```
> #Assignment12.2
> my_data <- read.delim("community.txt",sep = ",")
>
> #a Find top attributes having highest correlation (select only Numeric)
> options(max.print = 99999)
> nums <- unlist(lapply(my_data, is.numeric))
> numeric_attributes<- my_data[,nums]
> correlation <-as.data.frame(cor(numeric_attributes))
>
> #replacing all values of 1 with 0.
> correlation[correlation==1]<-0
> correlation1<-as.matrix(correlation)
>
> #a. Visualize the correlation between all variable in a meaningful
> #way, clear representation of correlations
> library(corrplot)
> #Positive correlations are displayed in blue and negative correlations in r
ed color
> corrplot(correlation1, method = "color")
> #The column X0.2.2 is ViolentCrimesPerPop which best identifies having more
crime.
> # Hence we will take the 3 highest correlations for this column and identif
y the 3 reasons.
> violent3<-correlation1[,102]
> violent3<-as.matrix(violent3)
> sort(violent3,decreasing = TRUE)
  [1]  0.73796471  0.63127917  0.57468959  0.55647177  0.55316446  0.52569861
0.52187160
  [8]  0.50422724  0.48823696  0.48340928  0.48282198  0.47450097  0.47103421
0.45289031
 [15]  0.44760441  0.42155395  0.41112107  0.38346973  0.37576608  0.36734559
0.36445578
 [22]  0.36308897  0.34878391  0.34026490  0.32502870  0.30457183  0.30000010
0.29556250
 [29]  0.29478403  0.29424191  0.29306223  0.29152480  0.28140359  0.26424179
0.25317516
 [36]  0.24831458  0.24812650  0.24802146  0.23074437  0.21602527  0.19685452
0.19436624
 [43]  0.17185474  0.15384464  0.15334517  0.14060739  0.11792479  0.09932653
0.08209145
 [50]  0.07551730  0.06713037  0.06385163  0.06043771  0.05370691  0.03760909
0.02121239
 [57]  0.00000000 -0.01946384 -0.02243101 -0.03272495 -0.03499534 -0.03980919
-0.04498121
 [64] -0.07149790 -0.07725173 -0.09082511 -0.09846907 -0.10995268 -0.12448456
-0.15051730
 [71] -0.15310812 -0.15540649 -0.15561743 -0.17216035 -0.19075628 -0.20926471
-0.21056961
 [78] -0.21158667 -0.23230769 -0.23988486 -0.24050482 -0.24144272 -0.24454495
-0.25184378
 [85] -0.27538110 -0.30549013 -0.31465626 -0.31901600 -0.33164891 -0.33909211
-0.35207176
 [92] -0.35737501 -0.42422017 -0.43910533 -0.47070124 -0.52550042 -0.57632919
-0.66159821
 [99] -0.66609375 -0.68478685 -0.70671300 -0.73844498
> #X.0.14,X.0.02,X.0.15.
> #Highest reason -      - PctRecImmig5- 0.73796471
> #Second highest reason - agePct65up  - 0.63127917
> #Third highest reason  - AsianPerCap - 0.57468959
> #b. What is the difference between covariance and correlation,
> #take an example from this dataset and show the differences if
> #any?
> #Attached pdf explains the detailed difference.
```

```
> #correlation  - refer to above calculated
> #covariance - as below
> a<- cov(correlation$X0.19,correlation$X0.34)
> a
[1] 0.01542169
```