STAT253 – Homework#<1>

Ufuk GURBUZ - <150113058>

8 Mayıs 2016

<<< PROJECT EXPLAIN >>>

Does the price of an appliance convey something about its quality? Forty-eight different dishwashers were ranked on characteristics ranging from an overall satisfaction score, washing (x1), energy use (x2), noise (x3), ease of use (x4), and cycle time (in minutes). The Bosch (SHE55M1[2]UC) had the highest performance score of 82 while the GE (GLD4408R[WW]) had the lowest at 53. Ratings pictograms were converted to numerical values for x1, ..., x4, where 5 = Excellent, 4 = Very good, 3 = Good, 2 = Fair, and 1 = Poor. Use a statistical computer package to explore the relationships between various pairs of variables in the table.

<<< DATA TABLE >>>

```
load("C:/Users/_SeriousBoy_/Desktop/Statistic Project 1/dishwashers.csv")
print(dishwashers)
```

```
##                          Model Price Score x1 x2 x3 x4 Cycle_time
## 1           Amana ADB1600AW[W]   350    61  4  4  2  3        130
## 2                    Asko D3531  1600    80  4  5  5  4        145
## 3             Asko D5233XXL[HS]  1500    56 NA  4  4  4        180
## 4                 Asko D5253XXL  1300    77  5  4  4  4        180
## 5           Bosch SHE55M1[2]UC   850    82  5  5  4  4        120
## 6           Bosch SHE6AP0[2]UC   600    75  5  4  3  4        135
## 7           Bosch SHX43P1[2]UC   800    77  5  5  3  4        115
## 8           Bosch SHX45P0[5]UC   900    79  5  5  4  4        115
## 9           Bosch SHX65P0[5]UC  1150    75  4  4  4  4        120
## 10          Bosch SHX6AP0[2]UC   700    77  5  5  4  4        110
## 11          Bosch SHX98M0[9]UC  1550    82  5  4  5  4        115
## 12       Dacor Epicure ED24[S]  1550    69  4  4  4  3        110
## 13       Electrolux Wave-Touch  1200    65  4  4  4  5        135
## 14      Frigidaire Gallery [1]   350    70  5  3  3  3        155
## 15      Frigidaire Gallery [2]   380    68  4  4  3  4        145
## 16      Frigidaire Gallery [3]   500    66  4  4  3  4        135
## 17             GE GDWF100R[WW]   600    70  5  4  3  4        120
## 18             GE GLD4408R[WW]   400    53  3  4  2  3        135
## 19             GE GLD7400R[WW]   600    62  4  4  3  5        110
## 20     GE Profile PDWT500R[WW]  1300    77  5  4  4  5        110
## 21       Hotpoint HDA3600R[WW]   300    53  3  5  1  2        115
## 22       Jenn-Air JDB3200AW[W]  1100    64  4  5  4  3        125
## 23             Kenmore 1318[2]   840    79  5  5  4  4        145
## 24             Kenmore 1324[2]   410    58  3  4  3  3        125
## 25             Kenmore 1344[2]   300    60  4  3  2  3        110
```

```
## 26              Kenmore 1348[2]   500    77  5  4  3  3         120
## 27              Kenmore 1374[2]   650    80  5  4  3  4         125
## 28              Kenmore 1389[2]   500    78  5  4  3  4         135
## 29 Kenmore Elite UltraWash [1]   780    79  5  5  4  4         140
## 30 Kenmore Elite UltraWash [2]  1100    81  5  5  4  4         145
## 31          Kenmore Pro 1317[3]  1280    79  5  5  5  4         145
## 32     KitchenAid KUDE50CV[SS]  1200    76  5  5  4  4         125
## 33     KitchenAid KUDE60FV[WH]  1340    78  5  5  4  4         135
## 34     KitchenAid KUDE70CV[SS]  1300    81  5  5  5  4         140
## 35     KitchenAid KUDS30IV[WH]   675    77  5  4  3  4         120
## 36     KitchenAid KUDS40CV[WH]   990    79  5  4  4  4         115
## 37               LG LDF6920[WW]   700    79  5  4  4  4         125
## 38        LG Steam LDF7932[ST]  1000    81  5  4  4  5         130
## 39          Maytag MDB7609AW[W]   450    75  5  4  3  2         120
## 40          Maytag MDB8959AW[W]   750    71  4  4  3  5         120
## 41          Maytag MTB4709AW[W]   400    69  5  4  2  3         125
## 42   Miele Inspira G2142SC[WH]  1150    76  5  5  4  3         145
## 43      Whirlpool DU1030XTX[Q]   350    68  5  4  2  3         130
## 44      Whirlpool DU1055XTV[Q]   400    76  5  4  3  3         125
## 45      Whirlpool DU1300XTV[Q]   420    72  5  3  3  3         140
## 46 Whirlpool Gold GU2300XTV[Q]   550    78  5  4  3  4         135
## 47 Whirlpool Gold GU2800XTV[Q]   700    77  5  5  3  4         155
## 48 Whirlpool Gold GU3600XTV[Q]   800    77  4  5  4  4         145
```

< - - QUESTION 1 - - >

Look at the variables Price, Score, and Cycle Time individually. What can you say about symmetry? About outliers?

~ The Five-Number Summary ~

- Let us find five-number summary of price datas.

```
summary(dishwashers$Price)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   300.0   487.5   725.0   814.9  1150.0  1600.0
```

- Let us find five-number summary of score datas.

```
summary(dishwashers$Score)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   53.00   68.75   76.50   72.90   79.00   82.00
```

- Let us find five-number summary of cycle time datas.

```
summary(dishwashers$Cycle_time)
```

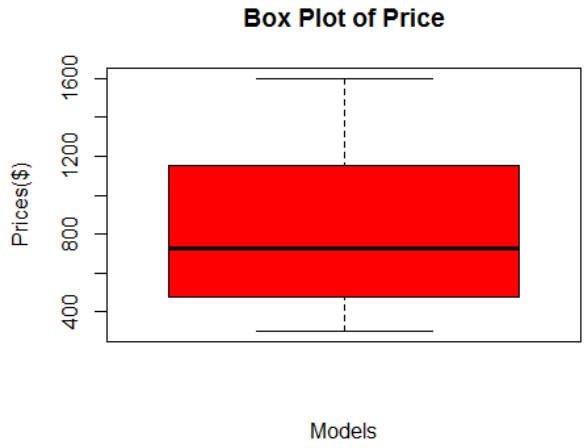
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   110.0   120.0   127.5   130.8   140.0   180.0
```

*** I examined "median" and "mean" of above data sets. The "median" and "mean" should be same for being symmetry. But; when i looked above data sets, i saw that nothing is same. So, there isn't symmetry for above data sets.

~ The Box Plots of Price, Score, Cycle Time ~

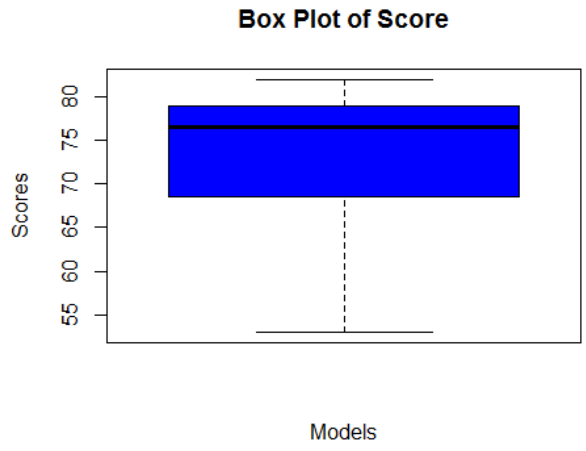
- Let us create box plot of price datas.

```
boxplot(dishwashers$Price, main="Box Plot of Price", width=250,
xlab="Models", ylab="Prices($)", col="red", outline = TRUE)
```



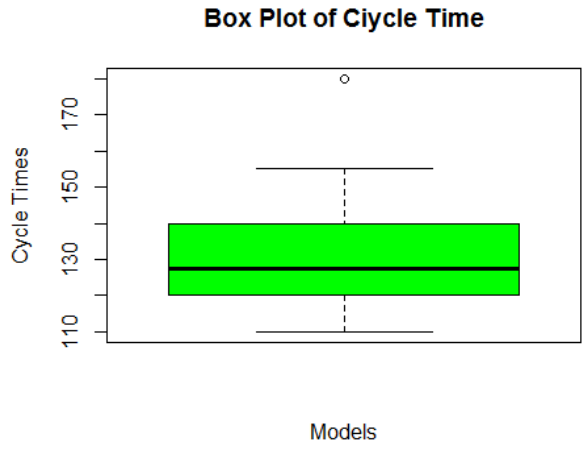
- Let us create box plot of score datas.

```
boxplot(dishwashers$Score, main="Box Plot of Score", width=250,
xlab="Models", ylab="Scores", col="blue", outline = TRUE)
```



- Let us create box plot of cycle time datas.

```
boxplot(dishwashers$Cycle_time, main="Box Plot of Ciycle Time", width=250,
xlab="Models", ylab="Cycle Times", col="green", outline = TRUE)
```



*** I examined "boxplots" of above data sets. There is no outlier in "boxplots" of 'price' and 'score' data sets. But, the "boxplot" of cycle time has a outlier.

< - - QUESTION 2 - - >

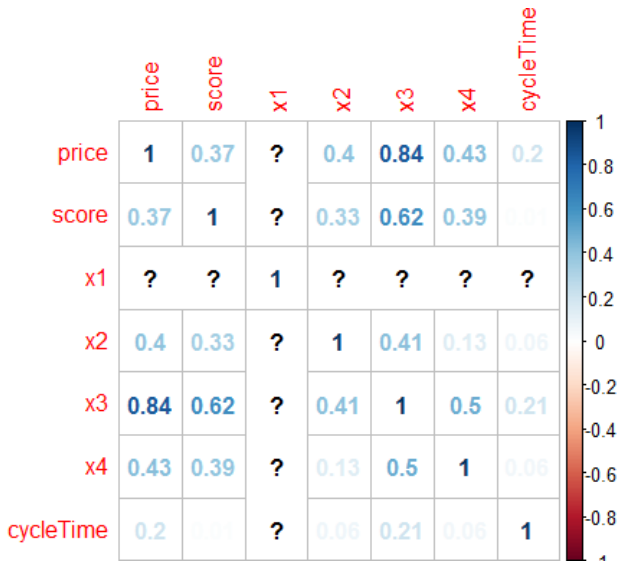
Look at all the variables in pairs. Which pairs are positively correlated? Negatively correlated? Are there any pairs that exhibit little or no correlation? Are some of these results counterintuitive?

```
#install.packages("corrplot")   # --> installing package for running
correlation plot
price <- dishwashers$Price
score <- dishwashers$Score
x1 <- dishwashers$x1
x2 <- dishwashers$x2
x3 <- dishwashers$x3
x4 <- dishwashers$x4
cycleTime <- dishwashers$Cycle_time

data <- data.frame(price, score, x1, x2, x3, x4, cycleTime)
#head.matrix(data, 48)

M <- cor(data)
round(M, 2)

##           price score x1   x2   x3   x4 cycleTime
## price      1.00  0.37 NA 0.40 0.84 0.43      0.20
## score      0.37  1.00 NA 0.33 0.62 0.39      0.01
## x1           NA    NA  1   NA   NA   NA        NA
## x2         0.40  0.33 NA 1.00 0.41 0.13      0.06
## x3         0.84  0.62 NA 0.41 1.00 0.50      0.21
## x4         0.43  0.39 NA 0.13 0.50 1.00      0.06
## cycleTime  0.20  0.01 NA 0.06 0.21 0.06      1.00
```

~ The Correlation Coefficient Plot ~

- Let us create correlation coefficient plot for all data sets.

```
library(corrplot)
corrplot(M, method="number")
```



*** I examined "values" on above data table as pairs. The correlation coefficient always is between '0' and '1' as theoretical and I saw it. There isn't 'negative' relationship between above variable pairs. But, there are 'positive'(numbers) and 'no'(NA-?) relationship. The number which has dark color shows "strong" relationship. Otherwise, the number shows "weak" relationship. There is some counterintuitive results. Example; there isn't anything logical correlation for "x1" variable or between price-score, score-x2, score-cycleTime, x3-x4, x4-cycleTime. Namely; although the score of appliance is low, the price of same appliance can be high. This is counterintuitive result. In addition to these, "x1" has correlation only with itself. Also, every variable have correlation coefficient number which is "1" with themselves.

< - - QUESTION 3 - - >

Does the price of an appliance, specifically a dishwasher, convey something about its quality? Which variables did you use in arriving at your answer?

*** The price of an appliance(dishwasher) doesn't convey information about its quality. Because, you can't make a generalization only by examining the price. You should examine necessary variable which determines price. Why is high the price, what is the cases which determine price or is it effect of brand? I used many variables in arriving at my answer. These variables are "overall satisfaction score", "washing (x1)", "energy use (x2)", "noise (x3)", "ease of use (x4)", and "cycle time. If the score and ease of use is high, the washing is quality, the energy usage, noise and cycle time is low; the appliance is so good. This quality determines price of appliance. But; if the price of a appliance is high, it isn't mean that the appliance is better than. Sometimes, we can pay money only 'brand name' of appliance. So, we should evaluate above variables and then, we should decide which we buy appliance.