

Data on attributes may be of two types - ordinal and nominal. They are ordinal when there is a clear ordering of the forms or categories of the attribute. For example, when the character 'education' is measured with categories primary school, high school, college and post-graduate, or the character 'economic status' is measured with categories poor, middle class and rich, there is a clear ordering of the categories though the absolute distances between them are unknown. Such data are common in the social sciences, in particular for measuring attitudes and opinion on various issues and status of various types. When the various forms of an attribute differ in nature, not in quantity, the data are nominal. For example, when we record the religions of persons as Hindu, Muslim, Christian, etc. or when the mother tongues of students are noted as Bengali, Hindi, Marathi, etc. the data are nominal. It is obvious that the order of listing of the different forms of an attribute is unimportant in this case.

The term variable (or variate) means a character of an item or an individual that can be expressed in numerical terms. It is also called a quantitative character and such characters can be measured

or constant. Weight of students in a school, ages of boys, family size etc. are characters of this type

### 1.3 P. 1.3 Population and Sample

In a statistical investigation, the interest lies in some characteristic relating to a group of individuals and such a group under study is referred to as the population or universe. The members of a population may be the employees of an industry, the apples in a basket, the various cultivable plots in a village, and so on. The number of members in a population represents its size. When complete information is taken for the whole population, it is called a census.

Quite often, it may not be practicable to study the whole population due to limitation of time, money and man-power, or due to the population being infinite, and, as such, we are to depend on the study of a part of the population for determination of the population characteristics. A part of the population which is meant to represent the whole population is called a sample and its selection is termed as sampling. If a sample is drawn from a population in such a manner that each member of the population has a definite pre-assigned probability

of being included in the sample, the sampling is called random sampling. A survey conducted on a suitable sample is called sample survey.

#### 1.4. Tabulation of data & diagrammatic representation of data

Tabular presentation or tabulation of data is an useful mode of exhibiting the data in a compact form. The systematic presentation of data in the structure of a table comprising some rows and columns, is called tabulation.

Graphs, charts, maps, pictures, etc. attractive and effective means for presentation of statistical data. Diagrams are readily capable of revealing some features of the exhibited data. It should be noted that the selection of an appropriate diagram depends mainly on the nature of the given data.

Following are the important merits and demerits of the diagrammatic mode of presenting data.

- Merits:
- (i) It is simple to understand even by laymen.
  - (ii) It is very essential for conveying statistical information to the general public in a short time.
  - (iii) In this approach, one may acquire some idea regarding the ~~importance~~ significance of the presented data at a glance.
  - (iv) This mode is ~~capable~~ capable of creating lasting impression.
  - (v) Two or more series of data can easily be compared.



- Demerits :
- (i) Diagrams fail to represent details; they only show the general nature of the given data.
  - (ii) Usually, a diagram represents the figures in approximate form and, as such, in most of the cases, precision of data has to be sacrificed.
  - (iii) Construction of a diagram requires sufficient time.
  - (iv) Only limited information can be presented in a diagram.

We now discuss some commonly used diagrams:

(a) Bar diagram: This mode of diagrammatic representation comprises a number of equidistant rectangles (turned as bars), each of ~~which~~ them being meant for some specific category of the available data. Bars of common width (conveniently chosen) are drawn on the base line, the length or height of a bar representing the value of the corresponding category. Bars may be vertical or horizontal, depending on the ~~the~~ nature of statistical data. Generally, vertical bars are drawn for time-series data and horizontal bars for data varying over space.

As already noted, bar diagrams are applicable to represent both time-series and spatial series data. Besides, these diagrams are useful in exhibiting data classified on the basis of quantitative characters.