UNIT 7 DIAGRAMMATIC AND GRAPHIC PRESENTATION

STRUCTURE

7.0 OBJECTIVES

After studying this Unit, you should be able to:

- explain the need and significance of visual presentation (diagrams and graphs) of the data in research work,
- describe various types of diagrams and illustrate how to present the data through an appropriate diagram,
- describe the principle of preparing a graph,
- present frequency distribution in the form of historigrams, histograms, frequency polygon and ogives to make decisions, and
- list out and distinguish between the major forms of diagrams and graphs.

7.1 INTRODUCTION

In the previous Unit 6, you have studied the importance and techniques of editing, coding, classification and tabulation that help to arrange the mass of data (collected data) in a logical and precise manner. Tabulation is one of the techniques for presentation of collected data which makes it easier to establish trend, pattern, comparison etc. However, you might have noticed, it is a difficult

and cumbersome task for a researcher to interpret a table having a large mass of numerical information. Sometimes it may fail to convey the message meaningfully to the readers for whom it is meant. To overcome this inconvenience, diagrammatic and graphic presentation of data has been invented to supplement and explain the tables. Practically every day we can find the presentation of cricket score, stock market index, cost of living index etc., in news papers, television, magazines, reports etc. in the form of diagrams and graphs. This kind of presentation is also termed as 'visual presentation' or ‘charting’.

In this unit, you will learn about the importance of visual presentation of research data and some of the reasons why diagrammatic and graphic presentation of data is so widely used. You will also study the different kinds of diagrams and graphs, which are more popularly used for presenting the data in research work. Also its principles on how to present the frequency distribution in the form of diagrams and graphs. As you are already familiar with graphs and diagrams, we will proceed with further discussions.

7.2 IMPORTANCE OF VISUAL PRESENTATION OF DATA

Visual presentation of statistical data has become more popular and is often used by the researcher and the statistician in analysis. Visual presentation of data means presentation of Statistical data in the form of diagrams and graphs. In these days, as we know, every research work is supported with visual presentation because of the following reasons.

1) **They relieve the dullness of the numerical data:** Any list of figures becomes less comprehensible and difficult to draw conclusions from as its length increases. Scanning of the figures from tables causes undue strain on the mind. The data when presented in the form of diagrams and graphs, gives a birds eye-view of the entire data and creates interest and leaves an impression on the mind of readers for a long period.

2) **They make comparison easy:** This is one of the prime objectives of visual presentation of data. Diagrams and graphs make quick comparison between two or more sets of data simpler, and the direction of curves bring out hidden facts and associations of the statistical data.

3) **They save time and effort:** The characteristics of statistical data, through tables, can be grasped only after a great strain on the mind. Diagrams and graphs reduce the strain and save a lot of time in understanding the basic characteristics of the data.

4) **They facilitate the location of various statistical measures and establish trends:** Graph makes it possible to locate several measures of central tendency such as Median, Quartiles, Mode etc. They help in establishing trends of the past performance and are useful in interpolation or extrapolation, line of best fit, establishing correlation etc. Thus, it helps in forecasting.

5) **They have universal applicability:** It is a universal practice to present the numerical data in the form of diagrams and graphs. In these days, it is an extensively used technique in the field of economics, business, education, health, agriculture etc.

6) **They have become an integral part of research:** In fact, now a days it is difficult to find any research work without visual support. The reason is that this is the most convincing and appealing way of presenting the data. You can find diagrammatic and graphic presentation of data in journals, magazines, television, reports, advertisements etc. After having understood about the importance of visual presentation, we shall move on to discuss about the Diagrams and graphs which are more frequently used in the area of business research.

7.3 DIAGRAMMATIC PRESENTATION

As you know, diagrammatic presentation is one of the techniques of visual presentation of statistical data. It is a fact that diagrams do not add new meaning to the statistical facts but they reveal the facts of the data more quickly and clearly. Because, examining the figures from tables becomes laborious and uninteresting to the eye and also confusing. Here, it is appropriate to state the words of M. J. Moroney, "cold figures are uninspiring to most people. Diagrams help us to see the pattern and shape of any complex situation." Thus, the data presented through diagrams are the best way of appealing to the mind visually. Hence, diagrams are widely used in practice to display the structure of the data in research work.

7.3.1 Rules for Preparing Diagrams

As we have discussed earlier, the prime objective of diagrammatic presentation of data is to highlight their basic hidden facts and relationships. To ensure that the presentation of numerical data is more attractive and effective, therefore, it is essential to keep the following general rules in mind while adapting diagrams in research work. Now, let us discuss them one by one.

1) You must have noted that the diagrams must be geometrically accurate. Therefore, they should be drawn on the graphic axis i.e., 'X' axis (horizontal line) and 'Y' axis (vertical line). However, the diagrams are generally drawn on a plain paper after considering the scale.
2) While taking the scale on 'X' axis and 'Y' axis, you must ensure that the scale showing the values should be in multiples of 2, 5, 10, 20, 50, etc.
3) The scale should be clearly set up, e.g., millions of tons, persons in Lakhs, value in thousands etc. On 'Y' axis the scale starts from zero, as the vertical scale is not broken.
4) Every diagram must have a concise and self explanatory title, which may be written at the top or bottom of the diagram.
5) In order to draw the readers' attention, diagrams must be attractive and well propotioned.
6) Different colours or shades should be used to exhibit various components of diagrams and also an index must be provided for identification.
7) It is essential to choose a suitable type of diagram. The selection will depend upon the number of variables, minimum and maximum values, objects of presentation.

Self Assessment Exercise A

List out the importance of visual presentation of statistical data.

.....

.....

.....

.....

7.4 TYPES OF DIAGRAMS

Generally, diagrams are classified on the basis of their length, width and shape. There are various types of diagrams namely, one dimensional diagrams, two dimensional diagrams, three dimensional diagrams, charts, pictograms, cartograms etc. However, in this unit, we will discuss the important types of diagrams, which are more frequently used in social science research in general, particularly in business research. Therefore, we have restricted ourselves to study only one dimensional bar diagrams, pie diagrams, and structure diagrams.

7.5 ONE DIMENSIONAL BAR DIAGRAMS

Bar refers to a thick line. Under this type of construction only one dimension i.e., length is taken into account for the purpose of comparison and observance of fluctuations in growth. The length of each bar is proportionate to the magnitude of the data. The width is not related to the magnitude of the data. Generally the width is given for the purpose of visual effect and attractiveness. The width of each bar and the gap between one bar to another bar must be uniform. Mention the respective figures at the top of every bar, particularly when the scale is too narrow, so that the reader knows the figures without consulting the scale of the diagram.

A large number of one dimensional diagrams are available for presenting data. Such as line diagram, simple bar diagram, multiple bar diagram, sub-divided bar diagram, percentage bar diagram, deviation bar diagram etc. We shall, however, study only the simple bar diagram, multiple bar diagram, and sub-divided bar diagram. Let us study these three kinds of diagrams with the support of relevant illustrations.

7.5.1 Simple Bar Diagram

In a Simple bar diagram, the data related to one variable is depicted. Such as, profits, investments, exports, sales, production etc.

This type of diagram may be drawn either vertically or horizontally. Both positive and negative values can be presented. In such a case, if bars are constructed vertically, the positive values are taken on the upper side of horizontal axis while the negative values are taken on its lower side. On the other hand if the bars are constructed horizontally, the positive values are taken on the right hand side of the vertical axis and the negative values are considered on its left side. These type of construction of bars are also called deviation bar diagram. The simple bar diagram is very easy to prepare and to understand the level of fluctuations from one situation to another. It should be kept in mind that, only length is taken into account and not width. Width should be uniform for all bars and the gap between each bar is normally identical. Let us consider the following illustrations and learn how to present the given data in the form of simple bar diagrams vertically and horizontally.

Illustration-1

Prepare a Simple Bar Diagram from the Following Data Relating to Tea Exports.

Year	**1995-96**	**1996-97**	**1997-98**	**1998-99**	**1999-00**	**2000-01**	**2001-02**
Exports (*In Million kgs.*)	167	209	410	316	192	215	160

Solution: The quantity of tea exported is given in million kgs. for different years. A simple bar diagram will be constructed with 7 bars corresponding to the 7 years. Now study the following vertical construction of bar diagram by referring the guide lines for construction of simple bars, as explained in section 7.5.1.

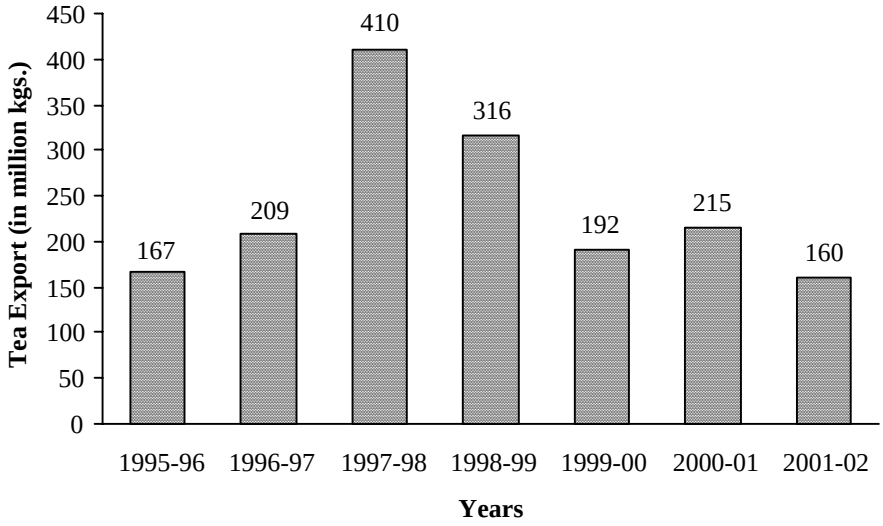

Figure 7.1: Simple Bar Diagram Showing the Tea Exports in Different Years.

Illustration-2

The following data relates to the Profit and Loss of different industries in 1999-2002. Present the data through simple bar diagram.

Industry :	Cement	Oil	Textile	Sugar	Garments
Profit/Loss (*Rs. In lakhs*)	48	25	–12	14	–24

Solution : The given data represents positive and negative values i.e., profit and loss. Let us draw the bars horizontally. Observe fig: 7.2 carefully and try to understand the construction of simple bars horizontally.

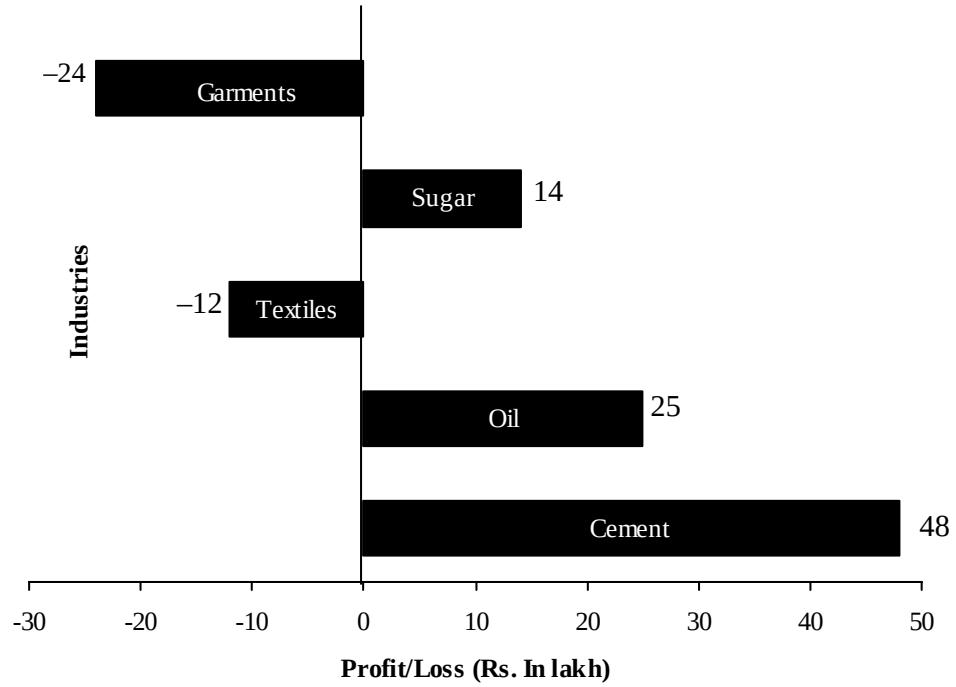

Figure 7.2: Simple Bar Diagram Showing the Profit and Loss of Different Industries During 1999-02

Self Assessment Exercise B

Represent the following data related to the surplus/deficit of Balance of Trade over a period, by simple bar diagram.

Years	1997	1998	1999	2000	2001	2002	2003
Surplus (+)/deficit (–) (*In million $*)	+ 34	+ 14	– 12	– 4	+ 6	– 12	– 20

.....

.....

.....

.....

.....

.....

.....

.....

.....

7.5.2 Multiple Bar Diagram

In this type of diagram, two or more than two bars are constructed side by side horizontally for a period or related phenomenon. This type of diagram is also called Compound Bar or Cluster Bar Diagram. The technique of preparing such a diagram is the same as that of simple bar diagram. This diagram, on the one hand, facilitates comparison of the values of different variables in a set and on the other, it facilitates the comparison of the values of the same variable over a period of time or phenomenon. To facilitate easy comparison, the different bars of a set may be coloured or shaded differently to distinguish

between them. But the Colour or shade for the bars representing the same variable in different sets should be the same.

Let us consider the following illustration and learn the method of presentation of the data in the form of a multiple bar diagram.

Illustration-3

Depict the following data in a multiple bar diagram.

Foreign Investment – Industry Wise Inflows

(Rs. in crores)

Industry	Years		
	1997-98	**1998-99**	**1999-2000**
Chemical	956	1580	523
Engineering	2155	1800	1423
Services	1194	1550	506
Food	418	78	525

Solution : The data relates to the Foreign Investment inflow of four industries during 1997-2000 (three years). Therefore, three sets of bars should be drawn, each set represents one year. In each set there should be four bars representing four sectors (Chemical, Engineering, Services and Food). Let us draw the multiple bars with the help of the procedure explained in subsection 7.5.2. Study the diagram carefully and learn how this type of diagram is drawn.

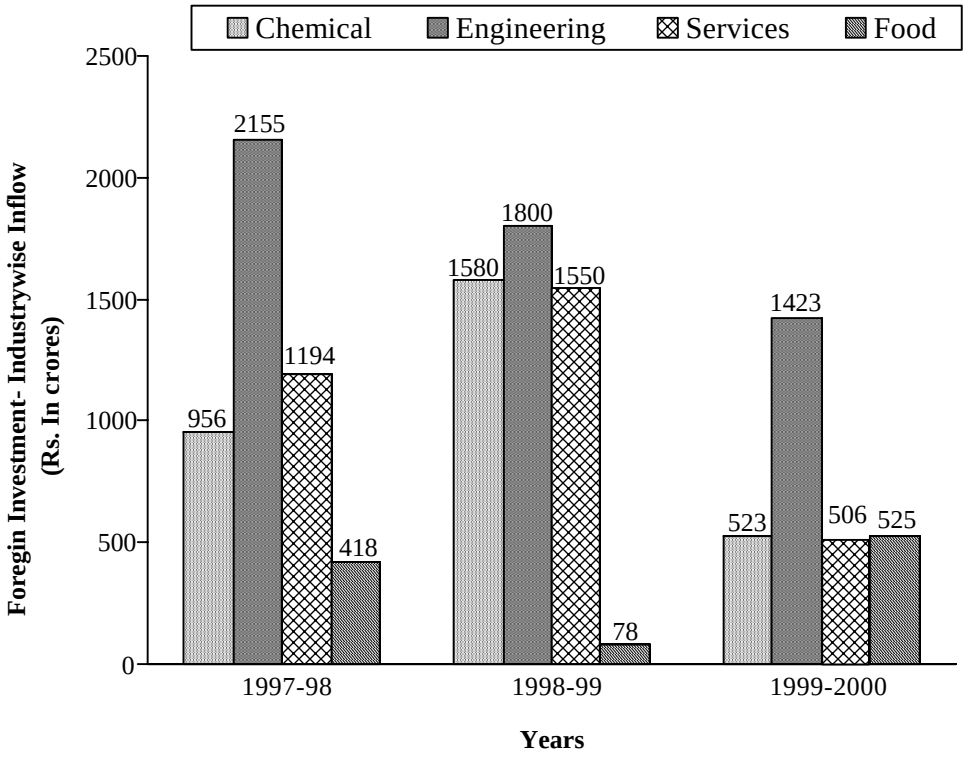

Figure 7.3: Multiple Bar Diagram Showing the Inflow of Foreign Investment in Selected Sectors During 1997-2000

Self Assessment Exercise C

The following table relates the Indian Textile Exports to different countries

Countries	Year		
	1997-98	1998-99	1999-2000
USA	746.13	759.36	882.41
Germany	366.01	300.46	338.88
UK	403.07	337.94	341.42
Italy	241.64	233.14	215.48
Korea (Republic)	127.00	88.30	185.13

i) Represent the data by Multiple bar diagram.

ii) Which aspects of the distribution does this diagram emphasize?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

7.5.3 Sub-divided Bar Diagram

In this diagram one bar is constructed for the total value of the different components of the same variable. Further it is sub-divided in proportion to the values of various components of that variable. This diagram shows the total of the variables as well as the total of its various components in a single bar. Hence, it is clear that the sub-divided bar serves the same purpose as multiple bars. The only difference is that, in case of the multiple bar each component of a variable is shown side by side horizontally, where as in construction of sub-divided bar diagram each component of a variable is shown one upon the other. It is also called a component bar diagram. This method is suitable if the total values of the variables are small, otherwise the scale becomes very narrow to depict the data. To study the relative changes, all components may be converted into percentages and drawn as sub-divided bars. Such a bar construction is called a **sub-divided percentage bar**. The limitation is that all the parts do not have a common base to enable us to compare accurately the various components of a set.

Let us take up an illustration to understand presenting of the data in the form of sub-divided bar diagram.

Illustration-4

The following data relates to India's exports of electronic goods to different countries during 1994-98. Represent the data by sub-divided bar diagram.

(Rs. in Crores)

Years	Country					Total
	USA	Hong Kong	Malaysia	Singapore	Germany	
1994-95	210	86	56	275	91	718
1995-96	378	105	159	467	118	1227
1996-97	789	189	221	349	93	1641
1997-98	880	248	175	327	90	1720
1998-99	900	220	200	350	130	1800

Solution : For construction of sub-divided bar diagram, first of all, we must obtain the total export value of the five countries in each year. However, in the above illustration of different countries, total exports in each year are given. Construct sub-divided bar diagram. Now study figure 7.4 carefully and understand the construction.

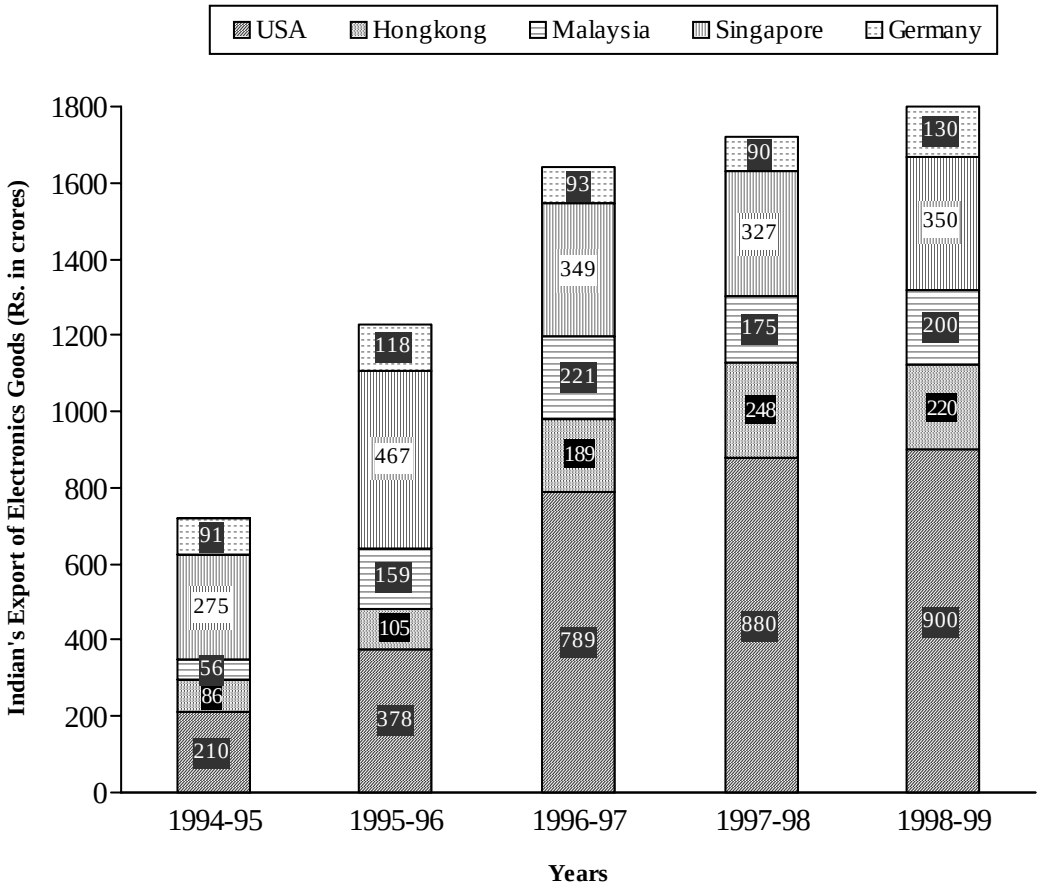

Figure 7.4: Sub-divided Bar Diagram Showing the India's Exports of Electronic Goods to Different Countries During 1994-99.

Self Assessment Exercise D

Draw sub-divided bar diagram for the following table. Do you agree that this diagram is more effective for comparison of figures rather than the Multiple bar diagram? Justify your opinion.

Item-wise Exports of Leather Products from India (1997-2000)

$ Million

Items	Year		
	1997-98	**1998-99**	**1999-2000**
Finished leather	296.19	268.38	239.00
Leather footwear	240.77	241.00	299.77
Leather goods	387.79	411.00	385.25
Leather garments	425.72	381.94	318.94
Others	26.00	33.00	36.00

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

7.6 PIE DIAGRAM

Pie diagrams are generally used to show per cent breakdowns. For instance, we can show how the budget is allocated under different heads. A pie diagram is a sub-divided circle. The area of different sub-divisions in pie diagrams are in the proportion of the data to be represented. While making comparision, pie diagrams should be used on a percentage basis and not on an absolute basis.

In constructing a pie diagram the first step is to convert the various values of components of the variable into percentages and then the percentages transposed into corresponding degrees. The total percentage of the various components i.e., 100 is taken as 360° (degrees around the centre of a circle) and the degree of various components are calculated in proportion to the percentage values of different components. It is expressed as:

$$\frac{360^{o}}{100} \times \text{component's percentage}$$

It should be noted that in case the data comprises of more than one variable, to show the two dimensional effect for making comparison among the variables, we have to obtain the square root of the total of each variable. These square

roots would represent the radius of the circles and then they will be sub-divided. A pie diagram helps us in emphasizing the area and in ascertaining the relationship between the various components as well as among the variables. However, compared to a bar diagram, a pie diagram is less effective for accurate interpretation when the components are in large numbers. Let us draw the pie diagram with the help of the data contained in the following table.

Illustration 5

A researcher made an enquiry about the sources of price information tapped from 550 sample farmers in a regulated agricultural market as given below. Present the data in the form of pie diagram and comment.

Source of Price Information	No. of farmers
Radio	50
Daily papers	60
Local traders	100
Co-framers	310
Personal visits	20
Market office	10

Solution : The number of farmers, who have expressed their sources of collecting information for price of agricultural products have to be converted into the corresponding percentages and then after that into degrees as shown below. Draw the circle and then measure points on the circumference representing the degrees of each source with the help of protractor. Let us first calculate the corresponding percentages and then convert into degrees in order to draw an appropriate pie chart.

Source of Price Information	No. of farmers	Percentage of No. of farmers	Degree of angle
Radio	50	9.1	33°
Daily wages	60	10.9	39°
Local traders	100	18.2	66°
Co-farmers	310	56.4	203°
Personal visits	20	3.6	13°
Market office	10	1.8	6°
Total	550	100.0	360°

After calculating the degrees of various components, depict them in a circle as shown in Figure 7.5.

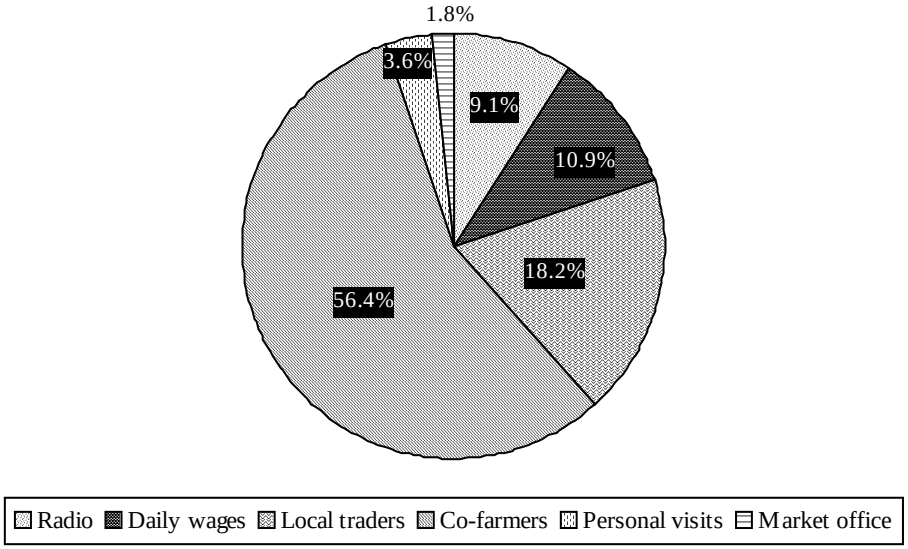

Figure 7.5: Sources of Price Information of Regulated Agricultural Market Tapped by the Farmers.

The pie diagram reveals that the majority of farmers seek price information from co-farmers, which represents about 56.4% among the sample farmers (550). Next to the source of co-farmers, they collect the information through local traders i.e., 18.2%. The study also reveals that the market office plays a very insignificant role which represents about 1.8% only.

Self Assessment Exercise E

Construct a pie diagram to describe the following data :

Reasons for Buying Face Cream

Reasons	No. of Respondents
Seen it advertised	280
Seen it on the counter	160
Reasonably priced	100
Scent appealed	70
Beneficial to skin	180
Recommendation	210
Total Respondents	1000

What features of this distribution does your pie diagram mainly illustrate?

7.7 STRUCTURE DIAGRAMS

There are several important diagram formats that are used to display the structural information (qualitative) in the form of charts. The format depends upon the nature of information. Under these type of diagrams we will discuss two different diagrams, i.e., (1) Organisational Charts and (2) Flow Charts.

7.7.1 Organisational Charts

These types of charts are most commonly used to represent the internal structure of organisations. There is no standard format for these kind of diagrams as the design of the diagram depends on the nature of the organization. A special format is used in the following illustration which relates to the organisational structure of the IGNOU. Study the Fig. 7.6 and try to understand the preparation of this kind of diagram relating to other organisations.

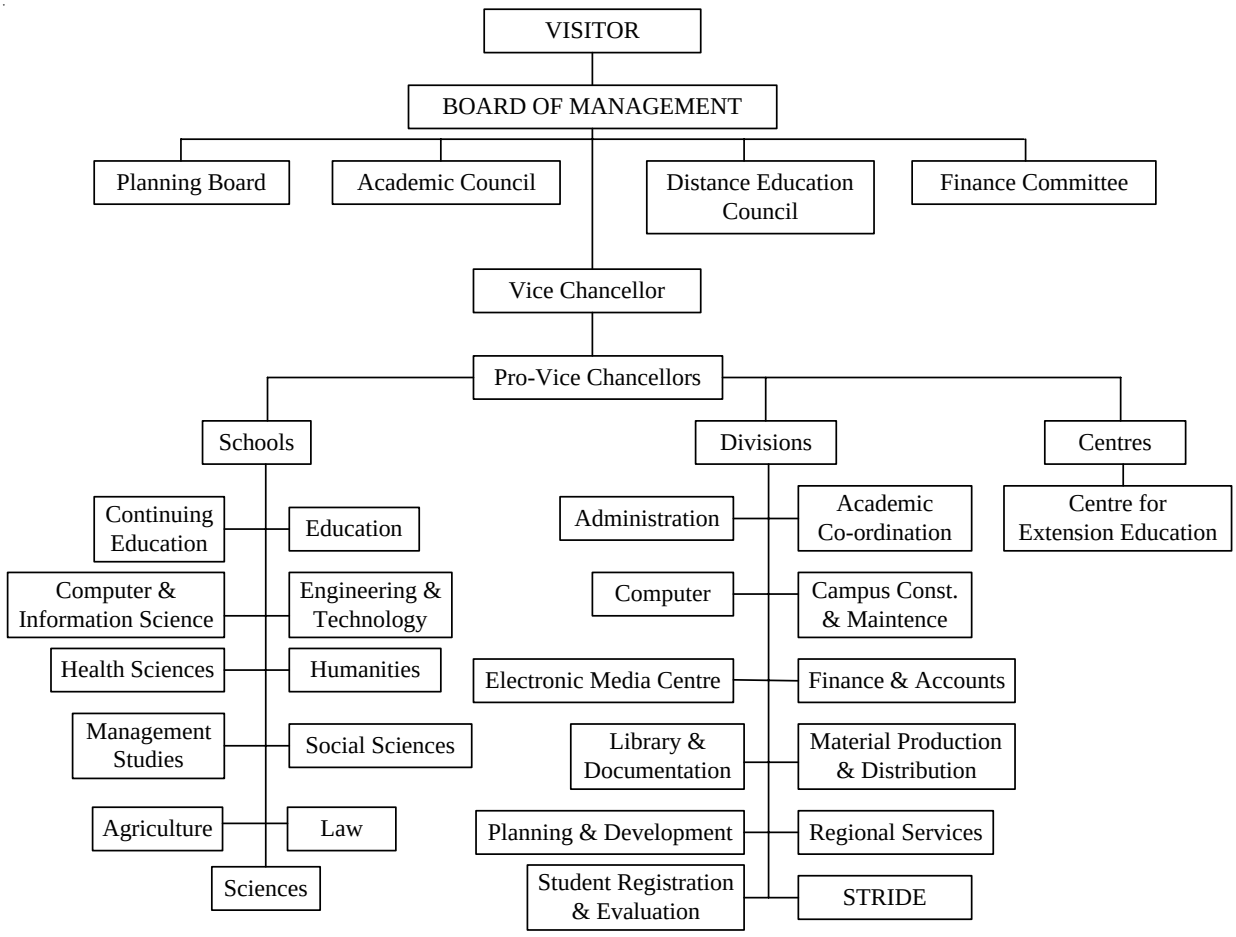

Figure 7.6: Organisational Set Up of the IGNOU

7.7.2 Flow Charts

Flow charts are used most commonly in any situation where we wish to represent the information which flows through different situations to its ultimate point. These charts can also be used to indicate the flow of information about various aspects i.e., material flow, product flow (distribution channels), funds flow etc.

The following Figure 7.7 relates to the marketing channels for fruits, which will give you an understanding about flow charts.

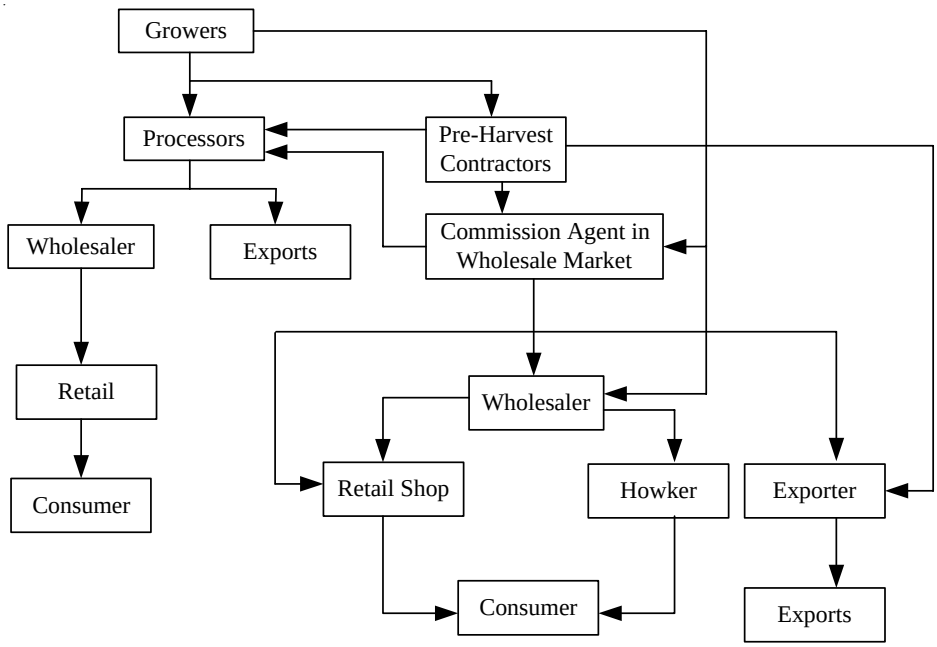

Figure 7.7 : Marketing channels for fruits

Self Assessment Exercise F

1) Prepare an organizational chart of any organization of your own choice.

2) Represent the different stages in processing of Data (Unit 6 of this course) through a flow chart.

7.8 GRAPHIC PRESENTATION

So far we have discussed about one of the techniques of visual presentation of data i.e., diagrammatic presentation. You will appreciate as to how such presentation eliminates the dullness of data and makes it more interesting, and also helps in comparison between two or more frequency distributions. Now, we will study another important technique of visual presentations of statistical data i.e., graphic presentation. You might have seen the graphic representation of stock index, cricket score, production trends etc., in various magazines and on television. Everybody, irrespective of whether he/she is a layman or an expert, has a natural fascination for appropriate graphical presentation of data which remains an essential part of research methodology. The graphic presentation of data leaves an impact on the mind of readers, as a result of which it is easier to draw trends from the statistical data.

The shape of a graph offers easy and appropriate answers to several questions, such as:

- The direction of curves on the graph makes it very easy to draw comparisons.
- The presentation of time series data on a graph makes it possible to interpolate or extrapotrate the values, which helps in forecasting.
- The graph of frequency distribution helps us to determine the values of Mode, Median, Quartiles, percentiles, etc.
- The shape of the graph helps in demonstrating the degree of inequality and direction of correlation

For all such advantages it is necessary for a researcher to have an understanding of different types of graphic presentation of data. In practice, there are a variety of graphs which can be used to depict the data. However, here we will discuss only a few graphs which are more frequently used in business research.

Broadly, the graphs of statistical data may be classified into two types, one is graphs of time series, another is graphs of frequency distribution. We will discuss both these types, after studying the parts of a graph.

Parts of a Graph

The foremost requirement for a researcher is to be aware of the basic principles for using the graph paper for presentation of statistical data graphically.

Conventionally, graphs are drawn on a graph paper. Two perpendiculars are drawn which intersect each other at right angles. This intersecting point is called the origin point or the ‘zero’ point. The horizontal line is known as ‘X’ axis (ordinate) on which independent variables are shown while the vertical line is known as ‘Y’ axis (abscissa) on which dependent variables are shown. The graph paper is thus divided into four parts, termed as “quadrants”. These quadrants are meant to depict the positive values and negative values of X variable and Y variable. By observing the following Chart 7.1. You will understand clearly about the purpose of quadrants of a graph.

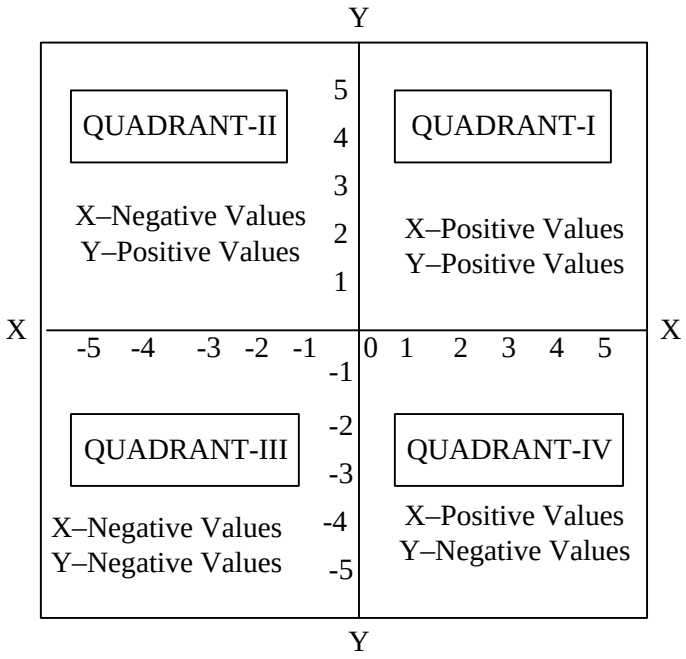

Chart 7.1 : Parts of a Graph

After understanding the above parts of a graph, let us study the different types of graphs.

7.9 GRAPHS OF TIME SERIES

A time series is a set of values of a variable or variables arranged over a period of time. For example, the data relating to the production, sales expenditure, exports etc., during the last ten years. Thus a graph of time series is prepared to show the values of one or more than one variables over a period of time. This type of graphs are also termed as time graphs or historigrams, because history is represented graphically. These graphs are helpful in studying the changes over a period of time and forecasting.

Historigrams can be constructed in two ways: 1) on a natural scale, (arithmatic scale), 2) on a ratio scale. The natural scale graph reflects the changes in absolute values over a period of time, where as the ratio scale graph reflects the relative changes over a period of time. In this unit, however, we study the historigarms on a natural scale which is generally used in business research.

Construction of Historigrams on Natural Scale

Natural scale graphs are used to show the absolute values or relative values in terms of percentages, such as index numbers of a time series.

The following Principles should be kept in mind while constructing historigrams so that the reader should not have to search through the text in order to understand a graph.

1) On X-axis we take the time as an independent variable and on Y axis the values of data as dependent variable. Plot the different points corresponding to given data; then the points are joined by a straight line in the order of time.
2) Equal magnitude of scale must be maintained on X-axis as well as on Y-axis.
3) The Y-axis normally starts with zero. In case, there is a wide difference between the lowest value of the data and zero (origin point), the Y-axis can be broken and a false base line may be drawn. However, it will be explained under the related problem in this section.

4) If the variables are in different units, double scales can be taken on the Y axis.

5) The scales adopted should be clearly indicated and the graph must have a self-explanatory title.

6) Unfortunately, graphs lend themselves to considerable misuse. The same data can give different graphical shapes depending on the relative size of two axes. In order to avoid such misrepresentations the convention in research is to construct graphs, wherever possible, such that the vertical axis is around 2/3 to 3/4 the length of the horizontal.

After having learnt about the principles for construction of historigrams, we move on to discuss the types of historigrams. There are various types which have been developed. Among them the frequently used graphs are one variable graphs and more than one dependent variable graphs. We will now look at the construction of these graphs.

7.9.1 Graph of one Dependent Variable

When there is only one dependent variable, the values of the dependent variable are taken on Y axis, while the time is taken on X-axis. Study the following illustration carefully and try to understand the method of construction for one dependent variable historigrams.

Illustration 6

The following data relates to India's exports to USA during the period of 1994-2000. Represent the data graphically.

Year	1994	1995	1996	1997	1998	1999	2000
Exports (*In million $*)	5310	5726	6170	7322	8237	9071	10687

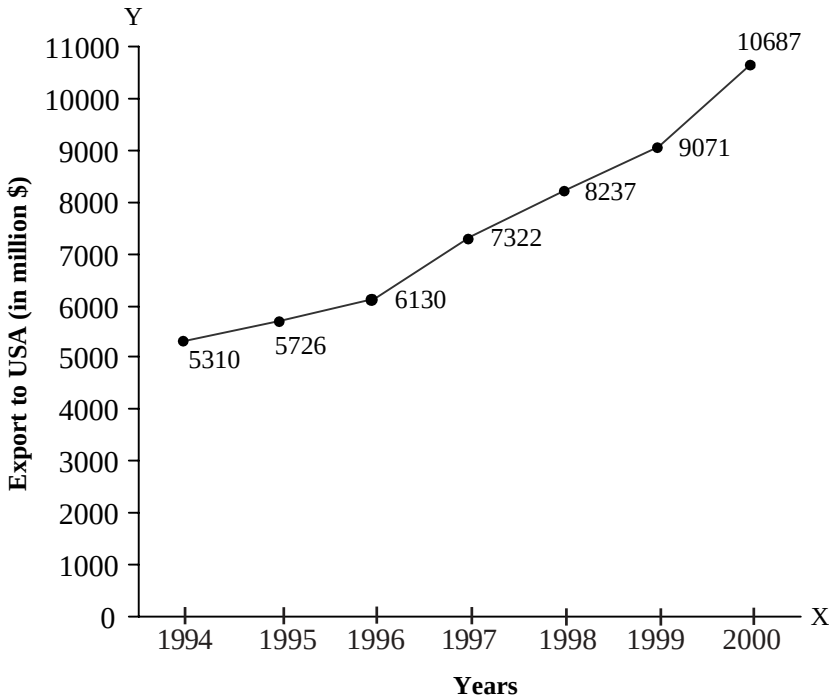

Figure 7.8: Historigram Showing India's Exports to USA During 1994-2000.

False base line

In the above graph (Figure 7.8), the scale on Y axis has been taken as 1cm = 1,000 million starting from the origin point i.e., zero. Consequently, the portion of the graph paper, on which lies the scale between zero and the smallest value of the data (5310) is omitted and only the above half of the graph paper is used to depict the data because, there is a wide difference between zero and the lowest value of the given data. Therefore, the curve which is drawn is not significant to understand the fluctuations. In such a situation, in order to use the space of the graph paper effectively, it is mandatory to draw a false base line. By using the false base line, minor fluctuations are amplified and they become clearly visible on the graph. The false base line breaks the continuity of the scale of Y axis from the origin point, i.e., zero by drawing a horizontal wave line in between the zero and the first centimeter on the scale of Y axis.

Let us consider the above illustration-6 to represent the data graphically by drawing a false base line, so that we can practically understand and appreciate the importance of the false base line for effective presentation of data on a graph sheet. Study Figure 7.9 carefully.

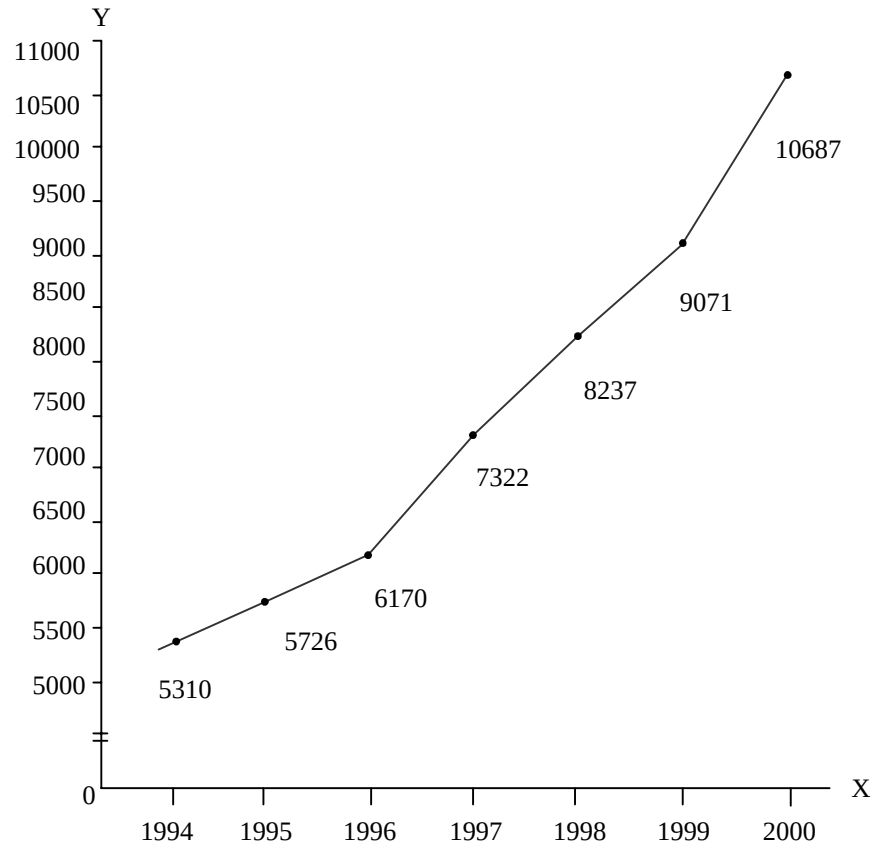

Figure 7.9 : Historigram Showing India's Exports to USA During 1994-2000.

7.9.2 Graph of more than one dependent variable

When the data of time series relate to more than one dependent variable, curves must be drawn for each variable separately. These graphs are prepared in the same manner as we prepare one dependent variable historigram. Let us consider the following data to construct historigrams. Study Figure 7.10 carefully and understand the procedure for preparation of this type of graph.

Illustration-7

Years	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Sales *(In lakh)*	31	58	42	65	75	80	72	96	83	98
Cost of Sales *(Rs. In Lakh)*	42	50	48	55	82	75	62	80	67	73
Profit/ Loss	–11	+8	–6	+10	–7	+5	+10	+20	+16	+25

Solution : The given data comprises of three variables, so, we have to draw a separate curve for each variable. In this graph, it is not necessary to draw false base line because the minimum value is close to the point of origin (zero). For easy identification, each curve is marked differently.

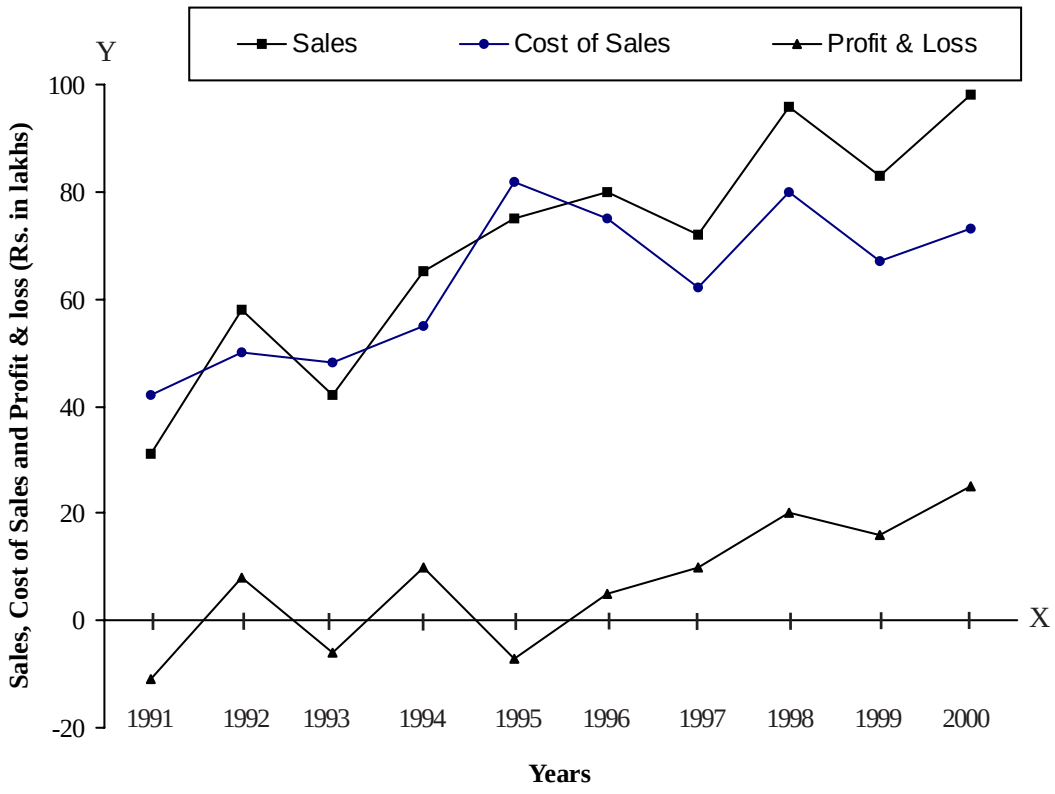

Figure 7.10 : Historigram Showing Sales, Cost of Sales and Profit/Loss of a Company During 1991-2000

The above graph clearly reveals that with passage of time the profits are rising after 1996, even though the sales are fluctuating slightly.

Self Assessment Exercise G

Represent the following data graphically by showing Exports, Imports and Balance of Trades.

Years	1994	1995	1996	1997	1998	1999	2000
Exports *(In lakh)*	52	46	62	58	72	89	92
Imports *(In Rs. Lakhs)*	48	52	69	51	60	80	96

7.10 GRAPHS OF FREQUENCY DISTRIBUTION

We have seen in Unit-6 the presentation of frequency distribution in the form of tables. Frequency distribution can also be presented in the form of graphs. Such graphs give a better understanding and provide illustrative information to readers than the data in tabular form. It is true that effective graphs can markedly increase a reader's comprehension of complex data sets. Compared to tables, graphs of frequency distribution are helpful in identifying the characteristics and relationships of the data. These graphs are also useful in locating the positional averages such as mode, median, qualities etc. In a continuous frequency distribution, class-limits/mid-values are taken on X axis and the frequency on the Y-axis. The vertical axis (Y-axis) is not broken, thus the false base line cannot be taken.

A frequency distribution can be portrayed by means of Histogram, frequency polygon, ogive curves and scatter diagram. However, the scatter diagram will be discussed in Unit-10 of this course.

Let us study the procedure involved in the preparation of these types of graphs.

7.10.1 Histogram and Frequency Polygon

Histogram: The graph usually drawn to represent a frequency distribution is called a Histogram. A histogram is a set of rectangles (vertical bars) each proportionate in width to the magnitude of a class interval and proportionate in area to the number of frequencies concerning the classes' intervals. In a histogram, the variables (class-intervals) are always shown on X-axis and the frequencies are taken on the Y-axis. In constructing a histogram there should

not be any gap between two successive rectangles, and the data must be in exclusive form of classes. However, we cannot construct histogram for distribution with open-end classes and it can be quite misleading if the distribution has unequal class intervals.

The value of mode can be determined from the histogram. The procedure for locating the mode is to draw a straight line from the top right corner of the highest rectangle (Modal Class) to the top right corner of the preceding rectangle (Pre Modal Class). Similarly, draw a straight line from the top left corner of the highest rectangle to top left corner of the succeeding rectangle (Post Modal Class). Draw a perpendicular from the point of intersection of these two straight lines to X-axis. The point where it meets the X-axis gives the value of mode. This is shown in Figure 7.11. However, graphic location of Mode is not possible in a multi-distribution.

Frequency Polygon: Polygon means 'many-angled' diagram. This is another way of depicting a frequency distribution graphically. It facilitates comparison of two or more frequency distributions. Frequency polygon can be drawn either from the histogram or from the given data directly.

The procedure for the construction of a frequency polygon by histogram is to first draw the histogram, as explained earlier, of the given data. Then, put a dot at the mid-point of the top horizontal line of each rectangle bar and join these dots by straight lines.

Another way of drawing frequency polygon is to obtain the mid-values of class intervals and plot them on X-axis. Mark frequency along the Y axis. Then, plot the frequency values corresponding to each mid point and connect them through straight lines. The area left outside is just equal to the area included in it. Hence, the area of a polygon is equal to the area of histogram. The difference between the histogram and the polygon is that the histogram depicts the frequency of each class separately where as the polygon does it collectively. The histogram is usually associated with the data of discrete series, while frequency polygon is for continuous series data.

Let us, now, take up an illustration to learn how to draw a histogram, and frequency polygon practically and also determine the mode. The data relates to the sales of computers by different companies.

Illustration-8

Sales *(Rs. In crores)*	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of Companies	8	20	35	50	90	70	30	15

Solution : For drawing histogram, as explained earlier, we have to show sales on X - axis and number of companies on Y-axis by selecting a suitable scale. For drawing frequency polygon, plot dots on the top middle of each rectangle, and join them by straight lines.

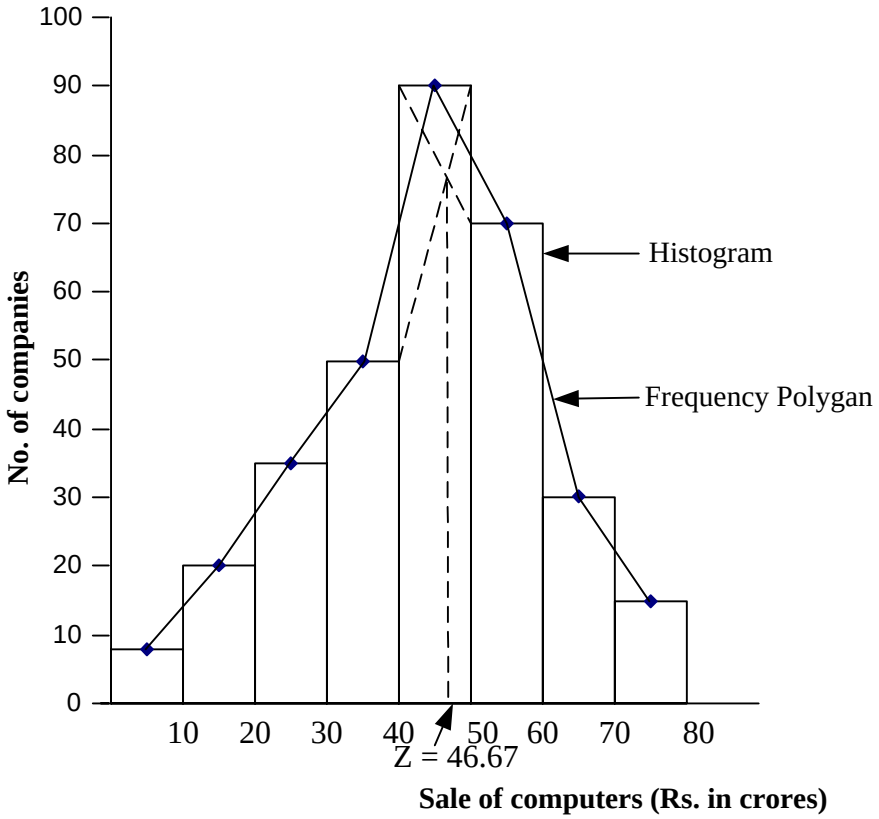

Figure 7.11: Histogram and Frequency Polygon for Computer Sales of Various Companies

Remark: The calculation of Mode will be discussed in Unit 8 of this block.

Self Assessment Exercise H

The monthly production of units by a sample of 200 workers in a bulbs manufacturing firm is given in the following table.

Output (Units)	200-225	225-250	250-275	275-300	300-325	325-350	350-375	375-400
No. of Workers	12	21	25	40	49	28	17	8

i) Draw a historgram and frequency polygon.

ii) Determine the mode graphically.

7.10.2 Cumulative Frequency Curves

Some times we are interested in knowing how many families are there in a city, whose earnings are less than Rs. 5,000 p.m. or whose earning are more than Rs. 20,000 p.m. In order to obtain this information, we have first of all to convert the ordinary frequency table into cumulative frequency table. When the frequencies are added they are called cumulative frequencies. The curves so obtained from the cumulative frequencies are called 'cumulative frequency curves', popularly known as "ogives". There are two types of ogives namely less than ogive, and more than ogive. Let us know about the procedure involved in drawing these two ogives.

In less than ogive, we start with the upper limit of each class and the cumulative (addition) starts from the top. When these frequencies are plotted we get less than ogive. In case of more than ogive we start with the lower limit of each class and the cumulation starts from the bottom. When these frequencies are plotted we get more than ogive. You should bear in mind that while drawing ogives the classes must be in exclusive form.

The ogives are useful to determine the number of items above or below a given value. It is also useful for comparison between two or more frequency distributions and to determine certain values (positional values) such as mode, median, quartiles, percentiles etc. Let us take up an illustration to understand how to draw ogives practically. Observe carefully the procedures involved in it.

Note: Mode and Median are explained in Unit 8. Similarly, quartiles are in Unit 9. This illustration can be better understood only after studying those units.

Illustration-9

The following data relates to the monthly operating expenses incurred by a sample of 200 small-scale industrial units in a city. You are required to draw ogives and locate the Q_1, Q_3 and Median (Q_2).

Operating Expenses (Rs. In thousands)	No. of Units
0-20	7
20-40	18
40-60	22
60-80	34
80-100	53
100-120	26
120-140	18
140-160	10
160-180	7
180-200	5

Solution : To depict "less than" and "more than" cumulative frequency curves (ogives), first, we have to convert the above distribution into "less than" and "more than" cumulative frequency distribution. Study carefully the procedure for conversion of ordinary frequency into cumulative frequencies as shown below:

"Less than" Method		"More than" Method	
Operating Expenses (Rs. In '000)	Frequency	Operating Expenses (Rs. In '000)	Frequency
Less than 20	7	More than 0	200
Less than 40	25	More than 20	193
Less than 60	47	More than 40	175
Less than 80	81	More than 60	153
Less than 100	134	More than 80	119
Less than 120	160	More than 100	66
Less than 140	178	More than 120	40
Less than 160	188	More than 140	22
Less than 180	195	More than 160	12
Less than 200	200	More than 180	5

The cumulative frequencies presented in the above table have the following interpretation.

The 'less than' cumulative frequencies are to be read against upper class limits. In contrast, the 'more than' cumulative frequencies are to be read against lower class boundaries. For instance, there are 7 units with operating expenses of less than Rs. 20,000, there are 160 units with operating expenses of less than Rs. 120,000. On the other hand, there are 153 units with operating expenses more than Rs. 60,000; no units with operating expenses more than or equal to Rs. 2,00,000.

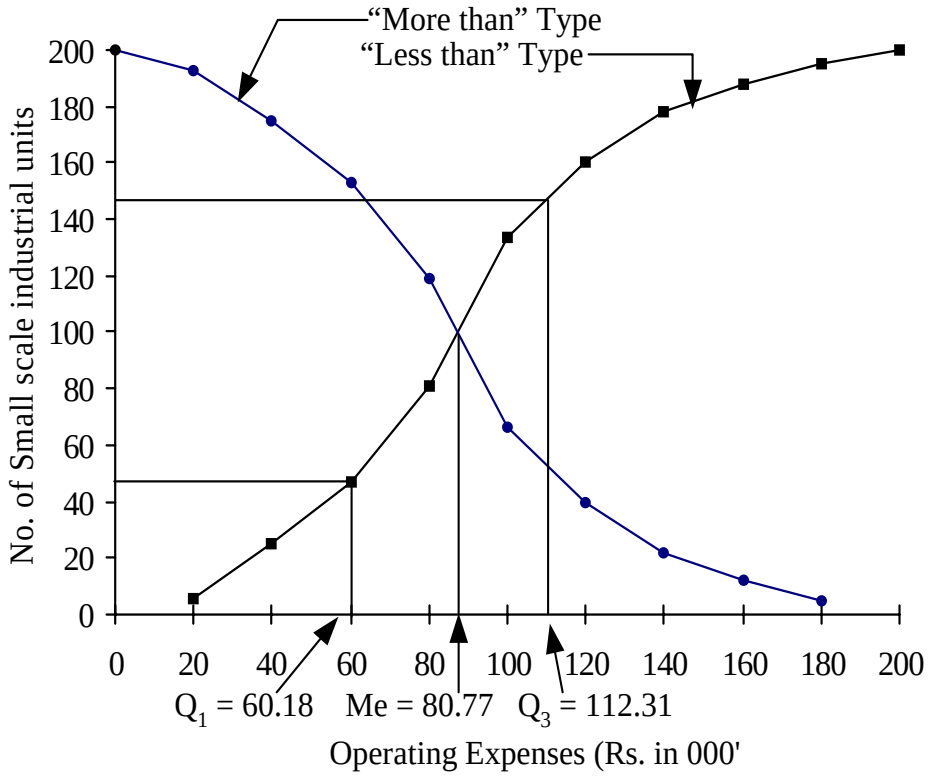

Fig 7.12: 'Less than' and 'More than' Cumulative Frequency Curves Showing the Operating Expenses (Rs. in' 000) of Small Scale Industrial Units.

Now, look at Figure 7.12 which shows both the cumulative curves on the same graph. Study carefully and understand the procedures for drawing ogives.

From the above ogives, the median can be located by drawing a perpendicular from the intersection of the two ogives to X-axis. The point where the perpendicular touches X-axis would be the Median of the distribution. Similarly, the perpendicular drawn from the intersection of the two curves to the Y-axis would divide the sum of frequencies into two equal parts. The values of positional averages like Q_1, D_6, P_{50}, etc., can also be located with the help of an item's value on the less than ogive. In the above figure determination of Q_1 and Q_3 are shown as an illustration.

Self Assessment Exercise I

The following data relates to the monthly expenditure on food incurred by a sample of 150 families in an institution.

Monthly Expenditure	No. of families
2,500 – 3,000	18
3,000 – 3,500	30
3,500 – 4,000	42
4,000 – 4,500	36
4,500 – 5,000	12
5,000 – 5,500	8
5,500 – 6,000	4

a) Draw more-than cumulative frequency curve and less than cumulative frequency curve.

b) Locate the median monthly expenditure using your ogive.

c) How many sample families are approximately spending less than 3,800 on food.

7.11 LET US SUM UP

Statistical data not only requires a careful analysis but also ensures an attractive and communicative display. The work of the researcher is to understand the facts of the data himself/herself and also to present them in such a form that their significance may be understandable to a common reader. In order to achieve this objective, we have, in this unit, discussed the techniques of diagrammatic and graphic presentation of statistical data. Besides, presenting the data in the form of tables, data can also be presented in the form of diagrams and graphs. Such visual presentation of data allows relation between numbers to be exhibited clearly and attractively, makes quick comparison between two or more data sets easier, brings out hidden facts and the nature of relationship, saves time and effort, facilitates the determination of various statistical measures such as Mean, Mode, Median, Quartiles, Standard deviation etc., and establishes trends of past performance. Hence, with the help of the diagrams and graphs the researcher can effectively communicate to readers the information contained in a large mass of numerical data.

We have discussed the method for constructing simple bar diagram, multiple bar diagram, sub-divided bar diagram, pie diagram and structure diagrams.

In graphs, we discussed graphs of time series (Historigrams), graphs of frequency distribution (Histograms, frequency polygon, and cumulative frequency curves). It is essential to keep in mind the basic principles while using the diagrams and graphs for presenting the data.

7.12 KEY WORDS

Bar : Is a thick line where the length of the bars should be proportional to the magnitude of the variable they represent.

Continuous Data : Data that may progress from one class to the next without a break and may be expressed by either fractions or whole number.

Discrete Data : Data that do not progress from one class to the next without break, i.e., where classes represent distinct categories or counts and may be represented by whole numbers only.

False Base Line : A line that is drawn between the origin point (zero) and the first c.m., by breaking Y-axis in case of historigrams. Hence the scale of Y-axis does not start at zero.

Flow Chart : Presents the information which flows through various situations to the ultimate point.

Frequency Polygon : A line graph connecting the mid-points of each class in a data set.

Historigram : A graph of time series.

Histogram : A graph of a frequency distribution, composed of a series of rectangles, each proportional in width to the range of a class interval and proportional in height to the number of observations falling in the class.

Organisational Chart : A diagram specially designed to show the structure of an organisation.

Ogive : A graph of a cumulative frequency distribution.

Pie Diagram : A circle divided into slices showing the relative areas of various components of the variable.

Structure Diagram : Displays the structural data (qualitative) in the form of charts.

7.13 ANSWERS TO SELF ASSESSMENT EXERCISES

A. 1) Relieves the dullness of the numerical data; 2) Facilitates comparison; 3) Saves time and effort; 4) Facilitates the location of various statistical measures and establishes trends; 5) Universal applicability; and 6) Is an integral part of research.

D. No, In sub-divided bar diagram, all items do not have a common base to enable one to compare acurately the various items of the data. Where as in multiple bar diagram, various items of a phenomena have a common base and are placed together side by side. Hence, comparison is very easy rather than in a sub-divided bar diagram.

E. Steps : 1) Find out the percentages of each reason for buying face cream. 2) Convert the percentages into degree of angle. 3) Then depict the percentages in a circle with the help of their respective degree of angles.

7.14 TERMINAL QUESTIONS/EXERCISES

1) Explain the significance of visual presentation of statistical data in research work.
2) Give a brief description of the different kinds of diagrams generally used in business research to present the data.
3) What are structure diagrams? Explain each with an illustration the method of representing the information by different structure diagrams.
4) Explain the principles of constructing a graph of time series. Under which situation the false base line will be used?
5) Survey your own statistics class in terms of the variables age, sex and income. Use the graphing techniques outlined in the unit to describe your results.
6) Represent the following data relating to exports of agricultural and allied products to Russia during 1996-2000 by a suitable diagram.

Year	1996-97	1997-98	1998-99	1999-2000	2000-01
Export *(In $ million)*	340	448	336	333	490

7) Draw a Multiple bar and sub-divided bar diagrams to represent the following data relating to the enrollment of various programmes in an open university over a period of four years and comment on it.

Programme	No. of Candidate enrolled			
	1998-99	1999-2000	2000-01	20001-02
MBA	1,565	2,356	1,924	3,208
M.Com	872	1,208	1,118	1,097
B.A.	1,600	1,220	1,090	987
B.Com	726	948	1,458	1,220

8) Construct a pie diagram to describe the following data which relates to the amount spent on various heads under Rural development programme.

Various heads	Rupees (in crores)
Agriculture	1,280
Rural Industries	450
Public Health	150
Transport	600
Education	325
Housing	40
Public Utilities	10

What features of this distribution does your pie diagram mainly illustrate?

9) The following table gives the Index numbers of wholesale Prices (Average) of Cereals, Pulses and oilseeds over a period of 7 yrs. Compare these prices through a suitable graph.

Years	Cereals	Pulses	Oilseeds
1997	433	398	529
1998	486	420	638
1999	520	415	829
2000	690	524	750
2001	430	415	858
2002	482	358	884
2003	624	494	866

10) Draw histogram and frequency polygon of the following distribution. Locate the approximate mode with the help of histogram.

Weekly wages *(In Rs.)*	100–120	120–140	140–160	160–180	180–200	200–220	220–240
No. of Workers	26	52	87	93	34	26	12

11) The following data relating to sales of 80 companies are given below

Sales (Rs.Lakhs)	No. of Companies
5-15	8
15-25	13
25-35	19
35-45	14
45-55	10
55-65	7
65-75	6
75-85	3

Draw cumulative frequency curves. Determine the number of companies whose sales are:
(i) more than 50 lakhs. (ii) Less than Rs. 30 lakhs (iii) Between Rs. 30 lakhs to Rs. 50 lakhs.

Note: These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

7.15 FURTHER READING

The following text books may be used for more indepth study on the topics dealt with in this unit.

Moskowitz; H. and G.P. Wright 1998 *Statistics for Management and Economics*, Charles E. Merill Publishing Company: Ohio, U.S.A.

Gupta, S.P. and M.P. Gupta, 2000. *Business Statistics*, Sultan Chand & Sons: New Delhi.

Sinha, S.C. and Dhiman, A.K. 2002. *Research Methodology*, Vol. 1. Ess Ess Publication, New Delhi.

George Argyrons. 2000. *Statistics for Social and Health Research with a Guide to SPSS*. Sate Publications. New Delhi.