students (Cumulative frequency)
More than 0	378
More than 9	375
More than 19	363
More than 29	327
More than 39	257
More than 49	154
More than 59	69
More than 69	30
More than 79	18
More than 89	6

Ex.10 Find the unknown entries (a,b,c,d,e,f,g) from the following frequency distribution of heights of 50 students in a class :

Class intervals (Heights in cm)	Frequency	Cumulative frequency
150 - 155	12	a
155 - 160	b	25
160 - 165	10	c
165 - 170	d	43
170 - 175	e	48
175 - 180	2	f
Total	g	

Sol. Since the given frequency distribution is the frequency distribution of heights of 50 students. Therefore,

$g = 50.$

From the table, we have

$a = 12,\ b + 12 = 25,\ 12 + b + 10 = c,$

$12 + b + 10 + d = 43,$

$12 + b + 10 + d + e = 48$ and

$12 + b + 10 + d + e + 2 = f$

Now, $b + 12 = 25$

$\Rightarrow\ b = 13$

$12 + b + 10 = c$

$\Rightarrow\ 12 + 13 + 10 = c \quad [\because\ b = 13]$

$\Rightarrow\ c = 35$

$12 + b + 10 + d = 43$

$\Rightarrow\ 12 + 13 + 10 + d = 43\ \ [\because\ b = 13]$

$\Rightarrow\ d = 8$

$12 + b + 10 + d + e = 48$

$\Rightarrow\ 12 + 13 + 10 + 8 + e = 48$

$[\because\ b = 13,\ d = 8]$

$\Rightarrow\ e = 5$

and, $12 + b + 10 + d + e + 2 = f$

$\Rightarrow\ 12 + 13 + 10 + 8 + 5 + 2 = f$

$\Rightarrow\ f = 50.$

Hence, $a = 12,\ b = 13,\ c = 35,\ d = 8,$

$e = 5,\ f = 50$ and $g = 50.$

Ex.11 The marks out of 10 obtained by 32 students are : 2, 4, 3, 1, 5, 4, 3, 8, 9, 7, 8, 5, 4, 3, 6, 7, 4, 7, 9, 8, 6, 4, 2, 1, 0, 0, 2, 6, 7, 8, 6, 1.

Array the data and form the frequency distribution

Sol. An array of the given data is prepared by arranging the scores in ascending order as follows :

0, 0, 1, 1, 1, 2, 2, 2, 3,3,3, 4,4,4,4,4,

5,5, 6,6,6,6, 7,7,7,7, 8,8,8,8, 9,9.

Frequency distribution of the marks is shown below.

Marks	Tally marks	Frequency
0	||	2
1	|||	3
2	|||	3
3	|||	3
4	~~||||~~	5
5	||	2
6	||||	4
7	||||	4
8	||||	4
9	||	2

Ex.12 Prepare a discrete frequency distribution from the data given below, showing the weights in kg of 30 students of class VI.

39, 38, 42, 41, 39, 38, 39, 42, 41, 39, 38, 38
41, 40, 41, 42, 41, 39, 40, 38, 42, 43, 45, 43,
39, 38, 41, 40, 42, 39.

Sol. The discrete frequency distribution table for the weight (in kg) of 30 students is shown below.

Weights (in kg)	Tally marks	Frequency
38	卌 |	6
39	卌 ||	7
40	|||	3
41	卌 |	6
42	卌	5
43	||	2
45	|	1

Ex.13 The class marks of a distribution are 82, 88, 94, 100, 106, 112 and 118. Determine the class size and the classes.

Sol. The class size is the difference between two consecutive class marks. $\therefore$ Class size = 88 – 82 = 6. Now 82 is the class mark of the first class whose width is 6. $\therefore$ Class limits of the first class are $82 - \frac{6}{2}$ and $82 + \frac{6}{2}$ i.e. 79 and 85. Thus, the first class is 79-85. Similarly, the other classes are 85–91, 91– 97, 97–103,
103 –109, 109 –115 and 115 –121.

Ex.14 The class marks of a distribution are 13, 17, 21, 25 and 29. Find the true class limits.

Sol. The class marks are 13, 17, 21, 25 and 29.

The class marks are uniformly spaced.

Class size = difference between two consecutive class marks

= 17 – 13 = 4

Half of the class size = $\frac{4}{2}$ = 2

To find the classes one has to subtract 2 from and add 2 to each of the class marks.

Hence, the classes are

11– 15

15 – 19

19 – 23

23 – 27

27 – 31

Since the classes are exclusive, the true class limits are the same as the class limits. So the lower class limits as well as the true lower class limits are 11, 15, 19, 23 and 27. The upper class limits as well as the true upper class limits are 15, 19, 23, 27 and 31.

Ex.15 Convert the given simple frequency series into a:

(i) Less than cumulative frequency series.

(ii) More than cumulative frequency series.

Marks	No. of students
0-10	3
10-20	7
20-30	12
30-40	8
40-50	5

Sol. (i) Less than cumulative frequency series

Marks	No. of students
Less than 10	3
Less than 20	10 (= 3 + 7)
Less than 30	22 (= 3 + 7 + 12)
Less than 40	30 (= 3 + 7 + 12 + 8)
Less than 50	35 (= 3 + 7 + 12 + 8 + 5)

(ii) More than cumulative frequency series

Marks	No. of students
More than 50	0
More than 40	5
More than 30	13
More than 20	25
More than 10	32
More than 0	35

Ex.16 Convert the following more than cumulative frequency series into simple frequency series.

Marks	No. of students
More than 0	40
More than 10	36
More than 20	29
More than 30	16
More than 40	5
More than 50	0

Sol. Simple frequency distribution table

Marks	Frequency
0-10	4 (= 40 – 36)
10-20	7 (= 36 – 29)
20-30	13 (= 29 – 16)
30-40	11 (= 16 – 5)
40-50	5 (= 5 – 0)

Ex.17 Drawn ogive for the following frequency distribution by less than method

Marks:	0-10	10-20	20-30	30-40	40-50	50-60
No. of Students	7	10	23	51	6	3

Sol. We first prepare the cumulative frequency distribution table by less than method as given below :

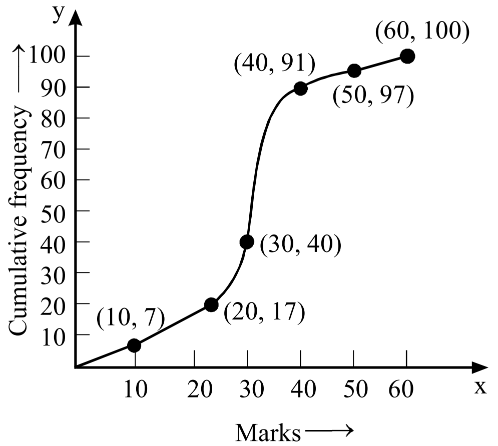
Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of Students	7	10	23	51	6	2
Marks less than	10	20	30	40	50	60
Cumulativef requency	7	17	40	91	97	100

Other thanthe given class intervals, we assume a class – 10-0 before the first class interval 0-10 with zero frequency.

Now, we mark the upper class limits (including the imagined class) along X-axis on a suitable scale and the cumulative frequencies along Y-axis on a suitable scale.

Thus, we plot the points

(0, 0), (10, 7), (20, 17), (30, 40), (40, 91), (50,97), and (60, 100)



Now, we join the plotted points by a free hand curve to obtain the required ogive.

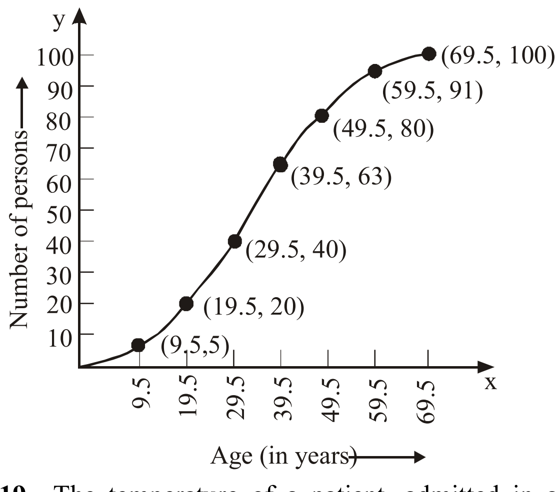
Ex.18 Draw a cumulative frequency curve for the following frequency distribution by less than method

Age (in years)	0-9	10-19	20-29	30-39	40-49	50-59	60-69
No. of persons:	5	15	20	23	17	11	9

Sol. The given frequency distribution is not continuous. So, we first make it continuous and prepare the cumulative frequency distribution as under :

Age (in years)	Frequency	Age less than	Cumulative frequency
– 0.5 - 9.5	5	9.5	5
9.5 - 19.5	15	19.5	20
19.5 - 29.5	20	29.5	40
29.5 - 39.5	23	39.5	63
39.5 - 49.5	17	49.5	80
49.5 - 59.5	11	59.5	91
59.5 - 69.5	9	69.5	100

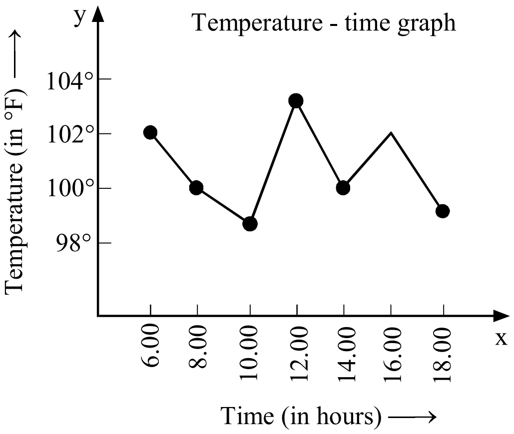
Now, we plot points (9.5, 5), (19.5, 20), (29.5,40), (39.5, 63), (49.5, 80), (59.5, 91) and (69.5, 100) and join them by a free hand smooth curve to obtain the required ogive as shown in Fig.



Ex.19 The temperature of a patient, admitted in a hospital with typhoid fever, taken at different times of the day are given below. Draw the temperature-time graph to reprents the data:

Time (in hours):	6:00	8:00	10:00	12:00	14:00	16:00	18:00
Temperature (in °F)	102	100	99	103	100	102	99

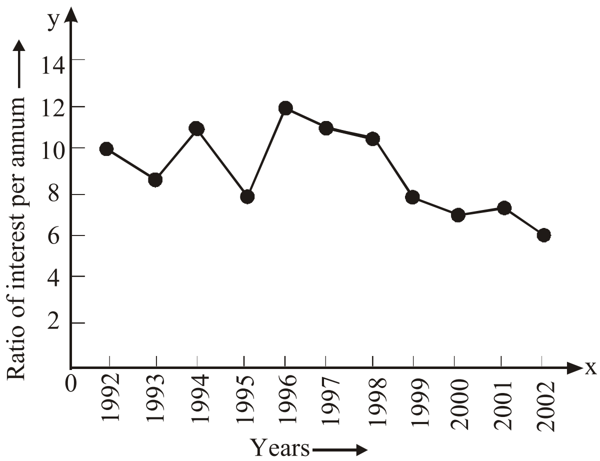
Sol. In order to draw the temperature-time graph, we represent time (in hours) on the x-axis and the temperature in ºF on the y-axis. We first plot the ordered pairs (6, 102), (8, 100), (10, 99), (12, 103), (14, 100), (16, 102) and (18, 99) as points and then join them by line segments as shown in Fig.



Ex.20 The graph shown in Fig. exhibits the rate of interest on fixed deposite upto one year announced by the reserve bank of india in different years. Read the graph and find.

(i) In which period was the rate of interest maximum?

(ii) In which period was the rate of interest minimum ?



Sol. In the graph, we find that years are represented on x-axis and the rate of interest per annum is along y-axis. From the graph, we find that

(i) The rate of interest was maximum (12%) in 1996.

(ii) The minimum rate of interest was 6.5% in the year 2002.

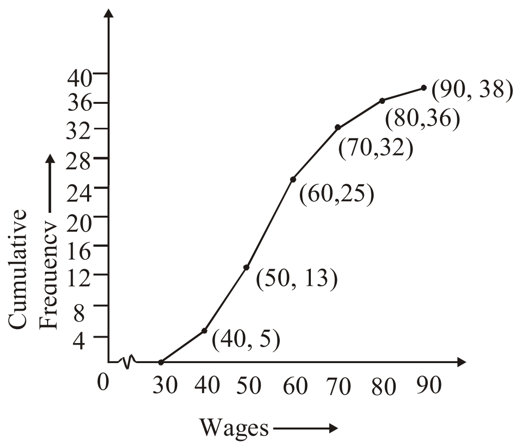
Ex.21 The following data represents the wages of 25 workers of a certain factory :

Weges (in rupees)	No. of workers
30-40	5
40-50	8
50-60	12
60-70	7
70-80	4
80-90	2

Sol. The cumulative frequency table is constructed as follows :

Wages (in rupees)	No. of workers	Cumulative frequency
30-40	5	5
40-50	8	13
50-60	12	25
60-70	7	32
70-80	4	36
80-90	2	38

The cumulative frequency curve is shown below:

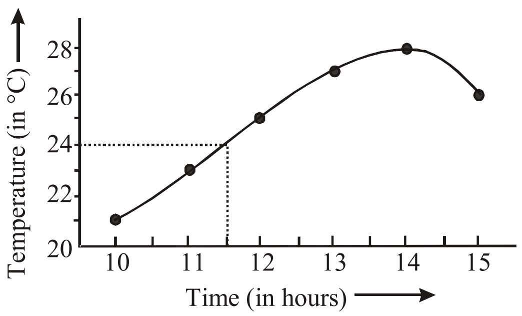


Ex.22 Draw the Time-Temperature graph from the following table

Time (in hour)	Temperature (in °C)
10-00	21
11-00	23
12-00	25
13-00	27
14-00	28
15-00	26

From the graph estimate the temperature at 11-30 a.m.

Sol. Time in hours is denoted along the X-axis and temperature (in °C) is inidicated along the Y-axis. The points are joined by drawing a freehand curve. From the graph, the temperature at 11-30 a.m. is found to be 24.0°C.



◈ **Disadvantages**

(i) It is highly affected by the presence of a few abnormally high or abnormally low scores.

(ii) In absence of a single item, its value becomes inaccurate.

(iii) It can not be determined by inspection.

❖ EXAMPLES ❖

Ex.23 If the mean of n observations $ax_1, ax_2, ax_3, ..., ax_n$ is $a\overline{X}$, show that

$$(ax_1 - a\overline{X}) + (ax_2 - a\overline{X}) + ... + (ax_n - a\overline{X}) = 0$$

Sol. We have

$$a\overline{X} = \frac{ax_1 + ax_2 + ... + ax_n}{n}$$

$\Rightarrow ax_1 + ax_2 + ... + ax_n = n(a\overline{X})$ ....(i)

Now, $(ax_1 - a\overline{X}) + (ax_2 - a\overline{X}) + ... + (ax_n - a\overline{X})$

$= (ax_1 + ax_2 + ... + ax_n) - \underbrace{(a\overline{X} + a\overline{X} + ... + a\overline{X})}_{\textbf{n – terms}}$

$= n(a\overline{X}) - n(a\overline{X}) = 0.$

Ex.24 The mean of n observations $x_1, x_2, ..., x_n$ is $\overline{X}$. If (a – b) is added to each of the observations, show that the mean of the new set of observations is $\overline{X} + (a - b)$

Sol. We have,

$$\overline{X} = \frac{x_1 + x_2 + ... + x_n}{n} \quad ....(i)$$

Let $\overline{X}'$, be the mean of $x_1 + (a - b)$, $x_2 + (a - b)$, ..., $x_n + (a - b)$. Then,

$$\overline{X} = \frac{\{x_1 + (a - b)\} + \{x_2 + (a - b)\} + ... + \{x_n + (a - b)\}}{n}$$

$$\Rightarrow \overline{X}' = \frac{x_1 + x_2 + ... + x_n + n(a - b)}{n}$$

$= \overline{X} + (a - b)$ (using (i))

Ex.25 Find the sum of the deviations of the variate values 3, 4, 6, 8, 14 from their mean.

Sol. Recall that the deviations of the values $x_1, x_2, x_3, ..., x_n$ about A are

$x_1 - A, x_2 - A, x_3 - A, ..., x_n - A.$

Let $\overline{X}$ be the mean of the values 3, 4, 6, 8, 14. Then,

$$\overline{X} = \frac{3 + 4 + 6 + 8 + 14}{5} = \frac{35}{5} = 7$$

Now, sum of the deviations of the values 3, 4, 6, 8, 14 from their mean $\overline{X} = 7$ is given by

$= (3 - 7) + (4 - 7) + (6 - 7) + (8 - 7) + (14 - 7) = -4 - 3 - 1 + 1 + 7 = 0.$

Ex.26 The mean of 40 observations was 160. It was detected on rechecking that the value of 165 was wrongly copied as 125 for computation of mean. Find the correct mean.

Sol. ∵ Here, n = 40, $\overline{X} = 160$

So, $\overline{X} = \frac{1}{n}\left(\sum x_i\right) \Rightarrow 160 = \frac{1}{40}\left(\sum x_i\right)$

$\Rightarrow \sum x_i = 160 \times 40 = 6400$

$\Rightarrow$ Incorrect value of $\sum x_i = 6400$

Now,

Correct value of $\sum x_i$

= Incorrect value of $\sum x_i$ – Incorrect item + Correct item

$\Rightarrow$ Correct value of $\sum x_i = 6400 - 125 + 165 = 6440$

$\therefore$ Correctmean

$$= \frac{\text{Correct value of} \sum x_i}{n} = \frac{6440}{40} = 161.$$

Ex.27 The mean of 10 numbers is 20. If 5 is subtracted from every number, what will be the new mean?

Sol. Let $x_1, x_2, ..., x_{10}$ be 10 numbers with their mean equal to 20. Then,

$$\overline{X} = \frac{1}{n}\left(\sum x_i\right)$$

$$\Rightarrow 20 = \frac{x_1 + x_2 + ... + x_{10}}{10}$$

$\Rightarrow x_1 + x_2 + ... + x_{10} = 200$ ....(i)

New numbers are $x_1 - 5, x_2 - 5, ..., x_{10} - 5$. Let $\overline{X}'$ be the mean of new numbers.

Then,

$$\overline{X}' = \frac{(x_1 - 5) + (x_2 - 5) + ... + (x_{10} - 5)}{10}$$

$$\overline{X}' = \frac{(x_1 + x_2 + ... + x_{10}) - 5 \times 10}{10} = \frac{200 - 50}{10}$$

[Using (i)]

$\overline{X}' = 15.$

Ex.28 The mean of 16 numbers is 8. If 2 is added to every number, what will be the new mean ?

Sol. Let $x_1, x_2, x_3, ..., x_{16}$ be 16 numbers with their mean equal to 8. Then,

$$\overline{X} = \frac{1}{n}\left(\sum x_i\right)$$

$$\Rightarrow 8 = \frac{x_1 + x_2 + ... + x_{16}}{16}$$

$\Rightarrow x_1 + x_2 + ... + x_{16} = 16 \times 8 = 128$ ...(i)

New numbers are $x_1 + 2, x_2 + 2, x_3 + 2, ..., x_{16} + 2$. Let $\overline{X}'$ be the mean of new numbers. Then,

$$\overline{X}' = \frac{(x_1 + 2) + (x_2 + 2) + ... + (x_{16} + 2)}{16}$$

$$\Rightarrow \overline{X}' = \frac{(x_1 + x_2 + ... + x_{16}) + 2 \times 16}{16} = \frac{128 + 32}{16}$$

[Using(i)]

$$\Rightarrow \overline{X}' = \frac{160}{16} = 10$$

Ex.29 If $x_1, x_2, ..., x_n$ are n values of a variable X such that

$$\sum_{i=1}^{n}(x_i - 2) = 110 \text{ and } \sum_{i=1}^{n}(x_i - 5) = 20.$$

Find the value of n and the mean.

Sol. We have,

$$\sum_{i=1}^{n}(x_i - 2) = 110 \text{ and } \sum_{i=1}^{n}(x_i - 5) = 20$$

$\Rightarrow (x_1 - 2) + (x_2 - 2) + ... + (x_n - 2) = 110$

and $(x_1 - 5) + (x_2 - 5) + ... + (x_n - 5) = 20$

$\Rightarrow (x_1 + x_2 + ... + x_n) - 2n = 110$ and

$(x_1 + x_2 + ... + x_n) - 5n = 20$

$$\Rightarrow \sum_{i=1}^{n} x_i - 2n = 110 \text{ and } \sum_{i=1}^{n} x_i - 5n = 20$$

$\Rightarrow S - 2n = 110$ and $S - 5n = 20$

Thus, we have

$S - 2n = 110$ ....(i)

and $S - 5n = 20$ ....(ii)

Subtracting (ii) from (i), we get

$3n = 90 \Rightarrow n = 30$

Putting n = 30 in (i), we get

$$S - 60 = 110 \;\Rightarrow\; S = 170$$

$$\Rightarrow \sum_{i=1}^{n} x_i = 170$$

$$\therefore \text{ Mean} = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) = \frac{170}{30} = \frac{17}{3}$$

Hence, n = 30 and mean $\frac{17}{3}$.

Ex.30 The sum of the deviations of a set of n values $x_1, x_2, ..., x_n$ measured from 50 is–10 and the sum of deviations of the values from 46 is 70. Find the values of n and the mean.

Sol. We have,

$$\sum_{i=1}^{n}(x_i - 50) = -10 \text{ and } \sum_{i=1}^{n}(x_i - 46) = 17$$

$$\Rightarrow \sum_{i=1}^{n} x_i - 50n = -10 \qquad ....(i)$$

$$\text{and } \sum_{i=1}^{n} x_i - 46n = 70 \qquad ....(ii)$$

Subtracting (ii) from (i), we get

$$-4n = -80 \;\Rightarrow\; n = 20$$

Putting n = 20 in (i), we get

$$\sum_{i=1}^{n} x_i = 50 \times 20 = -10 \;\Rightarrow\; \sum_{i=1}^{n} x_i = 990$$

$$\therefore \text{ Mean} = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) = \frac{990}{20} = 49.5$$

Hence, n = 20 and mean = 49.5

➢ ARITHMETIC MEAN OF UNGROUPED DATA

Arithmetic mean of raw data (when frequency is not given) :

The arithmetic mean of a raw data is obtained by adding all the values of the variables and dividing the sum by total number of values that are added.

Arithmetic mean

$$(\bar{x}) = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

The symbol $\sum_{i=1}^{n} x_i$ denotes the sum $x_1 + x_2 + ... + x_n$.

❖ EXAMPLES ❖

Ex.31 Neeta and her four friends secured 65, 78, 82, 94 and 71 marks in a test of mathematics. Find the average (arithmetic mean) of their marks.

Sol. Arithmetic mean or average

$$= \frac{65 + 78 + 82 + 94 + 71}{5} = \frac{390}{5} = 78$$

Hence, arithmetic mean = 78

Ex.32 The marks obtained by 10 students in physics out of 40 are 24, 27, 29, 34, 32, 19, 26, 35, 18, 21. Compute the mean of the marks.

Sol. Mean of the marks is given by

$$\bar{x} = \frac{24 + 27 + 29 + 34 + 32 + 19 + 26 + 35 + 18 + 21}{10}$$

$$= \frac{265}{10} = 26.50$$

Ex.33 The mean of 20 observations was found to be 47. But later it was discovered that one observation 66 was wrongly taken as 86. Find the correct mean.

Sol. Here, n = 20, $\bar{x} = 47$

We have, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \therefore 47 = \frac{\sum_{i=1}^{n} x_i}{20}$

$$\sum_{i=1}^{n} x_i = 47 \times 20 = 940.$$

But the score 66 was wrongly taken as 86.

$\therefore$ Correct value of $\sum_{i=1}^{n} x_i = 940 + 66 - 86 = 920$

$\therefore$ Correct mean $= \frac{920}{20} = 46$

Ex.34 If $\bar{x}$ denote the mean of $x_1, x_2, ..., x_n$, show that

$$\sum_{i=1}^{n} = (x_i - \bar{x})$$

Sol. $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}$

$= x_1 + x_2 + ... + x_n = n\bar{x}$ (i)

$= \Sigma(x_1 - \bar{x}) = (x_1 - \bar{x}) + (x_2 - \bar{x}) + ... + (x_n - x_1)$

$= (x_1 + x_2 + ... + x_n) - n\bar{x} = n\bar{x} - n\bar{x}$

$= 0$ (from (i))

Ex.35 If the mean of 5 observations is 15 and that of another 10 observations is 20, find the mean of all 15 observations

Sol. Let first five observations be $x_1, ..., x_5$

$\Rightarrow$ Mean $= \frac{x_1 + x_2 ... + x_5}{5}$

$\Rightarrow 15 = \frac{x_1 + x_2 ... + x_5}{5}$

$\Rightarrow x_1 + ... + x_5 = 75$ (i)

Let next ten observations be $y_1 + ... + y_{10}$.

$\Rightarrow$ Mean $= \frac{y_1 + ... + y_{10}}{10} \Rightarrow 20 = \frac{y_1 + ... + y_{10}}{10}$

$y_1 + ... + y_{10} = 200$ (ii)

The mean of all 15 observations will be

$\frac{(x_1 + ... + x_5) + (y_1 + ... + y_{10})}{15}$

$= \frac{75 + 200}{15}$ (from (i) and (ii))

$= 18.33$

If a variate X takes values $x_1, x_2 ... x_n$ with corresponding frequencies $f_1, f_2, f_3, ..., f_n$ respectively, then arithmetic mean of these values is

$$\overline{X} = \frac{f_1x_1 + f_2x_2 + ... + f_nx_n}{f_1 + f_2 + ... + f_n}$$

or $\overline{X} = \frac{\sum_{i=1}^{n} f_i x_i}{N}$, where $N = \sum_{i=1}^{n} f_i = f_1 + f_2 + ... + f_n$

Ex.36 Find the mean of the following distribution :

x :	4	6	9	10	15
f :	5	10	10	7	8

Sol. Calculation of Arithmetic Mean

x_i	f_i	f_ix_i
4	5	20
6	10	60
9	10	90
10	7	70
15	8	120
$N = \Sigma f_i = 40$		$\Sigma f_i x_i = 360$

$\therefore$ Mean $= \overline{X} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{360}{40} = 9.$

Ex.37 Find the mean of the following distribution :

x	10	30	50	70	89
f	7	8	10	15	10

Sol. Calculation of Mean

x_i	f_i	f_ix_i
10	7	70
30	8	240
50	10	500
70	15	1050
89	10	890
	$\Sigma f_i = N = 50$	$\Sigma f_i x_i = 2750$

$\therefore$ Mean $= \frac{\Sigma f_i x_i}{N} = \frac{2750}{50} = 55.$

Ex.38 Find the value of p, if the mean of following distribution is 7.5.

x :	3	5	7	9	11	13
y :	6	8	15	p	8	4

Sol. Calculation of Mean

x_i	f_i	f_ix_i
3	6	18
5	8	40
7	15	105
9	p	9p
11	8	88
13	4	52
	$N = Sf_i = 41 + p$	$Sf_ix_i = 303 + 9p$

We have, $\Sigma f_i = 41 + p$, $\Sigma f_ix_i = 303 + 9p$

$\therefore$ Mean $= \dfrac{\Sigma f_i x_i}{\Sigma f_i} \Rightarrow 7.5 = \dfrac{303+9p}{41+p}$

$\Rightarrow 7.5 \times (41 + p) = 303 + 9p$

$\Rightarrow 307.5 + 7.5p = 303 + 9p$

$\Rightarrow 9p - 7.5p = 307.5 - 303$

$\Rightarrow 1.5p = 4.5 \Rightarrow p = 3$

Ex.39 Find the missing frequencies in the following frequency distribution if it is known that the mean of the distribution is 1.46.

Number of accidents (x) :	0	1	2	3	4	5	Total
Frequency (f) :	46	?	?	25	10	5	200

Sol. Let the missing frequencies be f_t and f_2

Calculation of Mean

x_i	f_i	f_ix_i
0	46	0
1	f_1	f_1
2	f_2	75
3	25	$2f_2$
4	10	40
5	5	25
$N = 86 + f_1 + f_2$		$\Sigma f_ix_i = 140 + f_1 + 2f_2$

We have : N = 200

$\therefore 200 = 86 + f_1 + f_2 \Rightarrow f_1 + f_2 = 114$ ....(i)

Also, Mean = 1.46

$\Rightarrow 1.46 = \dfrac{\Sigma f_i x_i}{N} \Rightarrow 1.46 = \dfrac{140 + f_2 + 2f_2}{200}$

$\Rightarrow 292 = 140 + f_1 + 2f_2$

$\Rightarrow f_1 + 2f_2 = 152$ ....(ii)

Solving (i) and (ii) we get $f_1 = 76$ and $f_2 = 38$.

Ex.40 If the mean of the following data be 9.2, find the value of p.

x	4	6	7	p + 4	12	12
f	5	6	4	10	8	7

Sol. The table is rewritten as below :

x	f	f.x
4	5	20
6	6	36
7	4	28
p+4	10	10p + 40
12	8	96
14	7	98
Total	40	318 + 10p

Now, Mean $\bar{x} = \dfrac{\Sigma f.x}{\Sigma f} = \dfrac{318+10.p}{40}$

$\therefore 9.2 = \dfrac{318+10.p}{40}$

$\Rightarrow 318 + 10.p = 368 \Rightarrow 10p = 50 \Rightarrow p = 5$

Ex.41 The marks of 30 students are given below, find the mean marks.

Marks	Number of Students
10	4
11	3
12	8
13	6
14	7
15	2

Sol.

x	f	fx
10	4	40
11	3	33
12	8	96
13	6	78
14	7	98
15	2	30
	$\Sigma f = 30$	$\Sigma fx = 375$

Mean $= \dfrac{\Sigma fx}{\Sigma f} = \dfrac{375}{30} = 12.5$

GROUPED FREQUENCY DISTRIBUTION

There are 3 methods for calculation of mean :

1. Direct Method
2. Assumed mean deviation method
3. Step deviation method.

Direct Method for Calculation of Mean

mid-value	frequency	$f_i x_i$
x_1	f_1	$f_1 x_1$
x_2	f_2	$f_2 x_2$
:	:	:
x_k	f_k	$f_k x_k$
Total	N	$\sum_{i=1}^{k} f_i x_i$

According to direct method

$$\overline{x} = \frac{x_1 f_1 + x_2 f_2 + ... + x_k f_k}{f_1 + f_2 + ... + f_k} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum f_i} = \frac{1}{N}\sum_{1}^{k} f_i x_i$$

$[N = f_1 + f_2 + ... + f_k]$

❖ EXAMPLES ❖

Ex.42 Calculate the mean for the following distribution:

Variable	5	6	7	8	9
Frequency	4	8	14	11	3

Sol.

x	f	fx
5	4	20
6	8	48
7	14	98
8	11	88
9	3	27
Total	$N = \sum f = 40$	$\sum f x = 281$

$\therefore$ Mean $= \frac{\sum f x}{\sum f} = \frac{281}{40} = 7.025$

Ex.43

Mid-values	2	3	4	5	6
Frequencies	49	43	57	38	13

Mid values of class interval are given with their frequencies. Find the mean by direct method.

Sol.

Mid-values	Frequencies (f_i)	$f_i x_i$
2	49	98
3	43	129
4	57	228
5	38	190
6	13	78
Total	$N = \sum f_i = 200$	$\sum f_i x_i = 723$

By direct method.

Mean $= \frac{\sum f_i x_i}{\sum f_i} = \frac{723}{200} = 3.615$

Ex.44 Find the mean of the following frequency distribution :

Class Interval	Frequency
10 – 30	90
30 – 50	20
50 – 70	30
70 – 90	20
90 – 110	40

Sol.

Class interval	f	Mid value (x)	f x
10 – 30	90	20	1800
30 – 50	20	40	800
50 – 70	30	60	1800
70 – 90	20	80	1600
90 – 110	40	100	4000
	$\sum f = 200$		$\sum f x = 10000$

Mean $(\overline{x}) = \frac{\sum f x}{\sum f} = \frac{10000}{200} = 50$

Ex.45 Find the mean of the following frequency distribution :

Class Interval	Frequency
15 – 25	60
25 – 35	35
35 – 45	22
45 – 55	18
55 – 65	15

Sol.

Class interval	f	Mid value (x)	f x
15 – 25	60	20	1200
25 – 35	35	30	1050
35 – 45	22	40	880
45 – 55	18	50	900
55 – 65	15	60	900
	$\sum f = 150$		$\sum f x = 4930$

Mean($\bar{x}$)

$$= \frac{\sum f x}{\sum f} = \frac{4930}{150} = 32.8\bar{6} \text{ or } 32.87 \text{ (approx.)}$$

Ex.46 A survey was conducted by a group of students as a part of their environment awareness programme, in which they collected the following data regarding the number of plants in 20 houses in a locality. Find the mean number of plants per house.

Number of plants	0 – 2	2 – 4	4 – 6	6 – 8	8 – 10	10 – 12	12 – 14
No. of houses	1	2	1	5	6	2	3

Which method did you use for finding the mean and why ?

Sol.

Number of plants	Number of houses (f)	Mid value x	f x
0 – 2	1	1	1
2 – 4	2	3	6
4 – 6	1	5	5
6 – 8	5	7	35
8 – 10	6	9	54
10 – 12	2	11	22
12 – 14	3	13	39
	$\sum f = 20$		$\sum f x = 162$

$$\text{Mean} = \frac{\sum f x}{\sum f} = \frac{162}{20} = 8.1$$

Here, we have used direct method because numerical values of x and f are small.

Ex.47 Find the mean of the following distribution by direct method.

Class interval	0 – 10	11 – 20	21 – 30	31 – 40	41 – 50
Frequency	3	4	2	5	6

Sol.

Class interval	Frequency f	Mid value x	f x
0 – 10	3	5.0	15.0
11 – 20	4	15.5	62.0
21 – 30	2	25.5	51.0
31 – 40	5	35.5	177.5
41 – 50	6	45.5	273.0
	$\sum f = 20$		$\sum f x = 578.5$

$$\therefore \quad \text{Mean} = \frac{\sum f x}{\sum f} = \frac{578.5}{20} = 28.9$$

Ex.48 For the following distribution, calculate mean using all the suitable methods.

Size of Item	1 – 4	4 – 9	9 – 16	16 – 27
Frequency	6	12	26	20

Sol.

Size of item	Mid value (x_i)	Frequency (f_i)	$f_i x_i$
1– 4	2.5	6	15
4– 9	6.5	12	78
9– 16	12.5	26	325
16– 27	21.5	20	430
		$\sum f_i = 64$	$\sum f_i x_i = 848$

$$\text{Mean} = \frac{\sum f_i x_i}{\sum f_i} = \frac{848}{64} = 13.25$$

◈ Assumed Mean Method

$$\text{Arithmetic mean} = a + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}$$

Note : The assumed mean is chosen, in such a manner, that

1. It should be one of the central values.
2. The deviation are small.
3. One deviation is zero.

Working Rule :

Step 1 : Choose a number 'a' from the central values of x of the first column, that will be our assumed mean.

Step 2 : Obtain deviations d_i by subtracting 'a' from x_i. Write down hese deviations against the corresponding frequencies in the third column.

Step 3 : Multiply the frequencies of second column with corresponding deviations d_i in the third column to prepare a fourth column of $f_i d_i$.

Step 4 : Find the sum of all the entries of fourth column to obtain ∑fidi and also, find the sum of all the frequencies in the second column to obtain Σf_i.

❖ EXAMPLES ❖

Ex.49 The following table gives the distribution of total household expenditure (in rupees) of manual workers in a city.

Expenditure (in rupees)	100-150	150-200	200-250	250-300	300-350	350-400	400-450	450-500
Frequency	24	40	33	28	30	22	16	7

Sol. Let assumed mean = 275

Expenditure (in rupees)	Frequency (f_i)	Mid value (x_i)	$d_i = x-275$	$f_i d_i$
100–150	24	125	–150	–3600
150–200	40	175	–100	–4000
200–250	33	225	–50	–1650
250– 300	28	275	0	0
300– 350	30	325	50	1500
350– 400	22	375	100	2200
400– 450	16	425	150	2400
450– 500	7	475	200	1400
	$\Sigma f_i = 200$			$\Sigma f_i d_i = -1750$

$$\bar{x} = a + \frac{\Sigma f_i d_i}{\Sigma f_i} = 275 + \frac{-1750}{200} = \text{Rs } 266.25$$

Ex.50 Calculate the arithmetic mean of the following distribution :

Class Interval	Frequency
0 – 50	17
50 –100	35
100 –150	43
150–200	40
200– 250	21
250– 300	24

Sol. Let assumed mean = 175 i.e. a = 175

Class	Mid value (x_i)	$d_i = x_i - 175$	frequency f_i	$f_i d_i$
0–50	25	–150	17	–2550
50–100	75	–100	35	–3500
100–150	125	–50	43	–2150
150–200	175	0	40	0
200-250	225	50	21	1050
250–300	275	100	24	2400
			$\Sigma f_i = 180$	$\Sigma f_i d_i = -4750$

Now , a = 175

$$\bar{x} = a + \frac{\Sigma f_i d_i}{\Sigma f_i} = 175 + \frac{-4750}{180}$$

$$= 175 - 26.39 = 148.61 \text{ approx.}$$

Ex.51 Calculate the arithmetic mean of the following frequency distribution :

Class interval	50– 60	60–70	70–80	80–90	90– 100
Frequency	8	6	12	11	13

Sol. Let assumed mean = 75 i.e., a = 75

Class	frequency f_i	Mid value (x_i)	$d_i = x_i - 75$	$f_i d_i$
50–60	8	55	–20	–160
60–70	6	65	–10	–60
70–80	12	75	0	0
80–90	11	85	10	110
90–100	13	95	20	260
	$\Sigma f = 50$			$\Sigma f_i d_i = 150$

$a = 75, \Sigma f_i d_i = 150, \Sigma f_i = 50$

$$\text{Mean } (\bar{x}) = a + \frac{\Sigma f_i d_i}{\Sigma f_i} = 75 + \frac{150}{50} = 78$$

Ex.52 Thirty women were examined in a hospital by a doctor and the number of heart beats per minute were recorded and summarised as follows. Find the mean heart beats per minute for these women, choosing a suitable method.

Number of heart beats per minute	Frequency
65– 68	2
68–71	4
71–74	3
74–77	8
77– 80	7
80– 83	4
83– 86	2

Sol. Let assumed mean a = 75.5

No. of heart beats	No. of women f	Mid value (x)	d = x –a	fd
65–68	2	66.5	–9	–18
68–71	4	69.5	–6	–24
71–74	3	72.5	–3	– 9
74–77	8	75.5	0	0
77–80	7	78.5	3	21
80–83	4	81.5	6	24
83–86	2	84.5	9	18
	$\sum f = 30$			$\sum fd = 12$

$$\text{Mean} = a + \frac{\Sigma fd}{\Sigma f} = 75.5 + \frac{12}{30} = 75.5 + 0.4 = 75.9$$

❖ **Step Deviation Method**

Deviation method can be further simplified on dividing the deviation by width of the class interval h. In such a case the arithmetic mean is reduced to a great extent.

$$\Rightarrow \text{Mean } (\bar{x}) = a + \frac{\Sigma f_i u_i}{\Sigma f_i} \times h$$

Working Rule :

Step-1 : Choose a number 'a' from the central values of x(mid-values)

Step-2 : Obtain $u_i = \frac{x_i - a}{h}$

Step-3 : Multiply the frequency f_i with the corresponding u_i to get $f_i u_i$.

Step-4 : Find the sum of all $f_i u_i$ i.e., $\Sigma f_i u_i$

Step-5 : Use the formula $\bar{x} = a + \frac{\Sigma f_i u_i}{\Sigma f_i} . h$ to get the required mean.

❖ **EXAMPLES** ❖

Ex.53 To find out the concentration of SO_2 in the air (in parts per million, i.e.ppm), the data was collected for 30 localities in a certain city and is presented below :

Concentration of SO_2 (in ppm)	Frequency
0.00 – 0.04	4
0.04 – 0.08	9
0.08 – 0.12	9
0.12 – 0.16	2
0.16 – 0.20	4
0.20 – 0.24	2

Find the mean concentration of SO_2 in the air.

Sol. Let the assumed mean a = 0.10.

Concentration of SO_2 (in ppm)	Frequency f_i	Mid value x_i	$u_i = \frac{x_i - 0.10}{0.04}$	f_iu_i
0.00 – 0.04	4	0.02	–2	–8
0.04 – 0.08	9	0.06	–1	–9
0.08 – 0.12	9	0.10	0	0
0.12 – 0.16	2	0.14	1	2
0.16 – 0.20	4	0.18	2	8
0.20 – 0.24	2	0.22	3	6
	$\Sigma f_i = 30$			$\Sigma f_iu_i = -1$

By step deviation method

$$\text{Mean} = a + \frac{\Sigma f_i u_i}{\Sigma f_i} \times h$$

$$= 0.10 + \frac{-1}{30} \times 0.04$$

$= 0.10 - 0.0013$

$= 0.0987$

$= 0.099$ ppm

Ex.54 The weekly observation on cost of living index in a certain city for the year 2004–2005 are given below. Compute the mean weekly cost of living index.

Cost of Living index	Number of weeks
1400-1500	5
1500-1600	10
1600-1700	20
1700-1800	9
1800-1900	6
1900-2000	2

Sol. Let assumed mean be 1750 i.e., a = 1750

Cost of living index	Frequency f_i	Mid value x_i	$u_i = \frac{x_i - 1750}{100}$	f_iu_i
1400–1500	5	1450	–3	–15
1500–1600	10	1550	–2	–20
1600–1700	20	1650	–1	–20
1700–1800	9	1750	0	0
1800–1900	6	1850	1	6
1900–2000	2	1950	2	4
	$\Sigma f_i = 52$			$\Sigma f_iu_i = -45$

By step deviation method

$$\text{Mean } (\bar{x}) = a + \frac{\Sigma f_i u_i}{\Sigma f_i} h$$

$$= 1750 + \frac{-45}{52} \times 100$$

$= 1750 - 86.54$

$= 1663.46$

Hence, the mean weekly cost of living index = 1663.46

Ex.55 Find the mean marks from the following data by step deviation method

Marks	Number of students
Below 10	5
Below 20	9
Below 30	17
Below 40	29
Below 50	45
Below 60	60
Below 70	70
Below 80	78
Below 90	83
Below 100	85

Sol. Let assumed mean = 55 ⇒ a = 55

Class interval	Frequency f_i	Mid value x_i	$u_i = \frac{x_i - 55}{10}$	f_iu_i
0 – 10	5	5	–5	–25
10 – 20	4	15	–4	–16
20 – 30	8	25	–3	–24
30 – 40	12	35	–2	–24
40 – 50	16	45	–1	–16
50 – 60	15	55	0	0
60 – 70	10	65	1	10
70 – 80	8	75	2	16
80 – 90	5	85	3	15
90 –100	2	95	4	8
	$\Sigma f_i = 85$			$\Sigma f_iu_i = -56$

Here, a = 55, h = 10,

$\Sigma f_i = 85$, $\Sigma f_i u_i = -56$

$$\text{Mean } (\bar{x}) = a + \frac{\Sigma f_i u_i}{\Sigma f_i}$$

$$h = 55 + \frac{-56}{85} \times 10$$

$= 55 - 6.59 = 48.41$

Hence, mean mark = 48.41.

Ex.56 Find the mean age of 100 residents of a colony from the follwing data :

Age in years	Number of persons
Greater than 0	100
Greater than 10	90
Greater than 20	75
Greater than 30	50
Greater than 40	25
Greater than 50	15
Greater than 60	5
Greater than 70	0

Sol. Let assumed mean a = 35

Age (in years)	Number of persons	Mid value xi	$u_i = \frac{x_i - 35}{10}$	fiui
0 – 10	10	5	–3	–30
10 – 20	15	15	–2	–30
20 – 30	25	25	–1	–25
30 – 40	25	35	0	0
40 – 50	10	45	1	10
50 – 60	10	55	2	20
60 – 70	5	65	3	15
Total	$\Sigma f_i = 100$			$\Sigma f_i u_i = -40$

Here, a = 35, h = 10

$$\bar{x} = a + \frac{\Sigma f_i u_i}{\Sigma f_i} \times h$$

$$\Rightarrow \bar{x} = 35 + \frac{-40}{100} \times 10 = 31$$

Hence, the mean age = 31 years :

Ex.57 The following distribution show the daily pocket allowance of children of a locality. The mean pocket allowance is Rs. 18.00. Find the missing frequency f.

Class-Interval	11–13	13–15	15–17	17–19	19–21	21–23	23–25
Frequency	7	6	9	13	f	5	4

Sol. we have,

Class-interval	Frequency	Mid value x	f x
11 – 13	7	12	84
13 – 15	6	14	84
15 – 17	9	16	144
17 – 19	13	18	234
19 – 21	f	20	20f
21 – 23	5	22	110
23 – 15	4	24	96
	$\Sigma f = 44 + f$		$\Sigma fx = 752 + 20f$

$$\text{Mean } \bar{x} = \frac{\Sigma fx}{\Sigma f} \Rightarrow 18 = \frac{752 + 20f}{44 + f}$$

$\Rightarrow 18\,(44 + f) = 752 + 20f$

$\Rightarrow 752 + 20f = 792 + 18f$

$\Rightarrow 2f = 40$

$\Rightarrow f = 20$

Hence, the missing frequency is 20.

Ex.58 The arithmetic mean of the following frequency distribution is 50. Find the value of p.

Class-Interval	0–20	20–40	40–60	60–80	80–100
Frequency	17	P	32	24	19

Sol.

Class-interval	Frequency (f)	Mid value (x)	f x
0 – 20	17	10	170
20 – 40	P	30	30 P
40 – 60	32	50	1600
60 – 80	24	70	1680
80 – 100	19	90	1710
	$\Sigma f = 92 + P$		$\Sigma fx = 5160 + 30P$

Mean $\bar{x} = \frac{\Sigma fx}{\Sigma f} \Rightarrow 50 = \frac{5160+30P}{92+P}$

$\Rightarrow$ 50 (92 + P) = 5160 + 30 P

$\Rightarrow$ 4600 + 50 P = 5160 + 30P

$\Rightarrow$ 20 P = 560 $\Rightarrow$ P = 28

Ex.59 The mean of the following frequency distribution is 62.8 and the sum of all frequencies is 50. Compute the missing frequencies f_1 and f_2 :

Class-Interval	0–20	20–40	40–60	60–80	80–100	100–120	Total
Frequency	5	f1	10	f2	7	8	50

Sol.

Class-interval	Frequency (f)	Mid value (x)	f x
0 – 20	5	10	50
20 – 40	f_1	30	$30 f_1$
40 – 60	10	50	500
60 – 80	f_2	70	$70 f_2$
80 – 100	7	90	630
100 – 120	8	110	880
	$\Sigma f = 30 + f_1 + f_2 = 50$		$\Sigma fx = 2060 + 30f_1 + 70 f_2$

$30 + f_1 + f_2 = 50 \Rightarrow f_1 + f_2 = 20$ ....(1)

Mean $\bar{x} = \frac{\Sigma fx}{\Sigma f} \Rightarrow 62.8 = \frac{2060+30f_1+70f_2}{50}$

$\Rightarrow 62.8 = \frac{206+3f_1+7f_2}{5}$

$\Rightarrow 206 + 3f_1 + 7f_2 = 314$

$\Rightarrow 3f_1 + 7f_2 = 108$ ....(2)

$3f_1 + 3f_2 = 60$ ....(3)

[Multiplying (1) by 3]

On Subtracting (3) from (2), we get

$4f_2 = 48 \Rightarrow f_2 = 12$

Putting $f_2 = 12$ in (1), we get

$f_1 = 8$

➤ MEDIAN

Median of a distribution is the value of the variable which divides the distribution into two equal parts i.e. it is the value of the variable such that the number of observations above it is equal to the number of observations below it.

❖ If the values x_i in the raw data. are arranged in order of increasing or decreasing magnitude, then the middle, most value in the arrangement is called the median.

Algorithm :

Step I : Arrange the observations (values of the variate) in ascending or descending order of magnitude.

Step II : Determine the total number of observations, say, n.

Step III : If n is odd, then

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

If n is even, then

Median

$$= \frac{\text{Value of } \left(\frac{n}{2}\right)^{th} \text{ observation} + \text{Value of } \left(\frac{n}{2}+1\right)^{th} \text{ observation}}{2}$$

❖ The median can be calculated graphically, while mean cannot be.

❖ The sum of the absolute deviations taken from the median is less than the sum of the absolute deviations taken from any other observation in the data.

❖ Median is not affected by extreme values.

❖ EXAMPLES ❖

Ex.60 Find the median of the following data :

25, 34, 31, 23, 22, 26, 35, 28, 20, 32

Sol. Arranging the data in ascending order, we get20, 22, 23, 25, 26, 28, 31, 32, 34, 35

Here, the number of observations n = 10 (even).

$$\therefore \text{Median} = \frac{\text{Value of } \left(\frac{10}{2}\right)^{th} \text{ observation} + \text{Value of } \left(\frac{10}{2}+1\right)^{th} \text{ observation}}{2}$$

$$\Rightarrow \text{Median} = \frac{\text{Value of } 5^{th} \text{ observation + value of } 6^{th} \text{ observation}}{2}$$

$$\therefore \quad \text{Median} = \frac{26+28}{2} = 27$$

Hence, median of the given data is 27.

Ex.61 Find the median of the following values :

37, 31, 42, 43, 46, 25, 39, 45, 32

Sol. Arranging the data in ascending order, we have 25, 31, 32, 37, 39, 42, 43, 45, 46

Here, the number of observations n = 9 (odd)

$$\therefore \text{ Median} = \text{Value of } \left(\frac{9+1}{2}\right)^{th} \text{ observation}$$

= Value of 5th observation = 39.

Ex.62 The median of the observations 11, 12, 14, 18, x + 2, x + 4, 30, 32, 35, 41 arranged in ascending order is 24. Find the value of x.

Sol. Here, the number of observations n = 10. Since n is even, therefore

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

$$\Rightarrow \quad 24 = \frac{5^{th} \text{ observation} + 6^{th} \text{ observation}}{2}$$

$$\Rightarrow \quad 24 = \frac{(x+2)+(x+4)}{2}$$

$$\Rightarrow \quad 24 = \frac{2x+6}{2}$$

$\Rightarrow \quad 24 = x + 3 \quad \Rightarrow \quad x = 21.$

Hence, x = 21.

Ex.63 Find the median of the following data : 19, 25, 59, 48, 35, 31, 30, 32, 51. If 25 is replaced by 52, what will be the new median.

Sol. Arranging the given data in ascending order, we have 19, 25, 30, 31, 32, 35, 48, 51, 59

Here, the number of observations n = 9 (odd)

Since the number of observations is odd. Therefore.

$$\text{Median} = \text{Value of } \left(\frac{9+1}{2}\right) \text{ the observations}$$

$\Rightarrow$ Median = value of 5^{th} observation = 32.

Hence, Median = 32

If 25 is replaced by 52, then the new observations arranged in ascending order are :

19, 30, 31, 32, 35, 48, 51, 52, 59

$\therefore$ New median = Value of 5^{th} observation = 35.

Ex.64 Find the median of the following data

(i) 17, 27, 37, 13, 18, 25, 32, 34, 23

(ii) 24, 37, 19, 41, 28, 32, 29, 31, 33, 21

Sol. (i) The scores when arranged in ascending order are

13, 17, 18, 23, 25, 27, 32, 34, 37

Here, the number of scores n = 9 (odd)

$$\therefore \quad \text{Median} = t_{\frac{9+1}{2}} = t_5 = 25$$

(ii) The scores when arraged in ascending order are

19, 21, 24, 28, 29, 31, 33, 34, 37, 41.

Total number of scores = 10, which is even. So there will be two middle-terms which are $t_5 = 29$ and $t_6 = 31$.

$$\therefore \quad \text{Median} = \frac{t_5 + t_6}{2} = \frac{29+31}{2} = 30$$

Ex.65 Calculate the median for the following distribution

Weight (in kg)	Number of sudent
46	3
47	2
48	4
49	6
50	5
51	2
52	1

Sol. The cumulative frequency table is constructed as shown below :

Weights x_i	Number of students f_i	Cumulative frequency
46	3	3
47	2	5
48	4	9
49	6	15
50	5	20
51	2	22
52	1	23

Here, n = 23, which is odd

Median = $t_{\frac{23+1}{2}} = t_{12}$

= 49

(i.e. weight of the 12th student when the weights have been arranged in order)

Ex.66 Find the median of the following data :

(i) 8, 10, 5, 7, 12, 15, 11

(ii) 12, 14, 10, 7, 15, 16

Sol. (i) 8, 10, 5, 7, 12, 15, 11

These numbers are arranged in an order

5, 7, 8, 10, 11, 12, 15

The number of observations = 7 (odd)

$\Rightarrow$ Median = $\frac{7+1}{2}$ = 4th term

$\Rightarrow$ Median = 10

(ii) 12, 14, 10, 7, 15, 16

These numbers are arranged in an order

7, 10, 12, 14, 15, 16

The number of observations = 6 (even)

The medians will be mean of $\frac{6}{2}$ = 3rd and 4th terms i.e., 12 and 14

$\Rightarrow$ The median = $\frac{12+14}{2}$ = 13

Ex.67 The following data have been arranged in desending orders of magnitude 75, 70, 68, x + 2, x – 2, 50, 45, 40

If the median of the data is 60, find the value of x.

Sol. The number of observations are 8, the median will be the average of 4th and 5th number

$\Rightarrow$ Median = $\frac{(x+2)+(x-2)}{2}$

$\Rightarrow$ 60 = $\frac{2x}{2}$

$\Rightarrow$ x = 60

Ex.68 Find the median of 6, 8, 9, 10, 11, 12 and 13.

Sol. Total number of terms = 7

The middle terms = $\frac{1}{2}(7 + 1)$ = 4th

Median = Value of the 4th term = 10.

Hence, the median of the given series is 10.

Ex.69 Find the median of 21, 22, 23, 24, 25, 26, 27 and 28.

Sol. Total number of terms = 8

Median

= Value of $\frac{1}{2}\left[\frac{8}{2}\text{th term} + \left(\frac{8}{2}+1\right)\text{th term}\right]$

= Value of $\frac{1}{2}$[4th term + 5th term]

= $\frac{1}{2}[24 + 25] = \frac{49}{2}$ = 24.5

➢ MEDIAN OF GROUPED FREQUENCY DISTRIBUTION

$$\text{Median} = \ell + \frac{\frac{N}{2} - C}{f} \times h$$

where,

ℓ = lower limit of median class interval

C = cumulative frequency preceding to the median class frequency

f = frequency of the class interval to which median belongs

h = width of the class interval

N = $f_1 + f_2 + f_3 + ... + f_n$.

➢ WORKING RULE TO FIND MEDIAN

Step 1: Prepare a table containing less than type cumulative frequency with the help of given frequencies.

Step 2 : Find out the cumulative frequency to which $\frac{N}{2}$ belongs. Class-interval of this cumulative frequency is the median class-interval.

Step 3 : Find out the frequency f and lower limit ℓ of this median class.

Step 4 : Find the width h of the median class interval

Step 5 : Find the cumulative frequency C of the class preceding the median class.

Step 6 : Apply the formula,

$$\text{Median} = \ell + \left(\frac{\frac{N}{2} - C}{f}\right) h \text{ to find the median.}$$

❖ EXAMPLES ❖

Ex.70 Find the median of the followng distribution :

Wages (in Rs)	No. of labourers
200 – 300	3
300 – 400	5
400 – 500	20
500 – 600	10
600 – 700	6

Sol. We have,

Wages (in Rs)	No. of labours	Less than type cumulative frequency
200 – 300	3	3
300 – 400	5	8 = C
400 – 500	20 = f	28
500 – 600	10	38
600 – 700	6	44

Here, the median class is 400 – 500 as $\frac{44}{2}$ i.e. 22 belongs to the cumulative frequency of this class interval.

Lower limit of the median class = ℓ = 400

width of the class interval = h = 100

Cumulative frequency preceding median class frequency = C = 8

Frequency of Median class = f =20

$$\text{Median} = \ell + h\left(\frac{\frac{N}{2} - C}{f}\right) = 400 + 100\left(\frac{\frac{44}{2} - 8}{20}\right)$$

$$= 400 + 100\left(\frac{22-8}{20}\right) = 400 + 100\left(\frac{14}{20}\right)$$

$$= 400 + 70 = 470$$

Hence, the median of the given frequency distribution is 470.

Ex.71 Find the median for the following :

Class-Interval	0–8	8–16	16–24	24–32	32–40	40–48
Frequency	8	10	16	24	15	7

Sol.

Class interval	Frequency	Less than type cumulative frequency
0 – 8	8	8
8 – 16	10	18
16 – 24	16	34 = C
24 – 32	24 = f	58
32 – 40	15	73
40 – 48	7	80

Since $\frac{80}{2}$ = 40 lies in the cumulative frequency of the class interval 24 – 32, so 24 – 32 belongs to the median class interval.

Lower limit of median class interval = ℓ = 24.

Width of the class interval = h = 8

Total frequency = N = 80

Cumulative frequency preceding median class frequency = C = 34

Frequency of median class = f = 24

$$\text{Median} = \ell + \left(\frac{\frac{N}{2} - C}{f}\right) h$$

$$= 24 + \left(\frac{\frac{80}{2} - 34}{24}\right) 8 = 24 + \left(\frac{40-34}{24}\right) 8$$

$$= 24 + 2 = 26$$

Hence, the median of the given frequency distribution = 26.

Ex.72 The following table shows the weekly drawn by number of workers in a factory :

Weekly Wages (in Rs.)	0–100	100–200	200–300	300–400
No. of workers	40	39	34	30

Find the median income of the workers.

Sol.

Weekly Wages (in Rs.)	No. of workers	Less than type cumulative frequency
0–100	40	40
100–200	39	79 = C
200–300	34 = f	113
300–400	30	143
400 – 500	45	188

Since $\frac{188}{2}$ = 94 belongs to the cumulative frequency of the median class interval (200 – 300), so 200 – 300 is the median class.

Lower limit of the median class interval = ℓ = 200.

Width of the class interval = h = 100

Total frequency = N = 188

Frequency of the median class = f = 34

Cumulative frequency preceding median class

= C = 79

$$\text{Median} = \ell + \left(\frac{\frac{N}{2} - C}{f}\right) h = 200 + \left(\frac{\frac{188}{2} - 79}{34}\right) 100$$

$$= 200 + \left(\frac{94-79}{34}\right) 100 = 200 + 44.117$$

$$= 244.117$$

Hence, the median of the given frequency distribution = 244.12.

Ex.73 The following frequency distribution gives the monthly consumption of electricity of 68 consumers of a locality. Find the median and mode of the data and compare them.

Monthly consumption	Number of consumers
65 – 85	4
85 – 105	5
105 – 125	13
125 – 145	20
145 – 165	14
165 – 185	8
185 – 205	4

Sol.

Monthly consumption	Number of consumers	Less than type cumulative frequency
65 – 85	4	4
85 – 105	5	9
105 – 125	13	22 =C
125 – 145	20 = f	42
145 – 165	14	56
165 – 185	8	64
185 – 205	4	68

Since $\frac{68}{2}$ belongs to the cumulative frequency (42) of the class interval 125 – 145, therefore 125 – 145 is the median class interval

Lower limit of the median class interval = ℓ = 125.

Width of the class interval = h = 20

Total frequency = N = 68

Cumulative frequency preceding median class frequency = C = 22

Frequency of the median class = f = 20

$$\text{Median} = \ell + \left(\frac{\frac{N}{2} - C}{f}\right) h = 125 + \left(\frac{\frac{68}{2} - 22}{20}\right) 20$$

$$= 125 + \frac{12 \times 20}{20} = 125 + 12 = 137$$

The frequency of class 125 – 145 is maximum i.e., 20, this is the modal class,

$x_k = 125$, $f_k = 20$, $f_{k-1} = 13$, $f_{k+1} = 14$, $h = 20$

$$\text{Mode} = x_k + \frac{f - f_{k-1}}{2f - f_{k-1} - f_{k+1}}$$

$$= 125 + \frac{20 - 13}{40 - 13 - 14} \times 20$$

$$= 125 + \frac{7}{40 - 27} \times 20 = 125 + \frac{7}{13} \times 20$$

$$= 125 + 10.77 = 135.77$$

Ex.74 Compute the median from the marks obtained by the students of class X.

Marks	Number of Students
40 – 49	5
50 – 59	10
60 – 69	20
70 – 79	30
80 – 89	20
90 – 99	15

Sol. First we will form the less than type cumulative frequency distribution and we make the distribution continuous by subtracting 0.5 from the lower limits and adding 0.5 to the upper limits.

Marks	Number of students	Less than type cumulative frequency
39.5 – 49.5	5	5
49.5 – 59.5	10	15
59.5 – 69.5	20	35 = C
69.5 – 79.5	30 = f	65
79.5 – 89.5	20	85
89.5 – 99.5	15	100

Since $\frac{100}{2}$ belongs to the cumulative frequency (65) of the class interval 69.5 – 79.5, therefore 69.5 – 79.5 is the median class.

Lower limit of the median class = ℓ = 69.5.

Width of the class interval = h = 10

Total frequency = N = 100

Cumulative frequency preceding median class frequency = C = 35

Frequency of median class = f = 30

$$\text{Median} = \ell + h\left(\frac{\frac{N}{2} - C}{f}\right) = 69.5 + 10\left(\frac{\frac{100}{2} - 35}{30}\right)$$

$$= 69.5 + 10\left(\frac{50 - 35}{30}\right) = 69.5 + \frac{10 \times 15}{30}$$

$$= 69.5 + 5 = 74.5$$

Hence, the median of given frequency distribution is 74.50.

Ex.75 An incomplete frequency distribution is given as follows :

Variable	Frequency
10 – 20	12
20 – 30	30
30 –40	?
40 – 50	65
50 – 60	?
60 – 70	25
70 – 80	18
Total	229

Given that the median value is 46, determine the missing frequencies using the median formula.

Sol. Let the frequency of the class 30 – 40 be f_1 and that of 50 – 60 be f_2.

Variable	Frequency	Less than type cumulative frequency
10 – 20	12	12
20 – 30	30	12 + 30 = 42
30 –40	f_1	$42 + f_1$
40 – 50	65 = f	$107 + f_1$
50 – 60	f_2	$107 + f_1 + f_2$
60 – 70	25	$132 + f_1 + f_2$
70 – 80	18	$150 + f_1 + f_2 = 229$

From the last item of the third column, we have

$$150 + f_1 + f_2 = 229$$

$$\Rightarrow f_1 + f_2 = 229 - 150$$

$$\Rightarrow f_1 + f_2 = 79$$

Since, the median is given to be 46, the class 40 – 50 is median class

Therefore, ℓ = 40, C = 42 + f_1, N = 299, h = 10

Median = 46, f = 65

$$\text{Median} = \ell + \frac{h\left(\frac{N}{2} - C\right)}{f} = 46$$

$$46 = 40 + 10\frac{\left(\frac{229}{2} - 42 - f_1\right)}{65}$$

$$\Rightarrow 6 = \frac{10}{65}\left(\frac{229}{2} - 42 - f_1\right)$$

$$\Rightarrow 6 = \frac{2}{13}\left(\frac{229 - 84 - 2f_1}{2}\right)$$

$$\Rightarrow 78 = 229 - 84 - 2f_1 \Rightarrow 2f_1 = 229 - 84 - 78$$

$$\Rightarrow 2f_1 = 67 \Rightarrow f_1 = \frac{67}{2} = 33.5 = 34$$

Putting the value of f_1 in (1), we have

$$34 + f_2 = 79$$

$$\Rightarrow f_2 = 45$$

Hence, f_1 = 34 and f_2 = 45.

Ex.76 Recast the following cumulative table in the form of an ordinary frequency distribution and determine the median.

No. of days absent	No. of students
less than 5	29
less than 10	224
less than 15	465
less than 20	582
less than 25	634
less than 30	644
less than 35	650
less than 40	653
less than 45	655

Sol.

No. of days	No. of students	No. of days absent	No. of students	Less than type cumulative frequency
less than 5	29	0 – 5	29	29
less than 10	224	5 – 10	195	224 = C
less than 15	465	10 – 15	241 = f	465
less than 20	582	15 – 20	117	582
less than 25	634	20 – 25	52	634
less than 30	644	25 – 30	10	644
less than 35	650	30 – 35	6	650
less than 40	653	35 – 40	3	653
less than 45	655	40 – 45	2	655

Since $\frac{655}{2}$ belongs to the cumulative frequency (465) of the class interval 10 – 15, therefore 10 – 15 is the median class.

Lower limit of the median class = ℓ = 10.

Width of the class interval = h = 5

Total frequency = N = 655

Cumulative frequency preceding median class frequency = C = 224

Frequency of median class = f = 241

$$\text{Median} = \ell + h\left(\frac{\frac{N}{2} - C}{f}\right) = 10 + 5\left(\frac{\frac{655}{2} - 224}{241}\right)$$

$$= 10 + 5\left(\frac{327.5 \times 224}{241}\right) = 10 + \frac{5 \times 103.5}{241}$$

= 10 + 2.147 = 12.147

Hence, the median of given frequency distribution is 12.147.

➤ MODE

Mode is also known as norm.

◈ Mode is the value which occurs most frequently in a set of observations and around which the other items of the set cluster density.

Algorithm

Step I : Obtain the set of observations.

Step II : Count the number of times the various values repeat themselves. In other words, prepare the frequency distribution.

Step III : Find the value which occurs the maximum number of times i.e. obtain the value which has the maximum frequency.

StepIV : The value obtained in step III is the mode.

Ex.77 Find the mode from the following data :

110, 120, 130, 120, 110, 140, 130, 120, 140, 120.

Sol. Arranging the data in the form of a frequency table, we have

Value	Tally bars	Frequency
110	||	2
120	||||	4
130	||	2
140	||	2

Since the value 120 occurs maximum number of times i.e. 4. Hence, the modal value is 120.

Ex.78 Find the mode for the following series :

2.5, 2.3, 2.2, 2.2, 2.4, 2.7, 2.7, 2.5, 2.3, 2.2, 2.6, 2.2

Sol. Arranging the data in the form of a frequency table, we have

Value	Tally bars	Frequency
2.2	||||	4
2.3	||	2
2.4	|	1
2.5	||	2
2.6	|	1
2.7	||	2

We see that the value 2.2 has the maximum frequency i.e. 4.

So, 2.2 is the mode for the given series.

Ex.79 Compute mode for the following data

7, 7, 8, 8, 8, 9, 9, 10, 10, 10, 11, 11, 12, 13, 13

Sol. Here, both the scores 8 and 10 occurs thrice (maximum number of times). So, we apply the empirical formula.

Here,

$$\text{mean} = \frac{7\times2+8\times3+9\times2+10\times3+11\times2+12+13\times2}{2+3+2+3+2+1+2}$$

$$= \frac{14+24+18+30+22+12+26}{15} = \frac{146}{15} = 9.73$$

No. of scores = 15 (odd)

$\therefore$ Median = $t_{\frac{15+1}{2}} = t_8 = 10$

$\therefore$ Mode = 3 median – 2 mean

= 3 × 10 – 2 × 9.73 = 30 – 19.46 = 10.54

Ex.80 Find the mode of the following data :

6, 4, 7, 4, 5, 8, 4, 5, 5, 3, 2, 5

Sol. We write the data in tabular form :

x	f
2	1
3	1
4	3
5	4
6	1
7	1
8	1

We observe that 5 has maximum frequency which is 4

$\Rightarrow$ Mode = 5

Ex.81 The following table gives the weights of 40 men. Calculate mode.

Weights (in kg)	Number of men
54	6
72	6
80	1
64	2
62	6
60	5
58	5
56	4
63	5

Sol. Here, each of the scores 54, 72 and 62 occurs maximum number of times (six times). So we apply the empirical formula.

We construct the following table :

Weights x	No. of men f	Cumulative frequency	Product f.x
54	6	6	324
56	4	10	224
58	5	15	290
60	5	20	300
62	6	26	372
63	5	31	315
64	2	33	128
72	6	39	432
80	1	40	80
Total	40		2465

$$\text{Mean} = \frac{\Sigma f.x}{\Sigma f} = \frac{2465}{40} = 61.625$$

Here, No. of scores = 40 (even)

$$\text{Median} = \frac{t_{20} + t_{21}}{2} = \frac{60+62}{2} = 61$$

$\therefore$ Mode = 3 median – 2 mean

= 3 × 61 – 2 × 61.625

= 183 – 123.25 = 59.75

Thus, modal weight = 59.75 kg

➤ RELATIVE CHARACTERISTICS OF MEAN, MEDIAN AND MODE

(i) Mean is usually understood as arithmetic average, since its basic definition is given in arithmetical terms.

(ii) Mean is regarded as the true representative of the whole population since in its calculation all the values are taken into consideration. It does not necessarily assume a value that is the same as one of theoriginal ones (which other averages often do)

(iii) Mean is suitable for sets of data which do not have extreme values. In other cases, median is the appropriate measure of location.

(iv) Mode is the most useful measure of location when the most common or most popular item is required.

◈ Merits of Mode

(i) Mode is readily comprehensively and easy to calculate. It can be located in some cases morely by inspection.

(ii) Mode is not all affected by extreme values.

(iii) Mode can be coneniently even class interval of unequal magnitude.

◈ Demerits of Mode

(i) Mode is ill defined. In some cases we may come across two modes.

(ii) It is not based upon all the observations.

(iii) No further mathematical treatment is possible in case of mode.

(iv) Mode is affected to a greater extant by flutuations of sampling.

❖ **Relationship among Mean, Median and Mode :**

Following are the relations,

Mode = 3 Median – 2 mean

Median = Mode + $\frac{2}{3}$ (Mean – Mode)

Mean = Mode + $\frac{3}{2}$ (Median – Mode)

❖ **EXAMPLES** ❖

Ex.82 If mean = 60 and median = 50, find mode.

Sol. We have,

Mean = 60, Median = 50

Mode = 3 Median – 2 Mean

= 3 (50) – 2 (60) = 30

Ex.83 If mode = 70 and mean = 100, find median.

Sol. We have, Mode = 70, Mean = 100

Median = Mode + $\frac{2}{3}$ (Mean – Mode)

= 70 + $\frac{2}{3}$ (100 – 70)

= 70 + 20

= 90

Ex.84 If mode = 400 and median = 500, find mean.

Sol. Mean = Mode + $\frac{3}{2}$ (Median – Mode)

= 400 + $\frac{3}{2}$ (500 – 400)

= 400 + $\frac{3}{2}$ (100)

= 400 + 150

= 550

➢ CUMULATIVE FREQUENCY CURVE OR THE OGIVE

First we prepare the cumulative frequency table, then the cumulative frequencies are plotted against the upper or lower limits of the corresponding class intervals. By joining the points the curve so obtained is called a cumulative frequency curve or ogive.

There are two types of ogives :

1. **Less than ogive :** Plot the points with the upper limits of the class as abscissae and the corresponding less than cumulative frequencies as ordinates. The points are joined by free hand smooth curve to give less than cumulative frequency curve or the less than Ogive. It is a rising curve.

2. **Greater than ogive :** Plot the points with the lower limits of the classes as abscissa and the corresponding Greater than cumulative frequencies as ordinates. Join the points by a free hand smooth curve to get the "More than Ogive". It is a falling curve.

When the points obtained are joined by straight lines, the picture obtained is called cumulative frequency polygon.

❖ **EXAMPLES** ❖

Ex.85 Draw a less than ogive for the following frequency distribution :

I.Q.	Frequency
60 – 70	2
70 – 80	5
80 –90	12
90 – 100	31
100 – 110	39
110 – 120	10
120 – 130	4

Find the median from the curve.

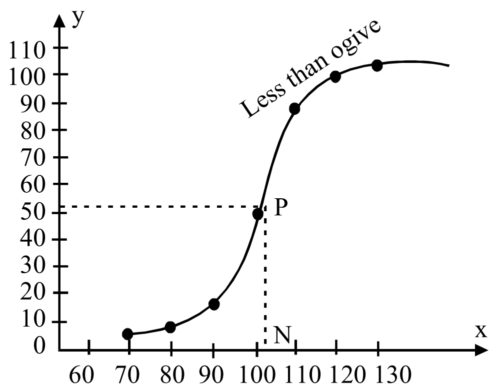
Sol. Let us prepare following table showing the cumulative frequencies more than the upper limit.

Class interval	Frequency (f)	Cumulative frequency
60 – 70	2	2
70 – 80	5	2 + 5 = 7
80 –90	12	2 + 5 + 12 = 19
90 – 100	31	2 + 5 + 12 + 31 = 50
100 – 110	39	2 + 5 + 12 + 31 + 39 = 89
110 – 120	10	2 + 5 + 12 + 31 + 39 + 10 = 99
120 – 130	4	2 + 5 + 12 + 31 + 39 + 10 + 4 = 103

Less than ogive :

I.Q. is taken on the x-axis. Number of students are marked on y-axis.

Points (70, 2), (80, 7), (90, 19), (100, 50), (110, 89), (120, 99), (130, 103), are plotted on graph paper and these points are joined by free hand. The curve obtained is less than ogive.



The value $\frac{N}{2}$ = 51.5 is marked on y-axis and from this point a line parallel to x-axis is drawn. This line meets the curve at a point P. From P draw a perpendicular PN to meet x-axis at N. N represents the median.

Here median is 100.5.

Hence, the median of given frequency distribution is 100.5

Ex.86 The following table shows the daily sales of 230 footpath sellers of Chandni Chowk.

Sales in Rs.	No. of sellers
0 – 500	12
500 – 1000	18
1000 – 1500	35
1500 – 2000	42
2000 – 2500	50
2500 – 3000	45
3000 – 3500	20
3500 – 4000	8

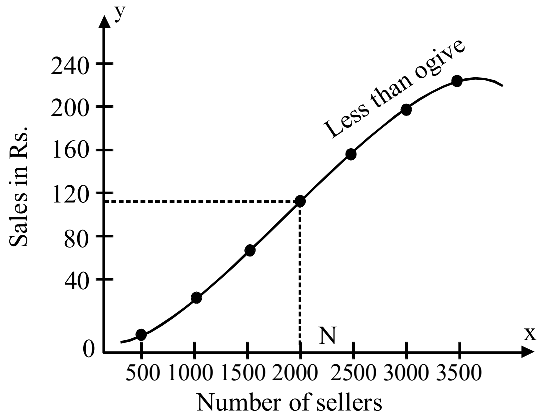
Locate the median of the above data using only the less than type ogive.

Sol. To draw ogive, we need to have a cumulative frequency distribution.

Sales in Rs.	No. of sellers	Less than type cumulative frequency
0 – 500	12	12
500 – 1000	18	30
1000 – 1500	35	65
1500 – 2000	42	107
2000 – 2500	50	157
2500 – 3000	45	202
3000 – 3500	20	222
3500 – 4000	8	230

Less than ogive :

Seles in Rs. are taken on the y-axis and number of sellers are taken on x-axis. For drawing less than ogive, points (500, 12), (1000, 30), (1500, 65), (2000, 107), (2500, 157), (3000, 202), (3500, 222), (4000, 230) are plotted on graph paper and these are joined free hand to obtain the less than ogive.



The value N/2 = 115 is marked on y-axis and a line parallel to x-axis is drawn. This line meets the curve at a point P. From P draw a perpendicular PN to meet x-axis at median. Median = 2000.

Hence, the median of given frequency distribution is 2000.

Ex.87 Draw the two ogives for the following frequency distribution of the weekly wages of (less than and more than) number of workers.

Weekly wages	Number of workers
0 – 20	41
20 – 40	51
40 – 60	64
60 – 80	38
80 – 100	7

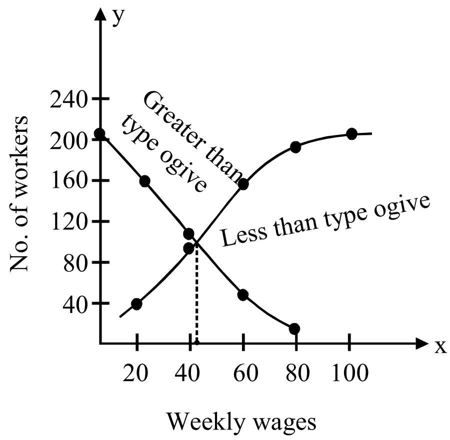
Hence find the value of median.

Sol.

Weekly wages	Number of workers	Cumulative (less than)	Frequency (More than)
0 – 20	41	41	201
20 – 40	51	92	160
40 – 60	64	156	109
60 – 80	38	194	45
80 – 100	7	201	7

Less than curve :

Upper limits of class intervals are marked on the x-axis and less than type cumulative frequencies are taken on y-axis. For drawing less than type curve, points (20, 41), (40, 92), (60, 156), (80, 194), (100, 201) are plotted on the graph paper and these are joined by free hand to obtain the less than ogive.



Greater than ogive

Lower limits of class interval are marked on x-axis and greater than type cumulative frequencies are taken on y-axis. For drawing greater than type curve, points (0, 201), (20, 160), (40, 109), (60, 45) and (80, 7) are plotted on the graph paper and these are joined by free hand to obtain the greater than type ogive. From the point of intersection of these curves a perpendicular line on x-axis is drawn. The point at which this line meets x-axis determines the median. Here the median is 42.652.

Ex.88 Following table gives the cumulative frequency of the age of a group of 199 teachers.

Draw the less than ogive and greater than ogive and find the median.

Age in years	Cum. Frequency
20 – 25	21
25 – 30	40
30 – 35	90
35 – 40	130
40 – 45	146
45 – 50	166
50 – 55	176
55 – 60	186
60 – 65	195
65 – 70	199

Sol.

Age in years	Less than cumulative frequency	Frequency	Greater than type
20 – 25	21	21	199
25 – 30	40	19	178
30 – 35	90	50	159
35 – 40	130	40	109
40 – 45	146	16	69
45 – 50	166	20	53
50 – 55	176	10	33
55 – 60	186	10	23
60 – 65	195	9	13
65 – 70	199	4	4

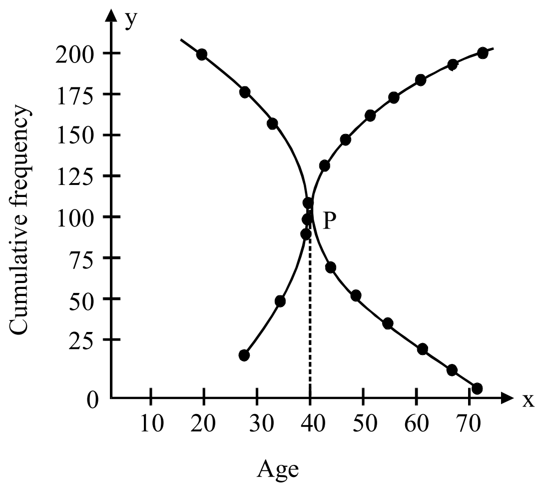
Find out the frequencies by subtracting previous frequency from the next frequency to get simple frequency. Now we can prepare the greater than type frequency. Ages are taken on x-axis and number of teachers on y-axis.

Less than ogive :

Plot the points (25, 21), (30, 40), (35, 90), (40, 130), (45, 146), (50, 166), (55, 176), (60, 186), (65, 195), (70, 199) on graph paper. Join these points free hand to get less than ogive.

Greater than ogive :

Plot the points (20, 199), (25, 178), (30, 159), (35, 109), (40, 69), (45, 53), (50, 33), (55, 23), (60, 13), (65, 4) on graph paper. Join these points freehand to get greater than ogive. Median is the point of intersection of these two curves.



Here median is 37.375.

➢ LESS THAN METHOD

To construct a cumulative frequency polygon and an ogive by less than method, we use the following algorithm.

Algorithm

Step 1 :

Start with the upper limits of class intervals and add class frequencies to obtain the cumulative frequency distribution.

Step 2 :

Mark upper class limits along X-axis on a suitable scale.

Step 3 :

Mark cumulative frequencies along Y-axis on a suitable scale.

Step 4 :

Plot the points (x_i, f_i) where x_i is the upper limit of a class and f_i is corresponding cumulative frequency.

Step 5 :

Join the points obtained in step 4 by a free hand smooth curve to get the ogive and to get the cumulative frequency polygon join the points obtained in step 4 by line segments.

➢ MORE THAN METHOD

To construct a cumulative frequency polygon and an ogive by more than method, we use the following algorithm.

Algorithm

Step 1 :

Start with the lower limits of the class intervals and from the total frequencysubtract the frequency of each class to obtain the cumulative frequency distribution.

Step 2 :

Mark the lower class limits along X-axis on a sutiable scale.

Step 3 :

Mark the cumulative frequencies along Y-axis on a suitable scale.

Step 4 :

Plot the points (x_i, f_i) where x_i is the lower limit of a class and f_i is corresponding cumulative frequency.

Step 5 :

Join the points obtained in step 4 by a free hand smooth curve to get the ogive and to get the cumulative frequency polygon join these points by line segments

Ex.89 Following is the age distribution of a group of students. Draw the cumulative frequency polygon, cumulative frequency curve (less than type) and hence obtain the median value.

Age	Frequency
5 – 6	40
6 – 7	56
7 – 8	60
8 – 9	66
9 – 10	84
10 – 11	96
11 – 12	92
12 – 13	80
13 – 14	64
14 – 15	44
15 – 16	20
16 – 17	8

Sol. We first prepare the cumulative frequency table by less then method as given below :

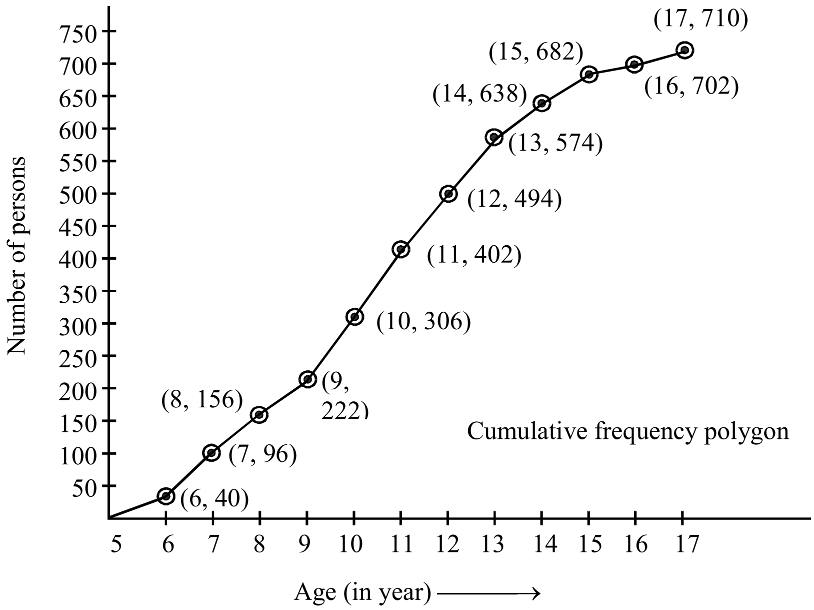
Age	Frequency	Age less than	Cumulative frequency
5 – 6	40	6	40
6 – 7	56	7	96
7 – 8	60	8	156
8 – 9	66	9	222
9 – 10	84	10	306
10 – 11	96	11	402
11 – 12	92	12	494
12 – 13	80	13	574
13 – 14	64	14	638
14 – 15	44	15	682
15 – 16	20	16	702
16 – 17	8	17	710

Other than the given class intervals, we assume a class 4-5 before the first class interval 5-6 with zero frequency.

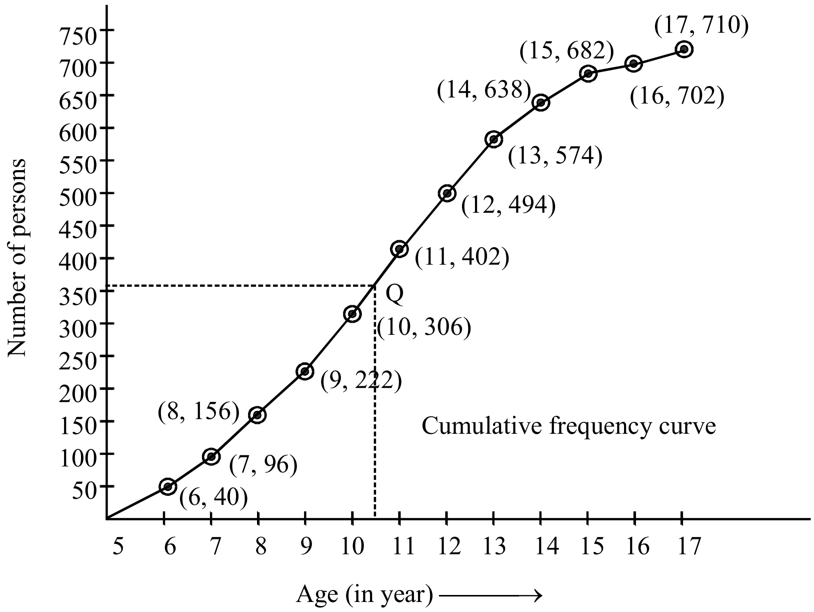
Now, we mark the upper class limits (including the imagined class) along X-axis on a suitable scale and the cumulative frequencies along Y-axis on a suitable scale.

Thus, we plot the points (5, 0), (6, 40), (7, 96), (8, 156), (9, 222), (10, 306), (11, 402), (12, 494), (13, 574), (14, 638), (15, 682), (16, 702) and (17, 710).

These points are marked and joined by line segments to obtain the cumulative frequency polygon shown in Fig.



In order to obtain the cumulative frequency curve, we draw a smooth curve passing through the points discussed above. The graph (fig) shows the total number of students as 710. The median is the age corresponding to $\frac{N}{2} = \frac{710}{2}$ = 355 students. In order to find the median, we first located the point corresponding to 355th student on Y-axis. Let the point be P. From this point draw a line parallel to the X-axis cutting the curve at Q. From this point Q draw a line parallel to Y-axis and meeting X-axis at the point M. The x-coordinate of M is 10.5 (See Fig.). Hence, median is 10.5.



Ex.90 The following observations relate to the height of a group of persons. Draw the two type of cumulative frequency polygons and cumulative frequency curves and determine the median.

Height in cms	140–143	143–146	146–149	149–152	152–155	155–158	158–161
Frequency	3	9	26	31	45	64	78
Height in cms	161–164	164–167	167–170	170–173	173–176	176–179	179–182
Frequency	85	96	72	60	43	20	6

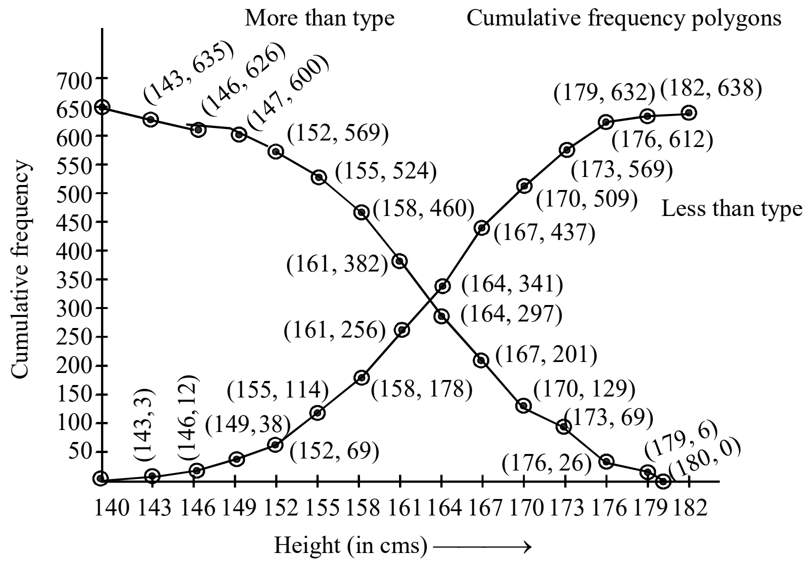
Sol. Less than method : We first prepare the cumulative frequency table by less than method as given below :

Height in cms	Frequency	Height less than	Frequency
140–143	3	143	3
143–146	9	146	12
146–149	26	149	38
149–152	31	152	69
152–155	45	155	114
155–158	64	158	178
158–161	78	161	256
161–164	85	164	341
164–167	96	167	437
167–170	72	170	509
170–173	60	173	569
173–176	43	176	612
176–179	20	179	632
179–182	6	182	638

Other than the given class intervals, we assume a class interval 137-140 prior to the first class interval 140-143 with zero frequency.

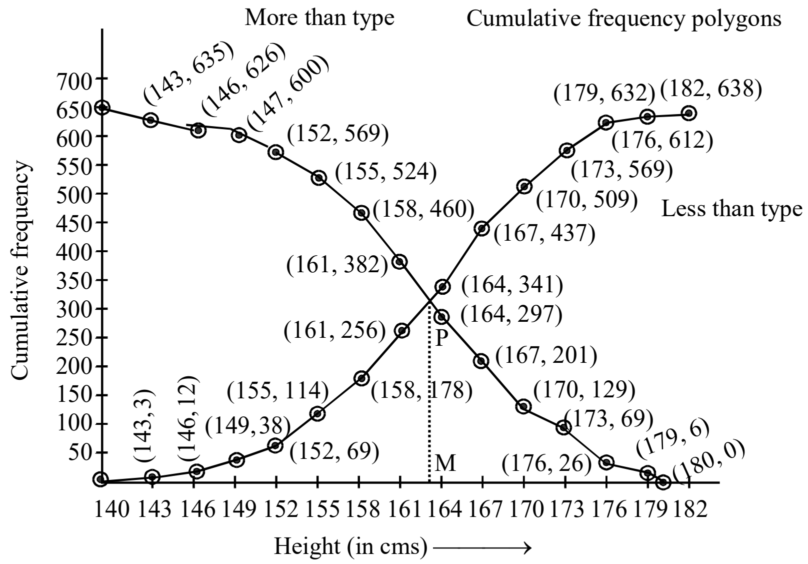
Now, we mark the upper class limits on X-axis and cumulative frequency along Y-axis on a suitable scale.

We plot the points (140, 0), (143, 3), (146, 12), (149, 38), (152, 69), (155, 114), (158, 178), (161, 256), (164, 341), (167, 437), (170, 509), (173, 569), (176, 612),.(179, 632) and 182, 638).



Cumulative Frequency polygon of less than and more than type

These points are joined by line segments to obtain the cumulative frequency polygon as shown in fig. and by a free hand smooth curve to obtain an ogive by less than method as shown in fig.



Less than and more than type cumulative frequency curves

More than method : We prepare the cumulative frequency table by more than method as given below :

Height in cms	Frequency	Height more than	Cumulative frequency
140–143	3	140	638
143–146	9	143	635
146–149	26	146	626
149–152	31	149	600
152–155	45	152	569
155–158	64	155	524
158–161	78	158	460
161–164	85	161	382
164–167	96	164	297
167–170	72	167	201
170–173	60	170	129
173–176	43	173	69
176–179	20	176	26
179–182	6	179	6

Other than the given class intervals, we assume the class interval 182-185 with zero frequency.

Now, we mark the lower class limits on X-axis and the cumulative frequencies along Y-axis on suitable scales to plot the points (140, 638), (143, 635), (146, 626), (149, 600), (152, 569), (155, 524), (158, 460), (161, 382), (164, 297), (167, 201), (170, 129), (173, 69), (176, 26) and (179, 6). By joining these points by line segments, we obtain the more than type frequency polygon as shown in fig. By joining these points by a free hand curve, we obtain more than type cumulative frequency curve as points by a free hand curve, we obtain more than type cumulative frequency curves as shown in fig.

We find that the two types of cumulative frequency curves intersect at point P. From point P perpendicular PM is drawn on X-axis. The value of height corresponding to M is 163.2 cm. Hence, median is 163.2 cm.

Ex.91 If the heights of 5 persons are 144 cm, 153 cm, 150 cm, 158 cm and 155 cm respectively, then find the mean height.

Sol. Mean Height = $\frac{144+153+150+158+155}{5}$

$= \frac{760}{5} = 152$ cm.

Ex.92 Find the arithmetic mean of the following frequency distribution :

x : 4 7 10 13 16 19

f : 7 10 15 20 25 30

Sol. The given frequency distribution is -

x_i	f_i	$f_i x_i$
4	7	28
7	10	70
10	15	150
13	20	260
16	25	400
19	30	570

$\sum f_i = 107$

$\sum f_i\ x_i = 1478$

$\overline{x} = \frac{\sum f_i\ x_i}{\sum f_i} = \frac{1478}{107} = 13.81$

Ex.93 The mean income of a group of persons is Rs.400. Another group of persons has mean income Rs.480. If the mean income of all the persons in the two groups together is Rs.430, then find the ratio of the number of persons in the groups.

Sol. $\bar{x} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

$\because$ $\bar{x}_1 = 400$, $\bar{x}_2 = 480$, $\bar{x} = 430$

$\therefore$ $430 = \dfrac{n_1(400) + n_2(480)}{n_1 + n_2}$

$\Rightarrow$ $30n_1 = 50n_2$

$\Rightarrow$ $\dfrac{n_1}{n_2} = \dfrac{5}{3}$

Ex.94 The number of runs scored by 11 players of a cricket team of school are 5, 19, 42, 11, 50, 30, 21, 0, 52, 36, 27. Find the median.

Sol. Let us arrange the value in ascending order

0, 5, 11, 19, 21, 27, 30, 36, 42, 50, 52

$\therefore$ Median M = $\left(\dfrac{n+1}{2}\right)^{th}$ value

$= \left(\dfrac{11+1}{2}\right)^{th}$ value = 6^{th} value

Now 6^{th} value in data is 27

$\therefore$ Median = 27 runs.

Ex.95 Find the mode of the data 3, 2, 5, 2, 3, 5, 6, 6, 5, 3, 5, 2, 5.

Sol. Since 5 is repeated maximum number of times, therefore mode of the given data is 5.

Ex.96 If the value of mode and mean is 60 and 66 respectively, then find the value of median.

Sol. Mode = 3 Median – 2 mean

$\therefore$ Median = $\dfrac{1}{3}$ (mode + 2 mean)

$= \dfrac{1}{3}(60 + 2 \times 66) = 64$

Ex.97 Mean of 25 observations was found to be 78.4. But later on it was found that 96 was misread 69. Find the correct mean.

Sol. Mean $\bar{x} = \dfrac{\sum x}{n}$

or $\sum x = n\bar{x}$

$\sum x = 25 \times 78.4 = 1960$

But this $\sum x$ is incorrect as 96 was misread as 69.

$\therefore$ correct $\sum x$ = 1960 + (96 – 69) = 1987

$\therefore$ correct mean = $\dfrac{1987}{25}$ = 79.47

$\therefore$

IMPORTANT POINTS TO BE REMEMBERED

1. Three measures of central value are :

(i) Mean (ii) Median and

(iii) Mode

2. Mean is computed by following methods :

(i) Direct method (ii) Short-cut Method

(iii) Step-deviation method

3. If a variate X takes values $x_1, x_2, ..., x_n$ with corresponding frequencies $f_1, f_2, ..., f_n$ respectively, then the arithmetic mean of these values is given by

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{n} f_i x_i \text{ , where } N = \sum_{i=1}^{n} f_i$$

Also, $\overline{X} = A + \frac{1}{N}\sum_{i=1}^{n} f_i d_i$, where $d_i = x_i - A$.

The number A is called the assumed mean.

If $u_i = \frac{x_i - A}{h}$, i = 1, 2, ... , n. Then,

$$\overline{X} = A + h\left\{\frac{1}{N}\sum_{i=1}^{n} f_i u_i\right\}$$

4. 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.

The median of a grouped or continuous frequency distribution may be computed by using the following formula :

$$\text{Median} = \ell + \frac{\frac{N}{2} - F}{f} \times h, \text{ where}$$

ℓ = lower limit of the median class

f = frequency of the median class

h = width of the median class

F = cumulative frequency of the class preceding the median class, and, $N = \sum_{I=1}^{n} f_I$

5. Mode is the value of the variable which has the maximum frequency. The mode of a continuous or grouped frequency distribution may be computed by using the following formula :

$$\text{Mode} = \ell + \frac{f - f_1}{2f - f_1 - f_2} \times h, \text{ where}$$

ℓ = lower limit of the modal class.

f = frequency of the modal class.

h = width of the modal class

f_1 = frequency of the class preceding the modal class.

f_2 = frequency of the class following the modal class.

6. Three measures of central value are connected by the following relation :

Mode = 3 median – 2 Mean

7. Ogive (s) can be used to find the median of a frequency distribution.