**1. Use the given link Data Set**

**Answer the below questions:**

**a. Find out top 5 attributes having highest correlation (select only Numeric features).**

Ans:

library(readr)

setwd("C:/Users/hemakumar/Downloads")

zipF<- "C:\\Users\\hemakumar\\Downloads\\united-states-crime-rates-by-county.zip"

outDir<-"crime\_data\_w\_population\_and\_crime\_rate.csv"

h<-unzip(zipF,exdir=outDir)

myData <- read.csv(h)

data<-data.frame(myData )

str(data)

data$county\_name<-as.numeric(data$county\_name)

c<-cor(data[,-1])

c1<-cor(data,method = "spearman")

c2<-cor(data,method = "kendall")

zdf <- as.data.frame(as.table(c))

mycorr <- .Last.value

which(mycorr > 0.991& mycorr != 1, arr.ind=TRUE)

top 5 attributes having highest correlation

#(FIPS\_ST,IDNO),(CPOPCRIM,CPOPARST),(population,CPOPCRIM),

(population,CPOPARST),(LARCENY,MODINDX)

**b. Find out top 3 reasons for having more crime in a city.**

Ans:

library(readr)

setwd("C:/Users/hemakumar/Downloads")

zipF<- "C:\\Users\\hemakumar\\Downloads\\united-states-crime-rates-by-county.zip"

outDir<-"crime\_data\_w\_population\_and\_crime\_rate.csv"

h<-unzip(zipF,exdir=outDir)

myData <- read.csv(h)

data<-data.frame(myData )

data$county\_name<-as.numeric(data$county\_name)

COR <- cor(as.matrix(data[,1]), as.matrix(data[,-1]))

COR

crime\_rate\_per\_100000 index EDITION PART IDNO CPOPARST CPOPCRIM

[1,] -0.001025064 -0.002339194 NA NA 0.04002424 0.01363593 0.0135424

AG\_ARRST AG\_OFF COVIND INDEX MODINDX MURDER RAPE

[1,] 0.01086486 0.005171146 -0.007700087 0.00638437 0.006256227 0.005876513 0.00564435

ROBBERY AGASSLT BURGLRY LARCENY MVTHEFT ARSON population

[1,] -0.004411223 0.01394834 0.006886243 0.004945489 0.01106265 0.002324574 0.01372858

FIPS\_ST FIPS\_CTY

[1,] 0.009569324 0.4978717

#by above we can say that top three reasons are FIPS\_CTY ,IDNO , AGASSLT

**c. Which all attributes have correlation with crime rate?**

COR <- cor(as.matrix(data[,2]), as.matrix(data[,-2]))

COR

county\_name index EDITION PART IDNO CPOPARST CPOPCRIM AG\_ARRST AG\_OFF

[1,] -0.001025064 -0.8880698 NA NA -0.1051217 0.2266388 0.2278254 0.2168906 0.2050074

COVIND INDEX MODINDX MURDER RAPE ROBBERY AGASSLT BURGLRY

[1,] 0.04927086 0.3390029 0.3003307 0.322898 0.3355057 0.2781407 0.3708294 0.3232806

LARCENY MVTHEFT ARSON population FIPS\_ST FIPS\_CTY

[1,] 0.2957901 0.2439377 0.2706489 0.2246898 -0.1043708 0.03708328