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

**Answer the below questions:**

**a. Visualize the correlation between all variables in a meaningful and clear way of representing. Find out**

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

Ans:

library(corrplot)

corrplot(c)

corrplot(c,method = "pie")

corrplot(c,method = "color")

corrplot(c,method = "number")

corrplot(c,method = "pie")

corrplot(c,type = "lower")

# top three reasons for having more crime in a city are FIPS\_CTY ,IDNO , AGASSLT .

**b. What is the difference between co-variance and correlation? Take an example from this dataset and**

**show the differences if any?**

Ans:

BASIS FOR COMPARISON COVARIANCE CORRELATION

Meaning Covariance is a measure Correlation is a statistical

indicating the extent to measure that indicates how

which two random variables strongly two variables

change in tandem. are related.

What is it? Measure of correlation Scaled version of covariance

Values Lie between -∞ and +∞ Lie between -1 and +1

Change in scale Affects covariance Does not affects correlation

Unit free measure No Yes

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Cov(X,Y)

Covariance vs. Correlation ρ(X,Y)= \_\_\_\_\_\_\_\_

σX σY

Where:

ρ(X,Y) – the correlation between the variables X and Y

Cov(X,Y) – the covariance between the variables X and Y

σX – the standard deviation of the X-variable

σY – the standard deviation of the Y-variable

cov(data$MURDER,data$BURGLRY)

cor(data$MURDER,data$BURGLRY)

#in above code we have covariance 45282.7 and correlation 0.8653636

mydata<-data[1:1000,]

cov(mydata$MURDER,mydata$BURGLRY)

cor(mydata$MURDER,mydata$BURGLRY)

# here we can see when scale is decreased we have covarience 129056.2 and correlation 0.8638122

# by this we can know the covarience is different and correlation in same