Quantitative Management Modelling - Assignment 1 part-I

1.Read a file for data :

- **Data :**

Math test results by grade and ethencity and other categories data of csv file format is downloaded from [link] https://catalog.data.gov/dataset

```
# Reading the dataset
NYS_Math_Test_Results_by_Grade <- read.csv("C:/Users/khush/Downloads/2006_-_2011_NYS_Math_Test_Results_

# Structure of the dataset
str(NYS_Math_Test_Results_by_Grade)
```

```
## 'data.frame':    168 obs. of  15 variables:
##  $ Grade           : chr  "3" "4" "5" "6" ...
##  $ Year            : int  2006 2006 2006 2006 2006 2006 2006 2007 2007 2007 ...
##  $ Category        : chr  "Asian" "Asian" "Asian" "Asian" ...
##  $ Number.Tested   : int  9768 9973 9852 9606 9433 9593 58225 9750 9881 10111 ...
##  $ Mean.Scale.Score: int  700 699 691 682 671 675 687 706 704 700 ...
##  $ Level.1..       : int  243 294 369 452 521 671 2550 156 209 211 ...
##  $ Level.1...1     : num  2.5 2.9 3.7 4.7 5.5 7 4.4 1.6 2.1 2.1 ...
##  $ Level.2..       : int  543 600 907 1176 1698 1847 6771 402 564 626 ...
##  $ Level.2...1     : num  5.6 6 9.2 12.2 18 19.3 11.6 4.1 5.7 6.2 ...
##  $ Level.3..       : int  4128 4245 4379 4646 4690 4403 26491 3886 3968 4257 ...
##  $ Level.3...1     : num  42.3 42.6 44.4 48.4 49.7 45.9 45.5 39.9 40.2 42.1 ...
##  $ Level.4..       : int  4854 4834 4197 3332 2524 2672 22413 5306 5140 5017 ...
##  $ Level.4...1     : num  49.7 48.5 42.6 34.7 26.8 27.9 38.5 54.4 52 49.6 ...
##  $ Level.3.4..     : int  8982 9079 8576 7978 7214 7075 48904 9192 9108 9274 ...
##  $ Level.3.4...1   : num  92 91 87 83.1 76.5 73.8 84 94.3 92.2 91.7 ...
```

```
# Top n rows of dataset
head(NYS_Math_Test_Results_by_Grade)
```

```
##   Grade Year Category Number.Tested Mean.Scale.Score Level.1.. Level.1...1
## 1     3 2006    Asian          9768              700       243         2.5
## 2     4 2006    Asian          9973              699       294         2.9
## 3     5 2006    Asian          9852              691       369         3.7
## 4     6 2006    Asian          9606              682       452         4.7
## 5     7 2006    Asian          9433              671       521         5.5
## 6     8 2006    Asian          9593              675       671         7.0
##   Level.2.. Level.2...1 Level.3.. Level.3...1 Level.4.. Level.4...1 Level.3.4..
## 1       543         5.6      4128        42.3      4854        49.7        8982
## 2       600         6.0      4245        42.6      4834        48.5        9079
## 3       907         9.2      4379        44.4      4197        42.6        8576
```

```
## 4      1176        12.2      4646        48.4      3332        34.7        7978
## 5      1698        18.0      4690        49.7      2524        26.8        7214
## 6      1847        19.3      4403        45.9      2672        27.9        7075
##   Level.3.4...1
## 1          92.0
## 2          91.0
## 3          87.0
## 4          83.1
## 5          76.5
## 6          73.8
```

By using read.csv command I can read the downloaded file from the system.

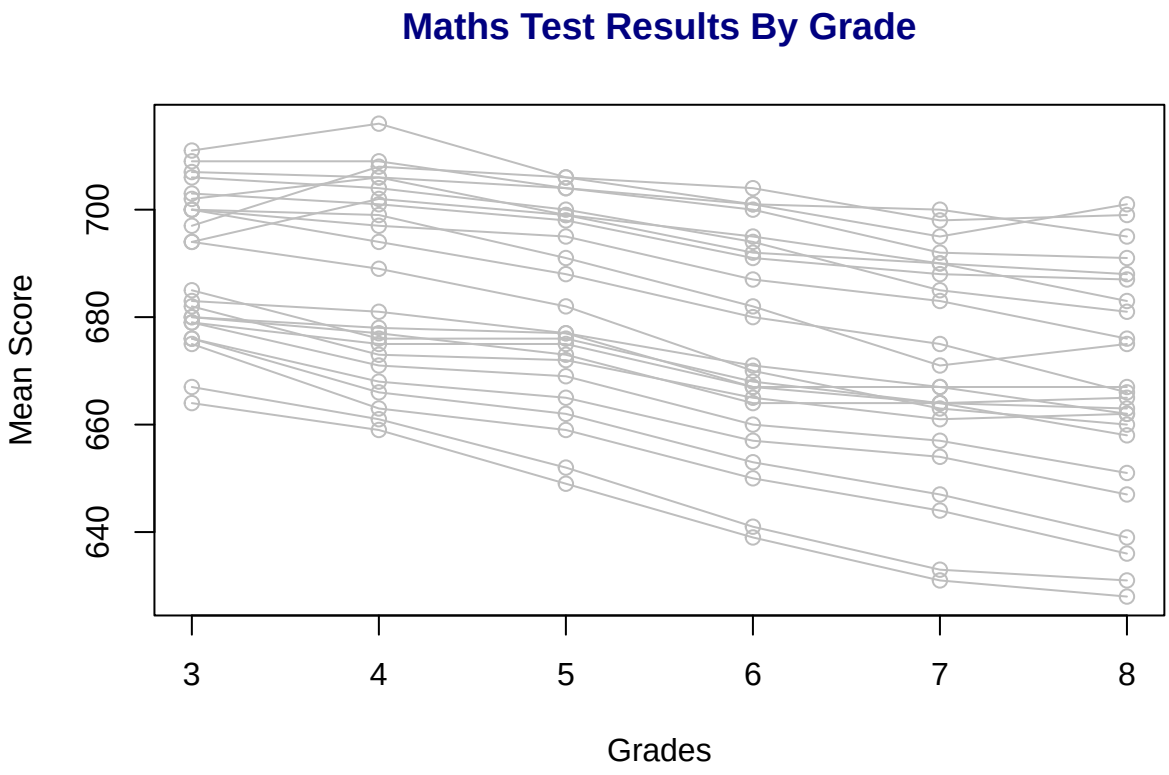
2.Produce summary statistics of the data : The output of the summary() function shows every variable a set of descriptive statistics, depending on the type of the variable: Numerical variables: summary() gives the range, quartiles, median, and mean. Numerical and factor variables: summary() gives number of missing values, if there are any.

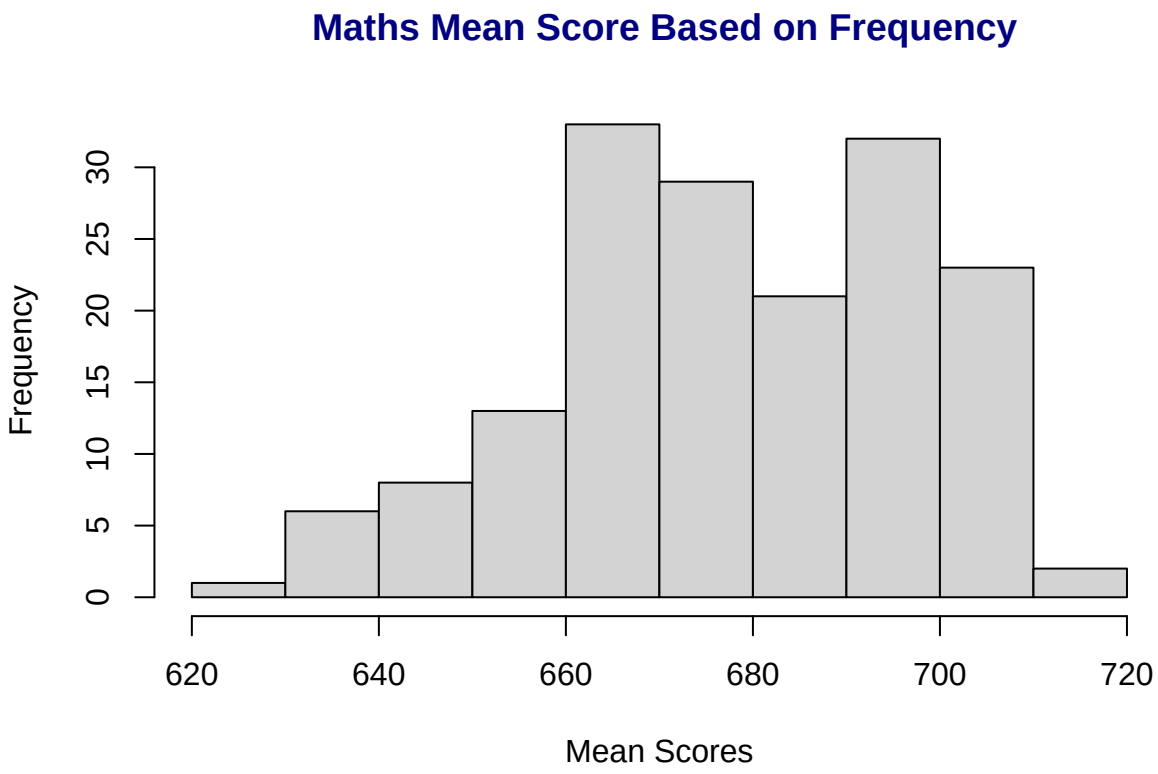
```
summary(NYS_Math_Test_Results_by_Grade)
```

```
##     Grade                Year        Category         Number.Tested
##  Length:168         Min.   :2006   Length:168         Min.   :  9433
##  Class :character   1st Qu.:2007   Class :character   1st Qu.: 10201
##  Mode  :character   Median :2008   Mode  :character   Median : 21127
##                     Mean   :2008                      Mean   : 30543
##                     3rd Qu.:2010                      3rd Qu.: 28593
##                     Max.   :2011                      Max.   :177382
##  Mean.Scale.Score   Level.1..      Level.1...1       Level.2..
##  Min.   :628.0    Min.   :   43   Min.   : 0.400   Min.   :  216
##  1st Qu.:664.0    1st Qu.:  333   1st Qu.: 2.875   1st Qu.: 1384
##  Median :677.5    Median : 1346   Median : 5.150   Median : 3992
##  Mean   :678.5    Mean   : 2876   Mean   : 7.532   Mean   : 7856
##  3rd Qu.:695.0    3rd Qu.: 3271   3rd Qu.:11.775   3rd Qu.: 9313
##  Max.   :716.0    Max.   :33091   Max.   :31.100   Max.   :70036
##   Level.2...1      Level.3..       Level.3...1      Level.4..
##  Min.   : 2.00   Min.   :  2762   Min.   :24.30   Min.   :  605
##  1st Qu.:11.78   1st Qu.:  4618   1st Qu.:36.27   1st Qu.: 2698
##  Median :18.60   Median :  7422   Median :42.95   Median : 4177
##  Mean   :21.65   Mean   : 13487   Mean   :43.79   Mean   : 6324
##  3rd Qu.:32.40   3rd Qu.: 14053   3rd Qu.:50.83   3rd Qu.: 5326
##  Max.   :49.60   Max.   :102188   Max.   :71.80   Max.   :33594
##   Level.4...1     Level.3.4..     Level.3.4...1
##  Min.   : 2.30   Min.   :  6491   Min.   :27.30
##  1st Qu.:12.45   1st Qu.:  8706   1st Qu.:54.23
##  Median :22.50   Median :  9976   Median :76.25
##  Mean   :27.02   Mean   : 19811   Mean   :70.81
##  3rd Qu.:42.60   3rd Qu.: 17366   3rd Qu.:86.05
##  Max.   :64.00   Max.   :132637   Max.   :97.60
```

By using the summary() , I can see the see summarized details about the dataset variables, such as mean,median,minimum-maximum ,25th and 75th quartiles values.

3.Produce a graph: Using **line chart** and **histogram** to analyse the data:





References: [link]https://catalog.data.gov/dataset

2006 - 2011 NYS Math Test Results by Grade - Citywide - by Race-Ethnicity