Assignment 12.2

1. Use the given link below:
https://archive.ics.uci.edu/ml/machine-learning-databases/communities/

Perform the below operations:

a. Visualize the correlation between all variable in a meaningful

way, clear representation of correlations. Find out top 3

reasons for having more crime in a city.

Corelation plots attached, since there are about 110 features i have split it into 5 and 5 corplots are attached

plot_zoom_file1.png,plot_zoom_file2.png,plot_zoom_file3.png,plot_zoom_file4.png,plot_zoom_file5.png

Education - PctLess9thGrade
Marital Status - MalePctNevMarr

Percent of Black Police : PctPolicBlack e.t.c are some of the main reasons for crime in city

b. What is the difference between covariance and correlation,

take an example from this dataset and show the differences if any

1. Covariance is nothing but a measure of correlation. On the contrary, correlation refers to the scaled form of covariance.
2. The value of correlation takes place between -1 and +1. Conversely, the value of covariance lies between -∞ and +∞.
3. Covariance is affected by the change in scale, i.e. if all the value of one variable is multiplied by a constant and all the value of another variable are multiplied, by a similar or different constant, then the covariance is changed. As against this, correlation is not influenced by the change in scale.
4. Correlation is dimensionless, i.e. it is a unit-free measure of the relationship between variables. Unlike covariance, where the value is obtained by the product of the units of the two variables.

Example

```
communityData_eg <-
communityData[,c('ViolentCrimesPerPop','PctPolicBlack','NumKindsDrugsSeiz')]

cov_det <- cov(communityData_eg, y = NULL, use = "all.obs",method = c("pearson", "kendall", "spearman"))
```

print(cov_det)

```
print(cov_det)
                    ViolentCrimesPerPop PctPolicBlack NumKindsDrugsSeiz
ViolentCrimesPerPop         0.076369646   0.035490004       0.007398431
PctPolicBlack               0.035490004   0.055823992       0.002728593
NumKindsDrugsSeiz           0.007398431   0.002728593       0.041281205
```

cor_det <- cor(communityData_eg, y = NULL, use = "all.obs", method = c("pearson", "kendall", "spearman"))

```
print(cor_det)
                    ViolentCrimesPerPop PctPolicBlack NumKindsDrugsSeiz
ViolentCrimesPerPop           1.0000000    0.54354464        0.13176590
PctPolicBlack                 0.5435446    1.00000000        0.05683974
NumKindsDrugsSeiz             0.1317659    0.05683974        1.00000000
```