**MODULE # 9 VISUALIZATION IN R**

To understand graphics in R, initially dataset is selected to understand it.

The dataset 'measles' is downloaded from the following link:[https://vincentarelbundock.github.io/Rdatasets/datasets.html](https://www.blogger.com/blog/post/edit/2482057455959808426/5701729624945142864)

The dataset 'measles' is imported in R by clicking on import dataset option below the environment tab in Rstudio followed by appearance of dialog box to choose data set from downloads.

To produce graphics in R, there are three types:

1. Base graphics

2. Lattice

3. ggplot2

***1. Base graphics:***

Base graphics already exists in the R installation. It allows to create a wide range of two dimensional plots, including scatterplots, line graphs, bar graphs, and more.

> plot(measles,pch=21,col=rainbow(12),main= " Deaths in London from measles", cex=1.0)

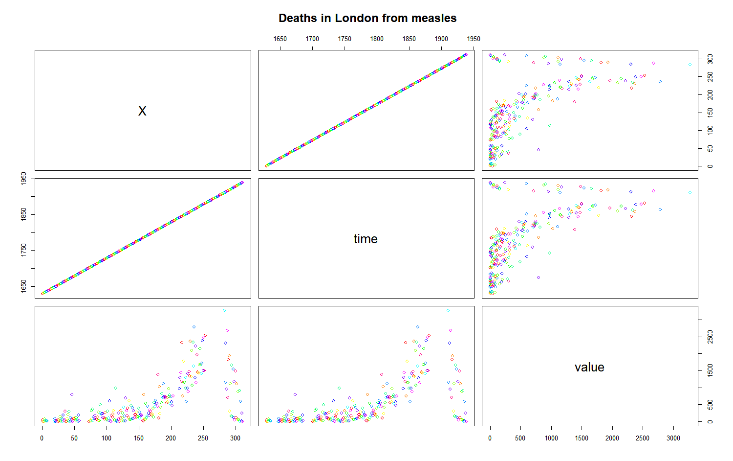
> plot(measles,pch=21,bg="green",col=cm.colors(12),main= " Deaths in London from measles", cex=1.0)

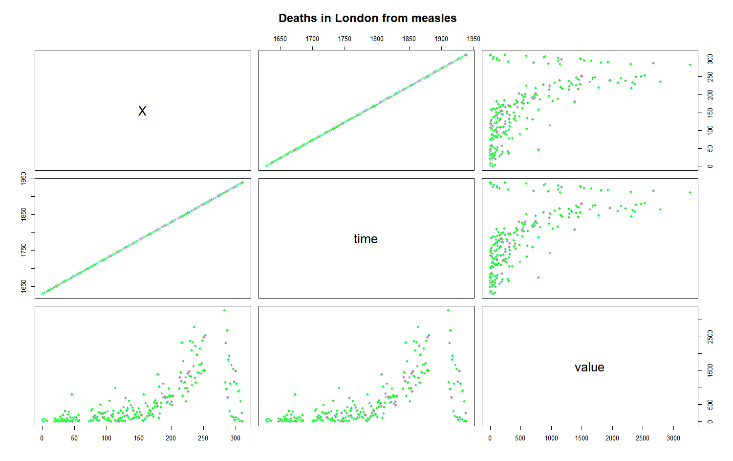
'plot' function is implemented  which creates scatterplot of a dataset. 'pch=21' is used to set the type of plotting to a filled circle with a solid fill. There are different color palletes which can be used, but for this code 'rainbow()' and 'cm.colors()' is utilized. The title of the plot can be added to the graph using 'main' function. 'bg' defines fill color of the circle. 'cex' controls the size of the points in the plot (character expansion). Finally, this code produces a scatter plot of the number of deaths in london over time with each point representing a particular time period.

*Disadvantages:*

Awkward outflow-User must deal with laying out the individual plots

No built in support for encoding additional information

[](https://blogger.googleusercontent.com/img/a/AVvXsEi_rmK62BGaEuPPkPp8FoQrqtOaZEcAX5PHQW1uDb23YYBUD47-tm3KixhR_xjRJPUvBbPw6r03DuoP8p0p-Jjagj3RkTaAdnOl3_1U_cUl2tFMnlBIwrwZ_EdJQeNt97sX1k786j0zBIlEH3mW0KWs4CZrdPQCziTV_6BsP8L0tvF1T0XBHIeLu7q9)

[](https://blogger.googleusercontent.com/img/a/AVvXsEjs_x1FcjsnqVH4gvHSUKzcq_Az2qkGo3RVqhtJXoQxsCq9PLUidYirnb4GikzA3G7EJK4ZoAjK6fUCa0bvMN5NikA2A5cPE-1YMmXOk5wb4E6fXzy9vWA7cNucAxZDC5NWHxAyPvkWJ8E735EwM2e6RB--0dTd8BO5RJP3U7Zsdf53qqrLTQ-gt9GX)

2. ***Lattice:***

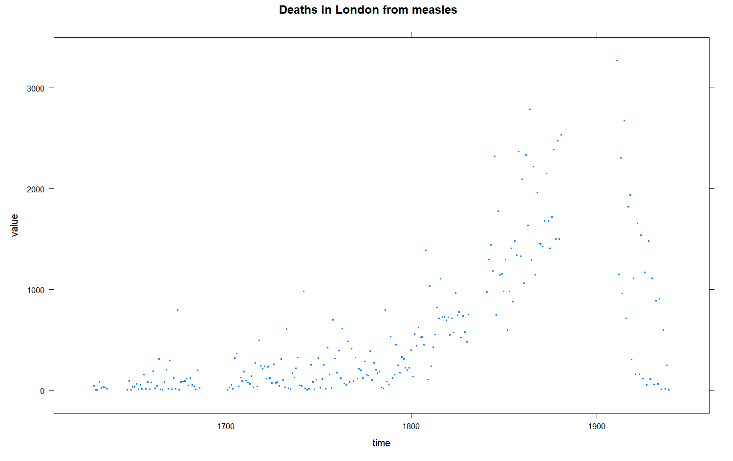
It provides functions to create high level graphics like trellis plots. It provides an additional layer of abstraction when compared to base graphics. In addition, they are very customisable.

> install.packages("lattice")

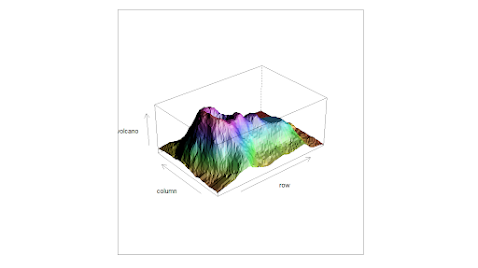
> library(lattice)

> xyplot(value~time, data=measles,main= " Deaths in London from measles", pch=".",cex=2.5)

Initially, lattice package is installed to create these graphics in R followed by loading the lattice package using library(). 'xyplot' function is used to plot the relationship between the time and the values i.e number of death due to measles in London. 'pch' function sets the type of plot.  'cex' controls the size of the points in the plot (character expansion).

[](https://blogger.googleusercontent.com/img/a/AVvXsEiBy_a6RFCjKc9DGh-B_nyynugPI85ElShcw0CN8sSIY0oneCeie2FlFl_0B0FoT8j0dC0eteSFpFjj8R6M6Aq5eTIHovvYpJqPZdYHF67webnxvXDSg1a0UMTKBB21eoaZnjOzKz-vji12QrTvEnhQg81NDYOoVFzo7Eab-ljM09Mdprcc3QHZahlI)

> wireframe(volcano, shade = TRUE, aspect = c(61/87, 0.4),light.source = c(10,0,10))

[](https://blogger.googleusercontent.com/img/a/AVvXsEigrIgyqwXuSCf_RtoqCllQHPJuve092lgC7zrQtr-LnsK1LPuEaqfMPc_uWYbsX1cRT86j34mez8JiZlfGKKRFZEeAuTeYFnrsb_zsJ78ZRMxOf84gMurC9LR2M-TUMc-8zCgCbydZIEmzwiV8w_M0wqvgWJc4kC2kqr9Rcd2p36Lj5Oj5yvRvFPHz)

'wireframe()' function creates 3D plot of dataset 'volcano'.'shade = TRUE' gives shade of the wideframe.'aspect' code sets the aspect ratio of the plot. 'light.source' gives the location of the light source for the plot.

*Advantages:*

They usually look better and can be extended in powerful ways. The resulting output can be annotated, edited, and saved.

3. ***ggplot2:***

Each component is added to the plot as a layer that never been done in statistical programming language. It creates complex and highly customizable visualizations.It takes two primary arguments, data and aes(). Aesthetic mapping to pass to be plotted.

> install.packages("ggplot2")

> library(ggplot2)

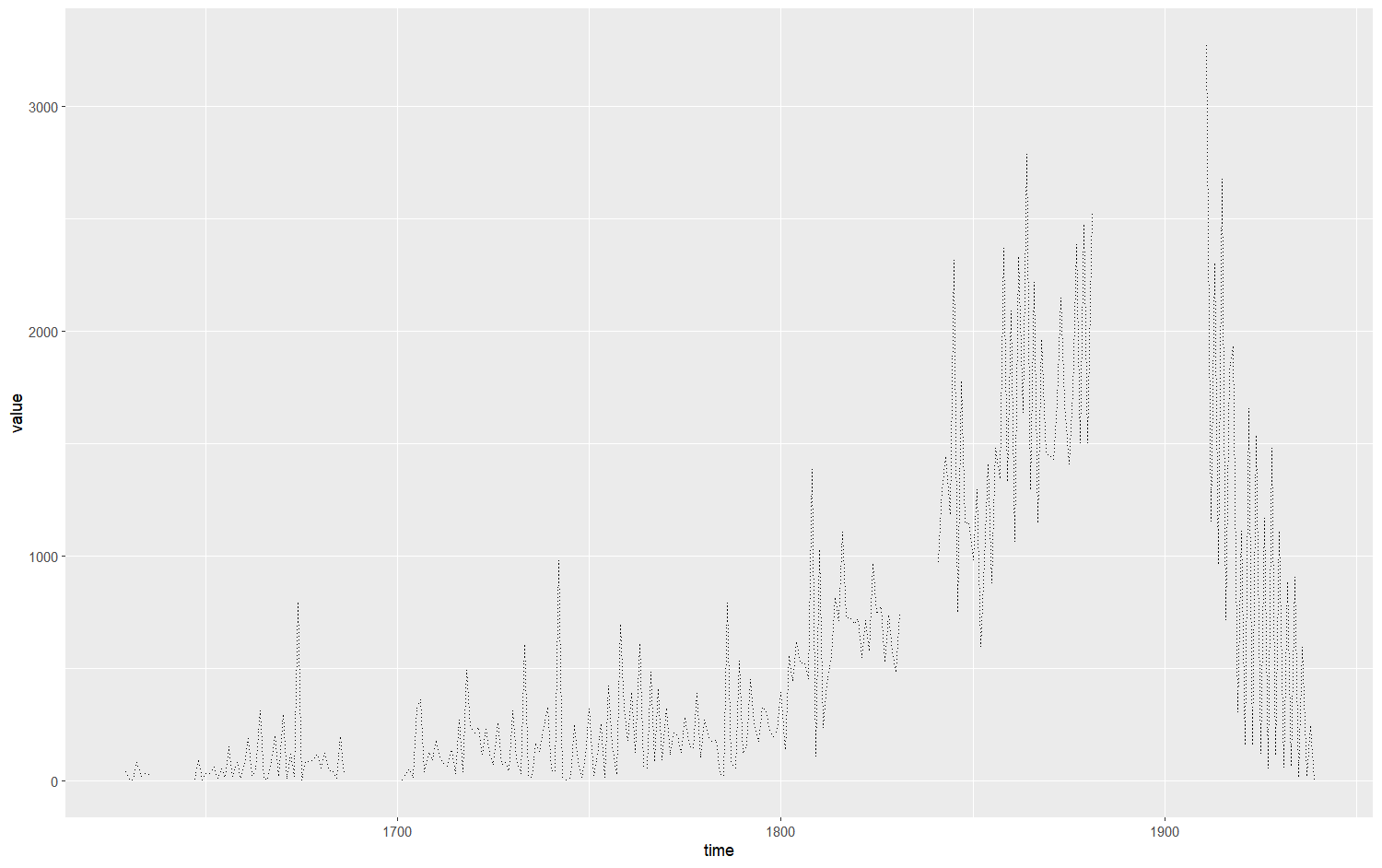
> ggplot(measles, aes(x=time,y=value))+geom\_line(linetype=3)

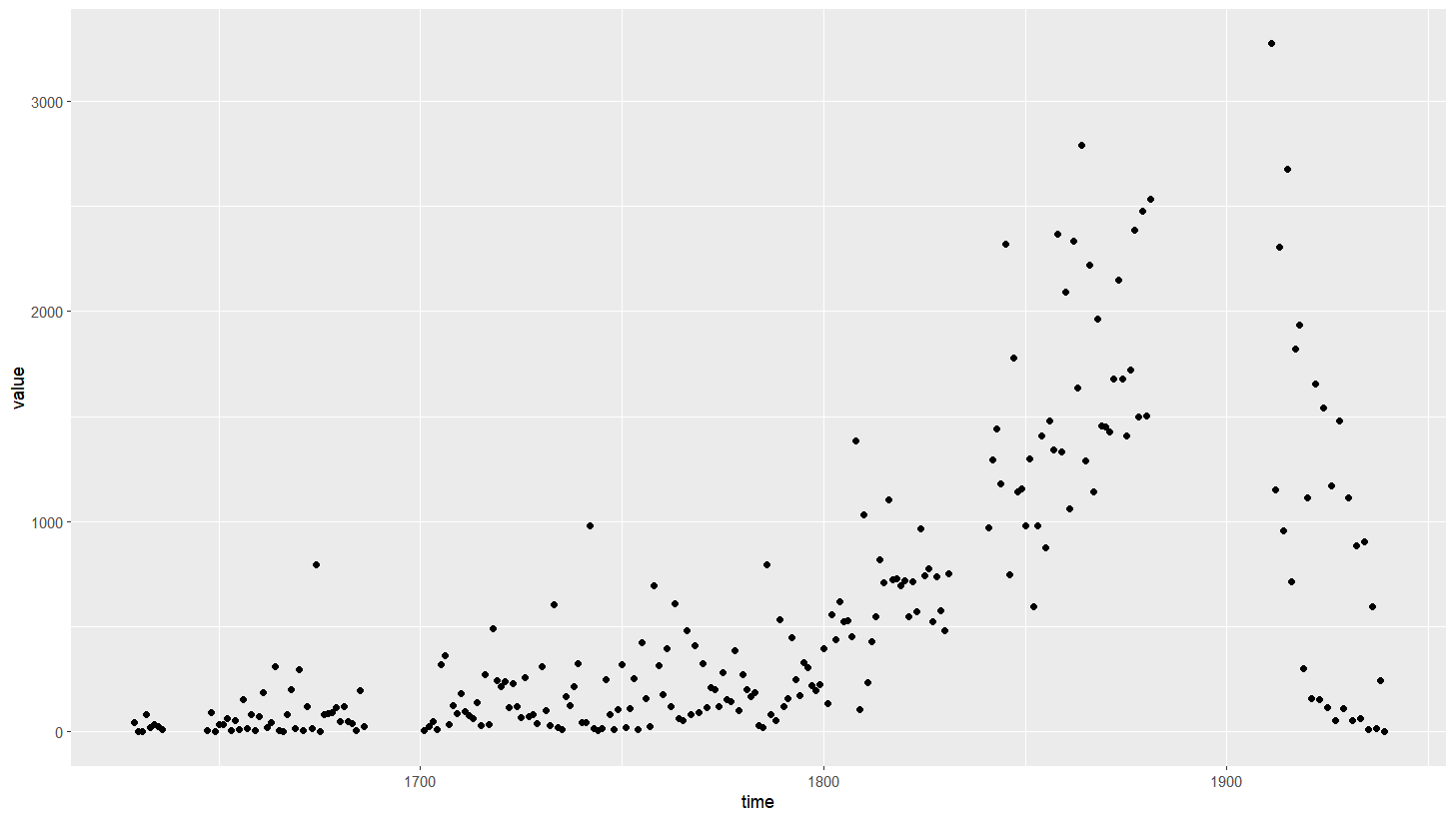
> ggplot(measles, aes(x=time,y=value))+geom\_point()

To create ggplot2 graphics in R, 'ggplot2' package is installed and loaded using 'library()' function.

'ggplot2' function is implemented to the 'measles' dataset.Plots convey information through various aspects of their aesthetics. 'x' is defined by 'time','y' is defined by 'value'. The sign '+' added to give a new layer with 'geom\_line()' which give line to plot. The code 'linetype=3' sets the line type to dashed line.

'geom\_point()' function adds points to the plot.

[](https://blogger.googleusercontent.com/img/a/AVvXsEjlLWsnhtrbeid9p94RpM3wryIT6r4joZP_qlR6EWnTSImIV69hOn7YbR_PubzT6ZvVI7bfrBBX-04XmrtUyqwHBJvFMVoWn8i7pewfXcDEcGuqD4H1qU3GTJMW5vTVpfe2xD1r2ajD5k6-wwFytL9GGtV7t3kb8DzrD8nSS0ltLdrTDvXgDII8wARQ)

[](https://blogger.googleusercontent.com/img/a/AVvXsEhQxsYiw3VEAND-MkitjsFlVjinzncXr6fHTNd7dqoAsybEBaCjBjZT-SL9BTtOEB7fjE75B6fVcS8wax7uNaYNnfkSAg1wc_9oYZtDe3xZ4p6VGqVZVbS5_VAXIIRfZUnSqnhcsQ5Z71CVWuhlqqUlnxQP530Fig59f_dY8rZ3EInTqVEwy4Zblwg3)

Through these graphics, the year in which more and less number of deaths due to measles in London can be determined and in addition, more clustering of the deaths in a particular year can also made out through these graphics. In this way, by importing datatsets and implementing codes for these graphics gives us information and broad perspective of how deaths are progressing and indirectly impact of disease on the population. These graphs are very helpful in pandemic too given the dataset.

**References**

Wickham, H. (2015). *R Packages.*

Matloff. N. (2011). *The Artof R programming*.