R-codes for the first 5 sessions

- Install R : http://www.r-project.org/
- Install R-studio : http://www.rstudio.com/
- Packages in R
  - Install using ***install.packages("package name"),*** select the CRAN Mirror closest to your location
  - Once installed we need to call the package in when needed using ***library("package name")***
- Import data
  - Text Files:
    - ***data<- read.table("C:/Users/xyz/Desktop/folderX/mydata.txt", header=TRUE)***
  - CSV Files:
    - ***data<- read.csv("C:/Users/xyz/Desktop/folderX/mydata.csv", header=TRUE)***
- Refer to a variable:
  - To refer to a particular variable, say 'v1' in the data, we use ***data$v1***
  - Or, attach the data using ***attach(data) ,***then call the variable using its name
- Define and Name a variable:
  - '<-' is used to assign names to variables
  - To create a new variable as some function of an existing variable use ***variable.new<-function(variable1)***
  - To rename variable1 as variable2 use ***variable2<-variable1***
- To save your workspace use 'ctrl+s' or 'save' from the drop down menu. This will restore your previous session.

Session 1 codes:

- Frequency table
  - ***table(variable)***
- Bar chart
  - ***barplot(table(variable))***
- Histogram
  - ***hist(variable)***
- Box-plot
  - ***boxplot(variable,horizontal=TRUE)***
- Summary stats
  - Mean: ***mean(variable)***
  - Variance: ***var(variable)***
  - Standard deviation: ***sd(variable)***
  - Skewness: ***skewness(variable)*** , needs package *'moments'*
  - Kurtosis: ***kurtosis(variable)***, needs package *'moments'*

Session 2 codes:

- For a variable following a Normal Distribution with **mean=mu** and **standard deviation=sigma** we have the following:
  - ***dnorm(data value, mu, sigma)*** : gives the density i.e. the function returns the height of the normal distribution, at some value along the x-axis
  - ***pnorm(data value, mu, sigma)*** : gives the distribution function
  - ***qnorm(quantile, mu, sigma)*** : gives the Quantile function for calculating critical values
  - ***rnorm(n,mu,sigma)***: generates 'n' samples from a Normal distribution with mean=mu and standard deviation=sigma
  - ***qqnorm(variable)*** : Without 'extRemes'' package will create a plot but without bands and with extRemes package will create a plot with bands.
  - ***qqPlot(variable,distribution="norm")*** :With 'car' package creates a normal probability plot with bands
  - ***qqline(variable,col="red")*** with or without extRemes package will draw a red line passing through the qqplot
  - To generate a new variable as a linear combination of two normal variables plot a qqplot, use the following commands:

    ```
    X1<-rnorm(500,15,3)
    X2<-rnorm(500,25,5)
    Y<-3*X1+4*X2
    library(car)
    qqPlot(Y)
    ```

- To understand these codes better, refer to
  - http://stat.ethz.ch/R-manual/R-devel/library/stats/html/Normal.html
  - http://ww2.coastal.edu/kingw/statistics/R-tutorials/prob.html

- To draw a Histogram with a Normal curve superposed on it:

  ```
  hist(variable,prob=TRUE)
  curve(dnorm(x,mean(variable),sd(variable)),col="red",add=TRUE)
  ```