**Classification and Tabulation**

The data collected for the purpose of a statistical inquiry some times consists of a few fairly simple figures, which can be easily understood without any special treatment. But more often there is an overwhelming mass of raw data without any structure. Thus, unwieldy, unorganised and shapeless mass of collected is not capable of being rapidly or easily associated or interpreted. Unorganised data are not fit for further analysis and interpretation. In order to make the data simple and easily understandable the first task is not condense and simplify them in such a way that irrelevant data are removed and their significant features are stand out prominently. The procedure adopted for this purpose is known as method of classification and tabulation. Classification helps proper tabulation.

“Classified and arranged facts speak themselves; unarranged, unorganised they are dead as mutton”.

 **Meaning of Classification**

Classification is a process of arranging things or data in groups or classes according to their resemblances and affinities and gives expressions to the unity of attributes that may subsit among a diversity of individuals.

 **Definition of Classification**

Classification is the process of arranging data into sequences and groups according to their common characteristics or separating them into different but related parts.

The process of grouping large number of individual facts and observations on the basis of similarity among the items is called classification.

- Stockton & Clark

**Characteristics of classification**

a) Classification performs homogeneous grouping of data

b) It brings out points of similarity and dissimilarities.

c) The classification may be either real or imaginary

d) Classification is flexible to accommodate adjustments

**Objectives / purposes of classifications**

* To simplify and condense the large data
* To present the facts to easily in understandable form
* To allow comparisons
* To help to draw valid inferences
* To relate the variables among the data
* To help further analysis
* To eliminate unwanted data
* To prepare tabulation

**Important types of classification**

* Geographical (i.e. on the basis of area or region wise)
* Chronological (On the basis of Temporal / Historical, i.e. with respect to time)
* Qualitative (on the basis of character / attributes)
* Numerical, quantitative (on the basis of magnitude)

***a) Geographical Classification***

In geographical classification, the classification is based on the geographical regions.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Ex***: Sales of the company (In Million Rupees) (region – wise)

|  |  |
| --- | --- |
| **Region** | **Sales**  |
| North  | 285  |
| South  | 300  |
| East  | 185  |
| West  | 235  |

 |

***b)Chronological Classification***

If the statistical data are classified according to the time of its occurrence, the type of classification is called chronological classification.

Ex. Sales reported by a departmental store

|  |  |
| --- | --- |
| **Month** | **Sales** **(Rs.) in lakhs** |
| January  | 22  |
| February  | 26  |
| March  | 32  |
| April  | 25  |
| May  | 27  |
| June  | 30  |

***c) Qualitative Classification***

In qualitative classifications, the data are classified according to the presence or absence of attributes in given units. Thus, the classification is based on some quality characteristics / attributes.

**Ex**: Sex, Literacy, Education, Class grade etc.

Further, it may be classified as

a) Simple classification b) Manifold classification

i) *Simple classification*: If the classification is done into only two classes then classification is known as simple classification.

***Ex:*** a) Population in to Male / Female

b) Population into Educated / Uneducated

ii) *Manifold classification*: In this classification, the classification is based on more than one attribute at a time.

***Quantitative Classification***:

In Quantitative classification, the classification is based on quantitative measurements of some characteristics, such as age, marks, income, production, sales etc. The quantitative phenomenon under study is known as variable and hence this classification is also called as classification by variable.

***Ex***: For a 50 marks test, Marks obtained by students as classified as follows

|  |  |
| --- | --- |
| **Marks**  | **No. of students**  |
| 0 – 10  | 5  |
| 10 – 20  | 7  |
| 20 – 30  | 10  |
| 30 – 40  | 25  |
| 40 – 50  | 3  |
|  |  |

In this classification marks obtained by students is variable and number of students in each class represents the frequency.

**Tabulation**

**Meaning and Definition of Tabulation**

Tabulation may be defined, as systematic arrangement of data is column and rows. It is designed to simplify presentation of data for the purpose of analysis and statistical inferences

**Major Objectives of Tabulation**

1. To simplify the complex data

2. To facilitate comparison

3. To economise the space

4. To draw valid inference / conclusions

5. To help for further analysis

**Differences between Classification and Tabulation**

1. First data are classified and presented in tables; classification is the basis for tabulation.

2. Tabulation is a mechanical function of classification because is tabulation classified data are placed in row and columns.

3. Classification is a process of statistical analysis while tabulation is a process of presenting data is suitable structure.

**Classification of tables**

Classification is done based on

1. Coverage (Simple and complex table)

2. Objective / purpose (General purpose / Reference table / Special table or summary table)

3. Nature of inquiry (primary and derived table).

***Ex:***

1. *Simple table*: Data are classified based on only one characteristic
2. *Complex table* Data are classified based on two or more characteristic