DS311 - R Lab Assignment

Miles Mendez

11/28/23

R Assignment 1

- In this assignment, we are going to apply some of the build in data set in R for descriptive statistics analysis.
- To earn full grade in this assignment, students need to complete the coding tasks for each question to get the result.
- After finished all the questions, knit the document into HTML format for submission.

Question 1

Using the **mtcars** data set in R, please answer the following questions.

```
# Loading the data
data(mtcars)

# Head of the data set
head(mtcars)
```

```
##                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

a. Report the number of variables and observations in the data set.

```
# Enter your code here!
?mtcars
```

```
## starting httpd help server ... done
```

```
dim(mtcars)
```

```
## [1] 32 11
```

```r
names(mtcars)
```

```
##  [1] "mpg"  "cyl"  "disp" "hp"   "drat" "wt"   "qsec" "vs"   "am"   "gear"
## [11] "carb"
```

```r
# Answer:
print("There are total of 11 numeric variables and 32 observations in this data set.")
```

```
## [1] "There are total of 11 numeric variables and 32 observations in this data set."
```

b. Print the summary statistics of the data set and report how many discrete and continuous variables are in the data set.

```r
# Enter your code here!
summary(mtcars)
```

```
##       mpg             cyl             disp             hp
##  Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0
##  1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5
##  Median :19.20   Median :6.000   Median :196.3   Median :123.0
##  Mean   :20.09   Mean   :6.188   Mean   :230.7   Mean   :146.7
##  3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0
##  Max.   :33.90   Max.   :8.000   Max.   :472.0   Max.   :335.0
##       drat             wt             qsec             vs
##  Min.   :2.760   Min.   :1.513   Min.   :14.50   Min.   :0.0000
##  1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000
##  Median :3.695   Median :3.325   Median :17.71   Median :0.0000
##  Mean   :3.597   Mean   :3.217   Mean   :17.85   Mean   :0.4375
##  3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000
##  Max.   :4.930   Max.   :5.424   Max.   :22.90   Max.   :1.0000
##        am              gear            carb
##  Min.   :0.0000   Min.   :3.000   Min.   :1.000
##  1st Qu.:0.0000   1st Qu.:3.000   1st Qu.:2.000
##  Median :0.0000   Median :4.000   Median :2.000
##  Mean   :0.4062   Mean   :3.688   Mean   :2.812
##  3rd Qu.:1.0000   3rd Qu.:4.000   3rd Qu.:4.000
##  Max.   :1.0000   Max.   :5.000   Max.   :8.000
```

```r
# Answer:
print("There are 5 discrete variables and 6 continuous variables in this data set.")
```

```
## [1] "There are 5 discrete variables and 6 continuous variables in this data set."
```

c. Calculate the mean, variance, and standard deviation for the variable **mpg** and assign them into variable names m, v, and s. Report the results in the print statement.

```r
# Enter your code here!
m <- mean(mtcars$mpg)
v<-var(mtcars$mpg)
s<-sd(mtcars$mpg)
```

```
print(paste("The average of Mile Per Gallon from this data set is ", m , " with variance ", v , " and s
```

```
## [1] "The average of Mile Per Gallon from this data set is  20.090625  with variance  36.324102822580
```

d. Create two tables to summarize 1) average mpg for each cylinder class and 2) the standard deviation of mpg for each gear class.

```
# Enter your code here!
m_mpg_cyl<-aggregate(mpg ~ cyl, mtcars, mean)
sd_mpg_gr<-aggregate(mpg ~ cyl, mtcars, sd)

print(m_mpg_cyl)
```

```
##   cyl      mpg
## 1   4 26.66364
## 2   6 19.74286
## 3   8 15.10000
```

```
print(sd_mpg_gr)
```

```
##   cyl      mpg
## 1   4 4.509828
## 2   6 1.453567
## 3   8 2.560048
```

e. Create a crosstab that shows the number of observations belong to each cylinder and gear class combinations. The table should show how many observations given the car has 4 cylinders with 3 gears, 4 cylinders with 4 gears, etc. Report which combination is recorded in this data set and how many observations for this type of car.

```
# Enter your code here!
cyl_gr_ctable<-table(mtcars$cyl, mtcars$gear)
print(cyl_gr_ctable)
```

```
##
##      3  4  5
##   4  1  8  2
##   6  2  4  1
##   8 12  0  2
```

```
print("The most common car type in this data set is car with 8 cylinders and 3 gears. There are total o
```

```
## [1] "The most common car type in this data set is car with 8 cylinders and 3 gears. There are total
```

Question 2

Use different visualization tools to summarize the data sets in this question.

a. Using the **PlantGrowth** data set, visualize and compare the weight of the plant in the three separated group. Give labels to the title, x-axis, and y-axis on the graph. Write a paragraph to summarize your findings.

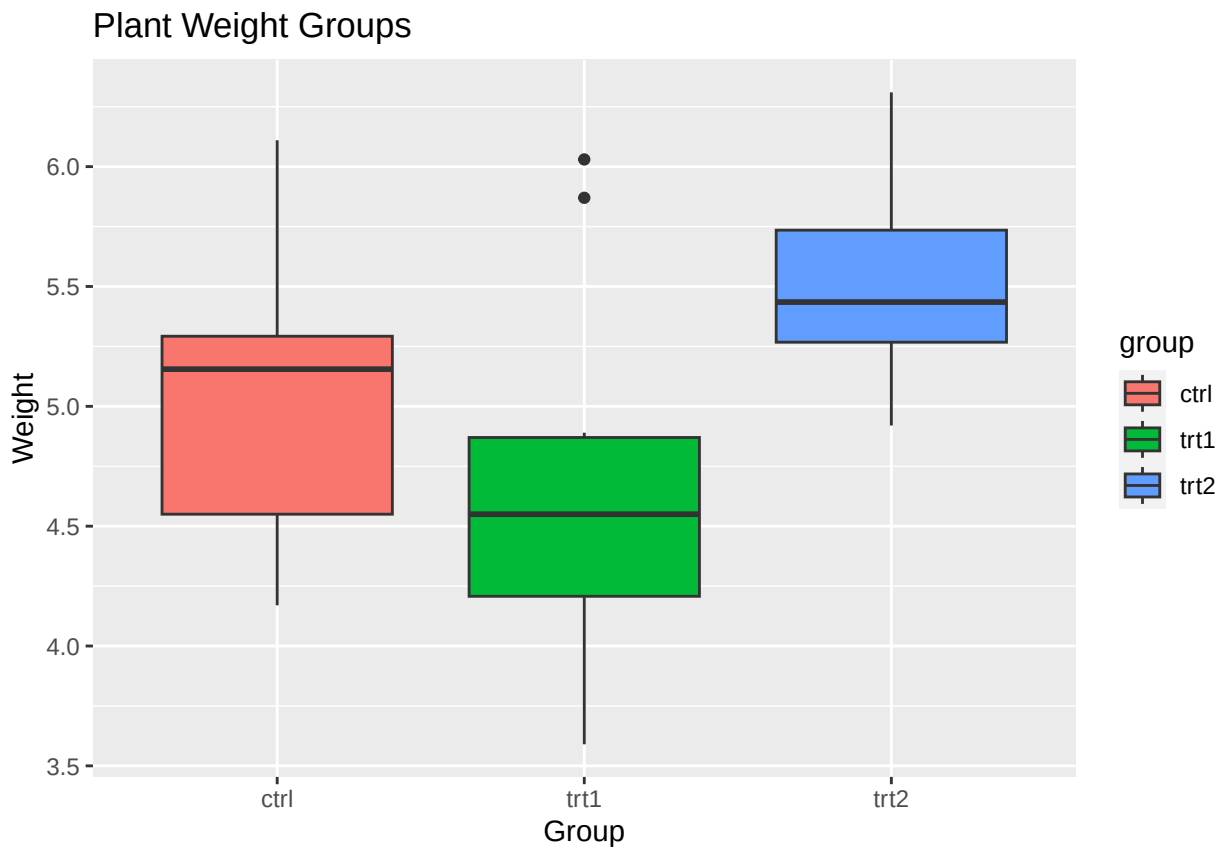
```
# Load the data set
data("PlantGrowth")

# Head of the data set
head(PlantGrowth)
```

```
##   weight group
## 1   4.17  ctrl
## 2   5.58  ctrl
## 3   5.18  ctrl
## 4   6.11  ctrl
## 5   4.50  ctrl
## 6   4.61  ctrl
```

```
# Enter your code here!

library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

```
ggplot(PlantGrowth, aes(x=group, y=weight, fill=group))+geom_boxplot()+
  labs(title="Plant Weight Groups", x="Group", y="Weight")
```

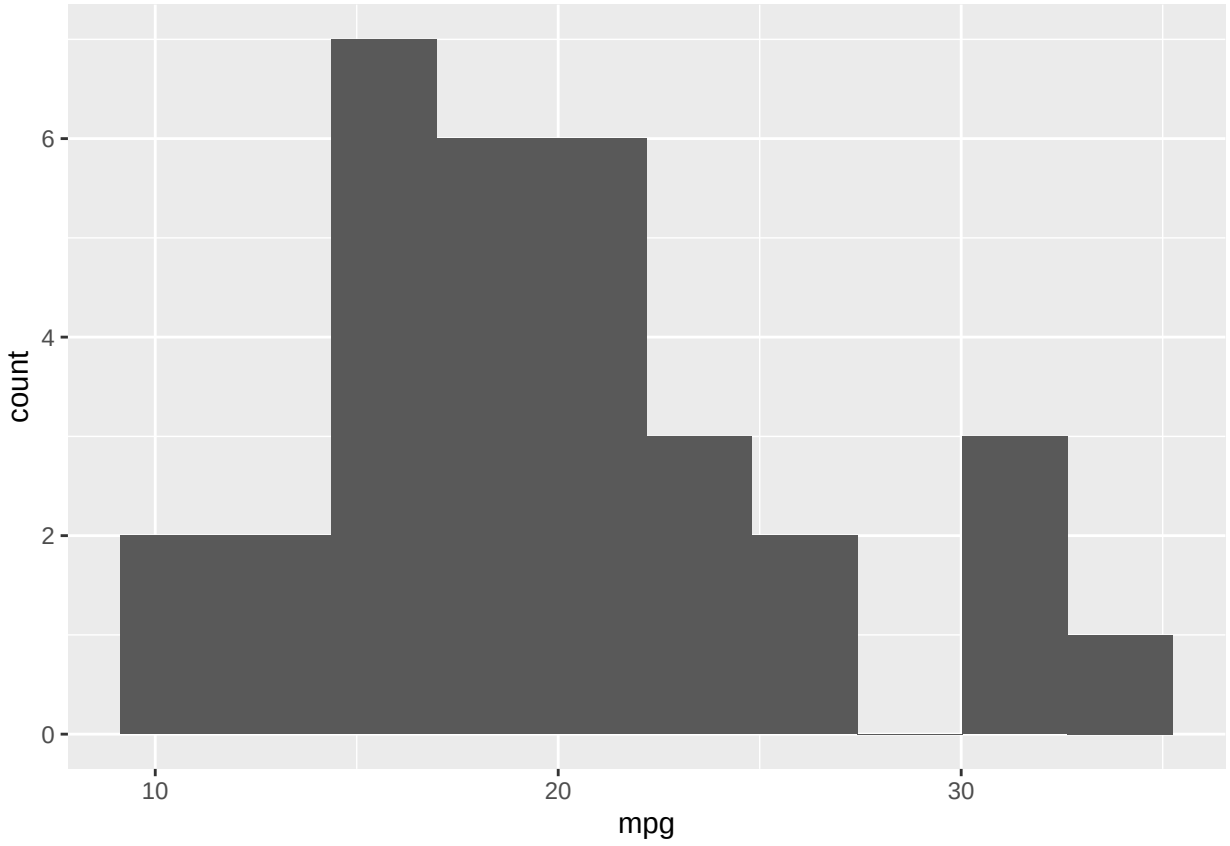


Result:

=> Report a paragraph to summarize your findings from the plot! The Plant weight groups reflect that the treatment groups provided vastly different results on both ends of the control group. Treatment 1 had the most similarities with the control group while treatment 2 was on the upper end of the control range barely intersecting at all. Treatment 2 produced bigger, healthier plants given this data. Treatment 2 had the highest median followed by the control and then treatment 1.The upper quartile for treatment 1 was very small with most plants in this group reflecting quartile 3 and 2. the median for the control group was roughly ~5.2, treatment 1 ~4.55, and treatment 3 at ~5.4.

b. Using the **mtcars** data set, plot the histogram for the column **mpg** with 10 breaks. Give labels to the title, x-axis, and y-axis on the graph. Report the most observed mpg class from the data set.

```
ggplot(data=mtcars)+geom_histogram(aes(x=mpg),bins=10)
```




```
print("Most of the cars in this data set are in the class of 15 mile per gallon.")
```

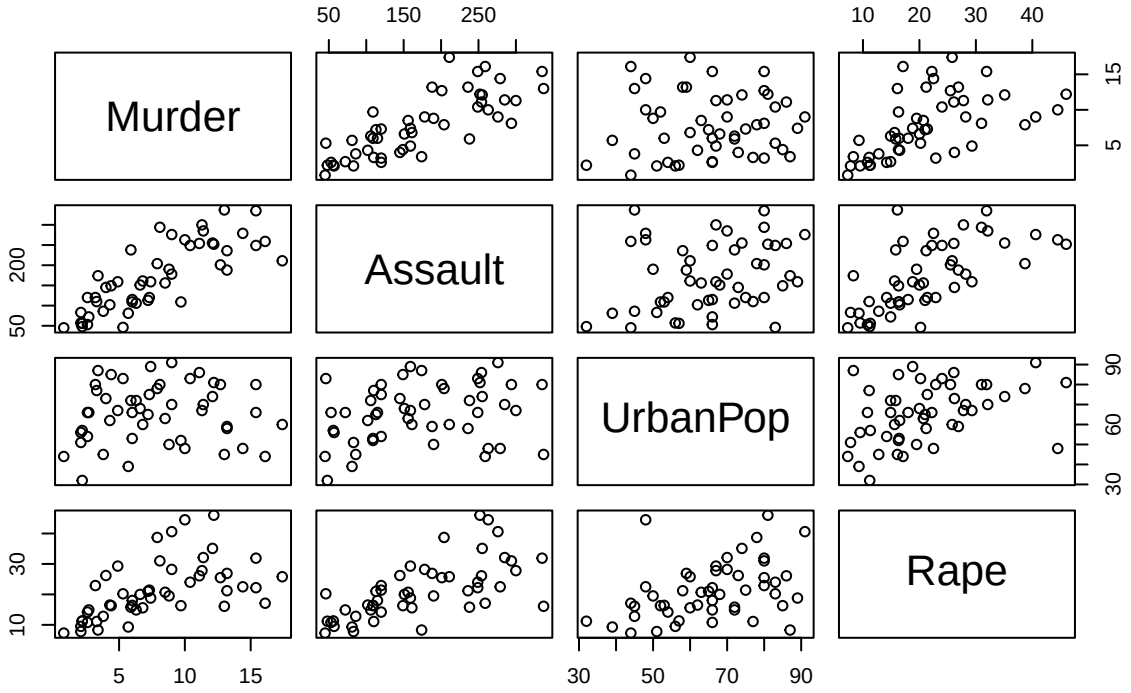
```
## [1] "Most of the cars in this data set are in the class of 15 mile per gallon."
```

c. Using the **USArrests** data set, create a pairs plot to display the correlations between the variables in the data set. Plot the scatter plot with **Murder** and **Assault**. Give labels to the title, x-axis, and y-axis on the graph. Write a paragraph to summarize your results from both plots.

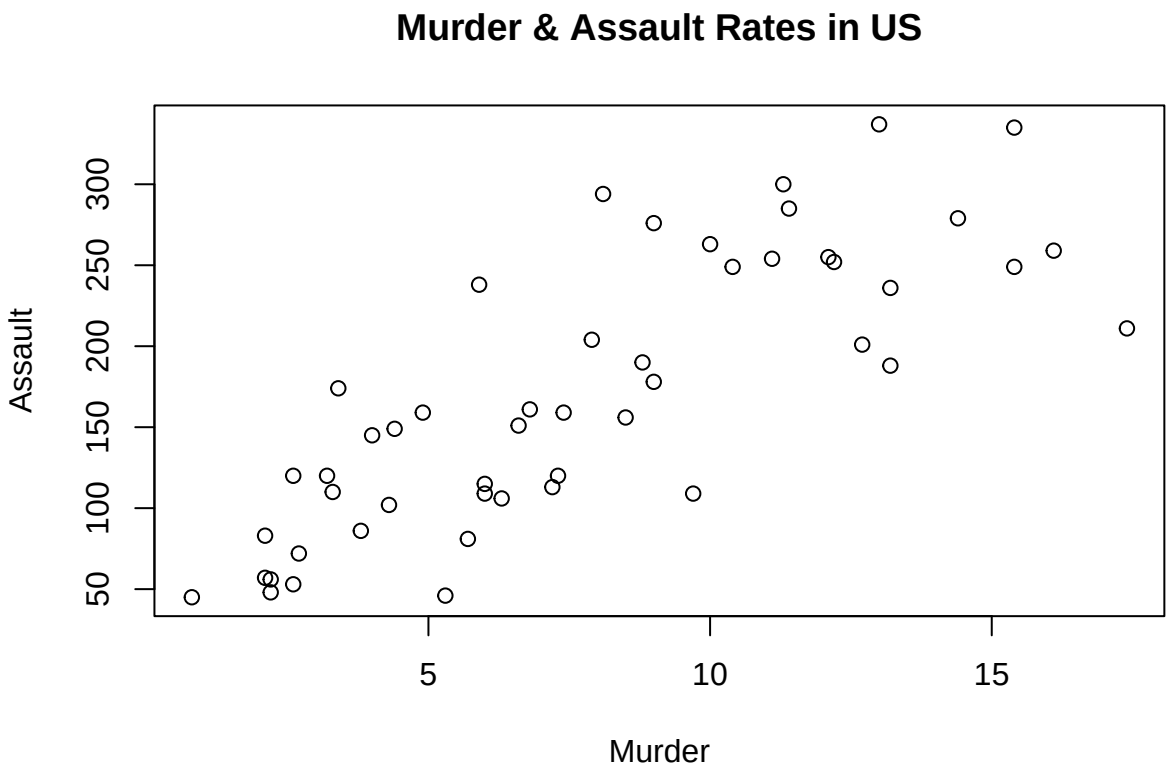
```
# Load the data set
data("USArrests")

# Head of the data set
head(USArrests)
```

```
##            Murder Assault UrbanPop Rape
## Alabama      13.2     236       58 21.2
## Alaska       10.0     263       48 44.5
## Arizona       8.1     294       80 31.0
## Arkansas      8.8     190       50 19.5
## California    9.0     276       91 40.6
## Colorado      7.9     204       78 38.7
```

```r
# Enter your code here!

pairs(USArrests, main="US Arrest Pair Plots")
```




```r
plot(USArrests$Murder, USArrests$Assault, xlab= "Murder", ylab="Assault", main="Murder & Assault Rates
```



Result:

=> Report a paragraph to summarize your findings from the plot!

Assaults are far more common than murder. As assaults increase, murders also increase showing a positive correlation.The trend also seems pretty linear in nature.

Question 3

Download the housing data set from www.jaredlander.com and find out what explains the housing prices in New York City.

Note: Check your working directory to make sure that you can download the data into the data folder.

a. Create your own descriptive statistics and aggregation tables to summarize the data set and find any meaningful results between different variables in the data set.

```
# Head of the cleaned data set
head(housingData)
```

```
##   Neighborhood Market.Value.per.SqFt      Boro Year.Built
## 1    FINANCIAL                200.00 Manhattan       1920
## 2    FINANCIAL                242.76 Manhattan       1985
## 4    FINANCIAL                271.23 Manhattan       1930
```

```
## 5      TRIBECA                247.48 Manhattan       1985
## 6      TRIBECA                191.37 Manhattan       1986
## 7      TRIBECA                211.53 Manhattan       1985
```

```
# Enter your code here!
summary(housingData)
```

```
##  Neighborhood       Market.Value.per.SqFt     Boro             Year.Built
##  Length:2530        Min.   : 10.66        Length:2530        Min.   :1825
##  Class :character   1st Qu.: 75.10        Class :character   1st Qu.:1926
##  Mode  :character   Median :114.89        Mode  :character   Median :1986
##                     Mean   :133.17                           Mean   :1967
##                     3rd Qu.:189.91                           3rd Qu.:2005
##                     Max.   :399.38                           Max.   :2010
```

```
MVSFmean<-aggregate(Market.Value.per.SqFt~Boro, data=housingData, FUN=mean)
MVSFvar<-aggregate(Market.Value.per.SqFt~Boro, data=housingData, FUN=var)
MVSFsd<-aggregate(Market.Value.per.SqFt~Boro, data=housingData, FUN=sd)

print(MVSFmean)
```

```
##            Boro Market.Value.per.SqFt
## 1         Bronx              47.93232
## 2      Brooklyn              80.13439
## 3     Manhattan             180.59265
## 4        Queens              77.38137
## 5 Staten Island              41.26958
```

```
print(MVSFvar)
```

```
##            Boro Market.Value.per.SqFt
## 1         Bronx              227.3035
## 2      Brooklyn              524.9035
## 3     Manhattan             3146.9929
## 4        Queens              947.1520
## 5 Staten Island               46.1433
```

```
print(MVSFsd)
```

```
##            Boro Market.Value.per.SqFt
## 1         Bronx             15.076589
## 2      Brooklyn             22.910773
## 3     Manhattan             56.098065
## 4        Queens             30.775835
## 5 Staten Island              6.792886
```

```
MVYBmean<-aggregate(Market.Value.per.SqFt~Year.Built, data=housingData, FUN=mean)
MVYBvar<-aggregate(Market.Value.per.SqFt~Year.Built, data=housingData, FUN=var)
MVYBsd<-aggregate(Market.Value.per.SqFt~Year.Built, data=housingData, FUN=sd)

print(MVYBmean)
```

```
##     Year.Built Market.Value.per.SqFt
## 1         1825              76.36000
## 2         1836             273.77000
## 3         1853             152.79000
## 4         1860             159.64500
## 5         1874             111.17000
## 6         1875             166.05000
## 7         1879             194.52000
## 8         1881             109.70500
## 9         1883             172.10000
## 10        1890             113.28750
## 11        1891              72.83000
## 12        1892              95.21000
## 13        1893             168.85000
## 14        1894             110.62000
## 15        1895             151.77500
## 16        1896             117.26500
## 17        1897              40.83000
## 18        1898              83.25000
## 19        1899             108.16000
## 20        1900             137.55908
## 21        1901             172.36778
## 22        1902             167.62167
## 23        1903             147.97000
## 24        1904             123.09333
## 25        1905             187.76583
## 26        1906             169.03364
## 27        1907             173.80000
## 28        1908             150.35000
## 29        1909             135.23667
## 30        1910             147.36257
## 31        1911             179.76067
## 32        1912             159.51636
## 33        1913             175.93500
## 34        1914             160.29286
## 35        1915             147.08673
## 36        1916             128.20714
## 37        1917              73.87000
## 38        1918             181.84000
## 39        1919              63.11000
## 40        1920             145.30862
## 41        1921             122.39125
## 42        1922             118.33250
## 43        1923             115.47625
## 44        1924             165.94091
## 45        1925             147.51316
## 46        1926             148.36423
## 47        1927             131.63357
## 48        1928             153.68375
## 49        1929             106.32121
## 50        1930             142.28936
## 51        1931             129.51731
## 52        1932              91.74333
## 53        1933              40.97000
```

```
## 54        1934            203.80000
## 55        1935            176.23000
## 56        1936             46.04333
## 57        1937             51.77250
## 58        1938             99.23857
## 59        1939             93.65083
## 60        1940            154.89857
## 61        1941            111.83733
## 62        1942            128.38600
## 63        1947            113.13500
## 64        1948            186.25000
## 65        1949             44.98000
## 66        1950            141.96182
## 67        1951            132.98833
## 68        1952             97.95143
## 69        1954             81.56500
## 70        1955            130.17538
## 71        1956            178.42786
## 72        1957            127.24091
## 73        1958            159.77000
## 74        1959            108.62692
## 75        1960            104.91200
## 76        1961            106.63000
## 77        1962            129.26294
## 78        1963            152.82937
## 79        1964            103.15000
## 80        1965            121.01313
## 81        1966             79.94375
## 82        1967             91.94000
## 83        1968            126.76000
## 84        1969            157.28000
## 85        1970            214.59000
## 86        1971             57.60000
## 87        1972            185.72000
## 88        1973            196.75500
## 89        1974            124.42500
## 90        1975            201.26667
## 91        1977            161.32250
## 92        1978            254.69000
## 93        1979            155.71333
## 94        1980            161.74500
## 95        1981            175.96800
## 96        1982            151.30364
## 97        1983            114.79917
## 98        1984            179.48333
## 99        1985            182.66868
## 100       1986            157.62328
## 101       1987            142.14055
## 102       1988            126.43686
## 103       1989            109.25390
## 104       1990             99.31500
## 105       1991            145.76105
## 106       1992             83.92333
## 107       1993             55.45000
```

```
## 108       1994              73.13500
## 109       1995              75.77375
## 110       1996             152.36750
## 111       1997             137.41364
## 112       1998             138.25125
## 113       1999             145.93217
## 114       2000             165.47296
## 115       2001             124.74295
## 116       2002             117.92442
## 117       2003             121.56193
## 118       2004             113.79702
## 119       2005             122.70817
## 120       2006             119.73598
## 121       2007             134.12665
## 122       2008             144.34935
## 123       2009              96.52619
## 124       2010              90.36667
```

```
print(MVYBvar)
```

```
##     Year.Built Market.Value.per.SqFt
## 1         1825                    NA
## 2         1836                    NA
## 3         1853                    NA
## 4         1860           3258.666450
## 5         1874                    NA
## 6         1875           4523.907200
## 7         1879              0.000000
## 8         1881            530.402450
## 9         1883                    NA
## 10        1890           1368.484825
## 11        1891                    NA
## 12        1892                    NA
## 13        1893            210.125000
## 14        1894            474.936200
## 15        1895           2060.178050
## 16        1896           2483.010450
## 17        1897                    NA
## 18        1898                    NA
## 19        1899            254.218900
## 20        1900           2716.073519
## 21        1901           2381.137494
## 22        1902           2614.890137
## 23        1903           7867.772743
## 24        1904           5116.211187
## 25        1905           2499.019790
## 26        1906           6565.315225
## 27        1907           6351.736060
## 28        1908           3073.581040
## 29        1909           5231.293747
## 30        1910           3950.707225
## 31        1911           3272.755964
## 32        1912           2638.376824
## 33        1913           3197.830642
```

```
## 34        1914          6420.206124
## 35        1915          1610.568502
## 36        1916          2994.437724
## 37        1917            15.456800
## 38        1918                   NA
## 39        1919                   NA
## 40        1920          3756.092403
## 41        1921          2509.008813
## 42        1922          4730.988130
## 43        1923          3883.373884
## 44        1924          3318.562269
## 45        1925          3276.497145
## 46        1926          4403.390081
## 47        1927          6225.151563
## 48        1928          4614.811816
## 49        1929          4420.556698
## 50        1930          4208.858741
## 51        1931          4071.365388
## 52        1932          6129.611907
## 53        1933           374.522950
## 54        1934                   NA
## 55        1935                   NA
## 56        1936             2.104633
## 57        1937           645.077936
## 58        1938          5686.392981
## 59        1939          5181.063045
## 60        1940          3650.905681
## 61        1941          5302.484250
## 62        1942          5352.763680
## 63        1947          8492.954450
## 64        1948          1497.139200
## 65        1949                   NA
## 66        1950         11349.892396
## 67        1951         10220.160137
## 68        1952          3400.728848
## 69        1954          2495.711250
## 70        1955          5940.094144
## 71        1956          3423.359649
## 72        1957          4052.401929
## 73        1958          3266.229900
## 74        1959          4968.902790
## 75        1960          5215.848089
## 76        1961          6241.680876
## 77        1962          3364.996822
## 78        1963          2937.430966
## 79        1964          4104.748978
## 80        1965          5268.727596
## 81        1966          1945.375284
## 82        1967          5328.842571
## 83        1968         16006.183200
## 84        1969          4120.063000
## 85        1970          2620.880000
## 86        1971                   NA
## 87        1972           526.825800
```

```
## 88        1973           3727.558550
## 89        1974           6998.628050
## 90        1975            108.480433
## 91        1977           5467.827692
## 92        1978            194.439200
## 93        1979           7735.403433
## 94        1980           7562.697310
## 95        1981           4973.035084
## 96        1982           7212.187853
## 97        1983           5247.755251
## 98        1984           6428.969261
## 99        1985           5291.172644
## 100       1986           6028.058382
## 101       1987           7639.721625
## 102       1988           5561.143562
## 103       1989           5939.763334
## 104       1990           3970.088329
## 105       1991           6200.911421
## 106       1992           5694.161267
## 107       1993            336.823600
## 108       1994            194.587710
## 109       1995           4108.392055
## 110       1996          10250.771692
## 111       1997           8088.192905
## 112       1998           7845.024158
## 113       1999           5529.888081
## 114       2000           7727.011606
## 115       2001           5202.246635
## 116       2002           4201.514013
## 117       2003           3690.549708
## 118       2004           3349.703398
## 119       2005           4368.012474
## 120       2006           3345.778459
## 121       2007           4816.248152
## 122       2008           4926.957866
## 123       2009           4305.694575
## 124       2010           1128.816133
```

```
print(MVYBsd)
```

```
##     Year.Built Market.Value.per.SqFt
## 1         1825                    NA
## 2         1836                    NA
## 3         1853                    NA
## 4         1860             57.084730
## 5         1874                    NA
## 6         1875             67.259997
## 7         1879              0.000000
## 8         1881             23.030468
## 9         1883                    NA
## 10        1890             36.993038
## 11        1891                    NA
## 12        1892                    NA
## 13        1893             14.495689
```

```
## 14        1894            21.793031
## 15        1895            45.389184
## 16        1896            49.829815
## 17        1897                   NA
## 18        1898                   NA
## 19        1899            15.944243
## 20        1900            52.115962
## 21        1901            48.796900
## 22        1902            51.135996
## 23        1903            88.700466
## 24        1904            71.527695
## 25        1905            49.990197
## 26        1906            81.026633
## 27        1907            79.697780
## 28        1908            55.439887
## 29        1909            72.327683
## 30        1910            62.854652
## 31        1911            57.208006
## 32        1912            51.365132
## 33        1913            56.549365
## 34        1914            80.126189
## 35        1915            40.131889
## 36        1916            54.721456
## 37        1917             3.931514
## 38        1918                   NA
## 39        1919                   NA
## 40        1920            61.286968
## 41        1921            50.090007
## 42        1922            68.782179
## 43        1923            62.316722
## 44        1924            57.606964
## 45        1925            57.240695
## 46        1926            66.358045
## 47        1927            78.899630
## 48        1928            67.932406
## 49        1929            66.487267
## 50        1930            64.875718
## 51        1931            63.807252
## 52        1932            78.291838
## 53        1933            19.352595
## 54        1934                   NA
## 55        1935                   NA
## 56        1936             1.450735
## 57        1937            25.398385
## 58        1938            75.408176
## 59        1939            71.979602
## 60        1940            60.422725
## 61        1941            72.818159
## 62        1942            73.162584
## 63        1947            92.157227
## 64        1948            38.692883
## 65        1949                   NA
## 66        1950           106.535874
## 67        1951           101.094808
```

```
## 68        1952             58.315768
## 69        1954             49.957094
## 70        1955             77.072006
## 71        1956             58.509483
## 72        1957             63.658479
## 73        1958             57.150940
## 74        1959             70.490445
## 75        1960             72.220829
## 76        1961             79.004309
## 77        1962             58.008593
## 78        1963             54.198072
## 79        1964             64.068315
## 80        1965             72.586001
## 81        1966             44.106409
## 82        1967             72.998922
## 83        1968            126.515545
## 84        1969             64.187717
## 85        1970             51.194531
## 86        1971                    NA
## 87        1972             22.952686
## 88        1973             61.053735
## 89        1974             83.657803
## 90        1975             10.415394
## 91        1977             73.944761
## 92        1978             13.944146
## 93        1979             87.951142
## 94        1980             86.963770
## 95        1981             70.519750
## 96        1982             84.924601
## 97        1983             72.441392
## 98        1984             80.180853
## 99        1985             72.740447
## 100       1986             77.640572
## 101       1987             87.405501
## 102       1988             74.573075
## 103       1989             77.069860
## 104       1990             63.008637
## 105       1991             78.745866
## 106       1992             75.459666
## 107       1993             18.352755
## 108       1994             13.949470
## 109       1995             64.096740
## 110       1996            101.246095
## 111       1997             89.934381
## 112       1998             88.572141
## 113       1999             74.363217
## 114       2000             87.903422
## 115       2001             72.126601
## 116       2002             64.819087
## 117       2003             60.749895
## 118       2004             57.876622
## 119       2005             66.090941
## 120       2006             57.842704
## 121       2007             69.399194
```

```
## 122       2008             70.192292
## 123       2009             65.617792
## 124       2010             33.597859
```

```r
MVNmean<-aggregate(Market.Value.per.SqFt~Neighborhood, data=housingData, FUN=mean)
MVNvar<-aggregate(Market.Value.per.SqFt~Neighborhood, data=housingData, FUN=var)
MVNsd<-aggregate(Market.Value.per.SqFt~Neighborhood, data=housingData, FUN=sd)

print(MVNmean)
```

```
##                 Neighborhood Market.Value.per.SqFt
## 1              ALPHABET CITY             148.35500
## 2       ARROCHAR-SHORE ACRES              57.75000
## 3                   ASTORIA              91.48167
## 4                 BATH BEACH              70.34000
## 5                  BAY RIDGE              68.03500
## 6                    BAYSIDE              71.42111
## 7       BEDFORD PARK/NORWOOD              38.24500
## 8         BEDFORD STUYVESANT              83.24172
## 9                    BELMONT              56.45000
## 10                BENSONHURST              71.70429
## 11               BERGEN BEACH              73.27000
## 12                BOERUM HILL              96.57600
## 13               BOROUGH PARK              64.10857
## 14                  BRIARWOOD              75.36250
## 15             BRIGHTON BEACH              81.91429
## 16              BRONX-UNKNOWN              32.06500
## 17                  BRONXDALE              28.94333
## 18          BROOKLYN HEIGHTS             114.11778
## 19             BUSH TERMINAL              60.95000
## 20                  BUSHWICK              76.13500
## 21                  CANARSIE              46.58000
## 22            CARROLL GARDENS              93.40556
## 23                   CHELSEA             215.94932
## 24                 CHINATOWN             154.17952
## 25               CITY ISLAND              40.83000
## 26              CIVIC CENTER             174.06696
## 27                   CLINTON             176.70032
## 28              CLINTON HILL              88.97385
## 29               COBBLE HILL             120.69800
## 30          COBBLE HILL-WEST              85.71125
## 31             COLLEGE POINT              65.05000
## 32              CONEY ISLAND              55.05750
## 33                    CORONA              94.20706
## 34              CROWN HEIGHTS              64.26286
## 35     DOWNTOWN-FULTON FERRY             103.26857
## 36      DOWNTOWN-FULTON MALL             132.42500
## 37        DOWNTOWN-METROTECH             122.48000
## 38             DYKER HEIGHTS              68.36000
## 39             EAST NEW YORK              36.99167
## 40              EAST TREMONT              72.33333
## 41              EAST VILLAGE             207.46115
## 42                  ELMHURST              69.80564
## 43              FAR ROCKAWAY              74.88500
```

```
## 44                   FASHION           194.81067
## 45                 FINANCIAL           199.30917
## 46           FLATBUSH-CENTRAL            65.71167
## 47   FLATBUSH-LEFFERTS GARDEN            46.27000
## 48              FLATBUSH-NORTH            54.00000
## 49                    FLATIRON           223.30311
## 50              FLUSHING-NORTH            80.16992
## 51              FLUSHING-SOUTH            89.62750
## 52       FLUSHING MEADOW PARK             58.59000
## 53                FOREST HILLS            70.20706
## 54                 FORT GREENE            81.76900
## 55                    GLENDALE            57.39667
## 56                     GOWANUS            82.45333
## 57                    GRAMERCY           188.68471
## 58                  GRANT CITY            47.60000
## 59                   GRAVESEND            75.63526
## 60                GREAT KILLS             33.74000
## 61                  GREENPOINT            86.18053
## 62  GREENWICH VILLAGE-CENTRAL           142.57767
## 63     GREENWICH VILLAGE-WEST           202.13667
## 64                GRYMES HILL             50.09000
## 65                    HAMMELS           139.07200
## 66              HARLEM-CENTRAL           102.79106
## 67                 HARLEM-EAST           139.93972
## 68               HARLEM-UPPER            79.25667
## 69                 HARLEM-WEST            95.20500
## 70  HIGHBRIDGE/MORRIS HEIGHTS             61.82000
## 71                   HILLCREST            53.95000
## 72                       HOLLIS           109.56000
## 73                HOWARD BEACH            55.06000
## 74                      INWOOD            62.05500
## 75             JACKSON HEIGHTS            47.79238
## 76                      JAMAICA           104.76600
## 77             JAMAICA ESTATES            79.69500
## 78               JAVITS CENTER           125.09000
## 79                  KENSINGTON            56.87500
## 80                KEW GARDENS             69.64300
## 81   KINGSBRIDGE HTS/UNIV HTS             23.86000
## 82    KINGSBRIDGE/JEROME PARK             58.37800
## 83                    KIPS BAY           191.31769
## 84                LITTLE ITALY           142.52308
## 85                 LITTLE NECK            65.85000
## 86            LONG ISLAND CITY           108.16667
## 87             LOWER EAST SIDE           173.56262
## 88                      MADISON            71.26000
## 89            MANHATTAN VALLEY           111.30043
## 90                      MASPETH            53.32750
## 91              MIDDLE VILLAGE             78.35857
## 92                 MIDTOWN CBD           234.36154
## 93                MIDTOWN EAST           211.04750
## 94                MIDTOWN WEST           222.06489
## 95                     MIDWOOD            79.50273
## 96         MORNINGSIDE HEIGHTS             74.63000
## 97        MORRIS PARK/VAN NEST             26.90000
```

```
## 98        MORRISANIA/LONGWOOD            44.21250
## 99     MOTT HAVEN/PORT MORRIS            30.96000
## 100              MURRAY HILL           206.26795
## 101             NEW BRIGHTON            41.47667
## 102   NEW BRIGHTON-ST. GEORGE            41.06000
## 103          NEW SPRINGVILLE            40.47000
## 104          OAKLAND GARDENS            66.94000
## 105               OCEAN HILL            37.92900
## 106      OCEAN PARKWAY-NORTH            76.51111
## 107      OCEAN PARKWAY-SOUTH            75.08000
## 108               OZONE PARK            54.10000
## 109               PARK SLOPE            88.01774
## 110         PARK SLOPE SOUTH            95.84200
## 111              PARKCHESTER            32.67500
## 112     PELHAM PARKWAY SOUTH            30.55000
## 113         PROSPECT HEIGHTS            79.16200
## 114                REGO PARK            62.13630
## 115                RIDGEWOOD            64.28667
## 116                RIVERDALE            57.10176
## 117            ROCKAWAY PARK            88.13600
## 118  SCHUYLERVILLE/PELHAM BAY            49.68000
## 119           SHEEPSHEAD BAY            79.79704
## 120              SILVER LAKE            35.80500
## 121                     SOHO           162.72473
## 122               SOUNDVIEW            43.40333
## 123         SOUTH OZONE PARK            40.78000
## 124              SOUTHBRIDGE           159.53333
## 125                SUNNYSIDE            61.61818
## 126              SUNSET PARK            80.58348
## 127              THROGS NECK            53.70667
## 128            TOMPKINSVILLE            35.81000
## 129                  TRIBECA           180.18473
## 130   UPPER EAST SIDE (59-79)           216.83715
## 131   UPPER EAST SIDE (79-96)           202.45179
## 132  UPPER EAST SIDE (96-110)           167.41600
## 133   UPPER WEST SIDE (59-79)           200.24391
## 134   UPPER WEST SIDE (79-96)           171.84515
## 135  UPPER WEST SIDE (96-116)           134.09353
## 136  WASHINGTON HEIGHTS LOWER            65.29600
## 137  WASHINGTON HEIGHTS UPPER            93.50833
## 138        WEST NEW BRIGHTON            39.69000
## 139               WHITESTONE            72.90000
## 140           WILLIAMSBRIDGE            42.46000
## 141      WILLIAMSBURG-CENTRAL            79.97017
## 142         WILLIAMSBURG-EAST            84.32605
## 143        WILLIAMSBURG-NORTH            84.10577
## 144        WILLIAMSBURG-SOUTH            82.27618
## 145           WINDSOR TERRACE            70.21200
## 146                 WOODHAVEN            38.61000
## 147                  WOODSIDE            80.52625
## 148           WYCKOFF HEIGHTS            84.93000
```

```
print(MVNvar)
```

```
##                 Neighborhood Market.Value.per.SqFt
## 1              ALPHABET CITY          1.432603e+03
## 2        ARROCHAR-SHORE ACRES                    NA
## 3                    ASTORIA          4.765799e+02
## 4                  BATH BEACH          4.725531e+02
## 5                   BAY RIDGE          2.747751e+02
## 6                     BAYSIDE          4.975028e+02
## 7         BEDFORD PARK/NORWOOD          1.786050e+00
## 8           BEDFORD STUYVESANT          1.686591e+02
## 9                     BELMONT                    NA
## 10                BENSONHURST          5.181109e+02
## 11                BERGEN BEACH                    NA
## 12                BOERUM HILL          5.717304e+02
## 13                BOROUGH PARK          4.663342e+02
## 14                   BRIARWOOD          7.877535e+02
## 15              BRIGHTON BEACH          1.340269e+03
## 16              BRONX-UNKNOWN          6.661250e+00
## 17                   BRONXDALE          1.252563e+01
## 18             BROOKLYN HEIGHTS          1.615134e+02
## 19                BUSH TERMINAL                    NA
## 20                    BUSHWICK          9.316125e+01
## 21                   CANARSIE          3.175682e+02
## 22             CARROLL GARDENS          3.974023e+02
## 23                     CHELSEA          1.402312e+03
## 24                   CHINATOWN          1.151165e+03
## 25                CITY ISLAND                    NA
## 26                CIVIC CENTER          1.351097e+03
## 27                     CLINTON          4.726897e+03
## 28                CLINTON HILL          1.530868e+03
## 29                 COBBLE HILL          3.879921e+03
## 30            COBBLE HILL-WEST          1.500731e+02
## 31               COLLEGE POINT          6.554524e+02
## 32                CONEY ISLAND          5.252351e+02
## 33                      CORONA          5.980170e+02
## 34                CROWN HEIGHTS          2.980943e+02
## 35      DOWNTOWN-FULTON FERRY          2.290867e+02
## 36       DOWNTOWN-FULTON MALL          7.916220e+02
## 37         DOWNTOWN-METROTECH          8.318775e+02
## 38               DYKER HEIGHTS                    NA
## 39               EAST NEW YORK          2.959206e+02
## 40                EAST TREMONT          2.349293e+01
## 41                EAST VILLAGE          1.170189e+03
## 42                    ELMHURST          7.810206e+02
## 43                FAR ROCKAWAY          1.150487e+03
## 44                     FASHION          3.449067e+03
## 45                   FINANCIAL          1.921830e+03
## 46            FLATBUSH-CENTRAL          6.878405e+02
## 47   FLATBUSH-LEFFERTS GARDEN                    NA
## 48              FLATBUSH-NORTH          2.764193e+02
## 49                    FLATIRON          1.061538e+03
## 50              FLUSHING-NORTH          6.733447e+02
## 51              FLUSHING-SOUTH          6.894765e+02
## 52       FLUSHING MEADOW PARK                    NA
## 53                FOREST HILLS          6.333463e+02
```

```
## 54                FORT GREENE         3.321076e+02
## 55                   GLENDALE         3.912623e+01
## 56                    GOWANUS         2.689536e+02
## 57                   GRAMERCY         1.562476e+03
## 58                 GRANT CITY         2.784800e+02
## 59                  GRAVESEND         1.678414e+02
## 60                GREAT KILLS                   NA
## 61                 GREENPOINT         7.792572e+01
## 62  GREENWICH VILLAGE-CENTRAL         1.450766e+03
## 63     GREENWICH VILLAGE-WEST         2.608190e+03
## 64                GRYMES HILL                   NA
## 65                    HAMMELS         1.267656e+04
## 66             HARLEM-CENTRAL         1.094418e+03
## 67                HARLEM-EAST         2.761448e+03
## 68               HARLEM-UPPER         3.118752e+02
## 69                HARLEM-WEST         4.500000e-04
## 70  HIGHBRIDGE/MORRIS HEIGHTS                   NA
## 71                   HILLCREST                   NA
## 72                      HOLLIS                   NA
## 73                HOWARD BEACH         6.694700e+00
## 74                      INWOOD         6.975113e+02
## 75            JACKSON HEIGHTS         4.740147e+01
## 76                     JAMAICA         3.536060e+02
## 77            JAMAICA ESTATES         1.457460e+03
## 78              JAVITS CENTER                   NA
## 79                 KENSINGTON         9.112500e-01
## 80                 KEW GARDENS         6.753258e+02
## 81   KINGSBRIDGE HTS/UNIV HTS                   NA
## 82    KINGSBRIDGE/JEROME PARK         2.349991e+02
## 83                    KIPS BAY         1.056490e+03
## 84                LITTLE ITALY         8.963593e+02
## 85                 LITTLE NECK                   NA
## 86           LONG ISLAND CITY         9.941855e+02
## 87            LOWER EAST SIDE         2.759558e+03
## 88                     MADISON         4.268941e+02
## 89           MANHATTAN VALLEY         2.697503e+03
## 90                     MASPETH         2.657642e+01
## 91              MIDDLE VILLAGE         1.217869e+03
## 92                 MIDTOWN CBD         1.418938e+03
## 93                MIDTOWN EAST         1.922460e+03
## 94                MIDTOWN WEST         1.586680e+03
## 95                     MIDWOOD         3.349567e+02
## 96         MORNINGSIDE HEIGHTS                   NA
## 97        MORRIS PARK/VAN NEST                   NA
## 98         MORRISANIA/LONGWOOD         1.801766e+02
## 99      MOTT HAVEN/PORT MORRIS                   NA
## 100               MURRAY HILL         1.686214e+03
## 101              NEW BRIGHTON         1.425333e-01
## 102   NEW BRIGHTON-ST. GEORGE         5.325120e+01
## 103           NEW SPRINGVILLE         1.439005e+01
## 104           OAKLAND GARDENS                   NA
## 105                OCEAN HILL         3.450968e+01
## 106       OCEAN PARKWAY-NORTH         6.847152e+02
## 107       OCEAN PARKWAY-SOUTH                   NA
```

```
## 108                OZONE PARK                  NA
## 109                 PARK SLOPE          3.330491e+02
## 110           PARK SLOPE SOUTH          5.562082e+01
## 111                PARKCHESTER          1.185845e+01
## 112       PELHAM PARKWAY SOUTH                  NA
## 113           PROSPECT HEIGHTS          7.666782e+02
## 114                  REGO PARK          5.097297e+02
## 115                  RIDGEWOOD          8.014532e+02
## 116                  RIVERDALE          1.593113e+02
## 117              ROCKAWAY PARK          9.844819e+02
## 118  SCHUYLERVILLE/PELHAM BAY                  NA
## 119            SHEEPSHEAD BAY          4.684936e+02
## 120               SILVER LAKE          5.000000e-05
## 121                      SOHO          2.408419e+03
## 122                 SOUNDVIEW          3.822607e+01
## 123           SOUTH OZONE PARK                  NA
## 124                SOUTHBRIDGE          1.416071e+02
## 125                  SUNNYSIDE          7.053084e+02
## 126                SUNSET PARK          2.718289e+02
## 127                THROGS NECK          2.675986e+02
## 128              TOMPKINSVILLE                  NA
## 129                    TRIBECA          2.375855e+03
## 130   UPPER EAST SIDE (59-79)          1.241306e+03
## 131   UPPER EAST SIDE (79-96)          9.892297e+02
## 132  UPPER EAST SIDE (96-110)          7.066487e+02
## 133   UPPER WEST SIDE (59-79)          2.063835e+03
## 134   UPPER WEST SIDE (79-96)          1.625411e+03
## 135  UPPER WEST SIDE (96-116)          2.660439e+03
## 136  WASHINGTON HEIGHTS LOWER          1.415999e+03
## 137  WASHINGTON HEIGHTS UPPER          8.103648e+02
## 138         WEST NEW BRIGHTON                  NA
## 139                WHITESTONE          1.243010e+03
## 140             WILLIAMSBRIDGE          1.132880e+01
## 141      WILLIAMSBURG-CENTRAL          1.297635e+02
## 142         WILLIAMSBURG-EAST          1.151922e+02
## 143        WILLIAMSBURG-NORTH          2.640817e+02
## 144        WILLIAMSBURG-SOUTH          2.925146e+02
## 145           WINDSOR TERRACE          6.247337e+01
## 146                 WOODHAVEN          2.566667e-02
## 147                  WOODSIDE          3.862802e+02
## 148           WYCKOFF HEIGHTS          1.765279e+03
```

```
print(MVNsd)
```

```
##                  Neighborhood Market.Value.per.SqFt
## 1                ALPHABET CITY          3.784975e+01
## 2         ARROCHAR-SHORE ACRES                  NA
## 3                     ASTORIA          2.183071e+01
## 4                   BATH BEACH          2.173829e+01
## 5                    BAY RIDGE          1.657634e+01
## 6                      BAYSIDE          2.230477e+01
## 7         BEDFORD PARK/NORWOOD          1.336432e+00
## 8           BEDFORD STUYVESANT          1.298688e+01
## 9                     BELMONT                  NA
```

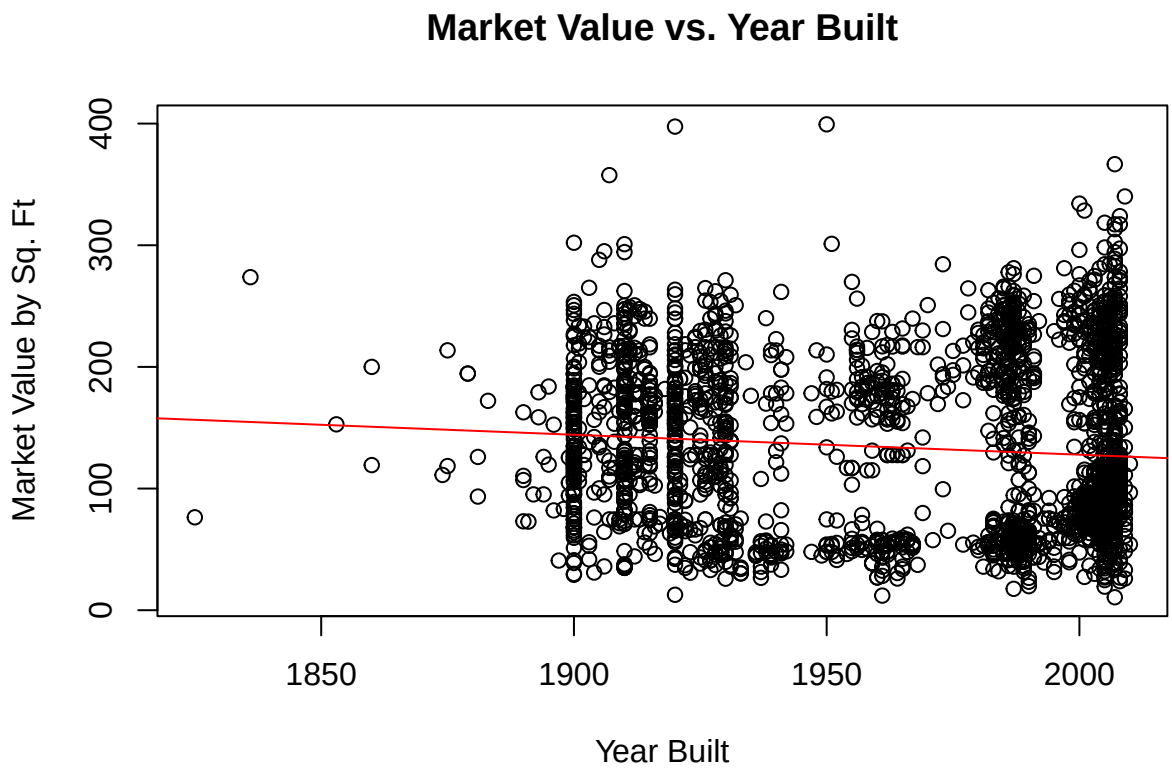
```
## 10               BENSONHURST         2.276205e+01
## 11              BERGEN BEACH                   NA
## 12               BOERUM HILL         2.391089e+01
## 13              BOROUGH PARK         2.159477e+01
## 14                 BRIARWOOD         2.806695e+01
## 15            BRIGHTON BEACH         3.660968e+01
## 16             BRONX-UNKNOWN         2.580940e+00
## 17                 BRONXDALE         3.539157e+00
## 18          BROOKLYN HEIGHTS         1.270879e+01
## 19             BUSH TERMINAL                   NA
## 20                  BUSHWICK         9.652008e+00
## 21                  CANARSIE         1.782044e+01
## 22           CARROLL GARDENS         1.993495e+01
## 23                   CHELSEA         3.744746e+01
## 24                 CHINATOWN         3.392883e+01
## 25               CITY ISLAND                   NA
## 26              CIVIC CENTER         3.675727e+01
## 27                   CLINTON         6.875243e+01
## 28              CLINTON HILL         3.912631e+01
## 29               COBBLE HILL         6.228901e+01
## 30          COBBLE HILL-WEST         1.225043e+01
## 31             COLLEGE POINT         2.560180e+01
## 32              CONEY ISLAND         2.291801e+01
## 33                    CORONA         2.445439e+01
## 34             CROWN HEIGHTS         1.726541e+01
## 35     DOWNTOWN-FULTON FERRY         1.513561e+01
## 36      DOWNTOWN-FULTON MALL         2.813578e+01
## 37        DOWNTOWN-METROTECH         2.884229e+01
## 38              DYKER HEIGHTS                   NA
## 39              EAST NEW YORK         1.720234e+01
## 40               EAST TREMONT         4.846951e+00
## 41               EAST VILLAGE         3.420803e+01
## 42                   ELMHURST         2.794675e+01
## 43               FAR ROCKAWAY         3.391882e+01
## 44                    FASHION         5.872876e+01
## 45                  FINANCIAL         4.383868e+01
## 46          FLATBUSH-CENTRAL         2.622671e+01
## 47  FLATBUSH-LEFFERTS GARDEN                   NA
## 48            FLATBUSH-NORTH         1.662586e+01
## 49                  FLATIRON         3.258126e+01
## 50            FLUSHING-NORTH         2.594889e+01
## 51            FLUSHING-SOUTH         2.625788e+01
## 52      FLUSHING MEADOW PARK                   NA
## 53               FOREST HILLS         2.516637e+01
## 54                FORT GREENE         1.822382e+01
## 55                   GLENDALE         6.255097e+00
## 56                    GOWANUS         1.639981e+01
## 57                   GRAMERCY         3.952817e+01
## 58                 GRANT CITY         1.668772e+01
## 59                  GRAVESEND         1.295536e+01
## 60                GREAT KILLS                   NA
## 61                 GREENPOINT         8.827554e+00
## 62 GREENWICH VILLAGE-CENTRAL         3.808892e+01
## 63    GREENWICH VILLAGE-WEST         5.107044e+01
```

```
## 64                GRYMES HILL                 NA
## 65                    HAMMELS        1.125902e+02
## 66             HARLEM-CENTRAL        3.308199e+01
## 67                HARLEM-EAST        5.254948e+01
## 68              HARLEM-UPPER        1.765999e+01
## 69                HARLEM-WEST        2.121320e-02
## 70  HIGHBRIDGE/MORRIS HEIGHTS                 NA
## 71                   HILLCREST                 NA
## 72                      HOLLIS                 NA
## 73               HOWARD BEACH        2.587412e+00
## 74                     INWOOD        2.641044e+01
## 75            JACKSON HEIGHTS        6.884872e+00
## 76                     JAMAICA        1.880442e+01
## 77            JAMAICA ESTATES        3.817670e+01
## 78              JAVITS CENTER                 NA
## 79                 KENSINGTON        9.545942e-01
## 80                KEW GARDENS        2.598703e+01
## 81   KINGSBRIDGE HTS/UNIV HTS                 NA
## 82    KINGSBRIDGE/JEROME PARK        1.532968e+01
## 83                    KIPS BAY        3.250369e+01
## 84                LITTLE ITALY        2.993926e+01
## 85                LITTLE NECK                 NA
## 86           LONG ISLAND CITY        3.153071e+01
## 87            LOWER EAST SIDE        5.253150e+01
## 88                     MADISON        2.066142e+01
## 89           MANHATTAN VALLEY        5.193749e+01
## 90                     MASPETH        5.155233e+00
## 91              MIDDLE VILLAGE        3.489799e+01
## 92                 MIDTOWN CBD        3.766879e+01
## 93                MIDTOWN EAST        4.384587e+01
## 94                MIDTOWN WEST        3.983315e+01
## 95                     MIDWOOD        1.830182e+01
## 96         MORNINGSIDE HEIGHTS                 NA
## 97        MORRIS PARK/VAN NEST                 NA
## 98         MORRISANIA/LONGWOOD        1.342299e+01
## 99      MOTT HAVEN/PORT MORRIS                 NA
## 100               MURRAY HILL        4.106353e+01
## 101              NEW BRIGHTON        3.775359e-01
## 102   NEW BRIGHTON-ST. GEORGE        7.297342e+00
## 103           NEW SPRINGVILLE        3.793422e+00
## 104           OAKLAND GARDENS                 NA
## 105                OCEAN HILL        5.874494e+00
## 106       OCEAN PARKWAY-NORTH        2.616706e+01
## 107       OCEAN PARKWAY-SOUTH                 NA
## 108                OZONE PARK                 NA
## 109                PARK SLOPE        1.824963e+01
## 110          PARK SLOPE SOUTH        7.457937e+00
## 111               PARKCHESTER        3.443610e+00
## 112      PELHAM PARKWAY SOUTH                 NA
## 113          PROSPECT HEIGHTS        2.768895e+01
## 114                 REGO PARK        2.257719e+01
## 115                 RIDGEWOOD        2.830995e+01
## 116                 RIVERDALE        1.262186e+01
## 117             ROCKAWAY PARK        3.137645e+01
```

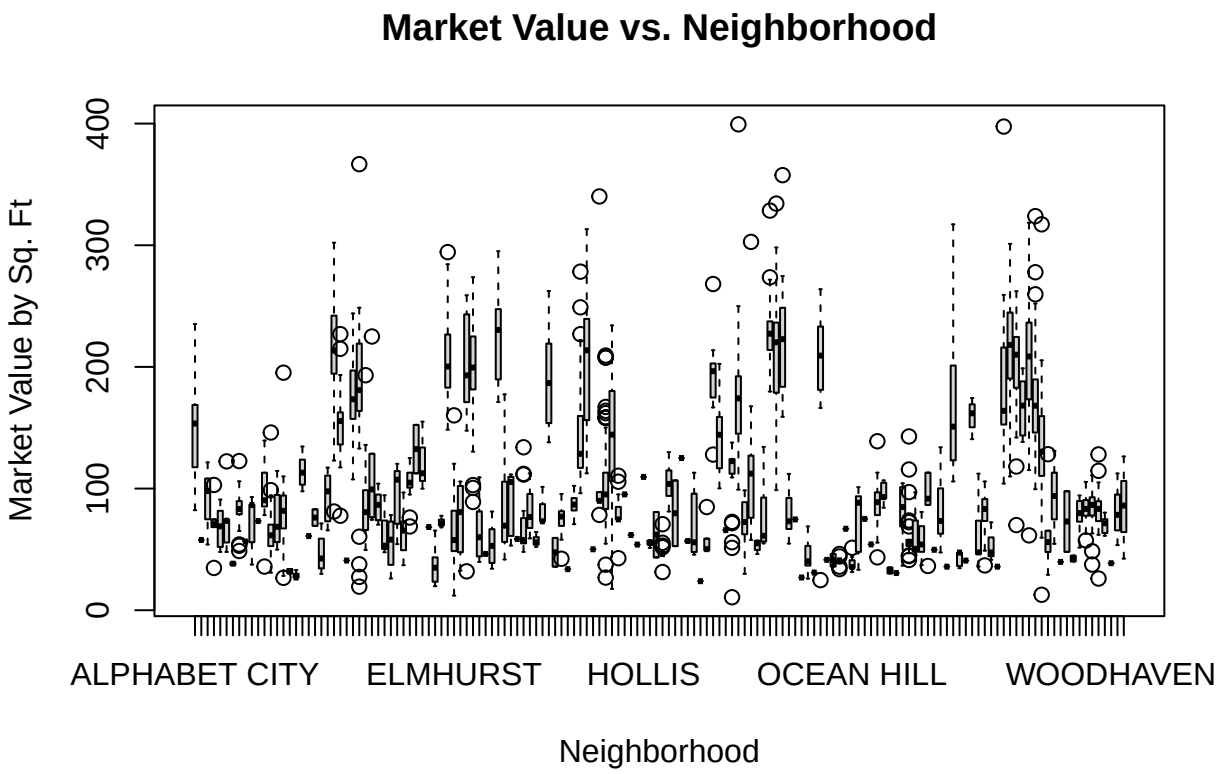
```
## 118  SCHUYLERVILLE/PELHAM BAY                  NA
## 119            SHEEPSHEAD BAY          2.164471e+01
## 120               SILVER LAKE          7.071068e-03
## 121                      SOHO          4.907565e+01
## 122                 SOUNDVIEW          6.182723e+00
## 123          SOUTH OZONE PARK                    NA
## 124               SOUTHBRIDGE          1.189988e+01
## 125                 SUNNYSIDE          2.655764e+01
## 126               SUNSET PARK          1.648724e+01
## 127               THROGS NECK          1.635844e+01
## 128             TOMPKINSVILLE                    NA
## 129                    TRIBECA          4.874274e+01
## 130   UPPER EAST SIDE (59-79)          3.523218e+01
## 131   UPPER EAST SIDE (79-96)          3.145202e+01
## 132  UPPER EAST SIDE (96-110)          2.658286e+01
## 133   UPPER WEST SIDE (59-79)          4.542946e+01
## 134   UPPER WEST SIDE (79-96)          4.031638e+01
## 135  UPPER WEST SIDE (96-116)          5.157945e+01
## 136  WASHINGTON HEIGHTS LOWER          3.762977e+01
## 137  WASHINGTON HEIGHTS UPPER          2.846691e+01
## 138         WEST NEW BRIGHTON                    NA
## 139                WHITESTONE          3.525634e+01
## 140            WILLIAMSBRIDGE          3.365828e+00
## 141      WILLIAMSBURG-CENTRAL          1.139138e+01
## 142         WILLIAMSBURG-EAST          1.073276e+01
## 143        WILLIAMSBURG-NORTH          1.625059e+01
## 144        WILLIAMSBURG-SOUTH          1.710306e+01
## 145           WINDSOR TERRACE          7.904010e+00
## 146                 WOODHAVEN          1.602082e-01
## 147                  WOODSIDE          1.965401e+01
## 148           WYCKOFF HEIGHTS          4.201522e+01
```

b. Create multiple plots to demonstrates the correlations between different variables. Remember to label all axes and give title to each graph.

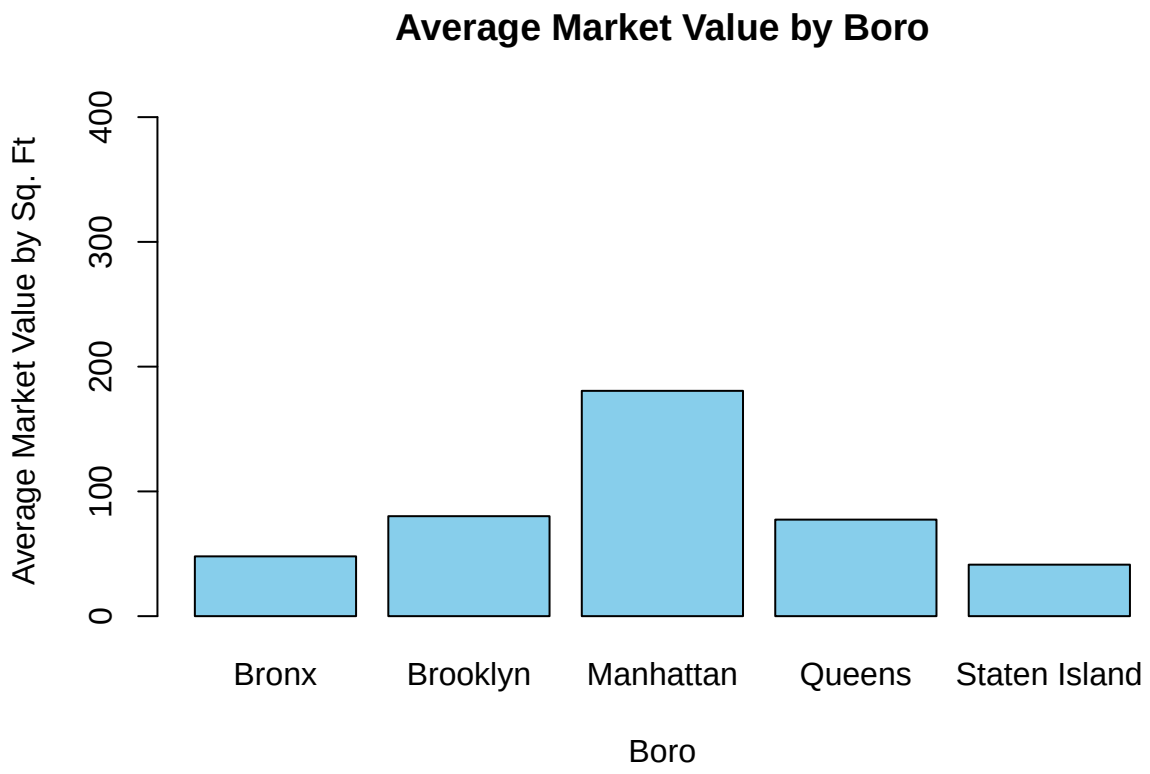
```r
# Enter your code here!
plot(housingData$Year.Built, housingData$Market.Value.per.SqFt,
     main = "Market Value vs. Year Built",
     xlab = "Year Built",
     ylab = "Market Value by Sq. Ft")

fit <- lm(Market.Value.per.SqFt ~ Year.Built, data = housingData)
abline(fit, col = "red")
```



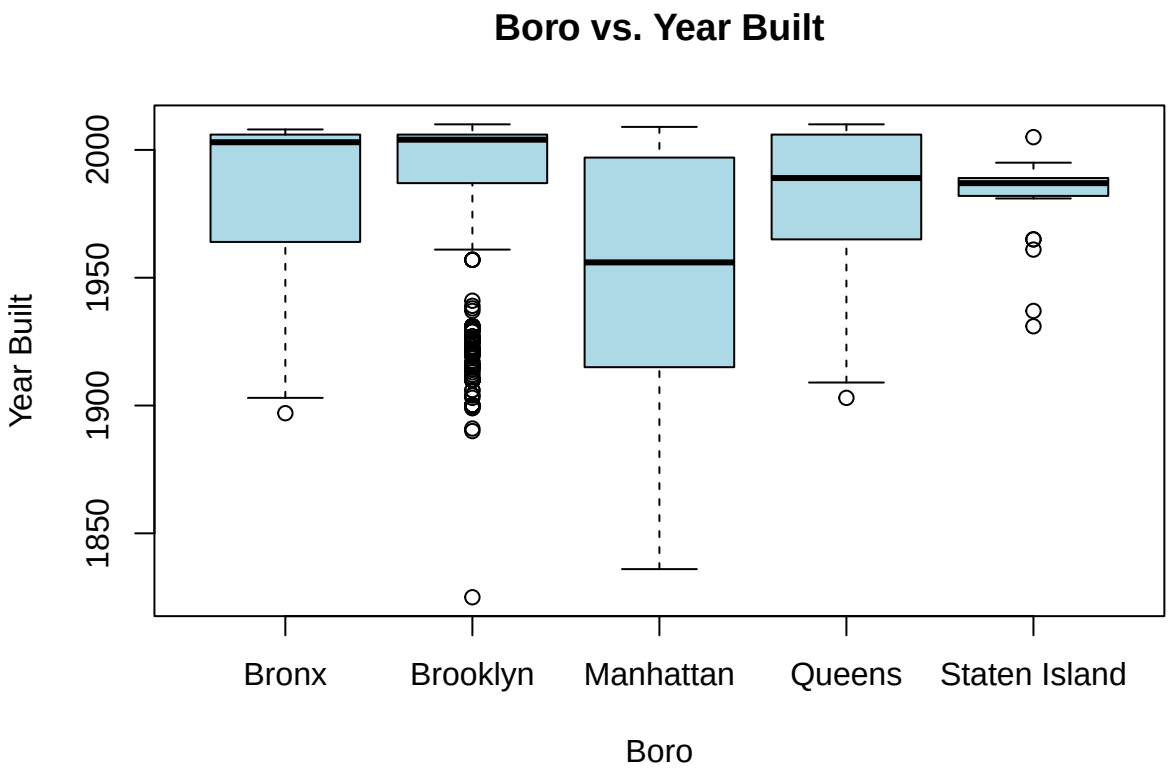

```
boxplot(Market.Value.per.SqFt ~ Neighborhood, data = housingData,
        main = "Market Value vs. Neighborhood",
        xlab = "Neighborhood",
        ylab = "Market Value by Sq. Ft")
```




```
barplot(tapply(housingData$Market.Value.per.SqFt, housingData$Boro, mean),
        main = "Average Market Value by Boro",
        xlab = "Boro",
        ylab = "Average Market Value by Sq. Ft",
        col = "skyblue",
        ylim = c(0, max(housingData$Market.Value.per.SqFt, na.rm = TRUE) + 10))
```




```r
boxplot(Year.Built ~ Boro, data = housingData,
        main = "Boro vs. Year Built",
        xlab = "Boro",
        ylab = "Year Built",
        col = "lightblue")
```



c. Write a summary about your findings from this exercise.

=> Enter your answer here!

Most of the homes were all built from 1980's to current. There aren't many buildings existing pre 1900 although there are few examples. Generally, newer houses command more market value per sq. ft despite there being many lower priced recent constructions. Manhattan commmands the highest avg. market value by a decent margin followed by Brooklyn and Queens which are very similar. The Bronx is one of the cheapest boros followed by Staten Island. The surprising thing when compared to previous rankings by boro is that Manhattan is older with its median falling around 1950. Despite it's older age, it still is high priced which may speak to the neighborhood safety, location or other factors. Staten Island sticks out as well due to a limited window of development around the 1980's which is still fairly new but is the cheapest by sq. foot. This could be due to location and inconvenience associated with this location. The Bronx and Queens were both developed around same time and also are priced similarly by sq. ft. Brooklyn has been developed most recently although houses contained in dataset also reflect examples tracing back to the 1900's.