# LECTURE 15 DIAGRAMS AND GRAPHS

DR. GAURAV SANKALP

## RULES FOR CONSTRUCTING DIAGRAMS

1. The diagrams should be simple.
2. Each diagram must be given a clear, concise and suitable title without damaging clarity.
3. A proper proportion between height and width must be maintained in order to avoid an unpleasant look.
4. Select a proper scale; it should be in even numbers or in multiples of five or ten. e.g. 25,50, 75 or 10, 20, 30, 40, .... etc. But there are no fixed rules.
5. In order to clear certain points, always put footnotes.
6. An index, explaining different lines, shades and colors should be given.
7. Diagrams should be absolutely neat and clean.

## ADVANTAGES AND DISADVANTAGES OF DIAGRAMS

### ADVANTAGES

1. Quick way for the audience to visualize what you are saying -- numbers, trends, up or down
2. Forceful -- emphasizes main point
3. Convincing -- proves a point, see and hear
4. Compact way to convey information
5. More interesting than just talk or print (Remember to use as many of the five senses as possible)

### DISADVANTAGES

1. Time consuming to make -- decisions must be made in advance for layout, color, materials, etc.
2. Technical in nature -- audience knowledge to interpret, or understand
3. Costly -- depending on the medium used (poster board, transfer letters, etc.)

## PICTOGRAM

A **pictogram**, also called a **pictogramme** or **pictograph**, is an ideogram that conveys its meaning through its pictorial resemblance to a physical object. Pictographs are often used in writing and graphic systems in which the characters are to a considerable extent pictorial in appearance.

## HISTOGRAM

A **histogram** is a graphical representation of the distribution of data. It is an estimate of the probability distribution of a continuous variable and was first introduced by Karl Pearson. A histogram is a representation of tabulated frequencies, shown as adjacent rectangles, erected over discrete intervals (bins), with an area equal to the frequency of the observations in the interval.

The height of a rectangle is also equal to the frequency density of the interval, i.e., the frequency divided by the width of the interval. The total area of the histogram is equal to the number of data. A histogram may also be normalized displaying relative frequencies. It then shows the proportion of cases that fall into each of several categories, with the total area equaling 1. The categories are usually specified as consecutive, non-overlapping intervals of a variable. The categories (intervals) must be adjacent, and often are chosen to be of the same size. The rectangles of a histogram are drawn so that they touch each other to indicate that the original variable is continuous.

Histograms are used to plot the density of data, and often for density estimation: estimating the probability density function of the underlying variable. The total area of a histogram used for probability density is always normalized to 1. If the lengths of the intervals on the x-axis are all 1, then a histogram is identical to a relative frequency plot. An alternative to the histogram is kernel density estimation, which uses a kernel to smooth samples. This will construct a smooth probability density function, which will in general more accurately reflect the underlying variable.

## GRAPHS VS DIAGRAMS

Sometimes, the information purported to be understood is too long and complex. To make it interesting and understandable in an exciting manner, different visual representations are used.

Graphs and diagrams are two of the common means to visually represent information that is either repetitive in nature or too complex. There are similarities in these techniques that confuse many to treat them as similar.

**DIAGRAMS**

We are too well aware of the use of diagrams to explain information and facts that are presented in the form of text. If you need to explain the parts of a machine or the principle of its working, it becomes difficult to make one understand the concept through text only. This is where diagrams in the form of sketches come into play. Similarly, diagrams are made heavy use of in biology where students have to learn about different body parts and their functions. Visual representation of concepts through diagrams has better chances of retention in the memory of students than presenting them in the form of text.

Diagrams are resorted to right from the time a kid enters a school as even alphabets are presented to him in a more interesting and attractive manner with the help of diagrams.

**GRAPHS**

Whenever there are two variables in a set of information, it is better to present the information using graphs as it makes it easier to understand the data. For example, if one is trying to show how the prices of commodities have increased with respect to time, a simple line graph would be a more effective and interesting way rather than putting all this information in the form of text which is hard to remember whereas even a layman can see how prices have gone up or down in relation to time.

Graphs make use of graph paper which has precise squares and presents the information in an accurate manner and the reader can see the effect of one variable on another in a very simple manner.

**DIFFERENCE BETWEEN GRAPHS AND DIAGRAMS**

1. All graphs are a diagram but not all diagrams are graph. This means that diagram is only a subset of graph.
2. Graph is a representation of information using lines on two or three axes such as x, y, and z, whereas diagram is a simple pictorial representation of what a thing looks like or how it works.

3. Graphs are representations to a scale whereas diagrams need not be to a scale
4. Diagrams are more attractive to look at which is why they are used in publicity whereas graphs are for the use of statisticians and researchers.
5. Values of mean and median can be calculated through graphs which is not possible with diagrams
6. Graphs are drawn on graph paper whereas diagrams do not need a graph paper
7. For frequency distribution, only graphs are used and it cannot be represented through diagrams

**TYPES OF GRAPHS:**

One goal of statistics is to present data in a meaningful way. It's one thing to see a list of data on a page; it's another to understand the trends and details of the data. Many times data sets involve millions (if not billions) of data values. This is far too many to print out in a journal article or sidebar of a magazine story. One effective tool in the statistician's toolbox is to depict data by the use of a graph.

They say a picture is worth a thousand words. The same thing could be said about a graph. Good graphs convey information quickly and easily to the user. Graphs highlight salient features of the data. They can show relationships that are not obvious from studying a list of numbers. Graphs can also provide a convenient way to compare different sets of data.

**LIST OF COMMON GRAPHS IN STATISTICS**

Different situations call for different types of graphs, and it helps to have a good knowledge of what graphs are available. Many times the type of data determines what graph is appropriate to use. Qualitative data, quantitative data and paired data each use different types of graphs.

**Seven of the most common graphs in statistics are listed below:**

1. **Pareto Diagram or Bar Graph** - A bar graph contains a bar for each category of a set of qualitative data. The bars are arranged in order of frequency, so that more important categories are emphasized.
2. **Pie Chart or Circle Graph** - A pie chart displays qualitative data in the form of a pie. Each slice of pie represents a different category.

3. **Histogram -** A histogram in another kind of graph that uses bars in its display. This type of graph is used with quantitative data. Ranges of values, called classes, are listed at the bottom, and the classes with greater frequencies have taller bars.
4. **Stem and Left Plot** - A stem and left plot breaks each value of a quantitative data set into two pieces, a stem, typically for the highest place value, and a leaf for the other place values. It provides a way to list all data values in a compact form.
5. **Dot Plot -** A dot plot is a hybrid between a histogram and a stem and leaf plot. Each quantitative data value becomes a dot or point that is placed above the appropriate class values.
6. **Scatter Plots** - A scatter plot displays data that is paired by using a horizontal axis (the *x* axis), and a vertical axis (the *y* axis). The statistical tools of correlation and regression are then used to show trends on the scatterplot.
7. **Time-Series Graphs** - A time-series graph displays data at different points in time, so it is another kind of graph to be used for certain kinds of paired data. The horizontal axis shows the time and the vertical axis is for the data values. These kinds of graphs can be used to show trends as time progresses.

**TIME SERIES GRAPHS**

Suppose that we want to study the climate of a region for an entire month. Every day at noon we note the temperature and write this down in a log. A variety of statistical studies could be done with this data. We could find the mean or the median temperature for the month. We could construct a histogram displaying the number of days that temperatures reach a certain range of values. But all of these methods ignore a portion of the data that we have collected.

The feature of the data that we may want to consider is that of time. Since each date is paired with the temperature reading for the day, we don‘t have to think of the data as being random. We can instead use the times given to impose a chronological order on the data. A graph that recognizes this ordering and displays the changing temperature as the month progresses is called a time series graph.

**CONSTRUCTING A TIME SERIES GRAPH**

To construct a time series graph, we must look at both pieces of our paired data set. We start with a standard Cartesian coordinate system. The horizontal axis is used to plot the date or time increments, and the vertical axis is used to plot the values variable that we are

measuring. By doing this each point on the graph corresponds to a date and a measured quantity. The points on the graph are typically connected by straight lines in the order in which they occur.

**USES OF A TIME SERIES GRAPH**

Time series graphs are important tools in various applications of statistics. When recording values of the same variable over an extended period of time, sometimes it is difficult to discern any trend or pattern. However, once the same data points are displayed graphically, some features jump out. Time series graphs make trends easy to spot. These trends are important as they can be used to project into the future.

In addition to trends, the weather, business models and even insect populations exhibit cyclical patterns. The variable being studied does not exhibit a continual increase or decrease, but instead goes up and down depending upon the time of year. This cycle of increase and decrease may go on indefinitely. These cyclical patterns are also easy to see with a time series graph.