Bar Graph and Histogram

Part A: Bar Graph

Definition 1: *Bar Graph:* A bar chart or a bar graph is a graph that uses vertical or horizontal bars to represent the frequencies of the categories in a data set.

The length of each bar is proportional to the amount that it represents.

There are 3 main types of bar charts.

- **Vertical bar chart**
- Horizontal bar chart
- Double bar chart

The **double bar chart** is used when we want to represent two sets of data on the same chart. We can put the bars side by side or we may put the bars of one set of data on top of the bars of the other set of data.

A bar graph is a valuable presentation tool, since it is effective at reinforcing differences in magnitude. Bar charts permit the visual comparison of data by displaying the magnitude of each category as a horizontal or vertical bar.

Bar charts are useful when the data set has categories (for example, hair color, gender, etc...) and data values that are qualitative in nature.

A bar chart is made up of columns plotted on a graph. Here is how to read a bar chart.

- The columns are positioned over a label that represents a categorical variable.
- The height of the column indicates the size of the group defined by the column label.
- The bars are equally separated and are equal in width.

Example 1: A sample of 300 college students was asked to indicate their favorite soft drink. The survey results are shown below. Display the information using a bar chart.

Soft Drink	Number of Students
Pepsi-Cola	92
Coca-Cola	78
Dr. Pepper	48
7-Up	42
Others	40

A horizontal bar chart for the same soft drink information is shown below.
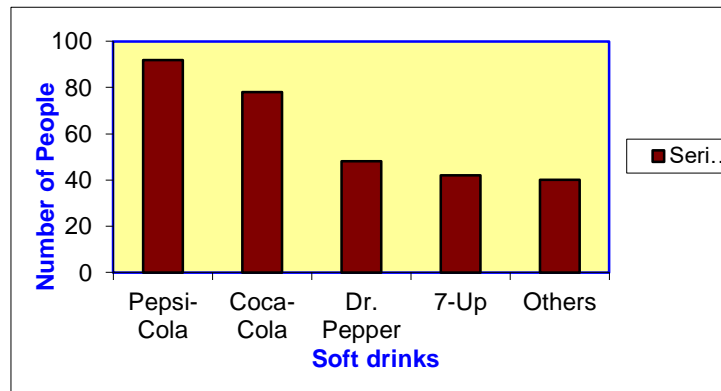
A vertical bar chart for the same soft drink information is shown below.
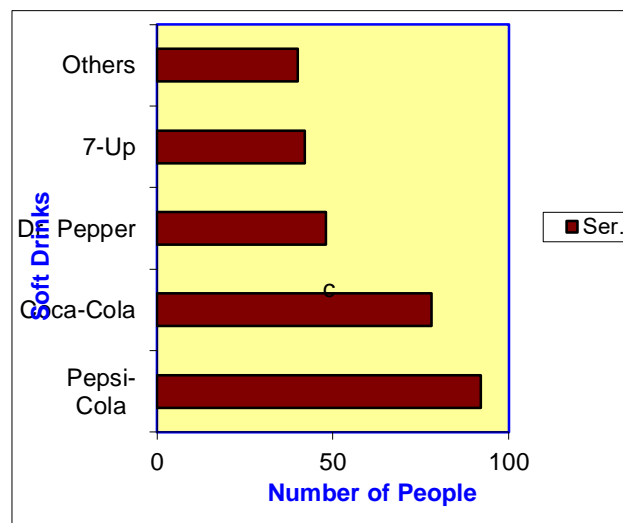
Remark 1: Remember that all graphs should contain the following:

- Labels for the vertical axes
- A title
- Evenly divided horizontal and vertical axes
- Accurate data

Advantages - Bar graphs can:

- show each nominal or ordinal category in a frequency distribution
- display relative numbers or proportions of multiple categories
- summarize a large data set in visual form
- clarify trends better than do tables or arrays
- estimate key values at a glance
- permit a visual check of the accuracy and reasonableness of calculations
- be easily understood due to widespread use in business and media

Disadvantages - Bar graphs can:

- require additional written or verbal explanation
- be easily manipulated to yield false impressions
- be inadequate to describe the attribute, behavior, or condition of interest
- fail to reveal key assumptions, norms, causes, effects, or patterns

Part B: Histograms

Histogram

A histogram is a graphical display of a frequency or a relative frequency distribution that uses classes and vertical bars (rectangles) of various heights to represent the frequencies.

The width of the bars represents the intervals, and the height of the bars represents the frequency.

Histograms are useful when the data values are quantitative. A histogram gives an estimate of the shape of the distribution of the population from which the sample was taken.

Like a bar chart, a histogram is made up of columns plotted on a graph. Here is how to read a histogram.

- The columns are positioned over a label that represents a quantitative variable.
- The height of the column indicates the size of the group defined by the column label.

Display the data with a histogram using seven classes.

A histogram with seven classes for the data is shown below

The histogram shows the frequency count for each class, with each class having a width of 10.

<u>COLUMNS SHOULD BE NEXT TO ONE ANOTHER</u>
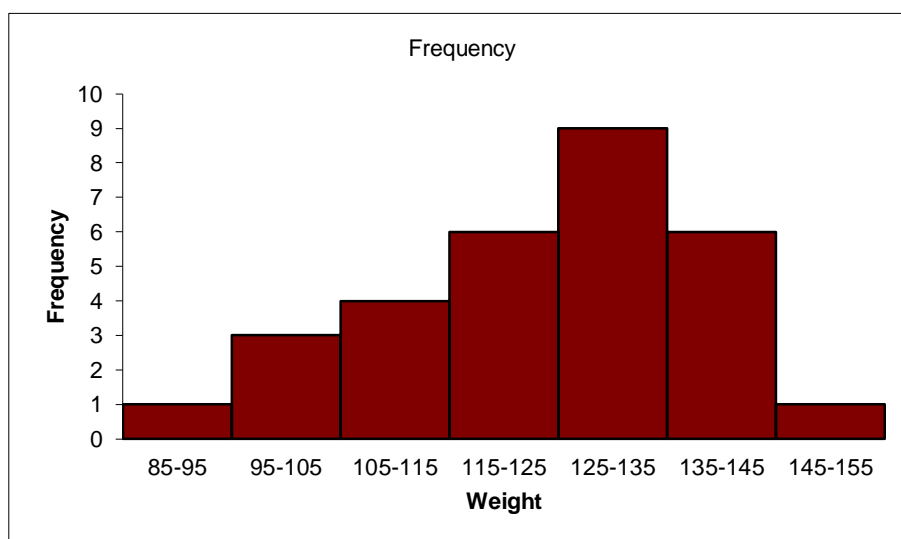
Observe the histogram; there is a frequency count of 1 for the interval 85-95, a frequency count of 3 for the interval 95-105, etc. From this information, we can construct the grouped frequency distribution.

If the relative frequencies are plotted along the vertical axis to produce a relative frequency histogram, the shape of the resulting histogram will be the same as that of a histogram in which the frequencies were plotted along the vertical axis. This is true because the relative frequencies are obtained by dividing the frequency values by the total number of values in the data set.

The Difference between Bar Charts and Histograms

Here is the main difference between bar charts and histograms. With bar charts, each column represents a group defined by a categorical variable; and with histograms, each column represents a group defined by a quantitative variable.

One implication of this distinction: it is always appropriate to talk about the skewness of a histogram; that is, the tendency of the observations to fall more on the low end or the high end of the x-axis.

In a histogram, however, the bars must touch. This is because the data elements we are recording are numbers that are grouped, and form a continuous range from left to right. There are no gaps in the numbers along the bottom axis. This is what makes a histogram.

With bar charts, however, the x-axis does not have a low end or a high end; because the labels on the x-axis are categorical - not quantitative. As a result, it is less appropriate to comment on the skewness of a bar chart.

Advantages - Histograms can:

- begin to show the central tendency and dispersion of a data set
- closely resemble the bell curve if sufficient data and classes are used
- show each interval in the frequency distribution
- summarize a large data set in visual form
- clarify trends better than do tables or arrays
- estimate key values at a glance
- permit a visual check of the accuracy and reasonableness of calculations
- be easily understood due to widespread use in business and the media
- use bars whose areas reflect the proportion of data points in each class

Disadvantages - Histograms can:

- require additional written or verbal explanation
- be easily manipulated to yield false impressions
- be inadequate to describe the attribute, behavior, or condition of interest
- fail to reveal key assumptions, norms, causes, effects, or patterns