Randall_Plyler_R_Ch3-4

Randall Plyler

1/22/2022

```
getwd()
```

```
## [1] "C:/Users/randa/Dropbox/Masters/Winter/TBANLT 560 Data Mining/Module2"
```

```
housing.df <- read.csv("C:/Users/randa/Dropbox/Masters/Winter/TBANLT 560 Data Mining/Files/DMBA-R-datas
```

```
#Randall Code write CSV
library(ggmap)
```

```
## Loading required package: ggplot2
```

```
## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.
```

```
## Please cite ggmap if you use it! See citation("ggmap") for details.
```

```
correlation_matrix <- cor(housing.df)
correlation_matrix
```

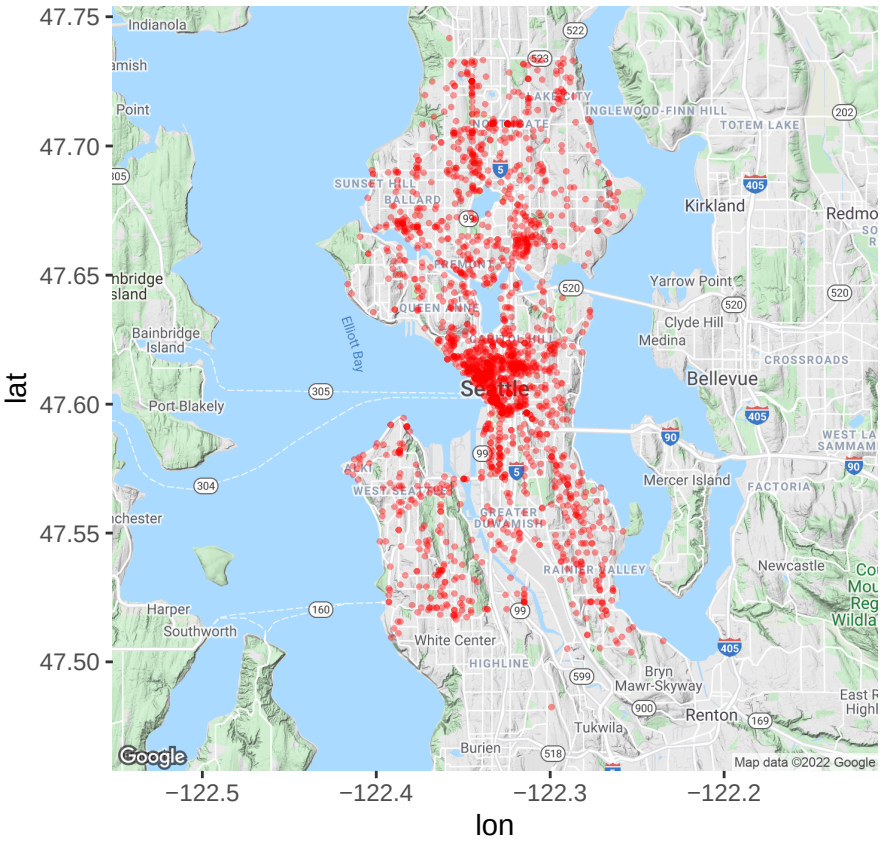
```
##                  CRIM          ZN       INDUS         CHAS         NOX
## CRIM       1.00000000 -0.20046922  0.40658341 -0.055891582  0.42097171
## ZN        -0.20046922  1.00000000 -0.53382819 -0.042696719 -0.51660371
## INDUS      0.40658341 -0.53382819  1.00000000  0.062938027  0.76365145
## CHAS      -0.05589158 -0.04269672  0.06293803  1.000000000  0.09120281
## NOX        0.42097171 -0.51660371  0.76365145  0.091202807  1.00000000
## RM        -0.21924670  0.31199059 -0.39167585  0.091251225 -0.30218819
## AGE        0.35273425 -0.56953734  0.64477851  0.086517774  0.73147010
## DIS       -0.37967009  0.66440822 -0.70802699 -0.099175780 -0.76923011
## RAD        0.62550515 -0.31194783  0.59512927 -0.007368241  0.61144056
## TAX        0.58276431 -0.31456332  0.72076018 -0.035586518  0.66802320
## PTRATIO    0.28994558 -0.39167855  0.38324756 -0.121515174  0.18893268
## LSTAT      0.45562148 -0.41299457  0.60379972 -0.053929298  0.59087892
## MEDV      -0.38830461  0.36044534 -0.48372516  0.175260177 -0.42732077
## CAT..MEDV -0.15198696  0.36529623 -0.36627559  0.108631150 -0.23250184
##                   RM         AGE         DIS          RAD         TAX
## CRIM      -0.21924670  0.35273425 -0.37967009  0.625505145  0.58276431
## ZN         0.31199059 -0.56953734  0.66440822 -0.311947826 -0.31456332
## INDUS     -0.39167585  0.64477851 -0.70802699  0.595129275  0.72076018
## CHAS       0.09125123  0.08651777 -0.09917578 -0.007368241 -0.03558652
## NOX       -0.30218819  0.73147010 -0.76923011  0.611440563  0.66802320
```

```
## RM          1.00000000 -0.24026493  0.20524621 -0.209846668 -0.29204783
## AGE        -0.24026493  1.00000000 -0.74788054  0.456022452  0.50645559
## DIS         0.20524621 -0.74788054  1.00000000 -0.494587930 -0.53443158
## RAD        -0.20984667  0.45602245 -0.49458793  1.000000000  0.91022819
## TAX        -0.29204783  0.50645559 -0.53443158  0.910228189  1.00000000
## PTRATIO    -0.35550149  0.26151501 -0.23247054  0.464741179  0.46085304
## LSTAT      -0.61380827  0.60233853 -0.49699583  0.488676335  0.54399341
## MEDV        0.69535995 -0.37695457  0.24992873 -0.381626231 -0.46853593
## CAT..MEDV   0.64126541 -0.19119589  0.11888651 -0.197924023 -0.27368672
##               PTRATIO      LSTAT       MEDV  CAT..MEDV
## CRIM        0.2899456  0.4556215 -0.3883046 -0.1519870
## ZN         -0.3916785 -0.4129946  0.3604453  0.3652962
## INDUS       0.3832476  0.6037997 -0.4837252 -0.3662756
## CHAS       -0.1215152 -0.0539293  0.1752602  0.1086312
## NOX         0.1889327  0.5908789 -0.4273208 -0.2325018
## RM         -0.3555015 -0.6138083  0.6953599  0.6412654
## AGE         0.2615150  0.6023385 -0.3769546 -0.1911959
## DIS        -0.2324705 -0.4969958  0.2499287  0.1188865
## RAD         0.4647412  0.4886763 -0.3816262 -0.1979240
## TAX         0.4608530  0.5439934 -0.4685359 -0.2736867
## PTRATIO     1.0000000  0.3740443 -0.5077867 -0.4434247
## LSTAT       0.3740443  1.0000000 -0.7376627 -0.4699108
## MEDV       -0.5077867 -0.7376627  1.0000000  0.7897888
## CAT..MEDV -0.4434247 -0.4699108  0.7897888  1.0000000
```

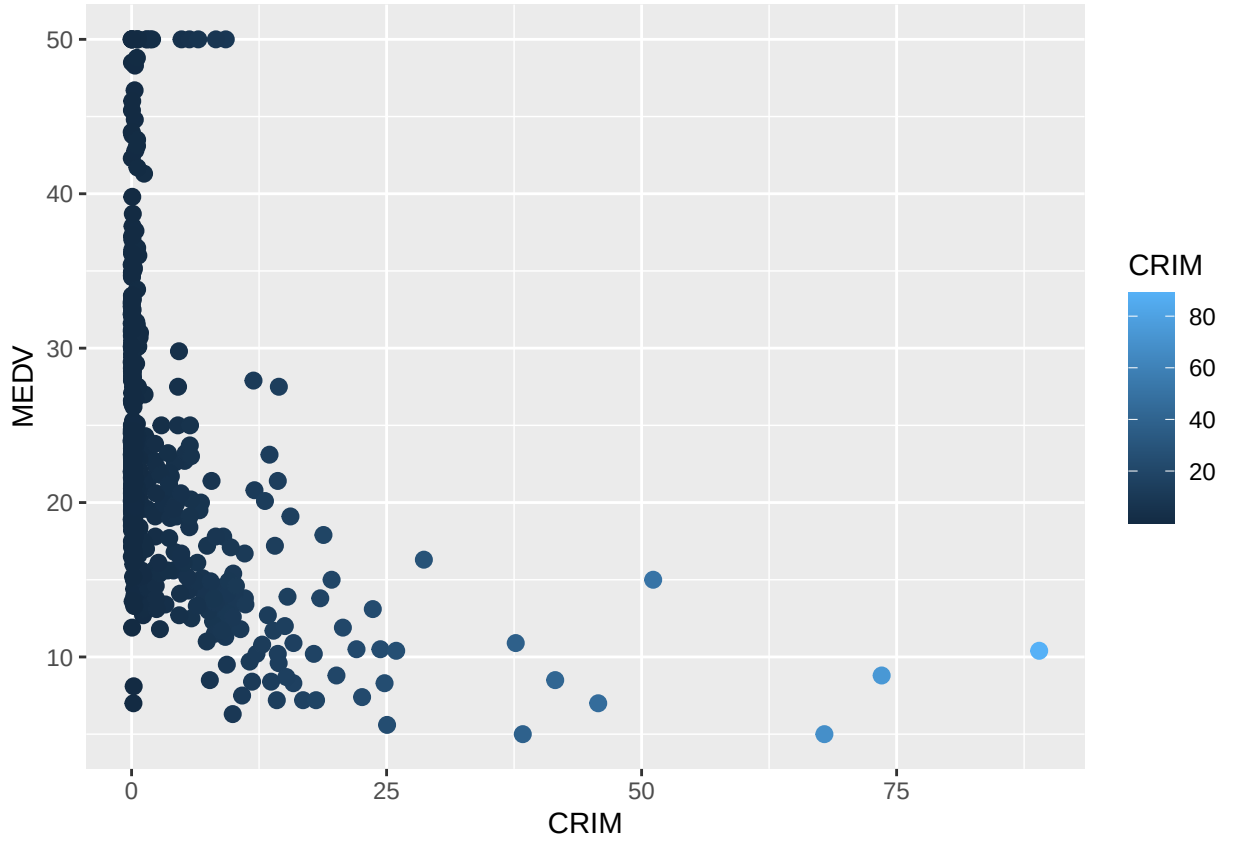
```
write.csv(correlation_matrix,
          "C:/Users/randa/Dropbox/Masters/Winter/TBANLT 560 Data Mining/Files/DMBA-R-datasets/DMBA-R-da

randall_correlation_matrix <- read.csv("C:/Users/randa/Dropbox/Masters/Winter/TBANLT 560 Data Mining/Fi

longitude_lattitude_data <- read.csv("C:/Users/randa/Dropbox/Masters/Winter/TBANLT 560 Data Mining/File

register_google(key = "AIzaSyBI3Nj5LhgKaOwcaRClCioZQelBFmd7ZOc")
Map <- get_map("Seattle, WA", zoom = 11)
```

```
## Source : https://maps.googleapis.com/maps/api/staticmap?center=Seattle,%20WA&zoom=11&size=640x640&sc

## Source : https://maps.googleapis.com/maps/api/geocode/json?address=Seattle,+WA&key=xxx
```

```
ggmap(Map) + geom_point(aes(x = Longitude, y = Latitude), data = longitude_lattitude_data,
                        alpha = 0.4, colour = "red", size = .5)
```




```
ggplot(housing.df, aes(x=CRIM, y=MEDV, colour=CRIM)) + geom_point(shape=19, size=2.5)
```




```
housing.df$CRIM <- ifelse(housing.df$CRIM > 1, 1, 0)
head(housing.df, 10)
```

```
##    CRIM   ZN INDUS CHAS   NOX    RM   AGE    DIS RAD TAX PTRATIO LSTAT MEDV
## 1     0 18.0  2.31    0 0.538 6.575  65.2 4.0900   1 296    15.3  4.98 24.0
## 2     0  0.0  7.07    0 0.469 6.421  78.9 4.9671   2 242    17.8  9.14 21.6
## 3     0  0.0  7.07    0 0.469 7.185  61.1 4.9671   2 242    17.8  4.03 34.7
## 4     0  0.0  2.18    0 0.458 6.998  45.8 6.0622   3 222    18.7  2.94 33.4
## 5     0  0.0  2.18    0 0.458 7.147  54.2 6.0622   3 222    18.7  5.33 36.2
## 6     0  0.0  2.18    0 0.458 6.430  58.7 6.0622   3 222    18.7  5.21 28.7
## 7     0 12.5  7.87    0 0.524 6.012  66.6 5.5605   5 311    15.2 12.43 22.9
## 8     0 12.5  7.87    0 0.524 6.172  96.1 5.9505   5 311    15.2 19.15 27.1
## 9     0 12.5  7.87    0 0.524 5.631 100.0 6.0821   5 311    15.2 29.93 16.5
## 10    0 12.5  7.87    0 0.524 6.004  85.9 6.5921   5 311    15.2 17.10 18.9
##    CAT..MEDV
## 1          0
## 2          0
## 3          1
## 4          1
## 5          1
## 6          0
## 7          0
## 8          0
## 9          0
## 10         0
```