# Classification of Data

The method of arranging data into homogeneous classes according to some common features present in the data is called classification. Raw data cannot be easily understood and it is not fit for further analysis and interpretation. This arrangement of data helps users in comparison and analysis. For example, the Population of town can be grouped according to sex, age, marital status etc.

## Objectives of Data Classification

- To consolidate the volume of data in such a way that similarities and differences can be quickly understood. Figures can consequently be ordered in a few sections holding common traits.
- To aid comparison.
- To point out the important characteristics of the data at a flash.
- To give importance to the prominent data collected while separating the optional elements.
- To allow a statistical method of the material gathered.

## Types of Classification

1. **Geographical Classification:** In this type of classification, data are organized in terms of geographical area or region. State-wise production of manufactured goods is an example of this type. Data collected from an all India market survey may be classified geographically. Usually the regions are arranged alphabetically or according to the size to indicate the importance.

2. **Chronological Classification:** When data is arranged according to time of occurrence, it is called chronological classification. Profit of engineering industries over the last few years is an example. In such a classification, data are classified either in ascending or in descending order with reference to time such as years, quarters, months, weeks, etc.

3. **Qualitative Classification:** Under this classification, data are classified on the basis of some attributes or qualities like honesty, beauty, intelligence, literacy, marital status etc. For example, Population can be divided on the basis of marital status as married or unmarried etc.

4. **Quantitative Classification:** This type of classification is made on the basis some measurable characteristics like height, weight, age, income, marks of students, etc.

5. **Simple Classification:** In the case of simple classification each class is divided into two sub classes and only one attribute is studied like user of a product or non-user of a product, married or unmarried, employed or unemployed, Brahmin or non-Brahmin etc.

6. **Manifold Classification:** In the case of manifold classification more than one attributes are considered. For example, the respondents in a survey may be classified as user of a particular brand of a product and non-user of particular brand of product. Both user and non-user can be further classified into male and female. Further one can classify male and female into two categories such as below 25 years of age and 25 and more years of age. We can further classify them as professionals at non-professionals.

**While developing the classes, the following decisions are taken:**

**1. Class Limits**

It refers to the lowest and the highest limit of the values that can be included in a class e.g. in class 100-200 the lowest limit is 100, below which there can be no item in the class and the upper limit is 200, above which there can be no item in the class.

**2. Type of Class**

The classes can be created in two ways:

i. **Exclusive method:** In this method, the upper limit of one class is the lower limit of the next class e.g.

| Income (₹) | Number of Persons |
|---|---|
| 5000-10,000 | 25 |
| 10,000-15,000 | 28 |
| 15,000-20,000 | 32 |

ii. **Inclusive method:** In this method, the upper limit of a class is included in that class itself:

| Income (₹) | Number of Persons |
|---|---|
| 5000-9,999 | 25 |
| 10,000-14,999 | 28 |
| 15,000-19,999 | 32 |

3. **Class Interval:** The difference between the upper and the lower limit of a class is termed as class interval e.g. 10,000 – 5000 = 5000. The class interval is fixed based on factors like range, the number of observations, etc.

**Mid-point of a class** = $\frac{\text{Upper limit of class} + \text{Lower Limit of class}}{2}$

# Tabulation of Data

It simply means presenting a data through tables. It is next to classification in the process of statistical investigation. In a broader sense, 'Tabulation is an orderly arrangement of data in columns and rows'.

Tabulation is the final stage in collection and compilation of data and it is a sort of stepping stone to the analysis and interpretation of figures. In deciding about the type of tabulation, one has to keep in mind the nature, scope and the object of inquiry. Tabulation of data should be done in such a form that it suits the nature and object of the research investigation.

The tabulation is used for summarization and condensation of data. It aids in analysis of relationships, trends and other summarization of the given data. The tabulation may be simple or complex. Simple tabulation results in one-way tables, which can be used to answer questions related to one characteristic of the data. The complex tabulation usually results in two way tables, which give information about two interrelated characteristics of the data; three way tables which give information about three interrelated characteristics of data.

Tabulation is the process of recording the classified facts in rows and columns. Therefore, after classifying the data into various classes, they should be in the tabular form.

For an ideal statistical table, certain guidelines are very helpful in its preparation. They are as follows:

- A good table present the data in as clear and simple manner as possible.
- The title should be brief and self-explanatory. It should represent the description of the contents of the table.
- Rows and columns may be numbered to facilitate easy reference.
- Table should not be too narrow or too wide. The space of columns and rows should be carefully planned, so as to avoid unnecessary gaps.
- Columns and rows which are directly comparable with one another should be placed side by side.

- Units of measurement should be clearly shown.
- All the column figures should be properly aligned. Decimal points and plus or minus signs also should be in perfect alignment.
- Abbreviations should be avoided in a table. If it is inevitable to use, their meanings must be clearly explained in footnote.
- If necessary, the derived data (percentages, indices, ratios, etc.) may also be incorporated in the tables.
- The sources of data should be clearly stated, so that the reliability of the data could be verified, if needed.

## Objectives of Tabulation

**1. To Simplify the Complex Data**

- It reduces the bulk of information i.e. raw data in a simplified and meaningful form so that it could be easily by a common man in less time.

**2. To Bring Out Essential Features of the Data**

- It brings out the chief/main characteristics of data.
- It presents facts clearly and precisely without textual explanation.

**3. To Facilitate Comparison**

- Presentation of data in row & column is helpful in simultaneous detailed comparison on the basis of several parameters.

**4. To Facilitate Statistical Analysis**

- Tables serve as the best source of organized data for further statistical analysis.
- The task of computing average, dispersion, correlation, etc. becomes much easier if data is presented in the form of a table.

**5. Saving of Space**

- A table presents facts in a better way than the textual form.
- It saves space without sacrificing the quality and quantity of data.

## Types of Tabulation

### 1. Simple Tabulation or One-way Tabulation

When the data are tabulated to one characteristic, it is said to be a simple tabulation or one-way tabulation.

For example, Tabulation of data on the population of the world classified by one characteristic like religion is an example of a simple tabulation.

### 2. Double Tabulation or Two-way Tabulation

When the data are tabulated according to two characteristics at a time, it is said to be a double tabulation or two-way tabulation.

For example, Tabulation of data on the population of the world classified by two characteristics like religion and sex is an example of a double tabulation.

### 3. Complex Tabulation

When the data are tabulated according to many characteristics, it is said to be a complex tabulation.

For example, Tabulation of data on the population of the world classified by three or more characteristics like religion, sex and literacy, etc. is an example of a complex tabulation.