QUANTITATIVE TECHNIQUES-I
(DBC014)
(BACHELOR OF COMMERCE)

ACHARYA NAGARJUNA UNIVERSITY

CENTRE FOR DISTANCE EDUCATION

NAGARJUNA NAGAR,

GUNTUR

ANDHRA PRADESH

Lesson: 1

Statistics- Scope and Importance

1.0 Objectives:

After going through the lesson you will be able to understand the following:

1. Meaning and definitions of Statistics
2. Functions of Statistics
3. Scope and importance of Statistics
4. Limitations of Statistics

Structure:

1.1: Meaning of Statistics:

The word statistics is generally used in two ways: one as '*data',* and the other as '*methods in statistics'*. In the case of the first one, statistics stands for data. The statistics (data) of rice production in India is an example of this type. Such statistics are found wherever records are collected and maintained in numerical and quantitative forms. Here the use of the word 'Statistics' is in a **plural sense** employed to denote only a collection of facts in figures.

In the second case also the word is used in plural form. It stands for all the principles and devices used in the collection, analysis and interpretation of quantitative statements of facts.

When the word statistics is used as a science of statistics, it is used in the **singular form**, denoting just a branch of applied mathematics. It is also customary to use the word 'statistics' which stands for a measure of formula employed in statistical studies, like an average, dispersion, coefficient of correlation etc.

1.2: Definitions of Statistics:

Statistics has been defined variously by different authors in different times. The following are some of the important definitions of Statistics.

"Science of Counting" — Bowley.

"Science of estimates and probabilities" — Boddington.

"Statistical methods are methods specially adapted to the elucidation of quantitative date effected by a multiplicity of causes" — Yule.

"Statistics is the method of judging collective natural or social phenomena from the results obtained by the analysis of an enumeration or collection of estimates" — W.I. King.

"The science of Statistics is a study of the methods applied in collecting, analyzing and interpreting quantitative data, effected by multiple causation in any department of enquiry" — Ghosh and Chaudhry.

"Classified facts respecting the condition of the people in a state …. especially those facts which can be stated in numbers or in tables of numbers or in any tabular or classified arrangement" — Webster.

"Statistics are numerical statements of facts in any department of enquiry placed in relation to each other" — Bowley.

"Statistics are measurements, enumerations or estimates of natural or social phenomena, systematically arranged so as to exhibit their inter-relations" — Connor.

"Statistics is concerned with scientific methods for collecting, organizing, summarizing, presenting and analyzing data, as well as drawing valid conclusions and making reasonable decisions on the basis of such analysis" — Murray R. Siegel.

The last two definitions given above can be said as reasonably adequate definitions. From the above definitions, we can understand that Statistics must possess the following characteristics:

1.3: Characteristics of Statistics:

1. Numerical statements of facts: Statistics are numerical facts. If they are described in qualitative manner they should be reduced to definite numerical quantities. For example, good, average and poor are qualitative terms. To understand in quantitative terms, they should be defined as – good students are those who secure over 60% marks, those securing between 40 and 60% are average students and those below 40% are poor.

2. Aggregates of facts: Statistics do not take into account individual cases. One student gets first class marks or that he is a good student, does not constitute Statistics unless the total number of students appearing in the examination is given out, of which so many passed, and in such and such divisions. Studies pertaining to individuals are not significant from statistical point of view, for conclusions cannot be drawn by means of comparison and also the figure cannot be treated otherwise. In order to advance the study it is necessary that other observations must be made available.

3. They should be capable of being related to each other: It is not significant as to how many students have passed in an examination unless it is known how many appeared, how these figures compare with similar figures of the previous years, and how do they compare with the figures of other sections of the same class, etc.

4. They must have certain objects behind them: Statistics must be collected for a pre-determined purpose. The figures must relate to a department of enquiry. Sets of figures without any object behind them are not capable of being placed in relation to each other. If in a school there are 500 students and 15 teachers, these figures may constitute statistics, because here the object may be to find the student-teacher ratio, but if instead of teachers we give the strength of class IV employees, there is obviously no object behind such a study. All aggregates of facts must pertain to a department of inquiry in order that they may be designed as Statistics.

5. They are affected to a marked extent by a large number of causes: There should not be only a single factor responsible for bringing about a change in the series. As the height increases, the weight also increases. It is a physical phenomenon. But the increase in weight is not caused by an increase in height alone; there are a large number of other factors also, viz., climate, diet, racial characteristics etc. If there is only one factor operating at a time, the study ceases to be significant from statistical point of view.

6. Reasonable standard of accuracy must be maintained in collection of statistics: Statistics deal with numbers. Sometimes they have to deal with very large numbers so mush so that it becomes impossible to observe each one of the items individually. It, then, becomes necessary to observe a sample and to apply the result to the entire group. We must be satisfied if the results of the smaller group are almost identical to those of the larger group. The term ‘reasonable standard’ is relative, depending upon the object of the enquiry and the resources available.

1.4: Functions of Statistics:

The following are the various functions of statistics.

1. **Measurement Phenomena:** Statistics provides measurement to social phenomena. In this respect it has two types of functions to perform. If there is already a scale of measurement we try to collect data according to it and if there is no standard scale of measurement we try to provide one through statistical analysis and evaluation of variables involved. Thus the first category of functions includes collection of all types of data. Some of the data can be collected by means of actual counting while others have to be estimated.

The other function of Statistics is providing standard scale of measurement where it does not exist. Most of the social phenomena are qualitative in nature and we do not have standard scale of measurement. For example, we generally say that the standard of living of a person is high or low, but we cannot give the exact measurement of it. Index numbers and scaling techniques of Statistics provide quantitative measurement.

2. Description of facts: Statistics provides description of fact by means of numbers. We can know about the magnitude of child marriages, or drinking through the statistics of these facts. Similarly, we can have a clear picture of the unemployment situation in the country only when we have the figures of unemployed people, duration of unemployment, the type of work that they can do size of their family, any supplementary source of income and so on. Statistics tries to introduce further clarity by means of the use of graph, diagrams, charts etc.

3. Objective valuation of phenomena: Qualitative descriptions are generally subjective in nature and may differ according to persons own idea of its magnitude. This gives rise to the lack of uniformity. Statistics, by providing standard scale helps eliminating element of subjectivity. Different people may give different impression regarding the crime situation in a country but when we express it in numbers there can be only one description. Statistics thus helps in objective and accurate valuation of a social phenomenon.

4. Trends and Estimates: Statistics tries to find out the direction and magnitude of change in a phenomenon over time. With the help of these we can find out its position in the near or distant future by projecting the trend further. For example, we generally find that the population of a country tends to rise regularly. By measuring the rate of growth we can forecast population on any future date.

5. Comparative study: Statistics provides the facility of comparative analysis. This comparison may be on the basis of time, place or facts. Comparison is made possible through quantitative measurement. For example, the health of two towns can be compared through death rate. Intelligence of two or students can be compared by means of Intelligent Quotient (I.Q). By giving the figures for the crime we can compare the administrative efficiency and police administration of two places. Statistics by providing a common measurement helps in the comparison. The change in the price level overtime can be compared by means of index numbers.

6. Degree of relationship: With the hope of statistical analysis we try to establish relationship between any two or more variables. This is done through various complicated statistical measures like coefficient of correlation, association of attributes, co-variance etc. The more important thing about statistical inference is that we not only find out that two variables are correlated but we can also locate the degree of relationship.

1.5: Scope and Importance of Statistics:

Statistics has become as wide as to include in its fold all quantitative studies and analysis relating to any department of enquiry. This, indeed, give the science of Statistics a very wide scope and one would thing that Statistics has almost an unlimited scope.

The chief importance of Statistics lies in providing the quantitative measurement to a phenomenon. Lord Kelvin rightly says, *"when you can measure what you are speaking about and express it in numbers, you know something about it, but when you cannot measure it, when you cannot express it in numbers your knowledge is of a meager and unsatisfactory kind"*. Quantitative measurement is the sign of the growth of particular discipline and our knowledge and control over the phenomenon. We are no longer satisfied by casual remark that the prices are rising, we must know how much they have risen. This we can do my means of index numbers of prices. It will no longer satisfy us to say that India has economically improved since independence. We would like to have exact measurement. This can be provided by figures of national income and per capital income.

Quantification of social phenomena is the basis of objective observation. Qualitative description is by nature haphazard, not standard and subjective. If two persons are asked to comment about the standard of living of a person, they are very likely to give different opinions. This can be avoided only when we have found out an exact measurement of the standard of living. Similarly, if intelligence of boys were to be expressed in the qualitative terms it would not give a clear picture to us. But if the same were to be expressed in terms of examination marks or I.Q. there will be no difficulty in understanding it and also there will not be variety of opinions about it.

Statistical analysis brings greater precision to our thinking. When facts are reduced to arithmetical figures all argument comes to an end and conclusion can be challenged only by counter Statistics. Figures never lie. They will put plain facts in the coldest and most detached way whatever may be the outcome.

As we are moving more towards social planning we have to base our policy upon aggregative figures. This is not much of consequence whether a person has committed suicide under some strange circumstances, what is important for social planners are the fact whether there has been a fall in the number of suicides. We can never remove suicides from the society. What is of consequence is therefore, whether the number of crimes and their seriousness in increasing or decreasing. Despite of our best efforts the accidents must occur. We as social planners are mainly concerned whether the accidents have shown a declining tendency.

Statistics is equally important in the evaluation of social reforms and nature and extent of social evils. Nothing can give clearer picture of the evil of drinking than the figures regarding the cases of suicides, indebtedness, high death rate and incidence of disease in the families of those who are drunkards. Similarly, the usefulness of prohibition can also be judged by the facts.

Statistical methods are becoming more and more popular among the social sciences. Successful attempts have been made at providing standard quantitative measures of phenomena which ahs hitherto remained qualitative in nature. We are moving more towards perfection and precision with the use of these refined tools of analysis.

1.6: Limitations of Statistics:

Like other sciences, statistics also has its limitations. They are as follows:

1. **Unable to express quantitatively:** Statistics cannot be applied to those facts which are not capable of being quantitatively expressed. Such facts should first be reduced to precise quantitative terms. For example, we cannot compare ‘culture’ of two countries unless we specify by ‘culture’ of two countries unless we specify by culture so many industries, hospitals, educational institutions, places of worship, law courts, etc. Statistical studies

cannot be brought to bear upon such phenomena unless we express them in definite mathematical quantities. Similarly, it is not possible to study 'prosperity', 'intelligence', 'honesty', 'youth' etc., unless we specify them as standing for certain requisite quantitative standards.

2. **Not applicable to studies of individuals:** Statistics does not take cognizance of individual items because they are aggregates of facts. It is unimportant as to what are the marks secured by a student in a certain class test, unless we know the marks of all the students and draw conclusions on that basis. Marks of one student do not constitute statistics, because one of the characteristics of Statistics is that they should be capable of being placed in relation to each other. Individual items cannot be placed in such a relationship

3. **Statistical laws are true only on an average and in the long run:** The quantitative nature of Statistics is true only on an average and in the long run. For example, the theory of probability says that if we toss a coin twice, one time it may fall head upward and a second time head downward. But it is possible that both the times it may come head upward or head downward. This possibility of 50 per cent times heads upward and 50 per cent times head downward will be approximately true if this experiment is repeated a larger number of times.

4. **Statistics often leads to false conclusions:** Statistics often leads to false conclusions, generally, in cases where Statistics are quoted without context or details. For example, in a certain competitive examination in the subject Computers the students of Andhra University have done better than those of Osmania University, it does not mean that the former University has a better standard. It is possible that the students of Andhra University may have been trained in special course in Computers while those of Osmania University may not have enjoyed such facility.

5. **Uniform data always not possible:** The statistical data must be uniform and its main characteristics must be stable throughout the study. It is not possible to compare the wages in two factories if the average wage is composed of adult wages in one, and of the wages of adults and children in the other. The data must be highly uniform and homogeneous.

6. **Only one among various methods:** Statistical methods are not the only method of finding the value of a group. There are other methods of studying a problem besides statistics.

7. **Wrong handling:** Statistics must always be handled by experts; otherwise, they give wrong results.

Distrust of Statistics: There is a popular feeling that statistics is undesirable. According to Gladstone, 'there are three degrees of comparison in lying – lies, damned lies and Statistics'. There can be no more damaging statement than this regarding the utility and seriousness of purpose of Statistics with which skilled students of the science works.

It is however, a mistake to apply these limitations to Statistics only. There are other sciences also which suffer from these limitations. Some of these limitations emerge from the very nature of the science. For example, statistics is applicable to quantitative studies only; so also is Mathematics, Astronomy etc. Naturally, therefore, when these limitations of Statistics are described, it is often forgotten that these are the features which distinguish Statistics from other sciences. Hence these should not be stated as its limitations. Similarly, the laws of Statistics are true on an average and not necessarily in all cases, just as there are exceptions of laws and rules in other social sciences. So far as the drawing of fallacious conclusions is concerned, such wrong conclusions are capable of being drawn in all cases where precise understanding of problems is lacking. Then Statistics requires the data to be uniform. Such uniformity is necessary in all cases where comparisons are to be made. Uniformity is required to render the data comparable on an equal footing. Similarly, this should not be considered as a limitation of Statistics that there are other methods besides Statistics to study a problem, just as there are several systems of medicine by which a particular disease can be cured. We do not say that the existence of the several systems of medicine is a limitation of each one of them. Thus, Statistics have limitations like any other science, and such limitations can be avoided if it is used by experts in this field.

1.7: Summary:

Statistics is inevitable for any type of quantitative measurement. It is characterized by numerical statements of facts, aggregate of facts etc. It functions as a measurement phenomenon, it describes the facts and it finds out the direction and magnitude of change in phenomenon overtime. Statistics has a wide scope and its importance is well known in all spheres of quantitative measurement. Though it has certain limitations like any other science, they can be avoided mostly, if statistical tools and methods are used by experts in the field.

1.8: Glossary:

Statistics: It is the science which deals with the collection, classification and tabulation of numerical facts as the basis of explanation, description and comparison of phenomena.

1.9: Self Assessment Questions:

1. Define Statistics and explain its characteristics.
2. “Statistics is the science of counting”. Give the functions of Statistics.
3. Explain the importance, scope and limitations of statistics.

- Dr.R.Jayaprakash Reddy.

LESSON 2

Statistical Enquiry - Collection of Data

2.0 OBJECTIVE

After studying this lesson you should be able to understand the following :

1. What is statistical enquiry.
2. How to collect the data.
3. Statistical System in India.

STRUCTURE OF LESSON

2.1 Introduction

2.2 Statistical Inquiry - Methods

2.3 Primary data - Methods of Collection of Data.

2.4 Drafting the Questionnaire for Collection of Data

2.5 Sources of Secondary data

2.6 Differences between Primary and Secondary Data.

2.7 Statistical System in India

2.8 Exercise

2.1 INTRODUCTION

Statistical enquiry means search for knowledge. It is also known as statistical investigation or survey. Statistical investigation is a technical job which requires specialized knwoledge and skill. It uses statistical methods. Statistical investigation provides answers to various management problems.

'Griffin'defined statistical enquiry as "Statistical enquires have always required considerable skill on the part of the statistician, rooted in a broad knowledge of the subject matter area and combined with considerable ingenuity in over coming practical difficulties.

Statistical enquiry is two types.

Statistical Enquiry

Enquiry for Special Purpose

Enquiry for General Purpose

Enquiry for Special Purpose : It related to that field in which we have special mission to fulfil.

Enquiry for General Purpose : It may relate to the fulfilment of any objective under consideration for which data are collected.

2.2 STAGES IN STATISTICAL INQUIRY

A statistical equiry is a comprehesive process which passes through the followig stages :

2.2.1 Planning the Statistical Inquiry

A proper planning is essential before a statistical investigation or inquiry is conducted. Careful planning of statistical investigation is essential to get the best results at the minimum cost and time. Following points should be considered in statistical inquiry.

1. Objective of the inquiry should be clear
2. Scope of the inquiry should be determined
3. Scope of the information should be decided
4. Unit of data collection should be defined
5. Source of data collection should be decided
6. Method of data collection should be decided
7. Reasonable standard should be fixed

2.2.2 Execution of an Inquiry

Execution should follow through out the following steps.

i) Collection of data

ii) Editing the data

iii) Presentation of data

iv) Analysis of data

v) Interpretation of data

vi) Presentation of final report

i) Collection of Data

The first step in the conduct of an investigation or inquiry is collection of data. The person who conducts the inquiry is known as an investigator. The persons from whom the information is collected are known as repondents. The persons who help the investigator in collecting data are called enumerators. The sources of collection of day may be primary or secondary. The data may be internal or external.

ii) Editing the Data

Editing the data refers to detect possible errors and irregularities committed during the collection of data. If the data are not edited then it may lead to wrong conclusions. Therefore, editing is essential to arrange the data in order.

iii) Presentation of Data

The collected data is presented through tables, series graph or diagrams. The classified data is to be presented in such a fashion that it becomes easily intelligible or understandable.

iv) Analysis of Data

Once the data is collected and presented, the next step is that of analysis. The main objective of analysis is to prepare data in such a fashion so as to arrive at certain definite conclusions.

v) Interpretation of Data

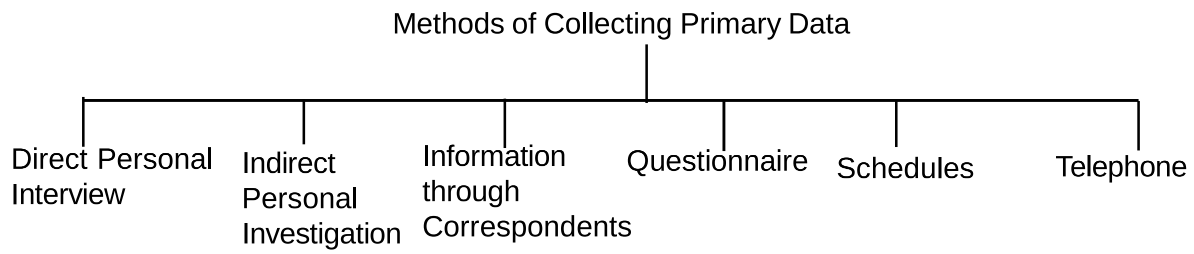
The next stage in statistical investigation is interpretation of data. It means to draw out conclusions from the collected and analysed data.

vi) Presentation of Final Report

The final report is prepared with the analysed data.

2.3 METHODS OF COLLECTING PRIMARY DATA

Primary data is one which is collected by the investigator for the first time. it is also known as first hand information. For instance if the extent of malaria in the city is to be computed, then the information regarding the facts collected by the investigators would be termed as primary data. In India agencies like National Sample Survey (NSS), State Level Economic and Statistical Departments collect Primary data. Following methods may be used to collect the primary data.



2.3.1 Direct Personal Interviews

Under this method of collecting data, there is a face-to-face contact with the persons from whom the information is to be obtained. The interviewer asks them questions pertaining to the survey and collects the desired information.

A. Merits : The advantages of personal interviews are

i) Response is encouraging because of personal approach

ii) The information obtained by this method is likely to be more accurate.

iii) It facilitate to collect supplementary information about the informant's personal characteristics.

iv) This system avoids inconvenience and misintepretation on the part of the informants.

B. Demerits : Important limitations of the Personal Interview method are :

i) It is very costly method of collection of data, if the number of persons to be interviewed is large and they are spread over a wide area.

ii) The chances of personal prejudice and bias are greater under this method.

iii) More time is required for collecting information by this method.

2.3.2 Indirect Oral Investigation

Under this method of collecting data, the investigator contacts third parties or witnesses capable of supplying the necessary information. This method is generally adopted by Government Committees. This methos is useful when the direct sources do not exist and cannot be relied upon.

Merits :

Following are the important merits of indirect oral investigation.

i) The investigator can take the help of expert enumerators to collect the data.

ii) Intensive and extensive investigation is possible

iii) It is economical

Demerits :

i) If the enumerator is not skilled then wrong data may be collected.

ii) The chances of personal bias are greater.

2.3.3 Information through Correspondents

Under this method the investigator does not collect the information from the persons concerned directly. He appoints local agents in different parts of the area under investigation. These local agents are called correspondents. These correspondents collect the information and pass it on to the investigator from time to time.

Merits :

i. It is cheap and economical

ii. It covers large area

iii. It is useful when regular information is required.

Demerits :

i. The chances for personal biase are greater.

ii. The collected data may not be uniform.

2.3.4 Questionnaire Method

In this method, the necessary information is collected from the respondents through a questionnaire. A questionnaire is a set of questions relating to the enquiry.

Merits :

i) Wide coverages is possible

ii) It is economical because no enumerators are required.

iii) It saves time.

i) It is unefected by the personal bias

Demerits :

i) It is costly because enumerators have to be paid

ii) It is time consuming

iii) It can be employed only by big organisations.

2.3.5 Schedules

Another method of collecting information is that of sending schedules through the enumerators or interviewers. The enumerators contact the informants, get replies to the questions contained in a schdule and fill them in their own handwriting in the questionnaire form. This method is free from most of the limitations of the mailed questionnaire method.

Merits :

i) It can be adopted in those cases where informants are illiterate.

ii) There is very little non-response

iii. Information received is more reliable as the accuracy of statements can be checked by supplementary questions wherever necessary.

Demerits :

i) It is costly

ii) It is time consuming

iii) It requires trained enumerators

iv) It can be employed only by big oraganizations.

2.3.6 Telephone Interview

The investigator may also obtain information on telephone. For instance the television viewers may be asked to comment on certain programmes on phone.

Merit :

i) This method is less expensive.

ii) The scope is wide.

Demerits :

i) A limited group can be approached

ii) Very few questions can be asked

iii) The respondents may give vougue and reckless answers.

2.4 DRAFTING THE QUESTIONNAIRE

Before framing the questionnaire it is essential to frame in detail the data which we desire from the answers to questionnaire. The success of the questionnaire method of collecting information depends largely on the proper drafting of the questionnaire. The following general principles may be helpful in framing a questionnaire.

a. The questions should not be too lengthy

b. A decent paper and printing is to be chosen.

c. The qustions asked should be well worded and shuld not be ambiguous.

d. The questions asked should be in proper sequence.

e. Irrelevant questions to the study should be avoided.

f. Questions should be free from personal bias and they should not injure the writing wok.

g. Necessary instructions and definitions should be given.

h. Questions involving the mathematical calculations should be avoided.

i. There should be guarantee to keep the answers secret and to use them only for the purpose of said investigation.

j. The covering letter.

2.4.1 Model Questionnaire

STUDY OF CHANGING PATTERN OF CORPORATE MANAGEMENT IN INDIA

1. Name of the Company ....
2. Registered Address ....
3. Line of Business ....
4. Total Piad up Capital ....
   (a) Number of Shares ....
   (b) Class of Shares ....
5. Shares held by Government financial institutions including Banks.
   Ans...................................
6. System of Mangement adopted by your company.
   Ans....................................
7. Where you a managing agency company or a company managed by a managing agent, please
   describe the activity in which the erstwhile managing agency company is now engaged in viz.
   (a) Trading (b) Manufacturing (b) Processing
   (d) Wound up or not (e) Investment (f) Consultancy Service
   (g) Miscellaneous
8. If the managing agency company is currently engaged in consultancy services - please state.
   (a) Whether it is rendering service only to the erstwhile managed company ?
   Ans....................................
   (b) Whether its consultancy service can be availed of by othe companies ?
   Ans....................................

(c) Please elaborate the services rendered by the consultancy company and number of qualified

expects on pay rolls.

Ans. ....

(d) Would you suggest any regulation of the consultancy services companies? If so how?

Ans...................................

9. (a) Do you believe that after the abolition of the managing agency system, a vacuum created in the management pattern and companies are finding it difficult to have suitable managerials to manage the companies ?

Ans...................................

(b) If the answer is 'Yes' what in your opinion should be done to develop the managerial talents in the company ?

Ans...................................

10. (a) Do you think that the provisions of the Companies Act, in relation to management of companies are very cumbersome and that management has to devote more time to comply with different legal requirements than to actual management of the company ?

Ans...................................

(b) If answer is 'Yes' what is your opinion are the cumbersome provisions ?

Ans...................................

(c) Do you think that these provisions are dropped or made less strict there would be no mis-management by those in charge?

Ans..................................

11. (a) Is there any labour participation in the management of your company ?

Ans...................................

(b) Is it possibel in India for labour to participate in management ?

Ans...................................

(c) If the answer is 'Yes' what suggestion you would make for such participation?

i) Labour representatives

ii) Others.

(d) Do you think that to give them some representation on the Board, employees will have some share holding in the company.

Ans..................................

12. Make your comments on the law relating to management of corporations in general and suggestions to make ti more efficient or effective.

13. Please supply one copy of :

(a) Articles of Association

(b) Memorandum of Association

(c) Latest Annual Report.

2.4.2 Diffirences between Questionnaire and Schedule

Questionnaire	Schedule
This method of collecting data can be easily adopted where the field of investigation is very vast.	This method can be adopted where the field of investigation is not very vast
It is less expensive	It is more expansive since it required trained staff
This method is useful only when informants are literate people.	This method is useful even the informants are illiterate people
It involves some uncertainty about the response	There may be no such uncertainty because of direct contact with informants

2.5 SOURCES OF SECONDARY DATA

The data which is not first hand information (primary data) is known as secondary data. Sometimes it is not possible to collect first hand information for want of resources in terms of money, time, etc., in that situation secondary data is used. This data is mainly classified into two categories. These are

a) Published data

b) Unpublished data.

2.5.1 Published Data

The published data may be obtained from various Intenational, National and Local Publications. Following are the main sources of Published DAta.

i) Internal Publications : Certain International Institutions publish reports from time to time regarding economic matters which are of great significance e.g. Annual Report, Balance of Payments published by IMF, Annual Reports of International Labour Organization (I.L.O.) or by World Bank (I.B.R.D.) etc.

ii) Official Publications of Central and State Governments : Generally State and Central Governments collect information regarding important economic variables like national income, savings, investment, employment, etc., and publish it after regular intervals e.g. Report on Currency and Finance, RBI Bulletin published by RBI, Census report published by Census department, Statisical Abstrafts are published by every state government at State level. The data published by Planning commission is also called Secondary data.

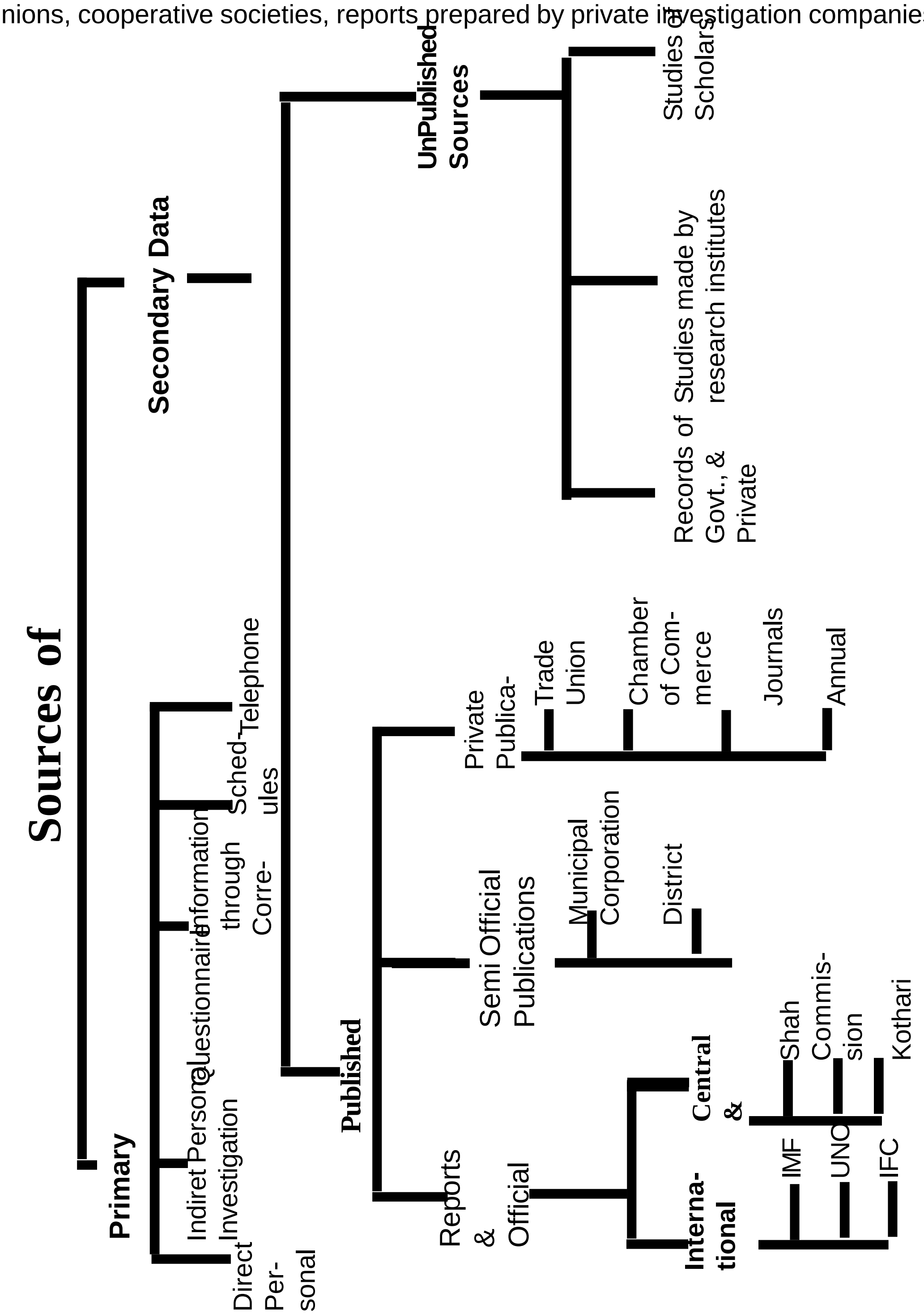
iii) Committee Reports : Sometimes the government appoints survey and enquiry commissions to get the expert views on matters of great importance e.g. Reports of Public Accounts Committee of Lok Sabha.

iv) Newspapers and Magazines : The newspapers like the Finanacial Express, The Economic Times and certain Periodicals like Economic and Political Weekly, Capital, Commerce, Money, etc. Publish the data regarding economic variables.

v) Individual Research Scholars : The various reports of research scholars and research institutions also contain data of economic significance.

2.5.2 Unpublished Data

When the data are collected by someone but which are not published and are taken by other persons for his investigation, they are known as Unpublished Secondary Data e.g. reports of trade unions, cooperative societies, reports prepared by private investigation companies etc.



2.6 Differences between Primary Data and Secondary Data

The investigator must decide whether he will use primary data or secondry data in his investigation. While choosing between the two types of data, following considerations should be kept in mind:

i) Nature and scope of the inquiry

ii) Availability of financial resources

Basis	Primary Data	Secondary Data
1. Cost Factor	Needs more funds	Needs comparateively less funds
2. Source	Investigating Agency collects the data	Some other investigating agency collects it for its own use.
3. Time Factor	Requires longer time for collection	Requiries less time for collection.
4. Reliability and Suitability	More reliable and suitable to the enquiry because the investigator himself collects it.	Less relaible and suitable as some-one else has done that job of collection which may not serve the purpose
5. rganisation Factor	Requires elaborate organisation	No need of any organisational set up.
6. Precautions	No extra precautions are required	Secondary data need more care and attention.

iii) Availability of time

iv) Degree of accuracy desired.

v) the status of the investigator i.e. Individual or corporation or government etc.

In actual practice most of the statistical analysis rests upon the secondary data. Primary data is used in those cases only where the secondary data does not provide an adequate basis for the analysis.

2.7 STATISTICAL SYSTEM IN INDIA

A national statistical system is required to organise the collection, compilation and publication of statistics as important aspects of national life regularly. The system determines the nature, scope and coverage of the statistics to be collected. The national statistical system coordinates the work of the various statistical offices in the country.

2.7.1 Types of Statistical Stystem

A Statistical system can be evaluated from various angles but according to the degree of centralisation there are five types of statistical systems which are given below :

1. Totally decentralised system

2. Minimum Coordination system

3. System decentralised by subject with co-ordinating Agency.

4. System with a central office for general statistics and a co-ordinating agency.

5. Centralised system.

2.7.2 Indian Statistical System

Systematic data collection in India started only with the advent of British rule. Before 1947, no serious attempt was made in our country to collect regular and reliable statistics. The present system of statistical organisation is decentralised in nature. At present each ministry in the centre has at least one statistical unit.

Thus the present statistical system in India is decentralised one, where the authority and responsibility for collection of statistics is divided betwee the Central Government and the State Governments on a subject-wise basis. The Central government acts as the co-ordinating agency for presentation of data on an All-India basis. At the Centre, the Central Statistical Organisation (CSO), a technical wing of the Department of Statistics located in the Cabinet Secretariate now shifted to the Ministry of Planning, New Delhi, acts as a co-ordinator at the national level of all the activities of the Central and State statistical agencies. At the state level, the State Satistical Bureaus attached to various departments in various State Governments, are charged with responsibility of co-ordination of all statistics at State Level.

2.7.3 Statistical Organisation at the Centre.

The Ministry of Statistics and progrmme implementation is the apex body in the official statistical system of the country. The ministry includes the following.

A) Central Statistical organisation (CSO) :

The CSO is located in New Delhi. it is responsible for formulation and maintenance of statistical standards. Its functions are as follows :

i) Perform work relating to National Accounts, Industrial Statistics, Consumer Price indices etc.

ii) Conduct of economic census and surveys.

iii) Training in official statistics

iv. Coordination of statistical activities under taken within the counry and liaising with international agencies in statistical mattes.

The CSO supplies statistical data in the following publications.

1) UN Statistical Year Book

2) Un National Accounts Year Book

3) UN Demographic Year Book

4) UN Monthly Bulletin of Statistics

5) Statistical Year Book of ECAFE

6) Statistical News Letter, etc.

B. National Sample Survey Organisation (NSS) :

National Sample Survey (NSS) was set up in 1950 for conducting large-scale surveys to provide data for estimation of national income and related aggregates especially in the unorganised sector of the economy and for planning and policy formaulation. It carries out annually socio-economic surveys covering various aspects of population. Now its personnel strength is abut 6000 in over 170 offices spread throughout the country.

The NSSO is headed by the Chief Executive officer who is also Member Secretary of the governing Council. Its head quarters are in Calcutta and Faridabad. Its activities are as follows.

i) Survey design

ii) Field Operations

iii) Processing of data collected and reporting of the results.

iv) The role of NSSO in agricultural statistics is to provide technical guidance to states for conducting crop estimation surveys and to keep continuous watch on quality of crop statistics collected by the state Governments.

v) the NSSO collects on monthly basis retail price data from selected shops and markets.

vi) Price indices for urban non-manual employees based on these data are compiled and published.

Survey results are published in the form of reports. About 480 reports are avilable in printed form. NSSO started a quarterly Journal 'Sarve Kshana' from July, 1977. It presents most of the results of NSSO.

At present each Central Ministry has statistical units which are responsible for collection, and compilation of statistics relating to its subject. Important statistical units of the main Central Ministries are as follows.

Ministry of Planning: There are four apex bodies, statistical units responsible for co-ordinating and administrative functions related to the collection of statistics in the country by different departments. These units are -

i) The Central Statistical Organisation (CSO).

ii) National Sample Survey Organisation (NSSO)

iii. Computer Centre.

iv) Programme Evaluation Organisation.

The CSO and NSSO because of their vital importance have alread been discussed.

Ministry of Home Affairs: Thirty statistical units are attached to this Ministry. The main publications of this office are : i) The Census of India Reports ii) Vital Statistics of India (Annual) and iii) Indian Population Bulletin (Biennial).

Ministry of Agriculture and Co-operation: 44 statistical units are attached this ministry. The most important unit is :

Directorate of Economics and Statistics: To compile and publish agricultural statistics on All-India basis this Directorate was established in 1947. The data covered relate to agriculture, live stock, fishery and forestry. The data are collected monthly by the State Governments. The directorate also serves the Central Government in an advisory capacity.

The important publications of the Directorate are

Annual Publications -

1. Indian Agricultural Statistics

2. Estimates of Area and Production of Principal crops in India

3. Indian Agriculture in Brief

4. Indian Livestock Statistics

5. Indian Forest Statistics.

6. Agricultural Prices in India

7. Agricultural Wages in India.

8. Bulletin on Food Statistics

9. Tea Statistics.

10. Coffee Statistics.

Monthly Publications -

Agricultural Situation in India.

Weekly Publications -

1. India Livestock Census

2. Indian Crop calendar

3. Bulletin on Commercial Crops

Ministry of Commerce : Eight statistical units are attached to this ministry important among them are

Directorate General of Commercial Intelligence and Statistics : It was set up in Calcutta in 1895 and the cenral statistical office was responsible for the collection, compilation and publication of important all-India statistical series till the Second World War. With the formation of statistical units in the various Ministries many of the former functions of this office were transferred to the appropriate Ministries. It's now responsible for commercial intelligence and foreign trade statistics. It's main publications are :

1. Indian Trade Journal (Weekly)

2. Indian Customs and Central Excise Traffic Vols I and II(Annual)

3. Annual Statement of Foreign Sea borne Trade of India.

4. Statistics of Maritime Navigation of India (Annual)

5. Accounts relating to the Inland (Rail and River borne) Trade of India (Monthly)

7. Monthly Statistics of Foreign Trade of India by Country and Currency Areas (Vol I and II Monthly)

Office of the Chiel Controller of Imports and Exports : This office publishes annual statistics on imports and exports (Annual Bulletin of Staistics of imports and exports) and annual reports on (Annual Administrative Reports) and weekly reports on licences relating to industrial exports an imports.

Ministry of Labour : The Labour Bureau was established in 1946 in the Ministry of Labour and Rehabilitation. It collects, compiles and publishes statistics of employment in respect of factories, mines, plantations, shops, commercial establishments etc., on an all-India basis. For the formulation of labour policy, it provides data after conducting research into the specific problems of labour. It brings out pamphlets on various aspects of labour legislation. it is also responsible for the construction and publication of consumer price index numbers, for industrial agriculural and rural labour. Its regular publications are :

1. Indian Labour year Book (Annual)
2. Large Industrial Establishments (Annual)
3. Statistics of Factories (Annual)
4. Report on Working of the Minimum Wages Act (Annual)
5. Working of the Trade Unions Act (Annual)
6. Indian Labour Journal (Monthly)

Ministry of Industrial Development : The ministry has seven statistical units. Main statistical unit is the office of the Economic Adviser to the Government of India which was established in 1938. Prior to the setting up of the CSO, it is functioned as the central co-ordinating authority in the field of statistics for the Government of India. Now it maintains wholesale price indices and price data in general and acts as the co-ordinator between various statistical units of the minisry. Its regular publiation is Monthly Statistics of Production of Selected Indusries.

Besides this, the Development Commission Small Scale Industries publishes yearly, monthly and half-yearly reports on the development of small scale industries.

Ministry of Defence: The Army Statistical Organisation (ASO) was set up in 1947 under the Ministry of Defence. it performs the following functions:

i) Maintenance of basic statistical records and the regular computation and supply of data regarding personnel, vehicles, armament, equipment, animals and accommodation etc.,

ii) Control of reports and returns coming from Army and Command Headquarters.

iii) Technical advice on statistics in the army.

iv) Design, conduct and analysis of sample surveys, experiments and investigation.

The ASO has one of the largest installations in India for mechanical tabulation of data. A research unit is concerned witht he development of survey methods and operations research techniques.

2.7.4 Statistical Orgnisation in the States

The apex statistical agency in each State or Union territory is a Statistical Bureau known by different names such as Directorate of Economics and Statistics, Buereau of Economics and Statistics, Directorate of Statistics and Ealuation, Economic and Statistical Organisation, Economic and Statistical Advisory to the State government etc. These are generally under the administrative control of the Finance or Planning Department of the concerned state. The main functions of the State Statistical Bureau are :

1. Systematic Collection, Compilation, analysis, co-ordination, and interpretation of the statistics relating to the States.

2. To act as an advisory body on economic issues referred to it.

3. Organising and conducting special enquiries and field surveys.

4. Liaison between staistical organisation of the Centre and other States.

5. Publication of an annual Statistical Abstract and monthly, quarterly bulleting including all essential statistics of the State.

6. Compilation of economic indicators and State Income Estimates.

7. Statistical Work relating to planning.

8. Publication of Socio-Economic Surveys of the State to be presented in the Budget Session of the State.

2.7.5 Non-Governmental Statistical Organisation

The following non-governmet organisations are working in the country.

1. Indian Statistical Institute, Calcutta.

2. Institute of Agricultural Research Statistics, New Delhi

3. Statisical Department of the Reserve Bank of India.

4. National Council of Applied Economic Research, New Delhi.

5. Institute of Economic Growth, Delhi.

6. Institute of Foreign Trade, New Delhi.

7. Gokhale Institute of Economics and Politics, Pune.

8. Tata Institute of Social Sciences, Bombay.

9. Institute of Labour Research, Bombay.

10. Economic Department of the Reserve Bank of India.

11. Universities in India.

2.7.6 National Statisical Commission (N.S.C.) :

The commission after examining the present system of collection of dissemination of satistics relating to different sectors of the economy adopted a five fold approach to bring about improvements.

1. Reform in the administrative structure of Indian Statistical System and upgrading its infrastructure so as to ensure its autonomy.

2. Improvement of present system of collection of data.

3. Exploration of alternative techniques, in relation to the existing statistics, if the present system for collecting data is under strain for whatever reasons.

4. Idenification of new data series that may be generated in keeping pace with the expanding economy.

5. Evolution of appropriate methodologies for collection of data in relation to new data requirements.

2.7.7. Features of National Statistical Commision

1. NSC has produced a comprehensive report on all aspects of the Indian Statistical System.

2. The commission's approach for improving and strengthening the statistical base has taken the obvious form of recommending about 10 census studies, over 60 types of sample studies and series of other data gathering acivities many of which would be fresh efforts.

3. It would cover not only myriad segments of unorganised or informal sectors but also organised sectors like private corporate sector, NBFC's and even registered factories sector.

4. The Commision has advised the government to exercise caution on enthusiasm shown by government depatments to engage private sector organisations as data collection agencies.

5. It has addressed all issues in their entirety.

2.7.8 Defects of National Statistical Commission

1. There is no sign of any innovation in it.

2. There is no vision of the possible course of changes taking place in Indian polity and the economic structure.

3. If NSC report not focuses on requirements with developmetal objecives.

4. Commission has failed to give proper attention to inadequacies in the estimation of domestic saving and investment.

2.7.9 Suggestions

1. The commission should emphasise on building of Regional Accounts not only at states level but also at an invariant regional grouping states.

2. There is need to break new ground in Industrial Statistics

3. NSC could have suggest the establishment of a system to monitor the progress made in new industrial investment taking place in private sector.

4. There is need to track progress in foreign Direct Investments through the requirement of regular data on projects implemented under FDI.

5. There is need for evolving data on lead in economic indicators.

2.8 APPRAISAL OF INDIAN STATISTICAL SYSTEM

The National Statistical System covers a wide spectrum of national accounts statistics, industrial statistics, export and import statistics, labour statistics, vital statistics, agricultural statistics, environmental statistics, meteorological statistics etc. The statistics are collected under collection of statistics Act, 1953 mainly for the Annual Survey of Industries conducted by the National Sample Survey Organisation (NSSO) and Census Act, 1948.

All is not well with the Nnational Statistical System (NSS), despite the recommendations of the Review Committee in 1979 for making improvements in the system. The Government of India had set up a National Advisory Board on Statistics on 1982, however, the statistical system is suffering from various deficiencies and gaps.

In order to revamp the statistical system in the country, the Government of India has taken two policy decisions:

1. The government of India has borrowed Rs. 850 cores from the World Bank for Revamping the national statistical system in order to bring it at part with the international standards.

2. The Government under the Ministry of Agriculture has set up a National Crop Forecasting Centre (NCFC) for preparing crop forecasts on scientific lines and enable the Government to take strategic decisions on the price front.

2.9 SUMMARY

The Indian Statistical System is still in the process of evolution. With a view to providing a sound statistical base and developing a system of continuous flow of information, the Planning Commission has continuous flow of information. The Planning Commission has constitued two committes, namely, (i) Standing Committee for Improvement of Data base for Planning and Policy Making and ii) Standing Committee for Improvement of Data Base for Dcentralised sectors, consisting of members from government and non-government organisations.

2.10 EXERCISE

A. Short Answer Questions

1. How to plan statistical inquiry

2. What is primary data.

3. Explain direct personal investigation

4. What is meant by C.S.O.

5. Wha are the functions of NSSO.

B. Essay Questions

1. What are the sources of Collection of data.

2. Explain differences between Questionnaire and Schedule.

3. What are the differences between primary data and secondary data.

4. Explain the Statistical System in India.

Dr. K. Kanaka Durga.

Lesson: 3

Classification and Tabulation

3.0 Objectives:

After going through the lesson you will be able to understand the following:

1. Presentation of data and its methods.
2. Classification of data, its need, types and methods.
3. Types of statistical series.
4. Tabulation of data, types of tables and rules for tabulation.

Structure:

3.1: Introduction:

Classification and tabulation of data occupy an important place in Statistics. Unless data are classified properly and tabulated attractively and meaningfully, they won't serve the purpose. In this lesson, all aspects relating to classification and tabulation are discussed. Further, importance of data presentation and statistical series are also discussed.

3.2: Presentation of data:

After the data have been collected and examined, they will have to be presented in a systematic manner either in their raw form as they emerge after editing, or they will have to be statistically treated before their final presentation to the people at large. Generally, data of a simple nature are presented in the form in which they emerge after collection and editing. But data of a more complex nature have to be treated statistically prior to their interpretation. The manner of presentation is very important. If data is not properly presented, it fails to attract due notice and its chief features are not adequately noticed. The data should be so presented that it may be within easy grasp and be swiftly available for easy reference. This is done effectively by graphical or pictorial methods. Whatever method may be employed the chief aim should be to enable one to grasp easily and readily significant proportions, differences or trends in the data.

3.2.1: Methods of presentation: The following methods of presentation are commonly used in Statistics:

1. **Presentation in the form of statements:** Presentation of data in the form of a statement consisting of text and figures is not always effective. It requires careful reading of the text before one is able to understand it. Then it has to be read over again and again as many times as one requires particular information. The main object of Statistics is to simplify complexities. On this score this method does not come up to the mark. The following abstract, taken from the 2007 Wipro Company's report of quarter 2, is an example of presentation of data in the form of statement.

"Wipro has reported a 35 per cent year-on-year revenues for the second quarter ended September 30, 2007 to Rs.4, 785 crore. The net profit stood at Rs.824 crore against Rs.700 crore in the corresponding quarter last year, an increase of 18 per cent. The company has announced an interim dividend of Rs.2 per share. Wipro's Global IT services and products revenue grew only 9.7 per cent sequentially to $796.5 million."

2. **Presentation in the form of classified statements:** When data are of such a nature that they can be broken into two or more parts according to their distinguishing features they may be so presented. A part of information contained in the clearance of Special Economic Zones issued by the government of India is as follows:

"The centre cleared 7 new Special Economic Zones (SEZs), they are—

One SEZ (ITeS) (TCS) - West Bengal

Two SEZs – Tamilnadu

One SEZ (ITeS), Indore	- Mandhya Pradesh
One Malwa IT Park Ltd, Bangalore	- Karnataka
One Perfect IT SEZ Pvt. Ltd, Noida	– Uttar Pradesh
One Calica Construction Impex Pvt, Ltd, Gandhinagar	- Gujarat".

The advantage of this mode of presentation is that figures which are considered as significant may be made to stand out prominently away from the statement. Explanatory notes may be included in the text of the statement. Isolating the figures from the statement makes the data more readily assailable, and avoids chances of confusion.

3. Presentation in the form of tables: This method of presentation makes the data more swiftly understandable as a mass of complex data is broken into several classes and consigned to appropriate columns in the tables. The title of the table gives brief account of its contents, and if the title is carefully selected it may become sufficiently self-explanatory. It gives the entire information intended to be conveyed in a brief and precise manner which it is very easy to scan. Particular attention can be invited to certain facts and figures by stating the facts in footnotes, and the figures in bold letters.

4. Diagrammatic and Graphic Presentation: This method is generally used as a visual aid, and is gradually coming into prominence. Its importance is being recognized as an effective mode of presenting data.

3.3: Classification of Data:

Classification is the process of dividing things into different classes or sequences according to the affinities of their character which exist among a diversity of features in them. The process of classification, if carried to its logical conclusion, means that there should be as many classes as items to be classified, because, while they will have some features in common, in several other respects they will be different from each other. Such a classification, then, would lose the very purpose for which it is made. It is, therefore, enough if we classify items according to the object in view. The object of inquiry will determine as to how facts should be separated into groups or classes according to characteristics needed to be studied for the purpose of the investigation.

3.3.1: Need for Classification:

The most important function of Statistics is to simplify complexities. A large mass of complex data is not capable of signifying anything unless it is presented in a proper manner, duly divided into groups with respect to some characters which are of a variable nature.

The chief object of classification, therefore, is to rid the data of its complex nature and render it easy to understand. Then, since classification is done according to affinity of character, another object of classification is to separate the similar from the dissimilar, and bring out the distinguishing features. Thus, it enables comparisons to be made and conclusions to be drawn without the necessity of considering directly hundreds of individual numbers. Then, since classification is a logical process, it ensures orderly arrangement of items, which is easy to follow and study further. It, thus, serves as mental and visual aid, and renders tabulation easy.

3.3.2: Types of Classification:

Classification may be of two types depending upon the nature of data. If the data is of a descriptive nature, possessing several qualifications which it is possible to classify according to some physical or natural characteristics, it can be classified according to attributes, for example, males and females, Indians and non-Indians, etc.

If, however, data are expressed in numerical quantities, they are classified according to class-intervals, for example, classification of persons according to age-groups as falling between ages 5 to 10, 10 to 25, 25 to 50 years, etc.

3.3.2.1: Classification according to Attributes:

When data relate to persons or things laying emphasis on their physical or natural characteristics, they can be classified according to their qualities. The process of classification according to qualities or attributes consists in isolating the similar from the dissimilar. Things or persons possessing the qualities common to each other are placed in one class. Classification according to attributes may be of two kinds.

1. **Simple Classification:** It is that where only one attribute is studies, for example, classification of persons according to their sex – males and females; according to literacy – literates and illiterates, etc. When one attributes is observed, it results in classification into two classes – one, consisting of those possessing the attribute, another, consisting of those not possessing the attribute. Thus the two classes are strictly exclusive of each other. A simple classification, where items are classified according to one attribute, forming two sub-classes, is also known as classification by dichotomy.

2. **Manifold Classification:** Where more than one attributes are observed, classification may lead to the formation of a number of classes and sub-classes, for example, students are classified as graduate and undergraduate students; among each of the broad classes there are again two sub-classes; males and females, or boys and girls; males and females are further sub-classified as Indians and non-Indians. There is no limit to which we can carry on this process of classification or sub-classification. In the above example the attributes observed are graduate and undergraduate students, their sex and their nationality. More attributes may be observed leading to the formation of further sub-classes.

3.3.2.2: Classification according to Class-Intervals:

When data are expressed in numerical characters and it is necessary to make them easy to comprehend, it is sub-divided into classes constructed out of limits formed either arbitrarily or on grounds of convenience. Such a classification is known as classification according to class-intervals. Sometimes, attributes not capable of precise description are defined by numerical notions, for example, tall and short is a classification according to attribute. But who is a tall person? If population of a town is to be studied for some statistical object, it will not serve any useful purpose if we classify the population as infants, children, young, middle aged, and old but, in order to make the data precise, we shall have to adopt some such numerical notations as: below 5 years (infants), 5-10 years (children), 10-35 years (young), 35-55 years (middle aged) and above 55 years (old).

It is necessary to study certain terms which are used in connection with classification according to class-intervals. Firstly, the classes (viz. below 5, 5-10, 10-15, 35-55 etc.) are known as *class-intervals*. The figures 5, 10, 35, 55 etc. are known as *limits* of the class-intervals. In the class-intervals given above, the first figures in each of them are the *lower limits* of the class-intervals and the second figures the *upper limits* of the class-intervals. The difference between the upper limit and the lower limit of a class-interval is known as the *magnitude* of the class-interval. If this difference is the same throughout the various classes in the class-intervals, the magnitude is known as *uniform,* example, 0-5, 5-10, 10-15, 15-20 etc. The difference is 5 in each case. But if the difference changes in various class-intervals it is known as *un-uniform* magnitude, example, 0-5, 5-10, 10-35, 35-55 and 55-80. Here the difference varies from class to class. In the first two cases it is 5 in each case, in the third it is 25, in the fourth 20 and in the last, again 25. The number of items belonging to each of the classes is called the *frequency* of the class-intervals. If for each of the class-intervals the frequencies given are aggregates of the preceding frequencies, they are known as *cumulative frequencies*, otherwise they are known as individual *class frequencies* or simply frequencies. The frequencies may be cumulated either from the top or from the bottom. The class-intervals are put accordingly.

3.3.3: Methods of framing class-intervals:

The method according to which the above class-intervals, viz. 0-5, 5-10, 10-15 etc., are framed is known as the '*exclusive method'*. Here the class-intervals overlap. In assigning items to various classes, the main difficulty which arises is as to what class should items falling on the limits be assigned, for example, whether '5' should be included in the first class or in the second class, and similarly whether '10' belongs to the second class or to the third class. In the exclusive method, an item which is identical to the upper limit of a class-interval is *excluded* from that class-interval and is included in the next class-interval. Hence it is called 'exclusive method'. An item, the measurement of which is exactly '5' will belong to the second class and not to the first, and so on. For all practical purposes, therefore, the class-interval '0-5' means from '0' to less than '5', '5-10' means from '5' to less than '10' and so on.

There is another method of framing the class-intervals, where the above ambiguity about items identical to a limit of the class-interval is sought to be removed. This method is known as *'inclusive method'*. The above class-intervals according to the inclusive method will read as: 0-4, 5-9, and 10-14 etc. To remove difficulty of an item which is not a complete number and falls between the upper limit of a class and the lower limit of the next class, the above class may be expressed according to inclusive method also as: 0-9.5, 5-9.5, 10-14.5 etc. or 0-4.9, 5-9.9, 10-14.9 etc.

It should, however, be noted that whether the upper limit of the first class is expressed as 5, or 4, o4 4.5 or as 4.9, it would always stand for 'less than 5' and the magnitudes of the class-interval will be 5.

3.3.4: Class-Intervals with Cumulative Frequencies:

Sometimes class frequencies are not given as individual class-frequencies but as cumulative class frequencies. When frequencies are cumulated, the measurement of class-intervals is also cumulated. Frequencies may be cumulated either from the top or from the bottom. The class intervals are not expressed in usual manner with their lower and upper limits, but only with the upper limits preceded by the word 'below', (or 'less than'), or 'above' (or 'more than') as the case may be according to as the frequencies are cumulated from the top or from the bottom.

Before treating such data statistically, it is necessary to convert them into usual class-intervals and individual class frequencies. The following example shows how frequencies cumulated from the top and from the bottom are converted into usual types of data:

1. Class frequencies cumulated from top

Marks	Number of Students
Below 5	10
Below 10	22
Below 15	37
Below 20	50
Below 25	55

2. Class frequencies cumulated from bottom

Marks	Number of Students
Above 0	55
Above 5	45
Above10	33
Above15	18
Above 20	5

The above data converted into usual type of class-intervals and individual class frequencies will read as follows:

Marks	No. of Students
0 – 5	10
5 – 10	12
10 – 15	15
15 – 20	13
20 – 25	5

General Considerations: It is for the statistician to decide about classifications, but some general considerations need to be taken care of:

1. The classification must be exhaustive. It should be possible to include each of the given values in one or the other class.
2. The classes must be mutually; exclusive i.e. they should not overlap. If, however they have to overlap as in the case of exclusive classes, the statistician must observe the rules of classification applicable to such classification.
3. The number of classes should be neither too large nor too small; for either of the practices is likely to undermine the purpose of classification, and upset the pattern of distribution of the frequencies. It is not possible to lay down the number of classes which may be applicable to all situations.
4. The magnitude of class-intervals should be uniform, if possible, throughout the classification, and the system of 'open' classes should be avoided.

3.4: Statistical Series:

According to L.R.Connor, “if two variable quantities cab be arranged side by side so that the measurable differences in the one correspond to the measurable differences in the other the result is said to form a statistical series”. In other words, any logical or systematic arrangement of items constitutes a series. When things or attributes are counted, measured, or weighed and placed one after the other in some orderly manner, they are said to form a series.

As discussed in the above pages, series or arrangement of data can be done on the basis of time, space, or some conditions. So far as time series and spatial series are concerned, there is no problem in their formulation. Frequencies can be noted down on the basis of time or space, however, when series are formed on the basis of changes in some condition like age, weight, marks, production etc., and the series can be either discrete or continuous. Let us discuss about them in detail.

Discrete Series:

When items are arranged in groups showing definite breaks from one point to another, and when they are exactly measurable, they constitute a discrete series. Items are arranged in ascending or descending order and opposite them the number of times each item occurs is mentioned. In a question in which the maximum marks were six, students secured marks as follows:

Marks	No. of Students	Marks	No. of Students
1	5	4	7
2	8	5	6
3	10	6	1

After each marks group 1, 2, 3 and so on there are definite breaks and the students seem to secure exact marks as 1, 2, 3, and not as fractions. Such a series is termed as a discrete, or a broken or a discontinuous series.

Continuous Series:

When items are arranged in groups or classes because they are not exactly measurable, they form a continuous series. Items which are capable of precise measurement should either be placed in a series of individual observations or in a discrete series. But when it is not possible to measure them in exact terms, or if it is possible to so measure them but the measurements, they are entered into classes or groups of measurements.

3.5: Tabulation of Data:

Tabulation of data is the last stage in the compilation of data, and forms the basis for its further statistical treatment. It is a systematic presentation of data in columns and rows. The following are the important definitions of tabulation.

“The logical listing of related quantitative data in vertical columns and horizontal rows of numbers with sufficient explanatory and qualifying words, phrases and statements in the form of titles, headings and explanatory notes to make clear the full meaning, context and the origin of the data” ——— Tuttle.

“Tabulation is the process of condensing classified data in the form of a table so that it may be more easily understood, and so that any comparisons involved may be more readily made” ——— D.Gregory and H.Ward.

3.5.1: Objectives of Tabulation:

The following are the main objectives of Tabulation.

1. **To simplify complex data:** In the process of tabulation of data, unnecessary details are avoided and data are presented systematically in columns and rows in a concise form. All tabular data are presented in such a manner that they become more meaningful and can be easily understood by a common man.

2. **To facilitate comparison:** Data presented in rows and columns facilitate comparison. Since a table is divided into various parts and for each part separate sub-totals and totals are given relationship between various items of the table can be easily understood.

3. **To economize space:** Economy of space is achieved by tabulation, as all unnecessary details and repetitions are avoided without sacrificing quality and utility of the data.

4. **To depict trend and pattern of data:** Tabulation of data depicts the trend of the information under study and reveals the patterns within the figures which cannot be understood in a descriptive form of presentation.

5. **To help reference:** When data are arranged in tables with titles and table numbers, they can be easily identified and made use of, as source reference for future studies.

6. **To facilitate statistical analysis:** After classification and tabulation, statistical data become fit for analysis and interpretation. Various statistical measures like averages, dispersion, correlation, etc., can be calculated easily from the data which are systematically tabulated.

3.5.2: Types of Tables:

From the standpoint of usage, statistical tables are of two types:

1. General Tables
2. Summary or Special purpose Tables

1. **General Tables:** These tables contain a mass of detailed information including all that is relevant to the subject-matter. Hence such tables are very large, extending over a number of pages. The main purpose of such tables is to present all the information available on a certain problem at one place for easy reference. They usually find their place in the appendix of reports or special studies of problems.

2. **Summary Tables:** These tables are designed to serve some specific purposes. They are smaller in size than general tables and seek to lay emphasis on some aspect of data. They are generally contained in the text. They are called summary tables because they are brief and are also called derivative tables because the general tables serve as the source from which they are derived or made. They aim at analysis and comparison of data and enable conclusions to be drawn.

3.5.3: Forms of Tables:

Tables may be simple or complex to form. Let us discuss about them.

1. **Simple Tabulation:** In this type of tabulation, a table contains information pertaining to only one set of related data and seeks to answer one or more groups of an independent investigation. Observe the following table:

Marks	No. of students	Marks	No. of students
0 – 5	15	15 – 20	25
5 – 10	17	20 – 25	20
10 – 15	22	25 – 30	13

Thus, a simple table has two factors placed in relation to each other.

2. **Complex Tabulation:** In this type of tabulation, a table contains information pertaining to a number of coordinate factors. If there are two coordinate factors, the table is called a double table; if the number of coordinate groups is three it is a case of treble tabulation; and if it is a case of more than three coordinate groups the table is known as multiple tabulation.

In the above table, if the students are further classified into groups according to residence, hostellers or day scholars, it will be a case of double tabulation. If the students falling into each of the two groups – hostellers and day scholars – are classified according to sex, it will be a case of treble tabulation. If they are again classified as belonging to different religions, states, nationalities etc., it will constitute an example of manifold tabulation. More than one factor makes the plan of a table slightly complex, and the larger the characteristics distinguished the more complex the table becomes.

3.5.4: Rules and Precautions for Tabulation:

There are no hard and fast rules for tabulation. Experience is the best guide and practice is the best teacher to enable good table to be drawn. The main consideration, however, is that a table should amply fulfill the purpose it is designed to and must make the data readily assailable. With this end in view certain rules of procedure are laid down for the guidance of statisticians.

1. The chief consideration should be to make the table as **simple** as possible, free from all avoidable confusion. Then only it may bring out its chief features and the required information quite easily. Clarity should not be sacrificed at any cost for that is the main function of tabulation. If it is necessary to include a mass of relevant information, it may often be found convenient to break it into two or more tables accompanied by a summary table. Every table must be a unit by itself, dealing with different groups or sections of information. Too many details in a table confuse the eye, and make comparisons and detection of errors more and more difficult.

2. Figures to be **compared** should be placed as near to each other as possible, and absolute figures as well as figures expressed in units of comparison, for example, averages, percentages etc., should be shown for easy comparison. Figures to be compared should be placed in vertical columns as far as possible so that they may be compared easily. Totals to be compared may be given in bold type if it is possible.

3. Every table should be preceded by a suitable **title or heading** describing the contents of the table. The title and the sub-heads etc., should be complete so that it may not be necessary to refer to anything else in order to understand the nature of the table and its objects and contents.
4. **Explanatory notes** should always be given as footnotes and must be complete so that it may not be necessary to refer to something else in order to understand them.
5. The **source** from which the data is obtained must be indicated in the footnote. This is not only courteous but is also helpful to those who use the data in forming their own estimates about the reliability of the data.
6. The **ruling** should be such that major items are separated by bold or double lines.
7. For all important and principal heads there should be **separate columns**, and minor heads may be placed in one column which may be called 'miscellaneous'. The miscellaneous column must contain only those items which are not of a widely varying nature.
8. If certain data are **not available** for inclusion in the table this fact must be mentioned in the footnote by giving a suitable 'mark' (like N.A for not available) in the appropriate place where such data ought to figure.
9. The columns should be properly **'ranged'** by putting thousands under thousands and hundreds under hundreds. This gives an orderly appearance to the table.
10. The **arrangement of items** in the table should follow some logical order. They may either be arranged in order of their magnitude, or in alphabetical, geographical, and chronological or in any other suitable arrangement.

3.6: Summary:

Classification of data serves the purpose of easy understanding. According to the requirement, data can be classified. The data thus classified should be arranged in a systematic manner called series. Later, the data should be tabulated for easy understanding and viewing. Some rules and precautions, if followed, the tables should be attractive and meaningful.

3.7: Glossary:

Classification – It is the process of dividing data into different classes or sequences according to the features in the data.

Series – It is a logical or systematic arrangement of items.

Tabulation – It is a scientific process involving the presentation of classified data in an orderly manner.

3.8: Self Assessment Questions:

1. What is classification? Describe the various bases of classification.

2. Explain various types and forms of tables.

3. What are the guiding principles in the construction of a table?

-Dr. R. Jayaprakash Reddy

Lesson - 4

DIAGRAMS & GRAPHS

OBJECTIVES:

By the study of this lesson you will be able to understand the importance and utility of diagrams and various types of diagrams. You will also be able to understand the importance and utility of graphs and various types of graphs.

STRUCTURE:

4.1 Introduction

4.2 Importance or utility of Diagrams

4.3 Rules or directions for making Diagrams

4.4 Limitations of Diagrams.

4.5 Types of Diagrams.

4.6 One Dimensional Diagrams

4.6.1 Line Diagrams

4.6.2 Simple Bar Diagrams

4.6.3 Multiple Bar Diagrams

4.6.4 Sub - divided Bar Diagrams

4.6.5 Percentage Bar Diagrams

4.6.6 Broken Bar Diagrams

4.7 Two Dimensional Diagrams

4.7.1 Rectangles

4.7.2 Squares

4.7.3 Sub divided Circular Diagrams

4.8 Graphs - Introduction

4.9 Uses of Graphs

4.10 Rules or Guidelines for the preparation of graphs

4.11 Constitution of Graph paper

4.12 Choice of scale

4.1. INTRODUCTION :

Although tabulation is very good technique to present the data, but diagrams are an advanced technique to represent data. As a layman, one cannot understand the tabulated data easily but with only a single glance at the diagram, one gets complete picture of the data presented. According to M.J. Moroney ,-Diagrams register a meaningful impression almost before we think".

4.2 IMPORTANCE OR UTILITY OF DIAGRAMS :

1. Diagrams give a very clear picture of data. Even a layman can understand it very easily and in a short time.
2. We can make comparison between different samples very easily. We don't have to use any statistical technique further to compare.
3. This technique can be used universally at any place and at any time. This techniqe is used almost in all the subjects and other various fields.
4. Diagrams have impressive value also. Tabulated data has not much impression as compared to Diagrams. A common man is impressed easily by good diagrams.
5. This technique can be used for numerical type of statistical analysis, *e.g.* to locate Mean, Mode, Median or other statistical values.
6. It does not save only time and energy but also is economical. Not much money is needed to prepare even good diagrams.
7. These give us much more information as compared to tabulation. Technique of tabulation has its own limits.
8. This data is easily remembered. Diagrams which we see leave their lasting impression much more than other data techniques.

9. Data can be condensed with diagrams. A simple diagram can present what even cannot be presented by 10000 words.

4.3. RULES OR DIRECTIONS FOR MAKING DIAGRAMS

While preparing the diagrams we must observe some rules to make these diagrams more impressive and useful.

1. It must be attractive.
2. Its presentation must be proportionate in height and width.
3. It must be Economical in terms of money, energy and time.
4. It must be Intelligible.
5. Scale must be presented along with diagram.
6. Size of figure should be such that it may occupy considerable portion of paper.
7. It must be self-explanatory. It must indicate nature, place and source of data presented.
8. It must be neat and clean.
9. Diagrams are of several types. The diagram drawn must be suitable to data.
10. If some points are to be clarified, foot notes may be given.
11. Different shades, colours can be used to make diagrams more easily understandable.
12. Vertical diagram should be preferred to Horizontal diagrams.
13. If possible, suitable title may be given.
14. It must be accurate. Accuracy must not be done away with to make it attractive or impressive.

4.4 LIMITATIONS OF DIAGRAMS :

1. Diagrams depict only approximate results. Those are not so accurate.
2. Due to above reasons these can't be put for further analysis.
3. If scales are different, two diagrams can't be compared.
4. For false base diagrams, a lay man may not make difference.

4.5 TYPES OF DIAGRAMS :

Diagrams can be classified into following categories :

(i) One-dimensional Diagrams.

(ii) Two-dimensional Diagrams.

(iii) Three Dimensional Diagrams.

(iv) Pictograms or Picture Diagrams.

(v) Cartograms or Maps.

4.6. ONE DIMENSIONAL DIAGRAMS :

In this case only the length dimension is given the importance. These diagrams are either Bar or Line Diagrams.

4.6.1 Line Diagrams

In these diagrams only line is drawn to represent one variable. These lines may be vertical or horizontal. The lines are drawn such that their length is in proportion to value of the terms or items so that comparison may be done easily.

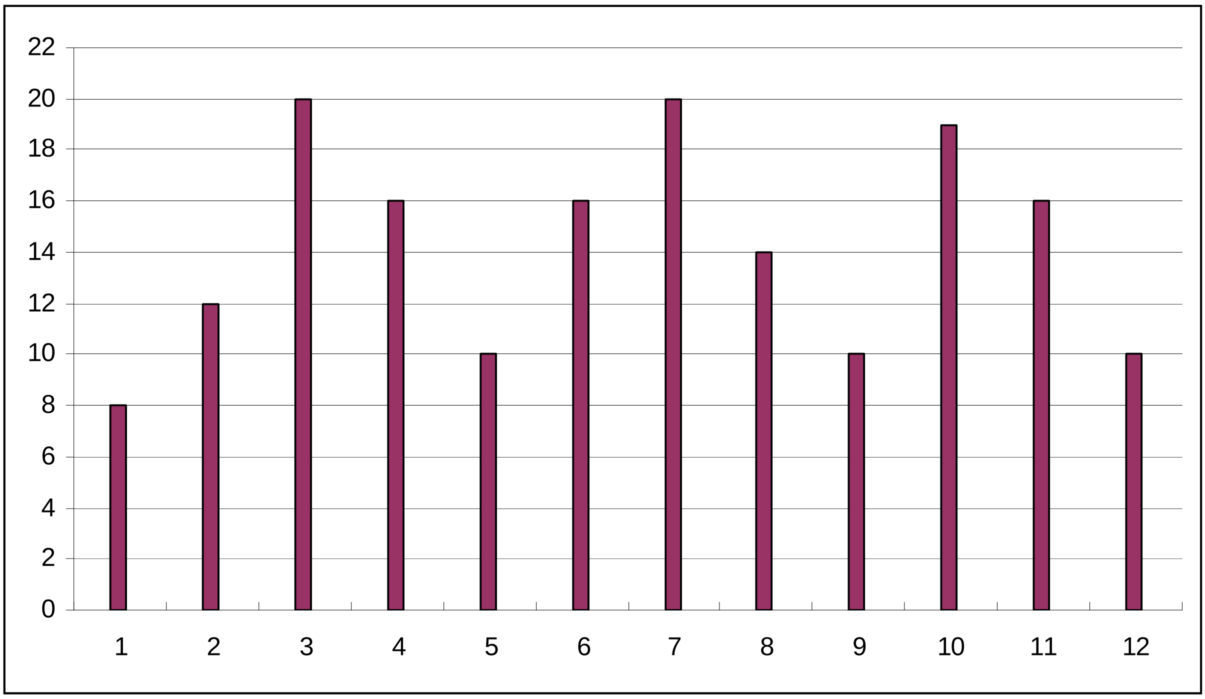
Example 1. No. of accidents in a city in a year is given below :

Month :	1	2	3	4	5	6	7	8	9	10	11	12
No of Accidents :	8	12	20	16	10	16	20	14	10	19	16	10

Solution : Prepare line diagram.

Scale Y - axis 1 cm. = 2 Accidents

On graph 1 Div. = 2 Accidents



4.6.2 Simple Bar Diagram

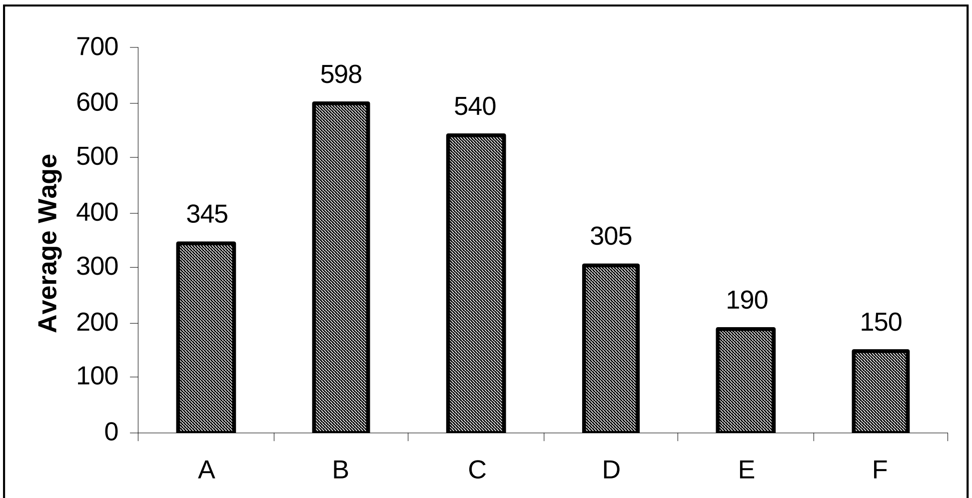
Like line diagrams these figures are also used where only single dimension *i.e.* length can present the data. Procedure is almost the same, only the thickness of lines is measured. These can also be drawn either vertically or horizontally. Breadth of these lines or bars should be equal. Similarly distance between these bars should be equal. The breadth and distance between them should be taken according to space available on the paper.

Example 2. Average wages of some firms are given below. Represent this by simple Bar Diagram.

Firm	:	A	B	C	D	E	F
Average wage	:	345	598	540	305	190	150

Solution :

Scale Y - axis 1 cm. = Rs. 100
On graph 1 Div. = Rs.10



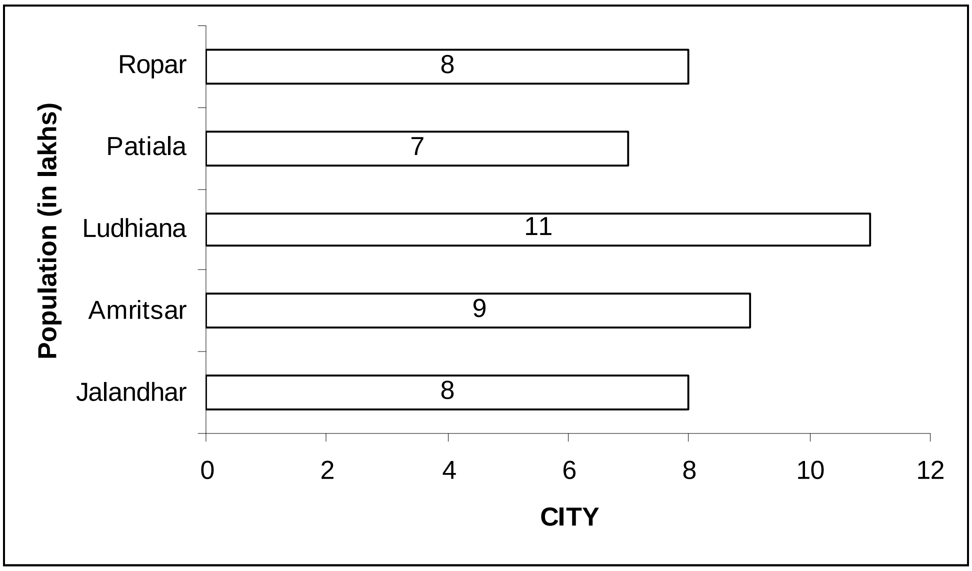
S.69

Example 3. Present the following data by horizontal bar diagram.

City	:	Jalandhar	Amritsar	Ludhiana	Patiala	Ropar
Population (Lakhs)	:	8	9	11	7	8

Solution :

Scale x axis 1 cm. = 2 lakhs,
on graph 1 Div. = 2. Lakh.



4.6.3 Multiple Bar Diagrams :

This diagram is used, when we have to make comparison between more than two variables. The number of variables may be 2, 3 or 4 or more. In case of 2 variables, pair of bars is drawn. Similarly, in case of 3 variables, we draw triple bars. The bars are drawn on the same proportionate basis as in case of simple bars.

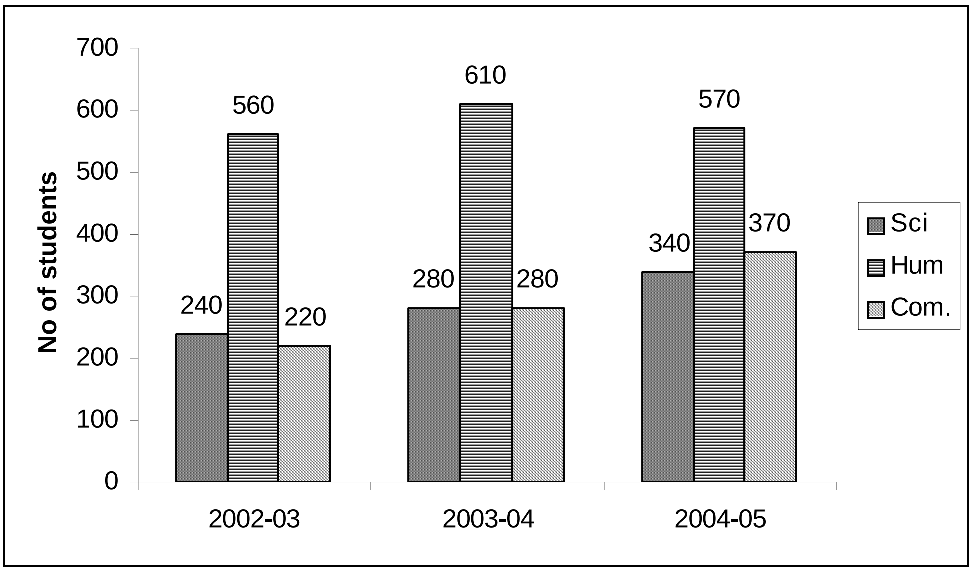
Example 4. No. of students in Postgraduate classes in a university is given below :

	Science	**Humanities**	**Commerce**
2002-03	240	560	220
2003-04	280	610	280
2004-05	340	570	370

Solution :

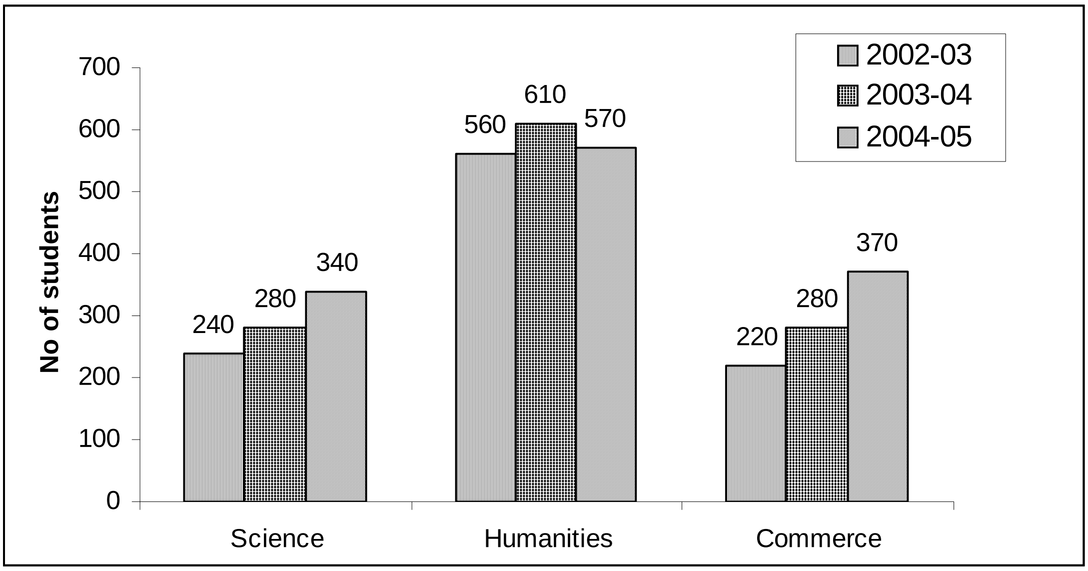
Here biggest item or term is 610, we should take last term in graph as 650 or 700.

Scale x axis 1 cm. = 100 students
on graph 1 Div. = 10 students



We can present this data by multiple bar diagram in the following manner according to requirements.

Scale x axis 1 cm. = 2 lakhs,
on graph 1 Div. = 2. Lakh.



4.6.4 Sub-divided Bar Diagram :

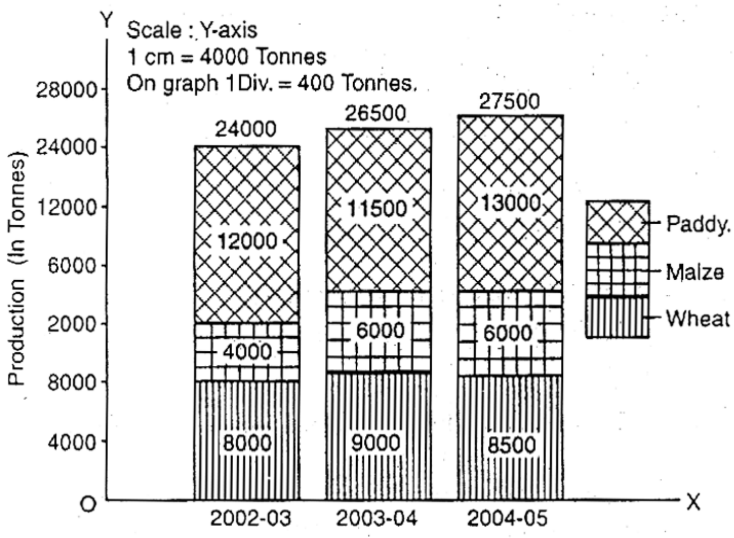
The data which is presented by multiple bar diagram can be presented by this diagram. In this case we add different variables for a period and draw it on a single bar as shown in the following examples. The components must be kept in same order in each bar. This diagram is more efficient if number of components is less i.e. 3 to 5.

Example 5. Production of grains in Punjab is as follows. Present the data by a suitable diagram

Production in Tonnes	Wheat	Maize	Paddy
2002-03	8000	4000	12000
2003-04	9000	6000	11500
2004-05	8500	6000	13000

Solution

	2002-03		2003-04		2004-05	
	Production (Tonnes)	Cumu lative	Production (Tonnes)	Cumu lative	Production (Tonnes)	Cumu lative
Wheat	8000	8000	9000	9000	8500	8500
Maize	4000	12000	6000	15000	6000	14500
Paddy	12000	24000	11500	26500	13000	27500
	24000		26500		27500	



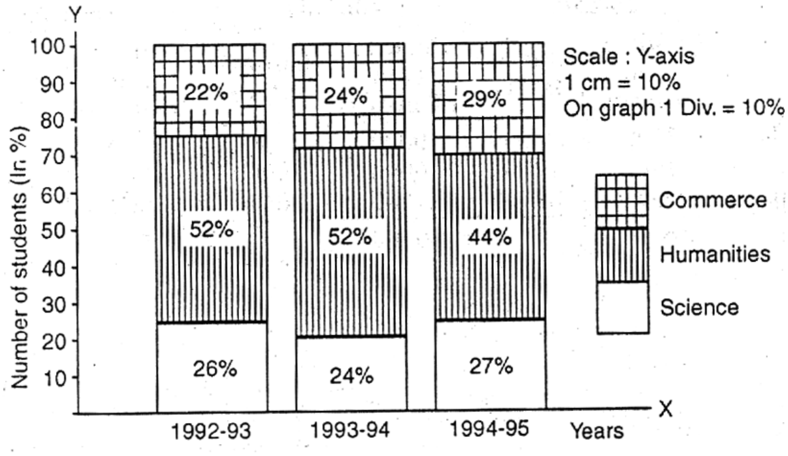
4.6.5 Percentage Bar Diagram :

Like sub-divided bar diagram, in this case also data of one particular period or variable is put on single bar, but in terms of percentages, Components are kept in the same order in each bar for easy comparison.

Example 6. Present data of Example 4 by percentage bar diagram.

Solution :

Subject	2002-03			2003-04			2004-05		
	No.	Cumu.	% Cumu.	No.	Cumu.	% Cumu.	No.	Cumu.	% Cumu.
Science	240	240	26	280	280	24	340	340	27
Humanities	560	800	78	610	890	76	570	910	71
Commerce	220	1020	100	280	1170	100	370	1280	100
	1020			1170			1280		



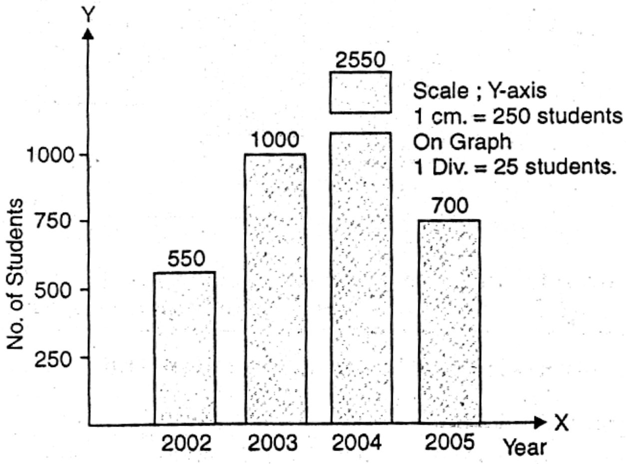
4.6.6 Broken Bar Diagram :

This diagram is used when value of some variable is very high or low as compared to others. In this case the bars with bigger terms or items may be shown broken.

Example 7 : Present the data given below by suitable diagram

Year	:	2002	2003	2004	2005
No.of students	:	550	1000	25550	700

Solution :



4.7 TWO DIMENSIONAL DIAGRAMS :

As in single bars it was mentioned that the width of each bar should be equal for a certain variable or items. But in this case not only the length but the width also is taken proportionately in case of rectangles.

But where the items are represented in square terms we use the technique of squares or circles. These are also known as Area Diagrams.

4.7.1 Rectangles :

As mentioned above, not only the length, but the breadth or width of each item is also taken proportionately.

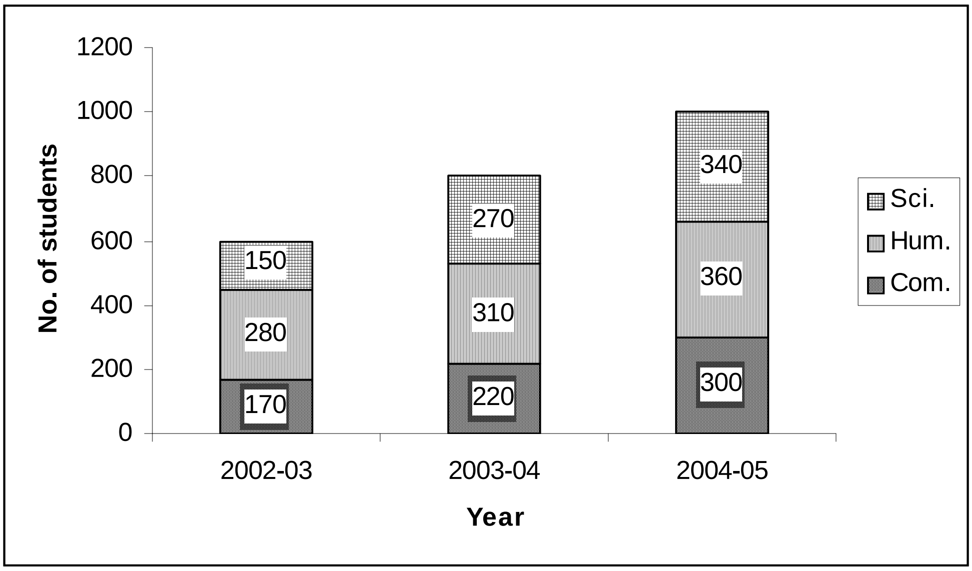
Example 8. Present the data given below by rectangle diagram.

No.of Students	Science	Humanities	Commerce	Total
2002-03	170	280	150	600
2003-04	220	310	270	800
2004-05	300	360	340	1000

Solution :

As the total number of students are in the ratio 3:4:5, we will take the width of bars in this ratio.

Scale Y-axis 1 cm. = 100 students
on graph 1 Div. = 10 students.



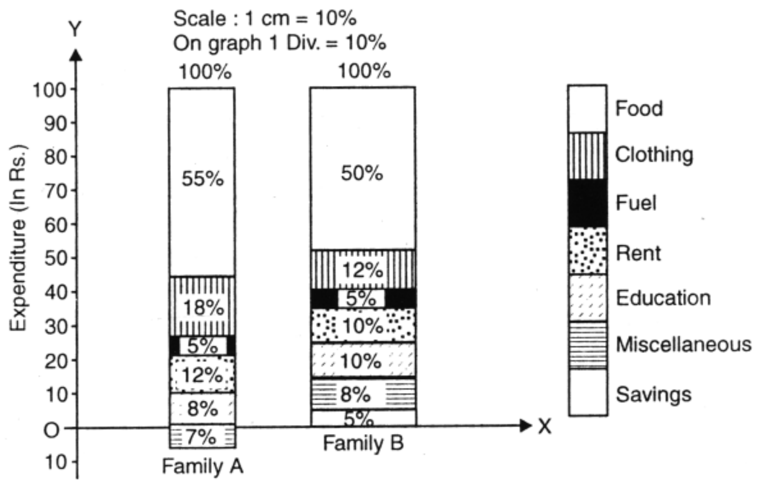
Example 9. Income of two families is Rs. 6000 and Rs. 9000 respectively. Present the following data by percentage rectangle diagram.

Expenditure	Food	Clothing	Fuel	Rent	Education	Miscellaneous
Family A	3300	1080	300	720	480	420
Family B	4500	1080	450	900	900	720

Solution :

As the income of two families are in the ratio 2:3 the width of bars should be in the ratio.

Total Income / Expenditure	Family A			Family B		
	Amount 6000	% age 100%	Cum. %age	Amount 9000	% age 100%	Cum. %age
Food	3300	55	55	4500	50	50
Clothing	1080	18	73	1080	12	62
Fuel	300	5	78	480	5	67
Rent	720	12	90	900	10	77
Education	480	8	98	900	10	87
Miscellaneous	420	7	105	720	8	95
Total	6300	105		8550	95	



4.7.2 Squares :

As told earlier, this technique can be used effectively when given items or terms are squares, preferably having two zeros (00) after every term.

Here we take square root of every item and then divide it by a suitable digit or number so as to get the size reduced to be put into the shape of a square on the given space.

It is also useful technique when difference between the numbers is large.

Example 10. Following is the population of some cities in thousands. Present by a suitable diagram.

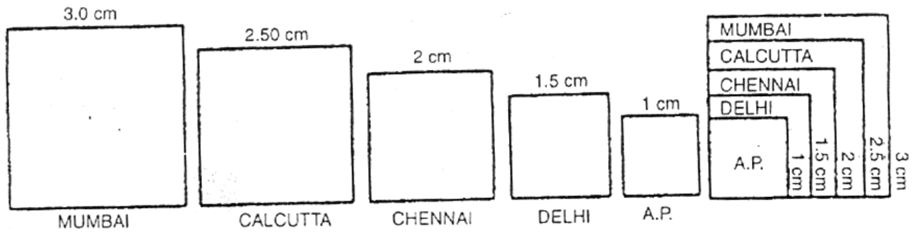
City	:	Mumbai	Calcutta	Chennai	Delhi	Andhra Pradesh
Population ('00)	:	3600	2500	1600	900	400

Solution :

The figures are the perfect squares, hence most suitable diagram will be square or circle. As side = $\sqrt{Area}$; in case of square.

So, we take square root of each term.

City	Population (00)	Square Root	Side (Dividing by 20)
Mumbai	3600	60	3.0 cms
Calcutta	2500	50	2.5 cms
Chennai	1600	40	2.0 cms
Delhi	900	30	1.5 cms
Andhra Pradesh	400	20	1.0 cms



4.7.3 Sub - divided circular diagram :

These are also called Pie or Angular Diagrams. We take the total of items and each item is given its proportionate angle taking the total as 360°. In this case we may have to compare in terms of totals also, if data belongs to two cases.

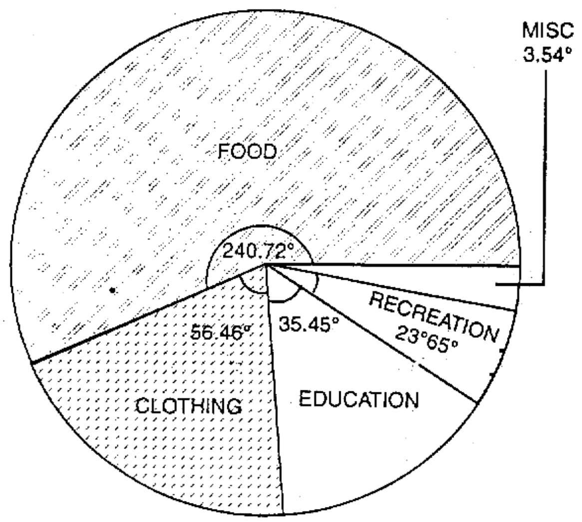
Example 11. Present the the following data through a pie - chart.

Items		**Food**	**Clothing**	**Education**	**Recreation**	**Misc.**
Expenditure (Rs)	:	5100	1200	750	500	75

Solution :

Items	Expenditure (Rs.)	Angle
Food	5100	$\frac{5100}{7625} \times 360^\circ = 240.72^\circ$
Clothing	1200	$\frac{1200}{7625} \times 360^\circ = 56.64^\circ$
Education	750	$\frac{750}{7625} \times 360^\circ = 35.45^\circ$
Recreation	500	$\frac{500}{7625} \times 360^\circ = 23.65^\circ$
Misc.	75	$\frac{75}{7625} \times 360^\circ = 3.54^\circ$
Total	7625	360°.00°

To determine the radius, we take the square - root of 7625, which is 87.32. Divide it by 43.66 to get the radius = 2 cm. Draw a circle with radius = 2 cm and plot the angles obtained in the table. This is called pie-presentation of data.



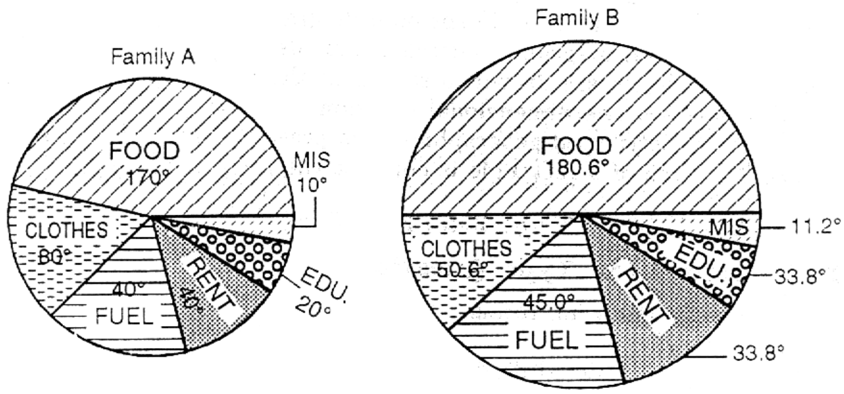
Example 12. Prepare Pie Diagram for the following data.

Items	**Food**	**Clothes**	**Fuel**	**Rent**	**Education**	**Misc.**
Family A:	1700	800	400	400	200	100
Family B :	3300	900	800	600	600	200

Solution :

Item	Family A		Family B	
	Expenditure	Angle	Expenditure	Angle
Food	1700	$\frac{1700}{3600} \times 360 = 170°$	3300	$\frac{3300}{6400} \times 360 = 180.6°$
Clothes	800	$\frac{800}{3600} \times 360 = 80°$	900	$\frac{900}{6400} \times 360 = 50.6°$
Fuel	400	$\frac{400}{3600} \times 360 = 40°$	800	$\frac{800}{6400} \times 360 = 45.0°$
Rent	400	$\frac{400}{3600} \times 360 = 40°$	600	$\frac{600}{6400} \times 360 = 33.8°$
Education	200	$\frac{200}{3600} \times 360 = 20°$	600	$\frac{600}{6400} \times 360 = 33.8°$
Miscellaneous	100	$\frac{100}{3600} \times 360 = 10°$	200	$\frac{200}{6400} \times 360 = 11.2°$
Total	3600	360°	6400	360°

Taking square roots of 3600 and 6400 we get, 60 and 80. We can divide it by a common denominator 40, to get their radii as 1.5 and 2.0 cms.



Note: If there is only single case, we may take any length of radius to suit our space.

GRAPHS :

4.8 INTRODUCTION :

Diagrams can present the data in an attractive style but still there is a method more reliable than this. Diagrams are often used for publicity purposes but are not of much use in statistical analysis. Hence graphic presentation is more effective and meaningful.

According to A.L. Boddington,

"The wandering of a line is more powerful in its effect on the mind than a tabulated statement; it shows what is happening and what is likely to take place, just as quickly as the eye is capable of working."

4.9 USES OR MERITS OR IMPORTANCE OF GRAPHS :

1. It is more effective than diagrams.
2. It is economical in terms of money, times and energy.
3. It gives us the picture in condensed form.
4. It is free from mathematical calculations.
5. It is most suitable for comparison.
6. It is helpful in forecasting.
7. It is also helpful in statistical analysis. We can determine Median and Mode by this method.

4.10 RULES OR GUIDELINES WHILE PREPARING A GRAPH

1. It must have proper heading.
2. Scale must be provided alongwith graph,
3. False base line may be used according to need.
4. If required footnotes may be given.
5. While choosing scale, size of the space must be kept in view,
6. If possible Y-axis should be about 50% more than X-axis,
7. Independent variables should be taken on X-axis and dependent variable on Y axis.
8. In time series graph, time should be shown on X-axis and other variable on Y-axis.
9. In frequency distribution, take value of variable on X-axis and frequency on Y-axis.

4.11 CONSTITUTION OF GRAPH PAPER :

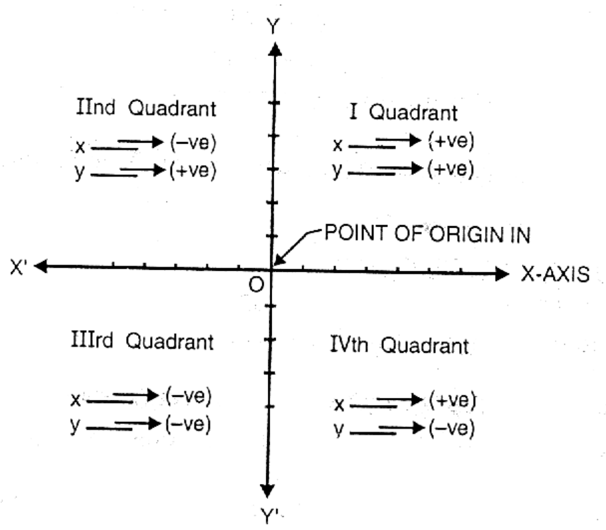
Graphs are drawn on a special type of paper known as graph paper. Graph papers are divided in small equal squares $\frac{1}{10}$ or $\frac{1}{10}$ cm.

For the construction of graph, two straight lines, are drawn which cut each other at 90r. The horizontal line is called 'X'-axis and is usually denoted by X'OX. The vertical line is called Y-axis and is usually denoted by Y'OY. The point where they cut each other is known as 'Origin'.

This- origin divides the graph paper in four parts. These parts are known as quadrants.

Zero value is taken on the point of origin for both lines. Positive values of X are taken towards fight side on horizontal line and of Y towards upper side on vertical line. Negative values of X are taken towards the left side on horizontal line and of Y towards the lower side on vertical line.

Positive and Negative Values :



As shown in above diagram in first quadrant X and Y both have positive values. In second, X is negative and Y positive. In third, X and Y both are negative and in fourth Quadrant X is positive and Y negative.

X-axis is also known as abscissa and Y-axis as Ordinate.

4.12 CHOICE OF SCALE :

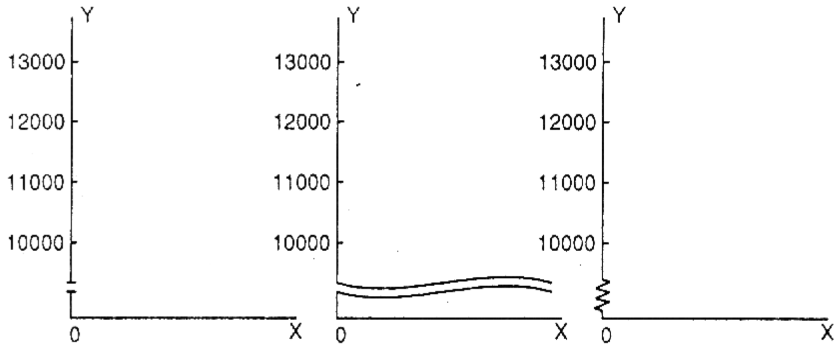
The scale indicates the unit of a variable that a fixed length of axis would represent. Scale may be different for both the axes.. It should be taken in such a way so as to accommodate whole of the data *on* a given graph paper in a lucid and attractive style.

4.13 FALSE BASE LINE :

Sometimes it is difficult to take zero at origin and proceed for the graph as is in the following example :

Year	:	2001	2002	2003	2004
No.of students	:	10320	10860	11400	11200

If we start with zero in this case, first fifty main divisions will remain empty, without any use. In will make the graph look clumsy. In such cases we use false base lines as shown below.



If required we can take flase base line on x - axis also.

4.14 TYPES OF GRAPHS :

There are two types of graphs.

1. Time series Graphs or Historigrams
2. Frequency Distribution Graphs.

4.14.1 Time series graphs or Historigrams :

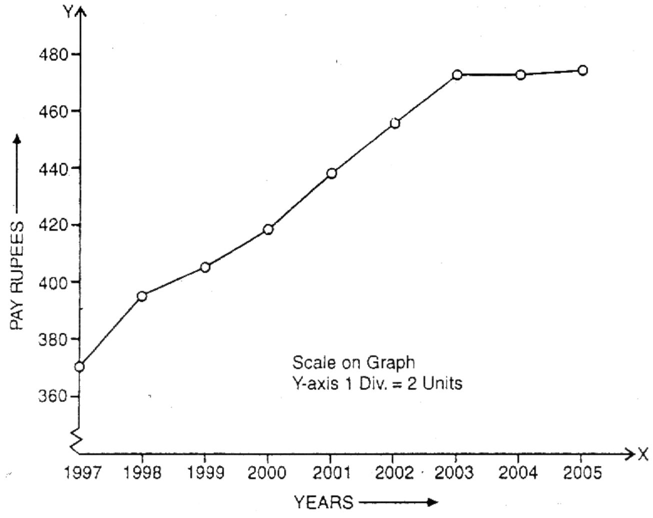
In this type of graphs, time is taken along X-axis and the other variables along Y-axis. The number of variables on Y-axis may be one or more than one. These are known as One Variable, Two Variables or Three Variables graphs.

Example 1. Present following data on a graph paper. (Single variables)

Year	:	1997	1998	1999	2000	2001	2002	2003	2004	2005
Pay/month	:	370	395	405	419	439	456	472	472	473

Solution :

We will take here false base line.

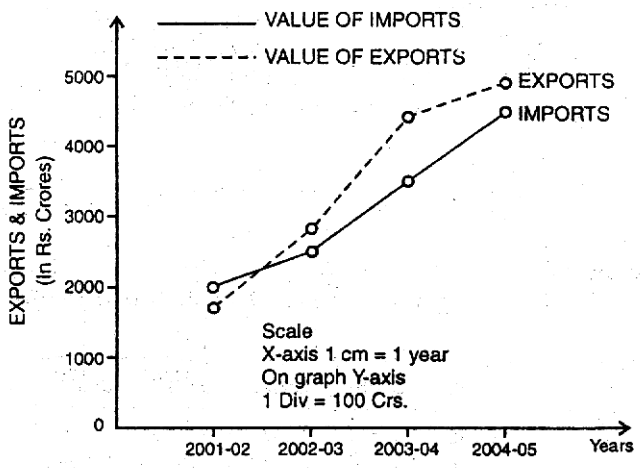


Example 2. Present following data on a graph paper. (Two variables)

Year	:	20001-02	2002-03	2003 -04	2004-05
Value of imports (in Rs. Crores)		2000	2500	3500	4500
Value of Exports (in Rs. Crores)		1700	2800	4400	4900

Solution :

We will take here false base line.



4.14.2 Graphs of frequency distribution :

A frequency distribution can be graphically presented in the following manner :

1. Histogram
2. Frequency Polygon
3. Smoothed frequency curves
4. Cumulative Frequency Curves or Ogives

4.14.2.1 . Histogram :

The term historgam should not be confused with the term historigram which represents time charts. Histogram or column diagram is the best way. of presenting graphically a simple frequency distribution. The classes are marked along the X-axis and by taking class-interval as the base rectangles are erected with heights proportional to the respective classes. Frequencies are measured along the Y-axis. With equal class intervals, all rectangles will have equal base and the area of each rectangle will be proportional to the frequency in that class. In case of unequal class intervals' the width of the rectangles will change and the heights of rectangles shall be proportional to the density of the frequency or the adjusted frequencies.

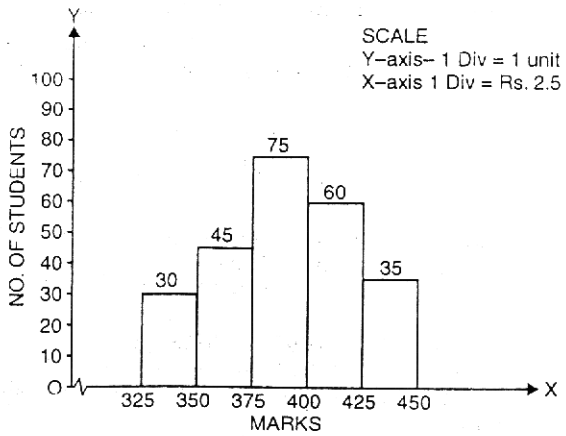
Construction of Histogram with equal class intervals

Example 8. Prepare a Histogram from the following data.

Marks :	325-350	350-375	375-400	400-425	425-450
No. of Students :	30	45	75	60	35

Solution :

Here we take false base line for OX, as smallest term is 325. We take scales for both the axes. For X-axis, 1 Div. = $\frac{25}{10}$ =2.5 marks and on Y-axis ; 1 Div. = $\frac{10}{10}$ = 1 student.



Construction of Histograms with Unequal class-Intervals :

When class intervals are unequal, it is necessary to make adjustment for varying magnitude of class intervals by determining frequency densities. First of all we should decide the class-interval in terms of which the frequency density is to be calculated. The most common interval is generally taken. Then we convert the frequencies of all those classes which have a larger or smaller class-interval to frequencies in terms of class-interval already decided.

Example 9. Prepare a Histogram from the following data.

Marks :	10-15	15-20	20-25	25-30	30-40	40-60	60-80
No. of Students :	7	19	27	15	12	12	8

Solution :

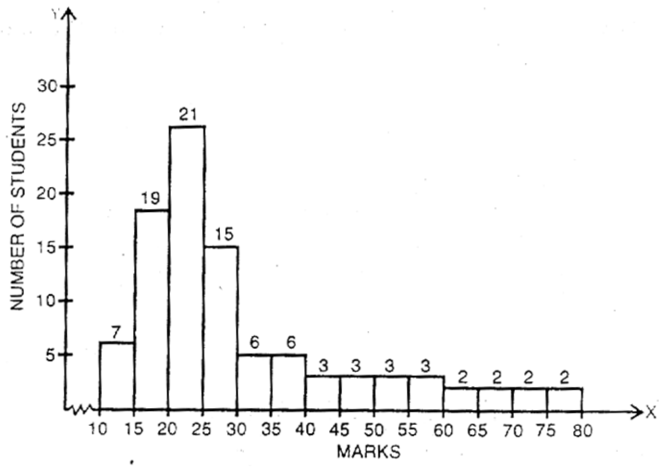
Since the class intervals are unequal, frequencies, must be adjusted. Here the most common smallest class interval is 5. We convert the interval 12 size 30-40 in two intervals and the frequency is divided by 2 *i.e.* $\frac{12}{2}$=6 Similarly 40-60 is divided into 4 and the frequency is divided by 4 i.e. $\frac{12}{4}$ = 3 and 60 -80 also into 4 and its frequency is also divided by 4 i.e. $\frac{8}{4}$ = 2.

Class Intervals (Marks)	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80
Frequency (No. of Students)	7	19	27	15	6	6	3	3	3	3	2	2	2	2

The Histogram is drawn on its basis is given below

4.14.2.2 Frequency Polygon :

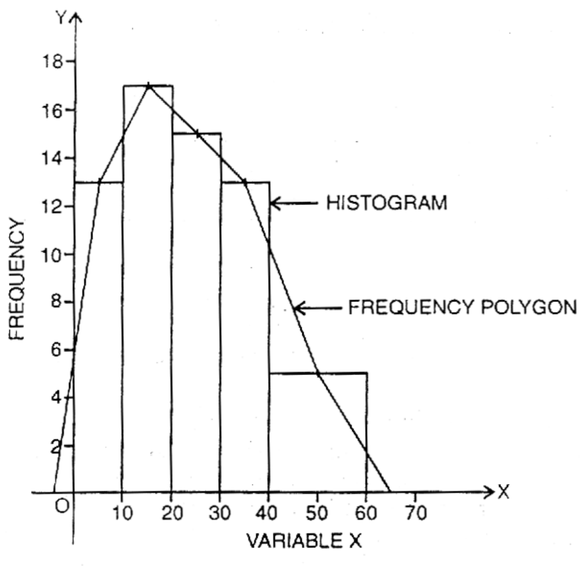
A frequency polygon is a curve representing a frequency distribution. If we join the middle points of the tops of the adjacent rectangles of the histogram, a frequency polygen is obtained. Here both the ends of the polygon are extended to the base line so that the area under the polygon is equal to the area under the histogram. The value of mode can easily be found by forming a frequency perpendicular from the apex of the polygon to the X-axis.



Example 10 : Prepare a histogram and frequency polygon from the following data.

X	f
0-10	13
10-20	17
20-30	15
30-40	13
40-60	10

Solution :



The class 40 -60 is presented with frequency 5 (10/2 = 5)

Example 11 : From the following data, determine the modal value graphically and verify the result by actual calculation.

Profit (Rs.)	Number of Shops	
0-100	12	
100-200	18	9 Δ_1
200-300	27	3 Δ_2
300-400	24	
400-500	10	
500-600	6	

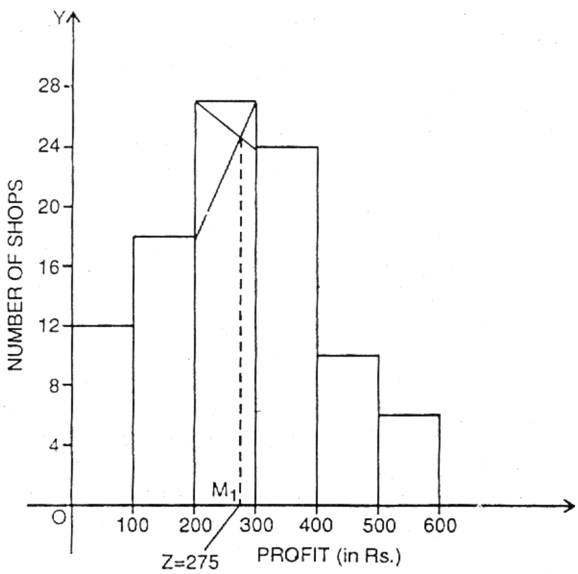
Solution :

Mode is calculated by using the equation

$$Z = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

200-300 is the modal class that is the class with the highest frequency. Substituting the values in the equation we have.

$$z = 200 + \frac{9}{9+3} x100 = 200 + 75 = Rs.\ 275$$



4.14.2.3 Smoothed Frequency Curve :

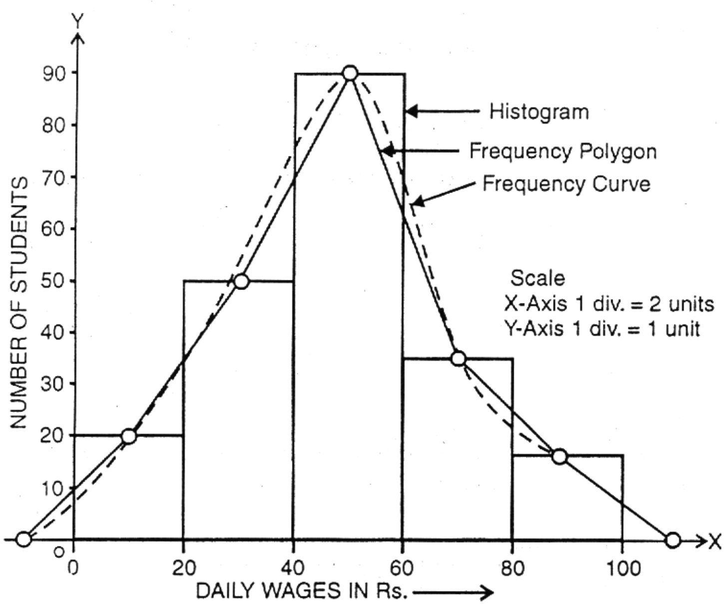
This curve emerges when the points of a frequency polygon are joined by free hand smoothed curves and not. by straight lines. The area of the frequency polygon is equal to that of the histogram. This curve should be based on samples and only continous series should be smoothed.

Example 12 . Draw a histogram, frequency polygram and frequency curve representing the following data.

Daily Wages (in Rs):	0-20	20-40	40-60	60-80	80-100
No. of Students :	20	50	90	38	15

Solution :

Here in addition to the construction of histogram, frequency polygon, frequency curve is drawn by smoothing the corners of the frequency polygon as shown below :



4.14.2.4 Cumulative Frequency Curves :

Sometimes it is necessary to know the number of items whose values are more than or less than a certain amount this case we have to change the form of frequency distribution from a simple to a cumulative distribution. The graphic representatian of cumulative frequency distribution is called the cumulative freqyency curve or Ogive.

There are two methods of drawing ogive -

(i) The less than method and

(ii) the 'more than method.

If we want to know the number of items that are 'less than' a particular size, the cumulation will start from the least to the greatest size and the series will be called 'less than' cumulative frequency distribution. When we want to know the number of items whose sizes are 'more than' a particular size, cumulation will commence from the greatest to the least and the series thus obtained shall be termed as 'more than' cumulative frequency distribution. Ogives - are used to-determine the number'or percentage of cases above or below a certain value. Ogives are also used to compare two or more frequency distributions. Ogives can also be used to determine graphically the values of median,quartile, deciles etc.

Example 13 . Draw the two gives from the following data and locate Median.

Class Interval :	100-200	200-300	300-400	400-500	500-600	600-700
Frequency :	12	18	30	42	68	78

Solution :

The cumulation frequency distributions are given bellow :

Class Interval	Cf	Class Interval	Cf
Less than 100	0	More than 100	240
" 200	12	" 200	228
" 300	30	" 300	210
" 400	60	" 400	180
" 500	102	" 500	138
" 600	162	" 600	78
" 700	240	700	0

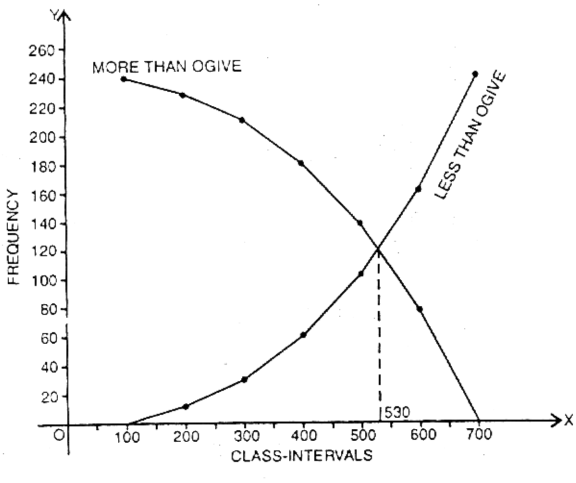
Table showing calculation of Median

Class Interval	f	Cf
100–200	12	12
200–300	18	30
300–400	30	60
400–500	42	102
500–600	60	162
600–700	78	240

Median Class

$= \frac{N}{2}$ th item

$= \frac{240}{2} = 120$th item



So median class is 500 -600

$$M = L + \frac{\frac{N}{2} - Cf}{f} \times i = 500 + \frac{120 - 120}{60} \times 100 = 500 + 30 = 530$$

14.15 SUMMARY :

Both Diagrams and Graphs are simple and attractive. Both can give condensed form to data and help to compare the variables. Even a layman or an illeterate person can easily understand the diagrams & graphs.

14.16 QUESTIONS :

1. What do you mean by a 'Diagram' ?
2. What are the limitations of diagrams ?
3. Define various types of diagrams.
4. Define One or Single dimension diagram.
5. Narrate merits of one dimension diagram.
6. Define line diagram. How will you draw it ?
7. What is simple bar diagram ? How will you draw it ?
8. What is multiple diagram ? When is it used ?
9. What is two dimensional diagram ? Define its various types.
10. When is a rectangle or a square or a circle is used to present a data ?
11. Explain the necessity of diagrams in statistics.
12. Explain the need and usefulness of diagrammatic representation of statistical data. What are the different types of diagrams you know ?
13. What is a pie diagram ? Draw a pie-Diagram with imaginary figures of children, adolescents, middle age and old age people in a particular place.
14. Describe the steps involved in the construction of a pie diagram.
15. What do you mean by a graph ?
16. How to choose scale for a graph ?

    Or

    What points should be taken on the base while selecting a scale for a graph ?

17. What do you mean by false base line ? When is it used and How ?

18. Define various types'of graphs.

19. Explain time series graph

    Or

    What is historigram ?

20. Define various types of frequency distribution graphs.

21. What is a histogram ? How to draw it ?

22. Define frequency polygon. How to draw it ?

23. Define Ogive. What are its types ? How to construct all these ?

24. (a) What is the difference between Diagrams and Graphs ?

    (b) Name the graph that are used to locate mode and, median respectively.

25. Define uses or merits or importance of Graphs.

26. What steps or guidelines should be followed to prepare a good graph

14.17 EXERCISES :

1. Draw line diagram to present the following data

(a)

Class	:M.Com.	M.Sc.	M.A	M.B.B.S	B.E.
No.of students :	220	180	340	80	120

(b)

Country :	U.S.A	U.K	Japan	India	Pakistan	France
Per capita income (in 000):	32	22	28	4	2	20

2. Present the following data by a bar diagram.

Country	**Production of sugar capita income (in 000)**
Cuba	32
Australia	30
India	20
Japan	5
Jawa	1
Egypt	1

3. Present the following data by multiple diagram.

(a)

Course	**No.of students**		
	2002-03	**2003-04**	**2004-05**
M.A	420	320	380
M.Sc.	200	240	360
M.Com.	140	300	480

(b)

Deptt.	**No.of students**		
	2002-03	**2003-04**	**2004-05**
Arts	600	550	500
Science	400	500	600
Commerce	200	250	300

4. Draw sub-divided bar diagrams for 3 (a) and (b)

5. Draw a suitable diagram to present the following information.

		Selling Price	**Qt. Sold**	**Wages**	**Material**	**Misc.**	**Totals**
Factory x	:	400	20	3200	2400	1600	7200
Factory y	:	640	30	6000	6000	9000	21000

(Hint : Preferably draw percentage bar diagram)

6. Present following data on a graph paper. (Single variables)

(a)

Year	:	2000	2001	2002	2003	2004	2005
No.of students	:	210	380	410	540	430	360

(b)

Year	:	2000	2001	2002	2003	2004	2005
No.of students	:	960	1080	1240	1160	1030	1260

(Hint : For (b), take false base)

7. Value of imports and exports is given; Draw graph.

Year	:	2000	2001	2002	2003	2004	2005
Imports	:	1080	1120	1240	1360	1040	1260
Exports (In Core Rs.)	:	640	120	980	1240	1340	1120

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson - 5

Averages - I

5.0 OBJECTIVE

After studying this lesson you should be able to understand the following.

1. Measure of Central Value
2. Objectives and features of Averages
3. Types of Averages
4. Merits and Demerits of Averages
5. Arithmetic Mean and its Calculations

STRUCTURE OF LESSON

5.1 Meaning of Average

5.2 Objectives of Average

5.3 Characteristics of Good Average

5.4 Types of Averages

5.5 Calculation of Arithmetic Mean

5.5.1 Arithmetic Mean- Individual Series

5.5.2 Arithmetic Mean - Discrete Series

5.5.3 Arithmetic Mean - Continuous Series

5.6 Combined Average

5.7 Weighted Arithmetic Mean

5.8 Merits and Demerits of Arithmetic Mean

5.9 Summary

5.10 Exercise

5.1 Meaning of Average

The main objective of statistical analysis is to arrive at one single value which represents the whole series. This value is called central value or an average. The value of average has a tendency towards centralisation. That means it lies in the middle of the data. It is the reason that averages are sometimes called measures of central tendency.

The word average is very commonly used in day-to-day conversation. For example, we often talk of average boy in a long, average height etc. It is defined by different statisticians.

a. According to Ya-Lun - Chou,

"An average is a typical value in the sense that it is sometimes employed to represent all the individual values in a series or of a variable".

b. According to 'Croxton and Cowden'.

"An average value is a single value within the range of the data that is used to reprsent all of the value in the series. Since an average is somewhere within the range of the data, it is something called a measure of central value".

5.2 Objectives of Average

There are two main objectives of the study of averages:

1. To get single value that describes the characteristic of the entire group.

2, Measures of Central value condenses the mass of data in one single value, enable us to get remember the dat easily.

3. Measures of Central values, by reducing the mass of data to one single figure, enable comparisons to be made. For example, we can compare the percentage results of the students of different colleges in a certain examination.

4. Averages are useful in decision making.

5.3 Requisites of a Good Average

A typical average should possess the following essentials or ideals to be a good average.

1. **It should be easy to understand:** Since statistical methods are designed to simplify complexity, it is desirable that an average be such that can be readily understood; otherwise, its use is bound to be very limited.
2. **It should be simple to compute** : An average should be simple to compute so that it can be used widely.
3. **It should be based on all the items**: The average should depend upon each and every item of the series so that if any of the items is dropped the average itself is altered.
4. **It should no be unduly affected by extreme observations** : If one or two very small or very large items unduly affect the averages i.e. either increase its value or reduce its value, the average cannot be reall typical of the entire series.
5. **It should be rigidly defined:** The average should be properly defined so that it has one and only one interpretation. It should preferably be defined by an algebraic formula so that if different people compute the average from the same figures they all get the same answer.
6. **It should be capable of further algebraic treatment :** We should prefer to have an average that could be used for further statistical comutations so that its utility is enhanced.

5.4 Types of Averages

The following are the important types of averages.

1. Arithmetic Mean

2. Median

3. Mode

4. Geometic Mean

5. Harmonic Mean

5.5 Arithmetic Mean

The most popular and widely used measure of representing the entire data by one value is what most laymen call an 'average' and what the statisticians call the arithmetic mean. Its value is obtained by adding together all the items and be dividing this total by the number of items. Arithmetic mean is two types.

Arithmetic Mean

Simple Arithmetic Mean | Weighted Arithmetic Mean

5.5.1 Arithmetic Mean - Individual Series

The process of computing mean in case of individual observations is very simple. Add together the various values of the variable and divide the total by the number of items.

Direct Method :

Arithmetic Mean = $\frac{\Sigma x}{N}$

Σx = Total of Values

N = Number of items.

Illustration 1 : Calculate Arihmetic Mean from the following data.

S.No.	:	1	2	3	4	5	6	7	8	9	10
Daily Wages	:	75	43	57	21	49	39	80	12	95	59

Solution :

Sno.	Daily Wages Rs.
1	75
2	43
3	57
4	21
5	49
6	39
7	80
8	12
9	95
10	59
10	**530**

Arithmetic Mean = $\frac{\Sigma x}{N}$

Σx = 530

N = 10

Arithmetic Mean = $\frac{530}{10}$ = 53

Illustration 2

Calculate Arithmetic Mean from the following Values.

Values 43 48 68 57 31 60 37 48 78

Values
43
48
65
57
31
60
37
48
78
526

Arithmetic Mean = $\frac{\Sigma x}{N}$

Σx = 526

N = 10

Arithmetic Mean = $\frac{526}{10}$ = 52.6

Arithmetic Mean = 52.6

Shortcut Method : The arithemetic mean can be calculated by using what is known as an arbitrary origin, when deviations are taken from the arbitrary origin, the formula for calculating arithmetic mean is -

$$\text{Arithmetic Mean} = X + \frac{\Sigma fdx}{N}$$

X = Assumed Mean

N = Number of items

Σfdx = Summation of multiples of deviations with their corresponding frequencies.

Illustration 3

Calculation Arithmetic Mean from the following information.

Values : 27 24 29 25 26 23 34 12 19 30 32

Solution

Values X	27	24	29	25	26	23	34	12	19	30	32	
dx	+2 (27-25)	-1 (24-25)	+4 (29-25)	0 (25-25)	+1 (26-25)	-2 (23-25)	+9 (34-25)	-13 (12-25)	-6 (19-25)	+5 (30-25)	+7 (32-25)	+28-22 = +6

$$\text{Arithmetic Mean} = X + \frac{\Sigma fdx}{N}$$

X = 25

N = 11

Σfdx = +6

$$\text{Arithmetic Mean} = 25 + \frac{6}{11}$$

Arithmetic Mean = 25 + .54 = 25.54

Arithmetic Mean = 25.54

Illustration 4 : From the following Marks calculate Arithmetic Mean.

Marks : 43 48 65 57 31 60 37 48 78 59

Solution

Marks	43	48	65	57	31	60	37	48	78	59	
dx	-14	-9	+8	0	+26	+3	-20	-9	+21	+2	**-78+34 = - 44**

$$\text{Arithmetic Mean} = X + \frac{\Sigma fdx}{N}$$

X = 57

N = 10

Σfdx = -44

$$\text{Arithmetic Mean} = 57 + \frac{-44}{10}$$

Arithmetic Mean = 57 - 4.4 = 52.6

Arithmetic Mean = 52.6

Illustration 5 : From the following data calculate Arithmetic Mean.

Family :	A	B	C	D	E	F	G	H	I	J
Salary Rs.:	85	70	10	75	500	8	42	250	40	36

Solution :

Family	A	B	C	D	E	F	G	H	I	J	
Salary (x) (Rs.)	85	70	10	75	500	8	42	250	40	36	
dx	+10	-5	-65	0	+425	-67	-33	+175	-35	-39	**-244 +610 =+366**

$$\text{Arithmetic Mean} = X + \frac{\Sigma fdx}{N}$$

X = 75

N = 10

Σfdx = +366

$$\text{Arithmetic Mean} = 75 + \frac{366}{10}$$

Arithmetic Mean = 75 + 36.6

Arithmetic Mean = 116.6

5.5.2 Arithmetic Mean - Discrete Series

In Discrete Series Arithmetic Mean may be computed by applying either Direct method or Short-cut method.

Direct Method :

$$\text{Arithmetic Mean} = \frac{\Sigma xf}{N}$$

Σxf = Multiply the frequency of each row with the variable and obtain the total of xf
Divide the total obtained by the number of observations. (i.e. total frequency)

Illustration 6

From the following data. Calculate Arithmetic Mean of 40 workers.

Wages (Rs.) :	3	5	8	10	12	15
No. of Workers:	4	10	12	8	4	2

Solution

Wages (X) (Rs.)	3	5	8	10	12	15	
No. of Workers (f)	4	10	12	8	4	2	**40**
xf	12	50	96	80	48	30	**316**

Arithmetic Mean = $\frac{\Sigma xf}{N}$

Σxf = 316

N = 40

Arithmetic Mean = $\frac{316}{40}$

Arithmetic Mean = 7.9

Illustration 7 : From the following data calculate Arithmetic mean of 100 employees.

Salary (Rs.) :	40	60	80	100	120	140	160	180	200
No.of Employees :	5	7	10	15	20	25	9	6	3

Solution

Salary (X) (Rs.)	40	60	80	100	120	140	160	180	200	
No. of Employees	5	7	10	15	20	25	9	6	3	**100**
xf	200	420	800	1500	2400	3500	1440	1080	600	**11,940**

Arithmetic Mean = $\frac{\Sigma xf}{N}$

Σxf = 11,940

N = 100

Arithmetic Mean = $\frac{11940}{100}$

Arithmetic Mean = 119.4

Short-cut Method

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = assumed mean

N = total of the frequency

Σfdx = Sum of multiples of deviations with their frequency

Illustration 8 :

Calculate Arithmetic Mean from the following data.

Values	1	2	3	4	5	6	7	8	9
Frequency	7	11	16	17	26	31	11	1	1

Solution

Values X	Frequency (f)	dx	f X dx fdx
1	7	-4	-28
2	11	-3	-33
3	16	-2	-32
4	17	-1	-17
5	26	0	0
6	31	+1	+31
7	11	+2	+22
8	1	+3	+3
9	1	+4	+4
			-110+60 = -50

$$\text{Arithmetic Mean} = X + \frac{\Sigma fdx}{N}$$

X = 5

N = 121

Σfdx = -50

$AM(\overline{X}) = 5 + \frac{-50}{121}$ = 5 + (-0.413)

AM = 4.587

Illustration 9 : Following is the data of 735 families. Calculate average number of children per families.

No. of Children :	0	1	2	3	4	5	6	7	8	9	10
No. of families :	96	108	154	126	95	62	45	20	11	6	5

Solution

No. of Children (X)	No. of families (f)	dx	fdx
0	96	-6	-576
1	108	-5	-540
2	154	-4	-616
3	126	-3	-1378
4	95	-2	-190
5	62	-1	-62
6	45	0	0
7	20	+1	+20
8	11	+2	+22
9	6	+3	+18
10	5	+4	+20
11	5	+5	+25
12	1	+6	+6
13	1	+7	+7
	735		**-2362 + 118= -2244**

$$\text{Arithmetic Mean} = X + \frac{\Sigma fdx}{N}$$

X = 6

N = 735

Σfdx = -2244

$AM(\overline{X}) = 6 + \frac{-2244}{735}$ = 6 + (-3.05)

AM = 2.95

Illustration 10

From the following data calculate Average Mark.

Marks	:	4	5	6	7	8	9
No. of Students	:	8	10	9	6	4	3

Solution

Marks (X)	No. of students (f)	dx	fdx
4	8	-2	-16
5	10	-1	-10
6	9	0	0
7	6	+1	+6
8	4	+2	+8
9	3	+3	+9
	40		**+23-26 = - 3**

$$\text{Arithmetic Mean} = X + \frac{\Sigma fdx}{N}$$

X = 6

N = 40

Σfdx = -3

$AM\,(\overline{X}) = 6 + \frac{-3}{40}$ = 6 + (-0.075)

AM = 5.925

Step Deviation Method

In the step deviation method the only additional point is that in order to simplify calculations we take a common factor from the data and multiply the result by the common factor.

$$\overline{X} = X + \frac{\Sigma fDx}{N} xC$$

X = assumed mean

N = total frequency

C = Common factor

ΣfDx = Sum of multiplies of Dx with frequency

Illustration 11

From the following information calculate Arithmetic Mean by using Step deviaion method.

Wages	40	60	80	100	120	140	160	180	200
No. of Workers	5	7	10	15	20	25	9	6	3

Solution

Wages (X)	No. of Workers (f)	dx	DX (20)	fDX (Dx X f)
40	5	-60	-3	-15
60	7	-40	-2	-14
80	10	-20	-1	-10
100	15	0	0	0
120	20	+20	+1	+20
140	25	+40	+2	+50
160	9	+60	+3	+27
180	6	+80	+4	+24
200	3	+100	+5	+15
	100			**-39+136 +97**

$$\overline{X} = X + \frac{\Sigma fDx}{N} xC$$

X = 100

N = 100

C = 20

ΣfDx = +97

$$\overline{X} = 100 + \frac{97}{100} x20 = 100 + 0.97 \text{ x } 20$$

= 100 + 19.4 = 119.4

$\overline{X}$ = 119.4

Illustration 12 : From the followign data calculate Arithmetic Mean.

Marks	5	15	25	35	45	55	65	75	
No. of Sudents		2	18	30	45	26	20	6	3

Solution

Marks (X)	No. of Students (f)	dx	DX (10)	fDX (Dx X f)
5	2	-30	-3	-6
15	18	-20	-2	-36
25	30	-10	-1	-30
35	45	0	0	0
45	26	+10	+1	26
55	20	+20	+2	40
65	6	+30	+3	18
75	3	+40	+4	12
	150			**-72+96 =+24**

$$\overline{X} = X + \frac{\Sigma fDx}{N} xC$$

X = 35

N = 150

C = 10

ΣfDx = +24

$\overline{X} = 35 + \frac{24}{150} x10$ = 100 + 0.97 x 20

= 35 + 0.16 x 10 = 35 + 1.6 = 36.6

$\overline{X}$ = 36.6

5.5.3 Arithmeic Mean - Continuous Series

In continuous series, arithmetic mean may be computed by applying any of the following methods.

1. Direct Method

2. Short-cut Method

3. Step Deviation Method

Direct Method

$$\text{Arithmetic Mean} = \frac{\Sigma fx}{N}$$

Σfx = Sum of multiples of variables with frequencies

N = Number of variables.

Illustration 13 : From the following data calculate Average Profit.

Profit per shop Rs. in Lakhs	No. of shops
0-10	12
10-20	18
20-30	27
30-40	20
40-50	17
50-60	6

Solution

Profit per shop Rs. in Lakhs(X)	No. of shops(f)	Mid Point of X	xf
0-10	12	5	60
10-20	18	15	270
20-30	27	25	675
30-40	20	35	700
40-50	17	45	765
50-60	6	55	330
	100		**2800**

$$\text{Arithmetic Mean} = \frac{\Sigma fx}{N}$$

Σfx = 2800

N = 100

$$\text{Arithmetic Mean} = \frac{2800}{100}$$ = 28 Lakhs

$\overline{X}$ = 28

Illustration 14

From the following wages of 40 workers calculate average wage.

Wage Rs.	No. of Workers
130-140	3
140-150	15
150-160	10
160-170	8
170-180	3
180-190	1

Solution

Wage Rs.(X)	No. of Workers(f)	Mid Point of X	xf
130-140	3	135	405
140-150	15	145	2175
150-160	10	155	1550
160-170	8	165	1320
170-180	3	175	525
180-190	1	185	185
	40		**6160**

$$\text{Arithmetic Mean} = \frac{\Sigma fx}{N}$$

Σfx = 6160

N = 40

$$\text{Arithmetic Mean} = \frac{6160}{40} = 154$$

$\overline{X} = 154$

Short Cut Method

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = assumed mean

Σfdx = sum of multiples of dx with frequencies

N = total of the frequency

Illustration 15 : Calculate Arithmetic Mean from the following data.

Values	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequencies	7	11	17	21	14	6	4

Values	Frequency	Mid Point of X	dx	fdx
0-10	7	5	-30	-210
10-20	11	15	-20	-220
20-30	17	25	-10	-170
30-40	21	35	0	0
40-50	14	45	+10	+140
50-60	6	55	+20	+120
60-70	4	65	+30	+120
	80			**-600+380** **-220**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 35

Σfdx = -220

N = 80

$$\overline{X} = 35 + \frac{-220}{80} = 35 + -2.75$$

Arithmeic Mean = 32.25

Illustration 16 : From the following data calculae Average Wage

Wages Rs.	130-140	140-150	150-160	160-170	170-180	180-190
No. of Workers	3	15	10	8	3	1

Solution :

Wages Rs.	No. of Wokers	Mid Point of X	dx	fdx
130-140	3	135	-20	-60
140-150	15	145	-10	-150
150-160	10	155	0	0
160-170	8	165	+10	+80
170-180	3	175	+20	+60
180-190	1	185	+30	+30
	40			**-210+170 =-40**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 155

Σfdx = -40

N = 40

$\overline{X} = 155 + \frac{-40}{40}$ = 155 + (-1) = 154

Arithmeic Mean = 154

Illustration 17 : From the following data calculae Arithmetic Mean

Class	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	12	18	27	20	17	6

Solution

Class	Frequency	Mid Point of X	dx	fdx
0-10	12	5	-30	-360
10-20	18	15	-20	-360
20-30	27	25	-100	270
30-40	20	35	0	0
40-50	17	45	+10	+170
50-60	6	55	+20	+120
	100			**-990+290 =-700**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 35

Σfdx = -700

N = 100

$$\overline{X} = 35 + \frac{-700}{100}$$ = 35 + (-7) = 28

Arithmeic Mean = 28

Step Deviation Method

$$\overline{X} = X + \frac{\Sigma fDx}{N} xC$$

X = assumed Mean

C = common factor

ΣfDx = sum of multiplies of Dx with frequencies

N = total frequency

Illustration 18 : Calculate Arithmetic Mean.

Class	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85
Frequency	7	8	12	26	32	42	42	15	17	9

Class	Frequency	Mid Point of X	dx	Dx / 5	fDx
35-40	7	37.5	-20	-4	-28
40-45	8	42.5	-15	-3	-24
45-50	12	47.5	-10	-2	-24
50-55	26	52.5	-5	-1	-26
55-60	32	57.5	0	0	0
60-65	42	62.5	+5	+1	+42
65-70	42	67.5	+10	+2	+84
70-75	15	72.5	+15	+3	+45
75-80	17	77.5	+20	+4	+68
80-85	9	82.5	+25	+5	+45
	210				**-102+284 =+182**

$$\overline{X} = X + \frac{\Sigma fDx}{N} xC$$

X = 57.5, C = 5

ΣfDx = +182

N = 210

$$\overline{X} = 57.5 + \frac{182}{210} x\,5$$

$$\overline{X} = 57.5 + \frac{910}{210} = 57.5 + 4.33$$

$$\overline{X} = 61.83$$

When Mid points are given :

If mid points are given take Midpoints directly and calculate arithmetic mean.

Illustration 19

Mid Points	1	2	3	4	5	6	7	8	9
Frequency	2	60	101	152	205	155	79	40	1

Solution

Mid Points	Frequency	dx	fDx
1	2	-4	-8
2	60	-3	-180
3	101	-2	-202
4	152	-1	-152
5	205	0	0
6	155	+1	+155
7	79	+2	+158
8	40	+3	+120
9	1	+4	+4
	795		**-105**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 5

Σfdx = -105

N = 795

$$\overline{X} = 5 + \frac{-105}{795}$$

$\overline{X}$ = 5 +(-0 .132) = 4.868

$\overline{X}$ = 4.868

Inclusive Method

When the data is given in inclusive form, then it is not necessary to adjust the classes for calculating arithmetic mean. it is becuase the mid value, remains the same whether the adjustment is made or not.

Illustration 20 : Calculate Arithmetic Mean from the following.

Marks	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65
No. of Students	4	8	27	48	57	81	86	77	49	36	20	5	2

Solution

Class	Frequency	Mid Point of X	dx	Dx 5	fDx
35-40	7	37.5	-20	-4	-28
40-45	8	42.5	-15	-3	-24
45-50	12	47.5	-10	-2	-24
50-55	26	<u>52.5</u>	-5	-1	-26
55-60	32	57.5	0	0	0
60-65	42	62.5	+5	+1	+42
65-70	42	67.5	+10	+2	+84
70-75	15	72.5	+15	+3	+45
75-80	17	77.5	+20	+4	+68
80-85	9	82.5	+25	+5	+45
	210				**-102+284 =+182**

Marks	No. of Students	Mid Points	dx	Dx (5)	fdx
1-5	4	3	-30	-6	-24
6-10	8	8	-25	-5	-40
11-15	27	13	-20	-4	-108
16-20	48	18	-15	-3	-144
21-25	57	23	-10	-2	-104
26-30	81	28	-5	-1	-81
31-35	86	33	0	0	0
36-40	77	38	+5	+1	+77
41-45	49	43	+10	+2	+98
46-50	36	48	+15	+3	+108
51-55	20	53	+20	+4	+80
56-60	5	58	+25	+5	+25
61-65	2	63	+30	+6	+12
	500				**-111**

$$\overline{X} = X + \frac{\Sigma fdx}{N} xC$$

X = 33, C = 5

Σfdx = -111

N = 500

$$\overline{X} = 33 + \frac{-111}{500} x5$$

$\overline{X}$ = 33 + (-1.11)

$\overline{X}$ = - 31.89

Unequal Classes

If the given classes are not equal, no need to change the class to calculate arithmetic mean.

Illustration 21

From the following data calculate Arithmetic Mean.

Revenue Rs. : Below 50 50-70 70-100 100-110 110-120 120-above

No.of Persons: 8 12 20 30 7 3

Solution

Revenue	No. of Persons	Mid Points	dx	fdx
30-50	8	40	-45	-360
50-70	12	60	-25	-300
70-100	20	85	0	0
100-110	30	105	+20	+600
110-120	7	115	+30	+210
120-130	3	125	+40	+120
	80			**-660+930 =+270**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 85, Σfdx = 270, N = 80

$$\overline{X} = 85 + \frac{270}{80}$$

$\overline{X}$ = 85 + 3.38

$\overline{X}$ = 88.38

Open-End Classes : Open-end classes are those in which lower limit of the first class and the upper limit of the last class are not known. In such case we cannot find out the Arithmetic Mean unless we make an assumption about the unknown limits. the assumption would naurally depend upon the class interval following the first class and preceding the last class.

Illustration 22

Class	below 50	50-100	100-150	150-200	200-250	above 250
Frequency	57	256	132	25	10	12

Solution

Class	Frequency	Mid Points	dx	fdx
Below 50	57	25	-100	-5700
50-100	256	75	-50	-12800
100-150	132	125	0	0
150-200	25	175	+50	1250
200-250	10	225	+100	1000
250-300	12	275	+150	1800
	492			**-18500+4050 =-14450**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 125, Σfdx = -14450, N = 492

$$\overline{X} = 125 + \frac{-14450}{492}$$

$\overline{X}$ = 125 + (-29.36) = 95.64

$\overline{X}$ = 95.64

Illustration 23

From the following data calculate Arithmetic Mean.

Income Rs.	35-40	40-45	45-50	50-55	55-60	60-75	75-90	90-100	100-120
No. of Persons	6	7	13	15	16	14	11	9	9

Solution

First Class 40

Second Class 40 - 45 Difference - 5

So take the first class difference also as 5.

Now the first class lower limit is = Upper limit - 5

= 40 - 5 = 35

Income Rs.	No. of Persons	Mid Points	dx	fdx
35-40	6	37.5	-17.5	-105.0
40-45	7	42.5	-12.5	-87.5
45-50	13	47.5	-7.5	-97.5
50-60	15	55.0	0	0
60-75	16	67.5	+12.5	200
75-90	14	82.5	+27.5	385
90-100	11	95.0	+40.5	440
100-120	9	110.0	+55.0	495
120-140	9	130.0	+75.0	675
	100			**+2195-290 =+1905**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 55, Σfdx = 1905, N = 100

$$\overline{X} = 55 + \frac{1905}{100}$$

$\overline{X}$ = 55 + 19.05 = 74.05

$\overline{X}$ = 74.05

Arithmetic Mean with Cumulative Frequency Disribution

When the data is given in the form of 'more than' or 'less than', 'above' or 'below' for all items, in the series, it is called comulative frequency distribution. To calculate arithmetic mean, construct class by taking difference of two given limits. Get general frequency by substracting cumulative frequency.

Less than Cumulative Frequency :

Illustration 24

Following are the marks of 80 B.Com., Sudents in Statistics. Calculate their average mark.

Marks	No. of Students
Less than 10	7
Less than 20	18
Less than 30	35
Less than 40	56
Less than 50	70
Less than 60	76
Less than 70	80

Solution

Marks	No. of Students	General Frequency	Mid Points	dx	fdx
Less than 10	7	7	5	-30	-210
Less than 20	18	11(18-7)	15	-20	-220
Less than 30	35	17(35-18)	25	-10	-170
Less than 40	56	21(56-35)	<u>35</u>	0	0
Less than 50	70	14(70-56)	45	+10	+140
Less than 60	76	6(76-70)	55	+20	+120
Less than 70	80	4(80-76)	65	+30	+120
		80			**-600+380 =-220**

$$\overline{X} = X + \frac{\Sigma fdx}{N}$$

X = 35, Σfdx = -220, N = 80

$$\overline{X} = 35 + \frac{-200}{80}$$

$\overline{X}$ = 35 + (-2.5) = 32.5

$\overline{X}$ = 32.5

Illustration 25 : From the following data of 240 students marks calculate Arithmetic Mean.

Marks	No. of Students
Less than 10	25
Less than 20	40
Less than 30	60
Less than 40	75
Less than 50	95
Less than 60	125
Less than 70	190
Less than 80	240

Solution

Marks	No. of Students	General Frequency	Mid Points	dx	Dx (C-10)	fDx
0-10	25	25	5	-30	-3	-75
10-20	40	15(40-25)	15	-20	-2	-30
20-30	60	20(60-40)	25	-10	-1	-20
30-40	75	15(75-60)	35	0	0	0
40-50	95	20(95-75)	45	+10	+1	+20
50-60	125	30(125-95)	55	+20	+2	+60
60-70	190	65(190-125)	65	+30	+3	+195
70-80	240	50(240-190)	75	+40	+4	+200
		240				**-125+475 =+350**

$$\overline{X} = X + \frac{\Sigma fDx}{N} x C$$

X = 35, ΣfDx = 350, N = 240, C = 10

$$\overline{X} = 35 + \frac{350}{240} x 10$$

$\overline{X}$ = 35 + 1.45 x 10 = 35 + 14.5 = 49.58;

$\overline{X}$ = 49.58

Illustration 26

Calculate Arithmeic Mean from the following data.

Marks	No. of Students
Less than 10	15
Less than 20	35
Less than 30	60
Less than 40	84
Less than 50	96
Less than 60	127
Less than 70	198
Less than 80	250

Solution

Marks	No. of Students	General Frequency	Mid Points	dx	Dx (C-10)	fDx
0-10	15	15	5	-30	-3	-45
10-20	35	20(35-15)	15	-20	-2	-40
20-30	60	25(60-35)	25	-10	-1	-25
30-40	84	24(84-60)	35	0	0	0
40-50	96	12(96-84)	45	+10	+1	+12
50-60	127	31(127-96)	55	+20	+2	+62
60-70	198	71(198-127)	65	+30	+3	+213
70-80	250	52(250-198)	75	+40	+4	+208
		250				**-110+495 = +385**

$$\overline{X} = X + \frac{\Sigma fDx}{N} x C$$

X = 35, ΣfDx = 385, N = 250, C = 10

$$\overline{X} = 35 + \frac{385}{250} x10$$

$\overline{X}$ = 35 + 1.54 x 10 = 35 + 15.4 = 50.4

$\overline{X}$ = 50.4

More than Cumulative Frequency

Illustration 27 :

Calculate average weight from the following data.

Weight (Pounds)	No. of Persons
More than 100	400
More than 110	300
More than 120	170
More than 130	100
More than 140	80
More than 150	50

Solution

Marks	No. of Students	General Frequency	Mid Points	dx	Dx (C-10)	fDx
100-110	400	100(400-300)	105	-20	-2	-200
110-120	300	130(300-170)	115	-10	-1	-130
120-130	170	70(170-100)	125	0	0	0
130-140	100	20(100-80)	135	+10	+1	+20
140-150	80	30(80-50)	145	+20	+2	+60
150-160	50	50	155	+30	+3	+150
		400				**-330+230 =-100**

$$\overline{X} = X + \frac{\Sigma fDx}{N} x C$$

X = 125, ΣfDx = -100, N = 400, C = 10

$$\overline{X} = 125 + \frac{-100}{400} x10$$

$\overline{X}$ = 125 + (-0.25) x 10 = 125 - 2.5 = 122.5

$\overline{X}$ = 122.5

Illustration 28 :

Calculate Arithmetic Mean from the following information

Height (cm)	Morethan 75	Morethan 85	Morethan 95	Morethan 105	More than 115	More than 125	More than 135	More than 145
No. of Persons	214	212	189	140	77	32	8	7

Solution

Heights	No. of Students	General Frequency	Mid Points	dx	Dx (C-10)	fDx
75-85	214	2(214-212)	80	-30	-3	-6
85-95	212	23(212-189)	90	-20	-2	-46
95-105	189	49(189-140)	100	-10	-1	-49
105-115	140	63(140-77)	110	0	0	0
115-125	77	45(77-32)	120	+10	+1	+45
125-135	32	24(32-8)	130	+20	+2	+48
135-145	8	1(8-7)	140	+30	+3	+3
145-155	7	7	150	+40	+4	+28
		214				**-162**

$$\overline{X} = X + \frac{\Sigma fDx}{N} x C$$

X = 110, ΣfDx = -162, N = 214, C = 10

$$\overline{X} = 110 + \frac{-162}{214} x10$$

$\overline{X}$ = 110 + (-0.75) x 10 = 110 - 7.5 = 102.5

$\overline{X}$ = 102.5

Illustration 29 :

Calculate Arithmetic Mean.

Class	Morethan 0	Morethan 10	Morethan 20	Morethan 30	Morethan 40
Frequency	40	36	28	15	5

Solution

Class	Frequency	General Frequency	Mid Points	dx	Dx (C-10)	fDx
0-10	40	4(40-36)	5	-20	-2	-8
10-20	36	8(36-28)	15	-10	-1	-8
20-30	28	13(28-15)	25	0	0	0
30-40	15	10(15-5)	35	+10	+1	+10
40-50	5	5	45	+20	+2	+10
		40				**-16+20 =+4**

$$\overline{X} = X + \frac{\Sigma fDx}{N} xC$$

X = 25, ΣfDx = +4, N = 40, C = 10

$$\overline{X} = 25 + \frac{4}{40} x10$$

$\overline{X}$ = 25 + 0.1 x 10 = 25 + 1 = 26

$\overline{X}$ = 26

Correcting Incorrent Values

It sometimes happens that due to an oversight or mistake in copying certain wrong items are taken while calculating mean. The problem is how to find out the correct mean. The process is from incorrect Σx deduct wrong items and add correct items and then divide the correct Σx by the number of observations. The result is correct mean.

Illustration 30:

It mean of 200 items was 50. Later on it was discovered that two items were misread as 92 and 8 instead of 192 and 88. Find the correct mean.

$$\overline{X} = \frac{\Sigma X}{N}$$

ΣX = N x $\overline{X}$

N = 200

$\overline{X}$ = 50

ΣX = 200 x 50

Incorrect ΣX = 1000

Correct ΣX = Incorrect ΣX - Wrong items + Correct Items

Correct ΣX = 10000 - (92 + 8) + (192 + 88)

= 9900 + 280 = 10180

Correct ΣX = 10180

Correct Mean = $\frac{10180}{200}$ = 50.9

$\overline{X}$ = 50.9

Illustration 31 :

Following are the results of 50 students who appeared for an examination.

Marks	4	5	6	7	8	9
No. of Students Passed	8	10	9	6	4	3

Average of 50 students marks are 5.16. Find out average marks of students who failed.

Solution

Marks(X)	No. of Students Passed (f)	xf
4	8	32
5	10	50
6	9	54
7	6	42
8	4	32
9	3	27
		237

Total marks of 50 students = 5.16 x 50 = 258.00

Total marks of 40 students = 237

Total marks of 10 students who failed = 258 - 237 = 21

Arithmetic Mean = $\frac{21}{10}$ = 2.1

Missing Figures

Illustration 32 :

From the following information find out missed value, where the average salary is Rs. 115.86.

Salary (X)	110	112	113	117	x	125	128	130
No. of Persons(f)	25	17	13	15	14	8	6	2

Solution

Salary (Rs.)	No. of Persons	xf
110	25	2750
112	17	1904
113	13	1469
117	15	1755
x	14	14x
125	8	1000
128	6	768
130	2	260
	100	**9906 + 14x**

$\overline{X} = \frac{\Sigma xf}{N}$

N = 100

Σxf = 9906 + 14x

Arithmetic Mean = 115.86

115.86 = $\frac{9906 + 14x}{100}$

14x = 11586 - 9906

x = $\frac{1680}{14}$ = 120

x = 120

Illustration 33 :

From the following data find out missed frequency. if the average income is Rs. 19.92.

Revenue (Rs.)	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	36-40
No. of Persons	11	13	16	14	x	9	17	6	4

Solution

Revenue (Rs.)	No. of Persons	Mid Points	xf
4-8	11	6	66
8-12	13	10	130
12-16	16	14	224
16-20	14	18	252
20-24	x	22	22x
24-28	9	26	234
28-32	17	30	510
32-36	6	34	204
36-40	4	38	152
	90 + x		**1772 + 22x**

Arithmetic Mean $\overline{X}$ = 19.92

N = 90 + x

Σxf = 1772 + 22x

$$\overline{X} = \frac{\Sigma xf}{N}$$

$$19.92 = \frac{1772 + 22x}{90 + x}$$

1772 + 22x = (19 + x) 19.92

1772 + 22x = 1792.80 + 19.92x

22x - 19.92x = 1792.80 - 1772

2.08x = 2080

$$x = \frac{2080}{2.08}$$

x = 10

5.6 Combined Average

If Arithmetic Mean and the number of items of two or more than two related groups are given, the combined average of these groups can be calculated by applying the following formula.

$$\bar{X}_{1\,2\,3\,....\,n} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2 + N_3\bar{X}_3 .... + N_n\bar{X}_n}{N_1 + N_2 + N_3 .... + N_n}$$

$\bar{X}$ 1 2 3 ... n = Combined mean of the groups

$\bar{X}_1$ = Arithmetic Mean of first group

$\bar{X}_2$ = Arithmetic Mean of Second group

$\bar{X}_3$ = Arithmetic Mean of third group

$\bar{X}_n$ = Arithmetic Mean of nth group

N_1 = Number of items in the first group

N_2 = Number of items in the second group

N_3 = Number of items in the third group

N_n = Number of items in the nth group

Illustration 34:

The Mean height of 25 male workers in a factory is 61 cm, and the mean height of 35 female workers in the same factory is 58 cm. Find the combined mean height of 60 workers in the factory.

Solution :

$$\bar{X}_{12} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$$

Where

$\bar{X}_1$ = 61

$\bar{X}_2$ = 58

N_1 = 25

N_2 = 35

$$\bar{X}_{12} = \frac{(25x61) + (35x58)}{25 + 35}$$

$$\bar{X}_{12} = \frac{1525 + 2030}{60}$$

$$\overline{X}_{12} = \frac{3555}{60}$$

$$\overline{X}_{12} = 59.25$$

Illustration 35 :

The mean of wages in factory A of 100 workers is Rs. 720 per week. The mean wages of 30 female workers in the factory was Rs.650 per week. Find out average wage of male workers in the factory.

Solution

$N = 100, N_1 = 30, \overline{X}_1 = 650 \ \overline{X}_2 = ?$

$N_1 + N_2 = N_{30} + N_2 = 100, N_2 = 70$

$\overline{X}_{12} = 720$

$$720 = \frac{30x650 + 70\,X_2}{100}$$

$72000 = 19500 + 70\ X_2$

$70\ X_2 = 52500$

$$X_2 = \frac{52500}{70}$$

$X_2 = 750$

5.7 Weighted Arithmetic Mean

One of the limitations of the arithmetic mean is that it gives equal importance to all the items. but there are cases where the relative importance of the different items is not the same. When this is so, we compute wiighted arithmetic mean. The term weight stands for the relative importance of the different items. The formula for computing weighted arithmetic mean is:

$$\overline{X}_w = \frac{\Sigma wx}{\Sigma w}$$

$\overline{X}_w$ = Weighted arithmetic mean

W = Weights

X = values

Illustration 36:

From the following data calculate weighted Arithmetic Mean.

Variables	80	75	67	86	35
Weights	2	3	4	5	6

Solution

S.No	X	W	XW
1	80	2	160
2	75	3	225
3	67	4	268
4	86	5	430
5	35	6	210
	343		**1293**

$$\overline{X}_w = \frac{\Sigma wx}{\Sigma w}$$

$$\overline{X}_w = \frac{1293}{20}$$

$$\overline{X}_w = 64.65$$

Illustration 37 :

From the following results of Three Universities calculate weighted Arithmetic Mean.

Examination	A No. of students		B No. of Students		C No. of Students	
	% passed	In hundreds	% Passed	in hundreds	% Passed	in hundreds
M.A.	70	5	75	4	75	6
M.Sc.	85	4	80	3	65	4
M.Com.	80	6	65	5	70	5
B.A.	75	7	85	6	80	7
B.Sc.	65	5	75	4	85	5
B.Com.	75	8	70	5	75	5

Solution

Examinati-on	University A No. of students			University B No. of Students			C No. of Students		
	% passed (X)	In hundreds (W)	XW	% passed (X)	in hundreds (W)	XW	% passed (X)	in hundreds (W)	XW
M.A.	70	5	350	75	4	700	75	6	450
M.Sc.	85	4	340	80	3	240	65	4	260
M.Com.	80	6	480	65	5	325	70	5	350
B.A.	75	7	525	85	6	510	80	7	560
B.Sc.	65	5	325	75	4	300	85	5	425
B.Com.	75	8	600	70	5	350	75	5	375
	450	**35**	**2620**	**450**	**27**	**2025**	**450**	**32**	**2420**

Weighted Arithmetic Mean $\overline{X}_w = \frac{\Sigma wx}{\Sigma w}$

University A = $\frac{2620}{35}$ = 74.86%

University B = $\frac{2025}{27}$ = 75.00%

University C = $\frac{2420}{32}$ = 75.64%

Illustration 38 : From the results of the College X and Y. State which of them is better and why.

Name of the Exam	College X Appeared	College X Passed	College Y Appeared	College Y Passed
M.A.	300	250	1000	800
M.Com.	500	450	1200	950
B.A.	2000	1500	1000	700
B.Com.	1200	750	800	500

Solution

Name of the Exam	College X Appeared	% of Pass	W_1X_1	College Y Appeared	% of Pass	W_1X_1
M.A.	300	$\frac{250}{300} \times 100 = 83.3$	24,990	1000	$\frac{800}{1000} \times 100 = 80.0$	80,000
M.Com.	500	$\frac{450}{500} \times 100 = 90.0$	45,000	1200	$\frac{950}{1200} \times 100 = 79.17$	95,004
B.A.	2000	$\frac{1500}{2000} \times 100 = 75.0$	150000	1000	$\frac{700}{1000} \times 100 = 70.0$	70,000
B.Com.	1200	$\frac{750}{1200} \times 100 = 62.5$	75000	800	$\frac{500}{800} \times 100 = 62.5$	50,000
	4000		2,94,990	4000		2,95,004

Weighted Arithmeic Mean for College X = $\frac{\Sigma w_1 x_1}{\Sigma w_1} = \frac{294990}{4000} = 73.747$

Weighted Arithmeic Mean for College Y = $\frac{\Sigma w_2 x_2}{\Sigma w_2} = \frac{295004}{4000} = 73.751$

Average pass percentage of College Y > College X.

So, College Y is better.

Merits and Limitations of Arithmetic Mean

Merits : Arithmetic Mean is most commonly used average in practice becuase of its advantages. Following are important Merits of Arithmetic Mean.

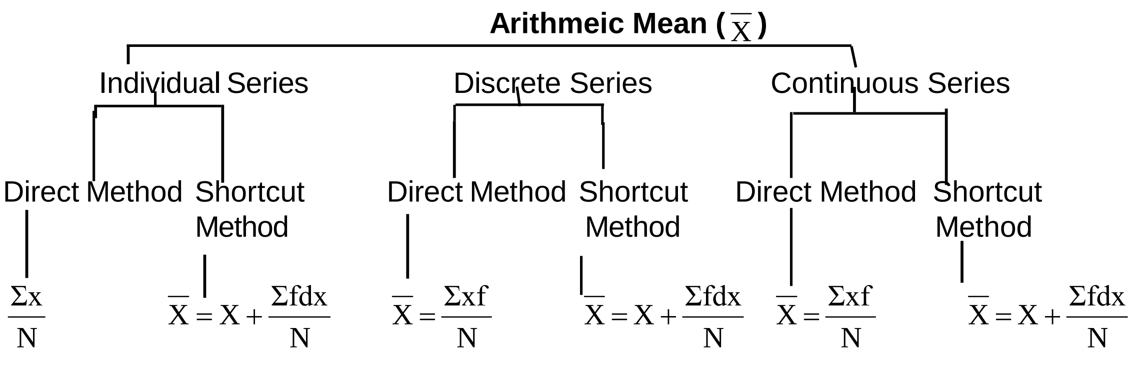
1. It is very simple to understand and calculate.
2. Arithmetic Mean is affected by the value of each and every item in the series.
3. Arithmetic Mean is defined by a rigid mathematical formula.
4. Arithmetic Mean is useful for algebraic tratment. it is better than median Mode, Geometric Mean and Harmonic Mean.
5. Arithmetic Mean is relatively stable. It does not fluctuate much when repeated samples are taken from one and the same population.
6. Arithmetic Mean is a calculated value and is not based on position in the series.

Demerits

1. The value of Arithmetic Mean depends on each and every item of the series. The value of average is affected by the extreme items, either very small or very large.
2. In open-end classes, the value of mean cannot be calculated without making assumption regarding the size of the class interval of the open - end classes.
3. In case the distribution is U shaped, then mean is not likely to serve a useful purpose. So it is not a good measure always.

5.9 SUMMARY

Thus, it is most widely used measure for representing the entire data. To a layman, it is average but for a statistician, it is called 'arithmetic mean'. It is calculated by adding values of all the items and dividing their total by the number of items. In case of discrete and continuous series, the values of the frequencies are taken into account. Following figure depicts the Calculation of arithmetic Mean.



Step Deviation Method $= \overline{X} = X + \frac{\Sigma fdx}{N} xC$

5.10 EXERCISE

1. What is an average? What are its objectives.
2. Explain requisites of a good average.
3. Define Arithmetic Mean and Explain its merits and demerits.
4. From the following data calculate Arithmetic Mean.

   Monthly Income Rs. : 200, 300, 330, 400, 500, 600, 400, 700, 740, 560, 440

   (Ans. : 470)
5. Find out Arithmetic mean.

Wages	3	5	8	10	12	15
No. of Workers	4	10	12	8	4	2

   (Ans. : 7.90)

6. Calculate Arithmetic mean.

Class	15-25	25-35	35-45	45-55	55-65	65-75
Frequencies	20	30	40	50	60	70

   (Ans. : 40.2)

7. Findout Arithmetic Mean.

Class	2-3	4-5	6-7	8-9	10-11	12-13
Frequency	20	43	50	30	18	10

   (Ans. : 6.66)

8. Calculate Arithmetic Mean.

Class	Frequency
Below 5	5
10	9
15	17
20	29
25	45
30	60
35	70
40	78
45	83
50	85

(Ans. 24.25)

9.Calculate Arithmetic Mean

Income	No. of Persons
More than 10	72
More than 20	67
More than 30	59
More than 40	50
More than 50	36
More than 60	21
More than 70	9
More than 80	3

(Ans. 49.03)

10. Calculate Arithmetic Mean.

Wages	No. of Workers
5	7
10	8
15	12
20	13
25	18
30	14
35	11
40	8
45	5
50	4

(Ans. : 2.25)

11. Calcualte the number of students against the class 30-40 of the following data where $\overline{X}$ = 28.

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of Students	12	18	27	2	17	6

(Ans. : 20)

12.Average weight of 150 students in a class is 60 kgs. Average weight of Boys of that class is 70 kgs, and girls average weight is 55 kgs. Find out number of boys and girls of that class. (Ans. : 50, 100)

13. Calculate Weighted Arithmetic Mean for the following data.

Salary per month Rs. :	1500	800	500	250	100
No. of Employees :	10	20	70	100	150

(Ans. : 302.86)

- Dr. K. Kanaka Durga

LESSON 6

AVERAGES - II MEDIAN

6.0 OBJECTIVE

After studying this lesson you should be able to understand -

1. What is Median.
2. Wht are it merits and limitations.
3. How to compute Median.

STRUCTURE

6.1 INTRODUCTION

The median is one of the measures of central value. One of the most important objects of statistical analysis is to get one single value that describes the characteristic of the entire mass of unwidely data such value is called th central value or an average. As distinct from the Arithmetic mean which is calculated from the value of every item in the series, the median is that is called a positional average. The term ‘position’ refers to the place of a value in a series. The place of the median in a series is such that an equal number of items lie on either side of it. For example, if the income of five persons is 2,800, 2820, 2880, 2885, 2890, then the median income would be Rs. 2,880. median is thus the central value of the distribution or the value that divides the distribution into two equal parts.

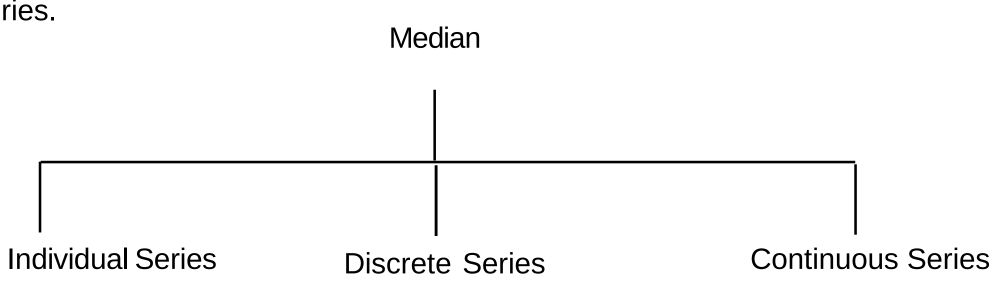
6.2 MEANING AND DEFINITION

The median by definition is the middle value of the distribution. Whenever the median is given as a measure, one half of the items in the distribution have a value.

Thus the median divides the distribution into two equal parts. if there are even number of items in a series there is no actual value exacly in the middle of the series and as such the median is indeterminate. In such a case the median is arbitrarily taken to be halfway betwen the two middle values.

6.3 CALCULATION OF MEDIAN

Median is claculated in three series such as Individual series, Discrete sereis and Continuous series.



6.3.1 Individual Series

Following is the procedure to calculate Median in Individual Series.

1. Arrange the data in ascending or descending order of magnitude

2. Apply the formula

M = size of $\frac{N+1}{2}$ th item.

N = No. of items

Illustration 1:

Obtain the value of median from the following data.

Values	391	384	407	672	522	777	753	2488	1490

Solution :

Values	384	391	407	522	<u>672</u>	753	777	1490	2488

Median = size of $\frac{N+1}{2}$ th item

N = 9 (Number of items)

Size of $\frac{9+1}{2}$ th item

= $\frac{10}{2}$ th item

5th item i.e. Median = 672

Illustration 2

From the wages of 11 workers, Calculate Median.

Values	1	2	3	4	5	6	7	8	9	10	11
Wages	60	55	45	70	75	80	50	90	95	100	85

Solution

Values	1	2	3	4	5	6	7	8	9	10	11
Wages	45	50	55	60	70	75	80	85	90	95	100

Median = size of $\frac{N+1}{2}$ th item

N = 11 (Number of items)

Size of $\frac{11+1}{2}$ th item

$\frac{12}{2}$ th item

6th item i.e. Median = 75

Illustration 3

Marks	30	27	26	35	37	40	25	45	47

Solution

Marks in Ascending Order	25	26	27	30	35	37	40	45	47

Median = size of $\frac{N+1}{2}$ th item

N = 9 (Number of items)

Size of $\frac{9+1}{2}$ th item

= $\frac{10}{2}$ th item

= 5th item i.e. Median = 35

If the number of items was odd and therefore, it is not possible to determine the middle value. When the number of items is even for example, if in the above case the numbe o items is 8 then median would be the value of $\frac{8+1}{2}$ = 4.5 th item for finding out the value of 4.5th item we shall take the average of 4th and 5th items i.e. Median = $\frac{\text{4th item} + \text{5th item}}{2}$

Illustration 4

Calculate median income from the following data.

Income Rs. : 891 884 991 907 1072 922 1277 1153 2488 1490

Solution

Income Rs.

884

891

907

Median = size of $\frac{N+1}{2}$ th item

N = 10 (Number of items)

Size of $\frac{10+1}{2}$ th item

222	$\frac{11}{2}$ th item
991	
1072	Median = $\frac{\text{5th item} + \text{6th item}}{2}$
1153	
1277	$\frac{991 + 1072}{2} = \frac{2063}{2}$ = 1031.5
1490	
2488	Hence Median Income is Rs.1031.5

Illustration 5

Calculate Median mark from the following data.

Marks : 40 45 31 75 81 57 63 92 35 21

Solution

Marks	Marks in Ascending Order
40	21
45	31
31	35
75	40
81	45
57	57
63	63
92	75
35	81
21	92

Median = size of $\frac{N+1}{2}$ th item

N = 10 (Number of items)

Size of $\frac{10+1}{2}$ th item

$\frac{11}{2}$ th item

Median = $\frac{\text{5th item} + \text{6th item}}{2}$

$\frac{45 + 57}{2} = \frac{102}{2}$ = 51

Median = 51

Illustration 6

From the following wages. Calculate Median.

Wages (Rs.) 60 55 45 70 75 80 50 90 95 100 95

Solution

Wages (Rs.)	Wages in Ascending Order (Rs.)
60	45
55	50
45	55
70	60
75	70
80	75
50	80
90	85
95	90
100	95
95	100
110	110

Median = size of $\frac{N+1}{2}$ th item

N = 12 (Number of items)

Size of $\frac{12+1}{2}$ th item

$\frac{13}{2}$ th item 6.5th item

Median = $\frac{\text{6th item} + \text{7th item}}{2}$

$\frac{75+80}{2} = \frac{155}{2}$ = 77.5

6.3.2 Discrete Series

Median = Size of $\frac{N+1}{2}$ th item

N = Total of the frequency.

Steps of calculate Median.

1. Arrange data in ascending order or descending order.
2. Find out Cumulative Frequency
3. Apply the formula.

Illustration 7

From the following data find out the Value of Median :

Income (Rs.)	No. of Persons
1600	24
1650	26
1580	16

1700	20
1750	6
1680	30

Solution

Income in Ascending Order (Rs.)	No. of Persons	Cumulative Frequency
1580	16	16
1600	24	40
1650	26	66
1680	30	96
1700	20	116
1750	6	122
	122	

Median = size of $\frac{N+1}{2}$ th item

N = 122

M = Size of $\frac{122+1}{2}$ th item

$\frac{123}{2}$ th item = 61.5 th item.

Size of 61.5th item = 1650. Hence Median 1650.

Illustration 8 :

From the following heights of 100 students calculate Median Height.

Height Cm. :	155	156	157	158	159	160	161	162	163	164
No. of Students	3	7	9	12	13	17	16	14	7	2

Solution :

Height cm.	No. of Students	Cumulative Frequency
155	3	3
156	7	10

157	9	19
158	12	31
159	13	44
160	17	61
161	16	77
162	14	91
163	7	98
164	2	100
	100	

Median = size of $\frac{N+1}{2}$ th item

N = 100

M = Size of $\frac{100+1}{2}$ th item

= $\frac{101}{2}$ th item = 50.5 th item.

Size of 50.5th item = 61 of Cumulative Frequenc.

Median = Corresponding value of 61 is 160

Median = 160

Illustration :

From the following Weights. Calculate Median Weight.

Weight (P)	70	100	180	150	80	120	200	250	170	90
No. of Persons		20	45	25	38	35	50	22	15	30

40

Solution :

Weight	No. of Persons	Culative Frequency
70	20	20
80	35	55
90	40	95
100	45	140
120	50	190

150	38	228
170	30	258
180	25	283
200	22	305
250	15	320
	320	

Median = size of $\frac{N+1}{2}$ th item

N = 320

Size of $\frac{320+1}{2}$ th item

$\frac{321}{2}$ th item = 160.5 th item.

Size of 160.5th item lies in 190.

Coresponding Value of 190 = 120.

Median = 120

6.3.3 Continuous Series

In continuous series calculation of Median follows the following steps.

1. Determine he particular class in which the value of median lies with the help of m = $\frac{N}{2}$.

2. Apply the Principle

$$M = l_1 + \frac{l_2 - l_1}{f_1} \text{ x } m - c$$

l_1 = Lower limit of the Median Class

l_2 = Upper limit of the Median Class

f_1 = Frequency of the Median Class

m = Value of $\frac{N}{2}$ nd item.

c = Cumulative Frequency of the class preceding the Median Class.

Illustration 10

From the following data calculate Median.

Values	0-20	20-40	40-60	60-80	80-100	100-120	120-140	140-160
Value Frequency	1	14	35	85	90	45	18	2

Solution

Values	Value Frequency	Cumulative Frequency
0-20	1	1
20-40	14	15
40-60	35	50
60-80	85	135
80-100	90	225
100-120	45	270
120-140	18	288
140-160	2	290
	290	

$$M = l_1 + \frac{l_2 - l_1}{f_1} \text{ x } m - c$$

$l_1 = 80$

$l_2 = 100$

$f_1 = 90$

m = 145

c = 135

$$M = 80 + \frac{100 - 80}{90} \text{ x } 145 - 135$$

$$M = 80 + \frac{20}{90} \text{ x } 10$$

$$M = 80 + \frac{206}{96}$$

M = 80 + 2.2 = 82.2.

M = 82.2

Illustration 11

Calculate Median from the following marks.

Marks	No. of Students
30-32	2
32-34	9
34-36	25
36-38	30
38-40	45
40-42	62
42-44	39
44-46	20
46-48	11
48-50	3

Solution

Marks	No. of Students	Cumulative Frequency
30-32	2	2
32-34	9	11
34-36	25	36
36-38	30	66
38-40	45	115
40-42	62	177
42-44	39	216
44-46	20	236
46-48	11	247
48-50	3	250
	250	

$m = \frac{N}{2}$

N = 250

$m = \frac{250}{2} = 125$

$M = l_1 + \frac{l_2 - l_1}{f_1} \times m - c$

$l_1 = 40$

$l_2 = 42$

$f_1 = 62$

m = 125

c = 115

$M = 40 + \frac{42 - 40}{62} \times 125 - 115$

$M = 40 + \frac{2}{62} \times 10$

$M = 40 + \frac{20}{62}$

M = 40 + 0.3 = 40.3

M = 40.3

Illustration 12

Calculate Median from the following data.

Class	Frequency	Cumulative Frequency
15-25	4	4
25-35	11	15
35-45	19	34
45-55	14	48
55-65	0	48
65-75	2	50
	50	

$m = \frac{N}{2}$

N = 50

$m = \frac{50}{2} = 25$

$M = l_1 + \frac{l_2 - l_1}{f_1} \times m - c$

$l_1 = 35$

$l_2 = 45$

$f_1 = 19$

$m = 25$

$c = 15$

$M = 35 + \frac{45 - 35}{19} \times 25 - 15$

$M = 35 + \frac{10}{19} \times 10$

$M = 35 + \frac{100}{19}$

$M = 35 + 5.2 = 40.2$

$M = 40.2$

Illustration 13

From the following data calculate Median Profit.

Profit(Rs.)	No. of Traders
1999.5-2999.5	20
2999.5-3999.5	45
3999.5-4999.5	70
4999.5-5999.5	50
5999.5-6999.5	28
6999.5-7999.5	22
7999.5-8999.5	15

Solution

Profit(Rs.)	No. of Traders	Cumulative Frequency
1999.5-2999.5	20	20
2999.5-3999.5	45	65
3999.5-4999.5	70	135
4999.5-5999.5	50	185
5999.5-6999.5	28	213
6999.5-7999.5	22	235
7999.5-8999.5	15	250
	250	

$m = \frac{N}{2}$

N = 250

$m = \frac{250}{2} = 125$

$M = l_1 + \frac{l_2 - l_1}{f_1} \times m - c$

$l_1 = 3999.5$

$l_2 = 4999.5$

$f_1 = 70$

m = 125

c = 65

$M = 3999.5 + \frac{4999.5 - 3999.5}{70} \times 125\text{-}65$

$M = 3999.5 + \frac{1000}{70} \times 60$

$$M = 3999.5 + \frac{6000}{7}$$

M = 3999.5 + 857.1 = 4856.6

M = 4856.6

6.3.3.1 Inclusive (Class) Series:

When the classes are in inclusive series, change the classes into exclusive form. To change into exclusive form take the diference between upper limit of first class and the lower limit of next class. Divide the difference by two. Subtract the difference from lower limits and add to the upper limits.

Example : 11-20

21-30

31-40

Difference between 21-20 is 1. It is divided by two i.e. $\frac{1}{2}$ = 0.5.

20 + 0.5 = 20.5 Upper limit

21 - 0.5 = 20.5 lower limit.

Illustration 14

From the following data calculate Median.

Class	Frequency
11-20	21
21-30	19
31-40	60
41-50	42
51-60	24
61-70	18
71-80	15

Solution

Class	Frequency	c.f
11-20	21	21
21-30	19	40
31-40	60	100
41-50	42	142
51-60	24	166
61-70	18	184
71-80	15	199

m = $\frac{N}{2}$ nd item

N = 199

m = $\frac{199}{2}$ = 99.5

M = $l_1 + \frac{l_2 - l_1}{f_1} \text{ x } m - c$

l_1 = 30.5

l_2 = 40.5

f_1 = 60

m = 99.5

c = 40

M = 30.5 + $\frac{40.5 - 30.5}{70}$ x 99.5 - 40

M = 30.5 + $\frac{10}{60}$ x 59.5

M = 30.5 + $\frac{595}{60}$

M = 30.5 + 9.9 = 40.4

Illustration 15

From the following Incomes of 9,990 persons. Calculate Median Income.

Revenue (Rs.)	No. of Persons
0-9	2756
10-19	2124
20-29	1677
30-39	1481
40-49	1021
50-59	610
60-69	245
70-79	67
80-89	6
90-99	3
	9990

Solution

Revenue (Rs.)	No. of Persons	Exclusive Class	Cumulative Frequency
0-9	2756	-0.5-9.5	2756
10-19	2124	9.5-19.5	4880
20-29	1677	19.5-29.5	6557
30-39	1481	29.5-39.5	8039
40-49	1021	39.5-49.5	9059
50-59	610	49.5-59.5	9669
60-69	245	59.5-69.5	9914
70-79	67	69.5-79.5	9981
80-89	6	79.5-89.5	9987
90-99	3	89.5-99.5	9990
	9990		

$m = \frac{N}{2}$

N = 9990

$m = \frac{9990}{2} = 4995$

$M = l_1 + \frac{l_2 - l_1}{f_1} \times m - c$

$l_1 = 19.5$

$l_2 = 29.5$

$f_1 = 1677$

m = 4995

c = 4880

$M = 19.5 + \frac{29.5 - 19.5}{1677} \times 4995 - 4880$

$M = 19.5 + \frac{10}{1677} \times 15$

$M = 19.5 + \frac{1150}{1677}$

M = 19.5 + 0.6 = 20.1

M = 20.1

6.3.3.2 Un equal Classes

When the class intervals are unequal the frequencies need not be adjusted to make the class intervals equal and the same formula for intepolation can be applied.

Illustration 16

Marks	No. of Students
0-10	5
10-30	15
30-60	30
60-80	8
80-90	2

Solution

Marks	No. of Students	Cumulative Frequency
0-10	5	5
10-30	15	20
30-60	30	50
60-80	8	58
80-90	2	60

Median = Size of $\frac{N}{2}$nd item

= $\frac{N}{2}$nd item = 30th item.

Median lies in the Class 30-60.

Median = $l_1 + \frac{l_2 - l_1}{f_1} \text{ x } m - c$

l_1 = 30

l_2 = 60

f_1 = 30

m = 30

c = 20

M = 30 + $\frac{60 \; - \; 30}{30}$ x 30 - 20

$$M = 30 + \frac{30}{30} \times 10$$

$$M = 30 + \frac{300}{30}$$

M = 30 + 10 = 40

M = 40

6.3.3.3 When Mid Points are given

When Mid points are given in the problem construct class by taking difference between two mid points.

Example :

Mid Point	Class		
115	110-120	(115-5)	(115+5)
125	120-130	(125-5)	(125+5)
135	130-140	(135-5)	(135+5)

125 - 115 = 10, 10 / 2 = 5, 115 - 5 = 110, 115 + 5= 120

Illustration 17

Compute Median from the following data.

Mid Value	Frequency
115	6
125	25
135	48
145	72
155	116
165	60
175	38
185	22
195	3

Solution :

Mid Value	Frequency	Cumulative Frequency
115	6	6
125	25	31
135	48	79
145	72	151
155	116	267
165	60	327
175	38	365
185	22	387
195	3	390
	390	

Difference between two mid points is 115 - 125 = 10

Divide 10 by 2 = i.e. 5

Mid Value -5 = Lower Limit = 115 - 5 = 110

Mid Value +5 = Upper limit = 115 + 5 = 120

Class is = 110 - 120

Median = Size of $\frac{N}{2}$ nd item

$= \frac{380}{2}$ th item = 195 th item

Median class = 150 - 160

Median = $l_1 + \frac{l_2 - l_1}{f_1} \text{ x } m - c$

$M = 150 + \frac{160 - 150}{116} \text{ x } 195 - 151$

$$M = 150 + \frac{10}{116} \times 44$$

$$M = 150 + \frac{440}{116}$$

M = 150 + 3.79; M = 153.79

Illustration 18

Mid Value	Frequency
1	3
2	60
3	101
4	152
5	205
6	155
7	79
8	40

Solution

Difference between two mid points is 1

Divide 1 by 2 = 0.5

Mid value - 0.5 = Lower Limit

Mid value + 0.5 = Upper Limit

1 - 0.5 = 0.5 - Lower Limit

1 + 0.5 = 1.5 Upper Limit

Class	Frequency	Cummulative Frequency
0.5 - 1.5	3	3
1.5 - 2.5	60	63
2.5 - 3.5	101	164
3.5 - 4.5	152	316
4.5 - 5.5	205	521
5.5 - 6.5	155	676
6.5 - 7.5	79	755
7.5 - 8.5	40	795
	795	

Median = Size of $\frac{N}{2}$ nd item

= $\frac{795}{2}$ th item = 397.5 th item.

Median = $l_1 + \frac{l_2 - l_1}{f_1}$ x m - c

M = 4.5 + $\frac{5.5 - 4.5}{205}$ x 397.5 - 316

M = 4.5 + $\frac{1}{205}$ x 81.5

M = 4.5 + $\frac{81.5}{205}$

M = 4.5 + 0.4 ; M = 4.9

6.3.3.4 Cumulative Frequency - Median (Less than, More than Methods)

When the data are given in the form of 'Less than', 'More than'. The given frequency is cumulative frequency. It is necessary to convert it into simple frequency distribution.

Illustration 19

From the following 655 Students. Calculate Median.

Values	Frequency
Less than 5	29
Less than 10	224
Less than 15	465
Less than 20	582
Less than 25	634
Less than 30	644
Less than 35	650
Less than 40	653
Less than 45	655

Solution

Values	Frequency	Cumulative Frequency	
0 - 5	29	29	29
5 - 10	195	(224-29)	224
10 - 15	241	(465-224)	465
15 - 20	117	(582-468)	582
20 - 25	52	(634-582)	634
25 - 30	10	(655-634)	644
30 - 35	6	(650-644)	650
35 - 40	3	(653-650)	653
40 - 45	2	655-653)	655
	655		

Median = Size of $\frac{N}{2}$ nd item

N = 655

m = $\frac{655}{2}$ = 327.5

Median = $l_1 + \frac{l_2 - l_1}{f_1} \times m - c$

l_1 = 10

l_2 = 15

f_1 = 241

m = 327.5

c = 224

M = 10 + $\frac{15 - 10}{241}$ x 327.5 - 224

M = 10 + $\frac{5}{241}$ x 103.5

M = 10 + $\frac{517.5}{241}$

M = 10 + 2.14. M = 12.14

Illustration 20

From the following Marks. Calculate Median Mark.

No.of Students	100	90	80	60	32	20	13	5
Marks	Less than 80	Less than 70	Less than 60	Less than 50	Less than 40	Less than 30	Less than 20	Less than 10

Solution

Marks	No. of Students	Marks	No. of Students	Cumulative Frequency
Less than 10	5	0-10	5	5
Less than 20	13	10-20	13	8
Less than 30	20	20-30	20	7
Less than 40	32	30-40	32	12
Less than 50	60	40-50	60	28
Less than 60	80	50-60	80	20
Less than 70	90	60-70	90	10
Less than 80	100	70-80	100	10
				100

Median = Size of $\frac{N}{2}$ nd item

N = 100

m = $\frac{100}{2}$ = 50th item

Median = $l_1 + \frac{l_2 - l_1}{f_1} \text{ x } m - c$

$l_1 = 40, l_2 = 50, f_1 = 28, m = 50, \quad c = 32$

M = 40 + $\frac{50-40}{28}$ x 50 - 32

M = 40 + $\frac{10}{28}$ x 18

M = 40 + $\frac{180}{28}$

M = 40 + 6.4

M = 46.42

More Than :

Illustration 21

From the following data. Calculate Median.

Class	Frequency
More than 90	51
More than 100	49
More than 110	49
More than 120	43
More than 130	37
More than 140	17
More than 150	5

Solution

Class	Frequency	Cumulative Frequency
90-100	2	2
100-110	0	2
110-120	6	8
120-130	6	14
130-140	20	34
140-150	12	46
150-160	5	51
	51	

Median = Size of $\frac{N}{2}$ nd item

N = 51

m = $\frac{51}{2}$ nd item

i.e. 25.5th item

Median = $l_1 + \frac{l_2 - l_1}{f_1} \times m - c$

l_1 = 130

l_2 = 140

f_1 = 20

m = 25.5

c = 14

M = 130 + $\frac{140 - 130}{20}$ x 25.5 - 14

M = 130 + $\frac{10}{20}$ x 11.5

M = 130 + $\frac{115}{20}$

M = 130 + 5.75

M = 135.75

6.4 CALCULATION OF MEDIAN BY GRAPHIC METHOD

Median can be calculated by Graphic Method. This is possible with the help of ogive curves which are also known as cumulative frequency curves. Cumulative frequency curves are two types.

a) Less than Curve: In order to draw these curves we have first of all to convert the ordinary frequencies into a cumulative frequency series. The frequency of all the preceding class intervals are summed up to the frequency of a class. We start with the upper limits of the classes and go on adding the frequencies. In this case of ogive curve has a rising trend.

b) More than Curve : In this case the frequencies of all the succeeding classes are added to the frequency of a class. We start with the lower limits of the classes and from the cumulative frequencies, we subtract the frequency of each class. In this case the ogive curve has a down ward trend.

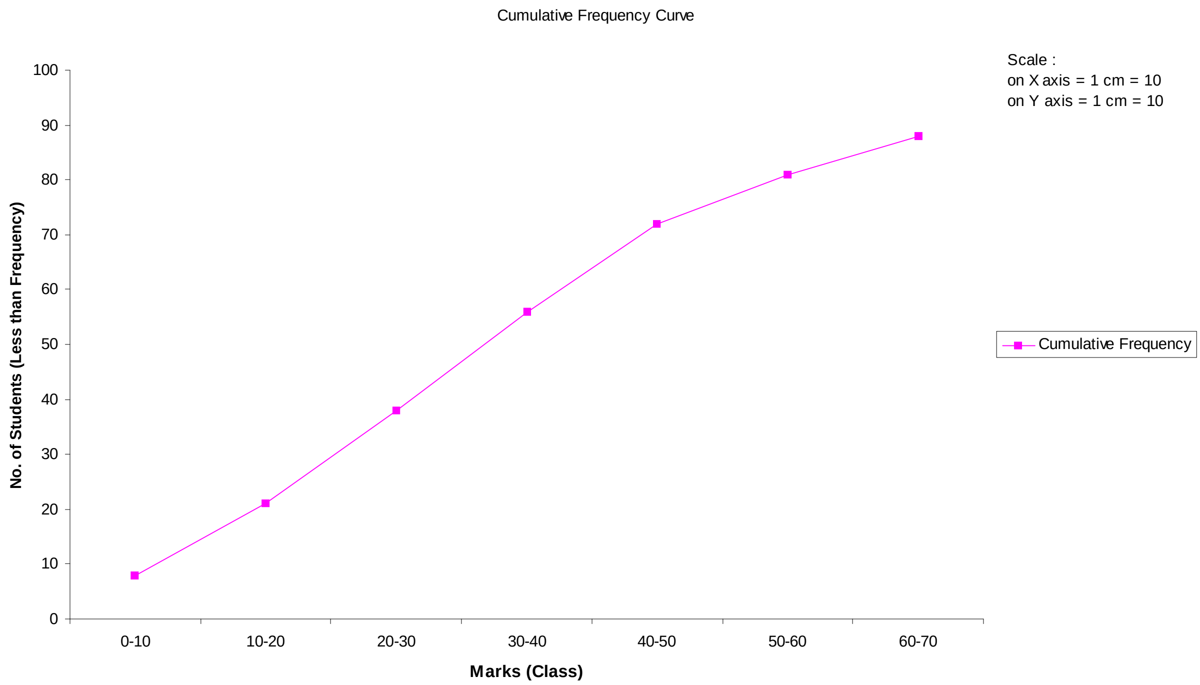
Illustration 22

From the following data. Locate Median through Graph.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Students	8	13	17	18	16	9	7

Solution

Marks	No. of Students	Cumulative Frequency
0-10	8	8
10-20	13	21
20-30	17	38
30-40	18	56
40-50	16	72
50-60	9	81
60-70	7	88



Note : Show less than frequency on 'Y' axis, and Marks on 'X' axis.

Illustration 23

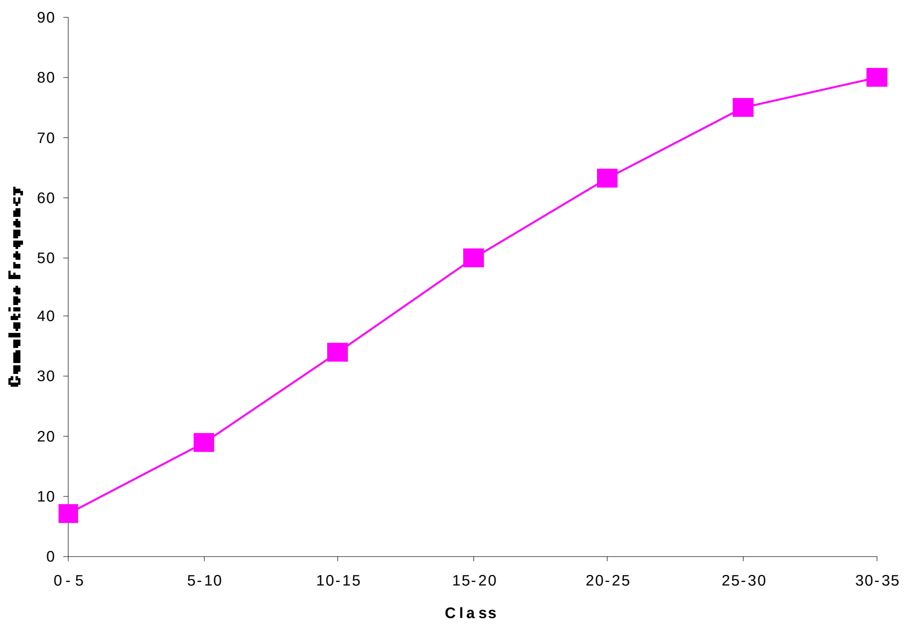
From the following data show Median through Graph.

Class	0-5	5-10	10-15	15-20	20-25	25-30	30-35
Frequency	7	12	15	16	13	12	5

Solution

Marks	No. of Students	Cumulative Frequency
0-5	7	7
5-10	12	19
10-15	15	34
15-20	16	50
20-25	13	63
25-30	12	75
30-35	5	80
	80	

Scale on X axis 1cm = 10, Y axis 1 cm = 10



Less than, More than Ogive Curves

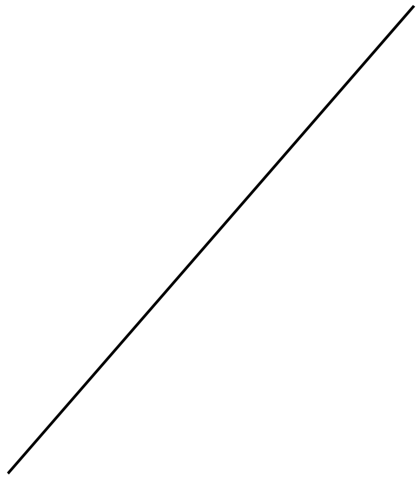


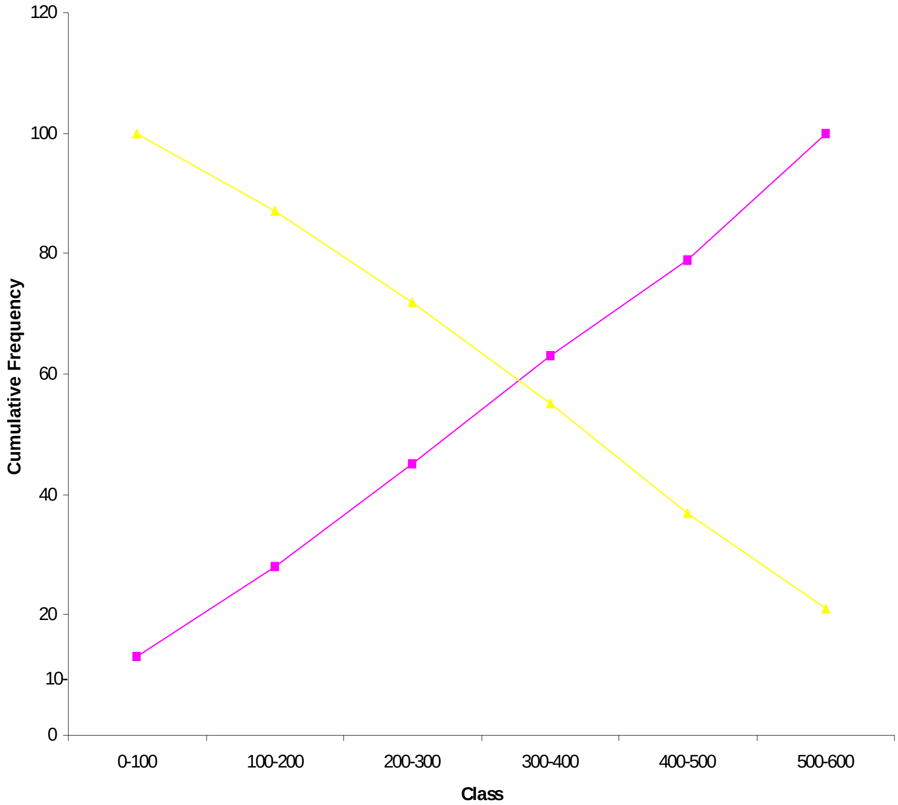
Illustration 24

From the following data show the Median with the help of two Ogive Curves.

Class	0-100	100-200	200-300	300-400	400-500	500-600
Frequency	13	15	17	18	16	21

Solution

X	f	c.f. Less than	c.f. More than
0-100	13	13	100
100-200	15	28	87
200-300	17	45	72
300-400	18	63	55
400-500	16	79	37
500-600	21	100	21
	N=100		



Here median is a point where two ogive curves interesect.

M = 327.78

6.5 MERITS OF MEDIAN

Following are the important merits or advantages of Median.

1 It is especially useful in case of open end classes since only the position and not the values of items must be known.

2 It is not influenced by the magnitude of extreme deviation from it.

3 In a markedly skewed distribution such as income distribution or price distribution where the arithmetic mean would be distorted by extreme values the median is especially useful.

4 It is most appropriate average in dealing with qualitative data.

5 The value of median can be determined graphically whereas the value of mean cannot be graphically as certained.

6.6 LIMITATIONS OF MEDIAN

Following are the limitations of Median.

1 For calculating median it is necessary to arrange th data. Other averages do not need any arrangements.

2 Since it is a positional average, its value is not determined by each and every observation.

3 It is not capable of algebraic treatment.

4 The value of median is affectd more by sampling fluctuaions.

5 When the numbe of items included in a series of data is even, the median is determined approximately as the mid-point of the two middle numbers.

6.7 SUMMARY

Thus Median is the value which divides the data into two parts. it is called a positional average. The term position refers to the place of a value in a series. If there are even number of items in a series there is no actual value eractly in the middle of the series and as such the median is indeterminate Median also can be derived with the help of graph.

6.8 EXERCISE

1. What is Median, Explain its merits and limitations.

2. From the following particulars calculate Median.

Marks : 75 24 42 57 63 49 91 12 8 20 35

3. From the following data find out Median.

No. of Children	0	1	2	3	4	5	6	7	8	9	10
No. of Families	7	12	75	89	80	47	35	23	12	13	5

4. Calculate Median from the following data?

Wages (Rs.)	31	32	33	34	35	36	37	38	40	
No. of Workers		3	7	8	13	16	15	14	5	2

5. From the following data calculate Median.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	
No. of Students		8	9	13	16	17	15	12	7	3

6. From the following data calculate Median.

Class	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	7	12	15	19	18	17	16	13

7. Calcualte Median from the following data.

Mid Values	12.5	13.0	13.5	14.0	14.5	15.0	15.5	16.0	16.5	17.0
Frequency	13	19	23	27	28	31	26	21	18	17

8. From the following data caluclate Median.

Marks	<10	<20	<30	<40	<50	<60	<70	<80	<90	<100
No. of Sudents	3	9	18	30	43	60	76	90	98	100

9. Calculate Median from the following data.

Values	>20	>30	>40	>50	>60	>70
Frequency	65	63	40	40	18	7

Ans.: 53.4

10. Calculate Median from the following data.

Class	>30.0	>32.5	>35.0	>37.5	>40.0	>42.5	>45.0	>47.5	>50.0	>52.5	>55.0
Frequency	940	903	825	646	655	271	186	103	38	6	1

(Ans. : 39.80)

11. From the following data find out Median.

Revenue (Rs.)	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99	
No. of Persons		2756	2124	1677	1481	1021	610	245	67	6	3

- Dr.K. Kanaka Durga

LESSON - 7

Averages - III

OTHER POSITIONAL MEASURES
OR
PATITION VALUES

OBJECTIVE

After studying this lesson you should be able to understand.

1. What are Positional Measures.

2. How to calculate Positional Measures.

STRUCTURE OF LESSON

7.1 Introduction

7.2 Quartiles

- **7.2.1 Quartiles - Individual Series**
- **7.2.2 Quartiles - Discrete Series**
- **7.2.3 Quartiles - Continuous Series**

7.3 Deciles

- **7.3.1 Deciles - Individual Series**
- **7.3.2 Deciles - Discrete Series**
- **7.3.3 Deciles - Continuous Series**

7.4 Percentiles

- **7.4.1 Percentiles - Individual Series**
- **7.4.2 Percentiles - Discrete Series**
- **7.4.3 Percentiles - Continuous Series**

7.5 Summary

7.6. Exercise

7.1 INTRODUCTION

Besides median, there are other measures which divide a series into equal parts. Important amongst these are quartiles, deciles, and percentiles.

7.2 QUARTLES

Quartiles are those values of the variate which divide the total frequences into four equal parts. There are three Quartiles denoted by Q. They are

1. Lower Quartile - Q_1
2. Upper Quartile - Q_3
3. Middle Quartile(Median) - Q_2

The procedure of computing quartiles is the same as the median.

7.2.1 Individual Series : Quartiles

While computing Quartiles in Individual Series we add 1 to N.

First Quartile : Q_1 (Lower Quartile)

$$Q_1 = \text{Size of } \frac{N+1}{4} \text{th item}$$

$$Q_3 = \text{Size of } \frac{N+1}{4} \text{ x 3rd item}$$

N = No .of items.

Illustration

From the following data calculate First Quartile and Third Quartile.

Wages Rs.	45	50	60	55	75	70	85	90	90	95	100

Solution :

Arrange data in order.

Wages 45 50 55 60 70 75 85 90 90 95 100

$$Q_1 = \frac{N+1}{4} \text{th item}$$

N = 11 (No. of items)

$$Q_1 = \frac{11+1}{4} \text{th item}$$

$$Q_1 = \frac{12}{4} \text{th item}$$

Q_1 = 3rd item

$\therefore\ Q_1 = 55$

Third Quartile or Upper Quartile : Q_3

$$Q_3 = \frac{N+1}{4} \text{ x 3rd item}$$

N = No. of items

N = 11

$$Q_3 = \frac{11+1}{4} \text{ x 3rd item}$$

$$Q_3 = \frac{12}{4} \text{ x 3rd item}$$

Q_3 = 3 x 3rd Item

Q_3 = 9th Item

$\therefore$ Q_3 = 90

Illustration 2

From the following data. Compute Q_1 and Q_3.

S.No.	1	2	3	4	5	6	7	8	9	10	11
Wages (Rs.)	61	64	66	67	68	69	70	73	74	75	76

Solution

$$Q_1 = \frac{N+1}{4} \text{th item}$$

N = 11 (No. of items)

$$Q_1 = \frac{11+1}{4} \text{th item}$$

$$Q_1 = \frac{12}{4} \text{th item}$$

Q_1 = 3rd item

$\therefore$ Q_1 = 66

Third Quartile or Upper Quartile : Q_3

$Q_3 = \frac{N+1}{4}$ x 3rd item

N = No. of items

N = 11

$Q_3 = \frac{11+1}{4}$ x 3rd item

$Q_3 = \frac{12}{4}$ x 3rd item

Q_3 = 3 x 3rd Item

Q_3 = 9th Item

$\therefore Q_3 = 74$

Illustration 3

S.No.	1	2	3	4	5	6	7	8	9	10	11	12
Wages (Rs.)	61	64	66	67	68	69	70	73	74	75	76	78

Solution

$\mathbf{Q_1 = \frac{N+1}{4}}$ **th item**

N = 12 (No. of items)

$Q_1 = \frac{12+1}{4}$ th item

$Q_1 = \frac{13}{4}$ th item

Q_1 = 3.25 th item

i.e. 3rd item + 0.25 x 4th item - 3rd item

66 + 0.25 x 67 - 66

66 + 0.25 x 1

66 + 0.25 = 66.25

$\therefore Q_1 = 66.25$

Third Quartile or Upper Quartile : Q_3

$Q_3 = \frac{N+1}{4}$ x 3rd item

N = No. of items

N = 12

$Q_3 = \frac{12+1}{4}$ x 3rd item

$Q_3 = \frac{13}{4}$ x 3rd item

Q_3 = 3.25 x 3rd Item

Q_3 = 9.75 th Item

Q3 = 9th item + 0.75 x 10th item - 9th item

= 74 + 0.75 x 75 - 74

= 74 + 0.75 x 1

$\therefore$ Q_3 = 74.75

7.2.2 Discrete Series

$Q_1 = \frac{N+1}{4}$ th item

$Q_3 = \frac{N+1}{4}$ x 3rd item

N = Total of the frequency

Illustration

From the following data calculate Q_1 and Q_3.

Values	2	3	4	5	7	9	11	12
Frequency	1	9	4	7	4	5	1	8

Solution

Values	Frequency	Cf
2	1	1
3	9	10
4	4	14
5	7	21
7	4	25
9	5	30
11	1	31
12	8	39

$Q_1 = \frac{N+1}{4}$ th item

N = 39

$Q_1 = \frac{39+1}{4}$ th item

$Q_1 = \frac{40}{4}$ th item

Q_1 = 10 th item

$\therefore\ Q_1 = 3$

Third Quartile or Upper Quartile : Q_3

$Q_3 = \frac{N+1}{4}$ x 3rd item

N = 39

$Q_3 = \frac{39+1}{4}$ x 3rd item

$Q_3 = \frac{40}{4}$ x 3rd item

Q_3 = 10 x 3rd Item

Q_3 = 30 th Item

$\therefore Q_3 = 9$

Illustration 5

Find out Quartiles.

X	0	1	2	3	4	5	6
f	13	54	75	90	64	21	15

Solution

X	0	1	2	3	4	5	6
f	13	54	75	90	64	21	15
Cf	13	67	142	232	296	317	332

$Q_1 = \frac{N+1}{4}$ th item

N = 332

$Q_1 = \frac{332+1}{4}$ th item

$Q_1 = \frac{333}{4}$ th item

Q_1 = 83.25 th item

$\therefore Q_1 = 2$

Third Quartile or Upper Quartile : Q_3

$Q_3 = \frac{N+1}{4}$ x 3rd item

N = 332

$Q_3 = \frac{332+1}{4}$ x 3rd item

$Q_3 = \frac{333}{4}$ x 3rd item

Q_3 = 83.25 x 3rd Item

Q_3 = 249.75 th Item

$\therefore$ Q_3 = 4

7.2.3 Continous Series

$q_1 = \frac{N}{4}$ th item

$Q_1 = l_1 + \frac{l_2 - l_1}{f_1} \times q_1 - C$

$Q_3 = \frac{N}{4}$ x 3rd item

$Q_3 = l_1 + \frac{l_2 - l_1}{f_1} \times q_3 - C$

l_1 = lower limit of Quantile class

l_2 = Upper limit of Quantile class

f_1 = frequency of Quantile class

q_1 = Value of q_1

C = Cumulative frequency of preceeding class of Quantile class.

Illustration 6

From the following data calculate Quantiles.

Age	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No. of Workers	50	70	100	180	150	120	70	60

Solution

Age	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	
No. of Workers	50	70	100	180	150	120	70	60	800
Cumulative Frequency	50	120	220	400	550	670	740	800	

$q_1 = \frac{N}{4}$ th item

N = 800

$q_1 = \frac{800}{4}$ th item i.e. 200th item

200 th item is in cumulative frequency of 220.

The corresponding class is 30 - 35.

$Q_1 = l_1 + \frac{l_2 - l_1}{f_1} x\ q_1 - C$

$l_1 = 30, l_2 = 35, f_1 = 100, q_1 = 200, C = 120$

$Q_1 = 30 + \frac{35-30}{100} x\ 200 - 120$

$Q_1 = 30 + \frac{5}{100} x\ 80$

$= 30 + \frac{400}{100}$

$Q_1 = 30 + 4 = 34$

$q_3 = \frac{800}{4}$ x 3rd item

= 200 x 3rd item = 600 th item

$Q_3 = l_1 + \frac{l_2 - l_1}{f_1} x\ q_3 - C$, Quartile Class = 45 - 50

$l_1 = 45, l_2 = 50, f_1 = 120, q_3 = 600, C = 550$

$Q_3 = 45 + \frac{50-45}{120} \text{ x } 600 - 550$

$= 45 + \frac{5}{120} \text{ x } 50 = 45 + \frac{250}{120} = 45 + 2.08 = 47.08$

$Q_3 = 47.08$

Illustration 7

From the following information Calculate Quartiles.

Values	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	5	25	40	70	90	40	20	10

Solution

Values	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	
Frequency	5	25	40	70	90	40	20	10	300
Cumulative Frequency	5	30	70	140	230	270	290	300	

$q_1 = \frac{N}{4}$ th item

N = 300

$q_1 = \frac{300}{4}$ th item i.e. 75 th item

$Q_1 = l_1 + \frac{l_2 - l_1}{f_1} \text{ x } q_1 - C$

$l_1 = 30, l_2 = 40, f_1 = 70, q_1 = 75, C = 70$

$Q_1 = \quad 30 + \frac{40-30}{70} \text{ x } 75 - 70$

$Q_1 = \quad 30 + \frac{10}{70} \text{ x } 5$

Q_1 = 30 + 0.71 = 30.71

$q_3 = \frac{300}{4}$ x 3rd item

= 75 x 3rd item = 225 th item

$Q_3 = l_1 + \frac{l_2 - l_1}{f_1} \times q_3 - C$, Quartile Class = 45 - 50

$Q_3 = 40 + \frac{50-40}{90} \times 225 - 145$

$= 40 + \frac{10}{90} \times 85 = \frac{850}{90} = 9.44$

Q_3 = 49.44

7.3 DECILES

Deciles divide the series into 10 equal parts. For any series, there are 9 deciles, as there are three quartiles for any series. Deciles range from D_1 to D_9.

7.3.1 Individual Series

$D = \frac{N+1}{10}$ x Required decile

N = Number of Items.

Illustration 8 :

From the following data calculate 8th decile.

Marks	11	12	14	18	22	26	30	32	35	41	45

Solution :

$D_8 = \frac{N+1}{10} \times 8$

N = 11

$D_8 = \frac{11+1}{10} \times 8$

$D_8 = \frac{12}{10}$ x 8 = $\frac{96}{10}$ = 9.6

9th item + 0.6 x 10th item - 9th item

35 + 0.6 x 41 - 35

35 + 0.6 x 6

35 + 3.6 = 38.6

D_8 = 38.6

7.3.2 Discrete Series

D = $\frac{N+1}{10}$ x Required decile

N = Number of Items.

Illustration 9

Calculate 7th decile from the following data.

Height (cm)	157	168	173	152	162	176
No. of persons	10	13	2	1	25	1

Solution :

Arrange data in order.

Height (cm)	152	157	162	168	173	176
No. of persons	1	10	25	13	2	1
Cumulative Frequency	1	11	36	49	51	52

$D_7 = \frac{N+1}{10}$ x Required decile

N = 7

$D_7 = \frac{52+1}{10}$ x 7 $= \frac{53}{10}$ x 7 = 37.1 item

D_7 = 168

7.3.3 Continuous Series

$d = \frac{N}{10}$ x Required Number

$D = l_1 + \frac{l_2 - l_1}{f_1}$ x d - C

Illustration 10

From the following data calculate 6th decile.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of Students	5	25	40	70	90	40	20	10

Solution

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of Students	5	25	40	70	90	40	20	10
Cumulative Frequency	5	30	70	140	230	270	290	300

$D = \frac{N}{10}$ x 6 th item, N = 300

$= \frac{300}{10}$ x 6 = 180

6th Decile lies in the class 40 - 50

$D_6 = 40 + \frac{50-40}{90}$ x 180 - 140

$= 40 + \frac{10}{90}$ x 40 = 40 + $\frac{400}{90}$ = 4.44

D_6 = 44.44

7.4 PERCENTILES

Percentiles divide the series into 100 parts. For any series, there are 99 percentiles. Percentiles is denoted by P. It ranges from P_1 to P_{99}.

7.4.1 Individual Series

$$P = \frac{N+1}{100} \text{ x Required Percentile}$$

N = No. of Items

Illustration 11

From the following data calculate 61st percentile.

Values	22	26	14	30	18	11	35	41	12	32

Solution :

Inorder

Values	11	12	14	18	22	26	30	32	35	41

$$P_{61} = \frac{N+1}{100} \text{ x } 61$$

N = 10

$$P_{61} = \frac{10+1}{100} \text{ x } 61$$

$$P_{61} = \frac{671}{100} = 6.71 \text{ th item}$$

6th item + 0.71 x 7th item - 6th item

26 + 0.71 x (30 - 26)

26 + 0.71 x 4

26 + 2.84 = 28.84

P_{61} = 28.84

7.4.2 Discrete Series

$$P = \frac{N+1}{100} \text{ x Required Percentile}$$

N = Total of the Frequency

Illustration 12:

From the following data calculate 95th Percentile.

Heights (cms)	152	155	157	160	162	164	168	170	171	172	173	176
No. of Persons	1	5	10	12	25	38	13	6	4	3	2	1

Solution

Heights (cms)	152	155	157	160	162	164	168	170	171	172	173	176
No. of Persons	1	5	10	12	25	38	13	6	4	3	2	1
Cumulative Frequency	1	6	16	28	53	91	104	110	114	117	119	120

$$P_{95} = \frac{N+1}{100} \times 95$$

N = 120

$$P_{95} = \frac{120+1}{100} \times 95 = 114.95 \text{ th item}$$

P_{95} = 172

7.4.3 Continous Series

$$p = \frac{N}{100} \times \text{Required Percentile}$$

$$P = l_1 + \frac{l_2 - l_1}{f_1} \times p - C$$

Illustration 13

Find out 85th Percentile.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of Students	8	12	20	32	30	28	12	4

Solution

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of Students	8	12	20	32	30	28	12	4
Cumulative Frequency	8	20	40	72	102	130	142	146

$p = \frac{N}{100} \times 85$, N = 146

$p = \frac{146}{100} \times 85$ = 124.1 th item

$P_{85} = l_1 + \frac{l_2 - l_1}{f_1} \times p - C$

$= 50 + \frac{60-50}{28} \times 124.1 - 102$

$= 50 + \frac{10}{28} \times 22.1 = 50 + 7.89 = 57.89$

$P_{85} = 57.89$

7.5 SUMMARY

Thus, besides Median,there are other positional measures which divide a series into equal parts. Important amongst these are quartiles, deciles and percentiles. In economics and business statistics quartiles are more widely used than deciles and percentiles. The deciles and percentiles are important in psychological and educational statistics concerning grades, ranks and rates etc.

7.6 EXERCISE

1. Explain two quartiles.

2. Explain percentiles.

3. From the following information fingout Q_1, Q_2, D_6, P_{20}.

Marks :	0 -10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 70	70 - 80
No. of Students :	5	25	40	70	90	40	20	10

(Ans. : Q_1 = 30.71, Q_3 = 49.44, D_6 = 44.4, P_{20} = 27.5)

4. From the following data compute 1st decile, 7th decile, 9th decile, 33rd percentile.

Marks : 35, 76, 63, 24, 12, 95, 47, 55, 85, 93, 3, 18, 29, 59, 69,

30, 29, 51, 68, 71, 80, 99, 8, 13, 41, 89, 73, 20, 9, 5.

(Ans. D_1 = 8.1, D_7 = 70.4, D_9 = 92.6, P_{33} = 10.23)

5. From the following data calculate Q_1, Q_3, D_6, P_3.

Marks	No. of Students
Less than 80	100
Less than 70	90
Less than 60	80
Less than 50	60
Less than 40	32
Less than 30	20
Less than 20	13
Less than 10	5

(Ans. : Q_1 = 34.25, D_6 = 50, P_3 = 6)

6. From the following data find out Q_1, Q_3, D_2, P_{90}.

Weight	Below 10	10 - 20	20 - 40	40 - 60	60 - 80	80 - 100
No. of Persons	8	10	22	25	10	5

(Ans. Q_1= 21.82, Q_3 = 56, D_2 = 18)

7. From the following data computer D_7, P_{85}.

Deposits (Rs.)	0 - 100	100-250	250-400	400-500	500-550	550-600	600-800	800-900	900-1000
No. of Deposits	25	100	175	74	66	35	5	18	2

(Ans. D_7 = 467.57, P_{85} = 538.64)

- Dr. K.Kanaka Durga.

LESSON 8

AVERAGES - IV

MODE

8.0 OBJECTIVE

After studying this lesson you should be able to understand.

1. What is Mode.
2. How to Calculate Mode
3. What are the merits and limitations
4. Geometric Mean
5. Harmonic Mean

STRUCTURE OF LESSON

8.1 Introduction

Mode like median is also a positional measure. Mode is useful in determining the stock of different goods. Since mode helps us in determining the popularity of a Commodity so it gives opportunity to the business men to stock such items as to get windfall gains.

8.2 Definition and Meaning

The most frequently occurring item of the series is known as mode. Mode is defined by “Croxton and Cowden” as “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”. According to ‘Zizek’, Mode is “The value occuring most frequently in a series of items and around which the other items are distributed most densely”.

The mode is the item which is repeated maximum times in the series will be the mode of the series. Thus in a given series, mode is the most popular and common item. This word is derived from the French word, *La mode* which means fashion or the most popular phenomenon.

Mode is the most specific or typical value of a series or the value around which maximum concentration of items occur.

For instance. A shirt maker would like to know the size of shirts that has the maximum demand. He will produce shirts of that size, which has the maximum demand.

8.3 Calculation of Mode

Mode can be calculated in Individual series, discrete series and Continuous series.

8.3.1 Individual Series

For determing mode count the number fo times the various values rpeat themselves and the value which occurs the maximum number of times is the model value.

Illustration

Find mode from the following data.

Values : 110 120 130 120 110 140 130 120 140 120

Since the value 120 occurs the maximum numbers of times. i.e., 4. Hence the modal value is 120.

When there are two or more values having the same maximum frequency one cannot say which is the model value and hence mode is said to be defined. Such a series is also known as bimodal or multimodal

Illustration

Find out Mode from the following data.

Income (Rs.) : 610 620 630 620 610 640 630 620 630 640

Solution

Size of Item	No. of Items it occurs
610	2
620	3
630	3
640	2

Here Mode is MUltiple Mode because 620, 630 repeated same number of times.

8.3.2 Discrete Series

In discrete series Mode is located by preparing a ‘grouping table’ and ‘analysis table’.

a) Grouping Table : A grouping table has six columns.

1. In column one the maximum frequency is marked.
2. In column two frequencies are grouped in two's.
3. In column three leave the first frequency and then group the remaining in two's.
4. In column four group the frequencies in three's.
5. In column five leave the first frequency and group in three's.
6. In column six leave the first two frequencies and then group the remaining in three's.

In each of these take the maximum total and mark it in a circle or by bold type.

b) Analysis Table : After preparing grouping table prepare analysis table, while preparing the grouping table. Put column number on the left hand side and the various probable. Values of mode on the right-hand side. The values against which frequencies are the highest are marked in the grouping table and then entered by means of a bar in the relevant box corresponding to values they represent.

Illustration

From the following data calculate Model wage.

Daily Wage Rs. :	41	42	43	44	45	46	47	48	49	50
No. of Workers :	8	17	20	22	19	14	10	8	5	3

Solution

Statement of Grouping

Daily Wage X	No. of Workers 1	2	3	4	5	6
41	8					
42	17	25				
43	20		37	45	59	
44	22	42				61
45	19		<u>41</u>	55		
46	14	33			43	32
47	10		24			
48	8	18		23	16	
49	5		13			
50	3	8				

Satement of Analysis

F \ V	41	42	43	44	45	46	47	48	49	50
1				✓						
2			✓	✓	✓					
3				✓	✓	✓				
4				✓						
5		✓	✓	✓						
6			✓	✓	✓					
		1	3	6	3	1				

44 repreated six times, so modal wage : 44

Illustration

Find out Mode from the following data.

Value	:	60	61	62	63	64	65	66
Frequency	:	27	146	435	398	210	128	98

Statement of Grouping

Values	Frequency 1	2	3	4	5	6
60	27					
61	146	173				
62	**435**		581	608	**979**	
63	398	**833**				
64	210		**608**	**736**		**1043**
65	128	338			436	
66	98		226			

Satement of Analysis

F \ V	60	61	62	63	64	65	66
1			✓				
2			✓	✓			
3				✓	✓		
4				✓	✓	✓	
5		✓	✓	✓			
6			✓	✓	✓		
		1	4	5	3	1	

Hence Mode is 63 because it is repeated 5 times.

Illustraiton

Weight (Pounds)	:	135	136	137	138	139	140	141	142	143
No. of Persons	:	4	16	20	18	10	4	25	3	2

Solution :

Statement of Grouping

Weights	No. of Persons 1	2	3	4	5	6
135	4					
136	16	20	36			
137	20			40		
138	18	38			54	48
139	10		28	32		
140	4	14			39	32
141	25		29			
142	3	28		30		
143	2		5			

Satement of Analysis

F \ V	135	136	137	138	139	140	141	142	143
1							✓		
2			✓	✓					
3		✓	✓						
4	✓	✓	✓						
5	✓	✓	✓						
6		✓	✓	✓					
	2	4	5	2			1		

Here Mode is 137 because it repeated 5 times.

8.3.3 Continuous Series

1. By preparing grouping table and analysis table ascetain the model class.

2. Determine the value of mode by applying the following formula.

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x (l_2 - l_1)$$

l_1 = Model class lower limit

l_2 = Model class upper limit

f_1 = General frequency of Model class

f_2 = General frequency of Succeeding class of Modal class

f_0 = General frequency of Preceeeding class of Modal class

Illustration

From the following data calculate Mode.

Values	0-5	5-10	5-15	15-20	20-25	25-30	30-35	35-40	40-45
Frequence	20	24	32	28	20	16	37	10	8

Statement of Grouping

Values X	Frequence 1	2	3	4	5	6
0-5	20	44		76		
5-10	24		56		84	
10-15	32	60				80
15-20	28		48	64		
20-25	20	36			73	
25-30	16		53			63
30-35	<u>37</u>	47		55		
35-40	10		18			
40-45	8					

Satement of Analysis

F \ V	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
1								✓	
2			✓	✓					
3		✓	✓						
4	✓	✓	✓						
5		✓	✓	✓					
6			✓	✓	✓				
	1	3	5	3	1			1	

Here Modal class is 10 - 15 because it is repeated 5 times.

Then apply the following principle to find out Mode.

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x l_2 - l_1$$

$l_1 = 10, l_2 = 15, f_1 = 32, f_2 = 28, f_0 = 24$

$$Z = 10 + \frac{32 - 24}{2x32 - 24 - 28} x15 - 10$$

$$Z = 10 + \frac{8}{64 - 24 - 28} x5$$

$$Z = 10 + \frac{8}{12} x5$$

$$Z = 10 + \frac{40}{12} = 10 + 3.33$$

$\therefore Z = 13.33$

Inclusive Method

From the following data calculate Mode.

Class	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45
Frequency	7	10	16	32	24	18	10	5	1

Solution :

Statement of Grouping

Values X	Frequence 1	2	3	4	5	6
0.5-5.5	7					
5.5-10.5	10	17	26	33		
10.5-15.5	16				58	72
15.5-20.5	<u>32</u>	48	56	74		
20.5-25.5	24	42				
25.5-30.5	18		28		52	33
30.5-35.5	10	15	6	16		
35.5-40.5	5					
40.5-45.5	1					

Satement of Analysis

F \ V	0.5-5.5	5.5-10.5	10.5-15.5	15.5-20.5	20.5-25.5	25.5-30.5	30.5-35.5	35.5-40.5	40.5-45.5
1				✓					
2			✓	✓					
3				✓	✓				
4				✓	✓	✓			
5			✓	✓					
6		✓	✓	✓	✓				
		1	3	6	3	1			

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x\, l_2 - l_1$$

$l_1 = 15.5$, $l_2 = 20.5$, $f_1 = 32$, $f_2 = 24$, $f_0 = 16$

$$Z = 15.5 + \frac{32 - 16}{2x32 - 16 - 24} x20.5 - 15.5$$

$$Z = 15.5 + \frac{16}{24} x5$$

$\therefore Z = 18.75$

Unequal Classes :

Illustration

Values	0-2	2-4	4-8	8-10	10-15	15-20	20-25	25-30	30-35	35-40	40-50	50-60	60-70
f	1	2	2	3	6	8	10	15	18	22	36	10	4

Statement of Grouping

Values X	Frequence 1	2	3	4	5	6
0-10	8					
10-20	14	22	39	47		
20-30	25				79	
30-40	40	<u>65</u>				
40-50	36		<u>76</u>	86		104
50-60	10	46				
60-70	4		14		50	

Satement of Analysis

F \ V	10-10	10-20	20-30	30-40	40-50	50-60	60-70
1				✓			
2			✓	✓			
3				✓	✓		
4				✓	✓	✓	
5			✓	✓			
6		✓	✓	✓	✓		
		1	3	6	3	1	

Modal Class = 30-40

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x l_2 - l_1$$

$l_1 = 30$, $l_2 = 40$, $f_1 = 40$, $f_2 = 36$, $f_0 = 25$

$$Z = 30 + \frac{40 - 25}{2x40 - 25 - 36} x40 - 30$$

$$Z = 30 + \frac{15}{19} x\,10$$

$$Z = 30 + \frac{150}{19} = 30 + 7.89$$

$\therefore Z = 37.89$

Less than - More than

1. Change the cumulative frequency into genral frequency.
2. Construct the class.

Illustration

From the following data calculate Modal mark.

Marks	No. of Students
Less than 5	29
Less than 10	224
Less than 15	465
Less than 20	582
Less than 25	634
Less than 30	644
Less than 35	650
Less than 40	653
Less than 45	655

Statement of Grouping

Marks X	No. of Students 1	General Frequence 2	3	4	5	6
0-5	29					
5-10	195	224	436	465		
10-15	<u>241</u>				553	
15-20	117	358	169			410
20-25	52			179		
25-30	10	62			68	
30-35	6		16			19
35-40	3	9	5	11		
40-45	2					

Satement of Analysis

F \ V	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
1			✓						
2			✓	✓					
3		✓	✓						
4	✓	✓	✓						
5		✓	✓	✓					
6			✓	✓	✓				
	1	3	6	3	1				

Modal Class - 10 - 15

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x l_2 - l_1$$

$l_1 = 10$, $l_2 = 15$, $f_1 = 241$, $f_2 = 117$, $f_0 = 195$

$$Z = 10 + \frac{241 - 195}{2x241 - 195 - 117} x\ 40 - 30$$

$$Z = 10 + \frac{46}{170} x\ 5$$

Z = 10 + 1.3

$\therefore$ Z = 11.35

Illustration

From the following data calculate Mode.

Mid Values	Frequency
Above 0	80
Above 10	77
Above 20	72
Above 30	65
Above 40	55
Above 50	43
Above 60	28
Above 70	16
Above 80	10
Above 90	8
Above 100	0

Solution:

Statement of Grouping

Marks X	No. of Students 1	2	3	4	5	6
0-10	3	8		15		
10-20	5		12		22	
20-30	7	17				29
30-40	10		22	37		
40-50	12	27			39	
50-60	15		27			33
60-70	12	18		20		
70-80	6		8		16	
80-90	2	10				10
90-100	8					
100-110	0					

Satement of Analysis

F \ V	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
1						✓				
2					✓	✓				
3						✓	✓			
4				✓	✓	✓				
5					✓	✓	✓			
6						✓	✓	✓		
				1	3	6	3	1		

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x l_2 - l_1$$

$$Z = 50 + \frac{15 - 12}{2x15 - 12 - 12} x\ 60 - 50$$

$$Z = 50 + \frac{3}{30 - 12 - 12} x10$$

$$Z = 50 + \frac{3}{6} x10$$

Z = 50 + 5

$\therefore$ Z = 55

Illustration

From the following data calculate Mode.

Mid Values	10	20	30	40	50	60	70
Frequency	7	12	17	29	31	5	3

Solution

Statement of Grouping

Class	Frequency 1	2	3	4	5	6
5-15	7	19		36		
15-25	12		29			
25-35	17	46			58	
35-45	29		60	65		77
45-55	31	36			39	
55-65	5		8			
65-75	3					

Satement of Analysis

F \ V	5-15	15-25	25-35	35-45	45-55	55-65	65-75
1				✓			
2		✓	✓				
3			✓	✓			
4			✓	✓	✓		
5	✓	✓	✓				
6		✓	✓	✓			
	1	3	5	4	1		

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x l_2 - l_1$$

$$Z = 35 + \frac{29 - 17}{2x29 - 15 - 31} x\ 45 - 35$$

$$Z = 35 + \frac{12}{10} x10$$

$$Z = 35 + \frac{12}{10} x10$$

Z = 35 + 1.2 x 10 = 35 + 12 = 47

$\therefore$ Z = 47

Here Mode should not lies in the Modal class so, Mode can be obtained with the following principle.

$$Z = l_1 + \frac{f_2}{f_0 + f_2} x l_2 - l_1$$

$$Z = 35 + \frac{31}{17 + 31} x45 - 35$$

$$Z = 35 + \frac{31}{48} x10$$

Z = 35 + 6.4

Z = 41.4

Multiple Mode

Statement of Grouping

Class	Frequency 1	2	3	4	5	6
90-100	3	5		23		
100-110	2		20		42	
110-120	18	40				61
120-130	22		43	62		
130-140	21	40			50	
140-150	19		29			32
150-160	10	13		15		
160-170	3		5			
170-180	2					

Satement of Analysis

F \ V	90-100	100-110	110-120	120-130	130-140	140-150	150-160	160-170	170-180
1				✓					
2			✓	✓	✓	✓			
3				✓	✓				
4				✓	✓	✓			
5			✓		✓	✓	✓		
6				✓	✓				
			2	5	5	3	1		

Here Mode is Multiple. In case of Multiple Mode apply the following principle to locate Mode.

Z = (3 x Median) - (2 x Mean)

Median :

Class	Frequency	cf
90-100	3	3
100-110	2	5
110-120	18	23
120-130	22	45
130-140	21	66
140-150	19	85
150-160	10	95
160-170	3	98
170-180	2	100

Median = $\frac{N}{2}$ nd item

= $\frac{100}{2}$ nd item

= 130 + $\frac{140-130}{21}$ x 50 - 45

= 130 + $\frac{10}{21}$ x 5

= 130 + $\frac{50}{21}$ x 2.38

Median = 132.38

Arithmetic Mean :

Class	Frequency	Mid Point	dx	fdx
90-100	3	95	-40	-120
100-110	2	105	-30	-60
110-120	18	115	-20	-360
120-130	22	125	-10	-220
130-140	21	135	0	0
140-150	19	145	+10	+190
150-160	10	155	+20	+200
160-170	3	165	+30	+90
170-180	2	175	+40	+80
	100			+560-760 =-200

(A.M.) $\overline{X}$ $= X + \frac{\Sigma fdx}{N}$

$= 135 + 135 + \left(\frac{-200}{100}\right)$

= 135 - 2 = 133

Z = (3 x Median) - (2 x A.M.)

Median = 132.38

A.M. = 133

Z = (2 x 132.38) - (2 x 133)

Z = 397.14 - 266 = 131.14

Z = 131.14

8.4 Locating Mode Graphically

In a frequency distribution the value of mode can also be determined graphically. The steps in calculation are:

1. Draw a histogram of the given data.
2. Draw two lines diagonally in the inside of the modal class bar, starting from each upper corner of the bar to the upper corner of the adjacent bar.
3. Draw a perpendicular line from the intersection of the two diagonal lines to the X-axis (horizontal scale) which gives us the modal value.

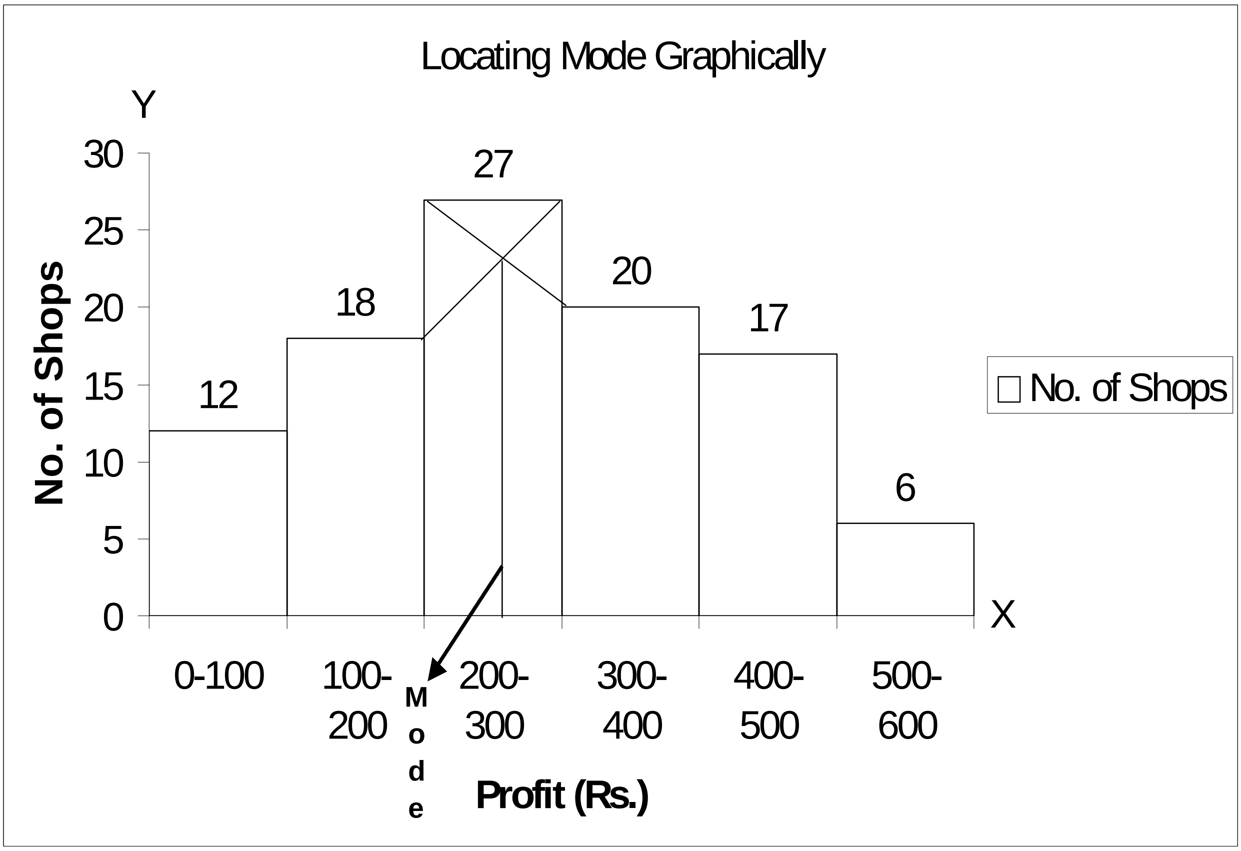
Illustration

The monthly profits in rupees of 100 shops are distributed as follows.

Profit per Shop	No. of Shops
0-100	12
100-200	18
200-300	27
300-400	20
400-500	17
500-600	06

Draw a histogram of the data and find out the Modal value. Check this value by direct calculation.

$$Z = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} x l_2 - l_1$$



l_1 = 20, l_2 = 300, f_1 = 27, f_2 = 20, f_0 = 18

$$Z = 200 + \frac{27-18}{2x27-18-20} x300-200$$

$$Z = 200 + \frac{9}{54-18-20} x\,100$$

$$Z = 200 + \frac{900}{16}$$ = 200 + 52.25

$\therefore$ Z = 256.25

Mode can also be determined from a frequeency polygon in which case a perpendicular is drawn on the base from the apex of the polygon and the point where it intersects the base given the modal value.

Graphic method of determing Mode cannot be determined if two or more classes have the same highest frequency.

8.5 Merits of Mode

The main merits of Mode are:

1. The mode is the most frequently occuring value. If the modal wage in a factory is Rs.4100 then more workers receive Rs.4,100 than any other wage. This wage is known as average wage or Modal wage.

2. It is not affected by extremely large or small items.

3. Its value can be determined in open end distributions without asceertaining the class limits.

4. It can be used to describe qualitative phenomenon.

5. The value of Mode can also be determined graphically.

Limitations

The following are the important limitations of Mode.

1. The value of Mode cannot always be determined, because in some cases we may have a bimodal series.

2. It is not capable of algebraic manipulations.

3. The value of Mode is not based on each and every item of the series.

4. It is not a rigidly defined measure.

8.6 Summary

Thus the value occuring maximum times is the modal value. This can be known by inspection in Individual Series. In discrete series, mode can be known by inspection. It means to look to that value of the series around which the items are most heavily concentrated. In continuous series after knowing Modal class, grouping and analysis principle is applied to know the Mode.

8.7 Exercise

1. Define Mode, How it is useful ?

2. How to locate Mode graphically ?

3. Describe Merits and Limitations of Mode.

4. Calculate Mode.

49, 35, 21, 46, 57, 67, 57, 13, 99 (Ans. 57)

5. From the following data. Calculate Mode.

Color Size (in inches)	No. of Persons
12.0	10
12.5	28
13.0	38
13.5	42
14.0	45
14.5	15
15.0	8
15.5	7

(Ans. 13.5)

6. From the following data calculate.

Values	:	2	3	4	5	6	7	8	9	10	11
Frequency	:	3	8	10	12	16	14	10	8	17	5

(Ans. : 6)

7. Calculate Mode from the following data.

Color Size (in inches)	No. of Persons
55	8
65	10
75	16
85	14
95	10
105	5
115	2

(Ans. 75)

8. Calculate Mode from the following wages of 50 workers working in a factory.

Daily Wages Rs.:	4	5	7	8	10	11	13	14	16
No. of Workers :	2	3	2	6	10	11	12	3	1

(Ans. : 11)

9. Find out Mode

Size of Items :	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Frequency :	4	10	16	18	24	28	28	30	22	26	18	14	8	6

(Ans. 7)

10. From the following data findout Modal size.

Values :	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22
Frequency :	5	48	189	303	522	980	981	794	515	474

(Ans. 18.005)

11. Findout Mode.

Wages	20-30	30-40	40-50	50-60	60-70	70-80	80-90	
No. of Workers	85	120	110	67	49	21	6	(Ans. 37.8)

12. Calculate Mode.

Class	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
Frequency	50	70	80	150	180	120	70	50

(Ans. 42)

13. Findout Mode.

Marks	No. of Students
More than 0	80
More than 10	77
More than 20	72
More than 30	65
More than 40	55
More than 50	43
More than 60	28
More than 70	16
More than 80	10
More than 90	8
More than 100	0

14. Calculate Mode.

Class	Frequency
More than 90	51
More than 100	49
More than 110	49
More than 120	43
More than 130	37
More than 140	17
More than 150	5

(Ans. 136.36)

15. Calculate Mode with the help of the Principal = Z = (3 x Median) - (2 x Mean)

Class	Frequency
Less than 5	29
Less than 10	224
Less than 15	465
Less than 20	582
Less than 25	634
Less than 30	644
Less than 35	650
Less than 40	653
Less than 45	655

(Ans. 10.69)

16. Find out Mode.

Profit Rs.	No. of Traders
Below 20	5
Below 30	14
Below 40	27
Below 50	48
Below 60	68
Below 70	83
Below 80	91
Below 90	94

(Ans. 52.25)

LESSON 9

AVERAGES : V

[Geometric Mean, Harmonic Mean]

9.0 OBJECTIVE

After studying this lesson you should be able to understand.

1. What is Geometric Mean, How to Calculate Geometric Mean

2. What is Harmonic Mean, How to Calculate Harmonic Mean

STRUCTURE OF THE LESSON

9.1 INTRODUCTION

There are two means other than Mean, Medin and Mode which are occassionally used in economics and business. These are Geometric Mean and Harmonic Mean. Averages are also called Ratio - Averages because these are more suitable when the data comprise rates, percentages of ratios insteaded of actual quantities.

9.2 Geometric Mean (G.M.) - Definiton and Meaning

Geometric Mean is defined as the nth root of the product of N items or values. If there are two items, we take the square root; if there are three items, the cube root; and so on. Symbolically

Geomatric Mean = $\sqrt[n]{(X_1)(X_2)......(X_n)} = [(X_1)(X_2)......(X_n)]^{1/n}$

Where X_1, X_2, X_n refers to the various items of the series.

When the number of items is three or more the task of multiplying the numbers and of exracting the root becomes excessively difficult. To simplify calculations logarithms are used. Geometric Mean is calculated as follows :

$$\text{Log Geometric Mean} = \log X_1 + \log X_2 + ... \log X_n$$

$$\text{Log Geometric Mean} = \frac{\Sigma \log X}{N}$$

$$\therefore \text{Geometric Mean} = \text{Antilog}\frac{\Sigma \log X}{N}$$

9.2.1 Properties of Geometric Mean

The following are the important mathematical properties of Geometric Mean.

1. The product of the value of series will remain unchanged when the value of Geometric Mean is substituted for each individual value.

For Example : The Geometric for series 2, 4, 8 is 4:

Therefore, we have 2 x 4 x 8 = 64 = 4 x 4 x 4

2. The sum of the deviations of the logarithms of the original observations above or below the logarithms of the geometricmean is equal. This also means that the value of the Geometric mean is such as to balance the ratio deviations of the observations from it.

From example : 2, 4, 8, is 4. We find that $\left(\frac{4}{2}\right)\left(\frac{4}{4}\right) = 2 = \left(\frac{8}{4}\right)$

3. A note worthy point there is that the Geometric Mean is always lower than arithmetic mean becuase it gives more weightage to small values.

4. If any item is '0', value of Geometric Mean is also '0'.

9.3 Calculation of Geometic Mean

Geometric Mena is calculated int he following three Series.

9.3.1 Individual Series

$$\text{Geometric Mean} = \text{Anti}\log\frac{\Sigma \log X}{N}$$

1. Take the logarithms of the variable X and obtain the total $\Sigma \log X$.

2. Divide $\Sigma \log X$ by N and take the antilog of the value so obtained.

Illustration 1

From the following data calculate Geometric Mean.

Item : 3, 12, 76, 115, 6, 9, 10, 100, 476, 96

Sno.	1	2	3	4	5	6	7	8	9	10	**10**
Item X	3	12	76	115	6	9	10	100	476	96	
log X	0.4771	1.0792	1.8808	2.0607	0.7782	0.9542	1.0000	2.0000	2.6776	1.9823	**14.8909**

$$\text{Geometric Mean} = \text{Anti}\log \frac{\Sigma \log X}{N}$$

$\Sigma \log X$ = 14.8909

N = 10

$$\text{Geometric Mean} = \text{Anti}\log \frac{14.8909}{10}$$

Geometric Mean = Anti log of 1.48901

Geometric Mean = 30.83

Illustration 2

Calculate Geometric Mean.

Values : 85, 70, 15, 75, 500, 8, 45, 250, 40, 36

Solution

Values (X)	85	70	15	75	500	8	45	250	40	36	**N = 10**
log X	1.9294	1.8451	1.1761	1.8751	2.6990	0.9031	1.0532	2.3979	1.6021	1.5563	**17.6373**

$$\text{Geometric Mean} = \text{Anti}\log \frac{\Sigma \log X}{N}$$

$\Sigma \log X$ = 17.6373

N = 10

$$\text{Geometric Mean} = \text{Anti}\log \frac{17.6373}{10}$$

Geometric Mean = Anti log of 1.76373

Geometric Mean = 88.29

Illustration 3

X : 3834, 382, 63, 9, 0.4, 0.03, 0.009, 0.0005

Solution

X	3834	382	63	9	0.4	0.03	0.009	0.0005	**N = 8**
log X	3.5837	2.5821	1.7993	0.9031	0.3979	1.5229	2.0458	2.3010	**9.53377**

$$\text{Geometric Mean} = \text{Anti}\log \frac{\Sigma \log X}{N}$$

$\Sigma \log X$ = 9.53377

N = 8

$$\text{Geometric Mean} = \text{Anti}\log \frac{9.53377}{8}$$

Geometric Mean = Anti log of 1.9172

Geometric Mean = 83.60

9.3.2 Discrete Series

$$\text{Geometric Mean} = \text{Anti}\log \frac{\Sigma \log xf}{N}$$

1. Find the logarithms of the variable x.

2. Multiplty logarithms with the respective frequencies and obtain the total $\Sigma \log xf$.

3. Divide $\Sigma \log xf$ by the total frequency and take the anti log of the value so obtained.

Illustration 4

From the following wages of 50 workers. Calculate Geometry Mean of Wages.

Wages	31	32	33	34	35	36	37	38
No. of Workers	5	7	8	13	9	4	3	1

Solution

Wages(x)	No. of Workers	log x	log x f
31	5	1.4914	7.4570
32	7	1.5052	12.5364
33	8	1.5185	12.1480
34	13	1.5315	19.9095
35	9	1.5450	13.9050
36	4	1.5563	6.2252
37	3	1.5682	4.7046
38	1	1.5798	1.5798
	50		**76.4655**

$$\text{Geometric Mean} = \text{Anti log} \frac{\Sigma \log xf}{N}$$

$\Sigma \log xf$ = 76.4655

N = 50

$$\text{Geometric Mean} = \text{Anti log} \frac{76.4655}{50}$$

Geometric Mean = Anti log of 1.52931

Geometric Mean = 33.88

Illustration 5

Calculate Geometric Mean of the following distribution.

Variable	8	9	10	11	12	13	14
Frequency	11	8	6	9	7	3	1

Solution

Variable(x)	Frequency(f)	log x	log x f
8	11	0.9031	9.9341
9	8	0.9542	7.6336
10	6	1.0000	6.0000
11	9	1.0414	9.3726
12	7	1.792	1.5544
13	3	1.1139	3.3417
14	1	1.1461	1.1461
	45		**44.9825**

$$\text{Geometric Mean} = \text{Anti log} \frac{\Sigma \log xf}{N}$$

$\Sigma \log xf$ = 44.9825

N = 45

$$\text{Geometric Mean} = \text{Anti log} \frac{44.9825}{45}$$

Geometric Mean = Anti log of 0.9996

Geometric Mean = 9.991

9.3.3 Continuous Series

$$\text{Geometric Mean} = \text{Anti log} \frac{\Sigma \log xf}{N}$$

1. Find out the Mid point of the classes.

2. Multiply logarithms with the respective frequencies of each class and obtain the total $\Sigma \log xf$.

3. Divide the total obtained by the total frequency and take the anti log of the value so obtained.

Illustration 6

Compute the Geometric Mean from the following data.

Marks	0-10	10-20	20-30	30-40	40-50
No. of Students	5	7	15	25	8

Solution

Marks(x)	No. of Students(f)	Mid Points x	log x	log xf
0-10	5	5	0.6990	3.4950
10-20	7	15	1.1761	8.2327
20-30	15	25	1.3979	20.9685
30-40	25	35	1.5441	38.6025
40-50	8	45	1.6532	13.2256
	60			**84.5243**

$$\text{Geometric Mean} = \text{Anti log}\frac{\Sigma \log xf}{N}$$

$\Sigma \log xf = 84.5243$

N = 60

$$\text{Geometric Mean} = \text{Anti log}\frac{84.5243}{60}$$

Geometric Mean = Anti log of 1.4087

Geometric Mean = 25.63

Illustration 7

From the following data. Calculate Geometric Mean.

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequency	4	8	9	13	7	6	3

Solution :

Class (x)	No. of Students(f)	Mid Points x	log x	log xf
0-10	4	5	0.6990	2.7960
10-20	8	15	1.1761	9.4088
20-30	9	25	1.3976	12.5811
30-40	13	35	1.5450	20.0850
40-50	7	45	1.6532	11.5724
50-60	6	55	1.7404	10.4424
60-70	3	65	1.8129	5.4387
	50			**72.3244**

$$\text{Geometric Mean} = \text{Anti}\log \frac{\Sigma \log xf}{N}$$

$\Sigma \log xf$ = 72.3244

N = 50

$$\text{Geometric Mean} = \text{Anti}\log \frac{72.3244}{50}$$

Geometric Mean = Anti log of 1.446488

Geometric Mean = 27.957

9.4 MERITS AND LIMITATIONS OF GEOMETRIC MEAN

Merits or Advantages

1. Geometric Mean is rigidly defined.
2. It is based on all observations.
3. It is suitable for further mathematical treatment.
4. Geometric is useful in fixation of prices etc.
5. it is not affected much by fluctuations of sampling.
6. Geometric mean is useful when data are in rates, ratios, percentages, etc.
7. It is useful for finding the compound rates of change.
8. It is used in the construction of Index numbers.

Demerits or Limitations

1. Geometric is not easy to understand and calculate.
2. If any item is zero, Geometric Mean becomes zero.

9.5 HARMONIC MEAN : Definition and Meaning

Harmonic Mean is used in Special types of problems. it is based on arithmetic mean or reciprocals of the values of the variable.

Harmonic Mean is defined as the " reciprocals of the values of the variables".

Symbolically $$\text{Harmonic Mean} = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} .... + \frac{1}{x_n}} = \frac{N}{\Sigma\left(\frac{1}{x}\right)}$$

X = Variable

Harmonic Mean is always less than not only the arithmetic mean but geometric mean as well. It is becuase that the average gives weightage to smaller items i.e. reciprocal of 3 is $\frac{1}{3}$ and that of 4 is $\frac{1}{4}$. Also number of value of the variable be zero.

9.6 CALCULATION OF HARMONIC MEAN

When the number of items is large the computation of Harmonic Mean is tedious. To simplify calculations we obtain reciprocal of the various items from the table and apply the Principle.

9.6.1 Harmonic Mean - Individual Series

$$\text{Harmonic Mean} = \frac{N}{\Sigma\left(\frac{1}{x}\right)}$$

1. Obtain reciprocals of given number.
2. Obtain arithmetic mean of the reciprocals.
3. Find the reciprocal of the arithmetic mean.

Illustration 8

Calculate Harmonic Mean

X	1238	178.7	89.9	78.4	9.7	0.989	0.874	0.012	0.008	0.0009

Solution

X	1 / X
1238.0	0.0008
178.7	0.0056
89.9	0.0111
78.4	0.0128
9.7	0.1031
0.989	1.0111
0.874	1.1442
0.12	83.3333
0.008	125.0000
0.0009	1111.1111
N = 10	**1321.7331**

$$\text{Harmonic Mean} = \frac{N}{\Sigma\left(\frac{1}{x}\right)}$$

N = 10

$\Sigma\left(\frac{1}{x}\right)$ = 1321.7331

$$\text{Harmonic Mean} = \frac{10}{1321.7331}$$

Harmonic Mean = 0.0076

Illustration 9

Calculate Harmonic Mean.

X 3834 382 63 8 0.4 0.03 0.009 0.0005

Solution

X	1 / X
3874	0.0003
382	0.0027
63	0.0159
8	0.1250
0.4	2.5000
0.03	33.3333
0.009	11.1111
0.0005	2000.0000
N = 8	**2147.0883**

$$\text{Harmonic Mean} = \frac{N}{\Sigma\left(\frac{1}{x}\right)}$$

N = 8

$\Sigma\left(\frac{1}{x}\right)$ = 2147.0883

$$\text{Harmonic Mean} = \frac{8}{2147.0883}$$

Harmonic Mean = 0.003726

9.6.2 Harmonic Mean - Discrete Series

$$\text{Harmonic Mean} = \frac{N}{\Sigma\left(\frac{1}{xf}\right)}$$

1. Take the reciprocal of the various items.
2. Multiply the reciprocals by respective frequencies.
3. Substitute the vlues of N and $\Sigma\left(\frac{1}{xf}\right)$.

Instead of finding out the reciprocals first and then multiplying them by frequencies it will be far more easier to divide each frequency by the respective value of the variable.

$$\text{Harmonic Mean} = \frac{N}{\Sigma\left(\frac{f}{x}\right)}$$

Illustration

From the following data compute the value of Harmonic mean.

Marks	10	20	25	40	50
No. of Students	20	30	50	15	5

Solution

Marks (X)	No. of Students (f)	f / X
10	20	2.000
20	30	1.500
25	50	2.000
40	15	0.375
50	5	0.100
	120	**5.975**

$$\text{Harmonic Mean} = \frac{N}{\Sigma\left(\frac{f}{x}\right)}$$

N = 120

$\Sigma\left(\frac{f}{x}\right)$ = 5.975

$$\text{Harmonic Mean} = \frac{120}{5.975}$$

Harmonic Mean = 20.08

9.6.3 Harmonic Mean - Continous Series

$$\text{Harmonic Mean} = \frac{N}{\Sigma\frac{1}{x}f}$$

1. Take the mid points of class.

2. Take the reciprocals of the mid points.

3. Multiply the reciprocals by respective frequencies.

4. Substitute the values in Principle.

Illustration:

Calculate Harmonic Mean from the following data.

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	8	12	13	15	17	16	14	5

Solution :

Class (X)	Frequency (f)	Mid Points	(1 / X)	$\frac{1}{X}$ x f
0-10	8	5	0.2000	1.0000
10-20	12	15	0.0667	1.0005
20-30	13	25	0.0400	1.0000
30-40	15	35	0.0286	1.0010
40-50	17	45	0.0222	0.9990
50-60	16	55	0.0182	1.0010
60-70	14	65	0.0154	0.8645
70-80	5	75	0.0133	0.9975
	N = 100			**7.8635**

$$\text{Harmonic Mean} = \frac{N}{\Sigma \frac{1}{x} f}$$

N = 100, $\Sigma \frac{1}{x} f$ = 7.8635

$$\text{Harmonic Mean} = \frac{100}{7.8635}$$

Harmonic Mean = 12.72

9.7 MERITS AND LIMITATIONS OF HARMONIC MEAN

Merits :

1. Harmonic Mean is rigidly defined.

2. It is based on all the items.

3. It is suitable for further algebraic treatment.

4. It gives greater weightage to smaller values. (because of reciprocal usage)

5. It is not affected by fluctuations of sampling.

6. It is useful in averaging special types of rates and ratios.

Demerits

1. It is not easy to calculate and understand.

2. If one of the items is zero Harmonic Mean can not be calculated.

3. It is hardly used in business problems. Because it is not a representative figure of the distribution unless the phenomenon needs greater importance to be given to smaller items.

9.8 SUMMARY

Thus the Geometric Mean and Harmonic Mean are two means which are occassionally used in economics and business. These are more suitable when the data comprises rates, percentages of ratios instead of actual quantities. Geometric Mean is also useful for finding the compound rates of chzanges like the rates of growth of population in a country over a period of time or the average rate of increase or decrease in the turnover of a business. Harmonic Mean would be representative when different rates of speed, for equal distances have to be averaged.

9.9 EXERCISE

1. What is Geometric Mean, what are its Merits and limitations.

2. What is Harmonic Mean.

3. When do we use Harmonic Mean.

4. What are the Merits and Demerits of Harmonic Mean.

5. Find the Harmonic Mean 2574, 46575.5, 0.8, 0.08, 0.005, 0.0009

(Ans. : 0.00604)zz

X 85, 70, 15, 75, 500, 8, 45, 250, 40, 36 (Ans. : 58.03)

7. Calculate Geometric Mean of the following data.

X 0.009, 0.005, 0.08, 0.8, 5, 75, 475, 2574 (Ans. : 1.841)

8. Find out Geometric Mean

X 10, 110, 135, 120, 50, 59, 60, 7 (Ans. : 46.56)

9. Find the Geometric mean from the following data.

X	2	3	5	6	4	
f	10	15	18	12	7	(Ans.: 3.850)

10. Compute Geometric Mean

X	10	15	18	20	25	
f	2	3	5	6	4	(Ans.: 18.2)

11. Calculate Geometric Mean

X	5	15	25	35	45	
f	5	7	15	25	8	(Ans. : 25.63)

12. Find out Geometric Mean

X	10-20	20-30	30-40	40-50	50-60	
f	5	10	15	7	4	(Ans. : 31.72)

13. Compute the Geometric Mean

X	0-10	10-20	20-30	30-40	40-50
f	10	5	8	7	20

14. The following marks are related to 60 students in Economics, Compute Geometric Mean.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	
No. of Students	3	8	15	20	10	4	(Ans. : 28.02)

15. Calculate Harmonic Mean.

X 10 20 40 60 120 (Ans. : 25)

16. Find out Harmonic Mean.

X	3834	382	63	0.8	0.4	0.03	0.009	0.0005

(Ans. : 0.00373)

17. Find out Geometric Mean

X	10	20	25	40	50	
f	20	30	50	15	5	(Ans. : 20.08)

18. Find out Harmonic Mean

Marks	0-10	10-20	20-30	30-40	40-50	
No. of Sudents	4	6	10	7	3	(Ans. : 16.03)

19. Find out Geometric Mean

Class Interval	10-20	20-30	30-40	40-50	50-60
Frequency	4	6	10	7	3

- Dr.K. Kanaka Durga

Lesson - 10

MEASURES OF DISPERSION - I

(Range, Quartile Deviation & Mean Deviation)

OBJECTIVES:

By the study of this chapter you will be able to understand the meaning of dispersion and three measures of dispersion (Range, Quartile Deviation & Mean Deviation). You will also be thorough with merits, demerits and method of computing all these three measures of dispersion.

STRUCTURE:

10.1 Introduction

10.2 Differencess between central tendency & Measures

10.3 Objectives of Measures of Dispersion

10.4 Types of measures of Dispersion

10.5 Range - Individual, Discrete & continueous serial - Examples

10.6 Merits of Range

10.7 Demerits of Range

10.8 Quartile Deviation - Introduction - All series with Example

10.9 Merits of Quartile Deviation

10.10 Demerits of Quartile Deviation

10.11 Mean Deviation - Introduction - All series with Examples

10.12 Merits of Mean Deviation

10.13 Demerits of Mean Deviation

10.14 Summary

10.15 Questions

10.16 Exercises

10.1 INTRODUCTION:

" Measures of Dispersion " or " Measures of Variation " are the "Average of second order" They are based on the average of deviations of the values obtained from the central tendencies i.e. Arithmetic Mean (a), Median (M), or Mode (z). The variability is the basic feature of the values of variables. Such type of variation or dispersion refers to the "lack of uniformity ".

10.2 DIFFERENCES BETWEEN CENTRAL TENDENCIES AND DISPERSION :

Following are the distinctions between the central tendencies and Dispersions -

Central Tendency	Dispersions
1. Average of the first order	1. Average of the second order
2. Do not throw light on the formation of series	2. Throw light on the fromation of series or distribution
3. Do not give detailed features of obseravations	3. Give detailed characteristics of obseravations
4. Do not establish relationship with the items	4. Establish relationship with the individual items
5. Do not reveal entire picture of distribution	5. Reveal the entire picture of the distribution
6. Give only the idea of concentration of item	6. Give the idea of deviation from central tendencies.

An average of second order is an average of the difference of all the items of the series from an average of those items. In averaging these differences or deviations, their irregularities are brushed off and a representative of dispersion results in.

All the distributions are not similar. They differ in numerical size of their average and in their respective formations. Let us observe the following series carefully

Series 1 :	30	30	30	30	30	30	30	30	30
Series 2 :	22	24	26	28	30	32	34	36	38
Series 3 :	6	12	18	24	30	36	42	48	54

Arithmetic mean and median in all the series are same i.e. : 30 but items in series differ widely. So the central tendencies fail to describe the scatterdness of the values. For measuring the nature of formation we require the average of second order in support of the first order.

10.3 OBJECTIVES :

The objectives of computing the second order averages are given below. -

a) To ascertain the suitability of the first order averages

b) To decide the consistency of performance and

c) To reveal the degree of uniformity in the series

In the three series as given above, constituted differently though their mean and median are the same. The first series is uniformly distributed and there is no dispersion at all. The second series is having same sort of dispersion from the central tendency and the uniformity is disturbed.

The third series shows a high degree of dispersion and there is no uniformity among the items. Thus it can be concluded that the larger the dispersion is, the lower will be the uniformity in the distribution.

10.4 TYPES OF MEASURES OF DISPERSION :

Measures of dispersion are mainly 4 types -

1. Range
2. Quartile Deviation or semi- inter quartile range
3. Mean deviation or Average deviation
4. Standard deviation

1. Range :

The difference between line Highest value (H) and least value (L) of a series is called the 'Range'. ' Range ' represents the difference between the extreme values. The values in between the two extremes are not at all taken into consideration.

Range (R) = H – L

Coefficient of Range = $\frac{H-L}{H+L}$ --------- Relative measure.

H = Highest value

L = Least value.

Individual series - Range : Range (R) = H – L

Coefficient of Range = $\frac{H-L}{H+L}$

Example 1

Compute the range and the coefficient of range of the series and state which one is more dispersed and which one is more uniform.

Values of Variables :

Series 1 :	13	14	15	16	17	(a =15)
Series 2 :	9	12	15	18	21	(a=15)
Series 3 :	1	8	15	22	29	(a=15)

" Central tendency is same but formation differ "

Solution :

	I	II	III
Range :	R = H – L	R = H – L	R = H – L
	= 17 - 13	= 21-9	= 29-1
	= 4	= 12	= 28
Coefficient of Range:	$\frac{H-L}{H+L}$	$\frac{H-L}{H+L}$	$\frac{H-L}{H+L}$
	$\frac{17-13}{17+13}$	$\frac{21-9}{21+9}$	$\frac{29-1}{29+1}$
	= 1.33	= 0.4	= 0.93

Series I is ' Less ' dispersed and more uniform.

Series II is ' Less ' Uniform and more dispersed

Discrete Series :

Range (R) = H – L

Coefficient of Range = $\frac{H-L}{H+L}$

Note : The frequencies are not to be taken into consideration in the computation of Range.

Example 2

From the following distribution find out the Range and its coefficient

Values of Variables :

Marks (x) :	1	2	3	4	5	6	7	8	9	10
No.of Students (f) :	4	7	12	13	18	16	14	9	5	2

Solution :

R = H – L

= 10 - 1

= 9

Coefficient of Range : $\frac{H-L}{H+L} = \frac{9}{11}$ = 0.81

Continuous Series :

Range (R) = H – L

Coefficient of Range = $\frac{H-L}{H+L}$

In finding out the Range in continuous series the frequencies are never taken into account. The upper limit of the Highest class (H) and lower limit of the least class (L) are only taken into account.

Example 3

Compute the Range and Coefficent from the following data

(x) :	10-12	12-14	14-16	16-18	18-20	20-22	22-24	24-26	26-28	28-30
(f) :	12	13	18	21	23	27	19	14	11	9

Solution :

R = H – L

= 30-10

= 20

Coefficient of Range : $\frac{H-L}{H+L}$

$= \frac{30-10}{30+10}$

$= \frac{20}{40} = 0.5$

10.6 MERITS OF RANGE :

Following are the merits of Range.

a) It is the simplest measure of dispersion

b) It is regidly defined and easiest measure of dispersion to compute

c) It is readily comprehensible and it requires very little calculations.

d) It is useful in statictical methods of quality control techniques

e) It is useful in studying the variations in the prices of share and stocks.

f) It is useful in studying weather conditions (weatheriology or meterology) where minimum and maximum temperature is identified

10.7 DEMERITS OF RANGE :

a) Unfortunately it is not a stable measure of dispersion, because it is affected by the extreme values only.

b) It is not suitable where the class intervals are open in the distribution.

c) It is completely depending upon the two extreme values but not on the other values.

d) It is not suitable for mathematical treatment

e) It is very sensitive to the fluctuations in the sampling size as the size of sample increase it tends to increase not in proportion.

10.8 QUARTILE DEVIATION OR SEMI - INTER QUARTILE RANGE :

Introduction : One of the demerits of the Range is that it is only affected by the extreme values. To over come this defect, the Quartile deviation is formulated with some modifications. It is similar to Range. For the computation of Quartile Deviation Q_3 and Q_1 will be taken as Highest and Least values. It means, the items below the Lower Quartile (Q_1) and the items above the Upper Quartile (Q_3) are not considered. It means only the middle may portion of the series will be considered. The range so obtained is divided by two to get the Quartile Deviation. Thus the Quartile Deviation measures the difference between the values of Q_1 and Q_3.

Quartile Deviation = Q.D = $\frac{Q_3 - Q_1}{2}$ -------Absolute Measure

Coefficient of Range = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$ --------- Relative measure.

Where Q_1 means first Quartile or Lower Quartile

Q_3 means Third Quartile or Upper Quartile

Individual series:

Q.D $= \frac{Q_3 - Q_1}{2}$

Coefficient of Range = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

$Q_1 = \frac{N+1}{4}$ th item

$Q_3 = \frac{N+1}{4}$ x 3rd item.

Note : The series must be arranged in an ascending order.

Example 4

From the following data, compute quartile Deviation and Co-efficient of Quartile Deviation.

Variables : 24,7,11,9,17,3,20,14,4,22,27

Solution :

Arranging the series in Ascending order :

S.No.	X
1	3
2	4
3	7
4	9
5	11
6	14
7	17
8	20
9	22
10	24
11	27
	N:11

$Q_1 = \frac{N+1}{4}$ th item

$= \frac{11+1}{4}$ th item $= \frac{12}{4}$ th item = 3rd item

3rd item = 7

$Q_3 = \frac{N+1}{4}$ x 3rd item

$= \frac{11+1}{4}$ x 3rd item $= \frac{12}{4}$ x3rd item

9th item = 22

Q.D $= \frac{Q_3 - Q_1}{2}$

$= \frac{22-7}{2} = 7.5$

Coefficient of Range = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

$= \frac{22-7}{22+7}$

$= \frac{15}{29} = 0.52$

Example 5

From the following Marks of 12 students compute the Quartile Deviation and its coefficient.

Marks : 43, 54, 67,80,89,84,72,61,48,30,25,37

Solution :

Ascending order -

S.No.	X
1	25
2	30
3	37
4	43
5	48
6	54
7	61
8	67
9	72
10	80
11	89
N:12	

$Q_1 = \frac{N+1}{4}$ th item

$= \frac{12+1}{4}$ th item = 3.25 th item = 3rd item + 25% of 6 (43-37)

37+1.5 = 38.5

$Q_3 = \frac{N+1}{4}$ x 3rd item

$= \frac{12+1}{4}$ x 3rd item

9.75th item =9th item + 75% of 8 (80-72)

= 72 + 6 = 78

$Q.D = \frac{Q_3 - Q_1}{2}$

$= \frac{78 - 38.5}{2} = 19.75$

Coefficient of Range = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

$= \frac{78 - 38.5}{78 + 38.5}$

$= \frac{39.5}{116.5} = 0.339$

Discrete Series - Quartile Deviation :

Q.D $= \frac{Q_3 - Q_1}{2}$

Coefficient of Range = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

$Q_1 = \frac{N+1}{4}$th item

$Q_3 = \frac{N+1}{4}$ x 3rd item.

These two items must be identified in the ' Cf ' and the corresponding variables shall be taken as Q_1 and Q_3

Example 6

compute the Quartile Deviation and its coefficient from the following data.

(x) :	21	22	23	24	25	26	27	28	29	30
(f) :	4	8	13	16	18	14	11	9	5	2

Solution :

X	f	cf
21	4	4
22	8	12
23	13	25
24	16	41
25	18	59
26	14	73
27	11	84
28	9	93
29	5	98
30	2	100
	N:100	

$Q_1 = \frac{N+1}{4}$ th item

$= \frac{100+1}{4}$ th item = 25.25 th item

It lies in the cf 41 and the corresponding variable is 24

$Q_1 = 24$

$Q_3 = \frac{N+1}{4}$ x 3rd item

$= \frac{100+1}{4}$ x 3rd item = 75.75 th item

It lies in the cf 84 and the corresponding variable is 27

$Q_3 = 27$

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$= \frac{24-27}{2} = \frac{3}{2} = 1.5$$

Coefficient of Range = $\dfrac{Q_3 - Q_1}{Q_3 + Q_1}$

$= \dfrac{24 - 27}{27 + 24}$

$= \dfrac{3}{51}$ = 0.0588

Continuous Series - Quartile Deviation :

Q.D $= \dfrac{Q_3 - Q_1}{2}$

Coefficient of Range = $\dfrac{Q_3 - Q_1}{Q_3 + Q_1}$

$Q_1 = \dfrac{N+1}{4}$ th item

$Q_3 = \dfrac{N+1}{4}$ x 3rd item.

These items must be identified in the ' cf ' and the corresponding classes shall the taken as Q_1 class Q_3 class. Then the following formula shall be applied to find the Q_1 and Q_3

$l + \dfrac{c \text{ x } i}{f}$

Where l = Lower limit of the class

c = difference between $\dfrac{N}{4}$ th item and the preceding cf

i = interval of the Quartile class

f = frequency of the quartile class

Note : Q_1 and Q_3 should not be calculated from the inclusive classes. They must be converted into Exclusive classes.

Example 7

compute the Quartile Deviation and its coefficient from the following data.

(x) :	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
(f) :	4	9	13	15	14	16	12	8	6	3

Solution :

X	**f**	**cf**
0-10	4	4
10-20	9	13
20-30	13	26
30-40	15	41
40-50	14	55
50-60	16	71
60-70	12	83
70-80	8	91
80-90	6	97
90-100	3	100
	N= 100	

Q_1 Position $= \frac{N}{4}$ th item

$= \frac{100}{4}$ th item = 25 th item

It lies in the cf 26 and corresponding class is 20-30

Thus Q_1 class = 20 - 30

$$Q_1 = l + \frac{c \times i}{f}$$

$$= 20 + \frac{(25 - 13) x 10}{13}$$

$= 20 + 9.232 = 29.232$

Q_3 Position $= \frac{N}{4}$ x 3rd item

$= \frac{100}{4}$ x 3rd item = 75 th item

H lies in the cf 83 and corresponding class is 60-70

Q_3 class = 60-70

$$Q_3 = l + \frac{c \times i}{f}$$

$$= 60 + \frac{(75-71)x10}{12}$$

= 60+3.33 = 63.33

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$= \frac{63.33 - 29.33}{2} = 17.05$$

$$\text{Coefficient of Range} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$= \frac{63.33 - 29.33}{63.33 + 29.33}$$

$$= \frac{34.10}{92.56} = 0.3684$$

Example 8

compute the Quartile Deviation and its coefficient from the following data.

(x) :	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
(f) :	3	8	12	13	19	18	17	14	9	5

Solution :

X	f	cf
0 - 9	3	3
10-19	8	11
20-29	12	23
30-39	13	36
40-49	19	55
50-59	18	73
60-69	17	90
70-79	14	104
80-89	9	113
90-99	5	118
	N= 118	

Q_1 Position $= \frac{N}{4}$ th item

$= \frac{118}{4}$ th item = 29.5 th item

It lies in the cf 36 and corresponding class is 30-39

Thus Q_1 class = 30-39 But it is an inclusive class. It must be converted into an exclusive class

Exclusive class = 29.5 - 39.5

$$Q_1 = l + \frac{c \times i}{f}$$

$$= 29.5 + \frac{(29.5 - 23)x10}{13}$$

$$= 29.5 + 5$$

$$= 34.5$$

Q_3 Position = $\frac{N}{4}$ x 3rd item

$= \frac{118}{4}$ x 3rd item = 88.5 th item

It lies in the cf 90. The corresponding class is 60-69

Q_3 class (Exclusive form) = 59.5 - 69.5

$$Q_3 = l + \frac{c \text{ x } i}{f}$$

$$= 59.5 + \frac{(88.5 - 73)x10}{17}$$

= 59.5 + 9.12

= 68.62

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$= \frac{68.62 - 34.5}{2} = \frac{34.12}{2}$$

= 17.06

$$\text{Coefficient of Range} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$= \frac{68.62 - 34.5}{68.62 + 34.5}$$

$$= \frac{34.12}{103.12} = 0.3311$$

10. 9 MERITS OF QUARTILE DEVIATION :

1. It is very easy to calculate and simple to understand.
2. It is not affected by extreme values of variable as it is concerned with the central half portion of distribution
3. It is not at all affected by open end class intervals.

10. 10 DEMERITS OF QUARTILE DEVIATION :

1. It ignores completely the portion below the lower quartile and above the upper quartile
2. It is not capable of further mathematical treatment
3. It is greatly affected by the fluctuation in the sampling
4. It is only a positional average but not mathematical average.

10.11 MEAN DEVIATION :

Introduction : The average of deviations taken from an average is called Mean Deviation(M.D) or Average Deviation. The base average may be either Mean or Median or Mode. But theoretically, the deviations of items are taken preferably from median instead that than form the Mean or the Mode. Mediam is supposed to be the suitable central tendency for calculating deviations because the sum of the deviations from the Median is less than the sum of deviations from the Mean. It is not a common pracitce to calculate the deviation from the mode as its value is sometimes not clearly defined.

In aggregating the deviations the algebric negative signs are not taken into account. It means all the deviations are treated as Positive ignoring the negative signs.

Individual series - Mean Deviation :

$$M.D = \frac{\in |dx|}{N}$$

Where MD = Mean Deviation

$\in$ |dx| = Total of the deviation taken from the average by ignoring the signs (+or –)

N= Number of variables.

first of all an average shall be calculated. It may be either mean or Median.

Arithmetic Mean (a) = $\frac{\in x}{N}$

Median = $\frac{N+1}{2}$ nd item (after arranging the series in an ascending order)

Deviations must be taken from the average to the other variables in the series by ignoring plus and minus.The total of these deviations must be devided with the number of deviations $\frac{\in |dx|}{N}$

Coefficient of Mean Deviation = $\frac{M.D}{Average}$

Example 9

Find out the Mean deviation from Mean and Median and also find out the coefficient.

(x) :	21	34	27	35	30	24	29	22	33	25

Solution : Computation of Mean Deviation from Mean

S.No.	X	|dx|
1	21	7
2	34	6
3	27	1
4	35	7
5	30	2
6	24	4
7	29	1
8	22	6
9	33	5
10	25	3
N=10	N= 280	42

No. = 10

$\in x = 280$

$$a = \frac{\in x}{N} = \frac{280}{10} = 28$$

$\in dx = 42$

$$M.D = \frac{\in |dx|}{N} = \frac{42}{10} = 4.2$$

$$\text{Coefficient} = \frac{M.D}{\text{Average}} = \frac{4.2}{28} = 0.15$$

Computation of M.D from Median : (Variables must be arranged in an ascending order)

S.No.	**X**	**|dx|**
1	21	7
2	22	6
3	24	1
4	25	7
5	27	2
6	29	4
7	30	1
8	33	6
9	34	5
10	35	3
N=10		42

Median (M) = $\frac{N+1}{2}$ nd item = $\frac{10+1}{2}$ nd item = 5.5th item

= 5th item + 50% of (29-27)

= 27 + 1 = 28

M.D = $\frac{\in|dx|}{N}$

= $\frac{42}{10}$

= 4.2

Coefficient = $\frac{M.D}{Average}$

= $\frac{4.2}{28}$ = 0.15

Example 10

Find out the M.D and its coefficient from Mean and Median from following data.

(x) :	3.2	6.7	4.5	9.4	8.6	6.8	1.3	0.9	4.1	2.0

Solution : Computation of Mean Deviation from Mean

S.No.	**X**	**|dx|**
1	3.2	1.55
2	6.7	1.95
3	4.5	0.25
4	9.4	4.65
5	8.6	3.85
6	6.8	2.05
7	1.3	3.45
8	0.9	3.85
9	4.1	0.65
10	2.0	2.75
N=10	47.5	25.00

$$a = \frac{\in x}{N} = \frac{47.5}{10} = 4.75$$

$$M.D = \frac{\in |dx|}{N} = \frac{25}{10} = 2.5$$

$$\text{Coefficient} = \frac{M.D}{\text{Average}} = \frac{2.5}{4.75} = 0.526$$

From Median : (Ascending order)

S.No.	X	|dx|
1	0.9	3.40
2	1.3	3.00
3	2.0	2.30
4	3.2	1.10
5	4.1	0.20
6	4.5	0.20]
7	6.7	2.40
8	6.8	2.50
9	8.6	4.80
10	9.4	5.10
N=10		24.50

Median (M) = $\frac{N+1}{2}$ nd item = $\frac{10+1}{2}$ nd item = 5.5th item

= 5th item + 50% of (4.5-4.1) = 4.1 + 0.2 = 4.3

M.D = $\frac{\in|dx|}{N} = \frac{24.50}{10} = 2.45$

Coefficient = $\frac{M.D}{Average} = \frac{2.45}{4.3} = 0.5698$

Discrete series - Mean Deviation :

M.D = $\frac{\in|fdx|}{N}$

Where MD = Mean Deviation

$\in$ |fdx| = Total of the deviations taken from the average by ignoring the signs (+and –) multiplied with the respective frequencies.

N= Total of the frequency

First of all an average (either mean or Median) must be calculated.

Arithmetic Mean (a) = $\frac{\in fdx}{N}$

Median (M) = $\frac{N+1}{2}$nd item

This item must be identified in the 'cf' and corresponding variable must be taken as Median

Then deviation must be taken from the average by ignoring plus and minus. The deviations must be multiplied with the respective frequencies (fdx). The total of this fdx ($\in$ fdx) must be devided with the total of the frequency.

Coefficient of M.D. = $\frac{M.D}{Average}$

Example 11

Find out the Mean Deviation from Mean & Median from the following data.

(x) :	21	22	23	24	25	26	27	28	29	30
(f) :	4	7	12	13	15	16	14	9	8	2

Solution : From Arithmetic Mean

X	f	dx	fdx	|dx|	|fdx|
21	4	-5	-20	4.39	17.56
22	7	-4	-28	3.39	23.73
23	12	-3	-36	2.39	28.68
24	13	-2	-26	1.39	18.07
25	15	-1	-15	0.39	5.85
26	16	0	0	0.61	9.76
27	14	+1	14	1.61	22.54
28	9	+2	18	2.61	23.49
29	8	+3	24	3.61	28.88
30	2	+4	8	4.61	9.22
	N=100		-61		187.78

$$a = x + \frac{\in fdx}{N} = 26 + \frac{-61}{100} = 26 - 0.61 = 25.39$$

$$M.D = \frac{\in | fdx |}{N} = \frac{187.78}{100} = 1.8778$$

$$\text{Coefficient} = \frac{M.D}{\text{Average}} = \frac{1.8778}{25.39} = 0.074$$

From Median :

X	f	Cf	|dx|	|fdx|
21	4	4	4	16
22	7	11	3	21
23	12	23	2	24
24	13	36	1	13
25	15	51	0	0
26	16	67	1	16
27	14	81	2	28
28	9	90	3	27
29	8	98	4	32
30	2	100	5	10
	N=100			187

$$\text{Median position} = \frac{N+1}{2}\text{nd item} = \frac{100+1}{2}\text{nd item} = 50.5\text{th item}$$

It lies in the of 'cf' 51 and the corresponding variable is 25

Then the median = 25

$$M.D = \frac{\in | fdx |}{N} = \frac{187}{100} = 1.87$$

$$\text{Coefficient} = \frac{M.D}{\text{Average}} = \frac{1.87}{25} = 0.074$$

Continuous series - Mean Deviation :

$$M.D = \frac{\in | fdx |}{N}$$

Where MD = Mean Deviation

$\in$ |fdx| = Total of the deviations taken from the average by ignoring the signs multiplied with the respective frequency

N= Total of the frequency

First of all an average (Arithermetic mean (a) or Median (M)) must be calculated. Deviations must be calculated from the average (by ignoring plus & Minus) to the other variables (|dx|). The deviations must be multiplied with respective frequencies (|f dx|). The total of this |fdx| must be devided with the total of the frequency (N)

$$\text{Arithmetic Mean (a)} = x + \frac{\in fdx}{N} \times i$$

Median (M) = $\frac{N}{2}$ nd item, this item must be identified in the 'cf' the corresponding class is to be taken as median class.

Then $M = l + \frac{c \times i}{f}$ is to be applied.

$$\text{Coefficient of M.D.} = \frac{M.D}{\text{Average}}$$

Example 12

Find out the Mean Deviation from Mean & Median and its Co-efficient.

(x) :	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
(f) :	3	6	8	13	16	18	15	12	6	3

Solution : From Arithmetic Mean

X	f	Mv	dx	fdx	|dx|	|fdx|
0-10	3	5	-5	-15	46.4	139.2
10-20	6	15	-4	-24	36.4	218.4
20-30	8	25	-3	-24	26.4	211.2
30-40	13	35	-2	-26	16.4	213.2
40-50	16	45	-1	-16	6.4	102.4
50-60	18	55	0	0	3.6	64.8
60-70	15	65	+1	15	13.6	204.0
70-80	12	75	+2	24	23.6	283.2
80-90	6	85	+3	18	33.6	201.6
90-100	3	95	+4	12	43.6	130.8
	N=100			-36		1768.8

$$a = x + \frac{\in fdx}{N} \times i$$

$$= 55 + \frac{-36}{100} \times 10$$

$$= 55 - 3.6$$

$$= 51.4$$

$$M.D = \frac{\in | fdx |}{N}$$

$$= \frac{1768.8}{100} = 17.688$$

$$\text{Coefficient} = \frac{M.D}{\text{Average}}$$

$$= \frac{17.688}{51.4} = 0.34$$

From Median :

X	f	cf	Mv	|dx|	|fdx|
0-10	3	3	5	47.22	141.66
10-20	6	9	15	37.22	223.32
20-30	8	17	25	27.22	217.76
30-40	13	30	35	17.22	223.86
40-50	16	46	45	7.22	115.52
50-60	18	64	55	2.78	50.04
60-70	15	79	65	12.78	191.70
70-80	12	91	75	22.78	273.36
80-90	6	97	85	32.78	196.68
90-100	3	100	95	42.78	128.34
	N=100				1762.24

Median position = $\frac{N}{2}$ nd item = $\frac{100}{2}$ nd item = 50th item

It lies in the 'cf' 64 and the corresponding class is 50-60

Median class = 50-60

$$M = l + \frac{c \times i}{f}$$

$$= 50 + \frac{(50-46) \times 10}{18}$$

= 50 + 2.22

= 52.22

$$M.D = \frac{\in|fdx|}{N} = \frac{1762.24}{100} = 17.6224$$

$$\text{Coefficient of MD} = \frac{M.D}{\text{Average}} = \frac{17.6224}{52.22} = 0.34$$

Example 13

Find out the M.D. and its Coefficient from mean & Median.

Marks	**No. of Students**
Less than 10	3
Less than 20	10
Less than 30	19
Less than 40	32
Less than 50	51
Less than 60	68
Less than 70	82
Less than 80	94
Less than 90	98
Less than 100	100

Solution : From Arithmetic Mean

X	**f**	**M v**	**dx**	**fdx**	**|dx|**	**|fdx|**
0-10	3	5	-4	-12	44.3	132.9
10-20	7	15	-3	-21	34.3	240.1
20-30	9	25	-2	-18	24.3	218.7
30-40	13	35	-1	-13	14.3	185.9
40-50	19	45	0	0	4.3	81.7
50-60	17	55	+1	+17	5.7	96.9
60-70	14	65	+2	+28	15.7	219.8
70-80	12	75	+3	+36	25.7	308.4
80-90	4	85	+4	+16	35.7	142.8
90-100	2	95	+5	+10	45.7	87.4
	100			43		1714.6

$$a = x + \frac{\in fdx}{N} x\ i$$

$= 45 + \frac{43}{100} \times 10$

$= 45 + 4.3$

$= 49.3$

$M.D = \frac{\in | fdx |}{N}$

$= \frac{1714.6}{100} = 17.146$

$\text{Coefficient} = \frac{M.D}{\text{Average}}$

$= \frac{17.146}{49.3} = 0.3477$

From Median :

X	f	cf	M v	|dx|	|fdx|
0-10	3	3	5	44.47	133.41
10-20	7	10	15	34.47	241.29
20-30	9	19	25	24.47	220.23
30-40	13	32	35	14.47	188.11
40-50	19	51	45	4.47	84.93
50-60	17	68	55	5.53	94.01
60-70	14	82	65	15.53	217.42
70-80	12	94	75	25.53	306.36
80-90	4	98	85	35.53	142.12
90-100	2	100	95	45.53	91.06
	N=100				1718.94

Median position = $\frac{N}{2}$ nd item = $\frac{100}{2}$ nd item = 50th item

It lies in the 'cf' 51 and the corresponding class is 40-50

Median class = 40-50

$$M = l + \frac{c \times i}{f}$$

$$= 40 + \frac{(50 - 32) \times 10}{19}$$

$$= 40 + \frac{186}{19}$$

= 40 + 9.473

= 49.473

$$M.D = \frac{\in|fdx|}{N} = \frac{1718.94}{100} = 17.1894$$

$$\text{Coefficient} = \frac{M.D}{\text{Average}} = \frac{17.1894}{49.473} = 0.3474$$

10. 12 MERITS OF MEAN DEVIATION :

1. It is rigidly defined easy to compute and understand.
2. It takes all the items into consideration and gives weight to deviation according to their size
3. It is less affected by extreme values of variables
4. It removes all the irregularities by obtaining deviation and provides a correct measure.

10. 13 DEMERITS OF MEAN DEVIATION :

1. It does not lend itself readily to algebraic treatment
2. It ignores the negative deviation and treats them as positive which is not justified mathematically
3. It is rarely used in social sciences
4. It is not suitable when the class intervals are open end.

10.14 SUMMARY :

Range is the difference between the Highest value and least value. But it is not a stable measure and has many limitations such as fluctuations of sampling . Quartile Deviation is better than Range. But here all the items are not taken into account. It also suffers from sampling instability. Mean Deviation is better than quartile Deviation. But it is also not capable of further Algebraic treatment, although it takes into account all the terms but still if the extreme values are big, it will desort the result. More over it ignores + signs.

10.15 QUESTIONS :

1. What is meant by measures of dispersion ?
2. What are the differences between the Measures of central tendency and Measures of Dispersion ?
3. What are the objectives of measuring the second order average ?
4. Define variation or Dispersion or scatterdness
5. Name various methods of measuring dispersion.
6. Define Range . Is it positional measure ? How ?
7. What is coefficent of Range ? Narrate the formula ?
8. What are the merits and demerits of Range ?
9. Define Semi - inter Quartile Range.
10. What are the objectives of computing the Quartile Deviation ?
11. What are the merits and demerits of Quartile deviation ?
12. Define Mean Deviation or Average Deviation
13. What is meant by Coefficient of Mean Deviation ?
14. Explain the method of calculation of M.D from Mean
15. Mean Deviation is free from all the short comings or Range Q.D. and hence is a super measure of variation. Discuss.
16. What are the merits and demerits of Mean Devaition ?

10.16 EXERCISES :

1. Find out of the Range and coefficient of Range from the following data.

   wages (Rs) = 27, 31, 32, 28, 40,39,37,34,30,29

2. The earnings of a worker in a week were as under. Find out Range & its co efficient .

Earnings (Rs) = 26,35,41,45,32,29.

3. Find the range and coefficient of Range for following data.

(x) :	5	10	15	20	25	30	35	40
(f) :	4	7	21	47	53	24	12	6

4. Calculate Range and its coefficient from the following data

(x) :	20-30	30-40	40-50	50-60	60-70	70-80
(f) :	4	9	16	21	13	6

5. Compute Quartile Deviation and its coefficient from the following data.

Marks of 11 students : 75, 43, 86, 21, 12, 3, 35, 57, 67, 94, 60

6. Calculate Quartile Deviation and its Coefficient from the data given below :

Wages 12 workess : 91, 98, 99 , 90, 89, 94, 93, 97, 96, 94, 95, 92

7. From the following data compute Q.D. and its coefficient

(x) :	81	82	83	84	85	86	87	88	89	90
(f) :	8	16	26	32	36	28	22	18	10	4

8. From the following data compute Q.D. and its coefficient

(x) :	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
(f) :	6	9	13	18	21	19	17	16	12	8

9. Find out Q.D and its coefficient from the following data

(x) :	0-2	2-4	4-6	6-8	8-10	10-12	12-14	14-16	16-18	18-20
(f) :	7	13	18	23	22	20	16	15	14	11

10. Find out Q.D and its coefficient from the following data

(x) :	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
(f) :	5	9	13	19	22	23	21	18	17	12

11. Find Q.D and its coefficient

Mid values of classes -	115	125	135	145	155	165	175	185	195	205
Frequencey -	6	12	13	21	25	24	23	21	19	7

12. Find Q.D and its coefficient

Mid values of classes -

2.5	7.5	12.5	17.5	22.5	27.5	32.5	37.5	42.5	47.5

Frequencey -

6	9	13	18	17	16	12	8	6	5

13. Find out Q.D and its coefficient

Marks	**No. of Students**
Less than 10	4
Less than 20	11
Less than 30	24
Less than 40	41
Less than 50	57
Less than 60	71
Less than 70	83
Less than 80	92
Less than 90	97
Less than 100	100

14. Find out Mean Deviation and it's coefficient from Mean and Median.

x -	75	64	79	67	70	61	68	82	63	71

15. Find out Mean Deviation at it's coefficient from Mean and Median.

x -12.6, 13.9,19.8, 14.7, 11.5,17.3, 16.2,10.1, 15.4,18.0, 16.7

16. Find Q.D and its coefficient

x -	5	10	15	20	25	30	35	40	45	50
f -	3	8	13	19	16	14	12	9	4	2

17. Find out mean Deviation from Median.

x -	11	12	13	14	15	16	17	18	19	20
f -	5	8	13	16	18	14	12	8	4	2

18. Find out Mean Deviation from Mean & Median Also find out the coefficent.

(x) :	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
(f) :	13	18	19	23	26	24	22	21	19	15

19. Find out Mean Deviation and its coefficient from Mean & Median.

(x) :	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
(f) :	5	9	12	13	14	16	13	9	7	2

20. Find out Mean Deviation and its coefficeint from the following data.

Marks	**No. of Students**
Above 0	100
Above 10	97
Above 20	89
Above 30	76
Above 40	60
Above 50	42
Above 60	28
Above 70	16
Above 80	7
Above 90	3

21. Find out Mean Deviation and its coefficient

(x) :	0-99	100-199	200-299	300-399	400-499
(f) :	2	8	13	15	18

(x) :	500-599	600-699	700-799	800-899	900-999
(f) :	14	11	9	7	3

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson - 11

MEASURES OF DISPERSION - STANDARD DEVIATION

OBJECTIVES:

By the study of this lesson, you will be able to understand the meaning and methods of computation of standard deviation, Coefficient of variation in individual , Discrete & Continuous series with examples.

STRUCTURE :

11.1. INTRODUCTION :

" Standard deviation " is the root of the sum of the squares of the deviations devided by their number. It is also called Mean " Error Deviation " " Mean square Error Deviation " or " Root Mean Square Deviation ". It is Second moment of a dispersion. Since the sum of the squares of the deviations from the Mean is minimum, the deviations are taken only from mean (but not from median or mode).

Standard Deviation is the root - mean - square average of all the deviations from the mean. It is proposed by " Prof - karl pearson " in 1893 and it is denoted by ' σ '(Sigma)

11.2. COMPUTATION OF STANDARD DEVIATION :

Individual Series :

$$\sigma = \sqrt{\frac{\in dx^2}{N} - \left(\frac{\in dx}{N}\right)^2}$$

Where

σ = Standard Deviation

$\in dx$ = Total of the deviations taken from the assumed mean

$\in dx^2$ = Total of the squares of the deviations taken from the assumed mean

N = Number of variables.

11.3. CO - EFFICIENT OF VARIATION :

It is the relative measure of dispersion in which the variation is expressed in percentage. It is often used to have the comparative study of the dispersion of two or more series in the same or different units. It is the percentage variation in the mean, where as the standard deviation is the total variation in the mean. This relative measure of dispersion implies the ratio of standard deviation to the mean signifying the percentage.

Co efficient of variation = $\frac{\sigma}{a}$x100

Where

σ = Standard deviation

a = Arithmetic mean $\left(a = x + \frac{\in dx}{N}\right)$

It is helpful in knowing the consistency of items of the series. If the value so arrived is greater (more thean 50%) the result signifies the lower degree of consistency. If the value so arrived is smaller (Less than 50%) the result signifies upper degree of consistency.

11.4. EXAMPLES :

Example 1 : Find out standard Deviation and coefficient of variation

x : 24 31 27 25 28 20 29 23 22 30

Solution :

X	dx	dx^2
24	-3	9
31	+4	16
27	0	0
25	-2	4
28	+1	1
20	-7	49
29	+2	4
23	-4	16
22	-5	25
30	+3	9
N=10	-11	133

$$\sigma = \sqrt{\frac{\in dx^2}{N} - \left(\frac{\in dx}{N}\right)^2}$$

$$= \sqrt{\frac{133}{10} - \left(\frac{-11}{10}\right)^2}$$

$$= \sqrt{13.3 - 1.21}$$

$$= \sqrt{12.09} = 3.476$$

$$a = x + \frac{\in dx}{N}$$

$$= 27 + \frac{-11}{10} = 27 - 1.1 = 25.9$$

Coefficient of variation = $\frac{\sigma}{a}$ x100

$$= \frac{3.476}{25.9} \text{x}100 = 13.42\ \%$$

Example 2 : The prices shares of 2 companies were as under. Which is more variable ?

Company A :	12	15	21	16	9	13	10	17	14	21	11	8
Company B :	107	109	100	111	97	93	96	104	101	108	106	105

Solution :

X	**dx**	**dx²**	**y**	**dy**	**dy²**
12	-3	9	107	+7	49
15	0	0	109	+9	81
21	+6	36	100	0	0
16	+1	1	111	+11	121
9	-6	36	97	-3	9
13	-2	4	93	-7	49
10	-5	25	96	-4	16
17	+2	4	104	+4	16
14	-1	1	101	+1	1
21	+6	36	108	+8	64
11	-4	16	106	+6	36
8	-7	49	105	+5	25
N=12	-13	217	N12	37	407

X - Series :

$$\sigma = \sqrt{\frac{\in dx^2}{N} - \left(\frac{\in dx}{N}\right)^2}$$

$$= \sqrt{\frac{217}{12} - \left(\frac{-13}{12}\right)^2}$$

$$= \sqrt{18.08 - (-1.08)^2}$$

$$= \sqrt{18.08 - 1.1664} = \sqrt{16.9136} = 4.1126$$

$$a = x + \frac{\in dx}{N}$$

$$= 15 + \frac{-13}{12} = 15 - 1.08 = 13.92$$

$$\text{Coefficient of variation} = \frac{\sigma}{a} \times 100$$

$$= \frac{4.1126}{13.92} \times 100 = 29.54\ \%$$

Y - Series :

$$\sigma = \sqrt{\frac{\in dy^2}{N} - \left(\frac{\in dy}{N}\right)^2}$$

$$= \sqrt{\frac{467}{12} - \left(\frac{-37}{12}\right)^2} = \sqrt{38.917 - (3.083)^2}$$

$$= \sqrt{38.917 - 9.505}$$

$$= \sqrt{29.412} = 5.423$$

$$a = y + \frac{\in dy}{N}$$

$$= 100 + \frac{37}{12} = 100 + 3.083$$

$$= 103.083$$

$$\text{Coefficient of variation} = \frac{\sigma}{a} \times 100$$

$$= \frac{5.423}{103.083} \times 100 = 5.2608\ \%$$

The coefficient of variation is more in companyA. There fore it can be said that the prices of shares of company A are more variable.

Example 3 : The marks secured by two students A&B in 10 examinations were as under. Find out who is more clever ? (or Find out who is more consistent)

Mark A :	42	70	36	30	48	45	34	50	60	25
Mark B :	55	95	42	20	60	50	48	70	80	10

Solution :

X	dx	dx^2	y	dy	dy^2
42	-3	9	55	+5	25
70	+25	625	95	+45	2025
36	-9	81	42	-8	64
30	-15	225	20	-30	900
48	+3	9	60	+10	100
45	0	0	50	0	0
34	-11	121	48	-2	4
50	+5	25	70	+20	400
60	+15	225	80	+30	900
25	-20	400	10	-40	1600
N=10	-10	1720	N=10	30	6018

X - Series :

$$\sigma = \sqrt{\frac{\in dx^2}{N} - \left(\frac{\in dx}{N}\right)^2}$$

$$= \sqrt{\frac{1720}{10} - \left(\frac{-10}{10}\right)^2}$$

$$= \sqrt{172 - 1} = \sqrt{171} = 13.08$$

$$a = x + \frac{\in dx}{N}$$

$$= 45 + \frac{-10}{10} = 45 - 1 = 44$$

Coefficient of variation = $\frac{\sigma}{a}$ x100

$= \frac{13.08}{44}$ x100

= 29.727 %

Y - Series :

$$\sigma = \sqrt{\frac{\in dy^2}{N} - \left(\frac{\in dy}{N}\right)^2}$$

$$= \sqrt{\frac{6018}{10} - \left(\frac{30}{10}\right)^2}$$

$$= \sqrt{601.8 - 9}$$

$$= \sqrt{592.8}$$

= 24.347

$$a = y + \frac{\in dy}{N}$$

$= 50 + \frac{30}{10}$ = 50 + 3

= 53

Coefficient of variation = $\frac{\sigma}{a}$ x100

$= \frac{24.347}{53}$ x100

= 45.94

The coefficient of variation in series x is smaller than Series Y. Therefore it can be concluded that Mr A (x series) is cleverer than Mr. B (Y series) or A is more consistant.

Discrete Series :

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \text{ x i}$$

Where

σ = Standard Deviation

$\in$ fdx = Total of the deviations taken from the assumed mean, multiplied with the respective frequencies.

$\in$ fdx^2 = Total of the squares of the deviations taken from the assumed mean, multiplied with the respective frequencies.

N = Number of variables.

i = Interval (common factor)

$$\text{Coefficient of variation} = \frac{\sigma}{a} \text{x100}$$

$$a = x + \frac{\in fdx}{N} \text{ x i}$$

Example 4 : Compute standard Deviation and its coefficient of variation from the following data.

x:	0	1	2	3	4	5	6	7	8	9	10
f :	3	5	8	11	13	16	14	12	9	6	3

Solution :

X	**f**	**dx**	**fdx**	**fdx^2**
0	3	-5	-15	75
1	5	-4	-20	80
2	8	-3	-24	72
3	11	-2	-22	44
4	13	-1	-13	13
5	16	0	0	0
6	14	+1	14	14
7	12	+2	24	48
8	9	+3	27	81
9	6	+4	24	96
10	3	+5	15	75
	N=100		10	598

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2}$$

$$= \sqrt{\frac{598}{100} - \left(\frac{10}{100}\right)^2}$$

$$= \sqrt{5.98 - 0.01}$$

$$= \sqrt{5.97} = 2.443$$

$$a = x + \frac{\in fdx}{N}$$

$$= 5 + \frac{10}{100} = 5.1$$

Coefficient of variation = $\frac{\sigma}{a}$ x100

$$= \frac{2.443}{5.1} \text{x}100$$

= 47.90%

Example 5 : Compute standard Deviation and its coefficient of variation

x:	5	10	15	20	25	30	35	40	45	50
f :	3	7	11	13	17	16	12	9	8	4

Solution :

X	f	dx	fdx	fdx²
5	3	-4	-12	48
10	7	-3	-21	63
15	11	-2	-22	44
20	13	-1	-13	13
25	17	0	0	0
30	16	+1	16	16
35	12	+2	24	48
40	9	+3	27	81
45	8	+4	32	128
50	4	+5	20	100
	N=100		51	541

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \text{ x i}$$

$$= \sqrt{\frac{541}{100} - \left(\frac{51}{100}\right)^2} \text{ x i}$$

$$= \sqrt{5.41 - (0.51)^2} \text{ x i}$$

$$= \sqrt{5.41 - 0.2601} \text{ x i}$$

$$= \sqrt{5.1499} \text{ x i} = 2.27 \text{ x } 5$$

$$= 11.35$$

$$a = X + \frac{\in fdx}{N} \text{ x i}$$

$$= 25 + \frac{51}{100} \text{ x } 5$$

$= 25 + \frac{255}{100}$ $= 25 + 2.55$ = 27.55

Coefficient of variation = $\frac{\sigma}{a}$ x100

= $\frac{11.35}{27.55}$ x100

= 41.2%

Continuous Series :

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \text{ x i}$$

Where

σ = Standard Deviation

$\in$ fdx = Total of the deviations taken from the assumed mean, multiplied with the respective frequencies.

$\in$ fdx^2 = Total of the squares of the deviations taken from the assumed mean, multiplied with the respective frequencies.

N = Number of variables.

i = Interval (common factor)

Note : In continuous series the classes must be converted into Mid values, assumed mean shall be taken from the Midvalues and the deviation shall be taken from the assumed mean to the other Mid values.

Coefficient of variation = $\frac{\sigma}{a}$ x100

a = $x + \frac{\in fdx}{N}$ x i

Example 6 : Compute standard Deviation and its coefficient

(x) :	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
(f) :	3	6	8	13	16	15	14	12	9	4

Solution : From Arithmetic Mean

X	f	Mv	dx	fdx	fdx^2
0-10	3	5	-4	-12	48
10-20	6	15	-3	-18	54
20-30	8	25	-2	-16	32
30-40	13	35	-1	-13	13
40-50	16	45	0	0	0
50-60	15	55	+1	+15	15
60-70	14	65	+2	+28	56
70-80	12	75	+3	+36	108
80-90	9	85	+4	+36	144
90-100	4	95	+5	+20	100
	N=100			76	570

X - Series :

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \text{ x i}$$

$$= \sqrt{\frac{570}{100} - \left(\frac{76}{100}\right)^2} \text{ x i}$$

$$= \sqrt{5.7 - 0.5776} \text{ x i}$$

$$= \sqrt{5.1224} \text{ x i}$$

= 2.2632 x 10 = 22.632

$$a = \frac{\sigma}{a} \text{x100}$$

$$= \frac{22.632}{52.6} \text{x100}$$

= 43.03%

Example 7 : Find out the standard Deviation and its coefficient of variation.

Marks	No. of Students
More than 0	100
More than10	97
More than 20	89
More than 30	77
More than 40	64
More than 50	57
More than 60	42
More than 70	28
More than 80	17
More than 90	5

Solution : From Arithmetic Mean

X	f	Mv	dx	fdx	fdx^2
0-10	3	5	-5	-15	75
10-20	8	15	-4	-32	128
20-30	12	25	-3	-36	108
30-40	13	35	-2	-26	52
40-50	7	45	-1	-7	7
50-60	15	55	0	0	0
60-70	14	65	+1	+14	14
70-80	11	75	+2	+22	44
80-90	12	85	+3	+36	108
90-100	5	95	+4	+20	80
	N=100			-24	616

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \text{ x i}$$

$$= \sqrt{\frac{616}{100} - \left(\frac{-24}{100}\right)^2} \text{ x i}$$

$$= \sqrt{6.16 - (-0.24)^2} \text{ x i}$$

$$= \sqrt{6.16 - 0.0576} \text{ x i}$$

$$= \sqrt{6.1024} \text{ x i}$$

= 2.47 x 10 = 24.7

$$a = x + \frac{\in fdx}{N} \text{ x i}$$

$$= 55 + \frac{-24}{100} \text{ x 10}$$

= 55 - 2.4 = 52.6

Coefficient of variation = $\frac{\sigma}{a}$ x100

$$= \frac{24.7}{52.6} \text{x100}$$

= 46.96%

Example 8 : Which of the following two series is more consistent in value ?

x :	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
f_1 :	13	17	19	23	27	25	22	21	19	14
f_2 :	15	16	21	24	28	25	23	18	16	14

Solution :

f_1

X	f	Mv	dx	fdx	fdx^2
0-5	13	2.5	-4	-52	208
5-10	17	7.5	-3	-51	153
10-15	19	12.5	-2	-38	76
15-20	23	17.5	-1	-23	23
20-25	27	22.5	0	0	0
25-30	25	27.5	+1	+25	25
30-35	22	32.5	+2	+44	88
35-40	21	37.5	+3	+63	189
40-45	19	42.5	+4	+76	304
45-50	14	47.5	+5	+70	350
	N=200			114	1416

f_2

X	f	Mv	dx	fdx	fdx^2
0-5	15	2.5	-4	-60	240
5-10	16	7.5	-3	-48	144
10-15	21	12.5	-2	-42	84
15-20	24	17.5	-1	-24	24
20-25	28	22.5	0	0	0
25-30	25	27.5	+1	+25	25
30-35	23	32.5	+2	+46	92
35-40	18	37.5	+3	+54	162
40-45	16	42.5	+4	+64	256
45-50	14	47.5	+5	+70	350
	N=200			85	1377

f_1- Series :

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \ x \ i$$

$$= \sqrt{\frac{1416}{200} - \left(\frac{114}{200}\right)^2} \ x \ i$$

$$= \sqrt{7.08 - 0.3249} \ x \ i$$

$$= \sqrt{6.7751} \ x \ i$$

= 2.599 x 5 = 12.995

$$a = x + \frac{\in fdx}{N} \ x \ i$$

$$= 22.5 + \frac{114}{200} \ x \ 5$$

$$= 22.5 + \frac{570}{200} \quad = 22.5 + 2.85 = 25.35$$

$$\text{Coefficient of variation} = \frac{\sigma}{a} x100$$

$$= \frac{12.995}{25.35} x100$$

= 51.26 %

f_2- Series :

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \ x \ i$$

$$= \sqrt{\frac{1377}{200} - \left(\frac{85}{200}\right)^2} \ x \ i$$

$$= \sqrt{6.885 - 0.1806} \ x \ i$$

$= \sqrt{6.7644}$ x i

= 2.5892 x 5 = 12.946

$a = x + \frac{\in fdx}{N}$ x i

$= 22.5 + \frac{85}{200}$ x 5

$= 22.5 + \frac{425}{200}$ =22.5 + 2.125= 24.625

Coefficient of variation = $\frac{\sigma}{a}$ x100

$= \frac{12.946}{24.625}$ x100 = 52.57 %

The coefficient of variation is small in f_1 series. Therefore, it can be said that the series f_1 are more consistant in value.

Example 9 : The profits and losses of 100 companines in an industry were as under.

Find out the standard Deviation and its coefficient of variation.

Profits & Losses	**No. of companies**
4000-5000	5
3000-4000	9
2000-3000	12
1000-2000	13
0-1000	17
-1000 - 0	16
-2000- -1000	14
-3000- -2000	8
-4000- -3000	4
-5000- -4000	2
	N=100

Solution :

X	f	M v	dx	fdx	fdx²
4000-5000	5	4500	+4	+20	80
3000-4000	9	3500	+3	+27	81
2000-3000	12	2500	+2	+24	48
1000-2000	13	1500	+1	+13	13
0-1000	17	500	0	0	0
-1000 - 0	16	-500	-1	-16	16
-2000- -1000	14	-1500	-2	-28	56
-3000- -2000	8	-2500	-3	-24	72
-4000- -3000	4	-3500	-4	-16	64
-5000- -4000	2	-4500	-5	-10	50
	N=100			-10	480

$$\sigma = \sqrt{\frac{\in fdx^2}{N} - \left(\frac{\in fdx}{N}\right)^2} \text{ x i}$$

$$= \sqrt{\frac{480}{100} - \left(\frac{-10}{100}\right)^2} \text{ x i}$$

$$= \sqrt{4.8 - 0.01} \text{ x i}$$

$$= \sqrt{4.79} \text{ x i}$$

= 2.1886061 x 1000 = 2188.61/-

$$a = x + \frac{\in fdx}{N} \text{ x i}$$

$$= 500 + \frac{-10}{100} \text{ x 1000}$$

= 500 – 100 = 400/-

$$\text{Coefficient of variation} = \frac{\sigma}{a} \text{x100}$$

$$= \frac{2188.61}{400} x100$$

= 547.15%

11.5 MERITS OF STANDARD DEVIATION :

a) It is based on all the observations given

b) It can be smoothly handled algebraically

c) It is a well defined and definite measure of dispersion.

d) It is of great importance when the comparison is made between variability of two item.

11.6 DEMERITS OF STANDARD DEVIATION :

a) It is difficult to calculate and understand

b) It gives more weight to extreme values as the deviations are squared.

c) It is not useful in economic studies.

11.7. SUMMARY:

Standard Deviation and Coefficient of variation possess all those properties, which a good measure of disperssion should possess. The process of squaring the deviations eliminates the negative signs and thus makes the mathematical manipulation of figures earh.

11.8. QUESTIONS:

1. Define standard Deviation ?
2. What is meant by Standard Deviation ?
3. What is meant by Coefficient of Variation ?
4. What are the merits and Demerits of Standard Deviation ?
5. Why the Standard Deviation is better than the other measures of dispersion ?

11.9. EXERCISES :

1. Find out Standard Deviation and its coefficient of variation .

(x) :	24	27	23	30	25	29	21	27	26	22

2. Find out Standard Deviation and its coefficient of variation .

(x) :	57	58	52	56	60	55	51	54	53	59

3. Compute Standard Deviation and its coefficient of variation .

(x) :	85	94	93	90	96	99	98	91	87	86

4. Calculate Standard Deviation and its coefficient of variation .

(x) : 345, 352, 341, 350, 355, 357, 354, 344, 348, 349, 341, 346,

5. The following 2 series were given to you : Which is more consistent in value. ?

(x) :	75	49	56	64	70	65	67	73	58	74
(f) :	160	170	173	164	167	161	175	177	172	169

6. The runs scored by two batsmen in 10 oneday matches were as under. Who is more consistent ?

a :	3	87	64	1	12	76	0	50	60	85	96	24
b :	75	86	63	47	55	60	49	21	13	70	49	2

7. The prices of the shares of two companies during last 12 months were as under. Which share is more variable in price.

Company x :	5	12	17	14	13	10	6	8	13	15	16	9
Company y :	45	54	49	50	51	53	42	47	48	55	54	52

8. Compute Standard Deviation and its coefficient of variation

x :	41	42	43	44	45	46	47	48	49	50
f :	5	9	13	17	18	16	14	7	8	3

9. Compute Standard Deviation and its coefficient

x :	2	4	6	8	10	12	14	16	18	20
f :	3	7	11	13	17	16	12	9	8	4

10. Compute Standard Deviation and its coefficient

(x) :	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
(f) :	4	7	8	12	18	17	13	11	6	4

11. Find out Standard Deviation and its coefficient

(x) :	20-24	24-28	28-32	32-36	36-40	40-44	44-48	48-52	52-56
(f) :	8	12	13	15	18	14	9	7	4

12. Compute Standard Deviation and its coefficient of variation

(x) :	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
(f) :	4	8	12	13	15	14	12	9	8	5

13. Compute Standard Deviation and its coefficient of variation

(x) :	0-99	100-199	200-299	300-399	400-499
(f) :	5	7	12	13	16
(x) :	500-599	600-699	700-799	800-899	900-999
(f) :	17	14	9	5	2

14. Find out Standard Deviation and its coefficient

(x) :	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
(f) :	13	17	19	23	27	25	22	21	19	14

15. Compute Standard Deviation and its coefficient of variation

Marks	**No. of Students**
Less than 10	4
Less than 20	11
Less than 30	20
Less than 40	33
Less than 50	50
Less than 60	66
Less than 70	81
Less than 80	93
Less than 90	98
Less than 100	100

16. Calculate Standard Deviation and its coefficient of variation

Marks	No. of Students
Less than 100	100
Less than 90	97
Less than 80	90
Less than 70	77
Less than 60	60
Less than 50	44
Less than 40	30
Less than 30	18
Less than 20	9
Less than 10	4

17. The profits and losses of 100 companines in an industry were as under. Find out the standard Deviation and its coefficient of variation.

Profits & Losses	No. of Companies
4000-5000	2
3000-4000	7
2000-3000	9
1000-2000	13
0-1000	16
-1000 - 0	17
-2000- -1000	14
-3000- -2000	12
-4000- -3000	7
-5000- -4000	3
	N=100

18. Find out the standard Deviation and its coefficient of variation.

x	f
500-600	20
400-500	30
300-400	60
200-300	20
100-200	10
0-100	8
-100 - 0	12
-200- -100	16
-300- -200	20
	196

19. Compute standard Deviation and its coefficient of variation.

x	f
-40 – -30	10
-30 – -20	28
-20 – -10	30
-10 – 0	42
0-10	65
10 –20	180
20 –30	10
	315

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson - 12

CO EFFICIENT OF CORRELATION

OBJECTIVES:

By the study of this lesson, you will be able to understand the meaning, definition of Karl Pearson's Coefficient of Correlation, probable error, uses of correlation and method of computation of coefficient of correlation.

STRUCTURE:

12.1 INTRODUCTION:

" Correlation " means a possible connection or relationship or interdependence between the values of two or more variables of the same phenomenon or individual series. It indicates the strength of the relationship. If we measure the heights and weights of 'n' individuals we assume two values - one relating to heights and the other relating to weights. Such distributions, in which each unit of the series assumes two values are called "Bivariate Distributions". If there are more than two variables in each unit such distributions are called " Multivariate Distributions " .

We can establish the relationship between the two or more values of the same series for the purpose of comparative study. Such a relationship can be established logically with some beliefs or assumptions or notions. It is purely a guess work. It does not relate to the establishment or cause and effect. However, there may or may not be the factor or causation. There may be third group of influencing factors of the changes in the values of variables. Thus sometimes, the existence of relationship is just purely a chance or accidental event.

12.2 TYPES OF CORRELATION :

Correlation is classified, into the following ways.

a) **Positive Correlation :** If the values of the two variables deviate in the same direction, it is said to be positive or Direct correlation.

b) **Negative Correlation :** When the values of two variables deviate in the opposite direction, it is said to be " Negative " or " Indirect " correlation.

c) **Partial Correlation :** When one variable is independent and the other variable is dependent on the former it is said to be " Partial correlation ".

d) **Simple Correlation :** When only two variables are studied, it is called " Simple Correlation ". It means the study involves only two variables which are changing either in the same or opposite direction.

e) **Multiple Correlation :** When three or more variables are studied, it is called a "Multiple Correlation ". The variables may change in the same direction or in different direction.

f) **Linear Correlation :** If for corresponding to a unit change in one variable there is a constant change in the other variable over the entire range of the values it is said to be a " Linear correlation ".

g) **Non - linear Correlation :** If the variables under study are graphed and the plotted points donot form a straight line. It is said to be a " Non- Linear correlation " or "Curvi- Linear correlation " The amount of change in one variable does not bear a constant change in the other variable.

12.3 METHOD OF COMPUTING CORRELATION :

Karl Pearson's Coefficient of correlation :

Karl Pearson (1807 - 1936) a great British Bio - metrician and statistician has propounded the formula for calculating the coefficient of correlation. The formula is based on arithmetic mean and Standard Deviation and it is most widely used.

The formula indicates whether the correlation is positive or negative. The answer lies between +1 and -1 (Perfect positive and Negative correlation respectively). Zero represents the absence of correlation. The formula is subject to algebraic manipulations and it is based on covariance is a highly useful concept in the statistical analysis. Karl pearson's coefficient of correlation is also known as the " Product Moment Coefficient ". It is denoted by ' γ '. It is a measure of association.

Karl Pearson's coefficient of correlation =

$$\gamma = \frac{\in dx\ dy \times N - (\in dx . \in dy)}{\sqrt{\in dx^2 \times N - (\in dx)^2} \times \sqrt{\in dy^2 \times N - (\in dy)^2}}$$

Where

γ = Coefficient of Correlation

$\in dx$ = Total of the deviations taken from the assumed mean in ' x ' series.

$\in dy$= Total of the deviations taken from the assumed mean in ' y' series.

$\in dx^2$ = Total of the squares of the deviations taken from the assumed mean in 'x'series

$\in dy^2$ = Total of the squares of the deviations taken from the assumed mean in 'Y'series

$\in dxdy$ = Total of deviations in x & y series multiplied by each other

N = Number of pairs

12.4 PROBABLE ERROR :

It is a difference resulting due to taking samples from the mass or population. According to "Secrist " the probable error of the correlation coefficient is an amount, which if added to and subtracted from the average correlation coefficient, produces amounts with in which the chances are even that a coefficient of correlations from a series selected at random will fall.

With the help of probable error, it is possible to determine the reliability of the value of the coefficient in so far as it depends on the conditions of random sampling. It is an old measure of testing the reliability of an observed value of correlation coefficient. It is based on the standard errors multiplied by the probable error. It is obtained by the formula.

Probable error P.E = $0.6745 \left(\frac{l - r^2}{\sqrt{N}} \right)$

12.5 EXAMPLES :

Example 1

From the following data compute Karl Pearson's coefficient of correlation

Wages :	100	101	102	102	100	99	97	98	96	95
Cost of living :	98	99	99	97	95	92	95	94	90	91

Solution :

X	dx	dx^2	y	dy	dy^2	dxdy
100	0	0	98	+3	9	0
101	+1	1	99	+4	16	+4
102	+2	4	99	+4	16	+8
102	+2	4	97	+2	4	+4
100	0	0	95	0	0	0
99	-1	1	92	-3	9	+3
97	-3	9	95	0	0	0
98	-2	4	94	-1	1	+2
96	-4	16	90	-5	25	+20
95	-5	25	91	-4	16	+20
N=10	-10	64	N=10	0	96	61

$$\gamma = \frac{\in dx\ dy \times N - (\in dx \,.\, \in dy)}{\sqrt{\in dx^2 \times N - (\in dx)^2} \times \sqrt{\in dy^2 \times N - (\in dy)^2}}$$

$$= \frac{61x10 - (-10x0)}{\sqrt{64x10 - (-10)^2} \, x \sqrt{96x10 - (0)^2}}$$

$$= \frac{610 - 0}{\sqrt{640 - 100} \, x \sqrt{960}}$$

$$= \frac{610}{\sqrt{540} \, x \sqrt{960}}$$

$$= \frac{610}{23.2379 \times 30.9838}$$

$$= \frac{610}{719.99}$$

= + 0.8472 Positive.

Example 2

Compute Karl Pearson's coefficient of correlation from the following data.

x :	27	21	35	44	29	30	32	42	41	36	28	26
y :	40	37	21	25	36	41	22	31	23	24	39	37

Solution :

X	dx	dx²	y	dy	dy²	dxdy
27	-3	9	40	+9	81	-27
21	-9	81	37	+6	36	-54
35	+5	25	21	-10	100	-50
44	+14	196	25	-6	36	-84
29	-1	1	36	-5	25	-5
30	0	0	41	+10	100	0
32	+2	4	22	-9	81	-18
42	+12	144	31	0	0	0
41	+11	121	23	-8	64	-88
36	+6	36	24	-7	49	-42
28	-2	4	39	+8	64	-16
26	4	16	37	+6	36	-24
N=12	31	637	N=12	4	672	-408

$$\gamma = \frac{\in dx\ dy \times N - (\in dx\ .\ \in dy)}{\sqrt{\in dx^2 \times N - (\in dx)^2} \times \sqrt{\in dy^2 \times N - (\in dy)^2}}$$

$$= \frac{-408 \times 12 - (31 \times 4)}{\sqrt{637 \times 12 - (31)^2} \times \sqrt{672 \times 12 - (4)^2}}$$

$$= \frac{-4896 - 124}{\sqrt{7644 - 961} \times \sqrt{8064 - 16}}$$

$$= \frac{-5020}{81.7496 \times 89.7106} = \frac{-5020}{7333.81} = -0.6845 \text{ Negative.}$$

Example 3

Compute Karl Pearson's coefficient of correlation from the following data.

x : 300 350 400 450 500 550 600 650 700

y : 800 900 1000 1100 1200 1300 1400 1500 1600

Solution :

X	dx	dx^2	y	dy	dy^2	dxdy
300	-4	16	800	-4	16	16
350	-3	9	900	-3	9	9
400	-2	4	1000	-2	4	4
450	-1	1	1100	1	1	1
500	0	0	1200	0	0	0
550	+1	1	1300	+1	1	1
600	+2	4	1400	+2	4	4
650	+3	9	1500	+3	9	9
700	+4	16	1600	+4	16	16
N=9	0	60	N=9	0	60	60

$$\gamma = \frac{\in dx\ dy \times N - (\in dx\ .\ \in dy)}{\sqrt{\in dx^2 x\ N - (\in dx)^2}\ x \sqrt{\in dy^2\ x\ N - (\in dy)^2}}$$

$$= \frac{60 \times 9 - (0 \times 0)}{\sqrt{60 \times 9 - (0)^2}\ x \sqrt{60 \times 9 - (0)^2}}$$

$$= \frac{540}{\sqrt{540}\ x \sqrt{540}}$$

$$= \frac{540}{540}$$

= + 1 Positive

Example 4

Compute Karl Pearson's coefficient of correlation

x :	20	40	60	80	100	120	140	160	180	200
y :	2000	1990	1980	1970	1960	1950	1940	1930	1920	1910

Solution :

X	dx	dx^2	y	dy	dy^2	dxdy
20	-4	16	2000	+5	25	-20
40	-3	9	1990	+4	16	-12
60	-2	4	1980	+3	9	-6
80	-1	1	1970	+2	4	-2
100	0	0	1960	+1	1	0
120	+1	1	1950	0	0	0
140	+2	4	1940	-1	1	-2
160	+3	9	1930	-2	4	-6
180	+4	16	1920	-3	9	-12
200	+5	25	1910	-4	16	-20
N=10	5	85	N=10	5	85	-80

$$\gamma = \frac{\in dx\ dy\ x\ N - (\ \in dx\ .\ \in dy)}{\sqrt{\in dx^2 x\ N - (\in dx)^2}\ x \sqrt{\in dy^2\ x\ N - (\in dy)^2}}$$

$$= \frac{-80\ x\ 10 - (5\ x\ 5)}{\sqrt{85\ x\ 10 - (5)^2}\ x \sqrt{85\ x\ 10 - (5)^2}}$$

$$= \frac{-800\ \ -25}{\sqrt{850 - 25}\ x \sqrt{850 - 25}}$$

$$= \frac{-825}{\sqrt{825}\ x \sqrt{825}} = \frac{-825}{825}$$

= -1 Negative

Example 5

Compute Karl Pearson's coefficient of correlation and probable Error.

Marks in Accounts:	50	60	58	47	49	33	65	43	46	68
Marks in Q.T. :	48	65	50	48	55	58	63	48	50	70

Solution :

X	dx	dx^2	y	dy	dy^2	dxdy
50	0	0	48	-7	49	0
60	+10	100	65	+10	100	100
58	+8	64	50	-5	25	-40
47	-3	9	48	-7	49	21
49	-1	1	55	0	0	0
33	-17	289	58	+3	9	-51
65	+15	225	63	+8	64	-120
43	-7	49	48	-7	49	49
46	-4	16	50	-5	25	20
68	18	324	70	+20	400	360
N=10	19	1077	N=10	10	770	579

$$\gamma = \frac{\in dx\ dy \times N - (\in dx\ .\ \in dy)}{\sqrt{\in dx^2 x\ N - (\in dx)^2}\ x \sqrt{\in dy^2\ x\ N - (\in dy)^2}}$$

$$= \frac{579 \times 10 - (19 \times 10)}{\sqrt{1077 x\ 10 - (19)^2}\ x \sqrt{770 \times 10 - (10)^2}}$$

$$= \frac{5790 - 190}{\sqrt{10770 - 361}\ x \sqrt{7700 - 100}}$$

$$= \frac{5600}{\sqrt{10409}\ x \sqrt{7600}}$$

$$= \frac{5600}{102.02 x 87.18}$$

$= \frac{5600}{8894.28}$

= +0.6296 Positive

Probable error $= 0.6745 \left(\frac{l - r^2}{\sqrt{N}} \right)$

$= 0.6745 \left(\frac{1 - 0.6296^2}{\sqrt{10}} \right)$

$= 0.6745 \left(\frac{1 - 0.3964}{3.162} \right)$

$= 0.6745 \left(\frac{0.6036}{3.162} \right)$

= 0.6745 (0.19)

$= \pm\ 0.128$

Example 6

The population and the number of persons partially or fully blind are given in the following table . Find out whether there is any correlation between their age and their blindness.

Age	**Population in ' 000**	**No.of persons blind**
0-10	100	55
10-20	60	40
20-30	40	40
30-40	36	40
40-50	24	36
50-60	11	22
60-70	6	18
70-80	3	15

Solution :

X	Mv	dx	dx²	y	dy	dy²	dxdy
0-10	5	-4	16	55	-45	2025	180
10-20	15	-3	9	67	-33	1089	99
20-30	25	-2	4	100	0	0	0
30-40	35	-1	1	111	11	121	-11
40-50	45	0	0	150	50	2500	0
50-60	55	+1	1	200	100	10000	100
60-70	65	+2	4	300	200	40000	400
70-80	75	+3	9	500	400	160000	1200
	N=8	-4	44	N=8	683	215735	1968

$$\gamma = \frac{\in dx\ dy \text{ x } N - (\in dx \ . \in dy)}{\sqrt{\in dx^2 \text{x } N - (\in dx)^2} \text{ x} \sqrt{\in dy^2 \text{ x } N - (\in dy)^2}}$$

$$= \frac{1968 \text{ x } 8 - (-4 \text{ x } 683)}{\sqrt{44 \text{ x } 8 - (-4)^2} \text{ x} \sqrt{215735 \text{ x } 8 - (683)^2}}$$

$$= \frac{15744 + 2732}{\sqrt{352 - 16} \text{ x} \sqrt{1725880 - 466489}}$$

$$= \frac{18476}{\sqrt{336} \text{ x} \sqrt{1259391}}$$

$$= \frac{18476}{18.33 \text{x} 1122.23}$$

$$= \frac{18476}{20570.48}$$

= +0.898 Positive

Thus it can be said that there is correlation between the age and blind ness.

Working Note :

0-10	For 100000 Population	55
10-20	$\frac{100000}{60000}$ x 40	67
20-30	$\frac{100000}{40000}$ x 40	100
30-40	$\frac{100000}{36000}$ x 40	111
40-50	$\frac{100000}{24000}$ x 36	150
50-60	$\frac{100000}{11000}$ x 22	200
60-70	$\frac{100000}{6000}$ x 18	300
70-80	$\frac{100000}{3000}$ x15	500

12.6 MERITS OF CO EFFICIENT OF CORRELATION :

1. **Counts all values :** It takes into account all values of the given data of x & y. Therefore it is based on all observations of the series.
2. **More practical and popular :** Karl Pearson's correlation is considered to be more practical method as compared to other mathematical methods used for 'γ'. It is also very popular and as such commonly used method.
3. **Numerical measurement of 'γ' :** It provides numerical measurement of Coefficient of correlation.
4. **Measures degree and direction :** This method measures both degree and direction of the correlation between the variables at a time.
5. **Facilitates comparison :** It is a pure number independent of units. Therefore the comparison between the series can be done easily.
6. **Algebraic treatment possible :** This technique can be easily applied for higher algebraic treatment.

12.7 DEMERITS OF CO EFFICIENT OF CORRELATION :

1. **Linear relationship :** It assumes linear relatoinship between the variables regard less of the fact whether that assumption is correct or not.
2. **More time consuming :** Compared with some other methods, this method is more time consuming.
3. **Affected by extreme items :** This method is affected by extreme items.
4. **Difficult to interpret :** It is not easy to interpret the significance of correlation efficient. It is generally misinterpreted.

12.8 SUMMARY :

Karl Pearson's coefficient of correlation method gives a precise and summary quantitative figure which can be meaningfully interpreted. It gives either positive or negative direction or degree of the relationship between the two variables.

12.9 QUESTIONS :

1. What is meant by coefficient of correlation ?
2. State the types of correlation
3. Explain the method of computing Coefficient of correlation
4. Explain about the probable error.
5. What are the merits and demerits of Co efficient of correlation ?
6. State the assumtion of Karl Pearson's Co efficient of correlation ?
7. What is meant by Linear and non- linear correlation ?

12.10 EXERCISES :

1. Compute Karl Pearson's Co efficient of correlation

Age of Husband :	25	22	28	26	35	20	22	40	20	18	19	25
Age of Wife :	18	15	20	17	22	14	16	21	15	14	15	17

2. Calculate Karl Pearson's Co efficient of correlation

Year :	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Price :	10	12	18	16	15	19	18	17	15	16
Supply :	30	35	45	44	42	48	47	46	44	45

3. Find out Karl Pearson's Co efficient of correlation and Probable Error.

Age of Husband :	23	27	28	29	30	31	33	35	36	39
Age of Wife :	18	22	23	24	25	26	28	29	30	32

4. Ascertain Karl Pearson's Co efficient of correlation and Probable Error.

x :	25	22	28	26	35	20	22	40	20	18	19	25
y :	18	15	20	17	22	14	16	21	15	14	15	17

5. Compute Karl Pearson's Co efficient of correlation and Probable Error.

x :	10	12	18	16	15	19	18	17	15	16
y:	30	35	45	44	42	48	47	46	44	45

6. In the following data, the population and the number persons partly or fully deaf, are given. Find out whether there is any relationship between their age and deafness.

Age	**Population in thousands**	**No.of persons deal**
0-10	100	60
10-20	80	45
20-30	60	43
30-40	40	42
40-50	30	33
50-60	20	30
60-70	10	26
70-80	5	24

7. Find out the Co efficient of correlation between the following two variables. Comment on the result through the Probable Error.

(x) :	6	8	12	15	18	20	24	28	31
(f) :	10	12	15	15	18	25	22	26	28

8. Calculate the Co efficient of correlation from the following data and calculate Probable Error.

Q.T. (x) :	30	60	30	66	72	24	18	12	42	6
Accounts (f) :	06	36	12	48	30	06	24	36	30	12

9. Calculate the Co efficient of correlation between income and weight from the following data. Comment on the result.

Income (Rs.) :	100	200	300	400	500	600
Weight (lbs) :	120	130	140	150	160	170

10. Calculate Karl Pearson's Co efficient of correlation from the following data

(x) :	150	200	250	300	350	400	450	500	550	600
(f) :	600	575	550	525	500	475	450	425	400	375

11. Calculate Karl Pearson's Co efficient of correlation of x and y variables.

(x) :	15	18	30	27	25	23	30
(f) :	7	10	17	16	12	13	9

12. Compute Karl Pearson's Co efficient of correlation for the following data

(Height in inches)

of Husband x :	60	62	64	66	68	70	72
of Wife Y :	61	63	63	63	64	65	67

13. Calculate Karl Pearson's Co efficient of correlation from the following data

(x) :	12	9	8	10	11	13	7
(f) :	14	8	6	9	11	12	3

14. Find out Co efficient of correlation from the following data

(x) :	3	5	6	7	9	12
(f) :	20	14	12	10	9	7

15. Calculate the Co efficient of correlation between Advertisement cost and sales as per the data given below :

Cost in thousands :	39	65	62	90	82	75	25	98	36	78
Sales in Lakhs:	47	53	58	86	62	68	60	91	51	84

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson - 13

SPEARMAN'S RANK CORRELATION

OBJECTIVES:

By the study of this lesson, you will be able to understand, the meaning definition and uses of Spearman's Rank correlation with examples.

STRUCTURE:

13.1 INTRODUCTION:

Charles Edward spearman, a British Psychologist, developed a formula to obtain the rank correlation coefficient in 1904. He has tried to establish the rank correlation coefficient between the Ranks of 'n' individuals in the two or more variables ' Accordingly, it is possible for a class teacher to arrange his students in an ascending order or in descending order of intelligence though intelligence cannot be measured quantitatively. In a similar way ranking can be made in a beauty contest and correlation can be established among the scores given by the different judge or selectors.

It is, however, possible to measure the degree of correlation between two sets of observations or between paired values when only the relative order of magnitude is given for each series. For example, suppose 10 students have appeared for two papess in a test and from actual marks obtained by them, their rankings can be determined. If we want to know whether their performances are correlated, we can use " Sperman's Rank corelation Coefficient " method. The formula is based on the ranks of the variables according to their sizes.

13.2 CIRCUMSTANCES WHEN THE RANK CORRELATION IS USED :

Following are the circumstances when the Rank Correlation coefficient is used.

i In a beauty contest, cooking contest, flower show contest and interview involving selections, we can use the rank correlation coefficient.

ii If the data are irregular or extreme items are erratic or in accurate, we can use the rank correlation coefficient.

In spearmen's coefficient of correlation we take the differences in Ranks, squaring them and finding out the aggregate of the squarred differences. Symbolically.

$$\gamma_s = 1 - \frac{6 \in D^2}{N\ (N^2 - 1)}$$

Where

γ_s = Coefficient of Correlation

$\in D^2$ = Total of the deviations between x & y items

N = No. of pairs.

13.3 TYPES OF RANK CORRELATION :

In Rank Coefficient of correlation three different cases must be studied.

Case I When Ranks are not given

Case II When Ranks are given

Case III When Ranks are equal.

When Ranks are not given :

Example 1

Compute Rank Correlation from the following data

x :	415	434	420	430	424	428
y :	330	332	328	331	327	325

Solution :

X	R_1	y	R_2	D $(R_1 - R_2)$	$-D^2$
415	6	330	3	3	9
434	1	332	1	0	0
420	5	328	4	1	1
430	2	331	2	0	0
424	4	327	5	-1	1
428	3	325	6	-3	9
N=6		N=6			$\in D^2$= 20

$$\gamma_s = 1 - \frac{6 \in D^2}{N(N^2 - 1)}$$

$$= 1 - \frac{6(20)}{6(6^2 - 1)}$$

$$= 1 - \frac{120}{210}$$

$= 1 - 0.571$

$= 0.429$

When Ranks are given :

Example 2

Compute Rank Correlation from the following data

x	:	415	434	420	430	424	428
y	:	330	332	328	331	327	325

Solution :

X	R_1	y	R_2	D $(R_1 - R_2)$	D^2
415	6	330	3	3	9
434	1	332	1	0	0
420	5	328	4	1	1
430	2	331	2	0	0
424	4	327	5	-1	1
428	3	325	6	-3	9
					$\in D^2 = \overline{\underline{20}}$

$$\gamma_s = 1 - \frac{6 \in D^2}{N(N^2 - 1)}$$

$$= 1 - \frac{6(20)}{6(6^2 - 1)}$$

$$= 1 - \frac{120}{210} = 1 - 0.571$$

$= 0.429$

Example 3

The Ranks given by 3 judges to 10 participants in a beauty contest were as under.

Judge A :	1	6	5	10	3	2	4	9	7	8
Judge B :	3	5	8	4	7	10	2	1	6	9
Judge C :	6	4	9	8	1	2	3	10	5	7

Solution :

R_1	R_2	R_3	D $(R_1 - R_2)$	D $(R_2 - R_3)$	D $(R_1 - R_3)$	D^2	D^2	D^2
1	3	6	-2	-3	-5	4	9	25
6	5	4	1	1	2	1	1	4
5	8	9	-3	-1	-4	9	1	16
10	4	8	6	-4	2	36	16	4
3	7	1	-4	6	2	16	36	4
2	10	2	-8	8	0	64	64	0
4	2	3	2	-1	1	4	1	1
9	1	10	8	-9	-1	64	81	1
7	6	5	1	1	2	1	1	4
8	9	7	-1	2	1	1	4	1
						200	214	60

Spearman's Coefficient of Rank Correlation

$$\gamma_s = 1 - \frac{6 \in D^2}{N\,(N^2 - 1)}$$

Between 1 and 2 $= 1 - \dfrac{6x\ 200}{10(10^2 - 1)}$

$= 1 - \dfrac{1200}{990}$

$= 1 - 1.212$

$= -0.212$

Between 2 and 3 $= 1 - \frac{6x\ 214}{10(10^2 - 1)}$

$= 1 - \frac{1284}{990}$

$= 1 - 1.296$

$= -0.296$

Between 1 and 3 $= 1 - \frac{6x\ 60}{10(10^2 - 1)}$

$= 1 - \frac{360}{990}$

$= 1 - 0.3637$

$= 0.6363$

Since the correlation between the judges 1&3 is positive value, it can be said that the pair 1st and 3rd judges have the nearest approach to common taste in beauty.

When Ranks are repeated :

$$\text{Spearman's Rank Correlation} = 1 - \frac{6\ [\in D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + - - - - - n]}{N\ (N^2 - 1)}$$

Example 4

Eight students have obtained the following marks in Accountancy and Economics. Calculate the rank coefficient of correlation.

Accountancy (x) :	25	30	38	22	50	70	30	90
Economics (y) :	50	40	60	40	30	20	40	70

Solution :

X	R_1	y	R_2	D $(R_1 - R_2)$	D^2
25	2	50	6	-4	16.00
30	3.5	40	4	-0.5	0.25
38	5	60	7	-2	4.00
22	1	40	4	-3	9.00
50	6	30	2	+4	16.00
70	7	20	1	+6	36.00
30	3.5	40	4	-0.5	0.25
90	8	70	8	0	0.00
N=8		N= 8			$\in D^2 =$ 81.5

$$R_s = 1 - \frac{6\left[\in D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + - - - - - n\right]}{N(N^2 - 1)}$$

Hence 30 is repeated twice in x series so m = 2
Hence 40 is repeated thrice in y series so m = 3

$$= 1 - \frac{6\left[81.5 + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(3^3 - 3)\right]}{8(8^2 - 1)}$$

$$= 1 - \frac{6\left[81.5 + 0.5 + 2\right]}{504}$$

$$= 1 - \frac{6\left[84\right]}{504}$$

$$= 1 - \frac{504}{504}$$

$$= 0$$

13.4 MERITS OF RANK CORRELATION METHOD :

1. It is easy to calculate and understand as compared to Karl Pearson's coefficient of correlation.

2. When the ranks of different values of the variables are given, it is then the only method left to calculate the degree of correlation.

3. When actual values are given and we are interested in using this formula, then we have to give ranks to calculate correlation.

4. This method is employed usefully when the data is given in a qualitative nature like beauty, honesty, intelligence etc.

13.5 DEMERITS OF RANK CORRELATION METHOD :

1. This method cannot be employed in a grouped frequency distribution.

2. If the items exceed 30, it is then difficult to find out ranks and their differences.

3. This method lacks precision as compared to pearson's coefficient of correlation as all the information concerning the variables is not used. It is just possible that the difference between Rank correlation and coefficient of correlation may be very insignificant.

13.6 SUMMARY :

Spearman's Rank correlation is based on the ranking of different items in the variable. This method is useful where actual item values are not given, simply their ranks in the series are known. Thus it is a good measure in cases where abstract quantity of one group is correlated with that of the other group.

13.7 QUESTIONS :

1. What is meant by Rank correlation ?

2. Write down spearman's formula for rank correlation co-efficient.

3. What are the Merits of Rank Correlation ?

3. What are the Limitations of Rank Correlation ?

18.6 EXERCISE :

1. In a beauty competittion two judges ranked 12 participants as follows.

Judge A :	3	4	1	5	2	10	6	9	8	7	12	11
Judge B :	6	10	12	3	9	2	5	8	7	4	1	11

2. Two ladies were asked to rank seven different brands of lipsticks as listed below.

Brands :	A	B	C	D	E	F	G
Lady 1 :	1	3	2	7	6	4	5
Lady 2 :	2	1	4	6	7	3	5

3. Ten participants in a beauty contest were ranked by three judges in the following order.

Judge 1 :	8	1	2	10	3	7	5	9	4	6
Judge 2 :	4	7	10	1	2	9	6	8	5	3
Judge 3 :	10	3	2	9	4	8	7	5	6	1

which of the 2 judges are agreeing with each other and who are against each other ?

4. In a contest, two judges ranked eight candidates in order of their performance as follows.

Judge 1 :	5	2	8	1	4	6	3	7
Judge 2 :	4	5	7	3	2	8	1	6

Find out the Rank Correlation.

5. Calculate the rank correlation coefficient for the following data.

x :	60	34	40	50	45	41	22	43	42	66	64	46
y :	75	32	35	40	45	33	12	30	36	72	41	57

6. From the marks obtained by 8 students in Accountancy and statistics, compute coefficient of correlation by rank difference method.

Marks in

Accountancy :	60	15	20	28	12	40	80	20
Statistics :	10	40	30	50	35	20	60	38

7. Ten competitors in a voice contest are ranked by three judges in the following order.

Judge 1 :	1	6	5	10	3	2	4	9	7	8
Judge 2 :	3	5	8	4	7	10	2	1	6	9
Judge 3 :	6	4	9	8	1	2	3	10	5	7

Which 2 judges have the nearest approach to common likings invoice ?

Which 2 judges have the opposite approach to common likings invoice ?

8. Find out spearman's rank correlation ?

x :	5	2	8	1	4	6	3	7
y :	4	5	7	3	2	8	1	6

9. Eight students have obtained the following marks in accountancy and statistics. Find out rank correlation.

x :	56	48	40	67	75	80	85	35
y :	75	43	56	94	71	92	76	54

10. Ten participants in a beauty contest were ranked by three judges in the following order.

Judge 1 :	8	1	2	10	3	7	5	9	4	6
Judge 2 :	4	7	1	1	2	9	6	8	5	3
Judge 3 :	10	3	2	9	4	8	7	5	6	1

Using rank correlation, determine which pair of judges have the nearest approach to common tastes in beauty.

11. Caculate the rank coefficient of correlation from the following data.

x :	80	78	75	75	68	67	60	59
y :	12	13	14	14	14	16	15	17

12. Caculate the rank coefficient of correlation from the data given below .

x :	91	97	102	103	103	105	110	114	116	124
y :	102	94	105	115	113	99	92	112	120	108

13. Find out Rank Correlation

x :	10	12	60	60	70
y :	15	20	20	20	50

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson - 14

REGRESSION ANALYSIS - I

OBJECTIVES:

By the study of this Chapter, you will be able to understand the meaning, Objectives, merits and limitations of Regression analysis and differences between correlation and Regression.

STRUCTURE:

14.1 INTRODUCTION :

' Regression ' means returning or stepping back to the average value with the help of values of one variable (independent) we can establish most likely values of other variable (dependent). On the basis of two available correlated variables, we can forecast the future data or events or values.

In statistics the term ' Regression ' means simply the ' Average Relationship '. We can predict or estimate the values of dependent variable from the given related values of independent variable with the help of a Regression Technique. The measure of Regression studies the nature of correlationship to estimate the most probable values. It establishes a functional relationship between the ' Independent ' and ' Dependent ' Variables.

The statistical technique of estimating or predicting the unknown value of a dependent variable from the known value of an independent variable is called regression analysis. Sir Francis Galton introduced the concept of ' Regression ' for the first time in 1877 where he studied the case of one thousand fathers and sons and concluded that the tall fathers tend to have tall sons and short fathers have short sons, but the average height of the sons of a group of tall fathers is less than that of the fathers and the average height of the sons of a group of short fathers is greater than that of the fathers.

The line showing this tendency to go back was called by Galton " Regression line ". The modern statisticians use the term ' estimating line ' instead of regression line as this concept is more classificatory now.

Sales depend on promotional expense. It is possible to predict sales for a given promotion expense. Regression is more useful for business planning and fore casting.

In Economics it is the basic tool for estimating the relationship among economic variables that constitute the essence of economic theory. If we know the two variables, price (x) and demand(y) are closely related, in that case we can find the most probable value of y for a given value of x.

14.2 DEFINITIONS :

According to Morris M.Blair " Regression is the measure of the average relationship between two or more variables in terms of the original units of the data.

According to Taro Yamane " one of the most frequently used techniques in economics and business research to find a relation between two or more variables that are related casually is regression analysis ".

Ya Lun Chou defines it as , " Regression analysis attempts to establish the nature of relationship between variables and there by provide a mechanism for prediction or forecasting".

14.3 OBJECTIVES OF REGRESSION ANALYSIS :

Regression analysis does the following

1. Explain the variations in the dependent variable as a result of using a number of independent variables.
2. Describe the nature of relationship in a precise manner by way of a regression equation.
3. It is used in prediction and forecasting problems.
4. It helps in removing unwanted factors.

14.4 DISTINCTION BETWEEN CORRELATION AND REGRESSION :

The correlation and the regression analysis help us in studying the relationship between the two variables yet they differ in their approach and objectives.

S.No	Correlation	S.No	Regression
1.	It preceds regression	1.	It succeeds correlation .
2.	It tests the closeness between the two variables.	2.	It studies the closeness between the two variables and estimates the values.
3.	It measures the degree of covariation	3.	It measures the nature of covariation.
4.	It is merely a tool of ascertaining the degree of relationship	4.	It is also a tool of studying cause and effect of relationship.
5.	The relationship may be purely a chance and it may not have practical relevance.	5.	There is a perfect relationship and it has practical relevance.
6.	There is no question of independent and dependent variables.	6.	There is an identification of independent and dependent variables.
7.	It is a two way average relationship	7.	It is a directional relationship with cause and effect.
8.	It establishes just a relationship	8.	It studies the functional relationship with the two equations of lines.

Both the techniques are based on different sets of assumptions in practice, the choice between the two techniques depends upon the purpose of investigation. The presence of correlationship does not imply causation, but the causation certainly implies correlationship. The association (correlation) need not imply causation (Regression) because a close association may be the result of pure chance. The causation (regression) definitely implies association (correlation), because cause and effect are based on relationship.

14.5 REGRESSION LINES :

A regression line is a graphic technique to show the functional relationship between the two variables x and y i.e dependent and independent variables. It is a line which shows average relationship between two variables x & y . Thus this is a line of average. This is also called an estimating line as it gives the average estimated value of dependent variable (y) for any given value of independent variable (x)

According to Galton " The regression lines show the average relationship between two variables ".

In the words of J.R. Stockton , " The device used for estimating the value of one variable from the value of the other consists of a line through the points drawn in such a manner as to represent the average relationship between the two variables. Such a line is called the line of regression ".

14.6 CLASSIFICATION OF REGRESSION ANALYSIS :

The regression analysis can be classified on the following basis -

1) Change in proportion and

2) Number of variables.

1. Basis of Change in Proportion :

On the basis of proportions the regression can be classified into the following categories.

i) Linear regression and

ii) Non - Linear regression.

i. Linear Regression Analysis Model : When dependent variable moves in a fixed proportion of the unit movement of independent variable it is called a linear regression. Linear regression when plotted on a graph paper forms a straight line. Mathematically the relation between x and y variables can be expressed by a simple linear regression equation as under .

$$Y_1 = a + b\, x_1 + e_1$$

Where a and b are known as regression parameters, e_1 donotes residual terms, x_1 presents value of independent variable and y_1 is the value of dependent variable. "a" expresses the intercept of the regression line of "y" on ' x' i.e value of dependent variable say ' y' when the value of independent variable that is ' x' is zero. Again 'b' denotes the slope of regression line of 'y' on ' x'. Again e_1 denotes the combined effect of all other variables (not taken in the model) on 'y'. This equation is known as classical simple linear regression model.

ii. Non - Linear Regression Analysis Model : Contrary to the linear regression model, in non - linear regression the value of dependent variable say ' y' does not change by a constant absolute amount for unit change in the value of the independent variable say 'x'. If the data are dotted on a plot, it would form a curve rather than a straight line. This is also called curvi - linear regression.

2. On the basis of number of variables :

On the basis of number of variables regression analysis can be classified as under -

i) Simple Regression

ii) Partial Regression

iii) Multiple Regression

i. **Simple Regression :** When only two variables are studied to find the regression relationship it is known as simple Regression analysis. Of these variables one is treated as an independent variable while the other as dependent one. Functional relation between price and demand may be noted as an example of simple regression.

ii. **Partial Regression :** When more than two variables are studied in a functional relationship but the realtionship of only two variables is analysed at a time, keeping other variables as constant, such a regression analysis is called partial regression.

iii. **Multiple Regression :** When more than two variables are studied and their relationships are simultaneously worked out it is a case of multiple regression. Study of the growth in the production of wheat in relation to fertilisers, hybrid seeds irrigation etc. is an example of multiple regression.

14.7 MERITS / UTILITIES / USES OF REGRESSION ANALYSIS :

The technique of regression is considered to be the most useful statistical tool applied in various fields of sociological and scintific disciplines. It is helpful in making quantitative predictions in the behaviour of the related variables. Following are some of the main uses of regression analysis.

1. Prediction of unknown value :

The regression analysis technique is very useful in predicting the probable value of an unknown vairable in response to some known related variable. For example the estimate of demand on a given price can be made if the demand and given price. are functionally related to each other.

2. Nature of relationship :

The regression device is useful in establishing the nature of the relationship between two variables.

3. Estimation of realtionship :

Regression analysis is extensively used for the measurement and estimation of the relationship among variables. It is an important statistical device which provides basis for analysis and interpretation in research studies.

4. Calculation of coefficient of determination :

The regression analysis provides regression coefficients which are generally used in calculation of coefficient of correlation. The square of co-efficient of correlation (r) is called the coefficient of determination which measures the degree of association that exists between two variables. The higher the value of r^2 the better are regression lines and more useful are the regression equations for prediction and estimation.

5. Helpful in calculation error :

Regression analysis is very helpful in estimating the error involved in using the regression line as a basis for estimation.

6. Policy formulation :

The prediction made on the basis of estimated inter relationship through the techniques of regression analysis provide sound basis for policy formulation in socio- economic fields.

7. Touch stone of hypothesis :

The regression tool is considered to be a pertinent testing tool in statistical methodology. It is used in testing the laws and theories of the social sciences as well as natural sciences where the inter relationship between the variables is involved.

14.8 LIMITATIONS OF REGRESSION ANALYSIS :

Despite all utilities the regression analysis too has various limitations. The following are some of the limitations of regression analysis.

1. Assumption of linear relationship :

Regression analysis is based on the assumption that there always exists linear relationship between related variables. The linear type of relationship doesnot always exist in the field of social sciences. In these fields non - linear or culvilinear relationship are most commonly found.

2. Assumption of static conditions :

While calculating the regression equations a static condition of relationship between the variables is presumed. It is supposed that the relationship has not changed since the regression equation was computed. Such type of assumption has made the regression analysis a static one and hence reduces its applicability in social fields.

3. Study of relationship in prescribed limits :

The linear relationship between the variables can only be ascertained with in limits. When prescribed limits are crossed the results become incorrect or inconsistent. Such a relation exists between price and profits. When prices are higher the profits are high to a certain limit. when the prices are abnormally high the profit may decline due to entry of new firms increasing there by the supply of the commodity.

14.9 SUMMARY :

Regression analysis measures the closelyness with which two or more variable co-vary in a given period of study. Similarly value of one variable can be estimated or predicted on the basis of functional relationship between them, given the value of another variable. Regression technique is considered to be a most statistical device.

14.10 QUESTIONS :

1. Define the term ' Regression '
2. What do you mean by the term ' Regression ' ?
3. What are the objectives of Regression analysis ?

4. What are the distinctions between Regression and Correlation ?
5. Explain the meaning of ' Regression lines ' ?
6. Explain the classification of Regression analysis.
7. Explain the term Simple, Partial and Multiple Regression .
8. What are the merits and demerits of Regression analysis ?

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson - 15

REGRESSION ANALYSIS - II

OBJECTIVES:

By the study of this Chapter, you will be able to understand the methods of calculating Regression equation with numerical examples in detail.

STRUCTURE:

15.1 INTRODUCTION :

There are Two Regression equations. They are -

1. Regression equation of x on y

2. Regression equation of y on x

15.2 REGRESSION EQUATION OF X ON Y :

x- a = bxy (y-a)

Where

x = x

a = Arithmetic Mean in x series $\left(x + \frac{\in dx}{N} \right)$

y = y

a = Arithmetic Mean in y series $\left(y + \frac{\in dx}{N} \right)$

$$bxy = \frac{\in dxdy \ x \ N - (\in dx. \ \in dy)}{\in dy^2 \ x \ N - (\in dy)^2}$$

15.2 REGRESSION EQUATION OF Y ON X :

y - a = byx (x-a)

Where

y = y

a = Arithmetic Mean in y series

x = x

a = Arithmetic Mean in x series

$$byx = \frac{\in dxdy \times N - (\in dx. \in dy)}{\in dx^2 \times N - (\in dx)^2}$$

15.3 CO EFFICIENT OF CORRELATION THROUGH REGRESSION EQUATION :

$$\gamma = \sqrt{bxy.\ byx}$$

15.4 REGRESSION EQUATION ON THE BASIS OF STANDARD DEVIATION (δ) AND COEFFICIENT OF CORRELATION (γ) :

Regression equation of x on y = x - a = $\gamma . \frac{\delta x}{\delta y}$ (y - a)

Regression equation of y on x = y - a = $\gamma . \frac{\delta y}{\delta x}$ (y - a)

Where δx = standard Deviation of x series

δy = standard Deviation of y series

γ = coefficient of correlation

15.6 EXAMPLES :

Example 1

Find out the regression equation of x on y and y on x

x :	6	2	10	4	8
y :	9	11	5	8	7

Solution :

X	dx	dx^2	y	dy	dy^2	dxdy
6	0	0	9	+1	1	0
2	-4	16	11	+3	9	-12
10	+4	16	5	-3	9	-12
4	-2	4	8	0	0	0
8	+2	4	7	-1	1	-2
N=5	0	40	N=5	0	20	-26

$$a = \left(x + \frac{\in dx}{N} \right) \qquad a = \left(y + \frac{\in dx}{N} \right)$$

$$= 6 + \frac{0}{5} \qquad = 8 + \frac{0}{5}$$

$$= 6 \qquad = 8$$

$$bxy = \frac{\in dxdy \text{ x } N - (\in dx. \ \in dy)}{\in dy^2 \text{ x } N - (\in dy)^2}$$

$$= \frac{-26x5 - (0x0)}{20x5 - (0)^2}$$

$$= \frac{-130}{100}$$

$$= -1.3$$

$$byx = \frac{\in dxdy \text{ x } N - (\in dx. \ \in dy)}{\in dx^2 \text{ x } N - (\in dx)^2}$$

$$= \frac{-26x5 - (0x0)}{40x5 - (0)^2}$$

$$= \frac{-130}{200}$$

$$= -0.65$$

Regression equation of x on y = x- a = bxy (y-a)

x - 6 = 1.3 (y - 8)

x - 6 = 1.3y + 10.4

x = -1.3y + 10.4 + 6

x = 1.3y + 16.4

Regression equation of y on x = y - a = byx (x-a)

y - 8 = -0.65 (x - 6)

y - 8 = -0.65x + 3.90

y = -0.65x + 3.90 + 8

y = -0.65x + 11.9

Example 2

Following data is given to you

x :	1	5	3	2	1	1	7	3
y :	6	1	0	0	1	2	1	5

Compute the regression equation of x and y. What is the value of x when the value of 'y' = 2.5

Compute the regression equation of y and x. What is the value of y when the value of 'x' = 5

Solution :

X	**dx**	**dx²**	**y**	**dy**	**dy²**	**dxdy**
1	-2	4	6	+4	16	-8
5	+2	4	1	-1	1	-2
3	0	0	0	-2	4	0
2	-1	1	0	-2	4	2
1	-2	4	1	-1	1	2
1	-2	4	2	0	0	0
7	+4	16	1	-1	1	-4
3	0	0	5	+3	9	0
N=8	-1	33	N=8	0	36	-10

$a = \left(x + \frac{\in dx}{N} \right)$ $\quad$ $a = \left(y + \frac{\in dy}{N} \right)$

$= 3 + \frac{-1}{8}$ $= 8 + \frac{0}{8}$

= 3 - 0.125 = 2

=2.875

$$bxy = \frac{\in dxdy \ x \ N- (\in dx. \ \in dy)}{\in dy^2 \ x \ N - (\in dy)^2}$$

$$= \frac{-10x8-(-1x0)}{36x8-(0)^2}$$

$$= \frac{-80}{288}$$

= - 0.28

$$byx = \frac{\in dxdy \ x \ N- (\in dx. \ \in dy)}{\in dx^2 \ x \ N - (\in dx)^2}$$

$$= \frac{-10x8-(-1x0)}{33x8-(-1)^2}$$

$$= \frac{-80}{264-1}$$

$$= \frac{-80}{263} \quad = \ - 0.30$$

Regression equation of x ony = x- a = bxy (y-a)

x - 2.875 = -0.28(y - 2)

x - 2.875 = -0.28y + 0.56

x = -0.28y + 0.56 + 2.875

x = -0.28y + 3.445

When the value of y = 2.5

x = -0.2.8 (2.5) + 3.445

= - 0.7 + 3.445

x = 2.745

Regression equation of y and x =y - a = byx (x-a)

$y - 2 = -0.3 (x - 2.875)$

$y - 2 = -0.3 x + 0.8625$

$y = -0.3x + 0.8625 + 2$

$y = -0.3x + 2.8625$

When the value of x = 5

$y = -0.3 (5) + 2.8625$

$= - 1.5 + 2.8625$

$y = 1.3625$

Example 3

The Arithmetic mean values of x series and y series are 65 and 67 respectively. Their standard Deviations are 2.5 and 3.5 respectively. Coefficeint of correlation of the two series is 0.8. Write down the two regression lines.

a) What is the value of x when the value of y = 70

b) What is the value of y when the value of x is equal to the value computed as per (a) above?

Solution :

Regression equation of x on y = $x - a = \gamma \cdot \frac{\delta x}{\delta y} (y - a)$

$$x - 65 = 0.8 \frac{2.5}{3.5} (y - 67)$$

$x-65 = 0.8 \times 0.714 (y-67)$

$x - 65 = 0.571 (y - 67)$

$x -65 = 0.5714 - 38.257$

$x = 0.571y - 38.257 + 65$

$x = 0.571y + 26.743$

When the value of y = 70

$x = 0.571 (70) + 26.743$

$= 39.97 + 26.743$

$x = 66.713$

Regression equation of y on x = $y - a = \gamma . \frac{\delta y}{\delta x} (y - a)$

$$y - 67 = 0.8 \frac{3.5}{2.5}(x - 65)$$

$$y - 67 = 0.8 \times 1.4 (x - 65)$$

$$y - 67 = 1.12 (x - 65)$$

$$y - 67 = 1.12x - 72.8$$

$$y = 1.12x - 72.8 + 67$$

$$y = 1.12x - 5.8$$

When the value of x = 66.173

$$y = 1.12 (66.173) - 5.8$$

$$y = 74.11 - 5.8$$

$$y = 68.31$$

Example 4

The following data is given to you

	Am	S.D
Yield of paddy	1000	60
Rain fall	20	2
Co efficient of correlation	0.7	

Compute the two regression equation

Estimate the yield of paddy when the annual rainfall is 15"

Estimate the annual rainfall when the yield of paddy is 1500 pounds.

Solution :

Regression equation of x on y = $x - a = \gamma . \frac{\delta x}{\delta y} (y - a)$

$$x - 1000 = 0.7 \frac{60}{2}(y - 20)$$

$$x - 1000 = 21 (y - 20)$$

x - 1000 = 21y - 420

x = 21y - 420 + 1000

x = 21y + 580

When the value of y = 15 (rain fall)

x = 21(5) + 580

= 315 + 580

x = 895

Regression equation of y on x = y - a $= \gamma = \frac{\delta\ y}{\delta\ x}$ (x - a)

y - 20 = 0.7 $\frac{2}{60}$ (x - 1000)

y- 20 = 0.023 (x - 1000)

y - 20 = 0.023 x 99

y = 0.023x - 23 + 20

y = 0.023x -3

When the value of x = 1500 (yield of paddy)

y = 0.023 (1500) - 3

y = 34.5 -3

y = 31.5

Example 5

From the following data, show the two regression lines and find out the coefficient of correlation with the help of equation.

x :	80	45	55	56	58	60	65	68	70	75	85
y :	82	56	50	48	60	62	64	65	70	74	90

Solution :

X	dx	dx^2	y	dy	dy^2	dxdy
80	+20	400	82	+12	144	240
45	-15	225	56	-14	196	210
55	-5	25	50	-20	400	100
56	-4	16	48	-22	464	88
58	-2	4	60	-10	100	20
60	0	0	62	-8	64	0
65	+5	25	64	-6	36	-30
68	+8	64	65	-5	25	-40
70	+10	100	70	0	0	0
75	+15	225	74	+4	16	60
85	+25	625	90	+20	400	500
N=11	57	1709	N=11	-51	1845	1148

$$a=\left(x+\frac{\in dx}{N}\right) \qquad a=\left(y+\frac{\in dx}{N}\right)$$

$$= 60+\frac{57}{11} \qquad = 70+\frac{-51}{11}$$

$$= 60 + 5.18 \qquad = 70 - 4.64$$

$$= 65.18 \qquad = 65.36$$

$$bxy = \frac{\in dxdy \times N - (\in dx. \ \in dy)}{\in dy^2 \times N - (\in dy)^2}$$

$$= \frac{1148x11 - (57x -51)}{1845 \times 11 - (-51)^2}$$

$$= \frac{12628 + 2907}{20295 - 2601}$$

$$= \frac{15535}{17694} = 0.88$$

$$byx = \frac{\in dxdy \ x \ N- (\in dx. \ \in dy)}{\in dx^2 \ x \ N - (\in dx)^2}$$

$$= \frac{1148 \ x \ 11 - (57 \ x \ -51)}{1709 \ x \ 11 - (57)^2}$$

$$= \frac{12628 + 2907}{18799 - 3249}$$

$$= \frac{15535}{15550} \quad = \quad 0.9997$$

Coefficient of correlation $= \sqrt{bxy. \ byx}$

$= \sqrt{0.88 \ x \ 0.9997}$

= 0.8797

Regression equation of x and y = x- a = bxy (y-a)

x - 65.18 = 0.88(y - 65.36)

x - 65.18 = 0.88 y - 57.52

x = 0.88y - 57.52 + 65.18

x = 0.88y + 7.66

Regression equation of y and x =y - a = bxy (x-a)

y - 65. 36 = 0.9997 (x - 65.18)

y - 65.36 = 0.9997x - 65.16

y = 0.9997x - 65.16+ 65.36

y = 0.9997x + 0.20

15.7 SUMMARY :

These are two regression equation. They can also be computed with the help of standard deviation and coefficient of correlation.

15.8 EXERCISE :

1. Find out the two Regression Equations.

x :	6	2	10	4	8
y :	9	11	5	8	7

2. Show the two regression lines.

x :	1	2	3	4	5	6	7	8	9
y :	9	8	10	12	11	13	14	16	15

3. The Height 10 father and sons were as under

Height of Fathers :	158	166	163	165	167	170	167	172	177	181
Height of Fathers :	163	158	167	170	160	180	170	175	172	175

Find out the two regression equation. Estimate the height of the son if the height of the father is 164 cm.

4. From the following data, find out the two regression equation.

	Am	S.D
Telugu (x)	40	10
English (y)	50	16

Coefficient correlation (γ) 0.3

Estimate the marks in English if the marks in Telugu are 50

Estimate the marks in Telugu, if the marks in English are 30.

5. The Heights and weights of 10 students were as under.

Height (x) in inches):	61	68	68	64	65	70	63	62	64	67
Height (y) (in pounds):	112	123	130	115	110	125	100	113	116	125

6. Following information is given to you -

Marks in

	Am	S.D
(x)	36	11
(y)	85	8
(γ)	0.66	

Write up two regression equation

Estimate the value of ' y ' if the value of x is 75

Estimate the value of ' x ' if the value of y is 75

7. Calculate the two regression equations. Compute the coefficient of correlation from the regression lines.

x :	100	101	k02	102	100	99	97	98	96	95
y :	98	99	99	97	95	92	95	94	90	91

8. Compute the two regression equations and find out Karl pearson's coefficient of correlation from the regression lines.

x :	23	27	28	29	30	31	33	35	36	39
y :	18	22	23	24	25	26	28	29	30	32

9. Draw up the two regression equation.

x :	27	21	35	44	29	30	32	42	41	36	28	26
y :	40	37	21	25	36	41	22	31	23	24	39	37

10. Compute two regression equation.

x :	10	12	18	16	15	19	18	17	15	16
y :	30	35	45	44	42	48	47	46	44	45

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson 16

SET THEORY

After studying this lesson you should be able to understand the following.

1. What is Set Theory

2. Types of Sets

3. Operations on Sets.

STRUCTURE OF LESSON

Set theory is applied in the study of difference mathematical concepts, statistical issues and general ecomic problems. It is important technique of the science of mathematics. It pays an important part in the quantitative analysis of socio-economic and business related problems.

Development of Set Theory

George Boole (1815-1864), on English Mathematician laid foundation to the technique of set theory through his book ‘ Investigation of Laws of Thoughts’.

16.1 SET - DEFINITION

Set is defined in various ways such as -

1. A set is a well defined collection of objects.

2. A set is any list, collection or aggregate of objects considered for a study.

3. A set is a collection of objects in which it is possible to decide whether a given object belongs to the collection.

4. “According to Tom M. Apostol “In mathematics the word set is used to represent a collection of objects viewed as a single entity”.

16.2 MEANING OF SET

In simple words, a set is a collection of objects or things. In mathetics set theory is a concept that deals with a phenomenon or a fact or inforformation in the form of a group or class called set.

Example :

1. Students in a class is one set

2. Books in the shelf

3. A team of doctors in the hospital

4. letters in our name.

Thus set is collection of elements or objects or things with some similarities.The objects are aggregated with a purpose, may be to study a phenomenon or to make a comparative analysis or any such idea. A set may be described by actually listing the objects belonging to it within enclosed brackets. This form of presenting objects is called “tabular form of a set”.

A set is always denoted by capital letters. Such as A, B, C. A = {35, 79}

Elements in set are denoted by small leters B = {a p p l e}

16.3 IMPORTANCE OF SET

The development of set theory influenced significantly the science mathematics and its applied branches. Some of the uses of set theory are

1. Set theory is an important technique in the quantitative analysis of different contemporary socio-economic and financial problems.

2. Set theoy contributed significantly for the development of the subject matter mathematics and its related branches.

3. Set theory has a special application in the study of relationship between mathematics and other social science.

4. Set theory makes it possible to develop workable methods and principles in any branch of sciences.

5. Set theory is more convenient to study, apply and analyse different issues in mathematics and statistics.

16.4 FEATURES OF SET THEORY

Following are the important features of set theory.

a. Set is a unified notion. It can be described more precisely than defining in a comprehensive way.
b. Set can be a collection of any group of objects. Any group implies that the scope of objects is extensive.
c. Set is a class of elements and each element is a unique one and a distinguished part of the set. Due to this feature it is defined as “collection of distinct objects”.
d. Unlike other mathematical expressions where the order of the listing of numbers and elements

are much to do with the ultimate result, elements in sets are not bound by such conditional arrangements.

e. Another important feature of sets is that nothing is assumed about the nature of the individual objects in the collection.

f. Abstract set theory deals with such collections of arbitrary objects, which make the theory on is, branch in mathematics.

Cordinality of a Set : It is also called order of set. Number of elements in a set is called Cordinality of set. it is denoted by n().

Elements : The individual objecs of the collection or set are called an element or a member of the set. They are denoted usually by small letters a, b, c, etc.

16.5 PRESENTATION OF SETS

A set is an aggregate of elements structured within brackets. The total structure of set gives the features of information to explain a phenomenon or a fact through a set. The related elements are to be arranged in a particular form, which states the features and meaning of elements and the total set. There ae two important ways of presentation of sets. They are 1) Roster method and 2) Set Builder method.

16.5.1 Roster Form or Tabulation method : Under the roster method all the elements of a set are written in a continuous rows. Each element is separated by comma. Total elements are enclosed in brackets.

Example 1 : A set of even numbes 10 and below 10 can be shown as : A = { 2, 4, 6, 8, 10 }

Example 2 : A set of numbers in a telephone number : B = { 2, 2, 5, 0, 7, 1, 3 }

Example 3: Set of natural numbers upto 1000. C = { 1, 2, 3, 4, ... 1000 }

Merits :

1. Common and simple way of presenting a set,

2. It is a direct method

3. A glance at the set gives the idea, meaning and features of set and its elements.

Demerits :

1. This method is not convenient to present lengthy and complex information.

2. This method is not convenient to elements that require additional explanation in the form of a statement.

16.5.2 Set Builder form or Rule Method : Under this method, a set may be specified by stating its properties. In this form of presentation, instead of writing the elements directly, the rule that governs all the elements or the common feature of elements or the properly of the elements in the set is stated in the form of a statement. It is important that all the elements of a set should satisfy the property and is within this common rule. Thus it is also knows as Rule Method.

Merits

1. This method is used when the elements to be listed in a set are large and infinite.

2. This method states the common property of the set, therefore one can get the feaures of the set more easily.

Demerit

This method takes time to understand and get the idea of actual elements of the set, because they arenot given directly in the set.

Exmaple :

1. Set of temples in West Godavari District of A.P. can be presented as,

A = { x : x is a temple in West Godavari Disrict of A.P. }

In the example set 'A' denotes all the temples in a district and this feature is expressed through an element x in such way that x is a temple in the district. Therefore all the other elements in set 'A' are temples in a district.

{ x : x or x/x to be read as such that }

16.6 TYPES OF SETS

Depending on the nature of elements and form of arrangement there are different types of sets. They are -

1. Finite Set : A set is defined as finite set, if the number of elements in the set is finite. It implies that the set consists of a specific number of elements. They are known and can be counted. The counting process will have an end.

Example : Set of even numbers below 11

A = { 2, 4, 6, 8, 10 }

2. Infinite Set : A set which is not finite is called infinite set. 'A' infinite set consists of countless numbers of elements which cannot be known and the counting process never ends.

Example : Set of days. A = { x : x is a day }

3. Null Set : It is also known as empty set or void set. A set which has no elements at all, is called a null set. it is denoted by Greak letter or by { }.

4. Singleton Set : Singleton set is also known as unit set or one element set. If a set contains only one element it is called Singleton Set.

Example : A = { a }

Even if a set contains same element repreatedly recorded in a set, it is defined as Singleton.

Example: Telehone number of 3333333

z = { 2, 2, 2, 2, 2, 2, 2 }

5. Universal Set : Universal set, as the name indicates is a set of all the elements of a specific issue or phenomena under consideration. It is denoted by U or 1 or Ω (omega) or μ.

Positive Integers above 1 and below 10 are considered for the study. Identify Universe Set.

A = { 3 5 7 9 } C = { 1 2 3 4 8 9 }

B = { 1 5 9 }   D = { 1 2 3 4 5 6 7 8 9 }

D is the Univesal set because is consists of all the integers between 1 and 10.

6. Sub Set : If A and B are two sets such that every element of A is also an element of B then A is a subset of B and 'B' is super set of A.

Example : $A \subseteq B$ = It denotes 'A' is a Sub set of B.

7. Proper Sub Set : A set is called a proper sub set of another set when every element of super set is in sub set and if the super set contains at least one element not in sub set.

Example: B = { 1, 2, 3, 4 }, A = { 2, 3, 4 }

A contains all elements of B except 1 therefore A is a prope Sub set of B.

8. Disjoint Sets : Two sets are called disjoint sets, if they do not have any common element between them, it implies that the elements of one set are totally different from other set.

Example: A { 1, 2, 3, 4, 5 }

B { 6, 7, 8, 9 }

A, B are disjoint sets because there is no element of A in set B and no element of B is found in set A.

9. Equal Sets : Two sets A and B are defined as equal sets if and only if every element of A is an element of B and also every element of B is an element of A.

It implies that in two sets elements are same and no set contains any element extra to other set or less to other set.

Example : A = { 1, 2, 3, 4 }   B = { 1, 2, 3, 4 }

A and B are equal sets.

10. Equivalent Sets : It A and B are two sets and if total number of elements of A and B are same they are called equivalent sets.

Example : A = { 3, 4, 2, 6, 7, 9 }   B = {1, 2, 3, 4, 5, 6 }

11. Comparable and Non-comparable sets : If A and B are two sets and if one of A and B is sub sets of anotehr set then A and B are comparable sets. if none of two sets are sub sets of another set then they are non-comparable sets.

12. Power Set : If A is a set, then the group of all possible sub sets of A is called power set of A. It is denoted by P(s).

If A = { 1, 2 }, it implies set A has 2 elements

the possible subsets of A are the elements of $P = (A) = 2^2 = 4$.

13. Class of Sets or Family of Sets : If elements of a set are sets themselves, such set is called class of sets or family of sets.

A = { (1), (2), (3), (4) }

14. Compliment of a Set : If set A is a subset of universal set U, set A contain element of U.

16.7 EXERCISE

1. What is a Set ? Explain its meaning.
2. Describe importance of Set Theory.
3. Explain Features of Set Theory.
4. What are the methods of presentation of Set Theory.
5. Explain different Types of Sets.

- Dr. K.Kanaka Durga

Lesson - 17

SET THEORY - II

17.0 OBJECTIVE

After studying this lesson you should be able to understand.

1. Operations on Sets.
2. Venn Diagram.
3. Applications of Set theory

STRUCTURE OF LESSON

17.1 Operations on Sets

Operations on sets lead to formation of new sets. The main operations of sets are the Union of sets intersection, complement of a set, difference of two sets. These operations can also be expressed with the help of Venn diagram.

17.1.1 Union of Sets

The union of two sets A and B means set of all elements contained in A as well as in B. It is denoted by $A \cup B$. It can be written in set builder form as :

$A \cup B = \{x : x \in$ B or both A & B$\}$

Example 1: A = {5, 6, 8, 11 }

B = {4, 5, 7, 9, 10, 11}

$A \cup B$ = {4, 5, 6, 7, 8, 9, 10, 11}

17.1.2 Intersection of Sets

Intersection of two given sets A and B refers to a set of elements common to set A and set B. It is denoted by $A \cap B$. It is expressed in the following form.

$A \cap B = \{x : x \in A \text{ and } x \in B\}$

$A \cap B = B \cap A$

Example 2 :

A = {1, 3, 5, 7}

$B = \{2, 3, 5, 6\}$

$A \cap B = \{3, 5\}$

Elements of 3 and 5 are common in both A and B sets.

17.1.3 Diference of Sets

Difference of two sets A and B is a set of all those elements of A which are not contained in B. It is denoted by A - B. It can be expressed in set builder form as

$A - B = \{x : x \in A \text{ and } x \notin B\}$ also

$B - A = x : x$ and x A

Example 3 : $A = \{1, 2, 3, 4\}$

$B = \{3, 4, 5, 6\}$

$A - B = \{1, 2\}$

$B - A = \{5, 6\}$

17.1.4 Complement of a Set

If U is the Universal set and A is a subset. Then complement of A is a set of all elements present in Universal Set except elements contained in A. it is denoted by A^c or A^1. it can be expressed in set-builder form as follows.

$A^1 = \{x : x \in U \text{ and } x \notin A\}$

or

$A \cup A^1 = U \qquad A \cap A^1 = \phi$

Associative Laws of Union and Intersection of Sets.

i) $(A \cup B) \cup C = A \cup (B \cup C)$

ii) $(A \cap B) \cap C = A \cap (B \cap C)$

Solution

i) $(A \cup B) \cup C = A \cup (B \cup C)$

L.H.S. $= (A \cup B) \cup C$

$= \{x : x \in (A \cup B) \text{ or } x \in C\}$

$= \{x : (x \in A \text{ or } x \in B) \text{ or } x \in C\}$

$= \{x : x \in A \text{ or } (x \in B \text{ or } x \in C)\}$

$= \{x : x \in A \text{ or } x \in (B \cup C)\}$

$= A \cup (B \cup C) =$ R.H.S.

ii) $(A \cap B) \cap C = A \cap (B \cap C)$

L.H.S. $= (A \cap B) \cap C$

$= \{x : x \in (A \cap B) \text{ and } x \in C\}$

$= \{x : (x \in A \text{ and } x \in B) \text{ and } x \in C\}$

$= \{x : x \in A \text{ and } (x \in B \text{ and } x \in C)\}$

$= \{x : x \in A \text{ and } x \in (B \cap C)\}$

$= A \cap (B \cap C)$ = R.H.S.

Example 4 :

Given that

A = {1, 2, 3, 6, 8}

B = {2, 3, 5, 8}

C = {3, 5, 7, 8}

Verify associative law of union and intersection of sets.

Solution

Associative law of union is

$(A \cup B) \cup C = A \cup (B \cup C)$

L.H.S. $= (A \cup B) \cup C$

$(A \cup B) = \{1, 2, 3, 5, 6, 8\}$

C = {3, 5, 7, 8}

$(A \cup B) \cup C = \{1, 2, 3, 5, 6, 7, 8\}$

R.H.S. $= A \cup (B \cup C)$

$(B \cup C) = \{2, 3, 5, 7, 8\}$

A = {1, 2, 3, 6, 8}

$A \cup (B \cup C) = \{1, 2, 3, 5, 6, 7, 8\}$

So $(A \cup B) \cup C = A \cup (B \cup C)$

Associative law of Intersection of Set

$(A \cap B) \cap C = A \cap (B \cap C)$

L.H.S. $= (A \cap B) \cap C$

$(A \cap B) = \{2, 3, 8\}$

C = {3, 5, 7, 8}

$(A \cap B) \cap C = \{3, 8\}$

R.H.S. $= A \cap (B \cap C)$

$(B \cap C) = \{3, 5, 8\}$

A $= \{1, 2, 3, 6, 8\}$

$A \cap (B \cap C) = \{3, 8\}$

So L.H.S. = R.H.S.

Distributive Law

i) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Solution

i) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

L.H.S. $= A \cap (B \cup C)$

$= \{x : x \in A \text{ and } x \in (B \cup C)\}$

$= \{x : x \in A \text{ and } (x \in B \text{ or } x \in C)\}$

$= \{x : (x \in A \text{ and } x \in B) \text{ or } (x \in A \text{ and } x \in C)\}$

$= \{x : x \in A \cap B \text{ or } x \in A \cap C\}$

$= (A \cap B) \cup (A \cap C)$

= R.H.S.

ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

L.H.S. $= \{x : x \in A\ (x \in B \text{ and } x \in C)\}$

$= \{x : (x \in A \text{ or } x \in B) \text{ and } (x \in A \text{ or } x \in C)\}$

$= \{x : x \in A \cup B \text{ and } x \in A \cup C\}$

$= (A \cup B) \cap (A \cup C)$

= R.H.S.

Example 5 :

Given the sets.

$A = \{1, 2, 3\}$

$B = \{1, 3, 5, 6\}$

$C = \{0, 3, 6\}$

Verify the distributive laws.

Solution

First rule of distributive laws.

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

L.H.S. $= A \cap (B \cup C)$

$B \cup C = \{0, 1, 3, 5, 6\}$

$A = \{1, 2, 3\}$

$A \cap (B \cup C) = \{1, 3\}$ ....(i)

R.H.S. $= (A \cap B) \cup (A \cap C)$

$(A \cap B) = \{1, 3\}$

$(A \cap C) = \{3\}$

$(A \cap B) \cup (A \cap C) = \{1, 3\}$ ....(ii)

From (i) and(ii)

L.H.S. = R.H.S.

Second rule of distributive law

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

L.H.S. $= A \cup (B \cap C)$

$B \cap C = \{3, 6\}$

$A \cup (B \cap C) = \{1, 2, 3, 6\}$ .... (i)

R.H.S. $= (A \cup B) \cap (A \cup C)$

$(A \cup B) = \{1, 2, 3, 5, 6\}$

$(A \cup C) = \{0, 1, 2\ 3, 6\}$

$(A \cup B) \cap (A \cup C) = \{1, 2, 3, 6\}$ .... (ii)

From (i) and (ii)

L.H.S. = R.H.S.

De Morgan's Laws

(i) This law states that Complement of union of two sets is equal to the intersection of complements of two sets.

i.e. $(A \cup B)^c = A^c \cap B^c$

Solution

$(A \cup B)^c = \{x \in U : x \notin A \cup B\}$

$= \{x \in U : x \notin A \text{ and } x \notin B\}$

$= \{x \in U : x \notin A^c \text{ and } x \notin B^c\}$

$= \{x \in U : x \notin A^c \cap B^c\}$

$= A^c \cap B^c$

(ii) Similarly Complement of intersection of two sets is euqal to the union of complements of two sets.

i.e. $(A \cap B)^c = A^c \cup B^c$

Solution

$(A \cap B)^c = \{x \in U : x \notin A \cap B\}$

$= \{x \in U : x \notin A \text{ or } B\}$

$= \{x \in U : x \notin A^c \text{ or } x \in B^c\}$

$= \{x \in U : x \in A^c \cup B^c\}$

$= A^c \cup B^c$

De Morgan's Law on Difference of Sets

If A, B, C are any three given sets, then

$A-(B \cup C) = (A-B) \cap (A-C)$

Solution

Let $A-(B \cup C) = \{x \in A-(B \cup C)\}$

$= \{x \in A \text{ and } x \notin (B \cup C)\}$

$= \{(x \in A \text{ and } x \notin B) \text{ and } (x \in A \text{ and } x \notin C)\}$

$= x \in (A - B) \cap (A - C)$

It means $A-(B \cup C) \subseteq (A-B) \cap (A-C)$

or we can say

$(A-B) \cap (A-C) \subseteq A-(B \cup C)$

$\therefore A-(B \cup C) = (A-B) \cap (A-C)$

Example 6:

Prove that $A - B = A \cap B^c$ and have show that

i) $A - B \cap C = (A - B) \cup (A - C)$

ii) $A - (A - B) = A \cap B$

Let $x \in (A-B)$

It means $x \in A$ and $x \in B$

or $x \in A$ and $x \in B^c$

or $x \in A \cap B^c$

(i) To show $A-(B \cap C) = (A-B) \cup (A-C)$

$= A \cap (B \cap C)^c$ [by property (1)]

$= A \cap (B^c \cup C^c)$ [De Morgan's Law]

$= (A \cap B^c) \cup (A \cap C^c)$ [Distributive law]

$= (A - B) \cup (A - C)$ [by property (1)]

(ii) To show $A - (A - B) = A \cap (A - B)^c$ [by property (1)]

$= A \cap (A \cap B^c)^c$ [by property (1)]

$= A \cap (A^c \cap (B^c)^c)$ [De Morgan's Law]

$= A \cap (A^c \cap B)$

$= (A \cap A^c) \cup (A \cap B)$ [Distributive law]

$= \phi \cup (A \cap B)$

$= (A \cap B)$

Example 7:

If A = {2, 3, 5 6}, B = {1, 3, 8}

E = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

Show that

(i) $(A \cup B)^c = A^c \cap B^c$, (ii) $(B^c)^c = B$

Solution

(i) L.H.S. $(A \cup B)^c$ $= \{1, 2, 3, 5, 6, 8\}^c$

$= \{4, 7, 9, 10, 11, 12\}$

R.H.S. $A^c \cap B^c$

A^c = {1. 4. 7. 8, 9, 10, 11, 12}

B^c = {2, 4, 5, 6, 7, 9, 10, 11, 12}

$A^c \cap B^c$= {4, 7, 9, 10, 11, 12}

Hence L.H.S. = R.H.S Proved

(ii) L.H.S. = $(B^c)^c$

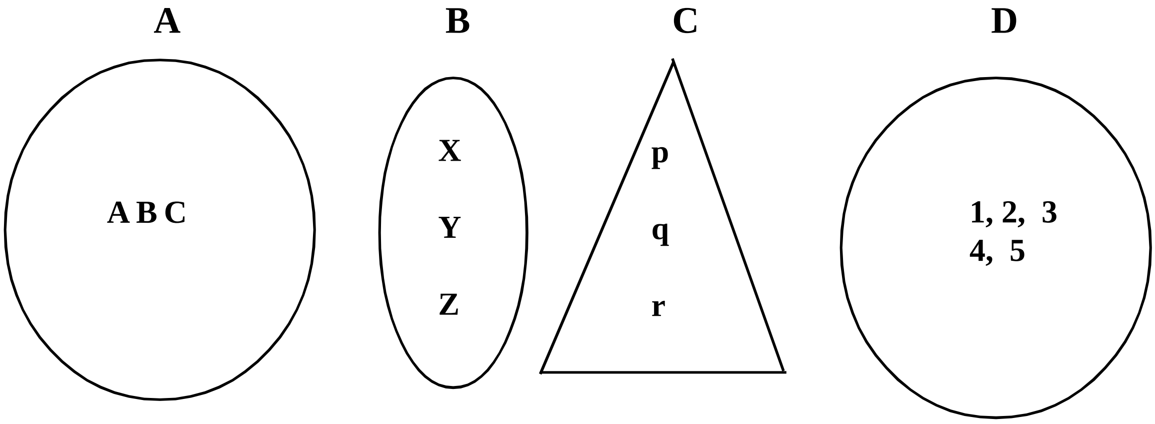
= $\{1, 2, 3, 5, 6, 7, 8, 10, 11, 12\}^c = \{1, 3, 8\} = B$

Hence L.H.S. = R.H.S.

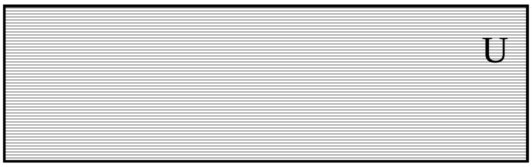
17.2 VENN DIAGRAMS

Simple closed diagrams are used to show sets. These diagrams are used first by an English Mathematician (1880) John Ven. Leward of Switzerland also used (1707 - 1783) these circle diagrams. On the name of John Venn. These simple closed diagrams are known as Venn diagrams.

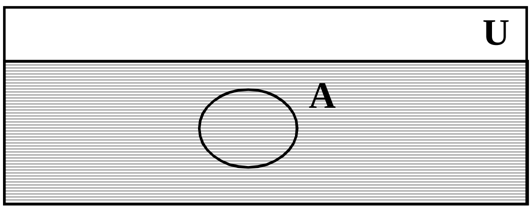
1. To show sets, the following diagrams can be used.



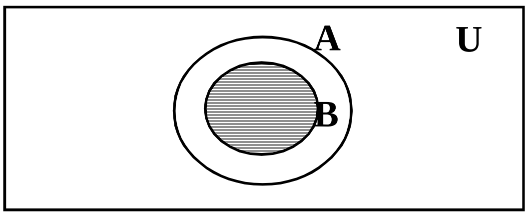
2. To the Universe set rectangle is used.



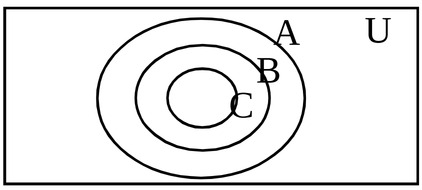
3. If $A \subset U$



4. If $U \supset A$ or $A \subset U$



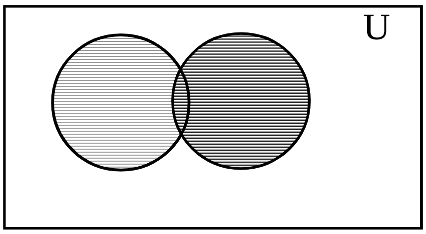
5. If $C \subset B$, $B \subset A$, $A \subset U$



Sets in Venn Diagrams

1. Union of Sets : The union of two sets A and B means set of all elements contained in A as well as in B. It is denoted by A $\cup$ B. It is showed in Venn diagram as:

Venn Diagram

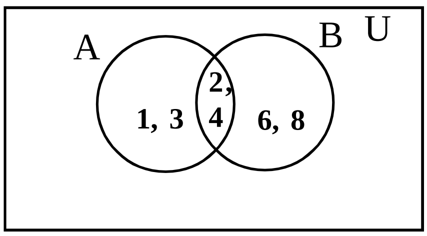


In the above diagram $\cup$ is Universal set represented by square from which two sets A and B are shown, the shaded area represent A $\cup$ B.

Example 8 : If A = {1, 2, 3, 4}

B = {2, 4, 6, 8} Find out A $\cup$ B

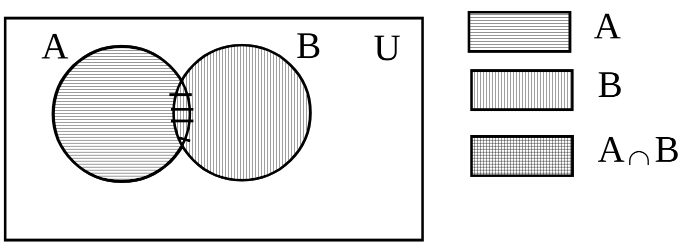
Solution : A $\cup$ B = {1, 2, 3, 4, 6, 8}



Example 9: Draw a Venn diagram to show A $\cup$ B

A = {1, 2, 3, 4} B = {1, 2, 3, 4, 5, 6, 7, 8}

Solution : A $\subset$ B, So, B is in A



2. Intersection of Sets

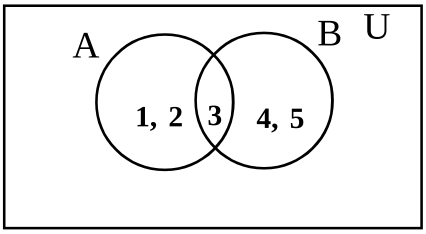
A, B are two sets, Intersection of A, B is denoted by

A $\cap$ B = {x/x $\in$ A and x $\in$ B}

It is showed in Venn Diagram.

Example 10: A = {1, 2, 3} B = {3, 4, 5} show the intersection of A and B in Venn Diagram.

Solution :



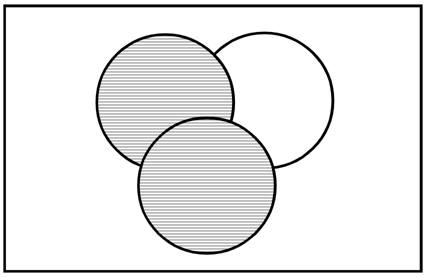
$A \cap B = \{3\}$

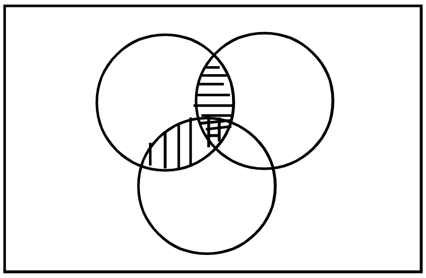
3. Distributive Laws

A, B, C are sets

i) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

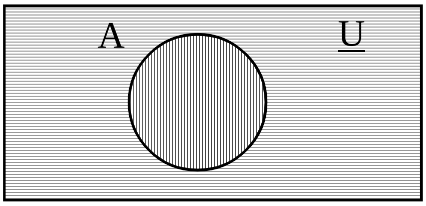
ii) Venn diagrams of $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

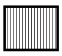
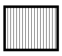




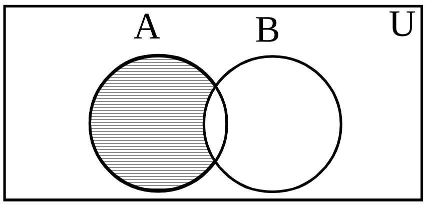
4. Complementary Sets

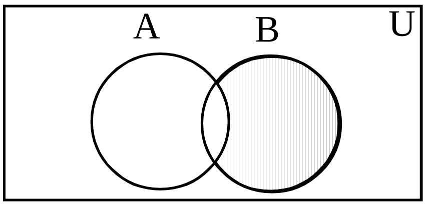
1. A

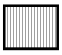
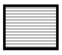


A =  A^C =  ; $A \cup A^C$ =  $\cup$  = $\cup$

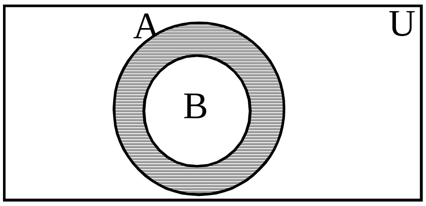
2. $A - B_1$, A - A Venn Diagram

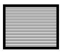




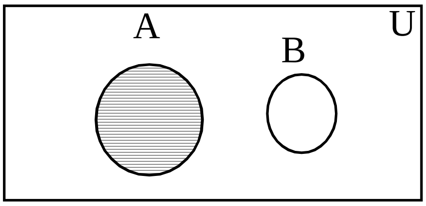
A - B =  B - A = 

3. If $B \subset A$, Venn Diagram



A - B = 

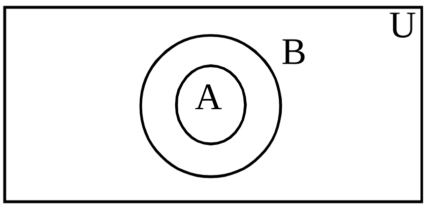
4. If $A \cap B = \phi$ or A, B's



A - B = A;

5. If A - B = ϕ ,

If $A \cap B$ A - B = ϕ



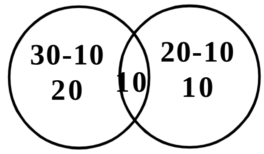
17.3 Applications of Set Theory

Example 11:

In a class of 50 student there are 30 students who play cards and 20 who play carroms. There are 10 students who play both games find number of students who play

1. only cards

2. only carroms



Cards Carroms
30 20

No. of students who play only cards - 20

No. of sudents who play only carroms - 10

Example 12:

In a survey of 100 students, it was found that 45 students studied economics 30 took up public administartion and 52 sociology, 6 students took up all the subjects, 17 studied economics and public administration, 15 took economics and sociology and 8 public administration and sociology. Findout no. of students who took.

1. Atleast one subject.

2 None of these subject.

3. Only Sociology

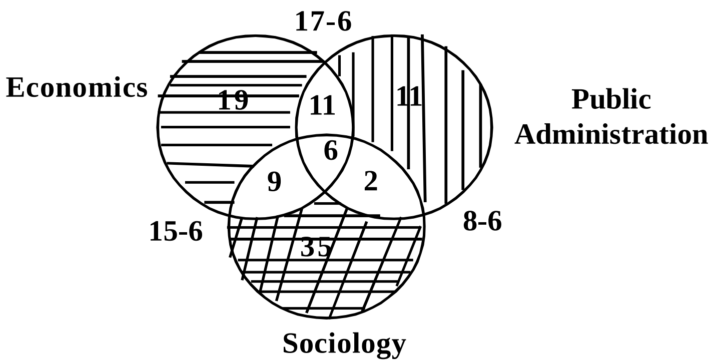
4. Only Public Administration.

5. Economics and Public Administration but not Sociology.

E = Economics

P = Public Administration

S = Sociology



$E \cap P \cap S = 6$

$E \cap P = 17$

$E \cap S = 15 - (19 + 6)$

$P \cap S = 8 - (6 + 2)$

Only Economics .... 45 11 + 9 + 6 = 19

Only Public Administration .... 30 11 + 6 + 2 = 11

Only Sociology .... 52 9 + 6 + 2 =35

Only Economics .... 45 - 26 = 19

Only Public Administration .... 30 (11 + 6 + 2) = 11

Only Sociology .... 52 (9 + 6 + 2) = 35

No. of Students studying atleast one subject = 93

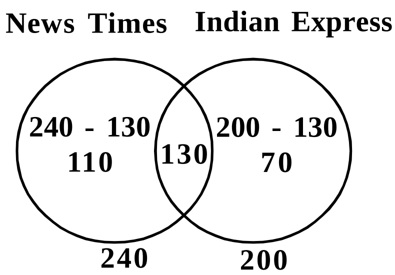
No. of Students studying none of these subjects = 100 - 93 = 7

Example 13:

In a sample survey of 400 families, 240 read the Newstimes, 200 families read the Indian Express and 130 families both the newspapers find how many families read.

1. Atleast one newspaper.

2. None of these news paper.



Total No. of families - 400

Let the number of families who read the News times = n(x)

Let the number of families who read the Indian Express = n(y)

No. of families who read both papers $n(x \cap y)$ = 168

i) No. of families who read atleast one = $n(x \cup y)$

$n(x \cup y)$ = $n(x) + n(y) - n(x \cap y)$

= 240 + 200 - 130

= 440 - 130

= 310

ii) The families who read none of two news papers

$n(U) - n(x \cup y)$

= 400 - 310

= 90

7.4 EXERCISE

1. Explain operation on Sets
2. Explain the procedure to show sets in Venn Diagram.
3. In a factory there are 100 workers, 45 workers operate on machine A while 52 workers operate machine B. There are 17 workers who can operate both machines. Find out number of workers who are operating neither of two machines.

   (Ans. : 20)
4. In a class of 40 students, 20 students have opted for Economics, 12 students have taken Civics but not Statistics. Find the number of students who have taken Economics and Statistics.

   (Ans. : 20)
5. In a survey of 200 college students, it was found that 90 take eggs, 60 take meat and 104 take fish. 34 students take both egges and meat, 30 take both eggs and fish while 16 students take meat and fish and 12 students take all three. Find

   1. The number of students who take at least one of the three things

   2. The number of students who take none of these things.

   (Ans. : 1.186, 2.14)

6. If $A = \{a, b, c\}$

   $B = \{b, c, d\}$

   $C = \{a, b\}$

   Compute

   1. $A \cup (B \cap C)$ 3. $A \cap (B \cap C)$

   2. $(A \cup B) \cap C)$ 4. $(A \cap B) \cup C$

7. If $A = \{1, 3, 5, 7\}$

   $B = \{1, 2, 3, 4, 5, 6\}$

   $C = \{5, 6, 7, 9\}$

   Compute

   i) $A \cup B$ ii) $B \cup A$ iii) $B \cup C$ iv) $C \cup A$ v) $A \cup C$

8. If $A = \{1, 2, 3, 4\}$

   $B = \{2, 4, 6\}$

   Compute

   i) A - B ii) B - A

9. If $X = \{1, 2, 3\}$ $Y = \{a, b, c\}$

   Find out X x Y

10. If $n(A \cup B) = 50$, $n(A) = 30$, $n(A \cap B) = 12$

    Findout n(B)

11. In a class of 300 students, they were given test in three subjects - economics, statistics and mathematics, 90 students failed in economics, 100 failed in statistics, 96 failed in maths, 60 failed in economics and statistics, 64 failed in statistics and maths, 70 failed in economics and maths, while 50 failed in all subjects. Find number of students who failed in at least one subject.

    (Ans. : 142)

12. Out of 300 workers in a factory, 150 workers take tea and 90 workers take tea but not coffee. Find (i) number of workers who take coffee.

    (Ans.: 150)

13. In a survey of 600 families, the following information is obtained:

    i) 360 families read Times of India.

    ii) 294 families read Indian Express.

    iii)168 families read both papers.

    Find (a) The no. of families who read atleast one newspaper

    (b) No. of families who read none of the two newspapers.

    (Ans. : 486, 114)

14. Out of 1200 students in a college, 336 played football, 360 played cricket, 504 played hockey, 96 played hockey and cricket, 120 played football and hockey, 60 played cricket and football, while 36 students played all the three games. Find :

    i) The no. of studens who played at least one game.

    ii) The no. of students who played no game.

    (Ans. : 960, 240)

15. If A and B are two subsets of a universal set U with n(U) = 500. if n(A) = 100, n(B) = 200 and $n(A \cap B) = 50$ find $n(A' \cup B')$.

    (Ans.: 250)

Dr.K.Kanaka Durga
Reader & H.O.D.
Dept. of Commerce & BBM
Hindu College, Guntur.

Lesson - 18

LAWS OF INDICES

OBJECTIVES:

By the study of this lesson, you will be able to understand, the meaning, analysis, chew digest the indices and various types of Indices.

STRUCTURE:

18.1 INTRODUCTION:

Algebra is the science of numbers where letters like, a,b,c, x,y,z.. are used to represent the numbers. Generally algerbraic methods are easier than arithmatical methods where symbols 1,2,3.... are often used. Indices are simple algebric opertaions with the help of which complex problems of solution are made easily under standable and comprehensible. The elementary knowledge of indices enable the readers to peep into the problem. Analyse, chew and digest for their pratical use.

18.2 DEFINITION :

Definition 1 :

If x is a real number and ' n ' is a positive integer, then

i x^n is defined as the product x,x,x,...... upto n factors.

ii x^{-n} is defined as reciprocal of x^n i.e. $x^{-n} = \frac{1}{x^n}$. $x \neq 0$. x^n is called "n"th power of x, n is called the index of this power.

iii x^0 is defind to be 1.

Definition 2 :

A number 'x' is called an 'm'th root of a real number 'a' if f $x^m = a$,m being a positive integer. We write it as.

$$x = a^{\frac{1}{m}} \quad \text{or} \quad x = \sqrt[m]{a}$$

When a is positive unless other wise mentioned $a^{\frac{1}{m}}$ i.e. m $\sqrt{a}$ .will mean the positive real 'm'th root of a,

For example $16^{\frac{1}{2}} = 4$, $27^{\frac{1}{3}} = 3$ and $243^{\frac{1}{5}} = 3$

When 'a ' is negative, may or may not have a real value, for example $(-1)^{1/3}$. has a real value $-1^{\frac{1}{2}}$ where as $(-1)^{1/2}$ does not have a real value. In case 'a' is negative and $a^{\frac{1}{m}}$ has a real value, then $a^{\frac{1}{m}}$ will usually stand for that real value.

For example = $(-1)^{1/3} = -1$ and $(-32)^{1/5} = -2$

Remark $a^{\frac{1}{2}}$ is usually written as $\sqrt{a}$.

Definition 3 :

If ' m ' is rational number, then $m = \frac{p}{2}$ where P and 2 are integers having no common factor and $2 \neq 0$ with out loss of generality we may suppose that 2 is positive. If x be real number then x^m is defined as $x^m = x^{\frac{p}{2}} = \left[x^{\frac{1}{2}}\right]^p = \left[x^p\right]^2$ It may be noted that $x \neq 0$ when p is negative

Remark $\left[x^p\right]^{\frac{1}{2}}$ is sometimes written as $\sqrt[2]{x^p}$

18.3 LAWS OF INDICES :

The laws governing algebraic operations are stated below -

Law 1 : When two factors with a common base are multiplied their powers are added i.e.

$x^m . x^n = x^{m+n}$

For example $x^2 . x^3 = (x.x) . (x. x. x)$

$= x.x.x .x.x = x^5 = x^{2+3}$

This is called the index law or the law of Indices. This is the fundamental law from which all other laws of indices can be derived. Therefore.

$x^P . X^2 . x^r = x^{P+2+r}$

[Since $x^P . X^2 . = x^{P+2} \therefore x^P . X^2 . x^r = x^{P+2} x^r = x^{P+2+r}$]

Law 2 : When any expression with some power is raised to any power, then the powers are multiplied i.e.

$(x^m)^n = x^{mn}$

for example $(x^2)^4 = x^2.x^2. x^2 = x^{2+2+2+2} = x^8 = x^{2x4}$

Law 3 : When two factors a common base are divided, their powers are subtracted i.e.

$\frac{x^m}{x^n} = x^{m-n}$

For examples $\frac{x^5}{x^2} = \frac{x.x.x.x.x}{x.x} = x^{5-2}$

$x.x.x = x^3$

In above put m = n , We get $\frac{x^m}{x^n} = x^{m-n}$ i.e $x^0 = 1$ (Where $x \neq 0$)

i.e. $x^0 = 1$ for all $x \neq 0$

i.e any quantity having power zero is equal to one.

Law 4 : $x^m = \frac{1}{x^{-m}}$ and $x^{-m} = \frac{1}{x^m}$

Law 5 : i) $(xy)^m = x^m . y^m$ provied x and y>0,

(ii) $\left[\frac{x}{y}\right]^m = \frac{x^m}{y^m}$

Law 6 : If $x^m = y^m$, then x = y , provided x and y>0 and $m \neq 0$ if $x^m = x^n$, then m=n provided $x \neq 1$ and x>0.

Law 7 : Meaning of $x^{\frac{1}{n}}$

Because $\left[x^{\frac{1}{n}}\right]^n = x^{\frac{1}{4} \times n} = x^1 = x$

Therefore taking 'n'th root of both sides we get

$x^{\frac{1}{n}} = \sqrt[n]{x}$

Hence $x^{\frac{1}{2}} = \sqrt{x},\ x^{\frac{1}{3}} = \sqrt[3]{x}, x^{\frac{1}{4}} = \sqrt[4]{x}$ and so on.

Law 8 : Meaning of $x^{\frac{m}{n}}$

Because $\left[x^{\frac{m}{n}}\right]^n = x^{\frac{m}{n} \times n} = x^m$

Therefore taking 'n'th root of both sides we get $x^{\frac{m}{n}} = \sqrt[n]{x^m}$

Also $x^{\frac{m}{n}} = \left[x^{\frac{1}{n}}\right]^m = \sqrt[n]{x^m}$

Note : Sign $\sqrt{\ }$ is called the redical sign and express root. Now we shall give some solved examples depending upon the laws of indices.

18.4 EXAMPLES :

Example 1 : Find the value of or evaluate $\sqrt{1\frac{9}{16}}$

Solution :

i. $\sqrt{1\frac{9}{16}} = \sqrt{\frac{25}{16}}$

$$= \left(\frac{25}{10}\right)^{\frac{1}{2}} \left(\frac{5^2}{4^2}\right)^{\frac{1}{2}}$$

$$= \left(\left(\frac{5}{4}\right)^2\right)^{\frac{1}{2}} = \left(\frac{5}{4}\right)^{2x\frac{1}{2}} = \left(\frac{5}{4}\right)^1 = \frac{5}{4} = 1\frac{1}{4}$$

Example 2 : Evaluate

Solution :

$$8^{\frac{2}{3}} = \left(2^3\right)^{\frac{2}{3}} = 2^{3x\frac{2}{3}} = 2^3 = 2^2 = 4$$

Example 3 : Evaluate

Solution :

$$\left[(243)^2\right]^{\frac{1}{5}} \left[\left(3^5\right)^2\right]^{\frac{1}{5}} = \left[3^{10}\right]^{\frac{1}{5}} = 3^{10x\frac{1}{5}} = 3^2 = 9$$

Example 4 : Evaluate

Solution :

$$\left[\frac{1}{27}\right]^{-\frac{2}{3}} = \left[\frac{1}{3^3}\right]^{-\frac{2}{3}}$$

$$= \left(3^{-3}\right)^{-\frac{2}{3}} = 3^{-3x-\frac{2}{3}} = 3^2 = 9$$

Example 5 : Simplify $\left[\frac{x^m}{x^n}\right]^l \left[\frac{x^n}{x^l}\right]^m \left[\frac{x^l}{x^m}\right]^n$

Solution :

The given expression

$$\frac{x^{lm}}{x^{nl}} \frac{x^{mn}}{x^{lm}} \frac{x^{nl}}{x^{nm}} = 1$$

Example 6 : Simplify $\frac{\left(x^{a+b}\right)^2\left(y^{a+b}\right)^2}{\left(xy\right)^{2a-b}}$

Solution :

$$\frac{\left(x^{a+b}\right)^2\left(y^{a+b}\right)^2}{\left(xy\right)^{2a-b}} = \frac{x^{2a+2b}.y^{2a+2b}}{x^{2a-b}.y^{2a-b}} = x^{3b}.y^{3b} = (xy)^{3b}$$

Example 7: If d $2x = 3y = 6^z$; prove that $z = \frac{xy}{x+y}$

Solution :

$2^x = 3^y = 6^z = k$ (say)

$2 = k^{\frac{1}{x}}.\ 3 = k^{\frac{1}{y}}.6 = k^{\frac{1}{z}}$

But 2 x 3 = 6

$k^{\frac{1}{x}}.\ k^{\frac{1}{y}}.\ k^{\frac{1}{z}}$ or $k^{\frac{1}{x}+\frac{1}{y}} = k^{\frac{1}{z}}$

or $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$

or $\frac{x+y}{xy} = \frac{1}{z}$

$z = \frac{xy}{x+y}$

Example 8 : If $a = xy^{P-1}$; $b = xy^{q-r}$; $c = xy^{r-1}$ show that $a^{q-r}.\ b^{r.p}\ .\ c^{P.2} = 1$

Solution :

L.H.S. of the result is :

$a^{2.r}.\ b^{r.p}\ .\ c^{P.2}$

Put values of a,b,c from the given relation we get

$= (xy^{P-1})^{q.r}.\ (xy^{q-1})^{r.p}\ .\ (xy^{r-1})^{P.2}$

$= x^{q-r}\,(y)^{(q-r)\,(p-1)}\ x^{r-p}\ (y)^{(q-1)\,(r-p)}\ x^{p-q}(y)^{(r-1)\,(p-q)}$

$= x^{q-r+\ r-p\ +p-q}\ (y)^{pq - Pr - q+r + rq-r+p=pq+ rp-p-rp+q}$

$= x^0.y^0 = 1.1. = 1$

Example9 : Show that

$$\frac{1}{1+x^{b-a}+x^{c-a}}+\frac{1}{1+x^{a-b}+x^{c-b}}+\frac{1}{1+x^{a-c}+x^{b-c}}=1$$

Solution :

The given expression

$$=\frac{1}{x^{a-a}+x^{b-a}+x^{c-a}}+\frac{1}{x^{b-b}+x^{a-b}+x^{c-b}}+\frac{1}{x^{c-c}+x^{a-c}+x^{b-c}}$$

$$=\frac{1}{x^{-a}\left[x^a+x^b+x^c\right]}+\frac{1}{x^{-b}\left[x^a+x^b+x^c\right]}+\frac{1}{x^{-c}\left[x^a+x^b+x^c\right]}$$

$$=\frac{1}{\left[x^a+x^b+x^c\right]}\left[\frac{1}{x^{-a}}+\frac{1}{x^{-b}}+\frac{1}{x^{-c}}\right]$$

$$=\frac{1}{x^a+x^b+x^c}\left(x^a+x^b+x^c\right)=1\left[\because\ \frac{1}{x^{-a}}=x^a;\frac{1}{x^{-b}}=x^b;\frac{1}{x^{-c}}=x^c\right]$$

Example 10 :

$$\frac{3^{2m+3n}.5^{m-1}\ .\ 10^{2n+1}.\ \ 14^{m+1}}{6^{m-2}.\ 7^{m+1}\ .\ 12^{n+2}.\ \ 15^{m+2n}}=1$$

Solution :

$$\frac{3^{2m+3n}.5^{m-1}\ .\ 10^{2n+1}.\ \ 14^{m+1}}{6^{m-2}.\ 7^{m+1}\ .\ 12^{n+2}.\ \ 15^{m+2n}}=1$$

Rewriting the equation with factors

$$= \frac{3^{2m+3n}.5^{m-1} . 2^{2n+1}. 5^{2n+1}. 2^{m+1}. 7^{m+1}}{3^{m-2}. 2^{m-2}.7^{m+1}. 2^{n+2}.3^{n+2}. 3^{n+2}. 5^{m+2n}.5^{m+2n}}$$

$$= \frac{3^{2m+3n-m+2-n-2-m-2n}.2^{2n+1+m+1-m+2-n-2-n-2}}{5^{m+2n-m+1-2n-1}. 7^{m+1-m-1}}$$

$$= \frac{3 . 2}{5 . 7} = \frac{1}{1} = \frac{1}{1} = 1 \text{ proved.}$$

Example 11 : If $x^{\frac{1}{3}} + y^{\frac{1}{3}} + z^{\frac{1}{3}} = 0$; prove that $(x+y+z)^3 = 27xyz$

Solution :

Put $x^{\frac{1}{3}} = a$, $y^{\frac{1}{3}} = b$; $z^{\frac{1}{3}} = c$

a+b+c = 0

or $a^3+b^3+ c^3 = 3abc$

or $x+y+z = 3x^{\frac{1}{3}} + y^{\frac{1}{3}} + z^{\frac{1}{3}}$

Clubing both sides

$(x+y+z)^3 = 27xyz$ proved

Example 12 : If $a^x = b$; $b^y =c$; $c^z = a$ prove that xyz = 1

Solution :

Given $a^x = b$ .... (i)

$b^y = c$ ....(ii)

Since $b^y = c$

or $(a^x)^y = c$

or $a^{xy} = c$

Since $C^z = a$

$\therefore [a^{xy}]^z = a = a^1$

$\therefore$ xyz = 1 ($\because$ When bases are the same, powers are equal) proved

18.5 SUMMARY :

Indices are simple algebraic operations with the help of which complex problem of solution are made easily understandable and comprehensible.

18.6 EXERCISE :

1. Find the value of

(i) $\sqrt{2\frac{7}{9}}$ (ii) $(81)^{-\frac{3}{4}}$ (iii) $\left(\frac{1}{625}\right)^{-\frac{3}{4}}$ (iv) $(27)^{-\frac{2}{3}}$

2. Simplify $\frac{(x^{a+b})^2(y^{a+b})^2}{(xy)^{2a+b}}$

3. Evaluation $2^{2^3} \div \left(2^2\right)^3$

4. Simplify $\left(\frac{x^a}{x^b}\right)^{a^2+ab+b^2} x \left(\frac{x^b}{x^c}\right)^{b^2+bc+c^2} x \left(\frac{x^c}{x^a}\right)^{c^2+ac+a^2}$

5. Find the value of $a^{\frac{11}{16}}\left[a\left\{a\left(a^{\frac{1}{2}}\right)^{\frac{1}{2}}\right\}^{\frac{1}{2}}\right]^{\frac{1}{2}}$ if a = 49

6. If $a^x = b^y = c^z$ and $b^2 = ac$, then prove $\frac{z}{y} = \frac{1}{x} + \frac{1}{z}$

7. (a) Divide $x^5.\ y^3 + x^4y^2 + x^3y^2 + 2x^2y + xbyx^2y + x$

(b) Divide $x^{\frac{2}{3}} - y^{\frac{2}{3}}$ by $x^{\frac{1}{3}} + y^{\frac{1}{3}}$

8. Simplify $\frac{3.27^{x+1} + 27.3^{3x}}{3.3^{3x+2}\ \frac{1}{3}27^{x+1}}$

9. (a) If $x = 3^{\frac{2}{3}} + 3^{\frac{1}{3}}$ then prove that x^3-9x-12 =0

(b) If $\frac{x}{b+c-a} = \frac{y}{c+a-b} = \frac{z}{a+b-c}$ prove that (b+c) x+(c-a)y+(a-b)z=0

10. Solve that equation $5^{1+x} + 5^{1-x} = 26$

11. (a) Simplify $\left[\frac{x^{\frac{1}{3}} \text{ x } y^{-\frac{2}{3}}}{z^{\frac{1}{2}}}\right]^{\frac{1}{2}} \text{ x } \left[\frac{y^{\frac{4}{5}} \text{ x } z^{\frac{3}{7}}}{x^{\frac{3}{2}}}\right]^{\frac{2}{3}} \div \left[\frac{y^{\frac{4}{5}} \text{ x } z^{\frac{1}{7}}}{x^{\frac{10}{3}}}\right]^{\frac{1}{4}}$

(b) Simplify $\frac{2^{3^m} . 3^{2m} . 5^m . 6^m}{8^m . 9^{3m} . 10^m}$

12. Solve for x and y from the equation

(a) $2^x + 3^y = 7$ and $2^{x+2} - 3^{y-1} = 15$

(b) $5^x+3 . 5^{x+2} = 76$

(c) $16^{x+1} = \frac{64}{4^x}$

13. $x^{2n} - y^{2n}$ by $x^{2n-1} - y^{2n-1}$

14. $\frac{1}{1+x^{b-a}+x^{c-a}} + \frac{1}{1+x^{a-b}+x^{c-b}} + \frac{1}{1+x^{a-c}+x^{b-c}} = 1$

15. Solve for x given $2^{2^x} = 16^{2^{3x}}$

16. Solve for x and y the equations

(i) $3^x + 2^y = 5$ and $2^{y+4} - 3^{x+1} = 41$

(ii) $3^x 9^y = 27$ and $2^{x+1} 4^{2y-1} = 1$

T. Nageswara Rao

Vice - Principal

Hindu College, Guntur.

LESSON 19

PROGRESSIONS

19.0 OBJECTIVE

After studying this lesson you should be able to understand Arithmetic, Geometric and Harmonic Progressions.

STRUCTURE OF LESSON

19.1 Introduction

19.2 Arithmetic Progressions

19.3 Geometric Progressions

19.4 Harmonic Progressions

19.5 Exercise

19.1 INTRODUCTION

A set of quantities, called terms, which are arranged according to some definite sequence and when a sequence is placed in summation form, it is called a series. In a given series, the successive terms (leaving the first term) are obtained either.

i) by adding or subtracting a particular number to the preceding term Arithmetic progression or

ii) by multiplying or dividing by a particular number the preceding term Geometric progression.

iii) by taking reciprocals of ther terms which form an Harmonic progresson.

19.2 ARITHMETIC PROGRESSION (A.P.)

A series in which terms increase or decrease by a constant difference is called an Arithmetic Progression.

19.2.1 A sequence of numbers are said to be in A.P. if the difference of consecutive terms is constant. The constant difference is called 'Common Difference'. Common difference is denoted by 'd'. The first term of any given series is denoted by 'a'.

19.2.2 Basic Concepts

i) Series is denoted by a, a+d, a+2d, a+3d, ...

ii) n^{th} term of above A.P. is $t_n = a + (n - 1) d$

iii) Sum to "n" terms of the A.P. is

$$S_n = \frac{n}{2} [2a + (n-1). d] = \frac{n}{2} (a+l)$$

a = first term l = last term

iv) If a, x, b are in A.P. then 'X' is called Arithmetic Mean of a, b and is

$$X = \frac{a+b}{2}$$

v) If a, a_1, a_2, a_3 . . . a_n b are in A.P., then

a, a_1, a_2, a_3 . . . a_n are n A.M.'s between a and b. Their sum is

$$a_1 + a_2 + a_3 + . \quad . \quad . a_n = \frac{n(a+b)}{2}$$

vi) Three numbers in A.P. are (a-d), a(a+d)

vii) Four numbers in A.P. are

a-3d, (a-d) (a+d), a(+3d)

viii) Five numbers in A.P. are

(a-2d) (a-d) (a) (a+d) (a+2d)

Example 1 : Find the 10th term of a given A.P. 2, 4, 6 ...

Solution : Here a = 2, d = 4 - 2 = 6 - 4 = 2

T_{10} = a + ad

= 2 + 9 x 2

= 20

Thus the 10th term is 20.

Example 2: The Third term of an A.P. is 18 and the seventh term is 30. find the 20th term.

Solution :

Given T_3 = (a+2d) = 18

T_7 = (a+6d) = 30

Solving there i & ii equations we get

4d = 12

d = 3

a = 12 and

The 20th term is T_{20} = a+19d

= 12 + 19 x 3 = 69

Example 3 :

Find the sum of 12 terms of an A.P., whose first term is 100 and common difference is - 10

Solution : To find S_{12}, Given a = 100, d = -10, n = 12

$S_{12} = \frac{n}{2}$ [2a + (n-1) d]

$S_{12} = \frac{12}{2}$ [2 x 100 + (12 - 1) (-10)]

= 6 [200 - 110]

= 6 (90) = 540

Therefore the sum of 12 terms = 540

Example 4 : A man borrows Rs, 1,000 and agrees to pay back with a total interest of Rs. 140 in 12 instalemnts, each instalment being less than the immediately preceding one by Rs. 10 what should be his first instalement.

Solution :

Borrowed Amount = Rs. 1000

Interest to be paid = Rs. 140

Total sum to be paid = s = Rs. 1140

Total number of instalments = 12

n = 12

each instalment is less than the preceding instalment by 10

d = -10

To find first instalment

i.e. a = ?

It is a problem of sum of terms in A.P.

$S = \frac{n}{2}$ [2a + (n-1) d]

$1140 = \frac{12}{2}$ [(2a + (12 - 1) (-10)]

or 2a + (-110) = $\frac{1140}{6}$

12a = 1800

a = 150

The first instalment is Rs. 150

19.3 GEOMETRIC PROGRESSION (G.P.)

A series is said to be a Geometric Progression when the ratio of any term to the preceding one is constant throughout . This Ratio is commonly known as common ratio and is denoted by ‘r’.

The first term of any given series is denoted ‘a’ and the common ratio by ‘r’. In this case, the series in G.P. becomes a, ar, ar^2, ar^3, . . .

19.3.1 Basic Concepts

i) nth term of a G.P. is $t_n = a(r^{n-1})$

ii) Sum to ‘n’ terms of a G.. is $S_n = \frac{a\, r^n - 1}{r - 1}$ if r > 1

$= \frac{a\, r^n - 1}{r - 1}$ if r < 1

= na if r = 1

iii) Infinite G.P. = sum to infinite terms of a G.P. exist if | r | < 1 and $S_\alpha = \frac{a}{1-r}$

iv) If a, b are in G.P. then X is called Geometric Mean of a, b and X = $\sqrt{ab}$

v) If $a_1, x_1, x_2, x_3, \ldots x_n$ b are in G.P. then $x_1, x_2, x_3, \ldots x_n$ are in G.M.’s between a, b.

And their product is $x_1, x_2, x_3, \ldots x_{n} = \sqrt[n]{ab}$

vi) Three numbers in G.P. are

$\frac{a}{r}$, a, ar

vii) Four numbers in G.P. are $\frac{a}{r^3}.\frac{a}{r}.ar, ar^3$

viii0 Five numbers in G.P. are $\frac{a}{r^2}.\frac{a}{r}.a,\ ar\ ar^2$

Example : Find the 7th and 11th terms of the series, 3, 9, 27, 81, . . .

Solution :

The given series is G.P.

Where a = 3, r = $\frac{9}{3} = \frac{27}{9} = \frac{81}{27} = 3$

$T_7 = a\, r^6 = 3 \times 3^6 = 3^7$

$T_{11} = a\,r^{10} = 3 \times 3^{10} = 3^{11}$

Example 2 : Find the first term, the common ratio and the series when the thrid term of a G.P. is 3 and the 6th term is 81.

Solution :

Given $T_3 = a\,r^2 = 3$

$T_6 = a\,r^5 = 81$

Dividing (ii) by (i) we get $\frac{a\,r^5}{a\,r^2} = \frac{81}{3}$ or $r^3 = 27 = 3^3$ or $r = 3$

If r = 3 from (i) $a = \frac{3}{3^2} = \frac{1}{3}$

$\therefore$ The first term $T_1 = a = \frac{1}{3}$ and Common Ratio r = 3

The series in G.P. is a, ar, ar^2, . . .

or $\frac{1}{3}, \frac{1}{3} \times 3, \frac{1}{3} \times 3^2, \ldots$

or $\frac{1}{3}$, 1, 3, 9, . . .

Example 3 : Which term of the series 2, 4, 8, . . . is 2048?

Solution : In the given series 2, 4, 8, . . .

a = 2, r = 2, $T_n = 2048$

But $T_n = a\,r^{n-1}$

putting values for a, r and T_n we get

$2048 = 2 \times 2^{n-1}$

or $2^{n-1} = 1024 = 2^{10}$

n - 1 = 10 or n = 11

$\therefore$ 11th term of the given series is 2048.

Example 4 : How many terms of the series 1+ 3+ 9+ 27+ . . . will sum up to 9841 ?

Solution : In the given series

a = 1, r = 3 and $S_n = 9841$. Here r > 1

$S_n = \frac{a\,r^{n-1}}{r-1}$

Putting values we get

$9841 = \frac{1(3^{n-1})}{3-1}$

or $3^{n-1} = 19682$

$3^n = 19683 = (3)^9$

$\therefore n = 9$

$\therefore$ The sum of 9 terms of the series 1 + 3 + 9 + 27 . . . will be 9841.

Example 5 : If Rs. 100 was invested at 12% compound interest.

Total amount at the end of first year = $100 + 100 \times \frac{12}{100} = 100(1+\frac{12}{100})$

Total amount at the end of 2nd year = $100(1+\frac{12}{100}) + 100\ (1+\frac{12}{100}) \times \frac{12}{100}$

$= 100\ (1+\frac{12}{100})\ (1+\frac{12}{100})$

$= 100\ (1+\frac{12}{100})^2$

Like that at the end of 3rd year = $100(1+\frac{12}{100})^3$

19.4 HARMONIC PROGRESSION (H.P.)

A series is said to be in Harmonic Progression if the reciprocals of its terms form on A.P. It is briefly denoted by the word H.P.

19.4.1 Basic Concepts

i) General H.P. is $\frac{1}{a}, \frac{1}{a+d}, \frac{1}{a+2d}, \ldots$

ii) nth term of a H.P. is

$t_n = \frac{1}{a+(n-1)d}$

iii) Sum to an terms of a H.P. does not exist.

iv) If a, x, b are in H.p. then X is called the

Harmonic Mean of a, b, And X = $\frac{2ab}{a+b}$

v) If a, x_1, x_2, x_3, . . . x_n b are in H.M. then x_1, x_2, x_3, . . . x_n are n H.M. is between a and b.

Example 6: Show that $\frac{1}{3}, \frac{1}{7}, \frac{1}{11}, \ldots$ are in H.P. and find 15th term of this H.P.

Solution : Given sequence is $\frac{1}{3}, \frac{1}{7}, \frac{1}{11}, \ldots$

Their Reciprocals are 3, 7, 11 . . . which are in A.P. with first term a = 3, d = 4.

Given numbers $\frac{1}{3}, \frac{1}{7}, \frac{1}{11}, \ldots$ are in H.P.

15th term in A.P. = a + 14d

= 3 + 14 x 4 = 3 + 56 = 59

$\therefore$ 15th term of H.P. = $\frac{1}{59}$

Example 7: Show that $\frac{1}{2},\frac{2}{5},\frac{1}{3},\frac{2}{7}\ldots$ are in H.P. and find 10th term of this H.P.

Solution : Given numbers are $\frac{1}{2},\frac{2}{5},\frac{1}{3},\frac{2}{7}\ldots$

Their reciprocals are $2,\frac{5}{2},\frac{3}{1},\frac{7}{2}.$

$t_2 - t_1 = \frac{5}{2} - 2 = \frac{1}{2}$; $t_3 - t_2 = 3 - \frac{5}{2} = \frac{1}{2}$

$\therefore$ 2, $\frac{5}{2}$, 3, $\frac{7}{2}$ . . . are in A.P.

$\therefore \frac{1}{2},\frac{2}{5},\frac{1}{3},\frac{2}{7}\ldots$ are in H.P.

10th term of A.P. = a + 9d

= 2 + 9 x $\frac{1}{2}$

$= 2 + \frac{9}{2} = \frac{13}{2}$

$\therefore$ 10th term of H.P. = $\frac{2}{13}$

Example 8 : Insert 2 H.M.'s between $\frac{1}{3}, \frac{1}{13}$.

Solution : Suppose x_1, x_2 are two H.M.'s between $\frac{1}{3}, \frac{1}{13}$

$\frac{1}{3}, x_1, x_2, \frac{1}{13}$

$3 \; \frac{1}{x_1} \; . \; \frac{1}{x_2}$, 13 are in H.P.

13 = 4th term of A.P.

= a + 3d

13 = 3 + 3. d

10 = 3d

d = $\frac{10}{3}$

$\frac{1}{x_1} = a + d = 3 + \frac{10}{3} = \frac{19}{3}$

$x_1 = \frac{3}{19}$

$\frac{1}{x_2} = \frac{1}{x_1} + d = \frac{19}{3} + \frac{10}{3} = \frac{29}{3}$

$x_2 = \frac{3}{29}$

$\therefore \frac{3}{19} \; . \; \frac{3}{29}$ are 2 H.M.'s between $\frac{1}{3}, \frac{1}{13}$.

19.5 EXERCISE

1. The sum of the three consecutive numbers in A.P. is 18 and their product is 192. Find the numbers.

(Ans. : 86.4 are in A.P.)

2. The third term of an A.P. is 18, and seventh term is 30. Find the 20th term.

(Ans.: -20)

3. Find the sum of 35 terms of the series in A.P. whose pth term is $\left\{\frac{P}{7}+2\right\}$.

(Ans. : 160)

4. A man borrows Rs. 840 and agrees to repay with a total interest of Rs. 240 in 12 instalments, each instalment being less than the preceding one by Rs. 8. What should be his first instalment.

(Ans. Rs.134)

5. Which term of the series $\frac{1}{128}, \frac{1}{64}, \frac{1}{32} \ldots$ is 1 ?

(Ans. : 9th term)

6. Find the sum of the series 2 + 4 + 8 + . . . to 10 terms.

(Ans. : 2046)

7. Find the sum of the series in G.P.

$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \ldots$ to 10 terms

(Ans. : $\frac{1023}{1536}$)

8. Find the 9th term of the H.P., 6, 4, 3 . . .

9. A person purchases a T.V. Set for Rs,. 3,200. Its life is estimated to be 50 years. Its price after 40 years is Rs. 640 only. Assuming the yearly depreciation to be at a constant rate, find the annual depreciation and its price ater 30 years.

(Ans. : Rs.64, Rs. 1280)

10. A man borrows Rs, 5115 to be paid in 10 montly instalments. If each instalment is double the value of the peceeding one, Find the value of the first and last instalments.

(Ans. : Rs. 5 and Rs. 2560)

- Dr. K.Kanaka Durga

Lesson - 20

MATRICES - I

OBJECTIVES:

By the study of this lesson you will be able to understand meaning the and definition of Matrices, Various types of matrices, operations in matrices with examples.

STRUCTURE:

20.1 Introduction

20.2 Definition of Matrices

20.3 Types of Matrices

20.4 Matrix operations

20.5 Exercises

20.6 Multiplication of Matrices

20.7 Process of Multiplication

20.8 Exercises

20.9 Summary

20.1 INTRODUCTION :

In economic analysis sets of equations show the relationship between variables. Matrix algebra which dates back to the works of Hamiton, CAYLEY and SYLESTER , provides a clear and concise notation for the formulation and solution of such problems which might be difficult to obtain with conventional algebraic notation. The techniques of matrix algebra are increasingly used in the problems of input - output analysis general equilibrium analysis, sector analysis, econometrics and mathematical economics.

20.2 MATRIX DEFINITIONS :

A system of mn elements, from a field F, arranged in the form of an ordered set of m horizontal lines (Called rows) and n vertical lines (called columns) is called an m x n matrix (to the read as m by n matrix) over F.

Note : Elements of a matrix are also called its entries. An m x n matrix is usually written as.

$C_1\ C_2\ C_3\ C_j\ C_n$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{1j} & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{2j} & a_{2n} \\ \hdashline a_{i1} & a_{i2} & a_{i3} & a_{ij} & a_{in} \\ a_{m1} & a_{m2} & a_{m3} & a_{mj} & a_{m} \end{pmatrix} \begin{matrix} R_1 \\ R_2 \\ R_i \\ R_m \end{matrix} \text{ Where each } a_{ij} \text{ E F}$$

Here ij is the entry in i th row and j th column of the matrix sign $\in$ means " Belongs to " Matrix A.

Note 1 : In short the above matrix is represented by A = [a_j] Where ivaries from 1 to m and j varies from 1 to n or simply by $[a_{ij}]_{mxn}$ for all $a_{ij} \in F$

Note 2 : It should be noted that matrix is not a number and it has got no value it is just an ordered collection of numbers arranged in the form of a rectangular array. By an ordered collection of numbers. We mean that in a matrix each number has a fixed position which cannot be altered.

Note 3 : If all the elements of a matrix are real numbers it is called a real matrix. If it consists of complex numbers, it is called a complex matrix.

Illustration :

1. $A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \\ 11 & -3 \end{bmatrix}_{3x2}$ is a matrix

of the type 3x 2 over the field C (Real numbers)

2. $B = \begin{bmatrix} 5+3i & 75 & 6 \\ -2+i & 6 & 0 \end{bmatrix}_{2x3}$ is a matrix

of the type 2x3 over the field R (Complex numbers)

20.3 TYPES OF MATRICES :

The arrangement of elements into different possibilities of ordered rows and columns given rise to different forms of matrices. The main important types of matrices are -

3.1 Square Matrix :

A matrix in which the number of rows are equal to the number of columns is called a square matrix. For example the matrix $A = [a_{ij}]_{mxn}$ Where m = n (m) denotes number of rows n number of columns is called a square matrix of order n.

Thus $A = (a_{11})$ $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

are square matrices of order 1,2,3 respectively.

3.2 Recetangular Matrix :

A matrix which is not a square matrix is called a rectangular matrix. In a rectangular matrix,

No of rows $\neq$ No. of columns i.e $m \neq n$ e.g $A = [a_{ij}]_{mn}$ is called is rectangular matrix if $m \neq n$

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 1 \end{bmatrix}_{2x3} \begin{matrix} R_1 \\ R_2 \end{matrix} \quad (\text{columns } C_1\ C_2\ C_3) \qquad \text{or} \qquad A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}_{2x3} \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix} \quad (\text{columns } C_1\ C_2)$$

are rectangular matrices of order 2x3 and 3x2 respectively.

3.3 Diagonal Matrix :

A diagonal matrix is a square matrix that has zeros every where except on the main diagonal, that is, the diagonal running from upper left of lower right.

Thus $A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}_{3x3}$

is a diagonal matrix of the order 3x3. The element a_{ij} of the matrix $A = [a_{ij}]_{mxn}$ for i = j are called diagonal elements and the line along which they lie is called the principal diagonal.

$$A = [8]_{1x1} \quad B = \begin{pmatrix} 7 & 0 \\ 0 & 5 \end{pmatrix}_{2x2} \qquad C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{pmatrix}_{3x3}$$

are examples of diagonal matrices.

3.4 Scalar Matrix :

A digonal matrix is called as scalar matrix if $= a_{11} = a_{22} = a_{33}$ i.e. a matrix in which all principal diaganol elements are equal is called a scalar matrix. In this case the element is called a scalar.

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}_{3x3} \quad \text{or } B = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}_{2x2} \quad C = [3]_{1x1}$$

Here $a_{11} = a_{22} = a_{33} = 3$. So the arrowed diagonal is a principal diagonal and 2 is called a scalar.

3.5 Identity (Unit) Matrix :

An identity or (unit) matrix is a diagonal matrix each of whose diagonal elements is postive one and is denoted by I. An nxn identity matrix is denoted by.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{3x3} \quad \text{or } B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{2x2} \quad C = [1]_{1x1}$$

are the 3x3 , 2x2 and 1x1 identity matrices respectively.

3.6 Null Matrix or Zero Matrix :

A null matrix is an mxn matrix all of whose elements are zeros. It is denoted by 0 or 0_{mxn}

$$0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{3x3} \quad 0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}_{2x2} \quad 0 = [0]_{1x1}$$

are the 3x3 , 2x2 and 1x1 identity matrices respectively.

3.7 Row Matrix :

A matrix which has only one row and any number of columns is called a row matrix e.g. a matrix of 1x n mean one row and n columns. Since it has one row only. It is called as row matrix.

$$A = \overset{\quad c_1 \quad c_2 \quad c_3}{[2 \quad 4 \quad 5]_{1x2}} R_1 \qquad B = \overset{\quad c_1 \quad c_2}{[2 \quad 4]_{1x2}} R_1$$

are examples of row matrices.

3.8 Column Matrix :

A matrix which contains only a single column and any number of rows is called a column matrix.

$$A = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix} \quad (C_1) \qquad B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{matrix} R_1 \\ R_2 \end{matrix} \quad (C_1)$$

are column matrices of the order 3x1 2x1 respectively

3.9 Transpose of a Matrix :

The transpose of a matrix A of mxn is a nxm matrix and is denoted by A^l whose rows are the colums of A and whose columns are the rows of A.

i.e If $A_{mxn} = (a_{ij})_{mxn}$ then the transpose of A is

$A^l = (a_{ij})$ nxm

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 0 & 3 & 1 \end{pmatrix} \quad \text{and} \quad A^l = \begin{pmatrix} 1 & 0 \\ -2 & 3 \\ 4 & 1 \end{pmatrix}$$

i.e Two rows of A become two columns in A^l and three columns of A become three rows of A^l .

$$A = \begin{pmatrix} 1 & -1 & 3 & 2 & 5 \end{pmatrix} \quad \text{then } A^l = \begin{pmatrix} 1 \\ -1 \\ 3 \\ 2 \\ 5 \end{pmatrix}$$

Here A matrix is order 1x5 (Row matrix) becomes 5x1 (column matrix) in A^l

Example If

$$A = \begin{pmatrix} 3 \\ 2 \\ 0 \\ -4 \end{pmatrix}_{4x1} \begin{matrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{matrix} \quad \text{then } A^l = \overset{\begin{matrix} C_1 & C_2 & C_3 & C_3 \end{matrix}}{\begin{pmatrix} 3 & 0 & -2 & -4 \end{pmatrix}_{1x4}}$$

Here A matrix is order 4x1 (column matrix) becomes 1x4 (Row matrix)

Example of $A = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{bmatrix} -1 & 4 & -5 \\ 2 & -3 & -2 \\ 1 & 5 & -4 \end{bmatrix}_{3x3} \end{array} \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix}$ then $A^I = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{bmatrix} -1 & 2 & 1 \\ 4 & -3 & 5 \\ -5 & -2 & -4 \end{bmatrix}_{3x3} \end{array} \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix}$

Here a matrix of order 3x3 (square matrix) gives A^I of order 3x3. All this shows that we can take the transpose of any matrix of any order. only thing is that the elements of first row are written in first column, element of second row in second column and so on.

Important properties of Transposed Matrices :

1. Transpose of transpose of a matrix is equal to the given matrix

   i.e. $(A^I)^I = A$

   Here the given matrix is A.

   Its transpose is A^I . Again taking its transpose i.e (A^I) will give us the given matrix i.e. A

2. Transpose of the sum of the matrices is equal to the sum of the transpose of the matrices.

   i.e $(A + B)^I = A^I + B^I$

   Let the given two matrices be A and B. Take the sum of the two i.e (A+B). Then take its transpose i.e $(A+B)^I$. This will be equal to the sum of the transposed value taken individually of the two matrices.

3. Transpose of the product of two matrices is equal to the transposes of the matrices taken in the reverse order.

   $$(AB) = B^I A^I$$

   It is to be remembered that the transpose of the product of two matrices is equal to the product of the transposed matrices A and B but the sequence is reversed i.e$B^I A^I$.

20.4 MATRIX OPERATIONS :

4.1 Addition and subtraction of Matrices :

Matrices can be added or subtracted if and only if they are of the sume order. The sum or difference of two mxn matrices is another matrix of order mxn whose elements are the sum or difference of the corresponding elements in the two matrices, thus if.

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 4 & 4 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & 6 & 5 \\ 0 & 1 & 1 \end{pmatrix}$$

the two matrices A and B are conformable for addition or subtraction as both are of the order 2x3. The new matrix C = A + B will be

$$C = A+B = \begin{pmatrix} 2+(-1) & 3+6 & 1+5 \\ 1+0 & 4+1 & 3+1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 9 & 6 \\ 1 & 5 & 4 \end{pmatrix}$$

and $$C = A-B = \begin{pmatrix} 2-(-1) & 3-6 & 1-5 \\ 1-0 & 4-1 & 3-1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & -3 & -4 \\ 1 & 3 & 2 \end{pmatrix}$$

Note : When two m atrices are of the same order and if each element of matrix A is exactly equal to the corresponding elements of matrix B, then the two matrices A and B are called equal matrices.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

is a case of equal matrices as a_{11} of A = b_{11}, a_{22} or A = b_{22} of B, a_{23} = b_{23} and so on. The sum or difference in this case will be

$$A+B = \begin{pmatrix} 1+1 & 2+2 & 3+3 \\ 4+4 & 5+5 & 6+6 \end{pmatrix} \text{ and } B = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$$

and $$A-B = \begin{pmatrix} 1-1 & 2-2 & 3-3 \\ 4-4 & 5-5 & 6-6 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Example 1

Given $$\begin{pmatrix} x+y-1 & z-t+3 \\ x-y+1 & z+t-3 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix}$$

Find x,y,z and t

Since the two matrices are given to be equal therefore each element of the first is equal to the corresponding element of the other. From definition, therefore.

x+y-1 = 2   z-t+3 = 4

x-y+1=3   z+t-3 = 5

Adding	Adding
2x = 5	2x = 9
x = 5/2	z= 9/2
Subtracting	Subtracting
2y-2 = -1	-2t+6 =-1
2y = 1	2t = 7
y = 1/2	+=7/2

Example 2

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 5 & 6 & 8 \\ 3 & 2 & -2 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

compute (A+B) and (A-B)

Solution :

$$A+B = \begin{pmatrix} 3+1 & 2+2 & 4+3 \\ 5+2 & 6+3 & 8+1 \\ 3+3 & 2+1 & -2+2 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 7 \\ 7 & 9 & 9 \\ 6 & 3 & 0 \end{pmatrix}$$

$$A-B = \begin{pmatrix} 3-1 & 2-2 & 4-3 \\ 5-2 & 6-3 & 8-1 \\ 3-3 & 2-1 & -2-2 \end{pmatrix} \text{ and } B = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 3 & 7 \\ 0 & 1 & -4 \end{pmatrix}$$

Example 3

Obtain the matrix resulting from the following operation.

$$\underset{A}{\begin{pmatrix} 2 & -3 & 6 \\ 5 & 4 & 5 \\ 0 & -1 & -9 \end{pmatrix}} - \underset{B}{\begin{pmatrix} 1 & -3 & 4 \\ 0 & -2 & 5 \\ 1 & 0 & -1 \end{pmatrix}} = \underset{A-B}{\begin{bmatrix} 2-1 & -3(-3) & 6-4 \\ 5-0 & 4-(-2) & 5-5 \\ 0-1 & -1-0 & -9+1 \end{bmatrix}}$$

Then $A\text{-}B = \begin{bmatrix} 1 & 0 & 2 \\ 5 & 6 & 0 \\ -1 & -1 & -8 \end{bmatrix}$

Example 4

Solve $\underset{A}{\begin{pmatrix} 6 & -1 & 0 \\ 4 & 2 & -1 \end{pmatrix}} + \underset{B}{\begin{pmatrix} 5 & 0 & 2 \\ 0 & -1 & 3 \end{pmatrix}} + \underset{C}{\begin{bmatrix} -2 & -1 & -3 \\ -4 & 1 & -1 \end{bmatrix}}$

Then $A+B+C = \begin{bmatrix} 6+5+-2 & -1+0-1 & 0+2-3 \\ 4+0-4 & 2-1+1 & -1+3-1 \end{bmatrix} = \begin{bmatrix} 9 & -2 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

Example 5

Solve x from the following relation

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{Here} \quad x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 2 \end{bmatrix}$$

Example 6

Prove that

$$\begin{bmatrix} 2 & 3 \\ 6 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 6 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix}$$

Applying addition and subtraction formula as each matrix is of the order 2x2 for which addition or subtraction is conformable.

$$\text{L.H.S.} \begin{bmatrix} 2 & 3 \\ 6 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 6 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 2+1-0 & 3+1-0 \\ 6+(-1)-6 & 4+2-4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix} \text{R.H.S.}$$

20.5 EXERCISE (A) :

1. (a) Define a matrix and give its four different types with examples.

   (b) What is a matrix ? Explain matrix operations with suitable examples.

2. Judge the types of each of the matrix.

   (a) $\begin{bmatrix} 3 & 0 & 1 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$

   (c) $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & -4 & 5 \\ 4 & 0 & 1 \\ 5 & -1 & 0 \end{bmatrix}$

   (e) $\begin{bmatrix} 0 & 8 & 7 \\ 8 & 1 & 5 \\ 7 & 5 & 3 \end{bmatrix}$ (f) $A = \begin{bmatrix} 8 & 7 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 8 & 7 & 1 \end{bmatrix}$

3. In matrix 2(d) above find a_{23}, a_{31}, a_{32}, a_{11}, a_{13}.

4. Find the transpose of the following matrices.

   (a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2x2}$ (b) $\begin{bmatrix} 5 & -2 & 1 \\ 9 & 7 & 5 \\ -6 & 8 & 0 \end{bmatrix}_{3x3}$

5. Given $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $B = \begin{bmatrix} 3 & 1 \\ 4 & 5 \end{bmatrix}$ find $(A+B)^{\prime}$

6. Given $\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$ $B = \begin{bmatrix} 3 & 1 \\ 2 & 5 \end{bmatrix}$ find $A^{\prime}+B^{\prime} = (A+B)^{\prime}$

ANSWERS (A) :

2. (a) Row Matrix (b) Square Matrix
(c) Column Matrix (d) Skew symmetric
(e) Symmetric Matrix (f) equal Matrix

3. $a_{23} = 1$, $a_{31} = 5$, $a_{32} = -1$, $a_{11} = 3$, $a_{13} = -5$.

4.

(a) $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 9 & -6 \\ -2 & 7 & 8 \\ 1 & 5 & 0 \end{bmatrix}$

5. $\begin{bmatrix} 4 & 7 \\ 3 & 9 \end{bmatrix}$

6. $\begin{bmatrix} 5 & 6 \\ 4 & 10 \end{bmatrix}$

20.6 MULTIPLICATION OF MATRICES :

Two matrics can be multiplied if and only if the number of columns in one matrix is equal to the number of rows in the other matrix. Take two matrices A and B. Then the matrix product AB is defind if and only if the number of columns in A is the same as the number of rows in B. In this case the matrices A and B are said to be conformable for multiplication (means multiplication is possible) and the product matrix will have the same number of rows as A and the same number of columns as B.

Thus if matrix A is of the other mxn and matrix B of nxp, then the product AB matrix will be of the order mxp

$$(A)_{mxn}\ (B)_{nxp} = (AB)_{mxP}$$

Here matrix A is called the pre - factor and matrix B as the post - factor. In matrix A or in Pre-factor the number of columns = n and in matrix B, the Post - factor, the number of rows = n. Since the number of columns in the Pre - factor = no.of rows in the post factor, therefore AB is conformable for multiplication. Product AB will be of the order of rows of matrix A and columns of matrix B.

In matrix multiplication the sequence in which multiplication is performed is very important. If matrix A is mxn and B is nxm, then it is possible to obtain both the product matrices AB and BA, as is evident from below.

If $(A)_{mxn} (B)_{nxm}$ then AB is conformable because the number of columns in matrix A is equal to n and the number of rows of matrix B are also n. In this case product AB will be of the order mxn. Again if we want to see whether product BA is conformable. Then we are to see whether the number of columns of B equals number of rows of A.

In the above case

$$\text{i.e } (B)_{nx\,m} (A)_{mxn} = (BA)_{nxn}$$

There fore in such a case both the product matrices are obtained How ever, in general $AB \neq BA$

Note : Even in case that both AB and BA are defined, AB and BA will give different results. In case of numbers.

$$\text{i.e} \quad 2x3 = 3x2$$

is true, but in case of matrices that AB = BA will be true, is wrong.

Example 1 . IF

$$AB = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{bmatrix} 1 & 3 & -1 \\ 2 & 0 & 0 \\ 0 & -1 & 6 \end{bmatrix}_{3x3} \end{array} \begin{matrix} R_1 \\ R_2 \end{matrix} \quad \begin{array}{c} \begin{matrix} C_1 & C_2 \end{matrix} \\ \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 1 & 3 \end{bmatrix}_{3x2} \end{array} \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix}$$

Find AB and BA

Since matrix A contains 3 columns and matrix B contains 3 rows it means AB is conformable and will be of 3x2 order.

$$AB = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{bmatrix} 1 & 3 & -1 \\ 2 & 0 & 0 \\ 0 & -1 & 6 \end{bmatrix}_{3x3} \end{array} \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix} \quad \begin{array}{c} \begin{matrix} C_1 & C_2 \end{matrix} \\ \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 1 & 3 \end{bmatrix}_{3x2} \end{array} \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix}$$

20.7 PROCESS OF MULTIPLICATOIN :

Pick up first row of matrix A and place it on first column of matrix B, multiply each element of the first row of matrix A to each element of column of B and add, then will give us the first element of the product matrix. For the second element (a_{12}) of the product matrix. Place first row of matrix. A on second columns of B, Multiply the corresponding elements and add them for the second row element of product matrix. Place second row of A on first and second columns of B, multiply the corresponding elements and add which will give the second row element of AB and so on.

$$AB = \begin{bmatrix} (1 \quad 3 \quad -1)\begin{bmatrix}1\\-1\\1\end{bmatrix} & (1 \quad 3 \quad -1)\begin{bmatrix}0\\2\\3\end{bmatrix} \\ (2 \quad 0 \quad 0)\begin{bmatrix}1\\-1\\1\end{bmatrix} & (2 \quad 0 \quad 0)\begin{bmatrix}0\\2\\3\end{bmatrix} \\ (0 \quad -1 \quad 6)\begin{bmatrix}1\\-1\\1\end{bmatrix} & (0 \quad -1 \quad 6)\begin{bmatrix}0\\2\\3\end{bmatrix} \end{bmatrix}$$

$$AB = \begin{bmatrix} 1x1+3x(-1)+(-1)(1) & 1x0+3x2+(-1)(3) \\ 2x1+0x(-1)+0(1) & (2)(x)+(0)(2)+(0)(3) \\ (0)(1)+(-1)(-1)+6(1) & 0x0+(-1)(2)+6x3 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1-3-1 & 0+6-3 \\ 2+0+0 & 0+0+0 \\ 0+1+6 & 0-2+18 \end{bmatrix}\begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix} = \begin{matrix} C_1 \quad C_2 \\ \begin{bmatrix} -3 & 3 \\ 2 & 0 \\ 7 & 16 \end{bmatrix} \end{matrix}$$

For BA $\quad B = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 1 & 3 \end{bmatrix}_{3x2} \quad A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 0 & 0 \\ 0 & -1 & 6 \end{bmatrix}_{3x3}$

For BA to be conformable for multiplication number of columns of B are not equal to number of rows of A. Therefore BA is not defined.

Example 2 :

$$A = \begin{bmatrix} 5 & -6 \\ -1 & 0 \\ 0 & 3 \end{bmatrix}_{3x2} \quad B = \begin{bmatrix} -1 & 8 & -3 \\ 0 & 10 & -4 \end{bmatrix}_{2x3}$$

Find AB and BA.

AB is conformable i.e No. of columns of A = No. of Rows of B .

$$AB = \begin{bmatrix} 5x(-1)+(-6)x0 & 5x8+(-6)(10) & 5(-3)+(-6)(-4) \\ (-1)(-1)+0x0 & -1x8+0(10) & (-1)(-3)+0x(-4) \\ 0x(-1)+3x0 & 0x8+3x10 & 0x(-3)+3(-4) \end{bmatrix}$$

$$= \begin{bmatrix} -5 & -20 & 9 \\ 1 & -8 & 3 \\ 0 & 30 & -12 \end{bmatrix}_{3x3}$$

BA is also conformable

Because number of column of B = number of rows of A.

for BA

$$B = \begin{bmatrix} -8 & 8 & -3 \\ 0 & 10 & -4 \end{bmatrix}_{2x3} \quad A = \begin{bmatrix} 5 & -6 \\ -1 & 0 \\ 0 & 3 \end{bmatrix}_{3x2}$$

$$BA = \begin{bmatrix} -1x5+8x(-1)+(-3)(0) & -1x(-6)+8x0+(-3)(3) \\ 0x5+10(-1)+(-4)(0) & 0x(-6)+10x0+(-4)(3) \end{bmatrix}$$

$$= \begin{bmatrix} -13 & -3 \\ -10 & -12 \end{bmatrix}_{2x2}$$

Example 3 :

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{3x3} \quad B = \begin{bmatrix} 3 & 4 & -4 \\ 2 & 1 & 2 \\ 3 & -1 & 3 \end{bmatrix}_{3x3}$$

Find AB and show that the product by an identity matrix A [Contrated by I] reproduces B matrix

Solution :

Poduct AB is conformable

$$\therefore AB = \begin{bmatrix} 1x3+0x2+0x3 & 1x4+0x1x(-1) & 1x-4+0x2+0x2 \\ 0x3+1x2+0x3 & 0x4+1x1+0x(-1) & 0x(-4)+1x2+0x2 \\ 0x3+0x2+1x3 & 0x4+0x1+1x(-1) & 0x(-4)+0x2+1x2 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 4 & -4 \\ 2 & 1 & 2 \\ 3 & -1 & 2 \end{bmatrix} = B$$

Matrix A is identiy matrix by definition . Its product with B reproduces the matrix B.

20.8 EXERCISE (B) :

1. Find 2A ; 3B ; A+B , A-B from the following matrices.

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 6 & 1 \\ 3 & 5 & 2 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 7 & 3 \\ 2 & 8 & 4 \\ -3 & -1 & 0 \end{bmatrix}$$

2. If $A = \begin{bmatrix} 1 & 2 & -2 \\ 3 & 0 & 5 \end{bmatrix}$ find 6A ; $\frac{3}{5}$A.

3. Find AB and BA when

(i) $A = \begin{bmatrix} -1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 2 & 3 \end{bmatrix}$

(ii) $A = \begin{bmatrix} 2 & -1 & 3 \\ -3 & 2 & 0 \\ 5 & 1 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 2 & -1 \\ 0 & 5 & 2 \\ 1 & -2 & 1 \end{bmatrix}$

Is BA = AB ? what conclusion do you draw ?

4. If $A = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}_{3x1}$ and $B = \begin{bmatrix} 3 & 1 & 0 & 2 \end{bmatrix}_{1x4}$

Find AB and BA which ever exist.

5. If $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 5 \end{bmatrix}$ find AB and BA which ever exist.

6. Express $4\begin{pmatrix} 1 & 3 \\ 1 & -4 \end{pmatrix} - \frac{1}{2}\begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix}$ as a single matrix.

7. If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ find A^2

8. Find AB and BA (if defined) where

$A = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}_{2x2}$ $B = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & -1 \end{bmatrix}_{2x3}$

9. If $A = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 2 & 4 & -1 & 3 \end{pmatrix}$ $B = \begin{bmatrix} 2 & 1 & 0 & 3 \\ 1 & -1 & 2 & 3 \end{bmatrix}$

find a 2x4 matrix such that A-2x = 3B

10. Show $\begin{bmatrix} 7 & -11 & 16 \\ -3 & 5 & -7 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & -3 \\ 2 & 5 & 1 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

11. If $A = \begin{bmatrix} 1 & 3 & 4 \\ 5 & -2 & 1 \\ 0 & 1 & 3 \end{bmatrix}$ $B = \begin{bmatrix} 1 & -1 & 0 \\ 5 & 1 & 2 \\ 0 & 2 & -2 \end{bmatrix}$ find AB

ANSWERS (A) :

1. $\begin{bmatrix} 2 & 4 & 10 \\ 0 & 12 & 2 \\ 6 & 10 & 4 \end{bmatrix} \begin{bmatrix} -3 & 21 & 9 \\ 6 & 24 & 12 \\ -9 & -3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 9 & 8 \\ 2 & 14 & 5 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} 2 & -5 & 2 \\ -2 & -2 & -3 \\ 6 & 6 & 2 \end{bmatrix}$

2. $\begin{bmatrix} 6 & 12 & -12 \\ 18 & 0 & 30 \end{bmatrix}$; $\begin{bmatrix} \frac{3}{5} & \frac{6}{5} & -\frac{6}{5} \\ \frac{9}{5} & 0 & 3 \end{bmatrix}$

3. (i) Mulitplication is not defined

   (ii) AB $\neq$ BA mulitplication is not commutative.

4. $AB = \begin{bmatrix} 3 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 9 & 3 & 0 & 6 \end{bmatrix}_{3 \times 4}$ BA does not exist.

5. AB does not exist BA $= \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 5 \end{bmatrix}$

6. $\begin{bmatrix} 0 & 10 \\ 2 & 20 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

8. $AB = \begin{bmatrix} 2 & 7 & 3 \\ -1 & -1 & -4 \end{bmatrix}_{2 \times 3}$ BA is not defined

9. $x = \begin{bmatrix} -\frac{5}{2} & -\frac{1}{2} & 0 & -\frac{5}{2} \\ -\frac{1}{2} & \frac{7}{2} & -\frac{7}{2} & -3 \end{bmatrix}$

11. $\begin{bmatrix} 16 & 10 & -2 \\ -5 & -5 & -6 \\ 5 & 7 & -4 \end{bmatrix}$

20.9 SUMMARY :

Matrix is an arrangement of group of numbers in the form of rows and columns. The Horizontal lines are called Rows and Vertical lines are called columns.

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.

Lesson - 21

MATRICES - II

OBJECTIVES:

By the study of this lesson you will be able to understand the meaning of determinants of a Matrix, Properties of determinants, Minors, Co- factors Adjoint of a matrix, Matrix inverse in detail with examples.

STRUCTURE:

21.1 DETERMINANT OF A MATRIX :

Determinant is a scalar quantity attached to a square matrix. This with every square matrix A, there is accociated a scalar quantity which is called the determinant of A. It is denoted by

det. A or |A|

Eg : We take a square matrix of the order 2x2 i.e $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

Its determinant is defined as

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \text{or} \quad |A| = a_{11} - a_{22} - a_{12} a_{21}$$

It is the product of the elements of the principal diagonal minus the product of elements of the cross diagonal.

Similarly the determinant of the 3x3 order matrix is

$$|A| = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Expanding the determinant w.r.t first row

$$|A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22}$$

Important Note :

We can also expand the given matrix with respect to the II row, Ist column, II column or III column. Thus total number of ways in which a given matrix can be expanded is equal to the total number of rows + total columns of that square matrix.

Eg. : In case of 3x3 matrix, we can expand in 3+3 = 6 ways and so on.

21.2 RULE TO EXPANDA DETERMINANT :

The rule for the determination of |A| by elements of first row (or column) is detailed below.

" Multiply each element of the first row (or column) of the determinant by a determinant obtained by deleting from the original determinant, the row and the column to which the element belongs, the signs being taken positve and negative alternatively.

Example 1 : Find the determinant of a given matrix.

$$A = \begin{pmatrix} 2 & 4 & -5 \\ -3 & 2 & -1 \\ 0 & 4 & 6 \end{pmatrix}$$

Solution :

Expand w.r.t Ist row

$$|A| = 2\begin{vmatrix} 2 & -1 \\ 4 & 6 \end{vmatrix} - 4\begin{vmatrix} -3 & -1 \\ 0 & 6 \end{vmatrix} -5\begin{vmatrix} -3 & 2 \\ 0 & 4 \end{vmatrix}$$

$|A| = 2(12+4) - 4(-18+0) - 5(-12+0)$

$= 32+72+60 = 164$

Example 2 :

Find the determinant of a given matrix.

$$A = \begin{pmatrix} 10 & 7 & 8 \\ 5 & 5 & 4 \\ 9 & 6 & 5 \end{pmatrix}$$

Solution :

Expand the given matrix w.r.t. II row

Then

$$|A| = -5\begin{vmatrix} 7 & 8 \\ 6 & 5 \end{vmatrix} + 5\begin{vmatrix} 10 & 8 \\ 9 & 5 \end{vmatrix} -4\begin{vmatrix} 10 & 7 \\ 9 & 6 \end{vmatrix}$$

$|A| = -5(35-48) +5 (50-72) - 4(60-63)$

$= -5(-13) + 5(-22) -4(-3)$

$= 65-110 + 12 = -33$

Note : we get the same result if we expand it w.r.t I row

$$\text{e.g } A = \begin{pmatrix} 10 & 7 & 8 \\ 5 & 5 & 4 \\ 9 & 6 & 5 \end{pmatrix}$$

Expand the a matrix w.r.t. I row

$$|A| = 10\begin{vmatrix} 5 & 4 \\ 6 & 5 \end{vmatrix} - 7\begin{vmatrix} 5 & 4 \\ 9 & 5 \end{vmatrix} +8\begin{vmatrix} 5 & 5 \\ 9 & 6 \end{vmatrix}$$

$|A| = 10(25-24) -7 (25-36) + 8 (30-45)$

$= 10(1) - 7(-11) + 8(-15) = 10+77 -120 = -33$

21.3 PROPERTIES OF DETERMINANTS :

Determinants have the following peculiar properties.

Property I: Transposing rows into columns of a given determinants does not change the value of the determinant e.g.

$$A=\begin{bmatrix}1 & 3\\ 2 & 5\end{bmatrix} \text{ Its Tranpose is } |A^{l}| = \begin{bmatrix}1 & 2\\ 3 & 5\end{bmatrix}$$

The value of $|A| = 1x5-2x3 = -1$ and $|A^{l}| = 1x5-3x2 = -1$

So the value of $|A^{l}|$ and $|A|$ are the same.

Property II: If two adjacent rows (or columns) are interchanged, the value of the determinant does not change numerically but the signs change.

$$A = \begin{vmatrix}1 & 0 & 2\\ 3 & 1 & 0\\ 1 & 4 & 4\end{vmatrix}$$

Its determinant is

$$|A| = 1\begin{vmatrix}1 & 0\\ 4 & 4\end{vmatrix} - 0\begin{vmatrix}3 & 0\\ 1 & 4\end{vmatrix} +2\begin{vmatrix}3 & 1\\ 1 & 4\end{vmatrix}$$

$= 4-0 + 2x11 = 4+22 = 26$

When rows are interchanged i.e IInd row becomes row I and row I becomes row II then.

$$|A| = \begin{vmatrix}3 & 1 & 0\\ 1 & 0 & 2\\ 1 & 4 & 4\end{vmatrix} = 3\begin{vmatrix}0 & 2\\ 4 & 4\end{vmatrix} - 1\begin{vmatrix}1 & 2\\ 1 & 4\end{vmatrix} + 0\begin{vmatrix}1 & 0\\ 1 & 4\end{vmatrix}$$

$= 3x-8-[4-2] +0 = -24-2 = -26$

Here we find that sign of the deteminant value changes only.

Property III: If all the elements of row or a column are zero, then the value of the determinant is zero.

$$\text{e.g.}\quad A = \begin{vmatrix} 1 & 4 & 5 \\ 0 & 0 & 0 \\ 2 & 3 & 1 \end{vmatrix}$$

Expanding w.r.t I row we get then

$$|A| = 1\begin{vmatrix} 0 & 0 \\ 3 & 1 \end{vmatrix} - 4\begin{vmatrix} 0 & 0 \\ 2 & 1 \end{vmatrix}\begin{vmatrix} 0 & 0 \\ 2 & 3 \end{vmatrix}$$

$$= 1(0) - 4(0) + 5(0) = 0$$

Property IV : If the elements of one row (or column) of a determinant are identical or proportional to the corresponding elements of another row or column) the value of the determinant is zero.

$$\text{e.g.}\quad A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2 & 4 \\ 1 & 5 & 6 \end{bmatrix}$$

Here row I and II are identical then

$$|A| = (12-20) - 2\,(6-4) + 4\,(5-2) = -8-4+12 = 0$$

Similarly if the elements of row I are twice the elements of row II, even then the determinant is zero.

$$\text{e.g.}\quad A = \begin{bmatrix} 2 & 8 & 10 \\ 1 & 4 & 5 \\ 2 & 3 & 1 \end{bmatrix}$$

then

$$|A| = 2\begin{vmatrix} 4 & 5 \\ 3 & 1 \end{vmatrix} - 8\begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} + 10\begin{vmatrix} 1 & 4 \\ 2 & 3 \end{vmatrix}$$

$$= 2(4-15) - 8(1-10) + 10\,(3-8)$$

$$= -22 + 72 - 50 = 0$$

Property V : If each in any row (or in any column) is multiplied by a scalar quantity λ the value of the whole determinant is multiplied by λ

$$\text{e.g. } |A| = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad |A| = \lambda \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Property VI : Generalisation of property i.e if each element of a determinant of (say 3x3 order) is multiplied by a scalar quantity K, the value of new determinant so obtained is K^3 times the value of the original determinant.

$$\text{e.g. } A = \begin{bmatrix} ka_{11} & ka_{12} & ka_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{bmatrix} \quad |A| = K^3 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Property VII : The addition of a constant multiple of one row (or column) to another row (or column) leaves the determinat unchanged.

$$\text{Thus } \begin{bmatrix} a_1 + \lambda b_1 & b_1 & c_1 \\ a_2 + \lambda b_2 & b_2 & c_2 \\ a_3 + \lambda b_3 & b_3 & c_3 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

21.4 MINORS OF DETERMINANTS :

The Determinants of any sub matrix is called minor of |A|

Let us take a 3x3 matrix say

$$A = \begin{bmatrix} a_{11} & b_{12} & c_{13} \\ a_{21} & b_{22} & c_{23} \\ a_{31} & b_{32} & c_{33} \end{bmatrix}$$

There are nine elements in this matrix and each element has its minor. So in all, there are 9 minors.

Minors may be defined as determinant obtained by deleting the row and the column from the given determinant to which the element belongs.

In a given matrix A above if we delete I row and I column, we get a 2x2 matrix which is called a sub-matrix of A.

$$\text{Minor of } a_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{32}a_{23}$$

Similarly

$$\text{Minor of } a_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = a_{21}a_{33} - a_{23}a_{31}$$

$$\text{Minor of } a_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{21}a_{32} - a_{22}a_{31}$$

$$\text{Minor of } a_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = a_{12}a_{33} - a_{13}a_{32}$$

$$\text{Minor of } a_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11}a_{33} - a_{13}a_{31}$$

$$\text{Minor of } a_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = a_{11}a_{32} - a_{12}a_{31}$$

$$\text{Minor of } a_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = a_{12}a_{23} - a_{13}a_{22}$$

$$\text{Minor of } a_{32} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{23} \end{vmatrix} = a_{11}a_{23} - a_{13}a_{21}$$

$$\text{Minor of } a_{33} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Example 1 :

Calculte the minors of the elements in given matrix.

$$A = \begin{bmatrix} 1 & 4 & 7 \\ -2 & 3 & 4 \\ 1 & 4 & 4 \end{bmatrix}$$

Solution : Compare the given matrix with

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The minors of

$$a_{11} = 1 = \begin{vmatrix} 3 & 4 \\ 4 & -4 \end{vmatrix} = -4x3-4x4 = -28$$

$$a_{12} = 4 = \begin{vmatrix} -2 & 4 \\ 1 & -4 \end{vmatrix} = (-2)(-4)-1x4 = 4$$

$$a_{13} = 7 = \begin{vmatrix} -2 & 3 \\ 1 & 4 \end{vmatrix} = (-2)(4) -3x1 = -11$$

$$a_{21} = -2 = \begin{vmatrix} 4 & 7 \\ 4 & -4 \end{vmatrix} = 4x(-4)-7x4 = -44$$

$$a_{22} = 3 = \begin{vmatrix} 1 & 7 \\ 1 & -4 \end{vmatrix} = 1x(-4)-1x7 = -11$$

$$a_{23} = 4 = \begin{vmatrix} 1 & 4 \\ 1 & 4 \end{vmatrix} = 1x4-1x4 = 0$$

$$a_{31} = 1 = \begin{vmatrix} 4 & 7 \\ 3 & 4 \end{vmatrix} = 4x4 - 7x3 = -5$$

$$a_{32} = 4 = \begin{vmatrix} 1 & 7 \\ -2 & 4 \end{vmatrix} = 1x4 -(-2)\ x7 = 18$$

$$a_{33} = -4 = \begin{vmatrix} 1 & 4 \\ -2 & 3 \end{vmatrix} = 1X3 -(-2)(4) = 11$$

21.5 CO- FACTOR OF DETERMINANTS :

Co - factors of the elements of a given determinant are defind as the of minors of the elements $(-1)^{i+j}$

Where i refers to the row and j the column position of the element whose co-factor is to be determined.

e.g in a given matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The co- factor of the element a_{11} is $A_{11} = (-1)^{1+1}$ (minor of a_{11})

$$\therefore \; A_{11} = (-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

The co- factor of the element a_{12} is $A_{12} = (-1)^{1+2}$ (minor of a_{12})

$$\therefore \; A_{12} = (-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

The co- factor of the element a_{13} is $A_{13} = (-1)^{1+3}$ (minor of a_{13})

$$\therefore \; A_{13} = (-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Co - factors of $a_{21} a_{22} a_{23}$ are $a_{31} a_{32} a_{33}$ are

$$A_{21} = (-1)^{2+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$$

$$A_{22} = (-1)^{2+2} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$$

$$A_{23} = (-1)^{2+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

$$A_{31} = (-1)^{3+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

$$A_{32} = (-1)^{3+2} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}$$

$$A_{33} = (-1)^{3+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$\therefore \text{ The co - factor matrix is } \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

Example 2 :

Find the co-factors of the elements a_{31}, a_{13} in a given matrix.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Solution :

The co- factor of the element a_{31} is

$$A_{31} = (-1)^{1+3} \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = (-1)^{4} (12-15) = (1) (-3) = -3$$

The co- factor of the element a_{13} is

$$A_{13} = (-1)^{1+3} \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} (-1)^{4} (32-35) = (1) (-3) = -3$$

21.6 ADJOINT OF A MATRIX :

The transpose of a co-factor matrix is called the Adjoint of a matrix.

Steps :

1. Find the co-factor of every element of the given matrix.
2. Form a new matrix with the values of co-factors which would be of the same order as the given matrix.
3. Take the transpose of co-factor matrix.

   The result would give the adjoint of a Matrix.

$$\text{Given } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\text{The Adj } A = \text{Transpose of } \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$$

Example 3 :

Find the adjoint of the following matrix.

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$$

Solution :

Now we have to find the co-factors of all the 9 elements of the given matrix.

$$\therefore \quad a_{11} = 0 \text{ and its co- factor } A_{11} = (-1)^{1+1} \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = -1$$

$$a_{12} = 1 \text{ and its co- factor } A_{12} = (-1)^{1+2} \begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} = 8$$

$$a_{13} = 2 \text{ and its co- factor } A_{13} = (-1)^{1+3} \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5$$

$$a_{21} = 1 \text{ and its co- factor } A_{21} = (-1)^{2+1} \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = 1$$

a_{22} = 2 and its co- factor $A_{22} = (-1)^{2+2} = \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} = -6$

a_{23} = 3 and its co- factor $A_{23} = (-1)^{2+3} = \begin{vmatrix} 0 & 1 \\ 3 & 1 \end{vmatrix} = 3$

a_{31} = 3 and its co- factor$A_{31} = (-1)^{3+1} = \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = -1$

a_{32} = 1 and its co- factor $A_{32} = (-1)^{3+2} = \begin{vmatrix} 0 & 2 \\ 1 & 3 \end{vmatrix} = 2$

a_{33} = 1 and its co- factor $A_{33} = (-1)^{3+3} = \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} = -1$

$$\therefore \text{Adj } A = \text{Transpose of } \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

Putting values

$$\text{The Adj } A = \text{Transpose of } \begin{bmatrix} -1 & 8 & -5 \\ 1 & -6 & 3 \\ -1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & -1 \\ 8 & -6 & 2 \\ -5 & 3 & -1 \end{bmatrix}$$

21.7 INVERSE OF A MATRIX :

Inverse of a matrix is denoted by A^{-1}

Let A ($a_{ij})_{nxn}$ be a given square matrix (of order n). Then n-square matrix A^{-1} is called of A if

$AA^{-1} = A^{-1} . A = 1$ i.e unit or identity matrix.

The inverse of A is calculated by the following formula

$$A^{-1} = \frac{\text{Adj } A}{|A|}$$

i.e inverse of A is equal to the Adj A divided by the determinant of (A)

21.8 STEPS TO CALCULATE INVERSE OF A MATRIX :

1. Calculate the co-factors of all the elements of a given square matrix.
2. Write down a co-factor matrix
3. Take the transpose (changing rows into columns and columns into rows) of the co-factor matrix.
4. Find the determinant of the given matrix
5. Divide the value obtained in step 3 by the value obtained in step 4. The result would be the value of A^{-1}

21.9 NECESSARY CONDITIONS TO FIND A^{-1} :

i. The given matrix whose inverse is to be found out should be a square matrix.

ii. The necessary and sufficient condition for a n-square matrix to possess its inverse is that $|A| \neq 0$. In other words, the given matrix should be non - singular matrix.

That matrix is called non- singular whose determinant is not equal to Zero. i.e. $|A| \neq 0$. If $|A|=0$, then the matrix is called singular.

Example 1 :

Find the adjoint and inverse of a matrix

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

Solution :

To find the adjoint and inverse of matrix, we have to find the co-factors of all the given elements and then take the transpose of the given co-factor matrix. This result would be called adjoint of A.

$$\text{compare } A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix} \text{ with } \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Co - factors of

$$a_{11} = 1 = (-1)^{1+1} \begin{vmatrix} 4 & 3 \\ 3 & 4 \end{vmatrix} = +1 \quad [16 - 9] = 7$$

$a_{12} = 3 = (-1)^{1+2} \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = -1 [4 - 3] = -1$

$a_{13} = 3 = (-1)^{1+3} \begin{vmatrix} 1 & 4 \\ 1 & 3 \end{vmatrix} = +1 [3-4] = -1$

$a_{21} = 1 = (-1)^{2+1} \begin{vmatrix} 3 & 3 \\ 3 & 4 \end{vmatrix} = -1 [12-9] = -3$

$a_{22} = 4 = (-1)^{2+2} = \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = 1 [4 - 3] = 1$

$a_{23} = 1 = (-1)^{2+3} = \begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix} = -1 [3 - 3] = 0$

$a_{31} = 1 = (-1)^{3+1} = \begin{vmatrix} 3 & 3 \\ 4 & 3 \end{vmatrix} = +1 [9-12] = -3$

$a_{32} = 3 = (-2)^{3+2} = \begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix} = -1 [3-3] = 0$

$a_{33} = 4 = (-1)^{3+3} = \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = 1 [4 - 3] = 1$

$\therefore$ Co - factor matrix is $\begin{bmatrix} 7 & -1 & -1 \\ -3 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$ or $A^1 = \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

The transpose of Co - factors matrix is called Adj A

$\therefore$ Adj. A $= \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

For Inverse of A :

Find |A| = determinant of (A)

$$|A| = 1\begin{vmatrix} 4 & 3 \\ 3 & 4 \end{vmatrix} - 3\begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} + 3\begin{vmatrix} 1 & 4 \\ 1 & 3 \end{vmatrix}$$

= 1x7 - 3(1) + 3(-1) = 7-3-3 = 1

$$\therefore \; A^{-1} = \frac{\text{Adj } A}{|A|} \text{ or } A^{1} = \frac{\begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}}{1} = \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

20.10 EXERCISE (A) :

1. Define the following with examples -

   (a) Determinant of a matrix

   (b) Minors

   (c) Co - factors

   (d) Adjoint of a matrix

   (e) Inverse of a matrix

2. Find the determinants of the following matrices.

(a) $\begin{bmatrix} 1 & 2 & -2 \\ -1 & 3 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 0 & 1 \\ 4 & 2 & 1 \\ 9 & -3 & 1 \end{bmatrix}$ (d) $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$

(e) $\begin{bmatrix} 1 & 2 & 5 \\ 7 & 3 & 4 \\ 5 & -1 & -6 \end{bmatrix}$ (f) $\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$ (g) $\begin{bmatrix} 5 & 1 \\ 2 & 2 \end{bmatrix}$ (h) $\begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & 4 \\ 6 & 1 & 1 \end{bmatrix}$

3. Show that the detereminant of matrix

$$|A| = \begin{vmatrix} 1 & 2 & 5 \\ 2 & 3 & 5 \\ 5 & 7 & 10 \end{vmatrix} = 0 \text{ and } A = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 0$$

4. Show that

(a) $\begin{bmatrix} ad+bc & bd-ac \\ ac-bd & ad+bc \end{bmatrix}$ = $(a^2+b^2)\ (c^2+d^2)$

5. Show that |A| $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ = -1

6. Expand the given matrix with respect to the first, second and third column.

$$A = \begin{bmatrix} 0 & -1 & 2 \\ -2 & 1 & 0 \\ 3 & -3 & -1 \end{bmatrix}$$

7. Show that the determinant to the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \\ 4 & 6 & 5 \end{bmatrix} = 1$$

8. Distinguish between 'minor' and 'co-factor'

Find the co-factor of all the elements of the given matrix.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \\ 4 & 6 & 5 \end{bmatrix}$$

9. Find the adjoint of the matrix

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$$

10. Show that transpose of A is one third of its adjoint, where A is a matrix given as under.

$$A = \begin{bmatrix} -1 & -2 & -2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$$

11. Find the inverse of the following matrix.

$$A = \begin{bmatrix} 4 & 3 \\ 2 & 8 \end{bmatrix}$$

12. Find the inverse of

$$A = \begin{bmatrix} 2 & -3 & 3 \\ 2 & 2 & 3 \\ 3 & -2 & 2 \end{bmatrix}$$

13. Verify $(AB)^{-1} = B^{-1} A^{-1}$ for the matrices

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 4 & 5 \\ 3 & 4 \end{bmatrix}$$

14. Prove that det of $A = \begin{bmatrix} a & b \\ c & \frac{1+bc}{a} \end{bmatrix} = 1$

ANSWERS (A) :

2. (a) 1 (b) -2 (c) -30 (d) $a_1b_2 - b_1a_2$

(e) 0 (f) $abc-af^1-bg^2-ch^2+2fgh$ (g) 7 (h) 39

6. 8

8. $A_{11} = -3$ $A_{12} = 2$ $A_{13} = 0$

$A_{21} = 8$ $A_{22} = -7$ $A_{23} = 2$

$A_{31} = -3$ $A_{32} = 3$ $A_{33} = -1$

9. Adj, $A = \begin{bmatrix} -1 & 1 & -1 \\ 8 & -6 & 2 \\ -5 & 3 & -1 \end{bmatrix}$

11. $A^{-1} = \begin{bmatrix} \frac{4}{13} & \frac{-3}{26} \\ \frac{-1}{13} & \frac{2}{13} \end{bmatrix}$

12. $\begin{bmatrix} \frac{-2}{5} & 0 & \frac{3}{5} \\ \frac{-1}{5} & \frac{1}{5} & 0 \\ \frac{2}{5} & \frac{1}{5} & \frac{-2}{5} \end{bmatrix}$

21.11 CRAMER'S RULE :

Consider the set of three simultaneous linear equations.

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

The above equations can be written in the matrix form as Ax = B

Where $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ $B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$\therefore$ x A^{-1}B, But $A^{-1} = \frac{\text{Adj } A}{|A|}$

$$\therefore X = \frac{\text{Adj}}{|A|} B$$

Now $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

Where A_{11} is co factor of a_{ij} in |A|

$$= \frac{1}{|A|}\begin{bmatrix} A_{11}b_1 + A_{21}b_2 + A_{31}b_3 \\ A_{12}b_1 + A_{22}b_2 + A_{32}b_3 \\ A_{13}b_1 + A_{23}b_2 + A_{33}b_3 \end{bmatrix}$$

$$X_1 = \frac{A_{11}b_1 + A_{21}b_2 + A_{31}b_3}{|A|}$$

or $X_1 = \frac{1}{|A|}\begin{bmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{bmatrix}$

$$X_2 = \frac{1}{|A|}\begin{bmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{bmatrix}$$

$$X_3 = \frac{1}{|A|}\begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{bmatrix}$$

This result is popularly known as Crammer's rule.

The above given example can be solved by crammer's rule :

Solution :

$2x_1 - x_2 + 3x_3 = 9$

$x_2 - x_3 = -1$

$x_1 + x_2 - x_3 = 0$

The given equation can be written in matrix form as A X = B

Where

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad B = \begin{bmatrix} 9 \\ -1 \\ 0 \end{bmatrix}$$

|A| = z = (-1+1) +1 (0+1) +3(0-1) = -2

By cramer's Rule

$$X_1 = \frac{1}{|A|}\begin{bmatrix} 9 & -1 & 3 \\ -1 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} = \frac{1}{-2}(-2) = 1$$

$$X_2 = \frac{1}{|A|}\begin{bmatrix} 2 & 9 & 3 \\ 0 & -1 & -1 \\ 1 & 0 & -1 \end{bmatrix} = \frac{1}{-2}(-4) = 2$$

$$X_3 = \frac{1}{|A|}\begin{bmatrix} 2 & -1 & 9 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} = \frac{1}{-2}(-6) = 3$$

$\therefore X_1 = 1, X_2 = 2, X_3 = 3$

EXERCISE (B) :

1. Use the method of determinants to solve the set of equations

   3x + 2y - z = 4

   -x - y +3z = 6

   5x - 3y + x = 2

2. Solve by inverse method

   x + y+ z = 6

   2x - 5y+ 5z = 27

   2x - 5y +11z = 45

3. Has the following matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \\ 4 & 8 & 4 \end{bmatrix}$$ its inverse ? Give reasons

4. Solve by using matrix inverse method

$9x_1 + x_2 = 13$

$8x_1 + x_2 = 16$

5. Solve by Crammer's rule

$x + 6y - z = 10$

$2x + 3y + 3z = 27$

$3x - 3y - 3z = -9$

6. Use the matrix method to solve the following set of equations

$x + 3y = 7$

$4x - y = 2$

7. Solve by Cramer's rule or other wise.

$x-y-z = 1 = y-z-x = z-x-y$

8. Find x,y,z by Cramer's rule

$x + 2y - 3z = 1$

$2x + 2y + 4z = 2$

$3x + 4y + 3z = 3$

9. Use matrix method to solve the following

$2x_1 - 2x_2 + 5x_3 = 1$

$2x_1 - 4x_2 + 8x_3 = 2$

$-3x_1 + 6x_2 + 7x_3 = 1$

10. Solve for x,y,z

$2x + y = 1$

$y + 2z = 7$

$3z + 2x = 11$

11. Define the following

    (a) Singular and Non - singular Matrixs

    (b) Orthogonal Matrix.

ANSWERS (B)

1. x=1 , y = 2, z = 3

2. x = x = y = 2, z = 3

3. No, as |A| = 0

4. $X_1 = 1$, $X_2 = 4$

5. x = 1 , y = 2 , z = 3

6. x = 1 , y = 2

7. x = -1 , y = -1 , z = -1

8. x = 1 , y = 0 , z = 0

9. $x_1 = \frac{4}{19}$, $x_2 = \frac{7}{38}$, $x_3 = \frac{4}{19}$

10. $x = \frac{4}{10}$, $y = \frac{1}{5}$, $z = \frac{17}{5}$

T. Nageswara Rao
Vice - Principal
Hindu College, Guntur.