

STATISTICS

INTRODUCTION

In various field, we need information in the form of numerical figure called data.

These data may relate to the marks obtained by the pupils of a class in a certain examination; the weights, height, ages, etc., of pupils in a class; the monthly wages earned by workers in a factory; the population of a town or the profits of a company during last few years, etc.

Evaluation of such data helps analysis study the various growth patterns all formulate future targets or policies or derive certain inferences.



IMPORTANT POINTS

- ◆ The word data means information. 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- ◆ Data which are not originally collected rather obtained from published or unpublished sources are known as **Secondary data**.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 them in tabular form in order to study their salient features. Such an arrangement is called **Presentation of data**.
- ◆ Raw data (if the data is arranged in the manner as it was collected, then it is called as 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:-

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.

Variate	Tally marks	Frequency
15		4
16		6
18		6
20		6
24		5
25		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.

$$\therefore \text{Range} = 112 - 14 = 98.$$

It is given that the class size is 10, and

$$\frac{\text{Range}}{\text{Classsize}} = \frac{98}{10} = 9.8$$

So, we should have 10 classes each of size 10. 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. 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. Now