Kunthavai Naachiyaar Govt. Arts College (W) Autonomous, Thanjavur

B.Stat Major **18K4SELO2**

Non Major Elective - II **Biostatistics** **Hrs:2**

Credit:2

Unit – I

Definition of Bio – Statistics, characteristics of Statistics. Data collection of primary and secondary data – Definition and methods of collecting primary and secondary data.

Unit – II

Processing of data – Classification – Objectives & types of classification. Tabulation – Objectives – Components of Tables and types of Tables. Formation of frequency distribution – discrete & continuous.

Unit – III

Diagramatic representation – definition, Rules for constructing diagrams and uses. Simple bar diagram, Component bar diagram, multiple bar diagram and piediagram. Use any one of the Agriculture data for practice.

Unit – IV

Measures of central tendency – Mean, Median, Mode. Measures of dispersion – Range and standard deviation – Simple problems. Use any one of the weather data for practice.

Unit – V

Correlation – definition,Types of correlation, Methods of studying correlation – Karl Pearson's coefficient of correlation , Rank Correlation (without repeated ranks),simple Regression lines(two variables only)-simple problems.Use any one of the medical data for practice.

Books for Study :

1. Statistics theory and practice- R.S.N.Pillai, Bagavathi
2. Bio-statistics – P.Ramakrishna
3. Statistical methods for Biologists – S. Palanichamy & M.Manoharan.
4. Bio-Statistics – Gurusamy.

B. Stat Major Code :- 18K4SEL02.

BIO STATISTICS

UNIT-I

Bio statistics:

Statistical processes and methods applied to the collection, analysis, and interpretation of biological data and especially data relating to human biology, health, and medicine.

Characteristics of Statistics:

1.) Statistics are aggregate of facts.
2.) Statistics must be numerically expressed.
3.) Statistics should be collected for a pre-determined purpose.
4.) Statistics should be collected in a systematic manner.

COLLECTION OF DATA:

Primary Data:-

Primary data are those which are collected for the first time and they are original in character.

Secondary data:

Secondary data are those which are already collected by someone for some purpose and are available for the present study. For instance, the data collected during census operations are primary data to the department of census and the same data.

choice of Methods:-

For the collection of primary data, following methods,

1.) Direct personal observation.

2.) Indirect oral interview.

3.) Information through agencies.

4.) Mailed questionnaires.

5.) Schedules sent through enumerator.

1.) Direct Personal observation :- Under this method, the data is collated by the investigator personally. The investigator must be a keen observer, tactful and courteous in behaviour. He asks or cross-examines the informant and collects necessary information. The enquiry is intensive, rather than extensive.

2.) Indirect Oral Interview :- When the informant is reluctant to supply information, the method of indirect oral interview investigation can be followed. Under this method the investigator approaches the witnesses or third parties, who are in touch with the informant. The enumerator interviews the people, who are directly or indirectly connected with the problem under study. For instance, we are asked to collect information relating to the gambling or drinking habits of people. In such cases, the informants will be reluctant to supply information relating to their own socially evil habits. On such occasion we may approach the dealers of liquor shops, friends, neighbours, etc.,

3.) Information through agencies :- Under this method local agents or

correspondents will be appointed. They collect the information and transmit it to the office or person. They do this according to their own ways and tastes. This system is adopted by newspapers, periodicals, agencies.

4.) Mailed questionnaires:- In this method, a questionnaire consisting of a list of questions pertaining to the enquiry is prepared. There are blank spaces for answers. This questionnaire is sent to the respondents, who are expected to write the answers in the blank spaces. A covering letter is also sent along with the questionnaire, requesting the respondents to extend their full cooperation by giving the correct replies and returning the questionnaire duly filled in time.

5.) Schedules sent through enumerators:- It is the most widely used method of collection of primary data. A number of enumerators are selected and trained. They are provided with standardised questionnaires. Specific training and instructions are given to them for filling up schedules. Each enumerator will be in charge of a certain area.

Sources of Secondary Data:

The various sources of secondary data can be divided into two broad categories:

1.) Published sources.

2.) Unpublished sources.

1.) Published Sources:- Various governmental, international and local egencies publish statistics data, and chief amoung them are.

(a) International Publications:- International agencies and international bodies publish regular and occasional reports on economic and statistics matters. They are the I.M.F., the I.B.R.D., the I.C.A.F.E., and U.N.O., etc.

(b) Official publications of central and state Governments:- Departments of the Union and state Governments regularly publish reports on a number of subjects. They gather additional information. Some of the important publications are: the Reserve Bank of India Bulletin, Census of India, Statistical Abstracts of States, Agricultural Statistics of India, Indian Trade Journal, etc.

(c) Semi-official publications:- Semi-Government Institutions, like Municipal corporation, District Board, Panchayat, etc., Publish reports.

(d) Publications of Research Institutions:- Indian Statistics Institution (I.S.I), Indian council of Agricultural Research (I.C.A.R) Indian Agricultural Statistics Research Institute (I.A.S.R.I)., Publish the findings of their research programmes.

(e) Publications of commercial and Finacial Institutions:-

(f) Reports of various committees and commissions appointed by the Government: For example, Wanchoo commission Report on Taxation, Pay commission Reports, Land Reforms committee Reports, etc., are sources of secondary data.

(g) Journals and Newspapers:- current and important materials on Statistics and socio-economic problems can be obtained from Journals and newspapers like, Economic Times, commerce, Capital, Indian Finance, Monthly Statistics of Trade, etc.

2) Unpublished Sources:-

Precautions in the use of secondary Data:-

"The degree of reliability of secondary source is to be assessed from the source, the compiler and his capacity to produce correct statistics and the users also, for the most part, tend to accept a series, particularly one issued by a government agency at its face value without enquiring its reliability."

(a) The suitability of data:- First, the investigator must satisfy himself that the data available are suitable for the purpose of enquiry. It can be judged by the nature and scope of the present enquiry with the original enquiry.

For example, if the object of the present enquiry is to study the trend in retail prices, and if the data provide only wholesale prices, such data are unsuitable.

(b) Adequacy of data:- If the data are suitable for the purpose of investigation, then we must consider whether the data are useful or adequate for the present analysis. It can be studied by the geographical area covered by the original enquiry. We must consider the time for which the data are available. In the above example, if our object is to study the retail price trend of India, and if the available data cover only the retail price trend in the state of Tamil Nadu, then it would not serve the purpose.

(c) Reliability of data:- The reliability of data can be tested by finding out the agency that collected such data. If the agency has used proper methods in collecting data, statistics may be relied upon.

UNIT-II

CLASSIFICATION

Classification:-

"The process of grouping a large number of individual facts or observations on the basis of similarity among the items, is called classification."

- Stockton and Clark.

Objects of classification:

The chief objectives of classification are:

1. to condense the mass of data.
2. to present the facts in a simple form.
3. to bring out clearly the points of similarity and dissimilarity.
4. to facilitate comparison.
5. to bring out the relationship.
6. to prepare data for tabulation.
7. to facilitate the statistical treatment of the data.
8. to facilitate easy interpretation.
9. to eliminate unnecessary details.

Types of classification:

The classification of data primarily depends on the purpose and objectives of the enquiry. There are four important types of classification. They are:

1. Geographical i.e., area wise or region wise or district wise.

2. Chronological or historical i.e., on the basis of time.

3. Qualitative by character or by attributes.

4. Quantitative or numerical or by magnitudes.

Tabulation of Data:-

According to Secrist, "classification is the process of arranging data into sequences and groups according to their common characteristics, or separating them into different but related parts." In the words of Prof. Neiswanger, "A statistical table is a systematic organisation of data in columns and rows." Tabulation is the process of presenting data in tables.

Objectives of tabulation:-

1. to clarify the object of investigation.
2. to simplify complex data.
3. to clarify the characteristics of data.
4. to present facts in the minimum of space.
5. to facilitate comparison.
6. to detect errors and omission in the data.
7. to depict trend and tendencies of the problem under consideration.
8. to facilitate statistical processing.

9. to help reference.

Parts of Tabulation :-

A good statistical table is an art, The following parts must be present in all tables.

1. Table number.
2. Title.
3. Head note.
4. caption.
5. Stubs.
6. Body of the table.
7. Foot-note.
8. Source-note.

UNIT-III

DIAGRAMMATIC PRESENTATION.

A diagram is a visual form for presentation of statistical data. Diagram refers to the various types of devices such as bars, circles, maps, pictorials, cartograms, etc. These devices can take many attractive forms. Strictly speaking, these are not graphic devices. Diagrams do not add any new meaning to the statistical facts, but they exhibit the results more clearly. An ordinary man can understant pictures and diagrams more easily than the figures. The use of diagrams is becoming more and more popular in the present time. Diagrams occupy an important place, because:

Rules for Making a Diagram:.

Diagrammatic presentation of a statistical table is simple and effective as photographic memory will last long in the mind than any other form. The construction of a diagram is an art, which can be acquired through pratise. However, the following guide line will help in making them more effective.

1. Heading:- Every diagram must have suitable title. The title, in bold letters, conveys the main facts depicted by the diagram. If needed, sub-headings can also be given. It must be brief, self-explanatory and clear.

2. Size:- The size of the diagram should neither be too big nor too small. It must match with the size of the paper. It should be in the middle of the paper.

3. Length and Breadth:- An appropriate proportion should be maintained between length and breadth. Lutz has suggested that proportions of length and breadth should be 2:1 or 1.414:1. If it is so, the diagram looks attractive. Care should be taken to ensure that the diagram does not wear ugly look.

4. Drawing: Since impression is needed, it should be drawn neatly and accurately with the help of drawing instruments. Each diagram should also be numbered for ready reference.

5. A proper scale:- A proper scale must be chosen for the diagram to look attractive and create a visual impact on the reader. It must suit the space available. Accuracy should not be sacrificed to attractiveness.

6. Selection of Appropriate Diagram:- The most important point is the selection of proper diagram to present a set of figures. All types of diagrams are not suitable for all types of data. A wrong selection of the diagram will distort the true characteristics of the phenomenon to the presented and might lead to very wrong and misleading interpretations.

7. Right method :- C.W. Lowe writes, "the important point, that must be borne in mind at all times, is that the pictorial presentation chosen for any situation, must depict the true relationship and point out the proper conclusion. Use of an inappropriate chart may distort the facts and mislead the reader. Above all, the chart must be honest."

8. Index :- When many items are shown in a diagram, through different colours, dottings, crossing, etc., an index must be given for identifying and understanding the diagram.

9. Sources :- If the data presented have been acquired from some external source, that fact should be indicated at the bottom.

10. Simplicity :- Diagram should be very simple. It must be so simple that even a lay man who does not have knowledge of mathematical or statistical background, can understand the diagram. If the data are very large draw more diagrams to represent the data.

Too much information presented in a diagram will be confusing. Therefore, it is suggested to draw several simple diagrams which are more effective than a complex one.

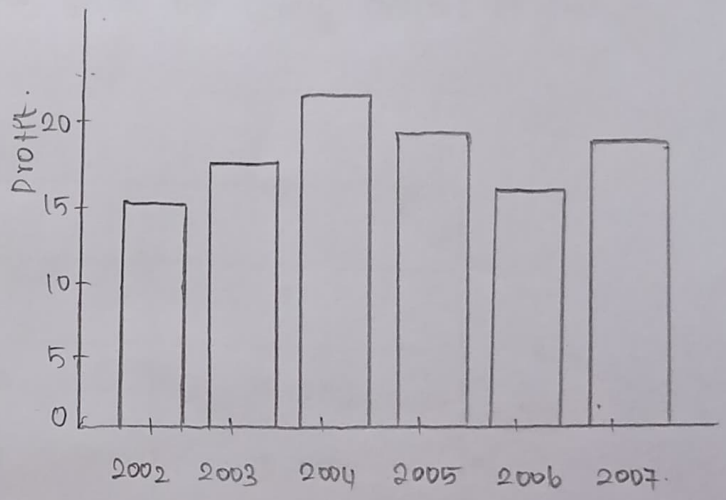
Simple bar diagram :- A simple bar diagram can be drawn either on horizontal or vertical base. Bars on horizontal base are more common. A bar diagram is simple to draw, and easy to understand. In business and economics it is commonly used.

Illustration 7.2 :

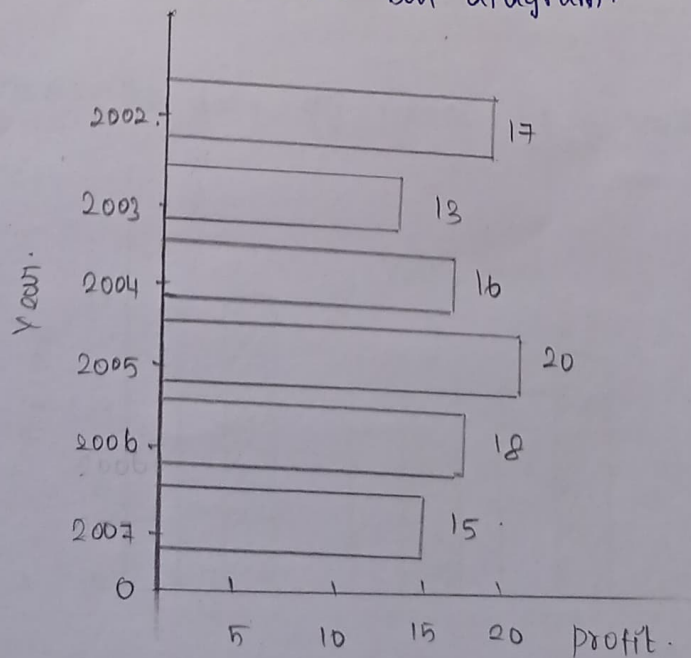
Draw a suitable bar diagram showing the following data.

Year	Profit ('000)
2002	15,000
2003	18,000
2004	20,000
2005	16,000
2006	13,000
2007	17,000

(a) Vertical bar diagram.



(b) Horizontal bar diagram.



Multiple bar diagram (compound bar diagram):

Multiple bar diagrams are used to denote more than one phenomenon, e.g., for import and export trend. Multiple bars are useful for direct comparison between two values. The bars are drawn side by side. In order to distinguish the bars, different colours, shades, etc., may be used and a key index to this effect be given to understand the different bar.

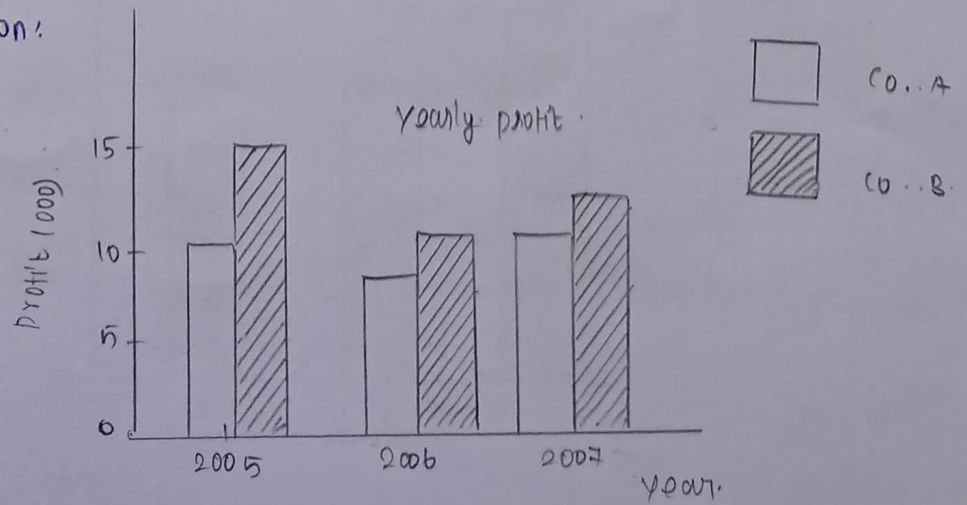
Illustration 7.3:

The data below gives the yearly profits of two companies A and B.

Year	Profit.	
	A	B
2005	10,000	15,000
2006	8,000	13,000
2007.	13,000	14,000.

Represent the data by means of a multiple bar diagram.

Solution:



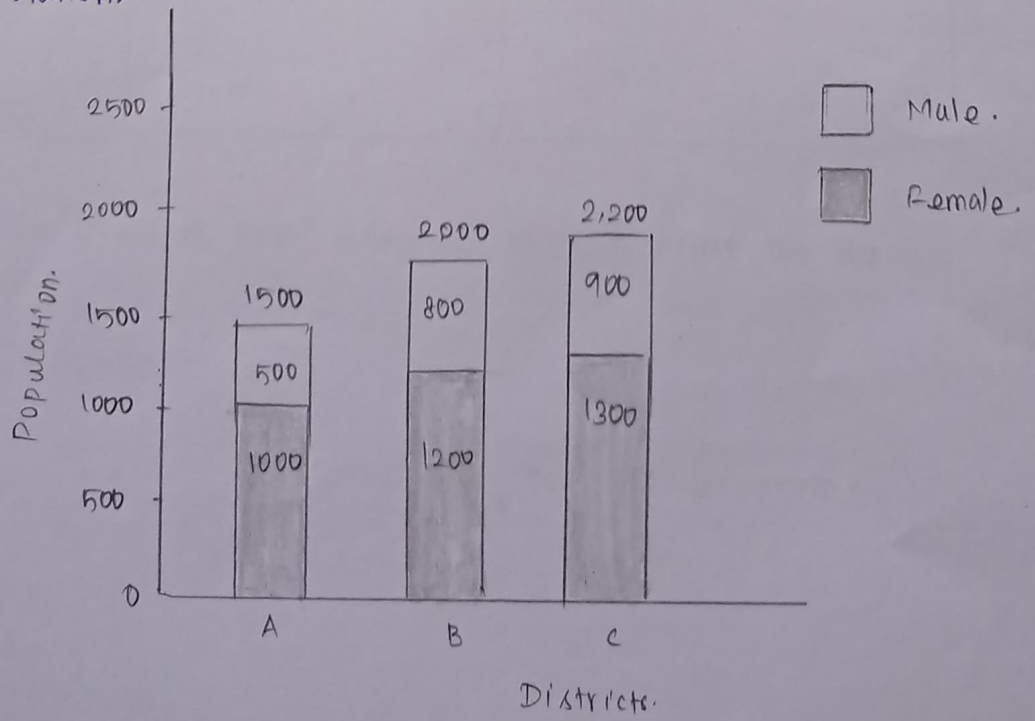
Sub-divided bar diagram (component bar diagram):

The bar is subdivided into various parts in proportion to the values given in the data and may be drawn on absolute figures or percentages. Each component occupies a part of the bar proportional to its share in the total. To distinguish different components from one another, different colour or shades may be given.

Illustration 7.4: Represent the following data in a suitable diagram.

Districts	A	B	C
Male Population	1,000	1,200	1,300
Female	500	800	900
	1,500	2,000	2,200

Solution:



Angular or Pie diagram:- The pie diagram ranks high in understanding just as we divide a bar or a rectangle to show its components, a circle can also be divided into sectors. As there are 360 degrees at the centre, proportionate sectors are cut taking the whole data equal to 360 degrees. This will be clear from the following illustration.

Illustration 7.8:

The following table shows the area in millions of square kilometres of the oceans of the world.

Ocean	Area (million sq. km)
Pacific	70.8
Atlantic	41.2
Indian	28.5
Antarctic	7.6
Arctic	4.8

Draw a pie diagram to represent the data.

Solution:-

Calculation for pie diagram.

Ocean	Area	Degrees.
Pacific	70.8	$\frac{70.8}{152.9} \times 360 = 167.$
Atlantic	41.2	$\frac{41.2}{152.9} \times 360 = 97.$
Indian	28.5	$\frac{28.5}{152.9} \times 360 = 67.$
Antarctic	7.6	$\frac{7.6}{152.9} \times 360 = 18.$
Arctic	4.8	$\frac{4.8}{152.9} \times 360 = 11.$
	152.9	360°

Pie diagram showing the area of oceans of the world.