SYLLABUS

Class – B.B.A. III Sem.

Subject – Business Statistics

UNIT – I	Introduction to statistics – purpose of statistics – kinds of numbers, measurements, variables, attributes, UNITs, discrete and continuous data. Frequency distributions – continuous frequency distribution – graphic representation – Histogram – frequency polygon and their uses.
UNIT – II	Measures of central tendency – Mean, Median, Mode merits and demerits – Measure of Dispersion – Range, mean deviation, standard deviation, coefficient of variations.

UNIT — I
STATISTICS

The word "Statistics" of English language has either been derived from the Latin word status or Italian word statistics and meaning of this term is "An organised political state.

Meaning: The science of collecting, analysing and interpreting such data or Numerical data relating to an aggregate of individuals.

E.g:- Statistics of National Income, Statistics of Automobile Accidents, Production Statistics, etc.

Definition: - "The classified facts relating the condition of the people in a state specially those facts which can be stated in members or in tables of members or in any tabular or classified arrangements."

-Webster

"Statistics may be regarded as (i) the study of population (ii) The study of variation (iii) The study of method of reduction of data"

-R.A. Fisher.

Nature /Features /Characteristics of statistics

- It is an aggregate of facts.
- Analysis of multiplicity of causes.
- It is numerically expressed.
- It is estimated according to reasonable standard of accuracy.
- It is collected for pre-determined purpose.
- It is collected in a systematic manner.

Division of Statistics

Theoretical	Statistical Methods	Applied

Theoretical: Mathematical theory which is the basis of the science of statistics is called theoretical statistics.

Statistical Methods: By this method we mean methods specially adapted to the elucidation of quantitative data affected by a multiplicity of causes.

Few Methods are:-

(1) Collection of Data (2) Classification (3) Tabulation (4) Presentation (5) Analysis (6) Interpretation (7) Forecasting.

Applied: - It deals with the application of rules and principles developed for specific problem in different disciplines.

Eg: - Time series, Sampling, Statistical Quality control, design of experiments.

Functions of Statistics:-

- It presents facts in a definite form.
- It simplifies mass of figures
- It facilitates comparison
- It helps in prediction
- It helps in formulating suitable & policies.

Scope of Statistics:-

1. Statistics and state or govt.
2. Statistics and business or management.
- Marketing
- Production

- Finance
- Banking
- Control
- Research and Development
- Purchases

3. Statistics and Economics
- Measures National Income
- Money Market analysis
- Analysis of competition, monopoly, oligopoly,
- Analysis of Population etc.

4. Statistics and science

5. Statistics and Research

Limitations:-
- (i)   It is not deal with items but deals with aggregates.
- (ii)   Only on expert can use it
- (iii)   It is not the only method to analyze the problem.
- (iv)   It can be misused etc.

Statistical Investigation

Meaning: In general it means as a statistical survey.

In brief. Scientific and systematic collection of data and their analysis with the help of various statistical method and their interpretation.

Stages of Statistical Investigation:-
→ Planning of Investigation
→ Collection of Data
→ Editing of Data
→ Presentation of Data
   (a) Classification
   (b) Tabulation
   (c) Diagrams
   (d) Graphs
→ Analysis of Data
→ Interrelation of Data or Report Preparation

Types of Statistical Investigation:-
1. Experiment or survey investigation
2. Complete or sample investigation
3. Official, semi-official, Non official investigation
4. Confidential or open investigation
5. General purpose and specific purpose investigation
6. Original or repetitive investigation.

PROCESS OF DATA COLLECTION

Data: - A bundle of Information or bunch of information.

Data Collection: Collecting Information for some relevant purpose & placed in relation to each other.

Types of Data:-

1. Raw Data:- When we collect data through schedules and questionnaires or some other method eg:- Classification, tabulation etc.

2. Processed Data:- When we use the above raw data for application of different methods of analysing of data. Like using correlation, Z-test, T-test on data. That will be known as processed data.

Sources of Data Collection:-

3. Internal Data: - When data is collected by problem the internal source for any specific

It purpose.
4. **External Data: -** This type of data collected by the external source.
5. **Primary Data: -** It is original and collected first time. it is like raw material and it is required large sum of money, energy and time.
6. **Secondary Data: -** Secondary data are those already in existence and which have been collected for some other purpose than answering of the question at hand.
7. **Qualitative Data: -** Which can not be measurable but only there presence and absence in a group of individual can be noted are called qualitative data.
8. **Quantitative Data: -** The characteristics which can be measured directly are known as quantitative data.

Collection of Data: - It means the methods that are to be employed for obtaining the required information from the units under investigations.

Methods of Data Collection:- (Primary Data)
- Direct Personal Interviews
- By observation
- By Survey
- By questionnaires

Difference between Primary and secondary data:-

Points	Primary Data	Secondary Data
1. **Originality**	Primary data are original i.e., collected first time.	Secondary data are not original, i.e., they are already in existence and are used by the investigator.
2. **Organisation**	Primary data are like raw material.	Secondary data are in the from of finished product. They have passed through statistical methods.
3. **Purpose**	Primary data are according to the object of investigation and are used without correction.	Secondary data are collected for some other purpose and are corrected before use.
4. **Expenditure**	The collection of primary data require large sum, energy and time.	Secondary data are easily available from secondary sources (published or unpublished).
5. **Precautions**	Precautions are not necessary in the use of primary data.	Precautions are necessary in the use of secondary data.

Preparation of Questionnaires:-
This method of data collection is quit popular, particularly in case of big enquires, it is adopted by individuals, research workers. Private and public organization and even by government also.
A questionnaires consists of number of question printed or type in a definite order on a form or set of forms. The respondents have to answer the question on their own.

Importance:-
  i. Low cost and universal
  ii. Free from biases.
  iii. Respondents have adequate time to respond
  iv. Fairly approachable

Demerits:-
(i) Low rate of return
(ii) Fill on educated respondents

(iii) Slowest method of Response

Preparation of Questionnaires: - It is considered as the heart of a survey operation. Hence it should be very carefully constructed. If it is not properly set up and carefully constructed.

Step I	:-	Prepare it in a general form.
Step II	:-	Prepare sequence of question.
Step III	:-	Emphasize on question formulation and wordings
Step IV	:-	Ask Logical and not misleading questions.
Step V	:-	Personal questions should be left to the end.
Step VI	:-	Technical terms and vague expressions should be availed classification and Tabulation of Data

Classification & Tabulation of Data

After collecting and editing of data an important step towards processing that classification. It is grouping of related facts into different classes.

Types of classification:-
   i.   **Geographical:-** On the basis of location difference between the various items. E.g. Sugar Cave, wheat, rice, for various states.
   ii. **Chronological:-** On the basis of time
        e.g.-

Year	Sales
1997	1,84,408
1998	1,84,400
1999	1,05,000

   iii. **Qualitative classification: -** Data classified on the basis of some attribute or quality such as, colour of hair, literacy, religion etc.

Population

   iv. **Quantitative Classification: -** When data is quantify on some units like height, weight, income, sales etc.

Tabulation of Data

A table is a systematic arrangement of statistical data in columns and Rows.

Part of Table:-
   1. Table number
   2. Title of the Table
   3. Caption
   4. Stub
   5. Body of the table
   6. Head note
   7. Foot Note

Types of Table:-

(i) Simple and Complex Table:-

(a) Simple or one-way table:-

Age	No. of Employees
25	10
30	7
35	12
40	9
45	6

(b) Two way Table

Age	Males	Females	Total
25	25	15	40
30	20	25	45
35	24	20	44
40	18	10	28
45	10	8	18
Total	**97**	**78**	**175**

2) General Purpose and Specific Purpose Table:- General purpose table, also known as the reference table or repository tables, which provides information for general use or reference.

Special purpose are also known as summary or analytical tables which provides information for one particular discussion or specific purpose.

METHODS OF SAMPLING

Meaning: - The process of obtaining a sample and its subsequent analysis and interpretation is known as sampling and the process of obtaining the sample if the first stage of sampling.

The various methods of sampling can broadly be divided into:
   i.   Random sampling method
   ii. Non Random sampling method

Random Sampling Method

I Simple Random Sampling: - In this method each and every item of the population is given an equal chance of being included in the sample.

(a) Lottery Method   (b) Table of Random Numbers

Merits:

Equal opportunity to each item.

Better way of judgment

Easy analysis and accuracy

Limitations:

Different in investigation

Expensive and time consuming

For filed survey it is not good

II Stratified Sampling:- In this it is important to divided the population into homogeneous group called strata. Then a sample may be taken from each group by simple random method.

Merit:- More representative sample is used.

Grater accuracy

Geographically Concentrated

Limitations: Utmost care must be exercised due to homogeneous group deviation. In the absence of skilled supervisor sample selection will be difficult.

III Systematic Sampling:- This method is popularly used in those cases where a complete list of the population from which sampling is to be drawn is available. The method is to be select k th item from the list where k refers to the sampling interval.

Merits: - It can be more convenient.

Limitation: - Can be Baised.

IV Multi- Stage Sampling: - This method refers to a sampling procedure which is carried out in several stages.

Merit: - It gives flexibility in Sampling

Limitation: - It is difficult and less accurate

Non Random Sampling Method:-
I. Judgment Sampling: - The choice of sample items depends exclusively on the judgment of the investigator or the investigator exercises his judgement in the choice of sample items. This is an simple method of sampling.
II. Quota Sampling: - Quotas are set up according to given criteria, but, within the quotas the selection of sample items depends on personal judgment.
III. Convenience Sampling: - It is also known as chunk. A chunk is a fraction of one population taken for investigation because of its convenient availability. That is why a chunk is selected neither by probability nor by judgment but by convenience.

Size of Sample:- It depends upon the following things:-
Cost aspects. The degree of accuracy desired. Time, etc. Normally it is 5% or 10% of the total population.

Limitation of overall sampling Method:-
Some time result may be inaccurate and misleading due to wrong sampling.
Its always needs superiors and experts to analyze the sample.
It may not give information about the overall defects. In production or any study.
It Becomes Biased due to following reason:-
(a) Faulty process of selection
(b) Faulty work during the collection of information
(c) Faulty methods of analysis etc.

UNIT-II
Measures of Central Tendency

The point around which the observations concentrate in general in the central part of the data is called central value of the data and the tendency of the observations to concentrate around a central point is known as Central Tendency.

Objects of Statistical Average:
➢   To get a single value that describes the characteristics of the entire group
➢   To facilitate comparison

Functions of Statistical Average:
➢   Gives information about the whole group
➢   Becomes the basis of future planning and actions
➢   Provides a basis for analysis
➢   Traces mathematical relationships
➢   Helps in decision making

Requisites of an Ideal Average:
➢   Simple and rigid definition
➢   Easy to understand
➢   Simple and easy to compute
➢   Based on all observations
➢   Least affected by extreme values
➢   Least affected by fluctuations of sampling
➢   Capable of further algebric treatment

ARITHMETIC MEAN ($\overline{X}$)

Arithmetic Mean of a group of observations is the quotient obtained by dividing the sum of all observations by their number. It is the most commonly used average or measure of the central tendency applicable only in case of quantitative data. Arithmetic mean is also simply called "mean".

Arithmetic mean is denoted by $\overline{X}$.

Merits of Arithmetic Mean:
- It is rigidly defined.
- It is easy to calculate and simple to follow.
- It is based on all the observations.
- It is readily put to algebraic treatment.
- It is least affected by fluctuations of sampling.
- It is not necessary to arrange the data in ascending or descending order.

Demerits of Arithmetic Mean:
- The arithmetic mean is highly affected by extreme values.
- It cannot average the ratios and percentages properly.
- It cannot be computed accurately if any item is missing.
- The mean sometimes does not coincide with any of the observed value.
- It cannot be determined by inspection.
- It cannot be calculated in case of open ended classes.

Methods of Calculating Arithmetic Mean:
- Direct Method

- Short cut method
- Step deviation method

Use of Arithmetic Mean:
Arithmetic Mean is recommended in following situation:
- When the frequency distribution is symmetrical.
- When we need a stable average.
- When other measures such as standard deviation, coefficient of correlation are to be computed later.

MEDIAN (M)
The median is that value of the variable which divides the group into two equal parts, one part comprising of all values greater and other of all values less than the median. For calculation of median the data has to be arranged in either ascending or descending order. Median is denoted by **M**.

Merits of Median:
- It is easily understood and easy to calculate.
- It is rigidly defined.
- It can sometimes be located by simple inspection and can also be computed graphically.
- It is positional average therefore not affected at all by extreme observations.
- It is only average to be used while dealing with qualitative data like intelligence, honesty etc.
- It is especially useful in case of open end classes since only the position and not the value of items must be known.
- It is not affected by extreme values.

Demerits of Median:
- For calculation, it is necessary to arrange data in ascending or descending order.
- Since it is a positional average, its value is not determined by each and every observation.
- It is not suitable for further algebric treatment.
- It is not accurate for large data.
- The value of median is more affected by sampling fluctuations than the value of the arithmetic mean.

Uses of Median:
The use of median is recommended in the following situations:
- When there are open-ended classes provided it does not fall in those classes.
- When exceptionally large or small values occur at the ends of the frequency distribution.
- When the observation cannot be measured numerically but can be ranked in order.
- To determine the typical value in the problems concerning distribution of wealth etc.

MODE (Z)
Mode is the value which occurs the greatest number of times in the data. The word mode has been derived from the French word **'La Mode'** which implies fashion. The Mode of a distribution is the value at the point around which the items tend to be most heavily concentrated. It may be regarded as the most typical of a series of values. Mode is denoted by **Z.**

Merits of Mode:
- It is easy to understand and simple to calculate.
- It is not affected by extreme large or small values.
- It can be located only by inspection in ungrouped data and discrete frequency distribution.
- It can be useful for qualitative data.

- It can be computed in open-end frequency table.
- It can be located graphically.

Demerits of Mode:
- It is not well defined.
- It is not based on all the values.
- It is suitable for large values and it will not be well defined if the data consists of small number of values.
- It is not capable of further mathematical treatment.
- Sometimes, the data has one or more than one mode and sometimes the data has no mode at all.

Uses of Mode:
The use of mode is recommended in the following situations:
- When a quick approximate measure of central tendency is desired.
- When the measure of central tendency should be the most typical value.

GEOMETRIC MEAN (G.M)
The geometric mean also called geometric average is the nth root of the product of n non-negative quantities. Geometric Mean is denoted by **G.M**.

Properties of Geometric Mean:
- The geometric mean is less than arithmetic mean, G.M<A.M
- The product of the items remains unchanged if each item is replaced by the geometric mean.
- The geometric mean of the ratio of corresponding observations in two series is equal to the ratios their geometric means.
- The geometric mean of the products of corresponding items in two series.

Merits of Geometric Mean:
- It is rigidly defined and its value is a precise figure.
- It is based on all observations.
- It is capable of further algebraic treatment.
- It is not much affected by fluctuation of sampling.
- It is not affected by extreme values.

Demerits of Geometric Mean:
- It cannot be calculated if any of the observation is zero or negative.
- Its calculation is rather difficult.
- It is not easy to understand.
- It may not coincide with any of the observations.

Uses of Geometric Mean:
- Geometric Mean is appropriate when:
    - Large observations are to be given less weight.
    - We find the relative changes such as the average rate of population growth, the average rate of intrest etc.
    - Where some of the observations are too small and/or too large.
- Also used for construction of Index Numbers.

HARMONIC MEAN (H.M)
Harmonic mean is another measure of central tendency. Harmonic mean is also useful for quantitative data. Harmonic mean is quotient of "number of the given values" and "sum of the reciprocals of the given values". It is denoted by **H.M.**

Merits of Harmonic Mean:
- It is based on all observations.
- It not much affected by the fluctuation of sampling.
- It is capable of algebraic treatment.
- It is an appropriate average for averaging ratios and rates.
- It does not give much weight to the large items and gives greater importance to small items.

Demerits of Harmonic Mean:
- Its calculation is difficult.
- It gives high weight-age to the small items.
- It cannot be calculated if any one of the items is zero.
- It is usually a value which does not exist in the given data.

Uses of Harmonic Mean:
- Harmonic mean is better in computation of average speed, average price etc. under certain conditions.