First, I prepared the environment and downloaded two packages required to complete these assignments – BSDA and exactRankTests using the following commands:

```
> rm(list=ls()) #Clear the environment
> setwd("YOUR_PATH") #Set working directory for the assignment
> getwd() #Check working directory
[1] "YOUR_PATH"
> #Download BSDA package
> install.packages("BSDA")
Installing package into 'C:/Users/NataliaWeakly/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/BSDA_1.2.0.zip'
Content type 'application/zip' length 875820 bytes (855 KB)
downloaded 855 KB

package 'BSDA' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\NataliaWeakly\AppData\Local\Temp\RtmpYvMzxl\downloaded_packages
> #Download exactRankTests package
> install.packages("exactRankTests")
Installing package into 'C:/Users/NataliaWeakly/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/exactRankTests_0.8-29.zip'
Content type 'application/zip' length 166456 bytes (162 KB)
downloaded 162 KB

package 'exactRankTests' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Natalia\AppData\Local\Temp\RtmpYvMzxl\downloaded_packages
```

Part 1. "Souperb" restaurant ratings

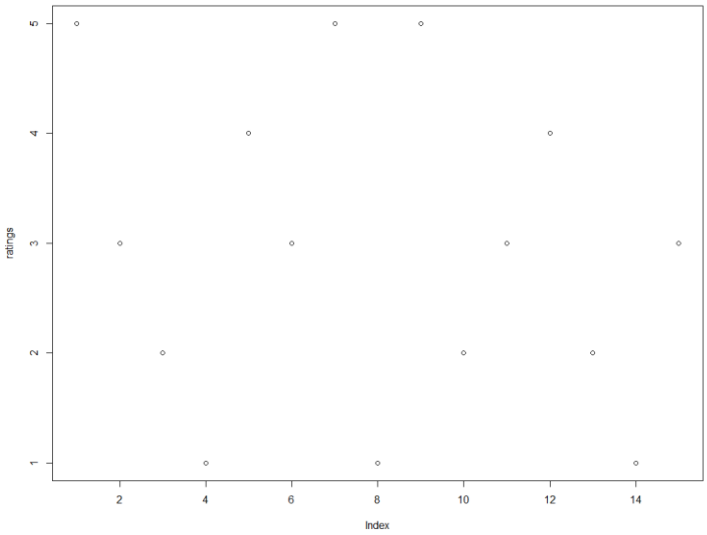
Then, I included the BSDA package and loaded the data to be analyzed in the first part of the assignment using the following code:

```
> #######Part 1. "Souperb" restaurant chain
>
> #Include BSDA package (for SIGN.test function)
> library(BSDA)
>
> #load ratings data
> ratings <- c(5,3,2,1,4,3,5,1,5,2,3,4,2,1,3)
```

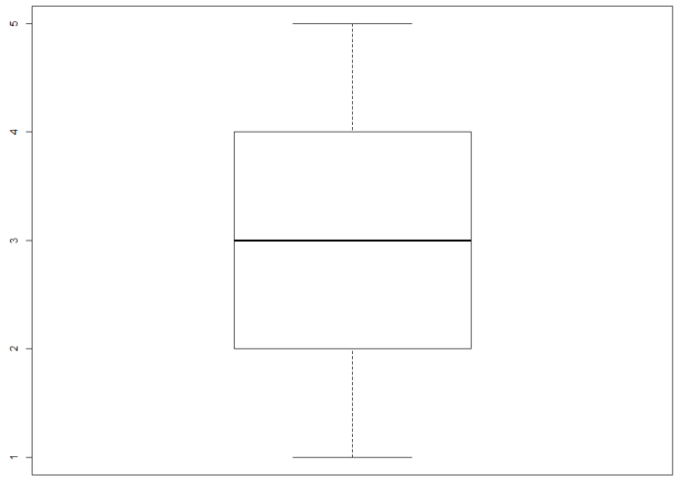
```
>#######Part 1. "Souperb" restaurant chain
>
> #Include BSDA package (for SIGN.test function)
> library(BSDA)
>
> #load ratings data
> ratings <- c(5,3,2,1,4,3,5,1,5,2,3,4,2,1,3)
```

To familiarize myself with the data, I performed limited EDA by visually inspecting the following graphs:

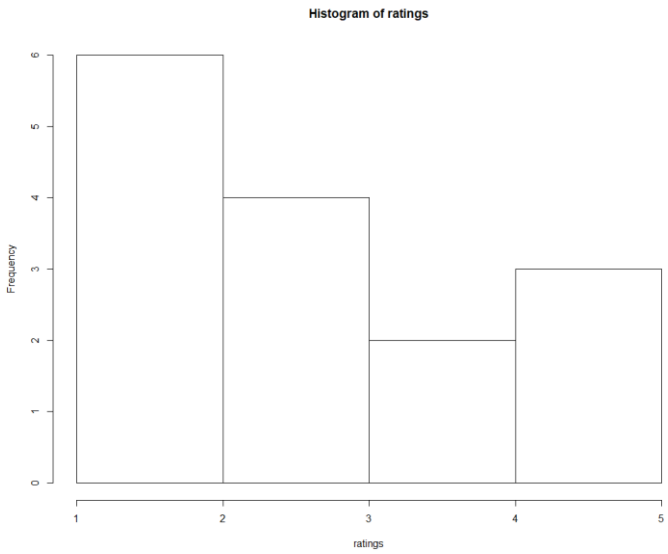
```
#EDA
> plot(ratings)
```




```
> boxplot(ratings)
```




```
> hist(ratings)
```



Overall, the restaurant ratings represent one sample ordinal data that does not follow any particular distribution. In addition, as the histogram shows, the distribution is not symmetric. So, we cannot use either parametric testing, nor Wilcoxon signed-rank test as the latter requires the distribution to be symmetric.

So, I proceeded with the Sign Test.

My hypothesis:

Ho: The true median rating is equal to 3.

Ha: The true median rating is less than 3.

I performed the test using the following commands:

```
> #Perform Sign Test (median = 3)
> SIGN.test(ratings, md=3, alternative = "less")

         One-sample Sign-Test

data:  ratings
s = 5, p-value = 0.5
alternative hypothesis: true median is less than 3
95 percent confidence interval:
 -Inf    4
sample estimates:
median of x
          3

Achieved and Interpolated Confidence Intervals:

                  Conf.Level L.E.pt U.E.pt
Lower Achieved CI     0.9408   -Inf      4
Interpolated CI       0.9500   -Inf      4
Upper Achieved CI     0.9824   -Inf      4
```

```
> #Perform Sign Test (median = 3)
> SIGN.test(ratings, md=3, alternative = "less")

        One-sample Sign-Test

data:  ratings
s = 5, p-value = 0.5
alternative hypothesis: true median is less than 3
95 percent confidence interval:
 -Inf    4
sample estimates:
median of x
          3

Achieved and Interpolated Confidence Intervals:

                  Conf.Level L.E.pt U.E.pt
Lower Achieved CI     0.9408   -Inf      4
Interpolated CI       0.9500   -Inf      4
Upper Achieved CI     0.9824   -Inf      4

  .
```

Conclusions:

The test returned the likelihood value (p-value) for the default median used (3) with the alternative hypothesis being median is less than 3 as p-value= 0.5. This p-value is substantially higher than the significance level of 0.05. It means that the test does not provide enough evidence to reject the null hypothesis (mean =3) in favor of the alternative hypothesis (true mean being less than 3). So, the “Souperb” restaurant did in fact did receive the median rating of at least 3.

Part 2. Comparing ratings for two operating systems - “M” and “W”

In order to compare the ratings, I loaded the exactRankTests package and the rating data for both operating systems using the following code:

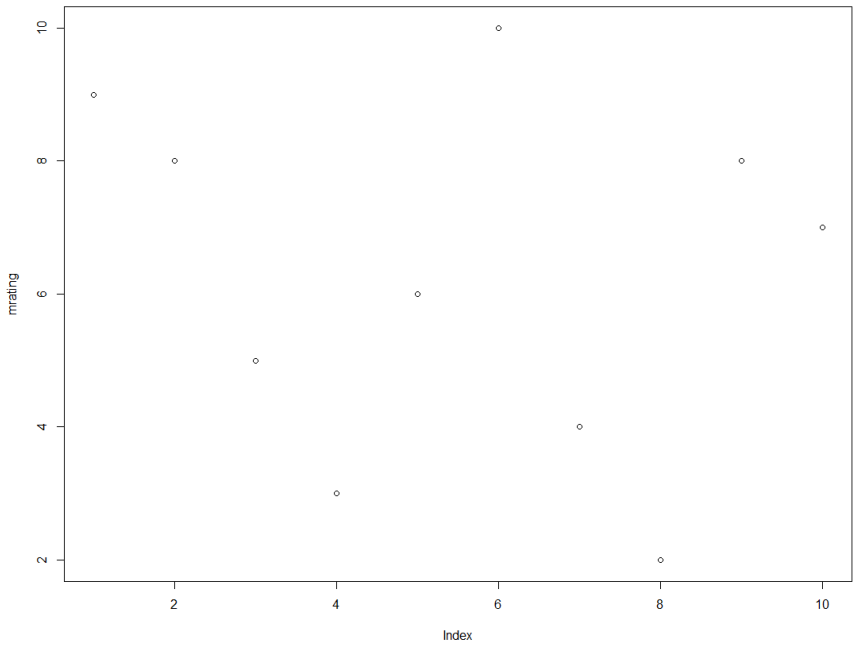
```
> #######Part 2. Operating systems "M" and "W"
>
> #Include exactRankTests package
> library(exactRankTests)
>
> #Load ratings data
> mrating <- c(9,8,5,3,6,10,4,2,8,7) #Ratings for "M" operating system
> wrating <- c(7,6,8,2,9,5,1,4,7,10) #Ratings for "W" operating system
```

```
> #######Part 2. Operating systems "M" and "W"
>
> #Include exacdtRankTests package
> library(exactRankTests)
>
> #Load ratings data
> mrating <- c(9,8,5,3,6,10,4,2,8,7) #Ratings for "M" operating system
> wrating <- c(7,6,8,2,9,5,1,4,7,10) #Ratings for "W" operating system
>
```

Then, visually inspected the data using the following graphs:

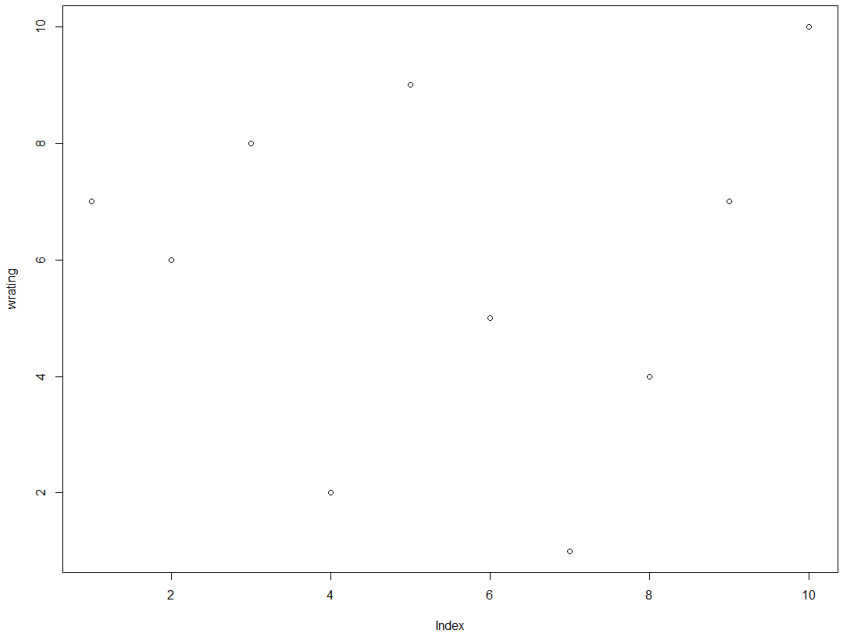
```
> #EDA
> plot(mrating)
```

“M” operating system ratings:

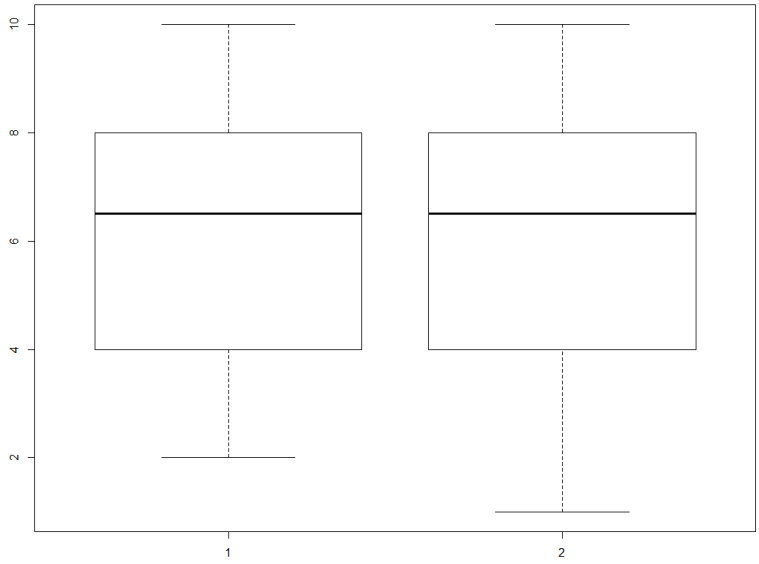



```
> plot (wrating)
```

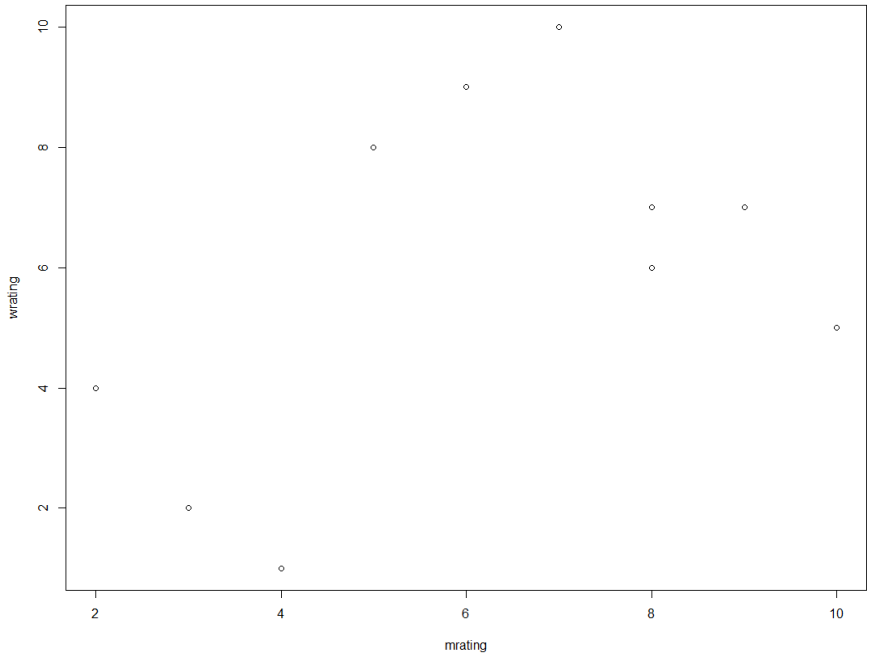
“W” operating system ratings:



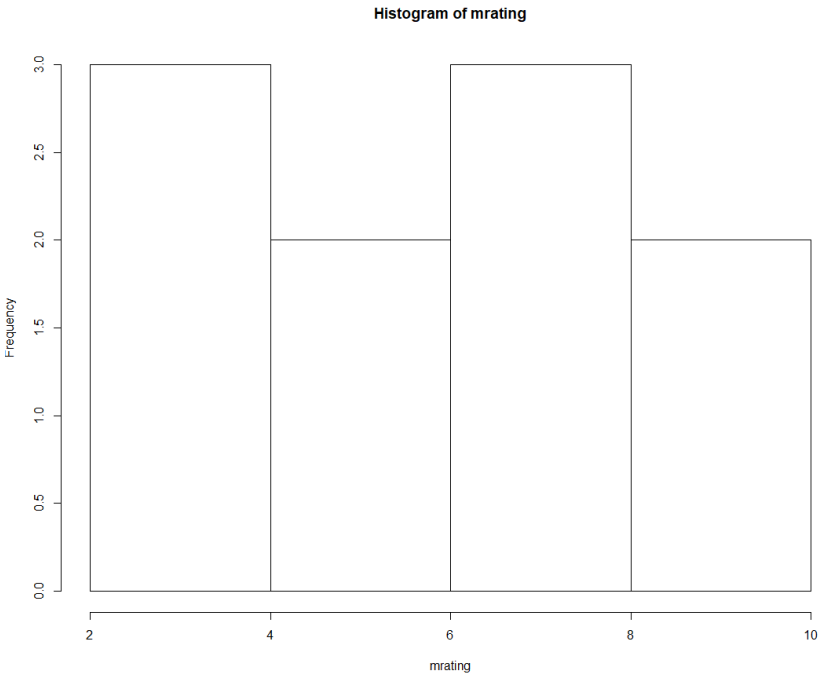

```
> boxplot(mrating, wrating)
```



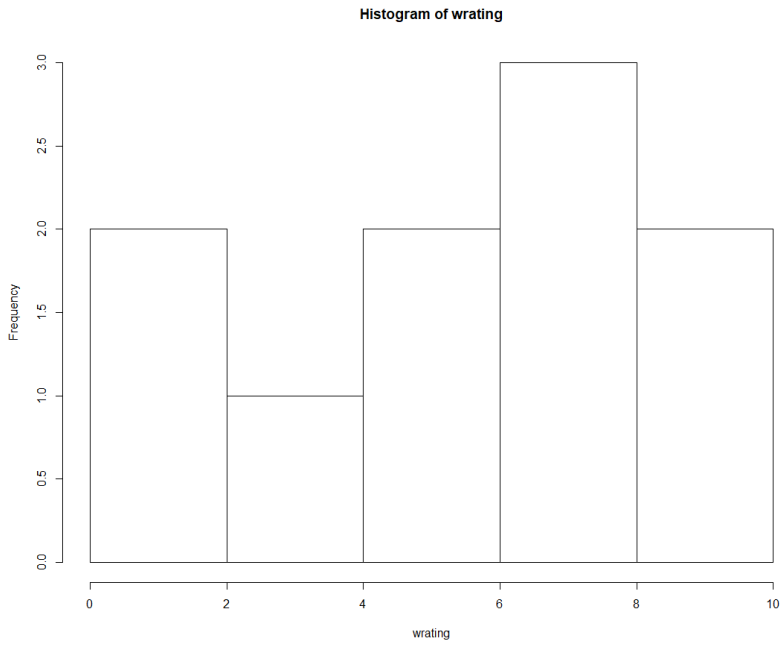

```
> plot(mrating, wrating)
```




```
> hist (mrating)
```




```
> hist(wrating)
```



This limited analysis did not yield any significant trends or abnormal findings. Neither of the ratings follow any particular distributions, and neither of them is symmetrical. Since we have to use a non-parametric test and, according to the assignments, the data is assumed to be paired, I chose Wilcoxon Rank-sum test for two samples of possible paired observations.

The hypothesis for this test:

Ho: Ratings for the operating systems M and W have the same distribution

Ha: Ratings for the operating systems M and W do not have the same distribution

Since this test is used on two paired samples, it tests that the null hypothesis in that the medians mratings – wratings equals mu.

I used the following commands:

```
> #Perform Wilcoxon Rank-Sum Test for a difference in mrating and wrating
> #since ratings are assumed to be paired
> wilcox.exact(mrating, wrating, paired = TRUE)

        Exact Wilcoxon signed rank test

data:  mrating and wrating
V = 28.5, p-value = 0.9414
alternative hypothesis: true mu is not equal to 0
```

```
> #Perform Wilcoxon Rank-Sum Test for a difference in mrating and wrating
> #since ratings are assumed to be paired
> wilcox.exact(mrating, wrating, paired = TRUE)

        Exact Wilcoxon signed rank test

data:  mrating and wrating
V = 28.5, p-value = 0.9414
alternative hypothesis: true mu is not equal to 0
```

Conclusion:

The above test results show the p-value of 0.9414 which is considerably higher than the statistical significance level of 0.05. It means that we do not have any evidence against the null hypothesis stating that the ratings for the two operating systems have the same distribution.