Statistics

15.01 Introduction :

The branch of science known as Statistics has been used in India form ancient times. Which is evident from the following examples.

In Mahabharata period, king Rituparna while going alone with King Nala for the Swaymbara of Damyanti, estimated accurately the number of leaves and fruits on the basis of sample of a tree. The description of administrative setup based on registration of births and deaths during the region of Chandragupta Maurya (324-300 B.C.) is found in Kautilya's Arthashtra. Similarly many examples of use of statistics in administsative setup and in wars are found. When we observe or notice in newspapers, channels of electronic media, magazines and communication, then we get the factual and comparative informations about the temperature of cities, position of rainfall.

In this chapter we shall try to study, in detail, about collection, classification, tabulation, mean, median and mode of such data. Situation of shares of different companies. We use the data in our whole life. It is very important for us that we know to produce meaningful information from the data.

The branch of mathematics in which we learn about these meaningful information is know as statistics.

15.02 Collection of data :

On the basis of method of collection, data can be divided into two categories :

(1) Primary data (2) Secondary data.

(1)   Primary data :

Data which are collected for the first time by the statical investigator or with the help of his workers is called primary data. e.g. The weight and height of the student in class 9th of school. We collect the primary data by following methods.

(i)   Direct personal investigation : In this method, the investigator himself contacts the person form whose he/she is to collect information.

(ii)   Indirect investigation : When the field of investigation is wide it is not possible

for the investigator to go and contact everyone personally. In such a situation he/she may try to collect the information in the following ways :

(a) **Schedule to be filled in by enumerators :** In this method, the inversigator prepares the schedules, gives it to the trained enumerators. Then the enumerator ask the questions to concerned informants and writes the answer in schedules.

(b) **Information through questionnaires to be filled in by informants :** In this method investigator prepares a questionnaire related to investigation and sends to the informants with the object to get information and assurance of information's secrecy.

(c) **Information through local sources or correspondents :** When collection of daily information is required form different places, then investigator appoints local persons or correspondents for reporting the informations.

(d) **Indirect investigation through experts :** In this method the information is not collected form the people directly involved in the investigation. Rather it is collected form people called witnesses who are indirectly involved. If the annual evaluation of the students is to be done without examination, then this can be done by the concerned teacher.

(2) **Secondary data :**

If this method the data are already collected by a person or a society and these may be in published or unpublished form. These are generally obtained form the following two sources :

(i) **Published sources :** Government, non-government and other investigators collect data on different subjects form time to time and get it published, there main sources are as follows :

(a) **International publication :** Organisations like Unite Nations Organisation (U.N.O.), International Labour Organisation (I.L.O.), International Monetary Fund (I.M.F.) collect the data related with them and get them published for their members.

(b) **Government publication :** A number of Ministries and Departments (The department of Statistics) of central and state government collect data concerning them and published them.

(c) **Semi Government publication :** Municipalities, Corporations, Zila Parishads, Village Panchayats publish data on Education, Electricity, Birth and death and revenue records form time to time.

Similarly Universities, Research Institutes and teachers organisations compile different types of data and get them published.

(ii) **Unpublished sources :** Sometimes data on important subjects is collected by government and non-government institutions but they remain unpublished. Such unpublished matter can be obtained form office files, records, registers or the diaries of research scholars.

Exercise 15.1

1. What are primary and secondary data ? Clarify the difference.
2. Describe the methods of collection of primary data.

15.03 Presentation of Data :

After collection of data, an investigator has to select the suitable method to present the data in a form which is meaningful, easy to understand and gives its main features at a glance.

Data collected in original form is called **'raw data',** e.g. marks obtained by 10 students in a test of science are :

| 62 | 75 | 65 | 40 | 35 | 70 | 25 | 20 | 36 | 55 |

Can you obtained maximum and minimum marks form this data? If we arrange the data in ascending order or descending order, then we can easily obtained maximum and minimum marks. Ascending order of data are follows :

| 20 | 25 | 35 | 36 | 40 | 55 | 62 | 65 | 70 | 75 |

Here, it is clear that minimum marks is 20 and maximum marks is 75. Difference of maximum and minimum marks of data is called **range.** Thus, here

$75 - 20 = 55$ is the range.

If data are large, then it should be written in tabular form rather than in ascending or descending order.

Presentation of Data in Tabular Form :

In a school, following are the marks obtained by 20 students of class X in half yearly examination, out of 10 marks :

| 5 | 4 | 3 | 4 | 5 | 8 | 4 | 3 | 8 | 2 |
| 4 | 5 | 4 | 3 | 8 | 5 | 3 | 4 | 2 | 8 |

By seeing this raw data, it is difficult to estimate the level of class, but if we arrange

the data in tabular form, then this task will became very easy. *The number of times an observation occurs in the given data, is called the frequency of observation.* It is denoted as "f".

Following is the method to prepare a table from 06 the above data :

(i) In the first column of the table, we write all the marks from lowest to highest.

(ii) We now look at the first value in the given raw data and put vertical line known as tally mark, in the second column opposite to it.

(iii) When four lines are made opposite any mark, don't make the fifth line in the same way but make a line across the first four ($\cancel{||||}$)

(iv) If a number occurs six times, then draw again a verticle line and repeat the process. ($\cancel{||||}|$)

(v) After completing the entries of second column, count the number of tally marks against each observation and write in the third column.

Table

Marks	Tally Marks	Frequency (f)
2	||	2
3	||||	4
4	$\cancel{||||}$|	6
5	||||	4
8	||||	4

$$\overline{\Sigma f = 20}$$

This type of frequency distribution is called **ungrouped frequency distribution table.**

If number of observations is very large, then we divide the data in groups. These groups are called **class** and their measure is known as **class-interval** or **class size** or **class width.** Lower number of each class is known as **lower class limit** and maximum number is **upper class limit**. In next example we will understand the presentation in tabular form of data.

During Van Mahotsava 50 plants had planted in each school out of 30 schools.

After one month, following was the number of remaining plants.

22	6	48	0	28	22	17	10	32	6
22	22	28	26	17	36	10	22	28	0
28	22	48	32	10	48	25	36	6	32

Table

No. of Remaining Plants	Tally Marks	No. of students (frequency)
0 — 10	⊬⊬⊦ ‖	8
11 — 20	‖	2
21 — 30	⊬⊬⊦⊬⊬⊦ ‖	12
31 — 40	⊬⊬⊦	5
41 — 50	‖‖	3
Total		$\overline{\underline{\sum f = 30}}$

Presentation of data, by this method is called **'Grouped frequency distribution'**.

We can easily estimate and conclude by inspection of this table.

Exercises 15.2

1. Given below are the weights (in kg.) of 30 students of class X of a school :

34 34 36 37 38 33 34 35 36 37 38 33 34 35 34
33 37 35 34 36 38 36 35 34 35 37 38 34 35 35

Prepare a frequency distribution table for the above data.

2. Given below were the weights (in kg.) of 30 children born in a village :

3.4	3.6	3.0	3.8	3.6	3.8	2.9	3.4	2.9	3.4
3.0	3.4	3.2	3.1	3.2	3.2	3.1	3.2	3.4	3.0
3.1	3.2	3.5	3.7	3.1	3.0	2.9	3.0	3.1	3.2

Prepare a frequency distribution table for the above data.

3. Three coins were tossed 30 times. Each time the number of heads occuring was noted down as follow :

0	1	2	2	1	2	3	1	3	0
1	3	1	1	2	2	0	1	2	1
3	0	0	1	1	2	3	2	2	0

Prepare a frequency distribution table for the data given above.

4. The blood groups of 30 students of class X are recorded as follows :

   AB   O O A B O A O B A O B A   O   O

   A AB O A A O O AB   B A O B A   B   O

   Present this data in the form of a frequency distribution table. Find out which is the most common and which is the rarest blood group among these students.

5. Following are the marks obtained by 30 students of class IX in an examination. Use these marks to prepare a frequency table of 5 class size 10.

   19 27 40 3 33 41 18 8 20 0 23 49   16 36   14

   39 6 12 29 28 22 24 37 10 23 38 35   9   49   23

6. Prepare a frequency distribution table for the following distribution of taking class interval 5.

13	11	8	19	0	44	27	10	8	35	13
27	30	17	43	23	19	43	17	7		

7. The value of π upto 50 decimal places is given below :

   3.14159265358979323846264338327950288419716939937510

   (i) Prepare a frequency distribution table of the digits from 0 to 9 which comes after the decimal point.

   (ii) What are the most and the least frequentely occuring digits ?

8. The distances form the residence of 40 engineers to their working palce (in km) are as follows :

5	3	10	20	25	11	13	7	12	31
19	10	12	17	18	11	32	17	16	2
3	18	15	12	5	3	8	7	9	7
14	12	9	2	15	6	7	15	12	6

   By taking first class interval 0-5 (5 not included) prepare a frequency distribution table of class size 5 for the above data . What main features do you observe from this tabular representation ?

9. It is asked to 30 students that how many hours they have studied in last week ? Result obtained are as follows :

2	3	5	8	6	9	8	7	14	12
6	17	1	15	8	2	12	4	3	10
3	2	6	1	12	5	8	5	8	4

(i) Prepare a frequency distribution table from the above data, taking class size as 5.

(ii) How many students have studied upto 15 hours or more than in one week ?

15.04 Graphical Representation of Data

After the creation of universe, man has developed science and mathematics according to his needs. In this developmenmt, statistics also developed and figures are being used for sending information which is known as graphical representation. Graphs are eye-catching and easier to understand than the actual data. We shall study the following three types of graphical representation of the data :

(i)   Bar Graph

(ii)   Histograms

(iii)   Frequency polygon

(i)   **Bar graphs :** A bar graph is a pictorial representation of the numerical data, in which bars or rectangles of uniform width are drawn with equal spaces in between them on the axis. The height of bar is proportional to the numerical data it represents, we can understand this by the following example.

Example 1. In class IX, 40 students were asked about the months of their birth and the following graph was prepared for the data so obtained :
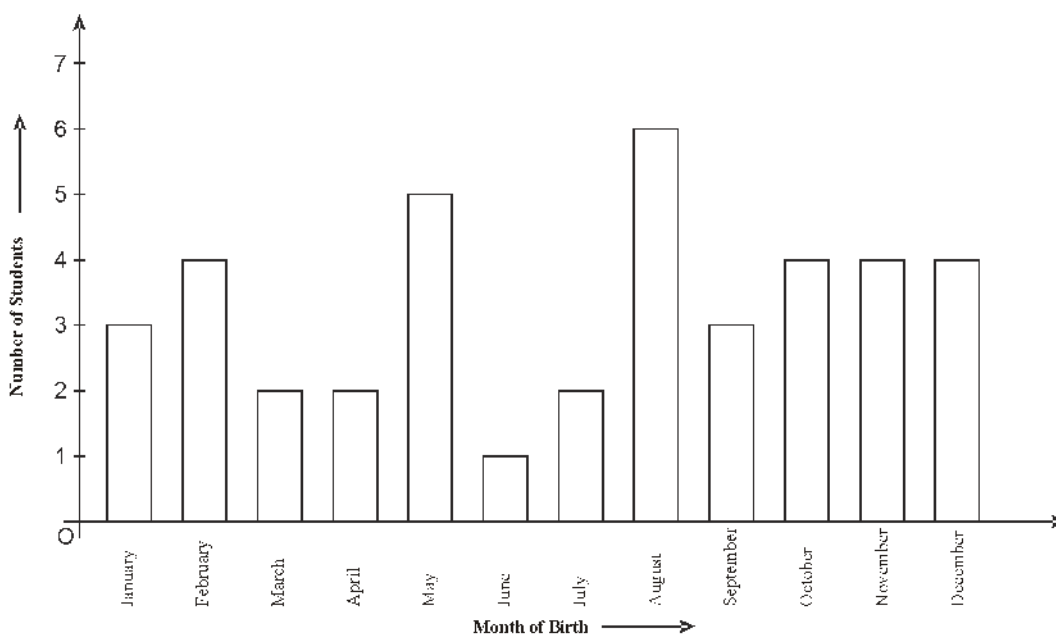
(i)   How many students were born in the month of May ?

(ii)   In which month maximum number of students were born ?

(iii)   In which month minimum number of students were born ?

Solution : Here variable is 'month of birth' and the value of variable is "number of students born".

(i) 5 students born in the month of May.

(ii) In August, the maximum number of students were born.

(iii) In June, the minimum number of students were born.

Example 2. The family has planned the following expenditures per month under various heads : monthly income of a family is 20 thousand.

Table - 3

Commodities (variable)	Expenditure (in thousand rupees)
Grocery	4
Rent	5
Children's education	5
Medicine	2
Fuel	2
Entertainment	1
Miscellaneous	1

Preapre a bar graph for the above data.

Solution : Following are the steps to prepare a bar graph for the above data :

1.   Take any scale and mark the variables on the horizontal axis. Since, the width of the bar is not important, but for clarity, we take equal width for all bars and maintain equal gaps in between. Let one variable be represented by one unit.

2.   We repersent the expenditure on the vertical axis. Since the maximum expenditure is 5 thousand rupees, so we can choose the scale as 1 unit = 1000 rupees.

3.   To represent first variable (Grocery), we draw a rectangle bar of width 1 unit and height 4 units.

4.   Similarly, other variable are represented by giving the unit space in between consecutive bars. (See Fig. 15.02)

[311]
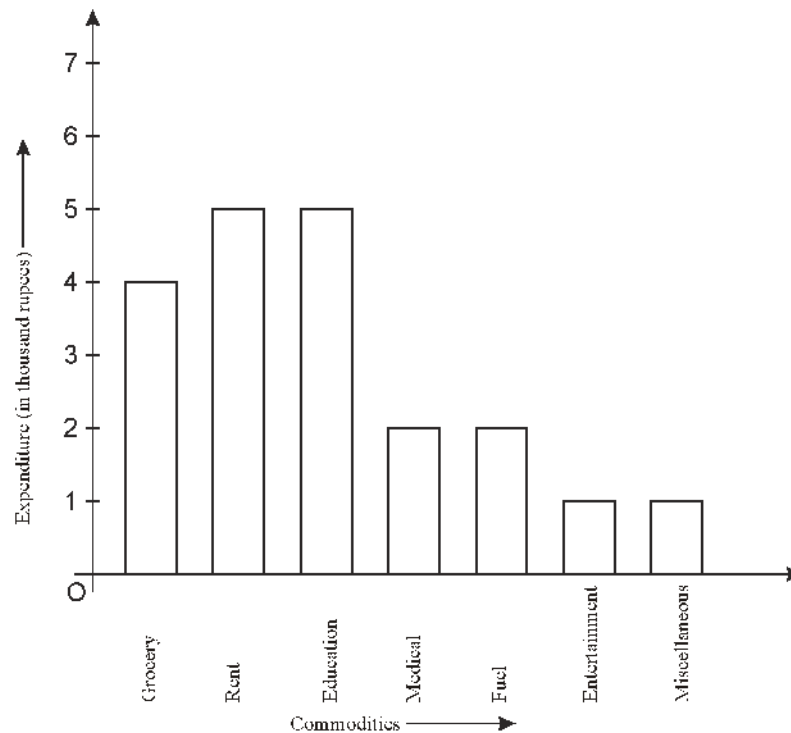
Here, we can easily compare the expenditures on different commodities. Thus, this is a best way instead of presenatation of data in tabular form.

(ii) Histogram

A histogram is a graphical representation of a frequency distribution in the form of rectangle with class intervals as bases and heights proportional to corresponsing frequencies. Choosing a suitable scale, mark class-limits on X-axis and frequencies on Y-axis, such that area of so formed rectangles should remain proportional to the corresponding frequencies. Here, we will study the construction of histograms related to four different types of frequency distribution.

(a)   When frequency distribution is grouped and continous and with equal class-intervals.

(b)   When frequency distribution is grouped and continuous but with unequal class-intervals.

(c)   When frequency distribution is grouped but not continuous.

(d)   When frequency distribution is ungrouped and mid-point of class-intervals are given.

Now,we will clarify the above statements by the following examples.

[312]

Example 3. Draw a histogram for the following frequency distribution:

Table-4

Class-interval (Age in Years)	0 — 5	5 — 10	10 — 15	15 – 20
No. of students	72	103	50	25

Solution : Here, frequency distribution is grouped and continuous and class-size is also same. Therefore, class-interval, *i.e.*, age in years (scale 1cm = 5 years) will be marked on *x*-axis.

Now, since number of students in class-interval 0-5 is 72. Taking class-interval 0-5 as base and the corresponding frequency as height, we construct rectangle. Similarly, we will construct rectangles B, C, D.
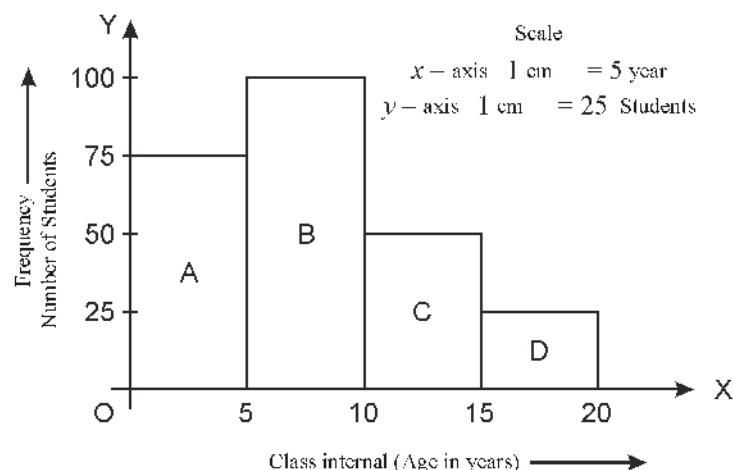
So, it is clear that all the rectangles have same base (1 cm) and height equal to frequency therefore area of rectangles should be proportional to frequency.

Example 4. The weekly wages of the workers of a company is given in the following table. Draw workers histogram for this data.

Table-5

Weekly Wages	1000 — 2000	2000 — 2500	2500 — 3000	3000 — 5000	5000 — 5500
Number of Workers	26	30	20	16	1

Solution : Here, frequency distribution is grouped and continuous but the class-interval is unequal, so to find heights of rectangles following method will be used, in which heights remains proportional to frequencies.

[313]

(a) Write minimum class size (h) of class-interval, here $h = 500$

(b) Compute the adjusted frequencies of each class by using the following formula :

$$\text{Adjusted frequency of a class} = \frac{h}{\text{class size}} \times \text{frequency of the class}$$

Thus, we obtain a new table as :

Table-6

Weekly Wages (in Rupees)	No. of Workers	Adjusted Frequencies (Adjusted No. of Workers)
1000 — 2000	26	$\dfrac{500}{1000} \times 26 = 13$
2000 — 2500	30	$\dfrac{500}{500} \times 30 = 30$
2500 — 3000	20	$\dfrac{500}{500} \times 20 = 20$
3000 — 5000	16	$\dfrac{500}{2000} \times 16 = 4$
5000 — 5500	1	$\dfrac{500}{500} \times 1 = 1$

[314]
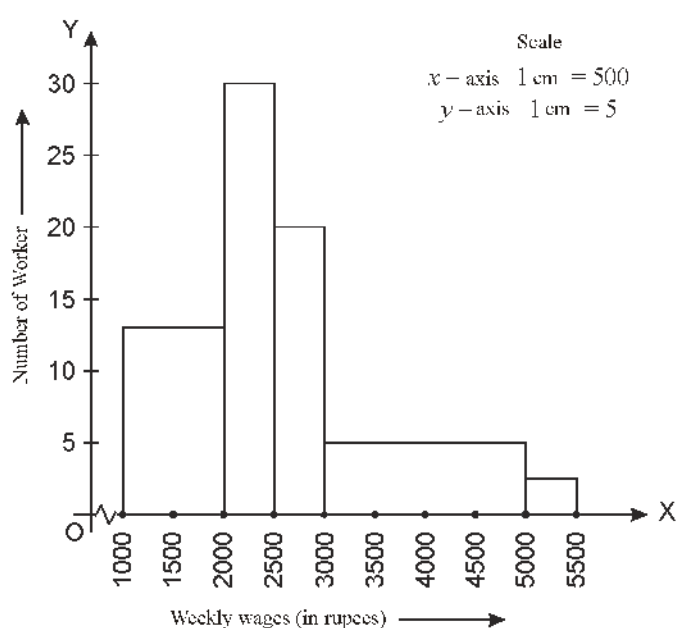
Draw class interval on *X*-axis (scale 1 cm = ₹ 500) and number of workers on *Y*- axis (scale 1 cm = 5 workers). Since, the first class interval (1000-2000) is starting form 1000 and not zero, so we show it on the graph by making a *kink* or a break on the axis. Now, we construct rectangles A, B, C, D, E with class-limits as bases and respective adjusted frequencies as heights.

Example 5. Draw a histogram for the frequency distribution :

| Class-interval	10 — 19	20 — 29	30 — 39	40 — 49	50 — 59
Frequency	2	3	5	7	3

Solution : Here, frequency distribution is grouped but not continuous so we first convert it into continuous frequency distribution. After making certain modifications, we get the following Table-8.

Table - 8

Class-interval	9.5 — 19.5	19.5 — 29.5	29.5 — 39.5	39.5 — 49.5	49.5 — 59.5
Frequency	2	3	5	7	3

Therefore, we will draw class-interval (scale 1cm = 10) on *x*-axis and frequencies on *y*-axis (scale 1 cm =1).

Since, the first class interval (9.5-19.5) is starting form 9.5 and not zero, so we show it on the graph by making a kink or a break on the axis. Thus, the required histogram is shown above.

[315]
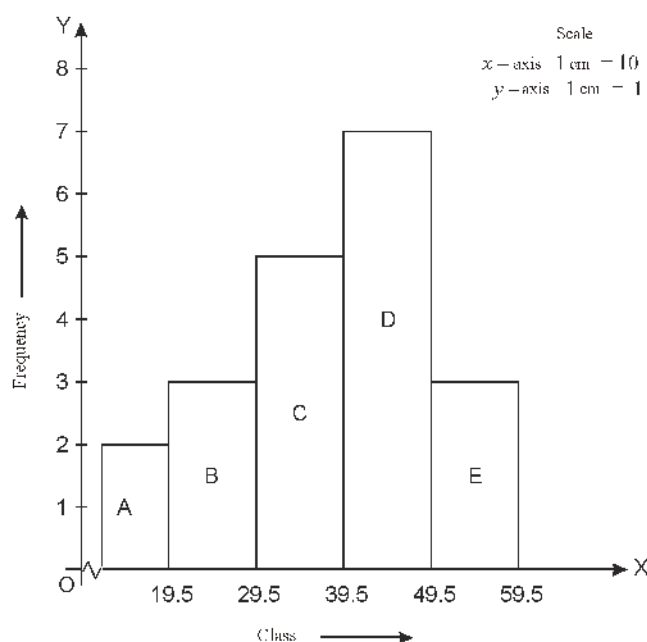
Example 6. Draw a histogram for the following frequency distribution :

Table - 9

Mid point	5	15	25	35	45
Frequency	4	7	6	2	1

Solution : Here frequency distribution is ungrouped and mid points of the distribution are given. So, after converting in grouped frequency distribution. We get the following table :

Table - 10

Class	0 — 10	10 — 20	20 — 30	30 — 40	40 — 50
Frequency	4	7	6	2	1

∴ This frequency distribution is grouped and continuous and class-size is also same. So, we will draw class-intervals on X-axis (scale : 1 cm = 10) and frequencies on Y-axis (scale : 1cm = 1) and obtained the follwing histogram.
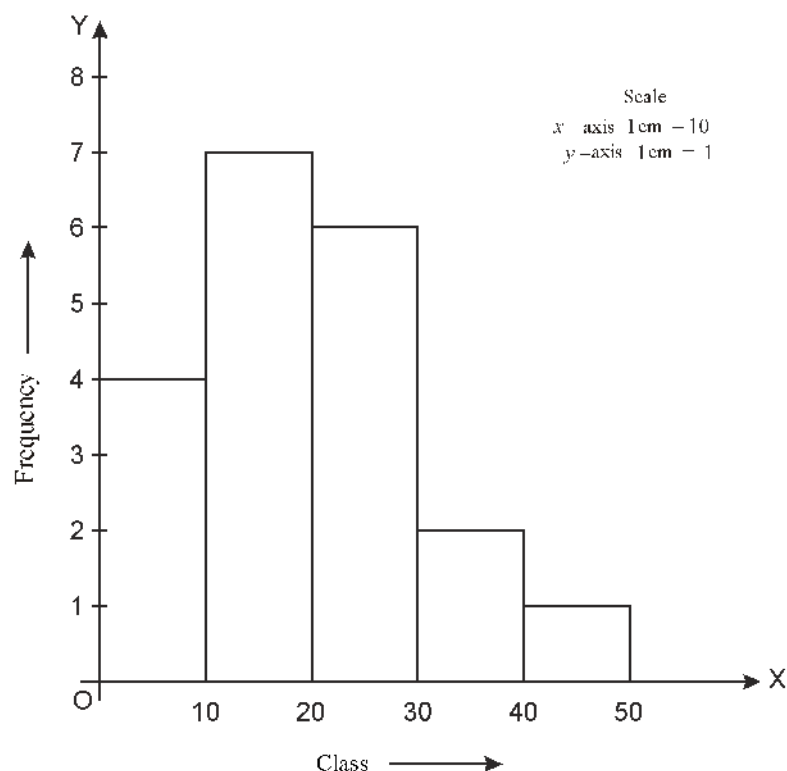
(3) Frequency Polygon

Frequency polygon is another method of representing frequency distributions graphically. Frequency polygon can be drawn in two ways :

1. By using histogram   2. Without using histogram

1. In order to draw a frequency polygon by using histogram, we may follow the following process :

(i)   Draw a histogram for the given frequency distribution.

(ii)   Obtain the mid points of the upper horizontal side of each rectangle.

(iii)   Join these mid points respectively, by straight lines.

(iv)   The mid points at each end are joined to the immediately lower or higher mid-point at zero frequency, *i.e.*, on the X-axis.

(v)   Join the mid point of last rectangle to its next possible mid point of class-interval on X-axis.

The figure so obtained is called a frequency polygon.

Remark :

If there is no possiblity of part (iv) & (v) for example, in case of marks obtained by the students in a test, we cannot go below zero and beyond maximum marks on the two sides. In such cases the extreme line-segments are only partly drawn and are brought down vertically so that they meet with the vertical sides of the first and last rectangles.

Example 7. Prepare a frequency polygon by making histogram for the following frequency distribution :

Table-11

Class-interval	20 — 30	30 — 40	40 — 50	50 — 60	60 — 70
Frequency	8	12	17	9	4

Solution : Given frequency distribution is grouped and continous so we draw histogram by the method mentioned earlier. Join the mid points P, Q, R, S, T of the upper horizontal side of rectangle A, B, C, D, E respectively, by straight lines.

Scale
x—axis 1cm = 10
y—axis 1cm = 2
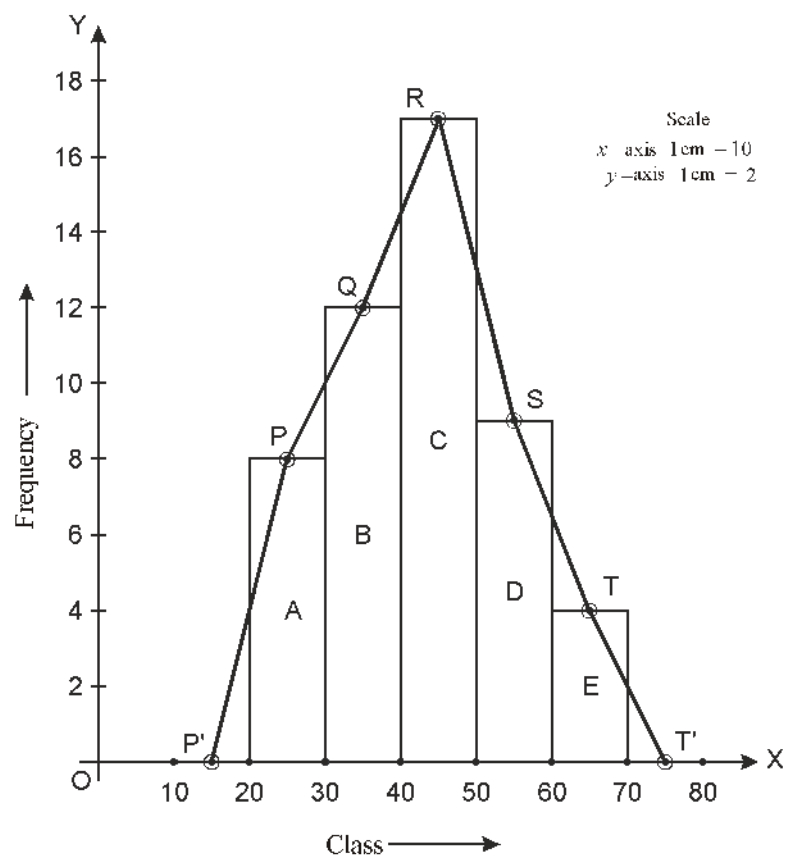
Since, here nothing is said about imagined classes so by joining mid points of class intervals from the begining and end points, we get frequency polygon P' $PQRSTT'$

Thus, the required frequency polygon is shown in fig. 15.7.

Example 8. **Make a frequency polygon by drawing a histogram for following frequency distribution.**

Table - 12

Marks	0 10	10 20	20 30	30 40	40 50	50 60	60 70	70 80	80 90	90 100
No. of Students	8	10	6	7	9	8	8	6	3	4

Solution : Given frequency distribution is grouped and continuous, so we draw histogram according to the method learnt earlier.

[318]
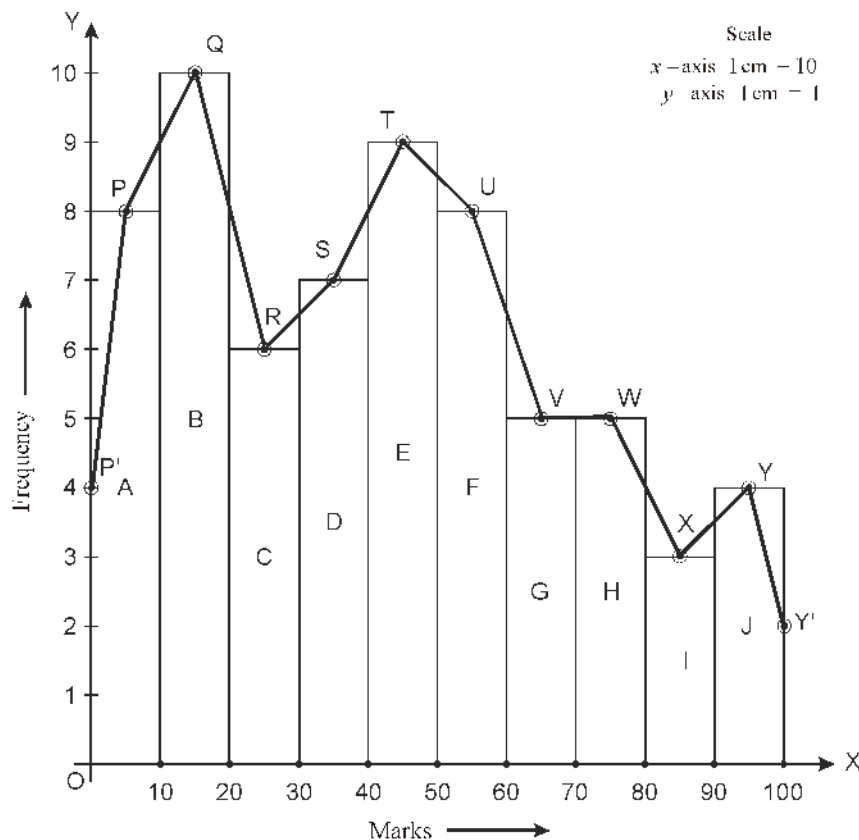
Join the mid-points *P, Q, R, S, T, U, V, W, X, Y* of the upper horizontal side of rectangles *A, B, C, D, E, F, G, H, I, J* resepectively, by straight lines.

Since, here marks obtained by students cannot below 0 and beyond 100 so imaginary intervals does not exist. The first mid point, i.e., P is joined to zero frequency the point where this line segment meets the vertical axis is marked as P'. Let Y be the mid-point of the class succeeding the last class of the given data. Thus, OP' PQRSTUVWXYY' is the frequency polygon, which is shown in Fig. 15.08.

2. Following is the working method to draw a frequency polygon without making histogram:

(i)   If frequency distribution is grouped, then find class mark. Now this frequency distribution will be converted into ungrouped.

(ii) Represent class marks on x-axis on a suitable scale.

(iii) Represent class marks on y-axis on a suitable scale.

[319]

(iv)　Plot the points (x_1, f_1), (x_2, f_2)...

(v)　Join these points by line segments.

(vi)　Join these ends by mid points of the imagined classes adjacent to them and thus obtained frequency polygon for given frequency distribution.

(v)　Join these points by line segments.

(vi)　Join these ends by mid points of the imagined classes adjacent to them and thus obtained frequency polygon for given frequency distribution.

Remark:

1.　For any type of frequency distribution, polygon can be made easily.

2.　If to make imaginary classes is not possible, then draw vertical lines at the starting point and end point of these class-intervals and join them with the mid point of class-interval. Thus, we will get required frequency polygon.

Example 9: Draw frequency polygon for given frequency distribution.

Table-13

Age (in years)	0 — 10	10 — 20	20— 30	30 — 40	40 — 50	50 — 60
Frequency	15	12	10	4	11	14

Solution : The frequency distribution is grouped and continuous. So, we obtained the following table on the basis of class.

Table-14

Age (in years)	0 — 10	10 — 20	20— 30	30 — 40	40 — 50	50 — 60
Class mark	5	15	25	35	45	55
Frequency	15	12	10	4	11	14

Now, mark the points (5, 15), (15, 12), (25, 10), (35, 4), (45, 11), (55, 14) assuming suitable scale.

[320]
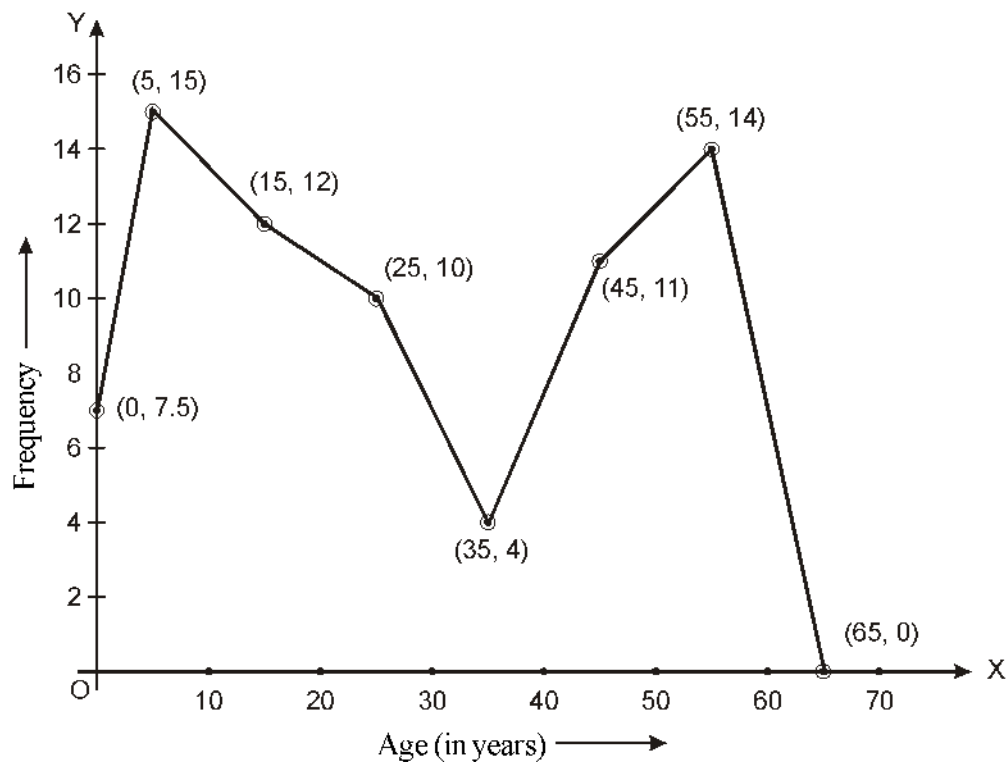
Since, age cannot be negative, so instead of imaginary classes, draw vertical lines at the starting point and end point of these class intervals and join them with the mid points of class interval. Thus, we will get required frequency polygon for given frequency distribution as shown in the figure.

Example 10. Draw frequency polygon for the following frequency distribution.

Table-15

Class	22 — 24	25 — 27	28— 30	31 — 33	34 — 36	37— 39	40 — 42
Frequency	12	7	5	8	11	6	4

Solution : Here, frequency distribution is grouped but not continuous because draw after making continuous class there is not change in class mark. Thus, instead of making continuous, we will find class mark and then draw frequency polygon.

Table - 16

Class	22 — 24	25 — 27	28— 30	31 — 33	34 — 36	37 — 39	40 — 42
Class mark	23	26	29	32	35	38	41
Frequency	12	7	5	8	11	6	4

[321]

By assuming suitable scale, we will mark the points (23, 12), (26, 7), (29, 5), (32, 8), (35, 11), (38, 6), (41, 4) on graph paper.
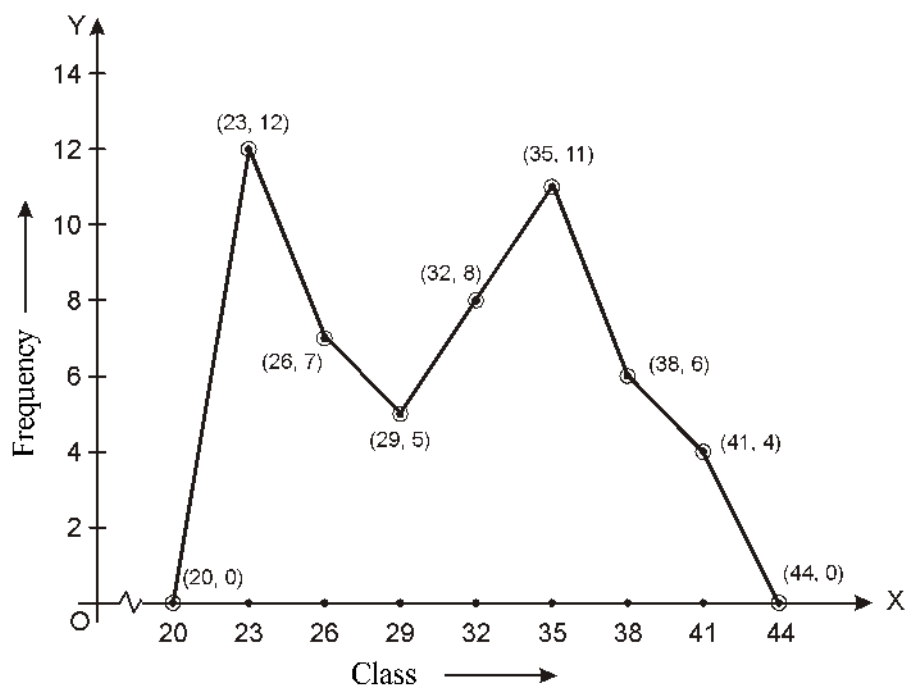
Join points (23, 12) to (20, 0) and (41, 4) to point (44, 0). Thus, we get the required frequency polygon.

Exercise 15.3

1. An organisation have conducted a survey in the whole world to find out the reason behind the diseases and death of females of age group 15-44 (in years) and obtained the following data (in %).

Table -17

S. No.	Reasons	Female death rate (%)
1.	Fertility health stage	31.8
2.	Nerve Psychiatric stage	25.4
3.	Loss	12.4
4.	Heart vessel stage	4.3
5.	Breathign stage	4.1
6.	Other reasons	22.0

(i) Express above informations by bar graph.

(ii) Which stage is the major cause of bad health and death of females in the whole world?

2. The following data on the number of girls (to the nearest ten) per thousand boys in different sections of Indain society is given below:

Table - 18

S.No.	Area	No. of Girls per thousand boys
1.	Scheduled Caste	940
2.	Scheduled Tribe	970
3.	Non SC/Tribe	920
4.	Backward districts	950
5.	Non-backward districts	920
6.	Rural	930
7.	Urban	910

(i) Represent the above informations by bar graph.

(ii) Discuss, what conclusion can be drawn from this graph?

3. In an assembly elections, results of winning seats by various parties are as follows:

Political Party   A   B   C   D   E   F

Winning Seats   75   55   37   29   10   37

(i) Draw a bar graph to represent the results of electron.

(ii) Which party won maximum seats?

From the following tables draw a histogram (4 to 8)

Table - 20

4.

Class	0 — 5	5 — 10	10 — 15	15 — 20	20 — 25
Frequency	18	15	14	8	10

Table - 21

5.

Class	0 — 20	20 — 40	40 — 60	60 — 80	80 — 100
Frequency	5	6	12	4	3

Table - 22

6.

Class	3 — 6	6 — 12	12 — 13	13 — 14	14 — 15
Frequency	150	420	100	110	50

Table - 23

7.

Class	5 — 9	10 — 14	15 — 19	20 — 24
Frequency	3	5	8	2

[323]

Table - 24

8.	Class	8	14	20	26	32
	Frequency	10	15	25	9	6

9. Draw a frequency polygon with the help of a histogram for the following distribution. Maximum number is 10.

Table - 25

Marks	0 — 5	5 — 10	10 — 15	15 — 20	20 — 25	25 — 30	30 — 35
Frequency	1	2	4	6	5	3	2

10. Draw a frequency polygon with the help of a histogram for the following distribution. Maximum number is 10.

Table - 26

Marks	0 — 2	2 — 4	4 — 6	6 — 8	8 — 10
Number of students	7	8	4	9	2

11. Draw a frequency polygon for the following frequency distribution.

Table - 27

Variable x	5	10	15	20	25	30
Frequency f	2	6	4	1	5	2

12. Draw a frequency polygon for the following frequency distribution.

Table - 28

Production (in tonnes)	0 — 10	10 — 20	20 — 30	30 — 40	40 — 50	50 — 60	60 — 70
Frequency	8	18	23	37	47	26	16

Measures of Central Tendency

In this chapter, we have studied the presentation of data in various forms - by frequency distribution tables, bar graphs, histograms and frequency polygons.

In order to make these data meaningful, there is always need to study all the data. This can be done with the help of measures of central tendency or averages.

Let us consider a situation, in which two students Parveen and Akash have received their test copies.

In test, there were five questions of 10 marks each. Following are the marks obtained by them.

Table - 29

Question Numbers	1	2	3	4	5
Makrs obtained by Praveen	10	8	9	8	7
Makrs obtained by Akash	4	7	10	10	10

Average marks obtained by Praveen $= \dfrac{42}{5} = 8.4$

Average marks obtained by Akash $= \dfrac{41}{5} = 8.2$

Since average marks of Praveen was greater than average marks of Akash. So Praveen says his performance is good but Akash was not agree with this. He kept marks obtained by both in ascending order and then obtained mid marks.

Table - 30

Marks obtained by Praveen	7	8	8	9	10
Marks obtained by Akash	4	10	10	10	10

Akash says his mid marks obtained was 10 whereas of Praveen's was 8. So his performance was better than Praveen, but Praveen was not agree with Akash to convince Praveen, Akash make at trick. He says, he had got 10 marks 3 times whereas Praveen have got only once. Thus his performance is better.

In first stage, Praveen has got average marks, that is **'Mean'.** Mid makrs which was used by Akash in his discussion is **Median.** In second discussion Akash has maximum marks got maximum times, is **mode.**

Now, we will discuss deeply on Mean.

The arithmetic mean of a set of observations is equal to their sum divided by the total number of observations and is denoted by $\bar{x}$ and read as (x bar).

Example 11. Daily wages of 5 workers are ₹ 250, ₹ 200, ₹ 225, ₹ 300, ₹ 275. Find their mean.

Solution : Mean of observations

$$(\bar{x}) = \frac{\text{Sum of all observations}}{\text{Total no. of observations}} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{5}$$

[325]

$$= \frac{250 + 200 + 225 + 300 + 275}{5} = \frac{1250}{5} = ₹250$$

Thus, average wages of 5 persons = ₹ 250

If we want to find the mean of 30 persons, we will write it as : $x_1 + x_2 + x_3...+x_{30}$ which is a lengthy process. We use Σ for summations. So, $x_1 + x_2 + x_3...+x_{30}$ can be written as $\sum\limits_{j=1}^{30} x_j$

$$\bar{x} = \frac{\sum\limits_{j=1}^{30} x_j}{30}$$

Similarly, if number of observation $= n$

$$\bar{x} = \frac{\sum\limits_{i=1}^{n} x_j}{n}$$

Median : Median of the given number of observations is the value of the variable which divides it exactly into two equal parts. So, when data is written in ascending or descending order, then median of ungrouped data is calculated as:

(i)   If number of observation (n) is odd, then

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

For example, if $n = 13$ then

$$\text{Median will be the value of} \left(\frac{13+1}{2}\right)^{\text{th}} \text{observation, i.e., value of 7th observation.}$$

(ii)   If number of observation (n) is even, then

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

For example, if $n = 16$, then

$$\text{Median} = \frac{\text{value of } 8^{th} \text{ observation} + \text{value of } 9^{th} \text{ observation}}{2}$$

Example 12 : Following are the heights (in cm) of 9 students of a class. Find the median of these data:

| 155 | 160 | 145 | 149 | 150 | 147 | 152 | 144 | 148 |

Solution : Arranging the given data in ascending order, we have

| 144 | 145 | 147 | 148 | 149 | 150 | 152 | 155 | 160 |

Here, number of students is 9 i.e., which is an odd number

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

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

$$= \text{value of } 5^{th} \text{ observation}$$

$$= 149 \text{ cm}$$

Example 13 : Following are the points obtained by a Kabaddi team in various matches:

| 17 | 2 | 7 | 27 | 15 | 5 | 14 | 8 | 10 | 24 |
| 48 | 10 | 8 | 7 | 18 | 28 |

Find the median of the points obtained by the team.

Solution : Arranging the given data in ascending order, we have

| 2 | 5 | 7 | 7 | 8 | 8 | 10 | 10 | 14 | 15 |
| 17 | 18 | 24 | 27 | 28 | 48 |

Here, number of observations is 16 which is an even number.

$$\therefore \quad \text{Median} = \frac{\left(\frac{16}{2}\right)^{th} \text{observation} + \left(\frac{16}{2} + 1\right)^{th} \text{observation}}{2}$$

$$= \frac{8^{th} + 9^{th} \text{observation}}{2}$$

$$= \frac{10 + 14}{2} = \frac{24}{2} = 12$$

Mode:

Mode is the value of the observation which occurs most frequently. It means a variable with maximum frequency is called mode.

This measure of central tendency is frequently used in readymade garment and shoe industry. With the help of mode, these industries take decision that propduction of which product whould be increased.

Example 14 : Find the mode of the marks obtained by 20 students (out of 10), from the followign data:

4, 6, 5, 9, 3, 2, 7, 7, 6, 5, 4, 9, 10, 10, 3, 4, 7, 6, 9, 9

Solution : Arranging the given data in ascending order, we have

2, 3, 3, 4, 4, 4, 5, 5, 6, 6, 6, 7, 7, 7, 9, 9, 9, 9, 10, 10

Here, 9 occurs maximum number of time, i.e., 4 times.

Thus, mode is 9.

Exercise 15.4

1. The following number of goals were scored by a team in a series of 10 matches:

   2   3   4   5   0   1   3   3   4   3

   Find the mean, median and mode of these scores.

2. In a mathematics test given to 15 students, the following marks (out of 100) are recorded:

   41   39   48   52   46   62   54   40   96   52

   98   40   42   52   60

   Find the mean, median and mode for above data.

3. The following observations are arranged in ascending order. If the median of the data is 63, then find the vlaue of x :

| 29 | 32 | 48 | 50 | x | $x + 2$ | 72 | 78 | 84 | 95 |

4. Find the mode of 14   25   14   28   18   17   18   14
   23   22   14   18

5. Find the mean salary of 60 workers, working in a factory, from the following table:

Table - 31

Salary (in ₹)	3000	4000	5000	6000	7000	8000	9000	10000
No. of workers	16	12	10	8	6	4	3	1

Miscellaneous Exericse - 15

1. In a distribution 5, 5, 6, 4, 9, 5, 3, 2, 7, 6, 3, 8, 4 frequency of class interval 3 - 5 is

   (A) 3   (B) 4   (C) 6   (D) 7

2. Range of the following distribution will be

   3.2   2.8   3.1   2.1   3.2   2.4   2.1   2.8   2.7   2.7

   (A) 2.7   (B) 3.1   (C) 2.4   (D) 1.1

3. In the following distribution, number of students of age less than 25 years is

Age (in years)	5-10	10 - 15	15 - 20	20 - 25	25 - 30
No. of students	3	6	8	8	2

   (A) 8   (B) 16   (C) 9   (D) 25

4. In a bar graph, height of rectangle is

   (A) inversely proportiional to frequency of class

   (B) proportional to frequency of class

   (C) proportional to class-interval

   (D) inversely proportional to class-interval

5. The examaination result of any class of a school can be comparatively studied:

   (A) by circular diagram   (B) by bar graph

   (C) by linear graph   (D) all of these

6. The range of distribution 6, 1, 2, 3, 9, 8, 3, 4, 8, 2, 3 will be

   (A) 4   (B) 8   (C) 7   (D) 6

[329]

7. In the distribution 5, 1, 5, 2, 3, 6, 5, 4 frequency of 5 will be
   (A) 1   (B) 2   (C) 3   (D) 4

8. The median of 11, 2, 7, 8, 9, 3, 5 will be
   (A) 7   (B) 9   (C) 5   (D) 11

9. The mean of 15, 0, 10, 5 will be
   (A) 15   (B) 10   (C) 5   (D) 7.5

10. In 4, 3, 4, 5, 4, 2, 4, 1 mode will be
    (A) 1   (B) 2   (C) 5   (D) 4

Very Short Answer Type Question

11. Write the frequency of class-interval 0 - 5 from the following distribution:
    3, 2, 0, 10, 8, 5, 13, 5, 6, 6, 0, 14

12. If 7 is the mean of the numbers 5, 8, 4, x, 6, 9 then find the vlaue of x

13. What is range?

14. What is histogram?

15. Make a frequency table of 9, 7, 9, 8, 3, 9, 8, 3, 5, 7, 5, 3.

16. If arithematic mean of a frequency distribution is 15 and $\Sigma f = 20$ then find the value of $\Sigma f x$

17. Write median of distribution 5, 2, 3, 7, 5, 4, 3, 2, 1

18. Find median of distribution 12, 1, 6, 4, 10, 8, 1, 4

19. Find mode of distribution 4, 3, 4, 1, 2, 4, 7, 5, 3

[330]

Importants Points

1. The data are of two types : (i) primary data (ii) secondary data.

2. The methods of collecting the primary data
   (i) By direct personal (individual) investigation, (ii) By indirect investigation.

3. Indirectly data is obtained by the following ways:
   (i) Filling the list by statisticians, (ii) Filling the list by informer,
   (iii) By local sources of correspondent,
   (iv) Direct oral investigation by the specialists.

4. Data which have already collected whether published or not, is called secondary data.

5. Secondary data may be published or not.

6. Following are the main sources of published data:
   (i) International organisation, (ii) Government publication,
   (iii) Semi-government publication, (iv) Publication of business organisation,
   (v) Publication of research centres, (vi) Paper-magazines,
   (vii) Published papers of research scholars.

7. Generally, x is used for marks is frequency distribution.

8. Range is the difference between maximum and minimum value of variable x.

9. Maximum and minimum value of a class is known as its upper and lower limit.

10. Average of upper and lower limit of a class is called mid vlaue of class or class mark and is represented as x

11. The following graphs are mainly used in statistical analysis:
    (a) Bar graph, (b) Histogram, (c) Frequency polygon

12. Three measures of central tendency for ungrouped data are:

    (i) **Mean :** The mean of a set of observation is equal to their sum divided by the total number of observations, i.e.,

    $$\bar{x} = \frac{\Sigma f x}{\Sigma f}$$

    (ii) **Median :** Middle most value (observation) is called median.

    (iii) **Mode :** Observation which occurs maximum times is called mode.

Answer

Exercise 15.1

1. **Primary data :** When an investigator collects data himself with a definite plan or design in his mind, it is called primary data.

   Secondary data : Data which are not originally collected rather obtained from published or unpublished sources is known as secondary data.

Exercise 15.2

1.

Variable (x)	33	34	35	36	37	38
Tally makrs	|||	⌒⌒|||	⌒⌒||	||||	||||	||||
Frequency (f)	3	8	7	4	4	4

2.

Variable (x)	2.9	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8
Tally makrs	|||	⌒⌒	⌒⌒	⌒⌒|	–	⌒⌒			||			||
Frequency (f)	3	5	5	6	–	5	1	2	1	2

3.

No. of ends	Tally makrs	Frequency (f)
0	⌒⌒|	6
1	⌒⌒⌒⌒	10
2	⌒⌒||||	9
3	⌒⌒	5
Total		30

4.

Blood group	Tally makrs	No. of students
A	⌒⌒||||	9
B	⌒⌒|	6
O	⌒⌒⌒⌒||	12
AB	|||	3
Total		30

5.

Class interval	Tally makrs	Frequency
0 — 10	⌒⌒	5
10 — 20	⌒⌒|	6

20 — 30	‖‖‖‖‖	9
30 — 40	‖‖‖‖	6
40 — 50	‖‖‖‖	4

6.

Class	Tally Makrs	Frequency (f)
0 — 5		1
5 — 10	|||	3
10 — 15	||||	4
15 — 20	||||	4
20 — 25		1
25 — 30	||	2
30 — 35		1
35 — 40		1
40 — 45	|||	3

7. (i)

Digits	0	1	2	3	4	5	6	7	8	9	Total
Frequency	2	5	5	8	4	5	4	4	5	8	50

(ii)   3 and 9 are the numbers whcih occurs maximum and number 0 ocurs minimum

8.

Distance (in km)	Tally Marks	Frequency
0 — 5	‖‖‖	5
5 — 10	‖‖‖‖‖‖	11
10 — 15	‖‖‖‖‖‖	11
15 — 20	‖‖‖‖‖	9
20 — 25		1
25 — 30		1
30 — 35	||	2

9.

Number of hours	Frequency
0 — 5	10
5 — 10	13
10 — 15	5
15 — 20	2
Total	30

    (ii)   2 Children

Exericse 15.3

1.   (ii)   Fertility health stage

3.   (ii)   Party A

Exercise 15.4

1.   Mean = 2.8, Median = 3, Mode = 3

2.   Mean = 54.8, Median = 52, Mode = 52

3. $x = 62$

4.   14

5.   Mean salary of 60 workers is ₹ 5083.33.

Miscellaneous Exericse - 15

1. (b)   **2.** (d)   **3.** (d)   **4.** (d)   **5.** (b)   **6.** (b)

7. (c)   **8.** (a)   **9.** (d)   **10.** (d)   **11.** −4   **12.** −10

13. Range is the difference between the maximum and the minimum valeu of variable x.

14. Histogram is a rectangular form of grouped and continuous frequency distribution.

15.

x	3	5	7	8	9
f	3	2	2	2	3

16. 300   **17.** 3   **18.** 5   **19.** 4