Chapter - II
DIAGRAMMATIC AND GRAPHIC PRESENTATION

Diagrammatic and Graphical Representation of Data

A picture is worth a thousand words. The impression created by a picture has much greater impact than any amount of detailed explanation. One of the important functions of Statistics is to simplify complex quantitative data and make than easily intelligible. Statistical data can be effectively presented in the form of diagrams and graphs. Graphs and Diagrams make complex data simple and easily understandable. They help to compare related data and bring out subtle difference with amazing clarity.

Importance of Diagrams

1. Diagrams reduce huge amount of data into simple figures. Thus, it is the best option for making short duration presentations to top management, Government, customer, etc who cannot go into details.
2. Diagrams bring out the essence of the underlying data with great clarity. A person going through the data may get lost in detail, but will get the message immediately if the same data is presented as a diagram
3. Diagrams help in highlighting the trends in underlying data that cannot be easily spotted by going through the data.
4. Diagrams create a positive and lasting impression. The audience can remember the details for a longer time.
5. Diagrams have the ability to reach a greater audience people who cannot read can also understand diagrams.
6. Diagrams facilitate comparison of data. Interesting conclusions can be drawn by merely juxtaposing two figures.
7. Diagrams have universal acceptance. There is not restriction of language, time or place.
8. Diagrams bring forth the characteristics of data, highlighting hidden facts and relationship.
9. Diagrams are helpful in garnering the attention of the user and create interest in the subject.
10. Diagrams save time, effort and costs. It is very economical to prepare and present diagrams.

Limitations of Diagrams

1. Diagrams are used to convey a message. The information that can be meaningfully presented by way of a diagram is limited. If too many details are packed in one diagram, have will be loss of clarity.
2. Diagrams have limited capability to highlight small differences in large measurements. For example, while plotting annual household incomes that can range from ₹10,000 to ₹10 cores, the difference between persons with annual household income of ₹ 50,000 and ₹ 1,00,000 will not get highlighted.
3. Diagrams can create visual illusions and can be miss-leading. For example, the visual impressions can be significantly influenced by use of a different scale.

4. Apart from simple diagrams, most diagrams are quite complex to draw as well as to understand. Thus, diagrams have utility only for experts.
5. A single diagram in isolation does not convey much. Diagrams are essentially tools for comparison
6. Diagrams are end products that are incapable of being further analyzed, except for mental analysis. They cannot be put through further statistical treatment.
7. Diagrams are not accurate. They disclose only approximate values. Precision gets sacrificed for want of better presentation.
8. Diagrams cannot be a substitute for tabular presentation. They can only supplement tabular presentation.

General Rules for Constructing Diagrams

1. **Title :** Each diagrams must be given a suitable title to convey the main idea it is intended to portray.
2. **Scale :** Scale should be selected consistent with the size of observations to be displayed. The diagram should be neither too small nor too large. The scale should specify the size of the unit and what it represents such as marks scored, in numbers price in rupees etc.
3. **Proportion between width and height :** A diagram should be displayed in the center of the page. It should ideally maintain a proportion of 1:1:4142 between the smaller side and the larger side respectively.
4. **Choice of Diagram :** The choice of a particular type of diagram has to be made depending on nature of data magnitude of observations and type of people who use the diagram. For example, data relating to frequency curves and time series is best represented by means of graphs.
5. **Index :** An index illustrating different types of shades, colors, lines, designs etc used for presenting the data should be provided for easy understanding of the diagram by the user.
6. **Footnotes :** Classifications regarding certain aspects of the diagram may be provided as a footnote.
7. **Source Note :** The source from where the data has been collected may be provided.
8. **Simplicity :** Diagrams are meant to make complex data simple. Hence, too much information should not be cluttered into a single diagram.
9. **Neatness :** Lastly, diagrams are visual aids and hence need to be neatly presented. A variety of devices such ad dots, broken lines etc are used to make the diagram attractive.
10. **Communicative :** Diagrams must be intelligible. They must convey a message.
11. **Self-Explanatory :** Any person looking at the diagram should be able to understand the underlying message. It must be self-explanatory.
12. **Reliable :** Diagrams should be reliable. They should not mislead the user, in an attempt to make the presentation more attractive data accuracy must be maintained.

Types of Diagrams :

Some of the most commonly used diagrams are explained hear under :

One Dimensional Diagrams

One-dimensional diagram means the construction of a diagram on the basis of one dimension. This could be height in case of vertical diagrams or length in case of horizontal diagrams. One dimensional diagrams can be line diagrams or bar diagrams.

Line Diagram

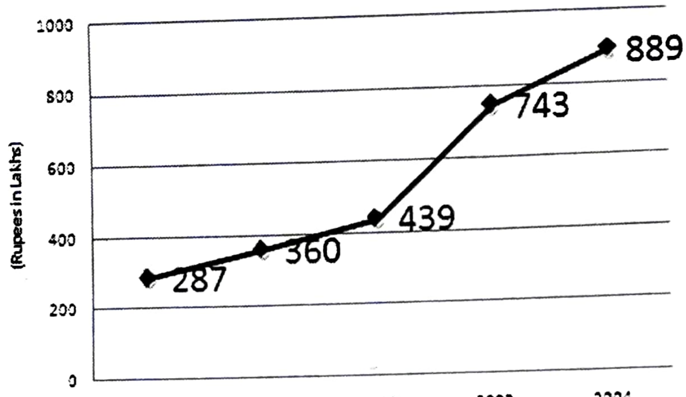
A line diagram involves drawing multiple vertical lines. Each vertical line represents a particular frequency. The variety (x) values are presented on X axis and the corresponding frequencies for each value of x are presented on Y axis. Both the axis are drawn as per a suitable scale. Let us understand the preparation with the help of an illustration.

Illustration -1 Present the following frequency distribution in the form of a line diagram

Year	2000	2001	2002	2003	2004
Sales (₹ in lakhs)	287	360	439	743	889

Solution :

X axis will represent Year and Y axis will represent sales in lakhs of rupees. One line needs to be drawn for each year in the ratio of the magnitude of the given data. Identical gap should be provided in between the different lines. The scale of one centimeter on Y-axis is equal to Rupees Two Hundred lakhs.



Bar Diagram :

Bar diagrams are one of the most commonly used diagrams for presenting data. This type of diagram is generally constructed when the data represent different values over a period or different situations. They are largely used by business and Government to present economy related data such as industrial / agricultural production inflation exports. Forex reserves etc,

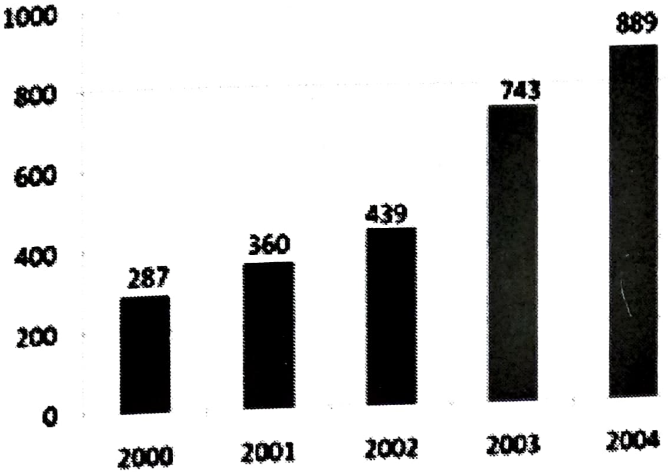
Bar diagrams are used specifically for categorical data or series. They consist of a group of equidistant rectangles. One for each group or category of data in which the value of magnitudes are represented by the length or height of the rectangles. The width of the rectangles immaterial. However, width should be kept uniform to prevent distorted image. Bars may be drawn vertically or horizontally The concept of Bar diagrams is similar to that of line diagram. However instead of lines data is presented in the form of bars. The bars are in proportion to the different figures they represent. As bar diagrams are one dimensional the height/ length of the bars alone is important. The height/ length of the bars should be in proportion to magnitude of observations. The bars can be of any width or thickness, but such thickness or width should make the diagram attractive. However, the width of all bars should be uniform. The space between any two bar should also be consistent with the scale of X axis. To elaborate, the space between bars should be equal, unless the variety (x) values are following a different trend. It is desirable to have the magnitude represented by each bar written on top of the bar. It is common practice to arrange the bars from left to right (in case of vertical diagrams) or from top to bottom (in case of horizontal diagrams)
Bar diagrams are of various types. Some of the common Bar diagram types are discussed below.

Simple Bar Diagrams

It is used for comparative study of two or more aspects of a single variable or single category of data. In this simple bar diagrams, single variable or simple classification or single category of data is represented. For example the data relating to profits sales, production exports or imports, population etc, for different periods may be represented by simple bar diagram.

Illustration -2 : Present the above illustration in the form of a simple bar diagram

Solution:



Sub-divided or Component Bar Diagram

Simple Bar diagrams are useful when data is homogenous and only one particular characteristic is being studied. However, if multiple aspects/ factors contributing to that one characteristic is to be studied, we will need to take the help of sub-divided bar diagram or component bar diagram. For example, if we want to study the gross sales of an organization in the last 5 years, a simple bar diagram will suffice. However, if we want to break the sales into Manufacturing Expenses. Administrative expense, selling expenses. Interest and profit we will need a sub divided bar diagram. Component diagram helps in comparison of different components amongst themselves as well as comparison of each component with the whole. However, if the number of components is very high, then it is advisable to use a pie diagram rather than a component diagram.

It is used for not only presenting several items of a variable or category graphically but also for making comparative study of different parts or components among themselves. It also helps in understanding the relationship between each component and the whole.

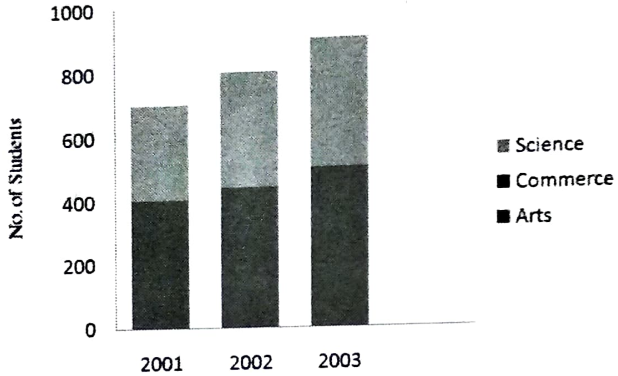
In order to construct a component diagram, the given variable is to be divided into various parts or sub-classes. A bar is to be drawn representing the given variable and then divided it into various segments, each segment representing a given sub-class of the total. Different shades colors, patterns etc are to be used to distinguish the various sub-divisions. An index is to be provided along with diagram to explain the differences

Illustration 3 : The number of students joining Art. Commerce and Science streams in the last 3 years is given below. Present the same in the form of a sub divided bar diagram.

	2001	2002	2003
Arts	250	280	300
Commerce	300	360	400
Science	150	160	200

Solution

First compute the cumulative figures. These work out to 700, 800 and 900 respectively for 2001, 2002 and 2003. Draw a simple bar diagram for this data. Now, for each year, take Arts and find out the point to which the bar diagram would have been drawn had it been only for Arts. For 2001, this figure will be 250. Draw a line there. The bottom part is the component representing Arts. Then take the sum of Arts and Commerce. For 2001 this works out to 550. Draw another line there. The component between the two lines drawn (at 250 and 550) represents Commerce. The remaining top part represents Science.



Percentage Bar Diagram

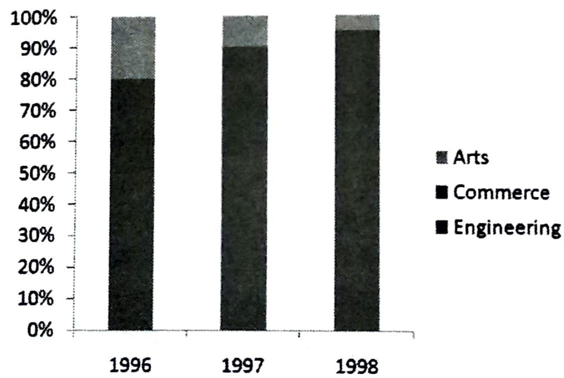
If subdivided bar diagrams are presented on percentage basis i.e., each component as a percentage of the whole, it is said to be a percentage bar diagram also. If the component diagram is presented in percentage, it is termed as 'percentage bar diagram'. It highlights the relative changes in the data. In order to construct a percentage bar diagram, the first step is to convert the given data of components into percentage of the whole. We can then proceed to prepare the component diagram as per the procedure described earlier. In percentage bar diagram, all the bars will be of same height, as all percentages will up to 100%

Illustration -4 The India Management Institute lists information on the educational background of students enrolling for its Management program.

Year	Engineering	Commerce	Arts
Batch of 1996	100	60	40
Batch of 1997	110	70	20
Batch of 1998	130	60	10

Present the same in the form of a Bar diagram

Solution :



Multiple Bar Diagram

In simple bar diagram, if is difficult to depict two or more variable. Thus, if two or more sets of data are to be presented simultaneously, multiple bar diagrams are used. In this diagram, two or more bars, which either represent different variables, or various components of the same variable are constructed adjoining each other. The method of drawing the two bars is exactly the same as described above. Like, simple bar diagram, the length of various bars varies in the ratio of magnitude of the given values. The width of different bars is identical. Different shades, colors or patterns are used to distinguish the different bars in the given set of data and an index is provided to explain the same. Multiple bar diagrams facilitate comparison of values of different variable and of the same variable over a period of time.

Multiple Bar diagrams are used if two or more sets of inter-related phenomena or variables are to the presented graphically. This is done by drawing a set of adjacent bars, each representing a different variable.

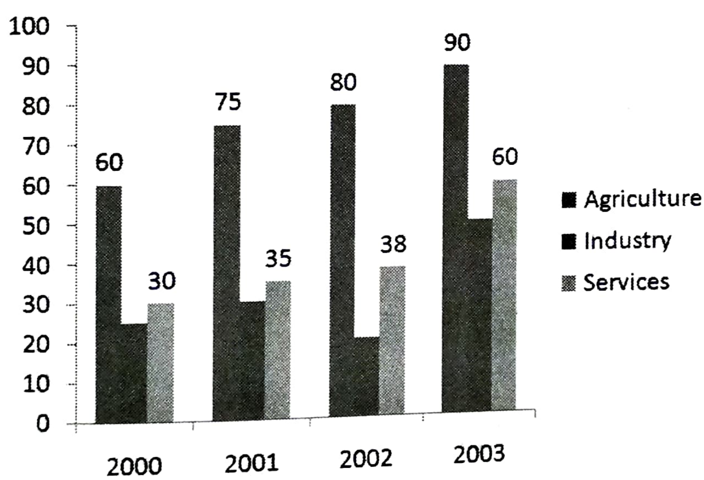
Illustration 5 : The following is the data in respect of number of persons joining the three sectors of employment, namely Agriculture, Industry and Services in the last 4 years. Present this data in a multiple bar diagram.

Number of new recruits

Year	**Agriculture**	**Industry**	**Services**
2000	60	25	30
2001	75	30	35
2002	80	20	38
2003	90	50	60

Solution

Consider only one set of data, say number of persons joining Agriculture in the last 4 years and draw a bar diagram for the same. However, leave sufficient space between the bars for each year. Now draw bar diagram in exactly the same way for the number of persons joining Industry. These Bars should be adjacent to the earlier bars. Now repeat the same with number of persons joining Services. You may kindly note that the bars are to be constructed adjoining each other with identical width. Identical gap is also to be left in between the sets. All the rules of drawing a bar diagram are applicable. The end result would be as under.



Deviation Bar Diagram

It is used for presentation of net quantities, which can be positive or negative. Net positive figures are presented by bars above the base line while negative figures are shown below the baseline.

Deviation diagrams are used to present the net quantities. When two sets of data are given and their net balance can be positive or negative. For example, if the Revenues and Expense of an organization are given it is possible that in some years, Revenue can be greater than expenses and in many cases. Revenue can be less than the expense. Setting off Revenues with Expenses can either result in profit or loss. In this diagram, these net figures are called deviations and they are represented by bars. It is important that each such deviation bar diagram should have a reference point. It is used to estimate plus (surplus) and minus (deficit) directions from the point of reference. Data of deviations can be presented either vertically or horizontally. If the bars are constructed vertically, the positive and negative values are taken on the upper and lower side of horizontal axis respectively. Similarly when the bars are constructed horizontally the positive and negative values are taken on the right and left hand side of the vertical axis respectively.

Duo-Directional Diagrams

Duo-directional diagrams, as the name suggests are drawn in 2 directions. These diagrams are used to present an 'aggregate result of different and opposite components of the same phenomenon. The total gross value of the phenomenon is broken down into 2 components, with one part being shown above the X axis (or to right of Y axis) and the other being shown below the X axis (or to left of Y axis). For example, Sales figures can be split into Expenses and Profit, as presented in the below illustration. This diagram is not the same as Deviation bar diagram where two sets of data are given and their net balance, positive or negative, is shown. Each reading will have a bar only on one side of the Axis. However, in case of duo directional diagrams, the bars are drawn on both sides of the axis, each side representing one aspect of the phenomenon.

Illustration 6 : Present given data in a suitable diagram

Year	**2000**	**2001**	**2002**	**2003**	**2004**
Revenue (₹ Crs)	120	135	148	165	195
Expenses (₹ Crs)	110	130	155	169	180
Net Profit / loss (₹ Crs)	10	5	-7	-4	15

Solution :

Having seen the deviations in the form of Net Profit / Loss, the samecan be presented as under :

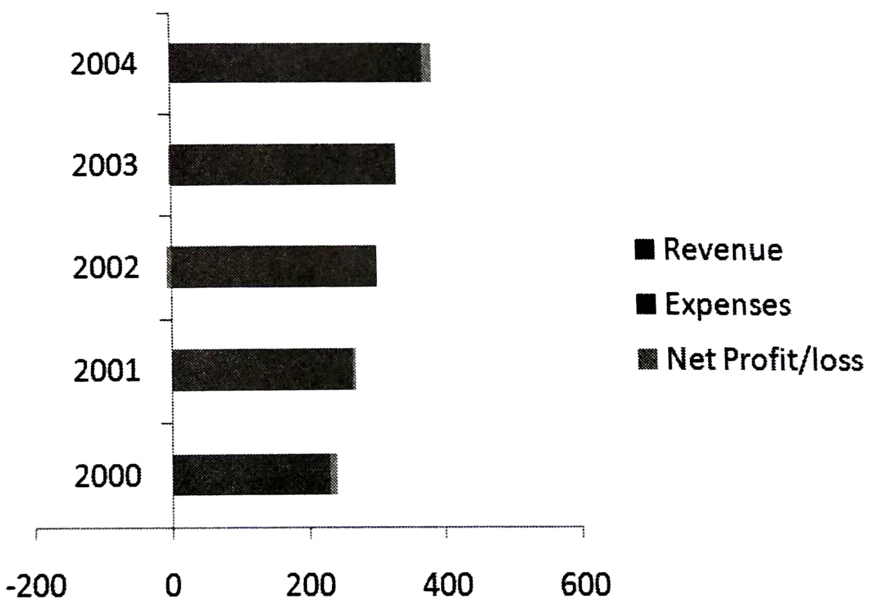
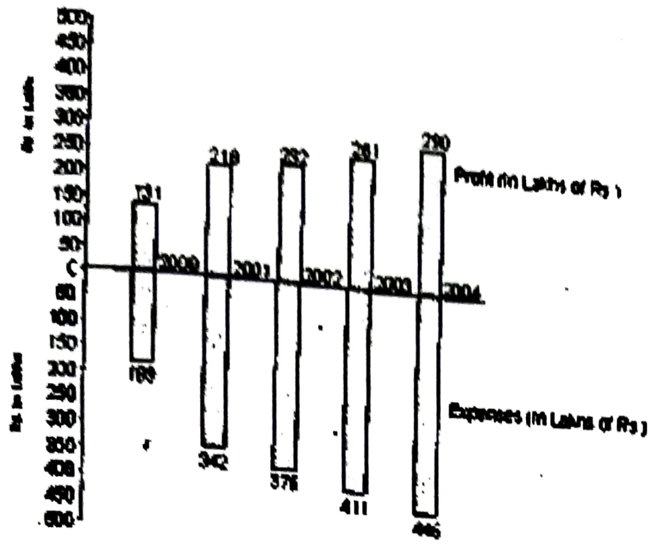


Illustration 7 : Present given data in a suitable diagram

Year	**Sales (₹ in Lakhs)**	**Expenses (₹ In Lakhs)**	**Profit (₹ In Lakhs)**
2000	320	189	131
2001	560	342	218
2002	608	376	232
2003	672	411	261
2004	736	446	290

Solution



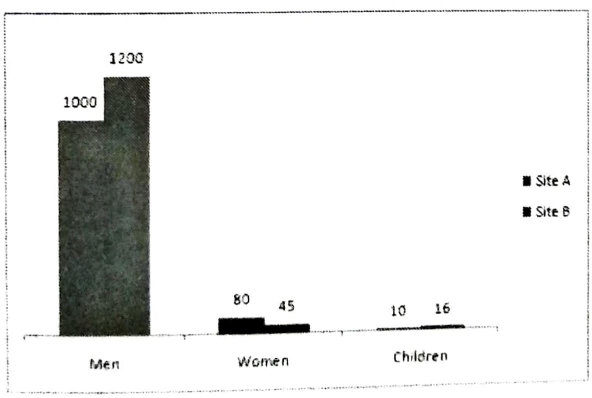
Broken Bar Diagram

It is used for presentation of data, which contain wide variations in values. Sometimes, the range of data is very wide. The data may contain very large observations along with small observation. If a normal bar diagram is drawn, the small bars will look too small and clumsy and may not disclose the true features of data. In such cases, the broken bar diagram is used. In case of broken bar diagrams, for providing adequate and reasonable shape for smaller ones, the larger bar (s) are split at the top. When we heave too large observations, instead of breaking the vertical axis, bars can be drawn with a 'false base line' for the vertical axis.

Illustration 8 : Show the following data in broken bar diagram

Labourers working on site	Site A	Site B
Men	1000	1200
Women	200	150
Children	10	16
Total	1210	1366

Solution



Pie Diagram

It enables us to show the breakup of a given total into various component parts. It is so called because the entire looks like a pie, and the components resemble slices cut from the pie. In order to construct a pie diagram, equate the total of all values to 360 deg by drawing a circle with a radius of any length. For each component, find the number of degrees that can be assigned to it, if the total is 360 degrees. For example, if total is 40,000 units and a component is 10,000 units, then it will be assigned 90 deg. {(10,000 /40,000) x Complete the angle with the original radius as one ray of the angle. Then take the next component, calculate the degrees assigned and complete the angle with any of the two radii. When all components are taken care of the circle would be split in the proportion of various components to the total.

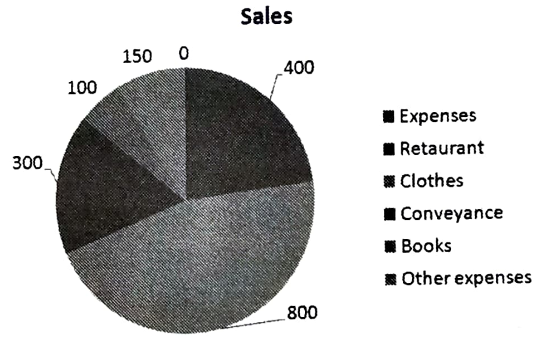
Illustration 9 : The following are the details of amount spent by a student in a month.

Sl.No	Expenses	₹
1	Restaurant	400
2	Clothes	800
3	Conveyance	300
4	Books	100
5	Other expenses	150
		1750

Present the same in the form of a Pie Chart

Solution

Degrees towards Restaurant = (400/1750*360)=82 degrees. Similarly other



Two –Dimensional Diagrams

In two-dimensional diagrams, both length and width of the observations are represented by bars. The area of a two-dimensional diagram is equal to the product of its length and width. So, two-dimensional diagrams are also called 'area diagrams'. The various types of two-dimensional diagrams are briefly explained below :

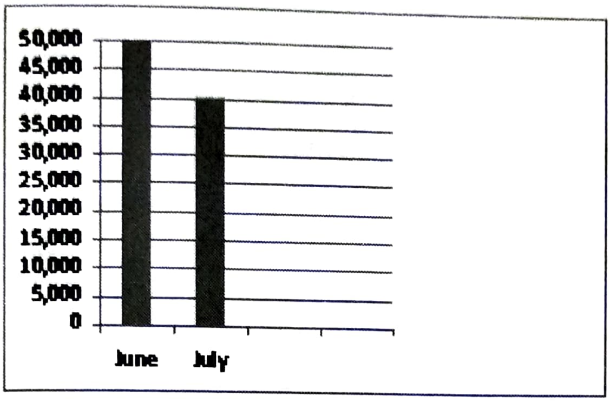
(a) **Rectangular Diagram :** A Rectangle diagram is a modified form of bar diagram. It provides more detailed information than is furnished by a bar diagram. The length of the rectangle represents one aspect of data while its width would represent another aspect of the data provided. For example, if an organization is selling 1000 units of an article per month at unit price of ₹ 10 and another organization is selling 1500 units of an article per month at a unit price of ₹ 8 the total sales of the two organization are ₹ 10,000 and ₹ 9,000 respectively. While this can be presented in the form of a bar diagram a Rectangular diagram can be used to presented a lot more information. The length of the rectangle will represent the number of units sold and the selling price per unit can be represented by width. The area of the rectangle drawn will represent total sales. Proper and equal spacing is given between different rectangles, placed side-by-side

Illustration 10 : Pravek Ltd furnishes the following data relating to production and cost per unit during june and july 2000. Construct rectangle diagram.

	June	July
Units produced	50,000	40,000
Per unit cost ₹	16	20

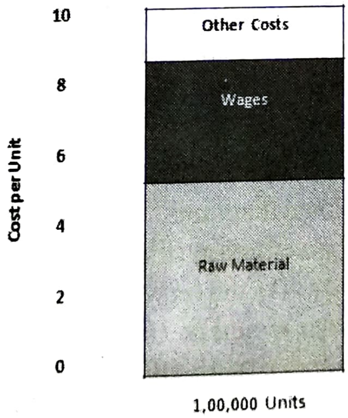
Solution

The rectangles are to be prepared for the given data. Represent the number of units produced on the vertical axis and per unit cost on horizontal axis. The length of rectangle is in the ratio of 5:4 while the width of is in the ratio of 4:5 On vertical line 5000 is taken as one centimeter. Look at the diagram 1.7 for the rectangle showing the cost of production.



Sub-divided Rectangle Diagram

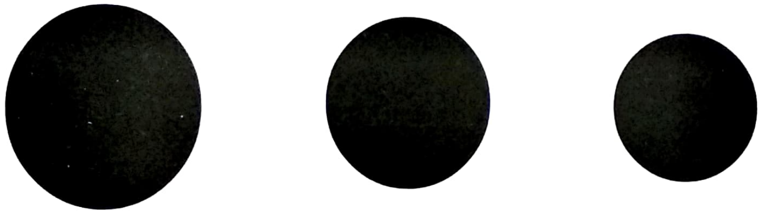
As in case of a bar diagram, a rectangle can also be subdivided or percentage sub-divided to represent multiple characteristics at the same time. The procedure for preparing a sub-divided or percentage rectangle (two dimensional) is the same as in case of sub-divided / percentage bar diagram. The rectangle is sub-divided according to the cumulative values of various components or sub-divisions. For example, if the cost of production per unit is to be broken down into cost or material, labor cost and other expenses, and the number of units produced is also to be compared, the total cost per unit can be depicted on Y axis and the number or units can be depicted on X axis. To illustrate if the cost of material labor and other costs per unit of a product are ₹ 6 ₹ 3 and Re.1 respectively and 10,000 units of the product have been produced, this data can be presented in the form of a sub-divided two dimensional diagram as under :



(b) **Square Diagrams :** A Square diagram also is a two-dimensional diagram representing area. Square diagrams are specifically and if we wish to compare values that differ significantly from one other. Since all sides of the square are the same and the square diagram is represented as area of square, the length as well as width of the square should be equal to or proportionate to the square root of the given data. If more than one square is drawn in a single diagram the bases of all the squares must lie on the same line.

For example, if we need to show the amount set aside in the union budget for Agriculture, Industry and Energy sectors which is say 8,100 crores. ₹ 3,600 crores and ₹ 400 crores respectively, we need to first calculate the square root of these amounts. This works out to 90, 60 and 20 respectively, we can now consider 1 cm =10 and draw square with sides of 9, 6 and 2 respectively. The three squares can be shown as under

(c) **Circle Diagrams :** Circle diagram is an alternative to square diagram. It is also used to present the values differing widely in their magnitudes. Since area of a circle that represents are given values is given by Πr^2(where =22/7 and r = radius of the circle) the radius of the circle should be equal to or proportionate to the square root of the given data. To illustrate, if the total spending on Agriculture is 8100 crores, we equate r^2 to 8100 crores. This translates to the value of r being approximately 50.76. Similarly, other data can be obtained and a circle diagram can be drawn. This is shown as under



Three-Dimensional Diagrams

Three –dimensional diagrams are diagrams in which three dimensions, namely, length breadth and height are taken into account. They are also termed as volume diagrams. The objective of presenting data in 3 dimensions is to highlight the magnitude of the data through volumes of the corresponding diagrams. Diagrams in the shape of cubes, spheres, cylinders, blocks etc are commonly used. Three dimensional diagrams are specially useful if there are very wide variations between the smallest and the largest magnitudes.

Cubes

Of the various three-dimensional diagrams, 'cubes' are the simplest and most commonly used methods of diagrammatic presentation of data. For example, if the smallest and the largest magnitudes to be presented are in the ratio of 1:1000, the bar diagrams cannot be used because the height of the biggest bar would be 1000 times the height of the smallest bar and thus they would look very disproportionate and clumsy. On the other hand, if square or circle diagrams are used then the sides (radii) of the squares (circles) will be in the ratio of the square roots viz. 1 1,000... i.e 31.63 i.e. 1.32 (approx), which will again give quite disproportionate diagrams. However, if cubes are used to present this data, then since the volume of cube of side x is x^3, the sides of the cubes will be in the ratio of their cube roots which is 1:........i.e., 1:10 which will give reasonably proportionate diagrams as compared to one-dimensional or two-dimensional diagrams.

Three-dimensional diagrams are useful if the range of data is very large. They are more appealing to the eye and make a lasting impression. However, they are difficult to construct and not easily understood. As magnitudes are represented by volumes, it is difficult to visualize and hence interpret them with precision. Hence, they are not very popular.

Similarly, Cylinders, Spheres and blocks are quite difficult to construct and are not actively used. However, use of projection techniques has enabled us to present even one – dimensional diagrams as three-dimensional diagrams, thus sprucing up the look and feel of the diagrams

Pictograms

Pictures are more attractive and appealing to the eye and have a lasting impression on the mind. Pictograms is the technique of presenting statistical data through appropriate pictures. It is a very popular device for presenting statistical facts to a layman who does not have any mathematical background. The magnitude of the particular phenomenon under study is of the pictures being proportional to the values of the different magnitudes to be presented. This technique is extensively used by Government for diagrammatic presentation of data relating to a variety of social, business or economic phenomenon primarily for display to the general public or common masses in fairs and exhibitions.

Pictograms are difficult and time-consuming to construct. In pictogram each pictorial symbol represents a fixed number of units like thousands millions or crores, etc. For instance, a pictogram food production in various states of India may represent 20 ml of production with I bag of grain, it is difficult to represent production that is in fractions of 20 mt. This proportionate representation introduces error and is quite difficult to visualise with precision.

Cartograms

Cartograms are used to present data pertaining to geographical regions with the help of maps. For example, the distribution of rainfall in different regions of a country can be shown with the help of maps or cartograms. The different geographical regions are depicted on a map and the quantum of rainfall in the regions may be shown by dots, different shades or colours etc. Cartograms are simple and elementary forms of visual presentation and are easy to understand. They are generally used when the regional or geographic comparisons are to be highlighted.

Choice of a Diagram

Visual presentation of a given set of data is not easy and requires great skill, intelligence and expertise. The choice will primarily depend upon the nature of the data, the objective of presentation and the type of the audience to whom the diagrams are to be presented. A wrong selection of the diagram will distort the true characteristics of the phenomenon to be presented and might lead to very wrong and misleading interpretations. Some special types of data, viz., the data relating to frequency curves an time series are best represented by means of graphs which we will discuss in the following sections.

Graphs

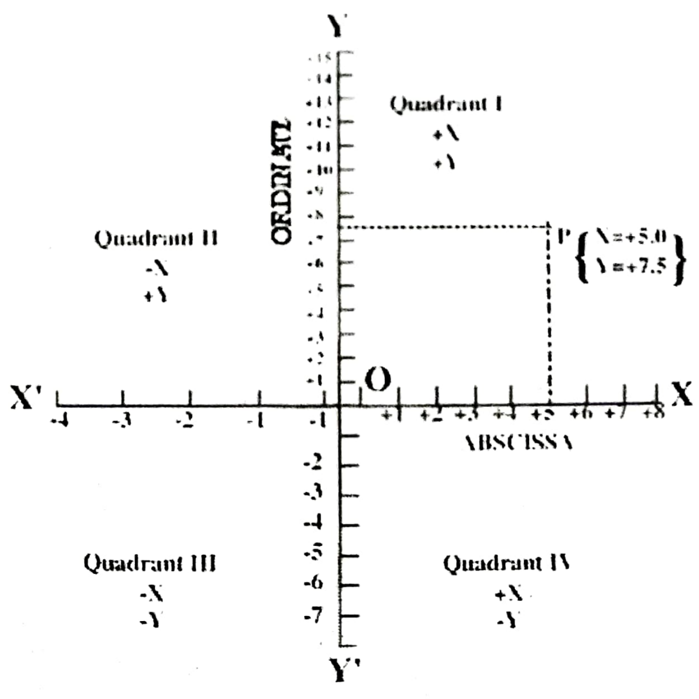
A student of Economics studying the relationship between demand and price (Law of Demand) observes the quantity of mangoes being demanded by a housewife at various prices. For example of housewife may be willing to buy 5 kg of mangoes if the price is ₹ 25 per kg. She may be restrained to buy only 2kg if the price is reduced to ₹ 30 per kg. She may limit her buying to just 1 kg if the price is increased to ₹ 50 per kg. The student can attempt to analyze this data to arrive at meaningful conclusions with the help of graphs.

Graphs are a specialized mode of representing data involving two variables. Data is represented with the help of two co-ordinates. Each co-ordinate representing the values of a variable. Each "point' represents value of a variable along with the corresponding value of the second variable. In our example (5.20) is a point which states that demand for mangoes is 5 kg at a price of ₹ 20 per kg. Similarly (2.30) and (1.50) are also points representing the two variables. Namely Demand and price. These points can be plotted on a graph paper for the purpose of further study.

Technique of construction of Graphs

Graphs are drawn on a special type of paper known as graphs which has a fine network of horizontal and vertical lines for smaller parts of the same. In a each division of a centimeter and thin lines for smaller parts of the same. In a graph of any size, two simple lines are drawn at right angle to each other, intersecting at point 'O' which is known as origin or zero reference. The two lines are known are co-ordinate axes. They divide the entire graph paper into four quadrants. The horizontal line is called XX or X axis and the vertical line YY or Y axis. The positive values of a variable on the X axis will be on the right side of the origin and the negative values will be on the left of the Origin. The positive values of a variable on the Y axis will be on the upper side of the origin and the negative values will be on the lower portion of the Origin. The top right hand corner will have both the variable positive and is called the first quadrant Normally, only the first quadrant is used. The remaining three quadrants are used only if negative values are also involved. The following are four quadrants showing how the points will figure in each quadrant with +or –signs:

Any pair of the values of the variables is represented by a point (x-y) wherein x always refers to the X-coordinate and usually represents the value of the independent variable and is shown along the X-axis y refers to the Y-coordinate and represent the value of the dependent variable and is shown along the Y-axis, The X co-ordinate is also known as abscissa and the y coordinate is also known as ordinate.



General rules for graphing

Apart from the various aspects discussed under general rules for constructing diagrams, the following points need to be kept in mind.

1. **Neatness :** One of the basic requirements of a graph is to present data in a manner that is visually pleasing. Neatness is of paramount importance in ensuring that the graph catches the attention of the intended audience.
2. **Title :** Every graph must have a clear and comprehensive title so that the user of the graph understands what is being presented by means of the graph.

3. **Footnote :** Any explanation that is necessary to explain and is not obvious needs to be explained with the help of a footnote. For example, if a graphs is representing the revenues of a corporate for the previous 5 years and the current year it may not catch the attention of the user that revenues on the last 5 years are audited. But the current year revenues may be unedited. This needs to be explained with the help of a footnote. It is common to use special symbols such as *, #, $ etc to indicate the place where a footnote is required.

4. **Structural Framework :** The position of the axes should be so chosen that the graph looks proportionate and attractive. In drawing the graph it is customary to plot the independent variable along the X-axis and the dependent variable along the Y-axis. For instance, if census details are being plotted on a graph paper, the time factor (Year) is taken along X-axis and the population of the country in different years is taken along Y-axis.

5. **Scale :** The scale along both the axes (X-axis and Y-axis) should be so chosen that the entire data can be accommodated in the available space without crowding. The ratio between the horizontal and vertical scales that need to be considered. The figure must be sufficiently small for the whole of it to be visible at once. If the range of data is very high. Minute accuracy can be sacrificed to some extent. All important variations must be clearly visible.

6. **False Base Line :** In a graph, the vertical scale normally starts with 0. However if the minimum value of the dependent variable is very far (greater than) from '0' and the fluctuations of the variable are relatively small, the fluctuations can be more effectively portrayed by using a false base line. In such a situation, the vertical scale is broken and the space between the origin '0' and the minimum value (or some convenient value near that) of the dependent variable is omitted drawing two zigzag horizontal lines above the base line. The scale along Y-axis is framed accordingly.

   False base line technique is quite extensively used for magnifying the minor fluctuations in a time series data. It also saves and enhances the look and feel of the graph. However, proper care should be taken to interpret graphs in which false base line is used.

7. **Line Designs :** If more than one variable is to be depicted on the same graph, the different graphs so obtained should be distinguished from each other by the use of different lines viz,. dotted lines, broken lines, dash-dot lines, thin or thick lines etc., and index to identify them should be given.

8. **Source Note :** The source note states the source from where the underlying data presented in the graph has been collected. The source note adds credibility to the data presented.

9. **Index :** If multiple shades, designs or colors have been used in the graph, they need to be explained to the user with the help of a legend.

10. **Simplicity :** The graph presented should be simple such that a layman can also understand the underlying message without having any mathematical / statistical background.

11. **Ratio or Logarithmic Scale :** These scales are used to display proportional changes in data. In order to display proportional or relative changes in the magnitude, the ratio or logarithmic scale should be used instead of natural or arithmetic scale is used to display absolute changes.

Types of Graphs

Graphs can be broadly classified under two types. The two types of graphs are

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

Time series graphs or Historigrams

Time series graphs have time as an independent variable on the X-axis (horizontal axis) and dependent variables on the Y-axis (vertical axis), Two-or more variables dependent on time can be conveniently presented on the same graph provided the magnitudes do not vary a great deal. A suitable scale has to be devised for presenting all the variables. For example time on X axis can be terms of years, months or even hours while Y axis be equal to 100 units or 10,000 units depending on the data to be presented.

The relationship between time and one or more variables can be presented in a number of ways. Depending on the type of data and the number or variables, Time Series graphs can be further classified into

- Horizontal Line Graphs
- Component or Band Graphs
- Silhouette or Net balance Graphs
- Range or variation Graphs and
- Z-curvein this case

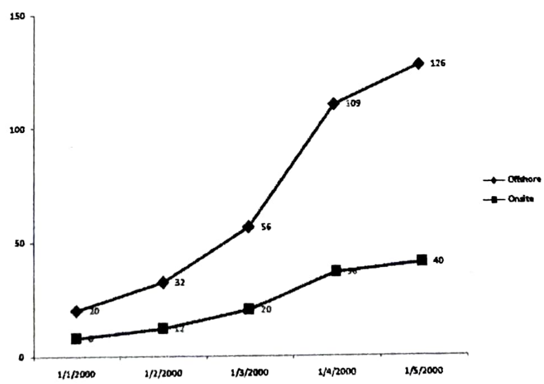
Horizontal Line Garphs

These are the simplest of Time series graphs. In case of Horzontal Lines Graphs, separate lines represent the variables with reference to the time.

Illustration 11 : Present the following data on a graph-paper

Year	**Jan-00**	**Feb-00**	**Mar-00**	**Apr-00**	**May-00**
Offshore	20	32	56	109	126
Onsite	8	12	20	36	40

Offshore Onsite Mix



It may be noted that in the graph more than one time series pertaining to some related phenomena and belonging to the same time period are shown and the scale on the Y-axis is kept same. The curves for each time series are shown by different types of lines.

Band Graphs

A Band graph is prepared when a variable has several components. Hence Band graph is also known as Component graph. The given data is converted into cumulative data and then the constituent parts of the variable are plotted one over the other. They are distinguished by different shades. Band graph is also known as Belt curve.

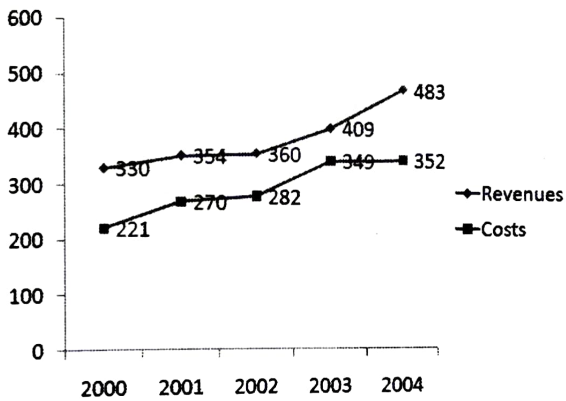
Silhouette or Net Balance Graph

This graph is used when two different but related variable need to be presented with clear emphasis on the difference between the two. In this graph, the net balance is highlighted with the help of distinct shades or such other depiction. The following illustration presents the data in the form of a Net Balance graphs.

Illustration 12 : Present the following data on a graph paper :

Year	**2000**	**2001**	**2002**	**2003**	**2004**
Revenues	330	354	360	409	483
Costs	221	270	282	349	352

Net Balance Graphs



Students may note that in the above example, if in any one year, the costs exceed revenues, then we still have a net balance, which represents loss. This needs to be depicted in a different style (different color, pattern or shade)

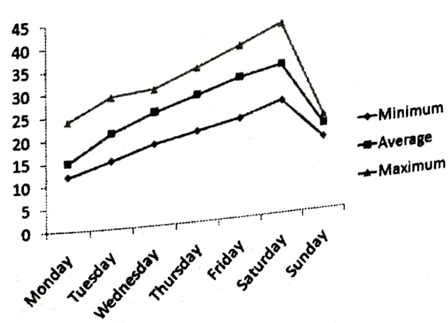
Range Graph

Range graph is used to depict the variation in data with reference to each period. The highest, the lowest and the average values of the variable being studied are plotted on the graph paper

Illustration 13 : Present the following data on graph by means of range graph :

	Temperature		
Minimum	**Average**	**Maximum**	
Monday	12	15	24
Tuesday	15	21	29
Wednesday	18	25	30
Thursday	20	28	34
Friday	22	31	38
Saturday	25	33	42
Sunday	16	19	21

Solution



This graph can also be prepared by drawing two different curves in a normal way-one for the maximum and another for the minimum and then the area between the two curves is shaded to show the range. It will then resemble the net balance graph, except that there will be only one shaded area, as at any time, maximum value of an observation will always be greater than the minimum value of the observation.

Z-chart or Z-curve

A Z-chart is a special kind or graph where figures are plotted at regular intervals over a period of one year. It is slightly complicated by the fact that it incorporates three graphs on the same chart and these curves or lines taken together tend to look like the letter Z: hence is name. The easiest approach is to build up the chart line by line.

Let us understand this with the help or an example.

Illustration 14 : Given below is data pertaining to Sales of a company for the last six months (January to June). Also given are Cumulative Sales and Estimated Sales for the six months assuming that monthly sales will be uniform each year. Plot the same on a graph paper

Monthly	**Cumulative Sales**	**Estd Sales**	**Sales (Monthly sales *6)**
January	15	15	90
February	14	29	84
March	11	40	66
April	10	50	60
May	12	62	72
June	15	77	90

Solution

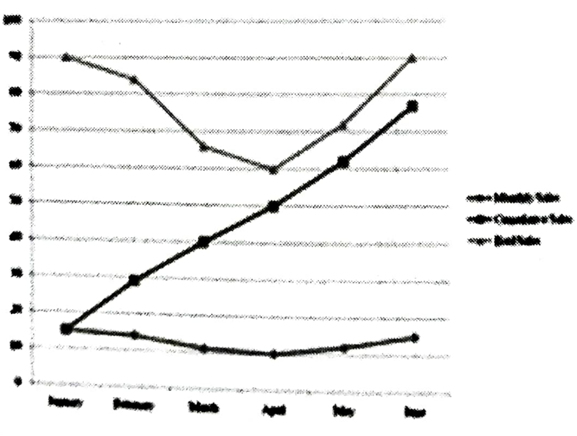
Step 1 : Draw the horizontal line graph for monthly sales

Step 2 : Draw the horizontal line graph for Cumulative sales

Step 3 : Draw the horizontal line graph for estimate six monthly sales

The resultant figure will look like the alphabet Z..

Z Chart-Sales of last 6 months



Graphs of Frequency distribution

Frequency graphs are designed to reveal clearly the characteristics features of a frequency data. In a frequency graph the size or the value of the item is presented on the horizontal axis and the frequency or the number of items on the vertical axis. The scale of measurement has to be indicated separately on both the axis. The following types of frequency graph are used:

- Vertical Line Graph
- Polygon
- Ogive
- Histogram
- Frequency curve
- Lorenz curve

Vertical Line Graph

A vertical line graph is used in case of discrete series. Wherein the phenomenon whose frequency is being studied is listed on the X axis and the frequency of the phenomenon is depicted in terms of the height of a straight line parallel to the Y axis. This is appropriate where a given frequency is associated with a given value only and not range of values.

Histogram

A histogram is a graphical method of presenting data. Where the observations are located on a horizontal axis and the frequency of those observations is depicted along the vertical axis. It is one of the most popular and commonly used devices for charting continuous frequency distributions.

Construction of Histogram

Construction of a histogram involves erecting a series of adjoining vertical rectangles on the sections of the horizontal axis (X-axis) with bases (sections) equal to the width of the corresponding class intervals and heights proportional to the corresponding frequencies of each class. The length and breadth of the rectangles should reflect the class intervals and corresponding frequencies. Construction of Histogram will vary depending on whether the class Intervals are equal or unequal.

(i) **When class intervals are equal :** Each class interval is drawn on the x-axis by a base proportional to the magnitude of the class interval. On each class interval base, erect a rectangle with height proportional to the corresponding frequency of the class. The series of adjacent rectangles so formed gives the histogram of the frequency distribution and its area represents the total frequency of the distribution as distributed across the different classes.

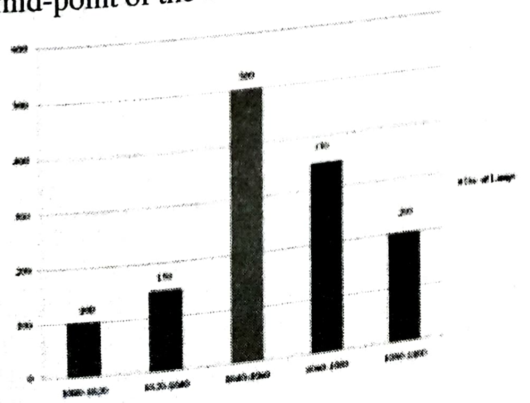
Illustration 15 : Draw a Histogram of the following frequency Distribution

Life of Electric Lamps (Hrs.)	No. of Lamps
1000-1020	100
1020-1040	150
1040-1060	500
1060-1080	350
1080-1100	200

Solution

Step 1 : List the class intervals on the x-axis

Step 2 : Erect rectangles with height proportional to frequency. Width of rectangle should be proportional to class Interval. This can be ensured by making the central point of the bar the mid-point of the class-interval



In the above case, area of the rectangle on each class-interval is exactly equal to the number of frequencies in that class multiplied by the range of the class interval and the total area of the figure comprising various rectangles is equal to the total number of frequencies multiplied by their respective class intervals.

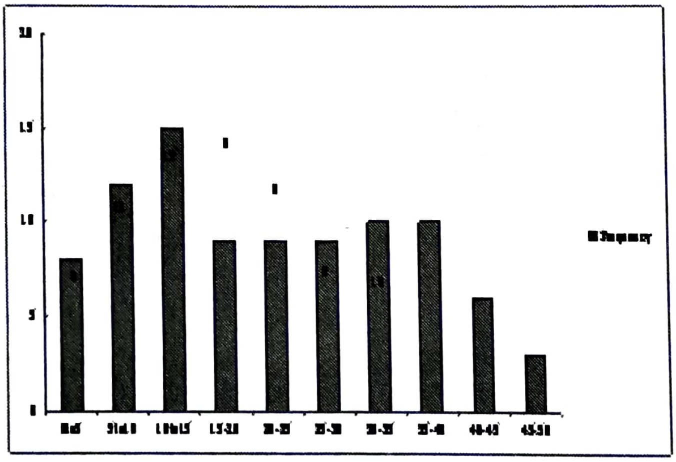
(ii) **When Class Intervals are Unequal** : In case of unequal class intervals, if the rectangles are erected on different class intervals as explained above the areas would not be proportional to the frequency and will give undue weight age to those having larger class intervals. In order to correct this situation, if class intervals are unequal, they should be converted into equal class intervals. For example if the first three class intervals are 0-5, 5-10 and 10-15 and the fourth interval is 15-30, with a frequency of 30, it is split into three classes of 15-20, 20-25 and 25-30 with frequency of 10 each. Once the Class intervals are equal, the problem is same as discussed above

Illustration 16 : Prepare Histogram for the following data :

Class Interval	0-5	5-10	10-15	15-30	30-40	40-45	45-50
Frequency	8	12	15	27	20	6	3

Solution

The Class Interval should be taken as 0-5, 5-10, 10-15, 15-20, 20-25, 25-30, 30-35, 35-40, 40-45 and 45-50



Points to Note

1. Histogram can be drawn only if the frequency distribution is continuous. In case the given series is an inclusive series, we can eliminate the gap between the limits of two adjacent classes by taking the mid-values of the class limits of two adjacent classes as the common limit of both the classes. In other words, we convert the series into an Exclusive series.
2. Sometimes, only the mid-values of different classes are given. In such a case the given distribution is converted into continuous classes by ascertaining the upper and lower limits of the various classes under the assumption that the frequency is uniformly distributed throughout the class intervals.
3. Histograms can also be used to represent discrete frequency distribution. This is done by considering the given values of the variable as the mid-points of continuous classes and then proceed as explained above.
4. Histograms can't be constructed for frequency distributions with open-end classes. The open and classes need to be converted into close ended class by assuming that the magnitude of the open class is same as that of its nearest class. In other words, if the first class in open, its class interval should be taken as equal to the Class interval of the succeeding (second) class. Similarly, if the last class is open its Class interval of the preceding (i.e. last but one) class.

Difference between Histogram and Bar diagram

(i) A histogram is two dimensional (area) diagram where both the width (base) and the length (height of the rectangle) are important whereas bar diagram is one dimensional diagram in which only length (height of the bar) matters while width is arbitrary.

(ii) In a histogram, the bars (rectangle) are adjacent to each other whereas in bar diagram proper spacing is given between different bars.

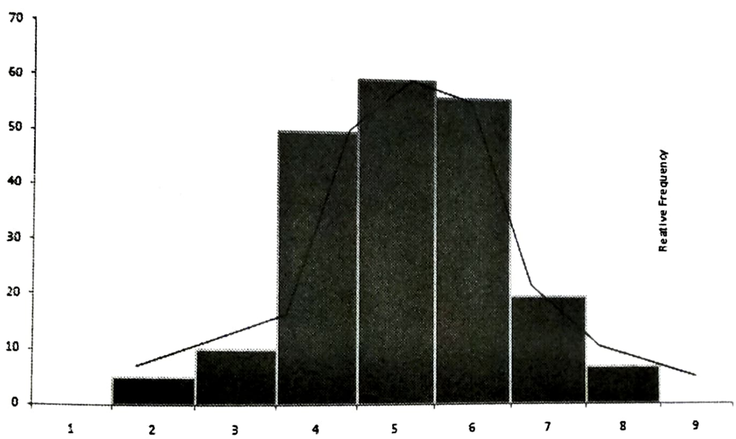
Frequency Polygon

Polygon is a figure with multiple, usually more than four side. Frequency polygon is graphical representation of frequency distribution in the form of a curve superimposed on a histogram. Generally frequency polygon is drawn only when class-intervals have a common width. The area of the polygon will be same as the area of histogram.

Construction of Frequency polygon : It is constructed in two ways

From Histogram

Mark the midpoints of each class interval on the tops (upper horizontal sides) of the adjacent rectangles of the histogram. Also mark the midpoint of two hypothetical classes is the class immediately before and first class and the class immediately after the last class. Since the frequency of both these classes is zero, the two points will be marked on the x-axis. Join all the marked points by drawing straight lines. The resulting figure is a frequency polygon.



Without Histogram

A Frequency Polygon is drawn on the presumption that the frequencies in a class interval are evenly distributed throughout the class and hence, their mid-points are representative in actual construction. Thus a frequency polygon can be drawn without drawing a Histogram. In order to draw the same, mark the mid values of all classes, including the two hypothetical classes discussed above, on the x-axis. For each of the classes, plot the frequencies (against the marked mid-value) along the y-axis. Join the points by drawing the straight lines to obtain frequency polygon.

Frequency polygon has an advantage over histogram as several distributions can be plotted on the same graph. This facilitates easy comparison. In case of histogram, each distribution will have to be presented in a separate graph. The frequency polygon is an improvement over histogram as it provides a continuous curve showing the slope of rise and fall in data.

On the other hand it is an approximation as it includes certain areas that do not form part of the actual data and leave out others that are part of it. However, these defects do not reduce the usefulness of the graph as the frequency distribution is by itself an approximation and the error involved is one of compensating type.

Frequency Curve

Frequency curve is a smooth free hand curve drawn through the vertices of frequency polygon. It is a smooth free hand curve drawn to fit into the histogram in a manner that it passes through the tops of rectangular blocks. The area included under the curve is approximately the same as that of the polygon.

The basic drawing a frequency curve is to present graphically the area covered by a histogram in a more symmetrical fashion. If the frequency curve is correctly drawn, it can be used for interpolation purposes.

The frequency curve represents the same frequency distribution which is represented by the histogram or frequency polygon, except that the lines joining the mid points are not straight but curved lines. The angularities are smoothed and the top is also rounded very carefully.

Types of Frequency Curves : There are different types of frequency curves:

1. **Symmetrical bell-shaped :** The curve is bell shaped and can be bisected at the center to result in two equal sides mirroring each other.
2. **Asymmetrical bell shaped :** The curve is bell shaped but cannot be bisected at the center. The peak is either towards the left or right of the center. If the peak is towards the left, if is termed as positively skewed. If the peak is towards the right, it is termed as Negatively skewed
3. **Upward Sloping or J-shaped :** This curve is not bell shaped. It keeps rising as we move from left it right. It loosely resembles the shape of the alphabet. J
4. **Downward Sloping or Reverse J-shaped :** This curve is also not bell shaped. It peaks in the beginning and then slowly keeps slopping down as we move from left to right. It loosely resembles the mirror image of the alphabet. J
5. **U Shaped :** As the name suggests, it loosely resembles the shape of the alphabet U. It is a downward sloping curve to begin with but gets converted into an upward sloping curve after some time.
6. **Bi-modal Curve :** This kind of Curve is obtained when there are two peak points (may not be equal) and frequencies are concentrated around them.
7. **Multi-modal (or Mixed):** This kind or curve is obtained when there are multiple peak points (not necessarily equal) and frequencies are concentrated around these multiple points.

Ogive

Ogive is a graphic representation of a cumulative frequency of a distribution. It is also known as Cumulative Frequency Polygon. It is pronounced as 'ogive'. The ogive is drawn by joining the cumulative frequencies plotted against the corresponding class boundaries. There are two types of ogives. It can be of 'less than' or 'more tan' type based on whether the cumulative frequencies used are of 'less than' or more than 'type.

1. **'Less than' Ogive** : The "less than olive" like an elongated S which starts from the lowest class boundary on the horizontal axis and gradually rises upward, ending at the highest class boundary corresponding to the total frequency of the distribution. To construct this curve, we plot the 'less than' cumulative frequencies against the upper class boundaries of the respective classes. The points plotted are joined by a smooth free hand curve. It is upward sloping from left to right.
2. **'More than' Ogive :** The "more than ogive" has the appearance of an elongated S turned upside down. It starts from the lowest class boundary on the horizontal axis and gradually slopes downward, ending at the highest class boundary corresponding to the total frequency of the distribution. To construct this curve, we plot the 'more

than' cumulative frequencies against the lower class boundaries of the respective classes. The points plotted are joined by a smooth free hand curve. It is down ward sloping from left to right.

Ogive is very useful for graphic computation of partition values, viz, Median, Quartiles, Deciles, Percentiles, etc. of which we will learn in the forthcoming chapters. It is alright to draw both the 'less than' ogive and 'more than' ogive on the same graph. If done so, they intersect at a point. A perpendicular from the point of intersection on the X-axis gives the value of median.

Ogive is used for estimating cumulative frequencies of values falling within the range of values represented on a graph. For example one can determine the total number of assesses paying not less than a certain sum of money, or the number of workers who are receiving wages that are not more than a certain amount, etc.

Ogive can be used with advantage over frequency curves. For each of the distribution, different ogive can be constructed on the same graph and they are generally less over-lapping than the corresponding frequency curves.

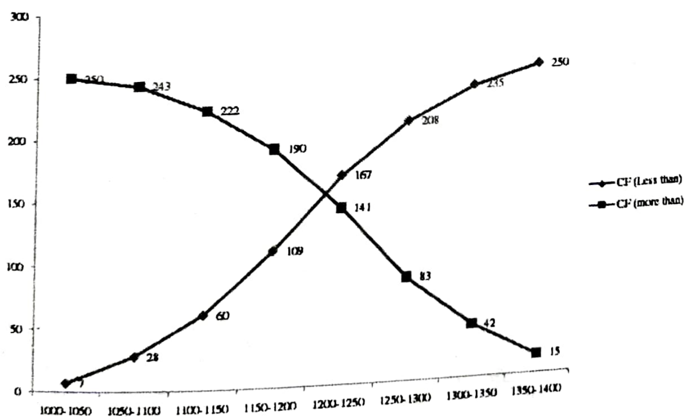
Illustration 17 : Following information is obtained on the number of telephone calls made by 246 companies for the months of May and June 98.

Telephone Calls	1001-1050	1051-1100	1101-1150	1151-1200
Companies	7	21	32	49
Telephone Calls	1202-1250	1251-1300	1301-1350	1351-1400
Companies	58	41	27	15

Construct (a) A More than Ogive (b) less than Ogive

Solution

Class Interval	F	Less than CF	More than CF
1000-1050	7	7	250
1050-1100	21	28	243
1100-1150	32	60	222
1150-1200	49	109	190
1200-1250	58	167	141
1250-1300	41	208	83
1300-1350	27	235	42
1350-1400	15	250	15



For drawing such a graph, it is preferable to use the percentage of cumulative frequencies instead of actual values, particularly when the actual values are very large. In fact use of the percentage of cumulative frequencies is very useful when comparing two ogives, each of which is based on a different data. Generally, instead of points on the graph being joined by straight lines, a smooth free-hand curve is drawn.

Lorenz curve

Lorenz Curve is a special type of graph designed to show how much a certain distribution varies from a completely uniform distribution. In fact it is a cumulative percentage curve, combining percentage of items under review with the percentage of the factor among the items.

Let us understand the concept and construction of Lorenz curve with the help of an example.

Consider the following data that shows the number of firms having their employee strength in a defined range and the output produced by them.

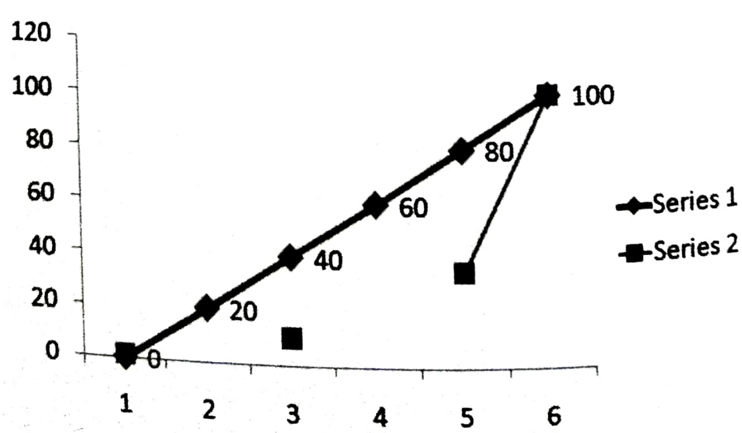
Average no. of employees	No. of firms	Net output (mt)
20 to 200	205	16
200 to 600	200	60
600 to 1000	35	18
1000 to 1500	30	26
1500 to 2000	20	26
2000 to 3000	10	54

For a given average strength of employees, one would normally assume that if 10% of firms fall in a particular category, they would also produce 10% of the total output. Let us see if this assumption holds good for the given data. This can be done by first calculating the percentage of firms falling under each category and the percentage of output produced by them. The next step is calculate cumulative percentages and compare the two cumulative percentages. This is done as under.

Average no. of employees	No.of firms	% of firms	Cumulative percentage	Net out	% of put (mt)	Cumulative putput percentage
20 to 200	205	41%	41%	16	8%	8%
200 to 600	200	40%	81%	60	30%	38%
600 to 1000	35	7%	88%	18	9%	47%
1000to1500	30	6%	94%	26	13%	60%
1500to2000	20	4%	98%	26	13%	73%
2000to3000	10	2%	100%	54	27%	100%
	500	100%		200	100%	

Let us now plot both the cumulative percentage series on to graph paper, we reach the following conclusion :

(To be drawn with hand)



Interpretation : The Lorenz curve is a graphic method of indicating whether a certain quantity is equally distributed throughout the population. In the worked example, the quantity is the net output and population is the number of firms. If there is an equal distribution of output between firms, then 20% of the firms would produce 20% of the output. 50% of the firms would account for 50% of the output etc. and all the points on the graph would be along a straight line. Departure from this line indicates the extent to which the distribution varies from uniformity.

In the example of this particular industry, we see that there is considerable inequality in the distribution. While the majority of about 90% of the firms (the smaller ones) produce 50% of the total output, the large sized firms, which are only 2% of the number, have 27% of the output while at the other end of the scale 41% of the firms account for 8% of the total output only.

The farther is the curve from the line of equal distribution, the less uniform is the distribution.

Difference Between Diagrams and Graphs

Diagrams	Graphs
Plain paper is used	Graph paper is used
Diagrams help in comparison of variables and not for study of mathematical relationship between them	Graphs help in study of mathematical relationship between the variable
Data is presented by devices such as bars, rectangles, squares, circles, cubes, etc	Points or lines of different kinds such as dots, dashes, dot-dash, etc, are used to present the data.
Diagrams furnish only approximate information	Graphs are more obvious, precise and accurate
Diagrams are not amenable to further mathematical treatment or statistical analysis	Graphs help the statistician in the study of slopes, rates of change and estimation.
Diagrams are useful in depicting categorical and geographical data	Graphs are used for the study of times series and frequency distributions.

EXERCISE – I

1. Present the following data in the form of a line diagram

Year	2000	2001	2002	2003	2004
Profit (₹ in Lakhs)	125	225	250	270	310

2. The following are 5 years sales of Lakshmi traders based on the data draw a simple bar Diagram

Year	2000	2001	2002	2003	2004
Sales	200	325	410	300	250

(*B.Com., A.N.U*)

3. For the following date draw a simple bar diagrm.

Production expenses	**Firm X**	**Firm Y**	**Firm Z**
Raw Material (₹)	20	40	25
Wages (₹)	10	10	10
Direct expenses (₹)	3	5	10
Indirect expenses (₹)	2	5	5
Total Cost	35	60	50

(*B.Com., A.N.U*)