Chapter 3

Organisation of Data

Organisation of Data

Organisation of data refers to the arrangement of figures in such a form that comparison of the mass of similar data may be facilitated and further analysis may be possible.

Classification

Classification is the process of arranging things in groups or classes according to their resemblances and affinities and gives expression to the unity of attributes that may exist amongst a diversity of individuals.

Objectives of Classification

- Simplification and Briefness
- Utility
- Distinctiveness
- Comparability
- Scientific arrangement
- Attractive and effective

Characteristic of a Good Classification

- Comprehensiveness
- Clarity
- Homogeneity
- Suitability
- Stability
- Elastic

Basis of Classification

- Geographical Classification This classification of data is based on the geographical or locational differences of the data.
- Chronological Classification When data are classified on the basis of time, it is known as chronological classification.
- Qualitative Classification This classification is according to qualities or attributes of the data.
  This classification may be of two types
    - Simple classification
    - Manifold classification
- Quantitative or Numerical Classification Data are classified in to classes or groups on the basis of their numerical values. Quantitative classification is also called classification by variables.
- Concept of Variable A characteristic or a phenomenon which is capable of being measured and changes its value overtime is called a variable. The variable may be either discrete or continuous
    - Discrete Variable These are those variables that increase in jumps or in compete numbers.
    - Continuous Variable Variable that assume a range of values or increase not in jumps but continuously or in fractions are called continuous variables.
- Raw Data A mass of data in its crude form is called raw data.

Types of Statistical Series Statistical series are of two types

- Individual Series These are those series in which the items are listed singly. These series may be presented in two ways
    - According to serial numbers
    - Ascending or descending order of data
- Frequency Series Frequency series may be of two types
    - Discrete Series or Frequency Array It is that series in which data are presented in way that exact measurement of items are clearly shown. In this series there are no class intervals and a particular item in the series.
    - Frequency Distribution It is that series in which items cannot be exactly measured. The items assume a range of values and are placed within the limits is called class interval.

Frequency distribution is also known as continuous series or series with class-intervals, or series of grouped data.

Types of Frequency Distribution

- Exclusive Series It is that series in which every class-interval excludes items corresponding to its upper limit.
- Inclusive Series An inclusive series is that series which includes all items upto its upper limit.
- Open End Series An open end series is that series in which lower limit of the first class-interval and the upper limit of last class- interval is missing like as below – 5, 20 and above
- Cumulative Frequency Series It is that series in which the frequencies are continuously added corresponding to each class-interval in the series.
  There are two ways of converting this series into cumulative frequency series
    - Cumulative frequencies may be expressed on the basis of upper class limits of the class-intervals.
    - Cumulative frequencies may b expressed on the basis of lower class limits of the class-intervals.
- Mid Values Frequency Series Mid value frequency series are those series in which we have only mid values of the class intervals and the corresponding frequencies.
- Univariate Distribution The frequency distribution of a single variable is called a univariate distribution.
- Bivariate Distribution A bivariate distribution is the frequency distribution of two variables.