Lesson 5
Graphic Presentation of Data

Objectives of the lesson:

To develop understanding among students about:

- The meaning of Graphic presentation of the data;
- Need and usefulness of Graphic presentation of the data
- Method of Graphic presentation;
- Various types of Graphs; and
- Appropriateness of various types of Graphs in use.

INTRODUCTION

Diagrams discussed in the last chapter are generally used for the purpose of publicity and propaganda. From the purely statistical point of view their importance is not much. They only give an approximate and rough idea about the level of a phenomenon. From the statistical point of view graphs and charts are much better than diagrams. Diagrams can be used only in those places where two or more quantities have to be compared. If the relationship between two variables is to be studied diagrams would be useless. Such studies can be made with the help of graphs only.

When we study the relationship between two variables the idea is either to study their cause and effect relationship or to study the extent of change in one variable if the other variable changes by a particular amount. Such studies cannot be made by diagrams; graphs are, however, very useful for studying such relationships. The special feature of graphs is that they are more obvious, accurate and precise than diagrams. Graphs are very useful for studying time series and frequency distributions.

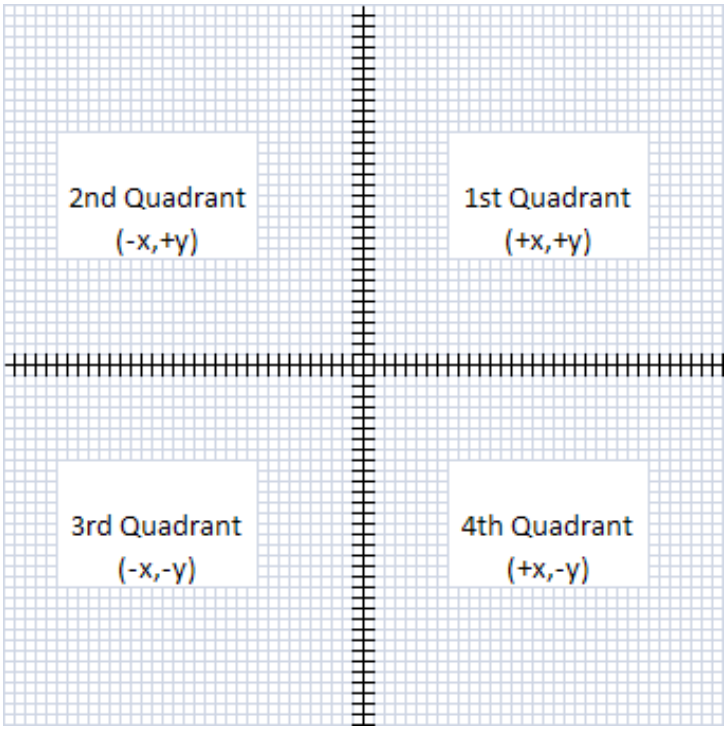
Usefulness of Graphic Presentation:

- Graphs provide an attractive and impressive view.
- Graphs simplify complexity of data.
- Apart from simplicity, it saves the time and energy of the statistician as well as the observer.
- Graphs provide easy comparison of two or more phenomena.

- Graphic method is probably the simplest method of presenting statistical data.
- Graphs need no special knowledge of mathematics to understand a graph.
- Graphs show any trend that may be present and the direction in which the trend may change.
- Graphs provide the basis to measure and locate the statistical measures, like median, mode, quartiles etc.

How to Plot a Graph?

For plot a graph graphs first draw two simple lines which cut each other at right angles. These lines are called axis. The horizontal line is called abscissa or X axis and the vertical line is called ordinate or Y axis. The point at which they cut each other is called the Point of Origin. The following figure gives two such lines.



In the above figure horizontal line is abscissa called X axis and the vertical line is ordinate called as Y axis. The point of their intersection is the point of origin. The figure shows that in first quadrant X and Y both are positive, so we depict positive value of X and Y. Second quadrant shows negative values of X and positive values of Y. In the same fashion, third and fourth quadrants show – X, + Y and + X, - Y respectively.

There are some Rules for Constructing Graphs. These are:

- Graph must have a clear title, indicating the facts presented by the graph.
- Independent variables should be plotted on the X axis and dependent variables on Y axis.
- Choose appropriate scale which may accommodate whole data.
- The graph must not be overcrowded with curves.
- Source of information should be mentioned as footnote.
- Appropriate Index should be given to show the scales and the meaning of different lines.
- If more than one variable is plotted on the same graph, it is necessary to distinguish them by different lines, viz., dotted lines, broken lines, dots, dash, thick, thin lines etc.
- Graph at vertical scale must start from zero. However, if fluctuations are quite small compared to the size of variables, there is no need of showing the entire vertical scale from the origin. The scale just sufficient for the need be shown and for this purpose and a false line may be used.

False Base Line: If the fluctuations in the values of a variable are very small as compared to the size of items, a false base line is used. By its use even minor fluctuations are magnified so that they are clearly visible on the graph. If the size of items is big and if the vertical scale begins from zero the curve would be mostly on the top of the paper and if the differences in the values of various items are not much, it would, more or less, be of the shape of a straight line. I

n false base line the scale from zero to the smallest value of the variable is omitted. Whenever false base line is used it should be very clearly indicated on the graph. Generally in such cases vertical scale is broken in two parts and some blank space is left between them. The lower part of the vertical scale is kept very small and it begins with zero. The upper part begins with a value equal or nearly equal to the smallest value of the variable. To make the breaking of vertical scale prominent usually saw-tooth lines are used.

TYPES OF GRAPHS

Graphs can broadly be classified in two categories:

- Graphs of Frequency Distribution, and
- Graphs of Time Series

(A) Graphs of Frequency Distribution: Frequency distribution can also be presented in the form of graphs in addition to table form. 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. A frequency distribution can be portrayed by means of Histogram, Frequency Polygon, Frequency Curve, and Ogives or Cumulative Frequency Curves.

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.

When the distribution has an unequal class interval, the frequency of each class should be adjusted. For making the adjustment, consider the lowest class interval and make the adjustment to the unequal classes. If one class interval is twice as wide as the lowest class interval, then divide its frequency by two; if three times, divide its frequency by three and so on.

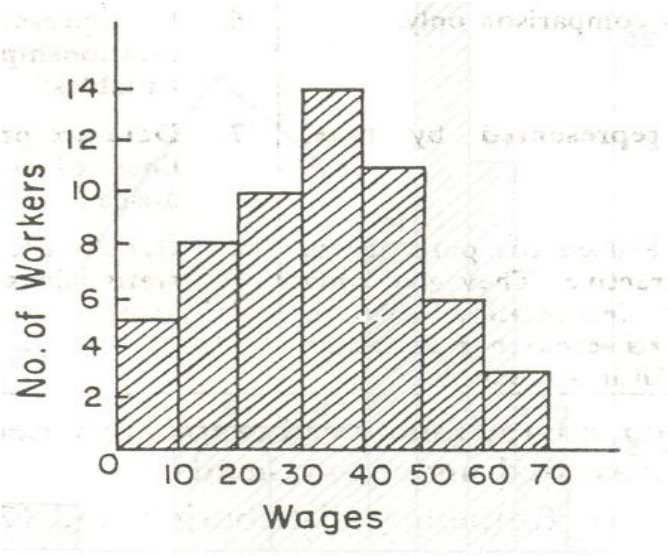
The value of mode can be determined from the histogram. The procedure for locating the mode is to draw a straight line from the top right corner of the highest rectangle (Modal Class) to the top right corner of the preceding rectangle (Pre Modal Class).

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.

Example 1: Draw a histogram for the following data related to hourly wage of workers

Wage (Rs.)	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Workers	5	8	10	14	11	6	3

Solution:

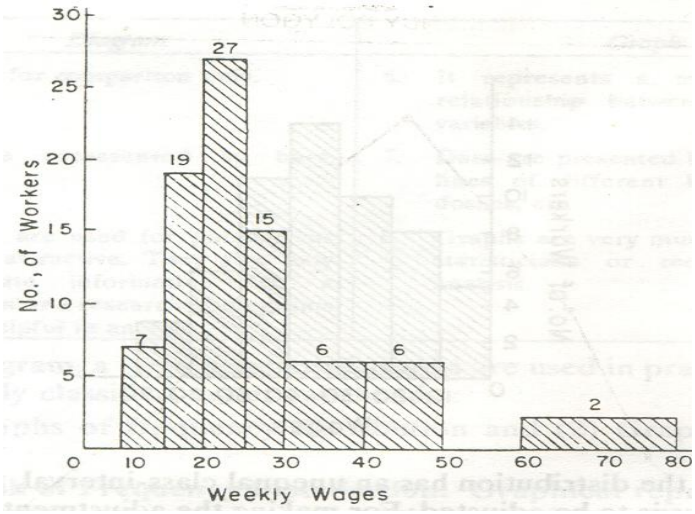


Example 2: Draw a Histogram for the following data related to weekly wage of workers.

Wage ('00 Rs.)	10-15	15-20	20-25	25-30	30-40	40-50	60-80
Workers	7	19	27	15	12	12	8

Solution: Here standard class interval is 5; so, adjust the frequencies according to density.

Wage (in '00 Rs.)	No. of workers	Adjustment (5 = 1)	Frequency density
10-15	7	1	7
15-20	19	1	19
20-25	27	1	27
25-30	15.	1	15
30-40	12	2	6
40-50	12	2	6
60-80	8	4	2



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.

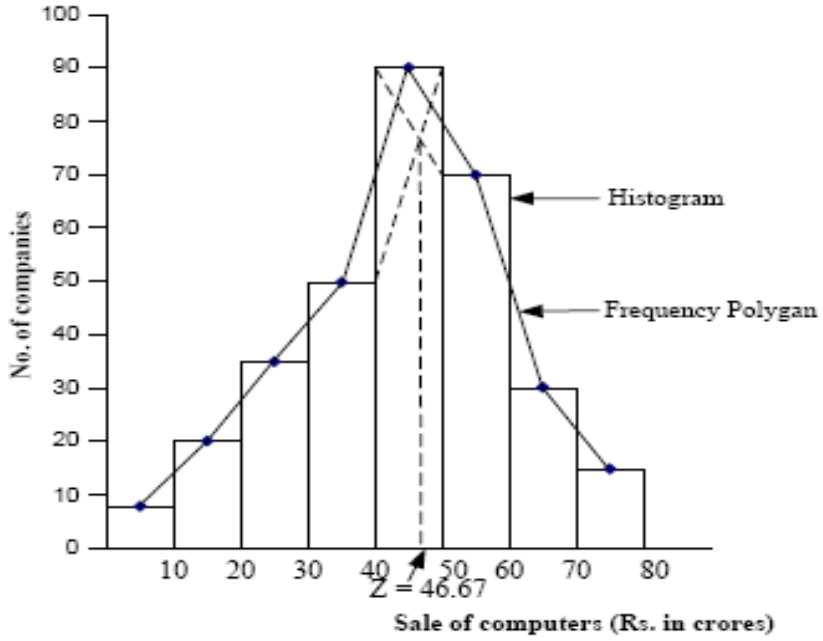
The procedure for the construction of a frequency polygon by histogram is to first draw the histogram, of the given data. Then, put a dot at the mid-point of the top horizontal line of each rectangle bar and join these dots by straight lines.

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.

Example 3: Draw a Histogram and Frequency Polygon for the following data related to sale of Computers. Also find the value of Mode graphically.

Sales (Crore Rs.)	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, we have to show sales on X - axis and number of companies on Y-axis by selecting a suitable scale.



Frequency curve: A frequency curve is drawn by smoothing the frequency polygon. It is smoothed in such a way that the sharp turns are avoided. A frequency polygon, if smoothed further, so as to minimize sudden changes, results into a continuous smooth curve known as frequency or smooth frequency curve. The curve should begin and end at the base line.

Ogives: 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 is 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". 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. We 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.

Example 4: Draw a cumulative frequency graph of the following distribution, showing the monthly wages of a group of workmen, and calculate the values of (i) the median, and (ii) the two quartiles.

Marks	20	21	22	23	24	25	26	27	28
No. of Students	8	10	11	16	20	25	15	9	6

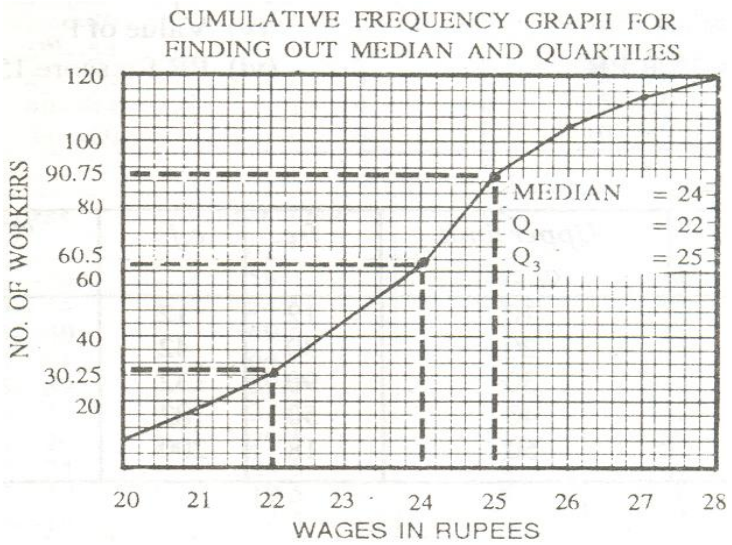
Solution: For finding out median and quartiles the frequencies will be cumulated:

Marks	20	21	22	23	24	25	26	27	28
No. of Students	8	10	11	16	20	25	15	9	6
Cum. Frequency	8	18	29	45	65	90	105	114	120

Here, Median = (120+1)/2 = Value of the 60.5th item

First Quartile = (120+1)/4 = Value of the 30.25th item

Third Quartile = 3(120+1)/4 Value of the 90.75th item

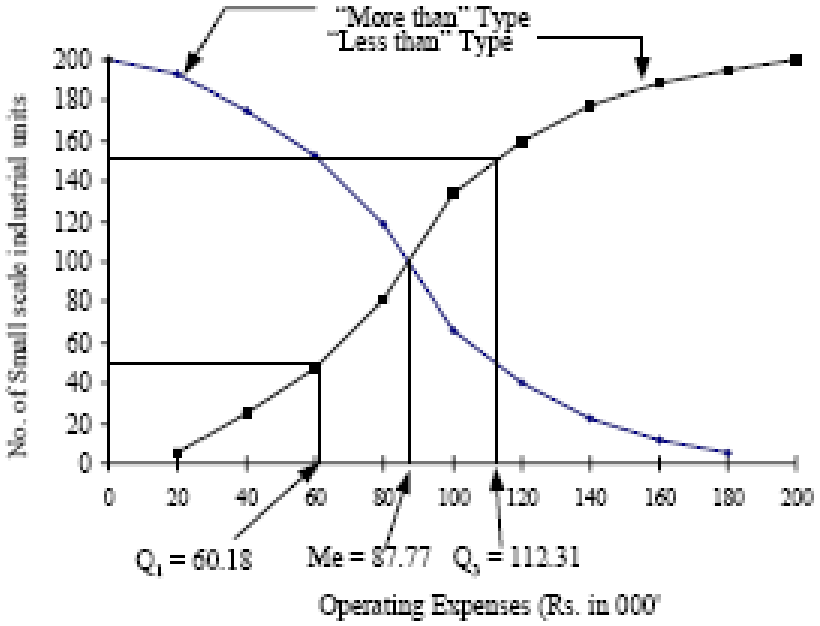


Example 5: The following data relates to the monthly operating expenses incurred by a sample of 200 small-scale industrial units in a city.

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

Operating Exp. (Thousand Rs.)	No. of Units (Frequency)	Less than Classes	Cumulative Frequency	More than Classes	Cumulative Frequency
0-20	7	Less than 20	7	More than 0	200
20-40	18	Less than 40	25	More than 20	193
40-60	22	Less than 60	47	More than 40	175
60-80	34	Less than 80	81	More than 60	153
80-100	53	Less than 100	134	More than 80	119
100-120	26	Less than 120	160	More than 100	66
120-140	18	Less than 140	178	More than 120	40
140-160	10	Less than 160	188	More than 140	22
160-180	7	Less than 180	195	More than 160	12
180-200	5	Less than 200	200	More than 180	5



(B) Graphs of Time Series: Most of the series relating to economics and business data are time series, e.g., population, bank deposits, price of articles, profits etc. Thus there are two variables; one of them is independent (i.e. time) and the other dependent (the phenomenon under study). The time units are placed on X-axis and the sizes of variables are measured on the Y-axis. If the actual values are plotted, the resulting graph would be called Absolute Histogram. When instead of actual values of series, their relative values in terms of percentages (index numbers) are plotted; the graph is called Index Histogram.

In making comparison, two or more variables may be represented on the same graph. Different types of lines may be used to avoid confusion. Various types of time series graph are Horizontal Line Graph or Historigram, Range Graph or Variation Graph, Net Balance Graph or Silhouette, and Band Graph.

Horizontal Line Graph or Historigram: In these graphs only one variable is plotted on the graph. The time variable is plotted on X-axis and value of the variables on the Y-axis. For the purpose of constructing graph firstly we mark a dot or point at the appropriate place, and then join all consecutive dots to give straight line. In case fluctuations are quite minute, compared to the size of variable, there is no need of showing the entire vertical scale from zero. In such a case the scale just sufficient for the purpose is shown and for unrepresented values a false base line is used. This omits portion of the scale which lies between zero and the smallest value of the variable.

Example:

Range Graph: This graph is also known as Variation graphs. In this graph the deviation or difference between the two extreme values are shown. For constructing Range graph we plot the maximum and minimum values of the variable and join them by straight lines to get range lines. This is particularly useful for high lighting the difference in, say for example, in price of commodity on different period of time, temperature on different days etc.

Example: The following are the share price quotations of a firm for five consecutive weeks. Present the data by an appropriate diagram:

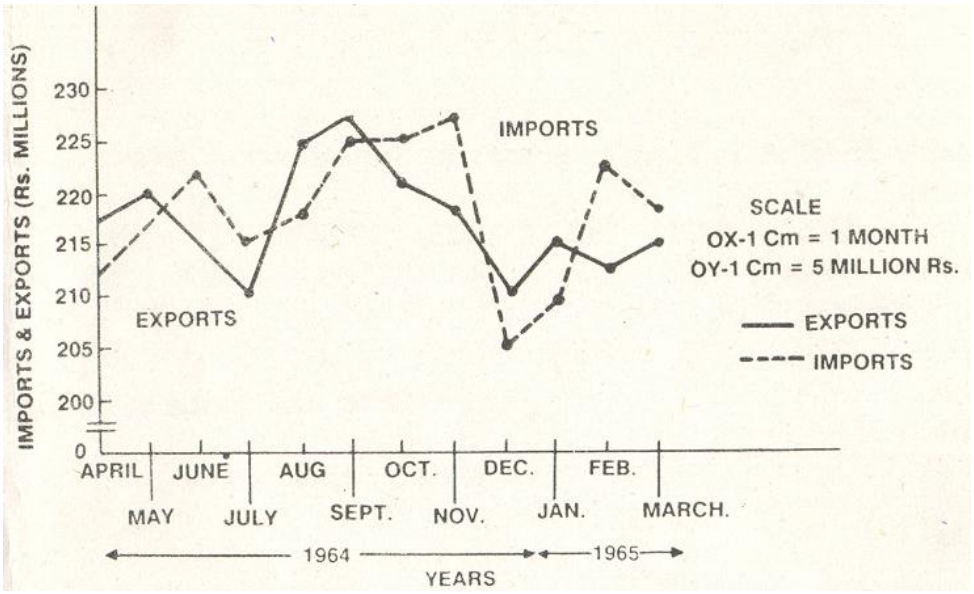
Net Balance Graph: This graph is used particularly to depict the difference or net balance between the two variables. For constructing these graphs we plot the maximum

and minimum values of the variable and then shade the area between two curves to show the net balance.

Example 6: Following were the import and export data of India during the financial year 1964-65. Prepare a graph showing imports, exports and the trade balance

Period	Apr, 1964	May,1964	Jun, 1964	Jul, 1964	Aug,1964	Sep, 1964
Imports (Million Rs.)	213	210	222	215	218	225
Exports (Million Rs.)	217	220	215	210	225	227
Period	Oct, 1964	Nov, 1964	Dec, 1964	Jan, 1965	Feb, 1965	Mar, 1965
Imports (Million Rs.)	225	218	210	215	212	215
Exports (Million Rs.)	220	218	210	215	212	215

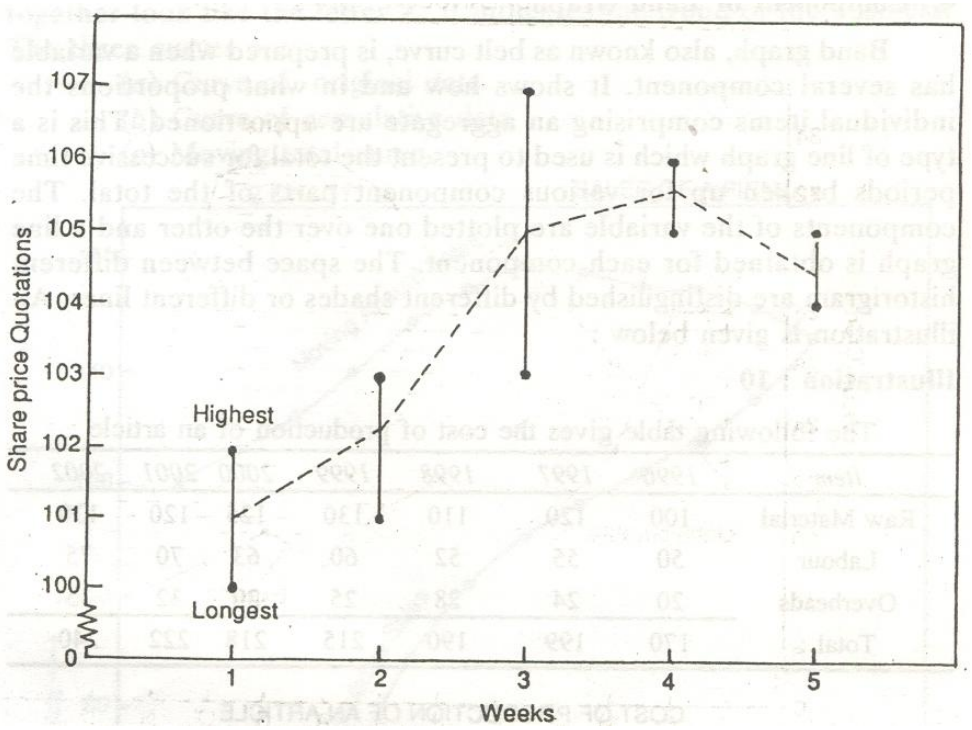
Solution:



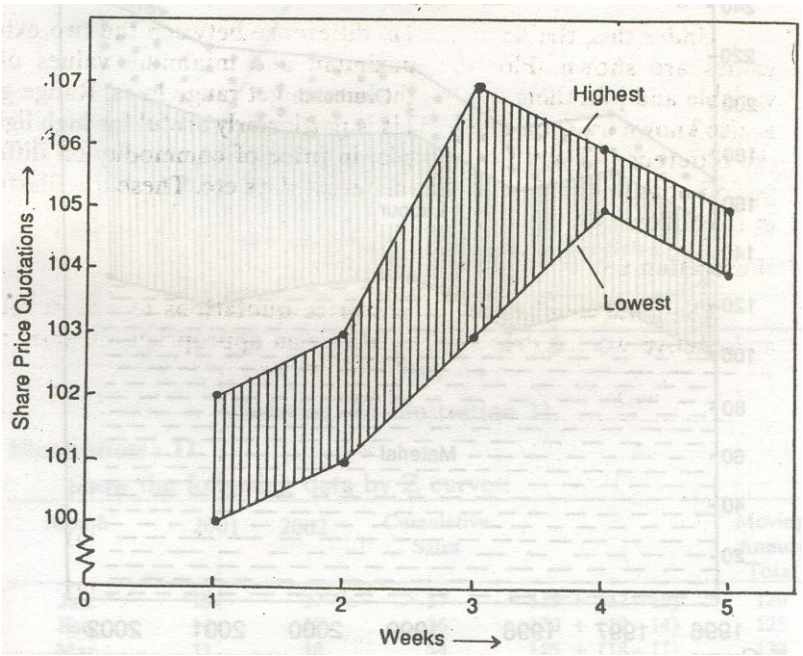
Example 7: Following are the stock quotations (high and low) of a company in last 5 weeks. Present the data by an appropriate graph.

Week	1	2	3	4	5
High	102	103	107	106	105
Low	100	101	103	105	104

Solution:



This graph can also be depicted as:



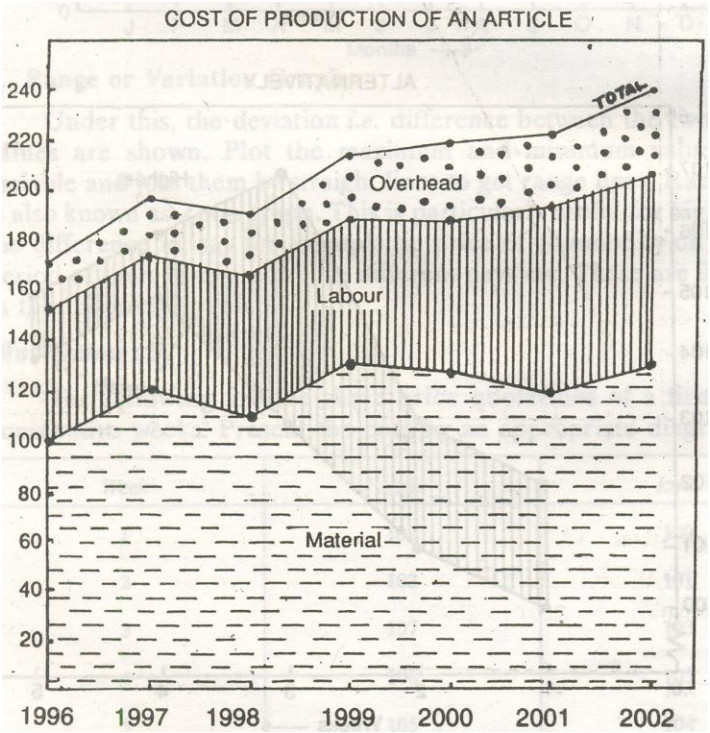
Band Graph: When a variable has several components Band graph is prepared. It shows how and in what proportions the individual items comprising an aggregate are apportioned. This type of line graph is used to present the total for successive time periods broken up for various components of the total. The components of the variable

are plotted one over the other and a line graph is obtained for each component. The space between different historigram is distinguished by different shades or different lines.

Example 8: The following table gives the cost of production of an article. Prepare a graph to show cost of various components and the total cost.

Item	1996	1997	1998	1999	2000	2001	2002
Raw Material	100	120	110	130	125	120	130
Labor	50	55	52	60	63	70	75
Overheads	20	24	28	25	30	32	35
Total	170	199	190	215	218	222	240

Solution:



SUMMARY

Statistical data not only requires a careful analysis but also ensures an attractive and communicative display. In order to achieve this objective, we have, in this lesson, discussed the techniques of graphic presentation of statistical data. Such graphic 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,

and Standard deviation etc. Hence, with the help of the graphs the researcher can effectively communicate to readers the information contained in a large mass of numerical data. In graphs, we discussed 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.

REVIEW QUESTIONS

1. Give a brief description of the different kinds of graphs generally used in business research to present the data.
2. What is the significance of graphic presentation? What precautions must a researcher take while plotting data on graph paper?
3. Explain with suitable examples: (i) Histogram, (i) Frequency Polygon, (iii) Ogives
4. Represent the following data relating to India's exports of Handicraft products to USA during 2004-09 by a suitable graph

Year	2004-05	2005-06	2006-07	2007-08	2008-09
Export (in US $)	320	450	385	310	425

5. Represent the following data by Histogram

Marks	No. of Students	Marks	No. of students
0-10	8	50-60	60
10-20	12	60-70	52
20-30	22	70-80	40
30-40	35	80-90	30
40-50	40	90-100	5

6. Present the following data through Histogram and calculate value of mode.

Class Interval	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	8	18	23	37	47	26	16	5

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

Daily Wage	100-120	120-140	140-160	160-180	180-200	200-220	220-240
No. of Workers	26	52	87	93	34	26	12

8. The following data relating to sales of 80 companies are given below. 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.

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

SUGGESTED READINGS

- Elhance DN: Fundamentals Of Statistics
- Gupta SP: Statistical Methods
- Gupta BN: Statistics
- Nagar KN: Fundamentals of Statistics
- Varshney RD: Fundamentals of Statistics
- Nagar AL: Fundamentals of Statistics