

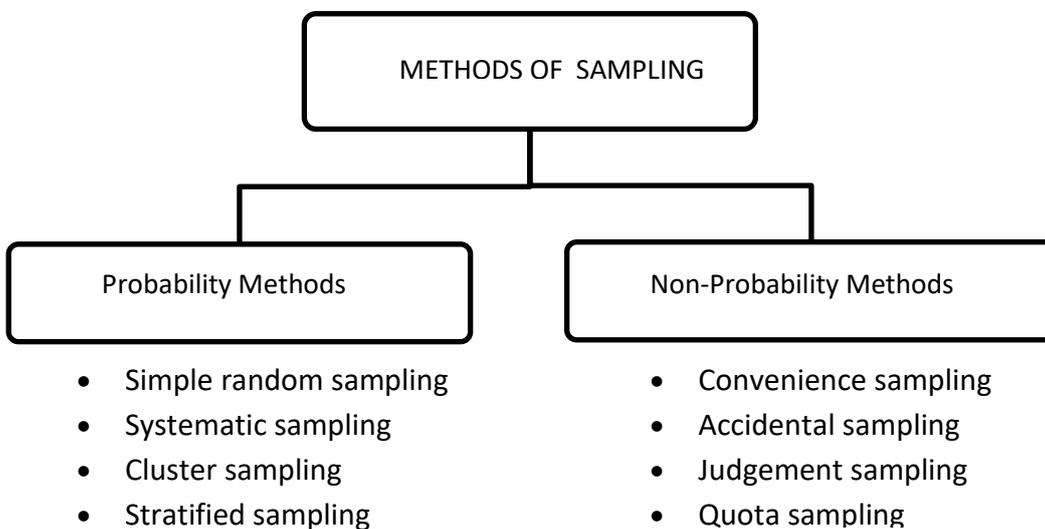
# **SUB : RESEARCH METHODOLOGY IN ACCOUNTING AND FINANCE.**

## **UNIT: 3**

### **DATA COLLECTION AND PROCESSING**

#### **METHODS OF SAMPLING**

The methods of sampling can be broadly divided into two groups:



#### **I. Probability Methods:**

Probability sampling is also known random sampling. Probability means possible chance. Therefore, each element of the population has known chance or opportunity of being selected or included in the sample. For example, is a sample of 100 students is to be selected from a population of 100 students, and then each student has a 10% chance of being selected.

##### **Features:**

1. Systematic and objective method
2. Accurate and reliable
3. Eliminating human biases.
4. Commonly used
5. Mechanical method

##### **Types/Methods:**

The various probability methods are explained as follows:

### 1. Simple random sampling:

This is the most popular method which is normally followed to collect research data. This technique provides every element or unit (such as household) an equal chance of being selected in the sample.

There are two sub- methods:

- a) Lottery methods
- b) Random tables

### 2. Systematic sampling:

It is a variation of sample random method. This technique is superior to simple random method. In this case, a sampling fraction is calculated based on the total population under study. After the sampling fraction is calculated, a certain number is selected at random and the process continues to select the sample size.

Formula: Skip Interval =  $\frac{\text{population size}}{\text{sample size}}$

### 3. Cluster sampling:

Cluster sampling is also called as 'Area sampling'. Under this method, instead of selecting individual units, the researcher divides the population into clusters or groups and accordingly sample is selected.

### 4. Stratified sampling:

The population is divided into various strata or segments based on income, occupation, age, religion, gender, etc. After dividing the population into different segments, certain number of sample members is selected from each stratum or segment. There are two types of stratified sampling:

- Proportionate stratified sampling
- Disproportionate stratified sampling

## II. Non-Probability Methods:

### 1. Convenience Sampling:

The sample is selected as per the convenience of the researcher. Under this method, every element of the population does not get a chance of being selected.

### 2. Accidental Sampling:

The researcher may select the sample by chance without following a systematic procedure. Every element of the universe does not get a chance of being selected.

### 3. Judgement Sampling:

The sample is selected based on the judgement of the researcher or some other person assisting or advising the researcher.

### 4. Quota Sampling:

Under this method, the researcher allocates certain quote (relating to respondents) to certain groups under study. The quotas may differ from each area depending upon certain factors like age, occupation, income, etc.

## Other Non-Probability Methods:

### **Sequential sampling method:**

Sequential sampling is a non-probability sampling technique wherein the researcher picks a single or a group of subjects in a given time interval, conducts his study, analyses the results then picks another group of subjects if needed and so on.

This sampling technique gives the researcher limitless chances of fine tuning his research methods and gaining a vital insight into the study that he is currently pursuing.

### **Snowball Sampling /Chain referral sampling:**

Snowball sampling is a non-probability sampling technique that is used by researchers to identify potential subjects in studies where subjects are hard to locate.

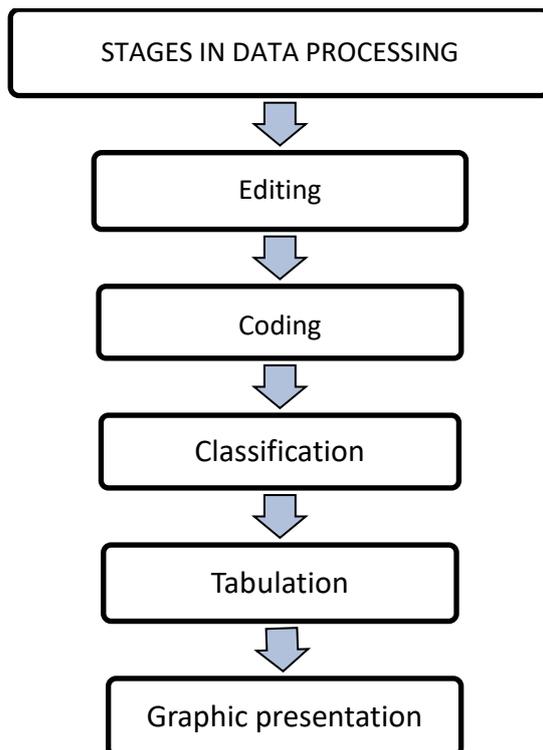
The process of snowball sampling is much like asking your subjects to nominate another person with the same trait as your next subject. The researcher then observes the nominated subjects and continues in the same way until the sufficient number of subjects is obtained.

## **STAGES IN DATA PROCESSING**

Processing of data is a process of editing, coding, classification, tabulation and graphic presentation of data. Data processing is required for the purpose of analysis and interpretation. Analysis and interpretation of data facilitates:

- **Research Findings**
- **Drawing Conclusions**
- **Recommendations**

The various stages in data processing are as follows:



### **1. EDITING OF DATA**

Editing is the process of checking errors and omissions in data collection, and making corrections, if required. Editing is required when:

- There is inconsistency in responses given by the respondents.
- Respondents may provide incorrect or false responses.
- Some vague/incomplete answers given by the respondents.
- No responses are provided by the respondents for certain questions.

#### **Types of Editing**

- a) Field Editing
- b) Central Editing

## **Essentials of Editing**

1. Accuracy
2. Avoid bias
3. Consistency
4. Completeness
5. Training to editors
6. Reliability
7. Uniformity
8. Economical

## **II. CODING OF DATA**

It is a process of assigning codes to the various statements or questions in the questionnaire. Coding is specially required when the sample size is large and there is large amount of data collection from respondents. Coding facilitates proper tabulation and analysis of data.

### **Types of codes**

- **Numerical codes such as code1, code2, code3 and so on.**
- **Alphabetical codes such as code A, code B, code C and so on.**
- **Alpha-Numerical codes such as code A<sub>1</sub>, A<sub>2</sub>, A<sub>3</sub>, B<sub>1</sub>, B<sub>2</sub>, B<sub>3</sub>, etc.**

### **Steps in Coding**

#### **1. Specify the Categories :**

The researcher or data processor must specify the categories into which the responses can be classified. For instance, the categories may include: Age, gender, Education, income, etc.

#### **2. Allocate Individual Codes:**

The researcher must allocate individual codes to each category of responses. For example: Males in the four age groups may be allocated codes as follows: M<sub>1</sub>, M<sub>2</sub>, M<sub>3</sub>, and M<sub>4</sub>.

## **III. CLASSIFICATION OF DATA**

It is the process of grouping of collected data into different categories. There for coding is an element of classification. The classification can be according to different categories: Age Group Wise, Gender Wise, Educational Level Wise, Income Group Wise, Occupation Wise, etc.

Each of categories can be further divide into sub-group – For example: the Age group can be further divided into different categories such as: Children, Teenagers, Young Adults, Middle Aged, and Senior Citizens.

## **IV. TABULATION OF DATA**

It refers to transferring the classified data in a tabular format for the purpose of analysis and interpretation. It involves sorting of data into different categories and counting the number of responses that belong to each category.

### **Variables in Tabulation**

1. Univariate
2. Multi-Variate

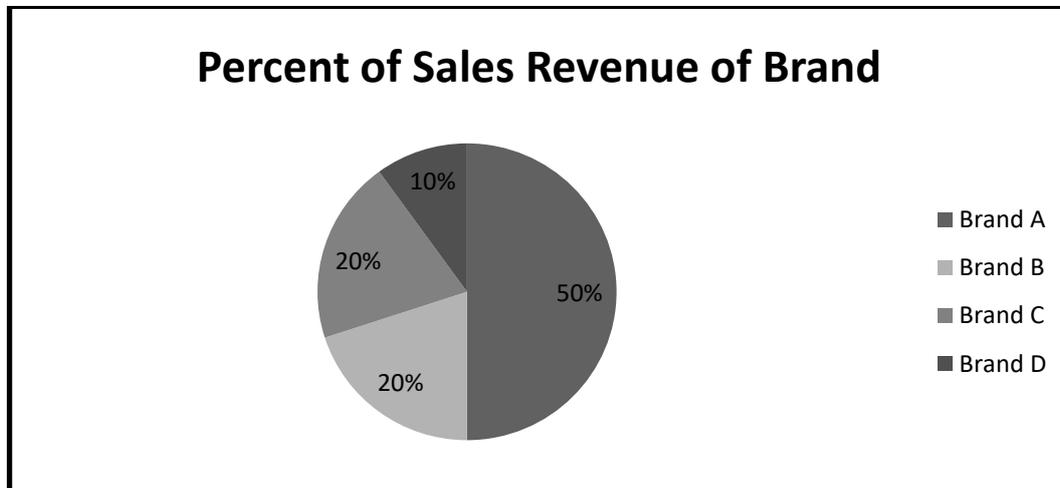
### **Methods of Tabulation**

1. Manual Tabulation
2. Mechanical Tabulation

## **V. GRAPHIC PRESENTATION OF DATA**

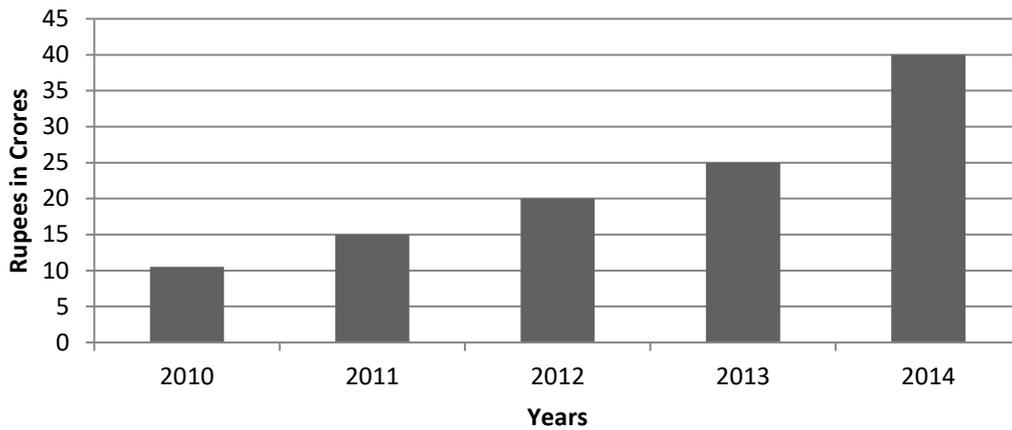
The research data needs to be presented effectively for quick and clear understanding. Bar graphs, pie charts, line graphs, histograms and other pictorial devices are an excellent means to present the data.

### **1. Pic Chat:**

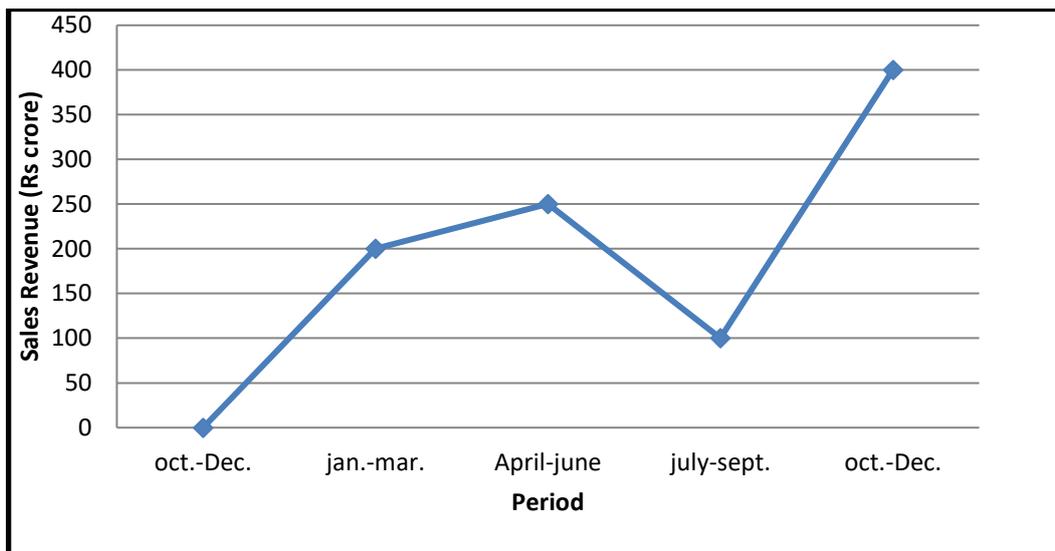


### **2. Bar Graphs:**

### Simple Bar Diagram



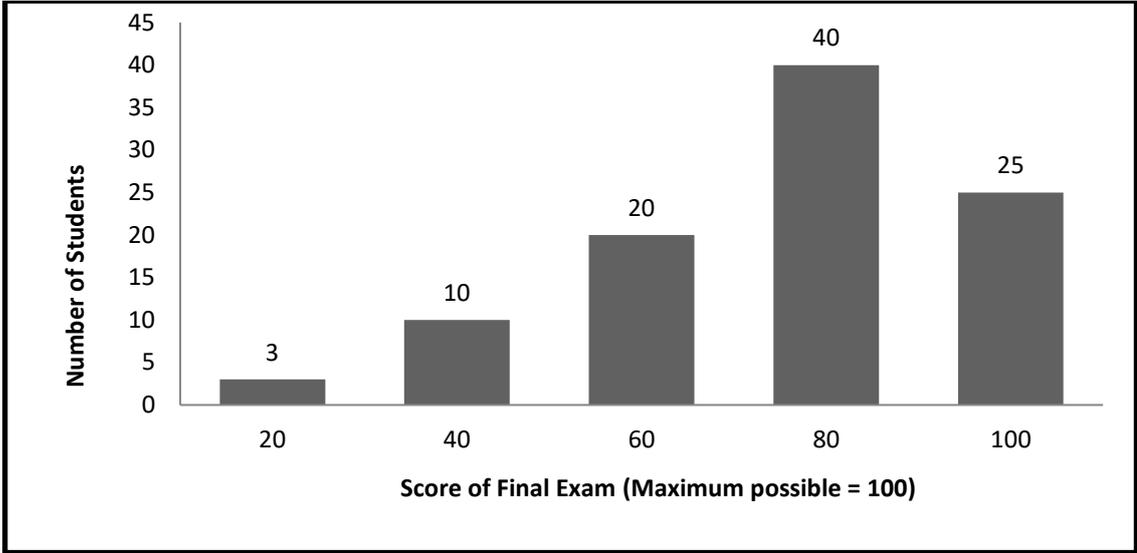
### 3. Line Graphs:



**4. Gantt chart:** A Gantt chart is a type of bar chart, developed by Henry Gantt in the 1910s, that illustrates a project schedule. For instance, a Gantt chart may consist of two horizontal or vertical bars for each period of time/activity. One bar indicates the planned/anticipated performance, and the other bar indicates the actual performance.

### 5. Histograms:

A histogram is a special kind of bar graph where the intervals are equal. In statistics, a histogram is a graphical representation of the distribution of data. It is an estimate of the probability distribution of a continuous variable and was first introduced by Karl Pearson.



# **MEASURES OF CENTRAL TENDENCY**

## **Meaning and Definitions**

It is difficult to remember the value of all items in the data. Therefore, it is better to have a value of a single item which is the representative of all the items in the data. Such a single value is called an **average or a measure of central tendency**. Thus, a single value within the range of the data that is used to reveal the general tendency and to represent the entire data is known as a measure of central tendency.

George Simpson and Fritz Kafka state that “**a measure of central tendency is a typical value around which the order values congregate**”.

## **Objectives of Averaging**

1. To arrive at a single value which is representative of the characteristics of the entire mass of data
2. To get a bird's eye view about a phenomenon under study
3. To facilitate comparison
4. To trace relationships
5. To know about the universe from a sample

## **Characteristics of a Good Measure of Central Tendency**

1. It should be simple to calculate and easy to understand
2. It should be rigidly defined
3. It should be based on all the observations
4. It should not be affected by extreme items
5. It should be capable of further algebraic treatment
6. It should have sampling stability
7. It can be easily calculated in the case of distributions containing open end class-intervals
8. It should be in the form of a mathematical formula

### **BIBLIOGRAPHY**

1. TEXT BOOK OF RESEARCH METHODOLOGY IN ACCOUNTING AND FINANCE,  
(AUTHOR: MICHAEL VAZ, PUBLISHER: MANAN PRAKASHAN)