General rules for constructing tables. For the construction of tables in general, the following rules prove quite useful:

1. Title of the table should be simple, concise and unambiguous. As a rule, it should appear on the top of the table.
2. The table should be suitably divided into columns and rows according to the nature of data and purpose. These columns and rows should be arranged in a logical order to facilitate comparisons.
3. The heading of each column or row should be as brief as possible. Two or more columns or rows with similar headings may be grouped under a common heading to avoid repetition and we may have sub-headings or captions.

4. Sub-totals for each separate classification and a general total for all combined classes are to be given. These totals should be given at the bottom or at the right of the concerned items (members).

5. The units (e.g. weight in kg, height in centimeters, percentage, etc.) in which the data are given must invariably be mentioned, preferably, in the headings of columns or rows.

6. Necessary footnotes for providing essential explanation of the points to ambiguous interpretation of the tabulated data must be given at the bottom of the table.

7. The source/sources from where the data have been received should be given at the end of the table. In case the primary source is not traceable, then the secondary source may be mentioned.

8. In tabulating long columns of figures, space should be left after every five or ten rows.

9. If the numbers tabulated have more than three significant figures, the digits should be grouped in threes. For example, we should write 5732981 as 5 732 981.

10. For all purposes and by all means, the table should be as simple as possible so that it may be studied by the readers with minimum possible strain and it should provide a clear picture and interpretation of the statistical data in the quickest possible time.

MEANING OF GRAPHICAL REPRESENTATION OF DATA

A graphic representation is the geometrical image of a set of data. It is a mathematical picture. It enables us to think about a statistical problem in visual terms. A picture is said to be more effective than words for describing a particular thing or phenomenon. Consequently, the graphic representation of data proves quite an effective and an economic device for the presentation, understanding and interpretation of the collected statistical data.

ADVANTAGES OF GRAPHICAL REPRESENTATION OF DATA

The advantages of graphic representation may be summarized as below:

1. The data can be presented in a more attractive and an appealing form.
2. It provides a more lasting effect on the brain. It is possible to have an immediate and a meaningful grasp of large amounts of data through such presentation.

3. Comparative analysis and interpretation may be effectively and easily made.
4. Various valuable statistics like median, mode, quartiles, may be easily computed. Through such representation, we also get an indication of correlation between two variables.
5. Such representation may help in the proper estimation, evaluation and interpretation of the characteristics of items and individuals.
6. The real value of graphical representation lies in its economy and effectiveness. It carries a lot of communication power.
7. Graphical representation helps in forecasting, as it indicates the trend of the data in the past.

MODES OF GRAPHICAL REPRESENTATION OF DATA

We know that the data in the form of raw scores is known as ungrouped data and when it is organized into a frequency distribution, then it is referred to as grouped data. Separate methods are used to represent these two types of data—ungrouped and grouped. Let us discuss them under separate heads.

Graphical Representation of Ungrouped Data

For the ungrouped data (data not grouped into a frequency distribution) we usually make use of the following graphical representation:

1. Bar graph or bar diagrams
2. Circle graph or Pie diagrams
3. Pictograms
4. Line graphs

Bar graph or bar diagram. In bar graphs or diagrams the data is represented by bars. Generally these diagrams or pictures are drawn on graph paper. Therefore these bar diagrams are also referred to as bar graphs.

These diagrams or graphs are usually available in two forms, vertical and horizontal. In the construction of both these forms, the lengths of the bars are in proportion to the amount of variables or traits (height, intelligence, number of individuals, cost, and so on) possessed. The width of bars is not governed by any set rules. It is an arbitrary factor. Regarding the space between two bars, it is conventional to have a space about one half of the width of a bar.

The data capable of representation through bar diagrams may be in the form of raw scores, total scores or frequencies, computed statistics and summarized figures like percentages and averages.

Graphical Representation of Grouped Data (Frequency Distribution)

There are four methods of representing a frequency distribution graphically:

1. The histogram
2. The frequency polygon
3. The cumulative frequency graph
4. The cumulative frequency percentage curve or ogive

Let us discuss these methods one by one.

Histogram. A histogram or column diagram is essentially a bar graph of a frequency distribution. The following points are to be kept in mind while constructing the histogram for a frequency distribution. (For illustration purposes, we can use the frequency distribution given in Table 2.1.)

1. The scores in the form of actual class limits as 19.5–24.5, 24.5–29.5 and so on are taken as examples in the construction of a histogram rather than the written class limits as 20–24, 25–30 and so on.
2. It is customary to take two extra intervals (classes) one below and the other above the given grouped intervals or classes (with zero frequency). In the case of frequency distribution

given in Table 2.1, we can take 14.5–19.5 and 69.5–74.5 as the two required class intervals.

3. Now we take the actual lower limits of all the class intervals (including the extra intervals) and try to plot them on the x-axis. The lower limit of the lowest interval (one of the extra intervals) is taken at the intersecting point of x-axis and y-axis.
4. Frequencies of distribution are plotted on the y-axis.
5. Each class or interval with its specific frequency is represented by a separate rectangle. The base of each rectangle is the width of the class interval (i) and the height is the respective frequency of that class or interval.
6. It is not essential to project the sides of the rectangles down to the base line.
7. Care should be taken to select the appropriate units of representation along the x-axis and the y-axis. Both the x-axis and the y-axis should not be too short or too long. A good general rule for this purpose as suggested by Garrett (1971) is:

   To select x and y units which will make the height of the figure approximately 75% of its width.

The above procedure may be properly understood through Figure 3.6 which shows the histogram of the frequency distribution given in Table 2.1.

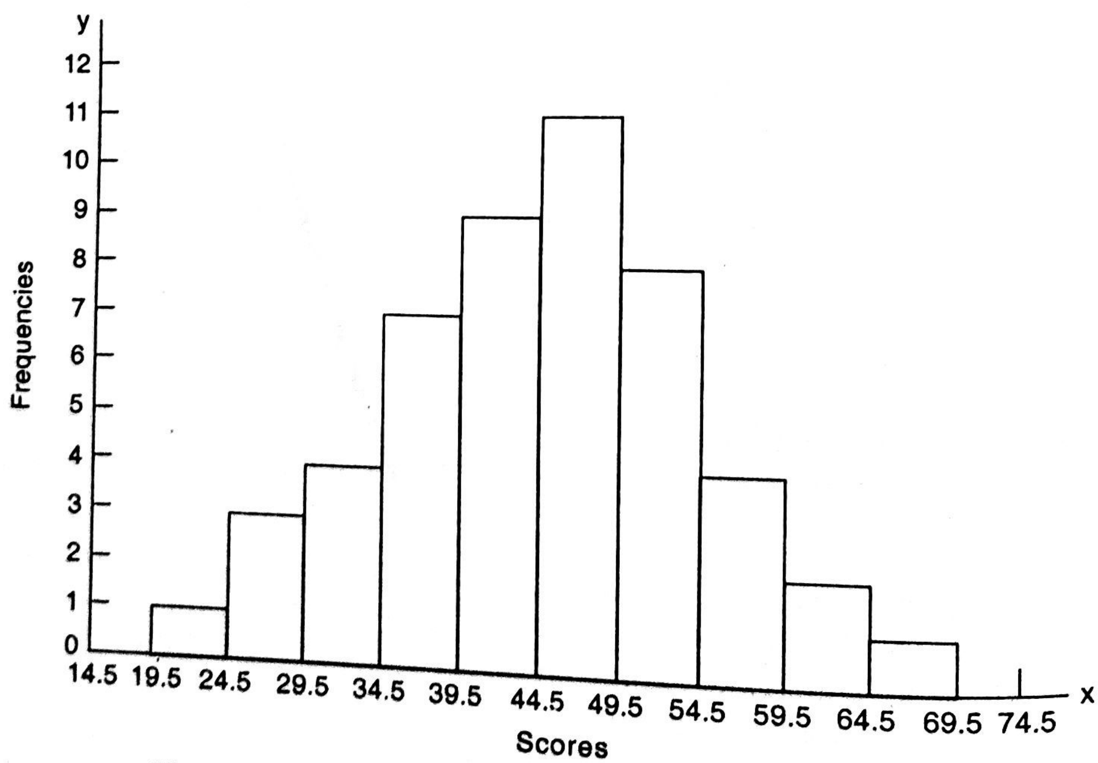

Figure 3.6 Histogram of frequency distribution.

Frequency polygon. A frequency polygon (Figure 3.7) is essentially a line graph for the graphical representation of the frequency distribution. We can get a frequency polygon from a histogram, if the midpoints of the upper bases of the rectangles are connected by straight lines. But it is not essential to plot a histogram first to draw a frequency polygon. We can construct it directly from a given frequency distribution. The following points are helpful in constructing a frequency polygon:

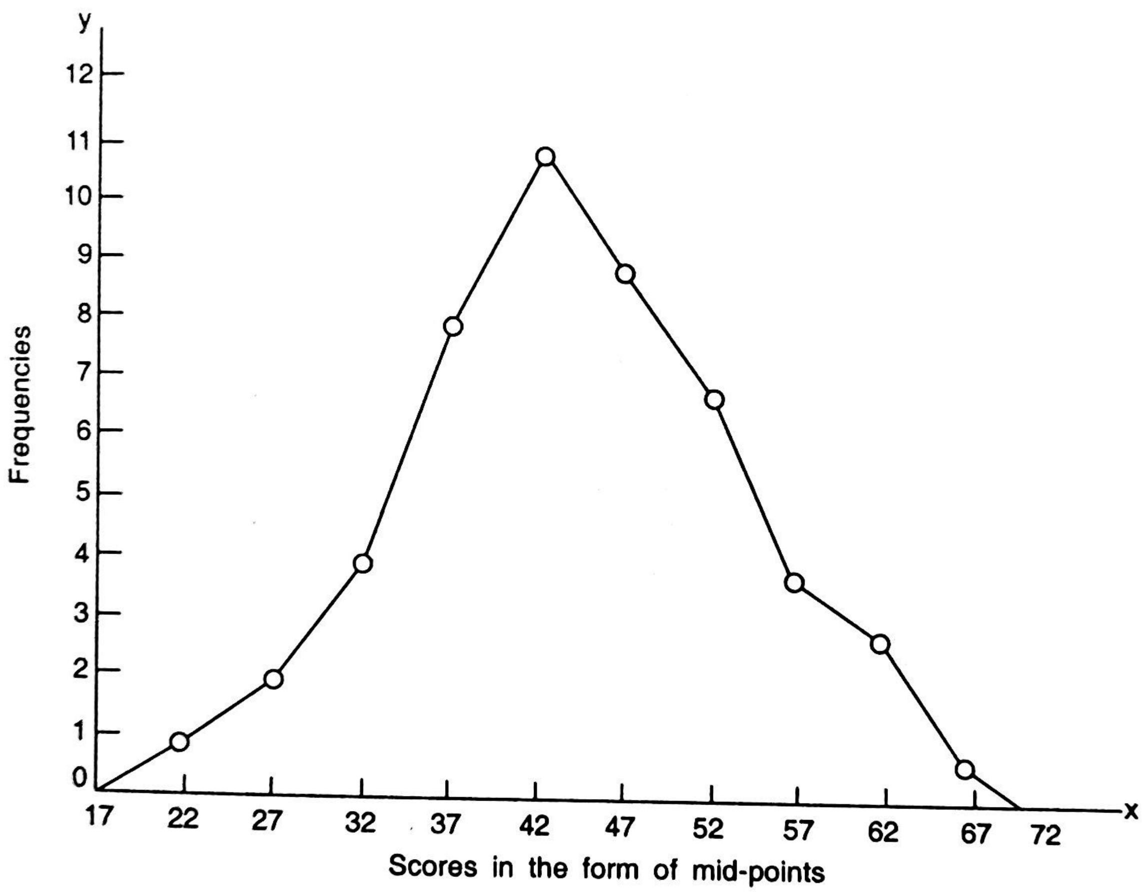

Figure 3.7 Frequency polygon of the frequency distribution given in Table 2.1.

1. As in the histogram, two extra intervals or classes, one above and the other below the given intervals are taken.
2. The mid-points of all the classes or intervals (including two extra intervals) are calculated.
3. The mid-points are marked along the *x*-axis and the corresponding frequencies are plotted along the *y*-axis by choosing suitable scales on both axes.
4. The various points obtained by plotting the mid-points and frequencies are joined by straight lines to give the frequency polygon.
5. For approximate height of the figure and selection of *x* and *y* units, the rule emphasized earlier in the case of histogram should be followed.

Difference between a frequency polygon and a frequency curve. A polygon is a many sided figure. It is essentially a closed curve while a frequency curve is not a closed curve. In a frequency curve we do not take two extra intervals or classes. But in a frequency polygon, we take these two extra classes in order to close the figure.

Comparison between the histogram and the frequency polygon. Although both histogram and frequency polygon are used for the graphic representation of frequency distribution and are alike in many respects, they possess points of difference. Some of these differences are cited below:

1. Where histogram is essentially the bar graph of the given frequency distribution, the frequency polygon is a line graph of this distribution.
2. In frequency polygon, we assume the frequencies to be concentrated at the mid-points of the class intervals. It points out merely the graphical relationship between mid-points and frequencies and thus is unable to show the distribution of frequencies within each class interval. But the histogram gives a very clear as well as accurate picture of the relative proportions of frequency from interval to interval. A mere glimpse of the figure answers such questions as:
   (a) Which group of class intervals has the largest or smallest frequency?
   (b) Which pair of groups or class intervals has the same frequency?
   (c) Which group has its frequency double that of another?
3. In comparing two or more distributions by plotting two or more graphs on the same axis, frequency polygon is more useful and practicable than the histogram.
4. In comparison to the histogram, frequency polygon gives a much better conception of the contours of the distribution. With a part of the polygon curve, it is easy to know the trend of the distribution but a histogram is unable to tell such a thing.