

BUSINESS STATISTICS

UNIT-I

Meaning and Definition of Statistics

The term **Statistics** is derived from the Latin word “*Status*” meaning **state** or **condition**. Originally, statistics was used for collecting data related to the state, such as population, revenue, and military strength. Over time, its scope expanded and today statistics is widely used in **business, economics, management, social sciences, and research**.

In simple terms, **Statistics deals with the collection, organization, presentation, analysis, and interpretation of numerical data** for decision-making.

Definitions of Statistics

According to A.L. Bowley, “Statistics are numerical statements of facts in any department of inquiry, placed in relation to each other.”

Statistics in Two Senses

1. **Statistics as Data** – Numerical facts (e.g., sales figures, profits, production data)
2. **Statistics as a Method** – Scientific techniques used to analyze data

Characteristics of Statistics

Statistics has certain unique characteristics that differentiate it from general numerical information.

1. Statistics are Aggregates of Facts

Statistics deals with **mass data** and not individual cases. A single observation does not represent statistics.

2. Statistics are Numerically Expressed

Only quantitative data can be studied statistically. Qualitative aspects must be converted into numbers before analysis.

3. Statistics are Affected by Multiplicity of Causes

Statistical data is influenced by various factors simultaneously. For example, sales depend on price, advertising, consumer income, competition, and economic conditions.

4. Statistics are Collected for a Definite Purpose

Data collection must have a clear objective. Without a purpose, data becomes meaningless.

5. Statistics are Collected Systematically

Scientific methods and standardized procedures must be used while collecting data.

6. Statistics are Capable of Comparison

Statistical data should be comparable over time, between regions, or among different groups.

7. Statistics are Approximate

Statistical conclusions are **not exact** and are true only on average.

Importance of Statistics in Business Management

1. Decision Making

Managers face uncertainty. Statistics helps analyze alternatives and select the best course of action.

2. Planning and Policy Formulation

Past data and trends assist managers in framing future policies and strategies.

3. Forecasting

Statistical tools help predict:

- Sales
- Demand
- Production
- Profits

4. Market Research

Statistics helps analyze consumer behavior, preferences, and satisfaction levels.

5. Quality Control

Statistical quality control techniques help minimize defects and maintain standards.

6. Performance Evaluation

Employee productivity and organizational performance can be measured statistically.

7. Risk Analysis

Probability and statistical analysis help in assessing and minimizing risks.

Scope of Statistics in Business Management

1. **Production Management** – Cost control, inventory management
2. **Marketing Management** – Demand analysis, pricing decisions
3. **Financial Management** – Budgeting, investment analysis
4. **Human Resource Management** – Wage analysis, performance appraisal
5. **Operations Management** – Process control, productivity analysis
6. **Economics and Business Research** – Data analysis and hypothesis testing

Major Limitations

1. **Statistics deals only with quantitative data**

Qualitative factors like honesty, morale, and leadership cannot be measured directly.

2. **Statistics does not give exact results**

Conclusions are approximate and based on probability.

3. **Statistics can be misused**

Wrong interpretation or manipulation of data can lead to misleading results.

4. **Statistics depends on the quality of data**

Poor data collection leads to inaccurate conclusions.

5. **Statistics is not a substitute for judgment**

Statistical results require managerial judgment and experience.

6. **Statistics does not reveal the whole truth**

It only highlights numerical aspects of a problem.

7. **Common people may misunderstand statistics**

Lack of statistical knowledge may lead to incorrect decisions.

Introduction to Sampling and Its Methods

Meaning of Sampling

Sampling is the process of selecting a **small group (sample)** from a **large group (population)** to represent the entire population.

Population and Sample

- **Population (Universe):** The entire group under study
- **Sample:** A part of the population selected for analysis

Need for Sampling

1. Saves time and cost
2. More practical for large populations
3. Better accuracy in many cases
4. Suitable for destructive testing

5. Enables quick decision-making

Methods of Sampling

I. Probability Sampling Methods

In probability sampling, **each unit of the population has a known and equal chance of being selected**. These methods are scientific, objective, and widely used in business and research studies.

1. Simple Random Sampling

Meaning

Simple Random Sampling is a method in which every unit in the population has an **equal and independent chance of selection**.

Methods of Selection

- 1. Lottery Method**

- Each unit is written on a slip and drawn randomly.

- 2. Random Number Table**

- Random numbers are used to select sample units.

Merits

- Simple and unbiased
- Easy to understand
- Suitable for homogeneous populations

Demerits

- Not suitable for large populations
- Time-consuming
- Requires complete population list

Example

Selecting 50 employees randomly from a company of 500 employees.

2. Stratified Sampling

Meaning

In stratified sampling, the population is divided into **homogeneous groups called strata** based on characteristics like age, income, or gender. Samples are then drawn from each stratum.

Types of Stratified Sampling

- 1. Proportionate Stratified Sampling**
 - Sample size proportional to stratum size
- 2. Disproportionate Stratified Sampling**
 - Sample size not proportional

Merits

- More representative
- Higher accuracy
- Suitable for heterogeneous populations

Demerits

- Complex method
- Requires detailed population information

Example

Dividing employees into departments and selecting samples from each department.

3. Systematic Sampling

Meaning

Systematic sampling involves selecting units at **regular intervals** from an ordered list.

Procedure

1. Calculate sampling interval:

$$k=N/n$$

2. Select a random starting point
3. Select every k th unit

Merits

- Simple and quick
- Economical
- Easy to apply

Demerits

- Risk of periodic bias
- Not truly random

Example: Selecting every 10th customer from a customer list.

4. Cluster Sampling

Meaning

The population is divided into **clusters** (geographical or natural groups). Entire clusters are randomly selected.

Merits

- Cost-effective
- Suitable for large and scattered populations

Demerits

- Less accurate
- Higher sampling error

Example

Selecting schools (clusters) and surveying all students in selected schools.

II. Non-Probability Sampling Methods

In non-probability sampling, **the probability of selection is unknown**. These methods are subjective and commonly used in exploratory research.

1. Convenience Sampling

Meaning

Samples are selected based on **ease of access**.

Merits

- Easy and fast
- Low cost

Demerits

- Highly biased
- Low reliability

Example: Surveying customers who visit a store on a particular day.

2. Judgment Sampling

Meaning

The researcher selects samples based on **personal judgment and expertise**.

Merits

- Useful for specialized studies
- Saves time

Demerits

- Subjective bias
- Not representative

Example: Selecting experienced managers for expert opinions.

3. Quota Sampling

Meaning

The population is divided into groups, and **fixed quotas** are assigned to each group.

Merits

- Ensures representation
- Easy to administer

Demerits

- Selection bias
- No statistical validity

Example: Selecting 50 male and 50 female respondents.

4. Snowball Sampling

Meaning

Initial respondents refer additional respondents, forming a **chain**.

Merits

- Useful for hidden populations
- Cost-effective

Demerits

- Bias due to referrals
- Limited generalization

Example: Research on drug users or rare professionals.

Advantages of Sampling

- Economical
- Time-saving
- Feasible for large populations
- Flexible and practical

Limitations of Sampling

- Possibility of sampling errors
- Requires expert knowledge
- Biased samples may lead to wrong conclusions

UNIT-II

Meaning and Definition of Data

The word **Data** is derived from the Latin word “*Datum*” which means **something given**. Data refers to **raw facts and figures** collected for analysis. By itself, data has **no meaning**, but when processed and analyzed, it becomes useful information for decision-making.

In statistics, data forms the **foundation of analysis**, interpretation, and conclusions.

Definitions of Data

According to **Oxford Dictionary**, “Data are facts or information used usually to calculate, analyze or plan something.”

Types of Data

Data can be classified on different bases:

A. Based on Source

1. **Primary Data**
2. **Secondary Data**

B. Based on Nature

1. **Qualitative Data**
 - Non-numerical data (e.g., opinions, satisfaction, attitude)
2. **Quantitative Data**
 - Numerical data
 - **Discrete Data** – Countable (number of employees)
 - **Continuous Data** – Measurable (height, weight, income)

C. Based on Time

1. Cross-Sectional Data

- Data collected at one point in time

2. Time Series Data

- Data collected over a period of time

Primary Data

Meaning of Primary Data

Primary data is **original data** collected **first-hand** by the investigator for a **specific purpose**

Methods of Primary Data Collection

1. Direct Personal Interview

Meaning

The investigator personally meets respondents and collects information.

Merits

- Accurate and reliable
- Suitable for intensive studies

Demerits

- Time-consuming
- Costly

2. Indirect Oral Investigation

Information is collected from knowledgeable persons when direct respondents are not available.

Merits

- Useful for sensitive issues

- Saves time

Demerits

- Chances of bias
- Less accuracy

3. Questionnaire Method

A list of questions is sent to respondents to be filled in.

Merits

- Economical
- Suitable for large areas

Demerits

- Low response rate
- Misinterpretation of questions

4. Schedule Method

Questionnaires are filled by trained enumerators through interviews.

Merits

- High response rate
- More accurate

Demerits

- Expensive
- Requires trained staff

5. Observation Method

Data is collected by observing events or behavior.

Merits

- Real-time data
- No respondent bias

Demerits

- Limited scope
- Observer bias

6. Telephonic / Online Surveys

Data collected using phone calls or digital platforms.

Merits

- Fast and economical
- Wide reach

Demerits

- Limited interaction
- Sampling bias

Secondary Data

Meaning of Secondary Data

Secondary data refers to data **already collected** by someone else and used by the investigator for analysis.

Sources of Secondary Data

I. Internal Sources of Secondary Data

Internal sources are data generated **within an organization** during its normal operations.

1. Accounting Records

These include:

- Profit and Loss Account
- Balance Sheet
- Cash Flow Statement
- Cost records

Uses:

- Financial analysis
- Profitability assessment
- Cost control

Merits:

- Highly reliable
- Easily accessible

Limitations:

- Limited to the organization
- Not useful for market analysis

2. Sales Records

Includes:

- Sales invoices
- Monthly/annual sales reports

- Customer purchase records

Uses:

- Sales forecasting
- Market trend analysis
- Customer behavior study

3. Production Records

Contains information related to:

- Output levels
- Machine utilization
- Defect rates
- Inventory levels

Uses:

- Production planning
- Efficiency analysis

4. Personnel Records

Includes:

- Employee attendance
- Wages and salaries
- Performance appraisal reports
- Employee turnover data

Uses:

- HR planning
- Productivity analysis

5. Past Research Reports

Internal market surveys, feasibility studies, and project reports conducted earlier.

Merits of Internal Sources

- Low cost
- Accurate
- Relevant to business decisions

Limitations of Internal Sources

- Limited scope
- Confidentiality issues

II. External Sources of Secondary Data

External sources are those **outside the organization** and can be further divided into **Published** and **Unpublished** sources.

A. Published Sources of Secondary Data

1. Government Publications

Government agencies are the **most important and reliable sources** of secondary data.

Examples:

- Census of Population
- Economic Survey
- National Sample Survey (NSS)
- Statistical Abstracts
- Reserve Bank of India (RBI) reports
- Ministry of Finance publications

Uses:

- Economic planning
- Market potential analysis
- Policy formulation

Merits:

- Reliable and authentic
- Large coverage

Limitations:

- Time lag
- May not suit specific research needs

2. Semi-Government and Public Corporations

Includes data published by:

- Public sector undertakings
- Development banks
- Regulatory authorities

Examples:

- SEBI reports
- NABARD publications
- World Bank and IMF reports

3. International Organizations

These organizations publish global data for international comparisons.

Examples:

- United Nations (UN)
- World Bank
- International Monetary Fund (IMF)
- World Health Organization (WHO)
- International Labour Organization (ILO)

Uses:

- Global economic studies
- Comparative research

4. Trade Associations and Chambers of Commerce

Includes:

- Industry reports
- Market surveys
- Trade journals

Examples:

- Confederation of Indian Industry (CII)
- Federation of Indian Chambers of Commerce and Industry (FICCI)

Uses:

- Industry analysis
- Competitive studies

5. Research Institutions and Universities

- Research papers
- Survey reports
- Academic publications

Examples:

- Universities
- Research councils
- Think tanks

6. Books, Journals, and Periodicals**Includes:**

- Textbooks
- Research journals
- Business magazines
- Newspapers

Uses:

- Theoretical background
- Market trends
- Case studies

7. Online and Digital Sources

- Government websites
- Research databases
- Company websites
- Online journals

Merits:

- Easily accessible
- Up-to-date

Limitations:

- Authenticity issues
- Data overload

B. Unpublished Sources of Secondary Data

1. Research Theses and Dissertations

Conducted by scholars and students.

2. Private Organization Records

- Market research agencies
- Consulting firm reports

3. Personal Documents

- Diaries
- Letters
- Memoirs

III. Precautions in Using Secondary Data

Before using secondary data, the researcher should ensure:

1. Reliability of source
2. Suitability of data
3. Accuracy and consistency
4. Adequacy of coverage
5. Timeliness

Merits of Secondary Data

- Economical
- Time-saving

- Easily available

Demerits of Secondary Data

- May not suit objectives
- Possibility of outdated data
- Accuracy cannot be ensured

Tabulation of Data

Meaning of Tabulation

Tabulation refers to the **systematic presentation of data in rows and columns** to make it easy to understand, compare, and analyze.

Importance of Tabulation

1. Simplifies complex data
2. Facilitates comparison
3. Saves time and space
4. Helps in statistical analysis
5. Forms basis for diagrams and graphs

Types of Tabulation

1. Simple (One-Way) Table

Shows data related to one characteristic.

2. Double (Two-Way) Table

Shows data related to two characteristics.

3. Multiple Table

Shows data related to more than two characteristics.

4. Complex Table

Contains detailed information with multiple variables.

Components of a Statistical Table

A statistical table is a **systematic arrangement of data in rows and columns**. For a table to be **complete, clear, and meaningful**, it must contain certain essential components.

1. Table Number

- Each table should be assigned a **distinct number**.
- It helps in **easy identification and reference**.
- Usually written at the **top of the table**.

Example:

Table 1: Sales of ABC Company

2. Title of the Table

- The title explains **what the table is about**.
- It should be **brief, clear, and self-explanatory**
- It usually mentions **subject, place, and time period**.

Example:

Sales of ABC Company in India during 2024

3. Head note (Unit of Measurement)

- Indicates the **unit in which data is expressed**.
- Placed just **below the title**.
- Avoids repetition of units in the table.

Example:

(Figures in ₹ Lakhs)

4. Captions (Column Headings)

- Captions refer to the **headings of columns**.
- They describe **what each column represents**.
- Placed at the **top of each column**.

Example:

Year, Production, Sales, Profit

5. Stubs (Row Headings)

- Stubs are the **headings of rows**.
- Placed in the **leftmost column**.
- Describe the **items or categories** listed in rows.

Example:

January, February, March

6. Body of the Table

- The body contains the **actual numerical data**.
- It lies at the **intersection of rows and columns**.
- It is the **most important part** of the table.

7. Footnotes

- Provide **additional explanations or clarifications**.
- Used when certain figures need special notes.
- Placed at the **bottom of the table**.

Example:

Note: Data for March is provisional.

8. Source Note

- Indicates the **source from which data is obtained**.
- Important for **authenticity and reliability**.
- Placed **below the footnote**.

Example:

Source: Company Annual Report

9. Totals and Sub-Totals (If Required)

- Used to show **summation of data**.
- Helps in comparison and analysis.
- Usually placed at the **end of rows or columns**.

Presentation of Data

Presentation of data means displaying data in a **visual and understandable form** using diagrams and graphs.

A. Types of Diagrams

1. Bar Diagrams

- Simple bar
- Multiple bar
- Component bar
- Percentage bar

Merits

- Easy to understand
- Useful for comparison

Demerits

- Limited information
- Not suitable for large data

2. Pie Diagram

Circular diagram showing data in proportion.

Merits

- Attractive
- Shows relative importance

Demerits

- Difficult for precise comparison
- Not suitable for many variables

3. Pictograms

Use symbols or pictures.

Merits

- Easy for laymen
- Visually appealing

Demerits

- Lack of accuracy
- Time-consuming

B. Types of Graphs

1. Line Graph

Shows trend over time.

Merits

- Best for time series
- Easy trend analysis

Demerits

- Not suitable for categorical data

2. Histogram

Shows frequency distribution of continuous data.

3. Frequency Polygon

Shows Graphical representation of frequency distribution.

4. Ogives (Cumulative Frequency Curves)

- Less than ogive
- More than ogive

Merits of Graphical Presentation

- Easy understanding
- Quick comparison
- Attractive presentation

Demerits of Graphical Presentation

- Lack of precision
- Requires skill
- Can be misleading

UNIT-III

Meaning and Definition of Central Tendency

Central tendency refers to the **single value** that represents or summarizes a large set of data by indicating its **central or typical value**. It gives an idea about where the data values are concentrated.

In statistics, measures of central tendency help in **condensing large data** into a meaningful and representative figure.

Definitions of Central Tendency

1. Clark

“An average is a figure that represents the typical value of a series.”

2. Croxton and Cowden

“Measures of central tendency are the statistical measures which describe the central position of a distribution.”

Essentials of a Good Average

A good average should possess the following qualities:

1. Rigidly Defined

The average should be clearly and precisely defined so that different people get the same result.

2. Easy to Understand and Calculate

A good average should be simple to compute and easy to interpret.

3. Based on All Observations

It should take into account all data values to be truly representative.

4. Not Unduly Affected by Extreme Values

A good average should not be influenced too much by very large or very small values.

5. Capable of Further Mathematical Treatment

It should be usable for further statistical analysis like variance and standard deviation.

6. Stable

A good average should show minimal variation from sample to sample.

7. Representative

It should reflect the true center of the data.

3. Measures of Central Tendency

The most commonly used measures are:

1. **Arithmetic Mean**
2. **Median**
3. **Mode**

A. Arithmetic Mean (AM)

Meaning

Arithmetic Mean is the **sum of all observations divided by the number of observations**. It is the most widely used measure of central tendency.

Formula

For Individual Series

$$M = \Sigma X / N$$

Where, M = Mean ΣX = Summation of scores in the distribution N = Total number of scores.

Let us now compute mean with the help of an example

The scores obtained by 10 students on psychology test are as follows: 58 34 32 47 74 67 35 34 30 39

Step 1: In order to obtain mean for the above data we will first add the marks to obtain ΣX : $58 + 34 + 32 + 47 + 74 + 67 + 35 + 34 + 30 + 39 = 450$

Step 2: Now using the formula, we will compute mean $M = \Sigma X / N$ $\Sigma X = 450$, $N = 10$ (Total number of students) Thus, $M = 450 / 10 = 45$ Thus, the mean obtained for the above data is 45

Advantages of Mean:

- 1) The definition of mean is rigid which is a quality of a good measure of central tendency.
- 2) It is not only easy to understand but also easy to calculate.
- 3) All the scores in the distribution are considered when mean is computed.
- 4) Further mathematical calculations can be carried out on the basis of mean.
- 5) Fluctuations in sampling are least likely to affect mean.

Limitations of Mean

- 1) Outliers or extreme values can have an impact on mean.
- 2) When there are open ended classes, such as 10 and above or below 5, mean cannot be computed. In such cases median and mode can be computed. This is mainly because in such distributions mid-point cannot be determined to carry out calculations.

3) If a score in the data is missing or lost or not clear, then mean cannot be computed unless mean is computed for rest of the data by not considering the lost score and dropping it all together.

4) It is not possible to determine mean through inspection. Further, it cannot be determined based on a graph.

5) It is not suitable for data that is skewed or is very asymmetrical as then in such cases mean will not adequately represent the data.

Median

Meaning

Median is the **middle value** of a data set when the observations are arranged in **ascending or descending order**.

Calculation

- If **odd number** of observations:
Median = Middle value
- If **even number** of observations:
Median = Average of two middle values

Advantages of Median

1) The definition of median is rigid which is a quality of a good measure of central tendency.

2) It is easy to understand and calculate.

3) It is not affected by outliers or extreme scores in data.

4) Unless the median falls in an open ended class, it can be computed for grouped data with open ended classes.

5) In certain cases it is possible to identify median through inspection as well as graphically.

Limitations of Median

1) Some statistical procedures using median are quite complex. Computation of median can be time consuming when large data is involved because the data needs to be arranged in an order before median is computed.

2) Median cannot be computed exactly when an ungrouped data is even. In such cases, median is estimated as mean of the scores in the middle of the distribution.

3) It is not based on each and every score in the distribution.

4) It can be affected by sampling fluctuations and thus can be termed as less stable than mean

C. Mode

Meaning

Mode is the **value that occurs most frequently** in a data set.

Types of Mode

- **Unimodal** – One mode
- **Bimodal** – Two modes
- **Multimodal** – More than two modes

Merits of Mode

1. Easy to locate
2. Not affected by extreme values
3. Suitable for qualitative data

Demerits of Mode

1. Not rigidly defined

2. Not based on all observations
3. Not suitable for further mathematical treatment

Business Applications

- Most popular product size
- Most common customer preference
- Fashion and marketing studies

4. Comparison of Mean, Median, and Mode

Basis	Mean	Median	Mode
Definition	Average	Middle value	Most frequent
Based on all values	Yes	No	No
Effect of extreme values	High	No	No
Mathematical treatment	Yes	Limited	No
Best used for	Symmetrical data	Skewed data	Popular value

5. Business Applications of Measures of Central Tendency

Measures of central tendency are widely used in business management for:

1. Sales Analysis

- Average sales help in forecasting and planning.

2. Cost and Profit Analysis

- Mean cost and profit help in pricing decisions.

3. Wage and Salary Administration

- Median wage avoids the effect of extreme salaries.

4. Market Research

- Mode identifies the most preferred product or service.

5. Financial Planning

- Average returns help in investment decisions.

6. Performance Measurement

- Average productivity helps evaluate employee efficiency.

UNIT IV

Introduction and Meaning of Dispersion

Introduction

While measures of central tendency indicate the **average or central value** of a dataset, they do not show how the data values are **spread or scattered** around that central value. Two datasets may have the same average but differ greatly in variability. To overcome this limitation, **measures of dispersion** are used.

Dispersion helps in understanding the **degree of variation, consistency, and reliability** of the data.

Meaning of Dispersion

Dispersion refers to the **extent to which data values differ from one another and from the central value**. It shows whether the data are **closely clustered** or **widely scattered**.

Definitions of Dispersion

1. Spiegel

“Dispersion is the degree of spread or variation of a set of observations.”

2. Business Interpretation

Dispersion indicates the **stability or risk** associated with business data such as sales, profits, or returns.

4.2 Objectives of Dispersion

The main objectives of studying dispersion are:

1. To Measure Variability

Dispersion shows how much the data values deviate from the average.

2. To Judge Reliability of Average

A smaller dispersion indicates that the average is more reliable and representative.

3. To Compare Two or More Series

Dispersion helps compare consistency between datasets (e.g., sales of two companies).

4. To Study Risk and Uncertainty

In business and finance, higher dispersion indicates higher risk.

5. To Aid Decision Making

Managers use dispersion to evaluate performance stability.

6. To Understand Data Distribution

Dispersion helps in identifying whether data is concentrated or spread out.

Measures of Dispersion

Measures of dispersion are broadly classified into **absolute** and **relative** measures. The commonly used **absolute measures** are:

1. Range
2. Quartile Deviation
3. Mean Deviation
4. Standard Deviation
5. Variance

A. Range

Meaning

Range is the **difference between the highest and lowest values** in a dataset.

$$\text{Range} = L - S$$

Coefficient of Range

$$\text{Co-efficient of Range} = \frac{L - S}{L + S}$$

Merits

- Simple to calculate
- Easy to understand
- Useful for quick comparisons

Demerits

- Based only on two extreme values
- Highly affected by extreme values
- Not reliable

Business Applications


- Stock price fluctuations
- Weather data
- Preliminary data analysis

B. Quartile Deviation (Semi-Interquartile Range)

Meaning

Quartile Deviation measures the **spread of the middle 50% of data**. It is half the difference between the third and first quartiles.

Coefficient of Quartile Deviation


$$\text{Quartile Deviation} = \frac{(Q_3 - Q_1)}{2}$$
$$\text{Coefficient of Quartile Deviation} = \frac{(Q_3 - Q_1)}{(Q_3 + Q_1)}$$

Merits

- Not affected by extreme values
- Suitable for skewed distributions
- Useful for open-ended distributions

Demerits

- Ignores 50% of the data
- Not suitable for algebraic treatment

Business Applications

- Income distribution
- Wage analysis
- Housing price studies

C. Mean Deviation

Meaning

Mean Deviation is the **average of the absolute deviations** of observations from a central value (Mean or Median).

Formula: Coefficient of Mean Deviation

Computation of Mean Deviation :	
Individual Series	Discrete & continuous series
a) Deviation from Mean : $\delta\bar{X} = \frac{\sum d\bar{x} }{N}, d\bar{x} = X - \bar{X} $ Coefficient of $\delta\bar{X} = \frac{\delta\bar{X}}{\bar{X}}$	a) Deviation from Mean : $\delta\bar{X} = \frac{\sum f d\bar{x} }{N}$
b) Mean Deviation from Median : $\delta M = \frac{\sum dM }{N}, dM = X - M $ Coefficient of $\delta M = \frac{\delta M}{M}$	b) Mean Deviation from Median : $\delta M = \frac{\sum f dM }{N}$
c) Mean Deviation from Mode : $\delta Z = \frac{\sum dZ }{N}, dZ = X - Z $ Coefficient of $\delta Z = \frac{\delta Z}{M}$	b) Mean Deviation from Mode : $\delta M = \frac{\sum f dZ }{N}$

Merits

- Based on all observations
- Less affected by extreme values (when calculated from median)

Demerits

- Absolute values limit mathematical use
- Rarely used in advanced analysis

Business Applications

- Price fluctuations
- Consumer spending analysis

D. Standard Deviation

Meaning

Standard Deviation measures the **average dispersion of values from the arithmetic mean**, expressed as the square root of variance.

Formula

<u>Sample Variance</u>	<u>Sample Standard Deviation</u>
$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$	$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$

Merits

- Most reliable measure
- Based on all observations
- Suitable for further statistical analysis

Demerits

- Complex to calculate
- Affected by extreme values

Business Applications

- Risk analysis
- Quality control
- Investment decisions

E. Variance

Meaning

Variance is the **average of squared deviations** from the mean.

Merits

- Foundation for advanced statistical techniques
- Widely used in economics and finance

Demerits

- Expressed in squared units
- Less intuitive

Business Applications

- Portfolio risk measurement
- Performance evaluation

UNIT V

Meaning of Index Number

Meaning

An **Index Number** is a **statistical measure** used to show changes in a variable or group of related variables over a period of time or across different places. It measures **relative change** with respect to a **base period**, which is usually taken as 100.

Index numbers are widely used to study changes in:

- Prices
- Quantities
- Values
- Cost of living
- Production and sales

Definitions of Index Number

1. Croxton and Cowden

“An index number is a device for measuring differences in the magnitude of a group of related variables.”

2. Bowley

“Index numbers are specialized averages designed to measure the change in a group of related variables over time.”

3. Business Interpretation

Index numbers help managers compare present performance with past performance.

Base Year and Current Year

- **Base Year:** The year used for comparison (index = 100)
- **Current Year:** The year whose change is measured

Advantages of Index Numbers

1. Measure of Economic Changes

Index numbers help measure changes in prices, production, and standard of living.

2. Simplification of Complex Data

Large and complex data is summarized into a single figure.

3. Useful for Comparison

Index numbers help compare data over different periods or regions.

4. Basis for Decision Making

Used by government, businesses, and economists for policy and planning.

5. Cost of Living Measurement

Helps in determining wage adjustments and dearness allowance.

6. Forecasting

Index numbers help in predicting future trends.

Disadvantages (Limitations) of Index Numbers

1. Approximate Results

Index numbers provide only approximate measurements.

2. Selection Bias

Improper selection of items, weights, or base year leads to misleading results.

3. Not Completely Reliable

Depends heavily on data quality and method used.

4. Changes in Quality

Quality changes in goods are difficult to measure accurately.

5. Subjective Nature

Personal judgment in selection of items and weights can affect results.

6. Limited Scope

Index numbers show relative changes but not absolute changes.

Methods of Calculating Index Numbers

Index numbers can be calculated using different methods. Broadly, they are classified into:

1. **Simple Index Numbers**
2. **Weighted Index Numbers**

A. Simple Index Number Methods

1. Simple Aggregate Method

Formula

$$P_{01} = \frac{\sum P_1}{\sum P_0} \times 100$$

Where:

P1= Prices in current year

P0 = Prices in base year

Merits

- Simple to calculate
- Easy to understand

Demerits

- Ignores relative importance of items
- Units of measurement may differ

2. Simple Average of Price Relatives

Price Relative Formula

Price Relative= $P_1/P_0 \times 100$

Simple average of Price Relatives

$$P_{01} = \frac{\sum \left(\frac{P_1}{P_0} \times 100 \right)}{N}$$

Merit

- Removes unit problem
- More representative than aggregate method

Demerits

- Ignores weights

- Time-consuming

B. Weighted Index Number Methods

Weighted methods consider **relative importance (weights)** of items.

1. Weighted Aggregate Method

(a) Laspeyre's Price Index

Formula

$$\text{Laspeyres Index} = \frac{\sum (\text{Observation Price} \times \text{Base Qty})}{\sum (\text{Base Price} \times \text{Base Qty})}$$

Merits

- Easy to compute
- Uses base year quantities

Demerits

- Overstates price rise
- Ignores substitution effect

(b) Paasche Price Index

Formula

$$\text{Paasche Price Index Formula} = \text{Sum} \frac{\text{Observation Price} \times \text{Observation Qty}}{\text{Base Price} \times \text{Observation Qty}}$$

Merits

- Uses current year quantities
- More realistic

Demerits

- Data collection difficult
- Understates price rise

(c) Fisher's Ideal Index

Formula

$$\text{Fisher's Index Number} = \sqrt{\frac{\sum P_1q_0}{\sum P_0q_0} \times \frac{\sum P_1q_1}{\sum P_0q_1}}$$

P_1 = Price of current year
 P_0 = Price of base year
 q_1 = Quantity of current year
 q_0 = quantity of base year

Merits

- Most accurate
- Satisfies time and factor reversal tests

Demerits

- Complex
- Data intensive

2. Weighted Average of Price Relatives

Formula

$$\text{Index} = \frac{\sum WR}{\sum W}$$

Where:

W = Weights

R = Price relatives

Uses of Index Numbers in Business

- Price comparison
- Inflation measurement
- Wage and salary revision
- Sales and production analysis
- Economic planning

Conclusion

Index numbers are powerful statistical tools for measuring changes over time. Although they have limitations, when constructed carefully, they provide valuable insights for economic and business decision-making.