**NAME: SANDEEP PATIL**

**Email: saashasandeep@gmail.com**

**DATA ANALYTICS WITH R, EXCEL and TABLEAU**

**Session 13 – Assignment – 13.1**

#1. Use the given link Data Set

#Answer the below questions:

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

Crimes <- read.csv("D:/BIG DATA/DATA ANALYTICS WITH R, EXCEL & TABLEAU/13 GENERALIZED LINEAR MODELS/Crimes\_-\_2001\_to\_present.csv", header=FALSE)

View(Crimes)

names(Crimes) <- c("Case", "Number", "Date", "Block", "IUCR", "Primary Type", "Description",

"Location Desc", "Arrest", "Domestic", "Beat", "District", "Ward", "Community Area",

"FBI Code", "X Coordinate", "Y Coordinate", "Year", "Updated On",

"Latitude", "Longitude", "Location")

head(Crimes)

tail(Crimes)

str(Crimes)

Crimes <- na.omit(Crimes)

names(Crimes)

c = as.numeric(Crimes)

sapply(Crimes, class)

cmain <- cor(Crimes[c(11,12,13,14,18,20,21)])

cmain

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

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