

UNIT-I

1.1 Meaning of Statistics

Introduction:

In the modern world of computers and information technology, the importance of statistics is very well recognized by all the disciplines. Statistics has originated as a science of statehood and found applications slowly and steadily in Agriculture, Economics, Commerce, Biology, Medicine, Industry, planning, education and so on. As on date there is no other human walk of life, where statistics cannot be applied.

Origin and Growth of Statistics:

The word 'Statistics' and 'Statistical' are all derived from the Latin word Status, means a political state. The theory of statistics as a distinct branch of scientific method is of comparatively recent growth. Research particularly into the mathematical theory of statistics is rapidly proceeding and fresh discoveries are being made all over the world.

Meaning of Statistics:

Statistics is concerned with scientific methods for collecting, organising, summarising, presenting and analysing data as well as deriving valid conclusions and making reasonable decisions on the basis of this analysis. Statistics is concerned with the systematic collection of numerical data and its interpretation. The word 'statistic' is used to refer to

- Numerical facts, such as the number of people living in particular area.
- The study of ways of collecting, analysing and interpreting the facts.

Definitions:

Statistics is defined differently by different authors over a period of time. In the olden days statistics was confined to only state affairs but in modern days it embraces almost every sphere of human activity. Therefore a number of old definitions, which was confined to narrow field of enquiry were replaced by more definitions, which are much more comprehensive and exhaustive. Secondly, statistics has been defined in two different ways – Statistical data and statistical methods. The following are some of the definitions of statistics as numerical data.

- Statistics are the classified facts representing the conditions of people in a state. In particular they are the facts, which can be stated in numbers or in tables of numbers or in any tabular or classified arrangement.
- Statistics are measurements, enumerations or estimates of natural phenomenon usually systematically arranged, analysed and presented as to exhibit important interrelationships among them.

Definitions by A.L. Bowley:

Statistics are numerical statement of facts in any department of enquiry placed in relation to each other.

- A.L. Bowley

Statistics may be called the science of counting in one of the departments due to Bowley, obviously this is an incomplete definition as it considers only the aspect of collection and ignores other aspects such as analysis, presentation and interpretation.

Bowley gives another definition for statistics, which states ‘statistics may be rightly called the scheme of averages’ . This definition is also incomplete, as averages play an important role in understanding and comparing data and statistics provide more measures.

Definition by Croxton and Cowden:

Statistics may be defined as the science of collection, presentation analysis and interpretation of numerical data from the logical analysis. It is clear that the definition of statistics by Croxton and Cowden is the most scientific and realistic one. According to this definition there are four stages:

- **Collection of Data:** It is the first step and this is the foundation upon which the entire data set. Careful planning is essential before collecting the data. There are different methods of collection of data such as census, sampling, primary, secondary, etc., and the investigator should make use of correct method.
- **Presentation of data:** The mass data collected should be presented in a suitable, concise form for further analysis. The collected data may be presented in the form of tabular or diagrammatic or graphic form.
- **Analysis of data:** The data presented should be carefully analysed for making inference from the presented data such as measures of central tendencies, dispersion, correlation, regression etc.,
- **Interpretation of data:** The final step is drawing conclusion from the data collected. A valid conclusion must be drawn on the basis of analysis. A high degree of skill and experience is necessary for the interpretation.

Definition by Horace Secrist:

Statistics may be defined as the aggregate of facts affected to a marked extent by multiplicity of causes, numerically expressed, enumerated or estimated according to a reasonable standard of accuracy, collected in a systematic manner, for a predetermined purpose and placed in relation to each other.

1.2 Scope of Statistics

Statistics is not a mere device for collecting numerical data, but as a means of developing sound techniques for their handling, analysing and drawing valid inferences from them. Statistics is applied in every sphere of human activity – social as well as physical – like Biology, Commerce, Education, Planning, Business Management, Information Technology, etc. It is almost impossible to find a single department of human activity where statistics cannot be applied. We now discuss briefly the applications of statistics in other disciplines.

a. Statistics and Industry:

Statistics is widely used in many industries. In industries, control charts are widely used to maintain a certain quality level. In production engineering, to find whether the product is conforming to specifications or not, statistical tools, namely inspection plans, control charts, etc., are of extreme importance. In inspection plans we have to resort to some kind of sampling – a very important aspect of Statistics.

b. Statistics and Commerce:

Statistics are lifeblood of successful commerce. Any businessman cannot afford to either by under stocking or having overstock of his goods. In the beginning he estimates the demand for his goods and then takes steps to adjust with his output or purchases. Thus statistics is indispensable in business and commerce.

As so many multinational companies have invaded into our Indian economy, the size and volume of business is increasing. On one side the stiff competition is increasing whereas on the other side the tastes are changing and new fashions are emerging. In this connection, market survey plays an important role to exhibit the present conditions and to forecast the likely changes in future.

c. Statistics and Agriculture:

Analysis of variance (ANOVA) is one of the statistical tools developed by Professor R.A. Fisher, plays a prominent role in agriculture experiments. In tests of significance based on small samples, it can be shown that statistics is adequate to test the significant difference between two sample means. In analysis of variance, we are concerned with the testing of equality of several population means.

For an example, five fertilizers are applied to five plots each of wheat and the yield of wheat on each of the plots are given. In such a situation, we are interested in finding out whether the effect of these fertilisers on the yield is significantly different or not. In other words, whether the samples are drawn from the same normal population or not. The answer to this problem is provided by the technique of ANOVA and it is used to test the homogeneity of several population means.

d. Statistics and Economics:

Statistical methods are useful in measuring numerical changes in complex groups and interpreting collective phenomenon. Nowadays the uses of statistics are abundantly made in any economic study. Both in economic theory and practice, statistical methods play an important role.

Alfred Marshall said, “ Statistics are the straw only which I like every other economist have to make the bricks”. It may also be noted that statistical data and techniques of statistical tools are immensely useful in solving many economic problems such as wages, prices, production, distribution of income and wealth and so on. Statistical tools like Index numbers, time series Analysis, Estimation theory, Testing Statistical Hypothesis are extensively used in economics.

e. Statistics and Education:

Statistics is widely used in education. Research has become a common feature in all branches of activities. Statistics is necessary for the formulation of policies to start new course, consideration of facilities available for new courses etc. There are many people engaged in research work to test the past knowledge and evolve new knowledge. These are possible only through statistics.

f. Statistics and Planning:

Statistics is indispensable in planning. In the modern world, which can be termed as the “world of planning”, almost all the organisations in the government are seeking the help of planning for efficient working, for the formulation of policy decisions and execution of the same.

In order to achieve the above goals, the statistical data relating to production, consumption, demand, supply, prices, investments, income expenditure etc and various advanced statistical techniques for processing, analysing and interpreting such complex data are of importance. In India statistics play an important role in planning, commissioning both at the central and state government levels.

g. Statistics and Medicine:

In Medical sciences, statistical tools are widely used. In order to test the efficiency of a new drug or medicine, t - test is used or to compare the efficiency of two drugs or two medicines, t - test for the two samples is used. More and more applications of statistics are at present used in clinical investigation.

h. Statistics and Modern applications:

Recent developments in the fields of computer technology and information technology have enabled statistics to integrate their models and thus make statistics a part of decision-making procedures of many organisations. There are so many software packages available for solving design of experiments, forecasting simulation problems etc.

SYSTAT, a software package offers mere scientific and technical graphing options than any other desktop statistics package. SYSTAT supports all types of scientific and technical research in various diversified fields as follows

- Archeology: Evolution of skull dimensions
- Epidemiology: Tuberculosis
- Statistics: Theoretical distributions
- Manufacturing: Quality improvement
- Medical research: Clinical investigations.
- Geology: Estimation of Uranium reserves from ground water.
-

1.3 Limitations of statistics

(i) **Statistics is not suitable to the study of qualitative phenomenon:** Since statistics is basically a science and deals with a set of numerical data, it is applicable to the study of only these subjects of enquiry, which can be expressed in terms of quantitative measurements. As a matter of fact, qualitative phenomenon like honesty, poverty, beauty, intelligence etc, cannot be expressed numerically and any statistical analysis cannot be directly applied on these qualitative phenomenon. Nevertheless, statistical techniques may be applied indirectly by first reducing the qualitative expressions to accurate quantitative terms. For example, the intelligence of a group of students can be studied based on their marks in a particular examination.

- (ii) **Statistics does not study individuals:** Statistics does not give any specific importance to the individual items, in fact it deals with an aggregate of objects. Individual items, when they are taken individually do not constitute any statistical data and do not serve any purpose for any statistical enquiry.
- (iii) **Statistical laws are not exact:** It is well known that mathematical and physical sciences are exact. But statistical laws are not exact and statistical laws are only approximations. Statistical conclusions are not universally true. They are true only on an average.
- (iv) **Statistics table may be misused:** Statistics must be used only by experts; otherwise, statistical methods are the most dangerous tools on the hands of the inexperienced. The use of statistical tools by the inexperienced and untraced persons might lead to wrong conclusions. Statistics can be easily misused by quoting wrong figures of data. As King says aptly 'statistics are like clay of which one can make a God or Devil as one pleases'.
- (v) **Statistics is only, one of the methods of studying a problem:** Statistical method do not provide complete solution of the problems because problems are to be studied taking the background of the countries culture, philosophy or religion into consideration. Thus, the statistical study should be supplemented by other evidence.

1.4 Primary Data

Any statistical data can be classified under two categories depending upon the sources utilized. There are two categories of data and they are

- Primary data
- Secondary data

Primary data is the one, which is collected by the investigator himself for the purpose of a specific inquiry or study. Such data is original in character and is generated by survey conducted by individuals or research institution or any organization.

Example: If a researcher is interested to know the impact of noon meal scheme for the school children, he has to undertake a survey and collect data on the opinion of parents and children by asking relevant questions. Such a data collected for the purpose is called primary data.

1.5 Methods of collecting Primary data

The primary data can be collected by the following five methods.

- (i) Direct personal interviews.
- (ii) Indirect Oral interviews.
- (iii) Information from correspondents.
- (iv) Mailed questionnaire method.
- (v) Schedules sent through enumerators.

Direct personal interviews:

The persons from whom information are collected are known as informants. The investigator personally meets them and asks questions to gather the necessary information. It is the suitable method for intensive rather than extensive field surveys. It suits best for intensive study of the limited field.

Merits:

- People willingly supply information because they are approached personally. Hence, more response noticed in this method than in any other method.
- The collected information are likely to be uniform and accurate. The investigator is there to clear the doubts of the informants.
- Supplementary information on informant's personal aspects can be noted. Information on character and environment may help later to interpret some of the results.
- Answers for questions about which the informant is likely to be sensitive can be gathered by this method.
- The wordings in one or more questions can be altered to suit any informant. Explanations may be given in other languages also. Inconvenience and misinterpretations are thereby avoided.

Limitations:

- It is very costly and time consuming.
- It is very difficult, when the number of persons to be interviewed is large and the persons are spread over a wide area.
- Personal prejudice and bias are greater under this method.

Indirect Oral Interviews:

Under this method the investigator contacts witnesses or neighbours or friends or some other third parties who are capable of supplying the necessary information. This method is preferred if the required information is on addiction or cause of fire or theft or murder etc., If a fire has broken out a certain place, the persons living in neighbourhood and witnesses are likely to give information on the cause of fire. In some cases, police interrogated third parties who are supposed to have knowledge of a theft or a murder and get some clues. Enquiry committees appointed by governments generally adopt this method and get people's views and all possible details of facts relating to the enquiry. This method is suitable whenever direct sources do not exist or cannot be relied upon or would be unwilling to part with the information.

The validity of the results depends upon a few factors, such as the nature of the person whose evidence is being recorded, the ability of the interviewer to draw out information from the third parties by means of appropriate questions and cross examinations, and the number of persons interviewed. For the success of this method one person or one group alone should not be relied upon.

Information from correspondents:

The investigator appoints local agents or correspondents in different places and compiles the information sent by them. Information to Newspapers and some departments of Government come by this method. The advantage of this method is that it is cheap and appropriate for extensive investigations. But it may not ensure accurate results because the correspondents are likely to be negligent, prejudiced and biased. This method is adopted in those cases where information are to be collected periodically from a wide area for a long time.

Mailed questionnaire method:

Under this method a list of questions is prepared and is sent to all the informants by post. The list of questions is technically called questionnaire. A covering letter accompanying the questionnaire explains the purpose of the investigation and the importance of correct

information and request the informants to fill in the blank spaces provided and to return the form within a specified time. This method is appropriate in those cases where the informants are literates and are spread over a wide area.

Merits:

- It is relatively cheap.
- It is preferable when the informants are spread over the wide area.

Limitations:

- The greatest limitation is that the informants should be literates who are able to understand and reply the questions.
- It is possible that some of the persons who receive the questionnaires do not return them.
- It is difficult to verify the correctness of the information furnished by the respondents.

With the view of minimizing non-respondents and collecting correct information, the questionnaire should be carefully drafted. There is no hard and fast rule. But the following general principles may be helpful in framing the questionnaire. A covering letter and a self addressed and stamped envelope should accompany the questionnaire. The covering letter should politely point out the purpose of the survey and privilege of the respondent who is one among the few associated with the investigation. It should assure that the information would be kept confidential and would never be misused. It may promise a copy of the findings or free gifts or concessions etc.,

Characteristics of a good questionnaire:

- (i) Number of questions should be minimum.
- (ii) Questions should be in logical orders, moving from easy to more difficult questions.
- (iii) Questions should be short and simple. Technical terms and vague expressions capable of different interpretations should be avoided.
- (iv) Questions fetching YES or NO answers are preferable. There may be some multiple choice questions requiring lengthy answers are to be avoided.
- (v) Personal questions and questions which require memory power and calculations should also be avoided.
- (vi) Question should enable cross check. Deliberate or unconscious mistakes can be detected to an extent.
- (vii) Questions should be carefully framed so as to cover the entire scope of the survey.
- (viii) The wording of the questions should be proper without hurting the feelings or arousing resentment.
- (ix) As far as possible confidential information should not be sought.
- (x) Physical appearance should be attractive, sufficient space should be provided for answering each question.

Schedules sent through Enumerators:

Under this method enumerators or interviewers take the schedules, meet the informants and filling their replies. Often distinction is made between the schedule and a questionnaire. A schedule is filled by the interviewers in a face-to-face situation with the informant. A questionnaire is filled by the informant which he receives and returns by post. It is suitable for extensive surveys.

Merits:

- It can be adopted even if the informants are illiterates.
- Answers for questions of personal and pecuniary nature can be collected.
- Non-response is minimum as enumerators go personally and contact the informants.
- The information collected are reliable. The enumerators can be properly trained for the same.
- It is most popular methods.

Limitations:

- It is the costliest method.
- Extensive training is to be given to the enumerators for collecting correct and uniform information.
- Interviewing requires experience. Unskilled investigators are likely to fail in their work.

Before the actual survey, a pilot survey is conducted. The questionnaire/Schedule is pre-tested in a pilot survey. A few among the people from whom actual information is needed are asked to reply. If they misunderstand a question or find it difficult to answer or do not like its wordings etc., it is to be altered. Further it is to be ensured that every questions fetches the desired answer.

Merits and Demerits of primary data:

- The collection of data by the method of personal survey is possible only if the area covered by the investigator is small. Collection of data by sending the enumerator is bound to be expensive. Care should be taken twice that the enumerator record correct information provided by the informants.
- Collection of primary data by framing a schedules or distributing and collecting questionnaires by post is less expensive and can be completed in shorter time.
- Suppose the questions are embarrassing or of complicated nature or the questions probe into personnel affairs of individuals, then the schedules may not be filled with accurate and correct information and hence this method is unsuitable.
- The information collected for primary data is more reliable than those collected from the secondary data.

1.6 Secondary Data

Secondary data are those data which have been already collected and analysed by some earlier agency for its own use; and later the same data are used by a different agency. According to W. A. Neiswanger, 'A primary source is a publication in which the data are published by the same authority which gathered and analysed them. A secondary source is a publication, reporting the data which have been gathered by other authorities and for which others are responsible'.

Sources of secondary data

In most of the studies the investigator finds it impracticable to collect first-hand information on all related issues and as such he makes use of the data collected by others. There is a vast amount of published information from which statistical studies may be made and fresh statistics are constantly in a state of production. The sources of secondary data can broadly be classified under two heads:

- (i) Published sources, and
- (ii) Unpublished sources.

Published Sources:

The various sources of published data are:

- a. Reports and official publications of
 - i. International bodies such as the International Monetary Fund, International Finance Corporation and United Nations Organisation.
 - ii. Central and State Governments such as the Report of the Tandon Committee and Pay Commission.
- b. Semi-official publication of various local bodies such as Municipal Corporations and District Boards.
- c. Private publications-such as the publications of –
 - i. Trade and professional bodies such as the Federation of Indian Chambers of Commerce and Institute of Chartered Accountants.
 - ii. Financial and economic journals such as ‘Commerce’, ‘Capital’ and ‘Indian Finance’ .
 - iii. Annual reports of joint stock companies.
 - iv. Publications brought out by research agencies, research scholars, etc.

It should be noted that the publications mentioned above vary with regard to the periodicity of publication. Some are published at regular intervals (yearly, monthly, weekly etc.,) whereas others are ad hoc publications, i.e., with no regularity about periodicity of publications. A lot of secondary data is available in the internet. We can access it at any time for the further studies.

Unpublished Sources

All statistical material is not always published. There are various sources of unpublished data such as records maintained by various Government and private offices, studies made by research institutions, scholars, etc. Such sources can also be used where necessary

Precautions in the use of Secondary data

The following are some of the points that are to be considered in the use of secondary data:

- How the data has been collected and processed.
- The accuracy of the data.
- How far the data has been summarized.
- How comparable the data is with other tabulation?
- How to interpret the data, especially when figures collected for one purpose is used for another.

With secondary data, people must compromise between what they want and what they can find.

Merits and Demerits of Secondary Data:

- Secondary data is cheap to obtain. Many government publications are relatively cheap and libraries stock quantities of secondary data produced by the government, by companies and other organisations.
- Large quantities of secondary data can be got through internet.
- Much of the secondary data available has been collected for many years and therefore it can be used to plot trends.
- Secondary data is of value to:

- The government – help in making decisions and planning future policy.
- Business and industry – in areas such as marketing, and sales in order to appreciate the general economic and social conditions and to provide information on competitors.
- Research organizations – by providing social, economical and industrial information.

1.7 Classification of Data

Classification:

The collected data, also known as raw data or ungrouped data are always in an unorganised form and need to be organised and presented in meaningful and readily comprehensible form in order to facilitate further statistical analysis. It is, therefore, essential for an investigator to condense a mass of data into more and more comprehensible and assimilable form. The process of grouping into different classes or sub classes according to some characteristics is known as classification, tabulation is concerned with the systematic arrangement and presentation of classified data. Thus classification is the first step in tabulation.

For Example, letters in the post office are classified according to their destinations viz., Delhi, Madurai, Bangalore, Mumbai etc.,

Objectives of Classification:

The following are main objectives of classifying the data:

- i. It condenses the mass of data in an easily assimilable form.
- ii. It eliminates unnecessary details.
- iii. It facilitates comparison and highlights the significant aspect of data.
- iv. It enables one to get a mental picture of the information and helps in drawing inferences.
- v. It helps in the statistical treatment of the information collected.

Types of classification:

Statistical data are classified in respect of their characteristics. Broadly there are four basic types of classification namely

- a) Chronological classification
- b) Geographical classification
- c) Qualitative classification
- d) Quantitative classification

a) Chronological classification:

In chronological classification, the collected data are arranged according to the order of time expressed in years, months, weeks, etc., The data is generally classified in ascending order of time. For example, the data related with population, sales of a firm, imports and exports of a country are always subjected to chronological classification.

Example: The estimates of birth rates in India during 1970 –76 are

Year	1970	1971	1972	1973	1974	1975	1976
Birth Rate	36.8	36.9	36.6	34.6	34.5	35.2	34.2

b) Geographical classification:

In this type of classification, the data are classified according to geographical region or place. For instance, the production of paddy in different states in India, production of wheat in different countries etc.,

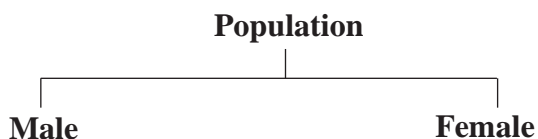
Example :

Country	America	China	Denmark	France	India
Yield of wheat in (kg/acre)	1925	893	225	439	862

c) Qualitative classification:

In this type of classification data are classified based on same attributes or quality like sex, literacy, religion, employment etc., Such attributes cannot be measured along with a scale. For example, if the population to be classified in respect to one attribute, say sex, then we can classify them into two namely that of males and females. Similarly, they can also be classified into ‘employed’ or ‘unemployed’ on the basis of another attribute ‘employment’. Thus, when the classification is done with respect to one attribute, which is dichotomous in nature, two classes are formed, one possessing the attribute and the other not possessing the attribute. This type of classification is called simple or dichotomous classification.

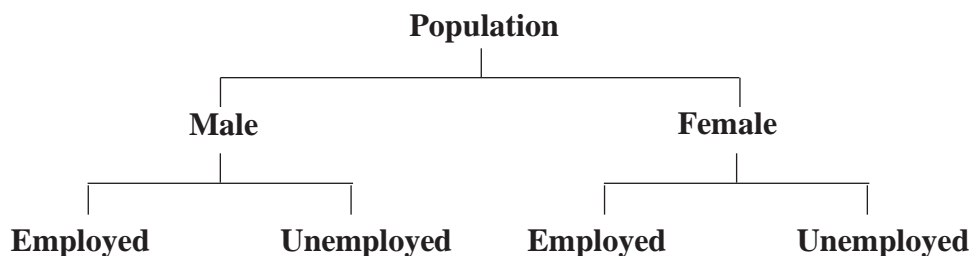
A simple classification may be shown as under



The classification, where two or more attributes are considered and several classes are formed, is called a manifold classification. For example, if we classify population simultaneously with respect to two attributes, e.g sex and employment, then population are first classified with respect to ‘sex’ into ‘males’ and ‘females’. Each of these classes may then be further classified into ‘employment’ and ‘unemployment’ on the basis of attribute ‘employment’ and as such Population are classified into four classes namely.

- (i) Male employed
- (ii) Male unemployed
- (iii) Female employed
- (iv) Female unemployed

Still the classification may be further extended by considering other attributes like marital status etc. This can be explained by the following chart



d) Quantitative classification:

Quantitative classification refers to the classification of data according to some characteristics that can be measured such as height, weight, etc., For example the students of a college may be classified according to weight as given below.

Weight (in lbs)	No. of Students
90-100	50
100-110	200
110-120	260
120-130	360
130-140	90
140-150	40
Total	1000

In this type of classification there are two elements, namely (i) the variable (i.e) the weight in the above example, and (ii) the frequency in the number of students in each class. There are 50 students having weights ranging from 90 to 100 lb, 200 students having weight ranging between 100 to 110 lb and so on.

1.7 Tabulation of Data

Tabulation is the process of summarizing classified or grouped data in the form of a table so that it is easily understood, and an investigator is quickly able to locate the desired information. A table is a systematic arrangement of classified data in columns and rows. Thus, a statistical table makes it possible for the investigator to present a huge mass of data in a detailed and orderly form. It facilitates comparison and often reveals certain patterns in data which are otherwise not obvious. Classification and 'Tabulation', as a matter of fact, are not two distinct processes. Actually, they go together. Before tabulation data are classified and then displayed under different columns and rows of a table.

Advantages of Tabulation:

Statistical data arranged in a tabular form serve following objectives:

- (i) It simplifies complex data and the data presented are easily understood.
- (ii) It facilitates comparison of related facts.
- (iii) It facilitates computation of various statistical measures like averages, dispersion, correlation etc.
- (iv) It presents facts in minimum possible space and unnecessary repetitions and explanations are avoided. Moreover, the needed information can be easily located.
- (v) Tabulated data are good for references and they make it easier to present the information in the form of graphs and diagrams.

Preparing a Table:

The making of a compact table itself an art. This should contain all the information needed within the smallest possible space. What the purpose of tabulation is and how the tabulated information is to be used are the main points to be kept in mind while preparing for

astatistical table. An ideal table should consist of the following main parts:

- (i) Table number
- (ii) Title of the table
- (iii) Captions or column headings
- (iv) Stubs or row designation
- (v) Body of the table
- (vi) Footnotes
- (vii) Sources of data

Table Number:

A table should be numbered for easy reference and identification. This number, if possible, should be written in the centre at the top of the table. Sometimes it is also written just before the title of the table.

Title:

A good table should have a clearly worded, brief but unambiguous title explaining the nature of data contained in the table. It should also state arrangement of data and the period covered. The title should be placed centrally on the top of a table just below the table number (or just after table number in the same line).

Captions or column Headings:

Captions in a table stands for brief and self explanatory headings of vertical columns. Captions may involve headings and sub-headings as well. The unit of data contained should also be given for each column. Usually, a relatively less important and shorter classification should be tabulated in the columns.

Stubs or Row Designations:

Stubs stands for brief and self explanatory headings of horizontal rows. Normally, a relatively more important classification is given in rows. Also a variable with a large number of classes is usually represented in rows. For example, rows may stand for score of classes and columns for data related to sex of students. In the process, there will be many rows for scores classes but only two columns for male and female students.

A model structure of a table is given below:

Table Number Title of the Table

Sub Heading	Caption Headings	Total
	Caption Sub-Headings	
Stub Sub-Headings	Body	
Total		

Foot notes:

Sources note:

Body:

The body of the table contains the numerical information of frequency of observations in the different cells. This arrangement of data is according to the description of captions and stubs.

Footnotes:

Footnotes are given at the foot of the table for explanation of any fact or information included in the table which needs some explanation. Thus, they are meant for explaining or providing further details about the data, that have not been covered in title, captions and stubs.

Sources of data:

Lastly one should also mention the source of information from which data are taken. This may preferably include the name of the author, volume, page and the year of publication. This should also state whether the data contained in the table is of 'primary or secondary' nature.

Requirements of a Good Table:

A good statistical table is not merely a careless grouping of columns and rows but should be such that it summarizes the total information in an easily accessible form in minimum possible space. Thus while preparing a table, one must have a clear idea of the information to be presented, the facts to be compared and the points to be stressed.

Though, there is no hard and fast rule for forming a table yet a few general points should be kept in mind:

- (i) A table should be formed in keeping with the objects of statistical enquiry.
- (ii) A table should be carefully prepared so that it is easily understandable.
- (iii) A table should be formed to suit the size of the paper. But such an adjustment should not be at the cost of legibility.
- (iv) If the figures in the table are large, they should be suitably rounded or approximated. The method of approximation and units of measurements too should be specified.
- (v) Rows and columns in a table should be numbered and certain figures to be stressed may be put in 'box' or 'circle' or in bold letters.
- (vi) The arrangements of rows and columns should be in a logical and systematic order. This arrangement may be alphabetical, chronological, or according to size.
- (vii) The rows and columns are separated by single, double, or thick lines to represent various classes and sub-classes used. The corresponding proportions or percentages should be given in adjoining rows and columns to enable comparison. A vertical expansion of the table is generally more convenient than the horizontal one.
 - The averages or totals of different rows should be given at the right of the table and that of columns at the bottom of the table. Totals for every sub-class too should be mentioned.
 - In case it is not possible to accommodate all the information in a single table, it is better to have two or more related tables.

Type of Tables:

Tables can be classified according to their purpose, stage of enquiry, nature of data or number of characteristics used. Based on the number of characteristics, tables may be classified as follows:

- a) Simple or one-way table
- b) Two-way table
- c) Mani-fold table

Simple or one-way Table:

A simple or one-way table is the simplest table which contains data of one characteristic only. A simple table is easy to construct and simple to follow. For example, the blank table given below may be used to show the number of adults in different occupations in a locality.

The number of adults in different occupations in a locality

Occupations	No. of Adults
Total	

Two-way Table:

A table, which contains data on two characteristics, is called a twoway table. In such case, therefore, either stub or caption is divided into two co-ordinate parts. In the given table, asan example the caption may be further divided in respect of ‘ sex’ . This subdivision is shown in two-way table, which now contains two characteristics namely, occupation and sex.

The number of adults in a locality in respect of occupation and sex

Occupation	No. of Adults		Total
	Male	Female	
Total			

Manifold Table:

Thus, more and more complex tables can be formed by including other characteristics. For example, we may further classify the caption sub-headings in the above table in respect of “marital status”, “ religion” and “socio-economic status” etc. A table, which has more than two characteristics of data is considered as a manifold table. For instance , table shown below showsthree characteristics namely, occupation, sex and marital status.

Occupation	No. of Adults						Total
	Male			Female			
	M	U	Total	M	U	Total	
Total							

Foot note: M Stands for Married and U stands for unmarried.

Manifold tables, though complex, are good in practice as these enable full information tobe incorporated and facilitate analysis of all related facts. Still, as a normal practice, not more than four characteristics should be represented in one table to avoid confusion. Other related tables may be formed to show the remaining characteristics.