Graphical Representation of Frequency Distribution

An attractive representation of a frequency distribution is graphical representation.

Graphical representation can be used for both the educated section and uneducated section of the society. Furthermore, any hidden trend present in the given data can be noticed only in this mode of representation.

We are going to consider the following types of graphical representation:

1. Histogram

A two dimensional graphical representation of a continuous frequency distribution is called a histogram.

In histogram, the bars are placed continuously side by side with no gap between adjacent bars.

That is, in histogram rectangles are erected on the class intervals of the distribution. The areas of rectangle are proportional to the frequencies.

Properties of the histogram:

(i) Frequencies are along the vertical axis and the scores (C.I.) are along the horizontal axis.

(ii) One assumes that the scores are evenly distributed within the class interval, thus giving us rectangular bars.

(iii) The frequencies within each interval of a histogram are represented by a rectangle, the size of the interval being the base and the frequency of that interval the height.

(iv) The area of each rectangle in a histogram corresponds to the frequency within a given interval, while the total area of a histogram corresponds to the total frequency (N) of the distribution.

(v) A histogram may be best constructed on a graph paper, which is ruled with equally spaced horizontal and vertical lines.

Advantages:

1. It is simple and easily made.
2. All the advantages of the graphic representation as shown earlier are applicable here.

Limitations:

1. It is difficult to superimpose more than one histogram on the same graph.

2. Comparisons of several frequency distributions cannot readily be made via histograms. Frequency polygons are much better suited for that purpose.

3. The assumption that the scores are evenly distributed within the C.I. produces a larger error when N is small than when N is large.

4. It cannot be smoothed.

Example 1:

Draw a histogram for the following table which represent the marks obtained by 100 students in an examination:

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Number of students	5	10	15	20	25	12	8	5

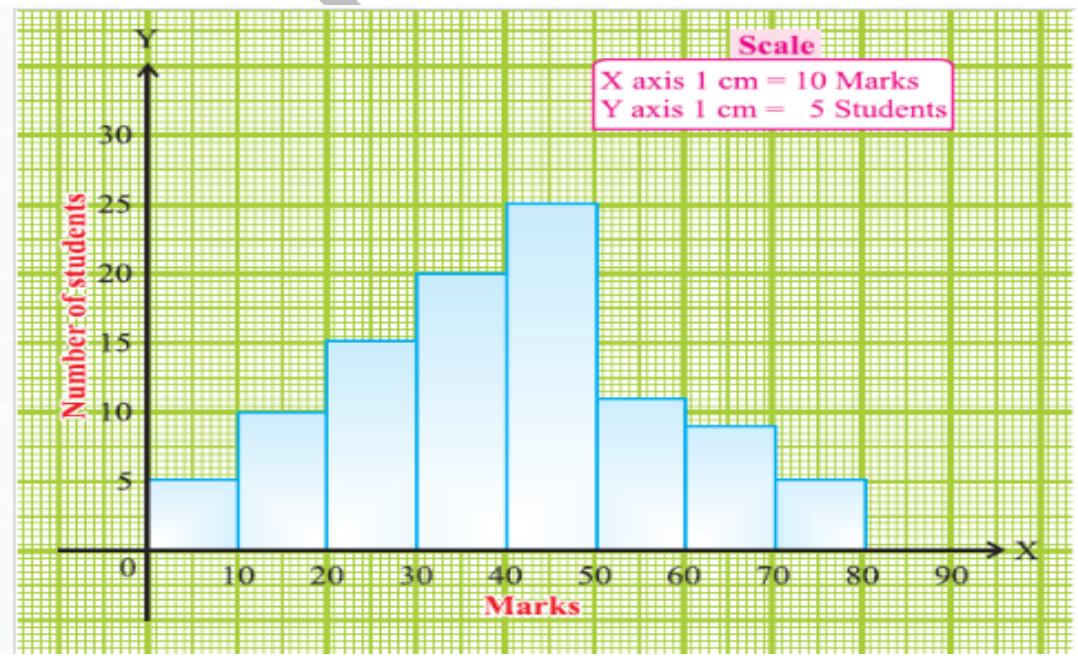
Solution:

The class intervals are all equal with length of 10 marks.

Let us denote these class intervals along the X-axis.

Denote the number of students along the Y-axis, with appropriate scale.

The histogram is given below.



2. Bar diagram

There are two types of bar diagrams namely, Horizontal Bar diagram and Vertical bar diagram. While horizontal bar diagram is used for qualitative data or data varying over space, the vertical bar diagram is associated with quantitative data or time series data.

Bars i.e. rectangles of equal width and usually of varying lengths are drawn either horizontally or vertically.

We consider Multiple or Grouped Bar diagrams to compare related series. Component or sub-divided Bar diagrams are applied for representing data divided into a number of components. Finally, we use Divided Bar charts or Percentage

Bar diagrams for comparing different components of a variable and also the relating of the components to the whole. For this situation, we may also use Pie chart or Pie diagram or circle diagram.

Advantages

- Show relationships between 2 or more variables
- Good visual impression of trends and changes
- Can show positive and negative values
- Simple to construct and read

Disadvantages

- Plotting too many bars makes it appear cluttered - less easy to interpret
- If there is a wide range of data it is difficult to read accurately
- Becomes more complicated if there are uneven class intervals
- Using too many or too few classes can mask important patterns in the data

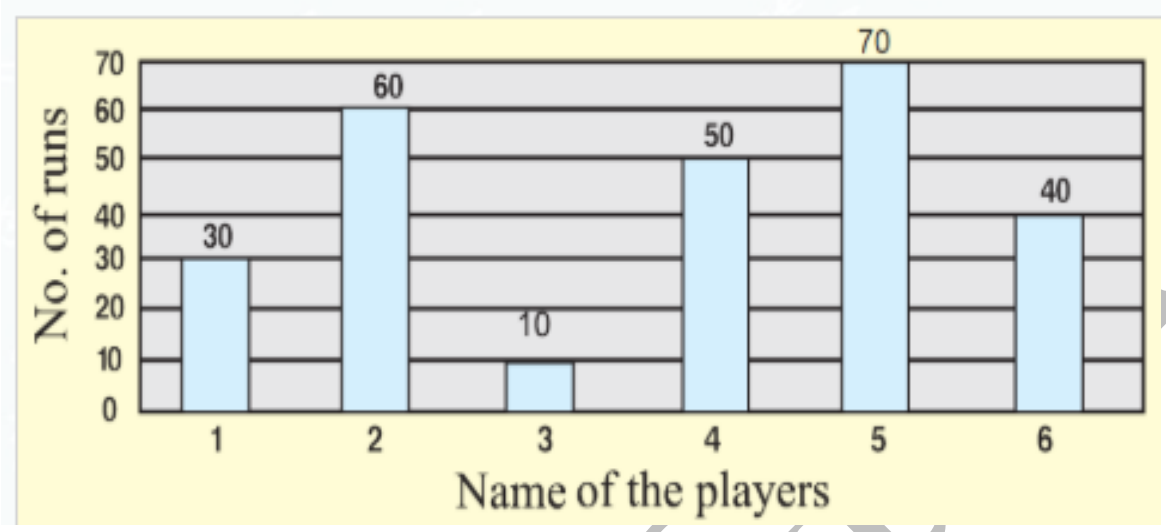
Example:

The total number of runs scored by a few players in one-day match is given.

Players	1	2	3	4	5	6
No. of runs	30	60	10	50	70	40

Solution:

Draw bar graph for the above data.



3. Pie chart

In a pie chart, the various observations or components are represented by the sectors of a circle and the whole circle represents the sum of the value of all the components .Clearly, the total angle of 360° at the center of the circle is divided according to the values of the components.

The central angle of a component is

= [Value of the component / Total value] x 360°

Sometimes, the values of the components are expressed in percentages. In such cases,

The central angle of a component is

= [Percentage value of the component / 100] x 360°

Advantages of a Pie Chart

- A simple and easy-to-understand picture.
- It represents data visually as a fractional part of a whole, which can be an effective communication tool for the even uninformed audience.
- It enables the audience to see a data comparison at a glance to make an immediate analysis or to understand information quickly.

- The need for readers to examine or measure underlying numbers themselves can be removed by using this chart.
- To emphasize points you want to make, you can manipulate pieces of data in the pie chart.

Disadvantages of a Pie Chart

- If too many pieces of data are used, pie chart becomes less effective.
- They themselves may become crowded and hard to read if there are too many pieces of data, and even if you add data labels and numbers may not help here.
- You need a series to compare multiple sets as this chart only represents one data set.
- To analyze and assimilate information quickly, this may make it more difficult for readers
- As the reader has to factor in angles and compare non-adjacent slices, it has its problems in comparing the data slices.
- To make decisions based on visual impact rather than data analysis leads readers to draw inaccurate conclusions
- Negative Pie / positive Pie cannot be understood until I hover the pointer on the pie. So when Negative data present, pie chart is a bad option

Example:

The number of hours spent by a school student on various activities on a working day, is given below. Construct a pie chart using the angle measurement.

Activity	Sleep	School	Play	Homework	Others
Number of hours	8	6	3	3	4

Draw a pie chart to represent the above information.

Solution:

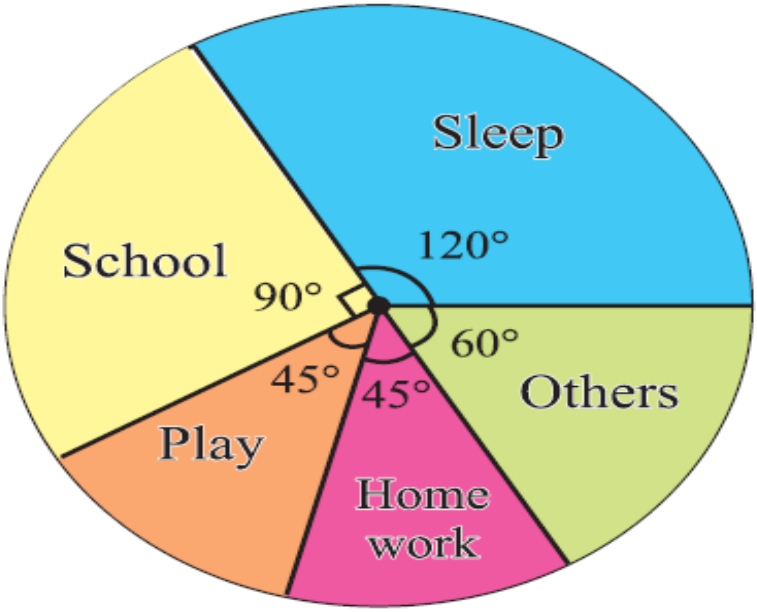
The central angle of a component is

= [Value of the component / Total value] x 360°

We may calculate the central angles for various components as follows:

Activity	Duration in hours	Central angle
Sleep	8	$\frac{8}{24} \times 360^0 = 120^0$
School	6	$\frac{6}{24} \times 360^0 = 90^0$
Play	3	$\frac{3}{24} \times 360^0 = 45^0$
Homework	3	$\frac{3}{24} \times 360^0 = 45^0$
Others	4	$\frac{4}{24} \times 360^0 = 60^0$
Total	24	360^0

From the above table, clearly, we obtain the required pie chart as shown below.



4. Line diagram

When the time series exhibit a wide range of fluctuations, we may think of logarithmic or ratio chart where "Log y" and not "y" is plotted against "t".

We use multiple line charts for representing two or more related time series data expressed in the same unit and multiple – axis chart in somewhat similar situations, if the variables are expressed in different units.

Advantages

- Show relationships between 2 or more variables
- Good visual impression of trends and changes
- Can show positive and negative values
- Simple to construct and read

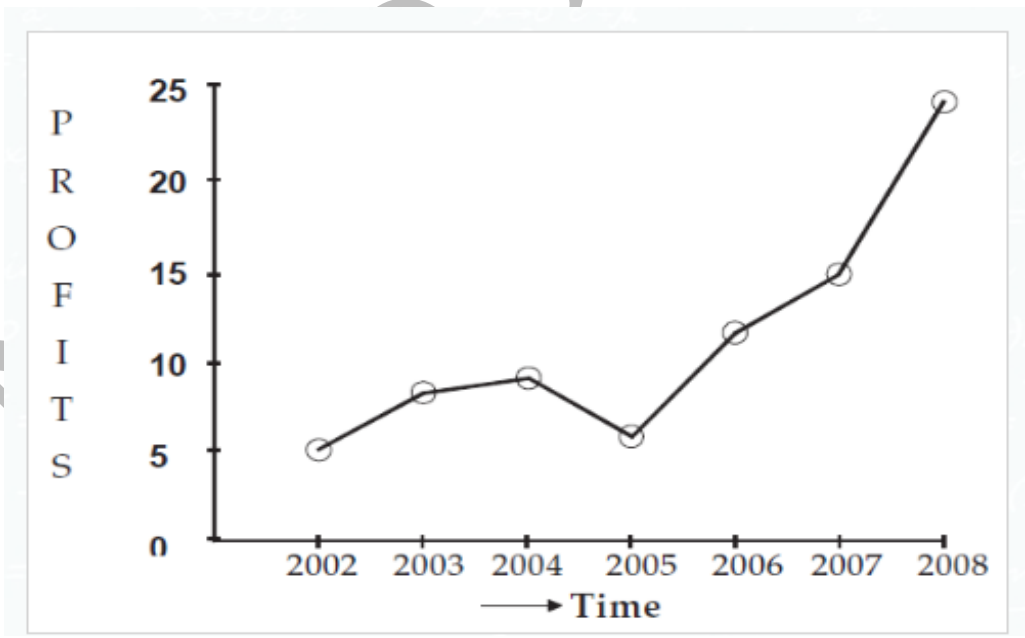
Disadvantages

- Plotting too many bars makes it appear cluttered - less easy to interpret
- If there is a wide range of data it is difficult to read accurately
- Becomes more complicated if there are uneven class intervals
- Using too many or too few classes can mask important patterns in the data

Example:

The profits in thousands of dollars of an industrial house for 2002, 2003, 2004, 2005, 2006, 2007 and 2008 are 5, 8, 9, 6, 12, 15 and 24 respectively. Represent these data using a suitable diagram.

Solution:



5. Frequency polygon

Frequency Polygon is another method of representing frequency distribution graphically.

Obtain the frequency distribution and compute the mid points of each class interval.

Represent the mid points along the X-axis and the frequencies along the Y-axis.

Plot the points corresponding to the frequency at each mid point.

Join these points, by straight lines in order.

To complete the polygon join the point at each end immediately to the lower or higher class marks (as the case may be at zero frequency) on the X-axis.

Example:

Draw a frequency polygon for the following data without using histogram.

Class interval	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
Frequency	4	6	8	10	12	14	7	5

Solution:

Mark the class intervals along the X-axis and the frequency along the Y-axis.

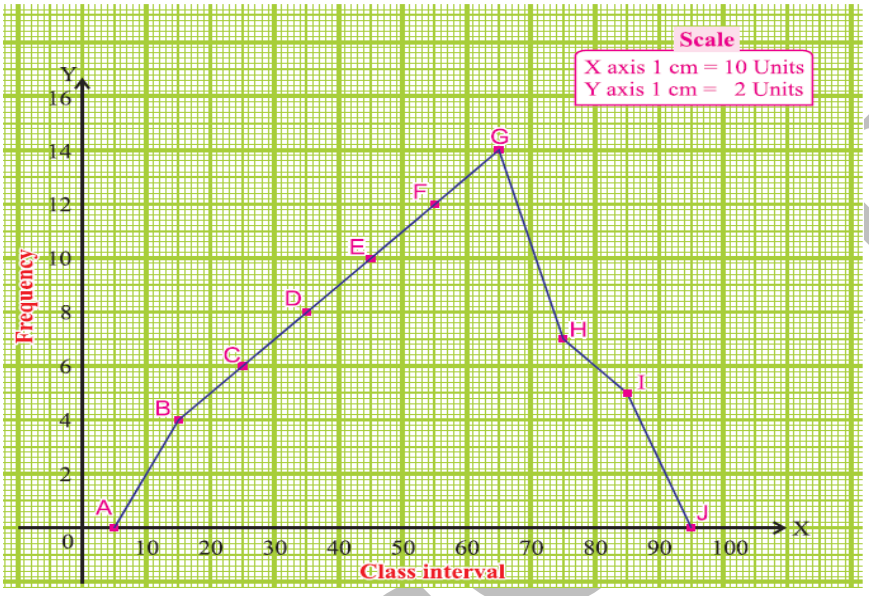
We take the imagined classes 0-10 at the beginning and 90-100 at the end, each with frequency zero.

We have tabulated which is given below.

Class interval	Midpoints	Frequency
0-10	5	0
10-20	15	4
20-30	25	6
30-40	35	8
40-50	45	10
50-60	55	12
60-70	65	14
70-80	75	7
80-90	85	5
90-100	95	0

Using the adjacent table, plot the points A (5, 0), B (15, 4), C (25, 6), D (35, 8), E (45, 10), F (55, 12), G (65, 14), H (75, 7), I (85, 5) and J (95, 0).

We draw the line segments AB, BC, CD, DE, EF, FG, GH, HI, IJ to obtain the required frequency polygon ABCDEFGHIJ, which is given below.



6. Ogives or Cumulative frequency graphs

By plotting cumulative frequency against the respective class boundary, we get ogives.

As such there are two ogives – less than type ogives, obtained by taking less than cumulative frequency on the vertical axis and more than type ogives by plotting more than type cumulative frequency on the vertical axis and thereafter joining the plotted points successively by line segments.

Advantage:

- Cumulative frequency is the useful way to determine the number of scores that occur until a certain value
- Original info from a grouped frequency distribution can be obtained from the CF curves
- It will show you whether it is a constant rate or if the speed up or slows down

Disadvantage:

- Difficult to compare the frequencies between each data group
- It requires high bandwidth

- It is also difficult to compare different data sets

Example:

Draw ogives for the following table which represents the frequency distribution of weights of 36 students.

Weights in kg (Class interval)	No. of students (Frequency)
43.50 - 48.50	3
48.50 - 53.50	4
53.50 - 58.50	5
58.50 - 63.50	7
63.50 - 68.50	9
68.50 - 73.50	8

Solution:

To draw ogives for the above frequency distribution, we have to write less than and more than cumulative frequency as given below.

Weight in kg (CB)	Cumulative Frequency	
	Less than	More than
43.50	0	33 + 3 or 36
48.50	0 + 3 or 3	29 + 4 or 33
53.50	3 + 4 or 7	24 + 5 or 29
58.50	7 + 5 or 12	17 + 7 or 24
63.50	12 + 7 or 19	8 + 9 or 17
68.50	19 + 9 or 28	0 + 8 or 8
73.50	28 + 8 or 36	0

Now, we have to write the points from less than and more than cumulative frequency as given below.

Points from less than cumulative frequency:

(43.50, 0), (48.50, 3), (53.50, 7), (58.50, 12), (63.50, 19), (68.50, 28) and (73.50, 36)

Points from more cumulative frequency:

(43.50, 36 (48.50, 33), (53.50, 29), (58.50, 24), (63.50, 17), (68.50, 8) and (73.50, 0)

Now, taking frequency on the horizontal axis, weights on vertical axis and plotting the above points, we get ogives as given below.