**Exercise 3**

In the previous exercise (2) we calculated the artificial columns {Clarity, Politeness, Satisfaction} for each entry in the data.

In the current exercise we filtered the data, leaving in only subjects with the ages of 18-49. We also filtered out the NA rows from the data.

In order to work with R for statistical analysis, we mapped the Hebrew in the given dataset to English identifiers. The relevant mappings to this exercise are as follows:

Comp\_Use\_Know:

|  |  |
| --- | --- |
| **ID** | **ORIGINAL VALUE** |
| F1 | ידע בינוני (למשל, מסוגל להתקין תוכנות בעצמי) |
| F2 | ידע בסיסי (למשל, יודע לגלוש באינטרנט ובפייסבוק) |
| F3 | ידע רב (משתמש בהרבה תוכנות שונות ויכול לפתור בעצמי הרבה בעיות במחשב) |
| F4 | מומחה (מבין בהרבה תחומים ומסוגל לאבחן ולפתור כמעט כל בעיה) |

Sex:

|  |  |
| --- | --- |
| **ID** | **ORIGINAL VALUE** |
| C1 | זכר |
| C2 | נקבה |

System:

|  |  |
| --- | --- |
| **ID** | **ORIGINAL VALUE** |
| S | מערכת תוכנה רגילה |
| C | תקשורת מתווכת מחשב |

**Part 1**:

For this section we split the data in two ways: (1) By System {S, C} (2) By Sex {C1, C2} Then we performed two unpaired t-tests for comparing two groups by their Politeness as follows:

We perform variance equality F test between the groups system S, system C and conclude that they have equal variance:

Test 1 for comparing the groups (1):

We performed unpaired t-test with equal variance and got the following results:

*t = 3.8093, df = 275 , p-value = 0.0001719*

As we can see from the results we reject which means that the factor System has a significant impact on Politeness.

We perform variance equality F test between the groups males, females and conclude that they have equal variance:

Test 2 for comparing the groups (2):

We performed unpaired t-test with equal variance and got the following results:

*t = -0.2531, df = 275, p-value = 0.8004*

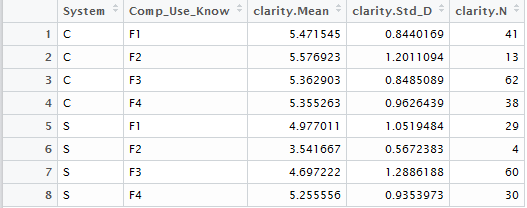
As we can see from the results we cannot reject which means that the factor Sex has no significant impact on Politeness.

**Part 2:**

In this section we conducted a two-way ANOVA with the input:

We have groups according to the possible combinations of the two factors: .

Descriptive statistics:



In order to test if the population variances are equal we performed Levene's test and got the following results:

Df Sum Squares Mean Square F-statistic p-value

Between Groups 7 6.78302 0.969 2.74904 0.00898

Within Groups 269 94.81927 0.35249

Total 276 101.60229

P-value is less than 0.05 therefore we reject the null hypothesis where we have an equal variance and conclude that the variances are not equal.

Giving this situation we used a significance level of 0.01 instead of 0.05 in the two-way ANOVA. (Theoretically we should be performing here an a-parametric test like B-F and Welch ANOVA fixes).

We got the following results from the two-way ANOVA:

Df Sum Sq Mean Sq F value Pr(>F)

System 1 20.4 20.402 19.649 1.36e-05 \*\*\*

Comp\_Use\_Know 3 3.8 1.265 1.219 0.3033

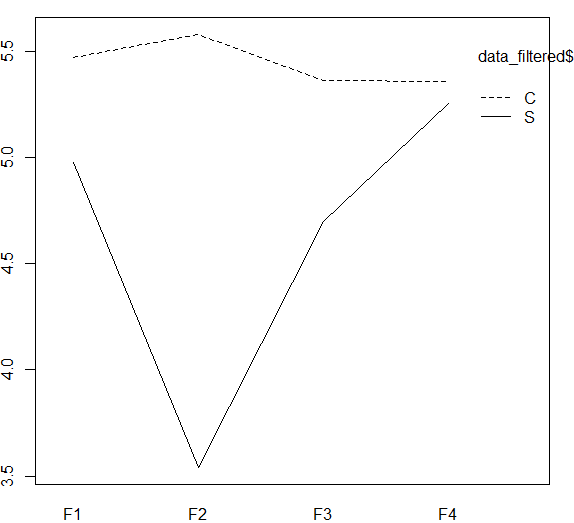
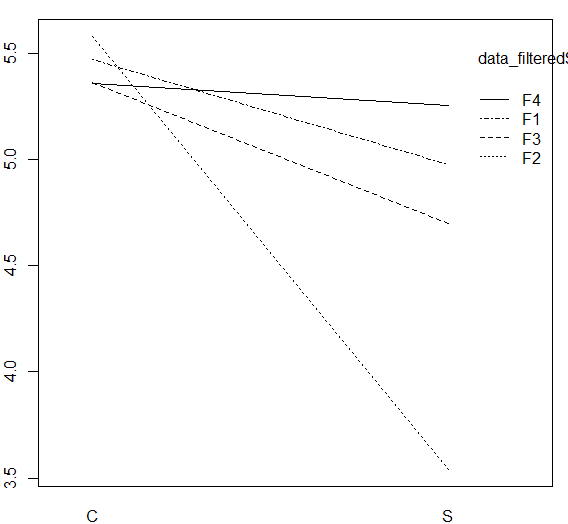
System:Comp\_Use\_Know 3 10.6 3.535 3.404 0.0182 \*

Residuals 269 279.3 1.038

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

As it can be seen from the results, the factor System is strongly significant which means that its impact is very high on Clarity.



**System**

**Comp\_Use\_Know**

**System**

**Comp\_Use\_Know**

**Clarity**

**Clarity**

We cannot conclude that there is an interaction between System and Comp\_Use\_Know because we use a significant level of 0.01.

According to the interaction plots we see some possible interaction between System and Comp\_Use\_Know. However, the results from the two-way ANOVA showed that this is not sufficient for the significance level of 0.01.

Because the p-value of the interaction variable is close to 0.01 we decided to do an interaction effects analysis using the following 6 tests:

1. One-way ANOVA of Clarity~Comp\_Use\_Know where System = "S".

* Levene's test:

Df Sum Squares Mean Square F-statistic p-value

Between Groups 3 3.19109 1.0637 2.36255 0.07471

Within Groups 119 53.57765 0.45023

Total 122 56.76874

* One-way ANOVA with equal population variances (P-value = 0.05):

Df Sum Sq Mean Sq F value Pr(>F)

Comp\_Use\_Know 3 13.63 4.544 3.482 0.0181 \*

Residuals 119 155.30 1.305

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

1. One-way ANOVA of Clarity~Comp\_Use\_Know where System = "C".

* Levene's test:

Df Sum Squares Mean Square F-statistic p-value

Between Groups 3 1.52606 0.50869 1.85014 0.14055

Within Groups 150 41.24162 0.27494

Total 153 42.76768

* One-way ANOVA with equal population variances (P-value = 0.05):

Df Sum Sq Mean Sq F value Pr(>F)

Comp\_Use\_Know 3 0.77 0.2561 0.31 0.818

Residuals 150 124.01 0.8267

1. T-test of Clarity~System where Comp\_Use\_Know = "F1".

* Equal Variances F test:

F = 0.64374, num df = 40, denom df = 28, p-value = 0.1983

95 percent confidence interval:

0.3143798 1.2627297

* T-test with equal variances:

t = 2.1793, df = 68, p-value = 0.03278

95 percent confidence interval:

0.04170721 0.94735923

1. T-test of Clarity~System where Comp\_Use\_Know = "F2".

* Equal Variances F test:

F = 4.4837, num df = 12, denom df = 3, p-value = 0.2429

95 percent confidence interval:

0.3127443 20.0607888

* T-test with equal variances:

t = 3.2247, df = 15, p-value = 0.00567

95 percent confidence interval:

0.6899923 3.3805206

1. T-test of Clarity~System where Comp\_Use\_Know = "F3".

* Equal Variances F test:

F = 0.43357, num df = 61, denom df = 59, p-value = 0.001441

95 percent confidence interval:

0.2598205 0.7219369

* T-test with unequal variances:

t = 3.3584, df = 101.6, p-value = 0.001104

95 percent confidence interval:

0.2725097 1.0588523

1. T-test of Clarity~System where Comp\_Use\_Know = "F4".

* Equal Variances F test:

F = 1.0591, num df = 37, denom df = 29, p-value = 0.8823

95 percent confidence interval:

0.518248 2.100629

* T-test with equal variances:

t = 0.42939, df = 66, p-value = 0.669

95 percent confidence interval:

-0.3639100 0.5633252

The results we got from the 6 tests match the interaction plots.

From the first two One-way ANOVA tests we conclude the following:

When System = "S" the value of Comp\_Use\_Know effect Clarity significantly with p-value = 0.0181. This result can be seen visually on the left interaction plot where the values of Clarity for each value of Comp\_Use\_Know are highly distinct. When System = "C" those values are more close to each other which means that Comp\_Use\_Know has less effect power on Clarity.

From the later four t-tests we conclude the following:

When Comp\_Use\_Know is F1, F2, F3 the t-test results show that p-value is significant which can be seen on the right interaction plot as the difference between Clarity values of System = "S" and System = "C".

When Comp\_Use\_Know = F2 the differences of Clarity values are the largest. this makes up for the relatively small size of the data of Comp\_Use\_Know = F2 according to the descriptive statistics.

When Comp\_Use\_Know = F4 the t-test results are not significant which can be seen on the plot as the smallest difference on Clarity between the two systems.

Finally, we performed three post-hoc Scheffe tests in order to find out if there are differences between each two values of the factor Comp\_Use\_Know as follows:

clarity std r Min Max

F1 5.266667 0.9603240 70 2.166667 7

F2 5.098039 1.3907667 17 2.833333 7

F3 5.035519 1.1332745 122 2.000000 7

F4 5.311275 0.9449633 68 2.666667 7

alpha: 0.05 ; Df Error: 269

Critical Value of F: 2.638161

Means with the same letter are not significantly different.

Groups, Treatments and means

a F4 5.311

a F1 5.267

a F2 5.098

a F3 5.036

clarity std r Min Max

F1 4.977011 1.0519484 29 2.166667 6.166667

F2 3.541667 0.5672383 4 2.833333 4.000000

F3 4.697222 1.2886188 60 2.000000 7.000000

F4 5.255556 0.9353973 30 2.666667 7.000000

alpha: 0.05 ; Df Error: 119

Critical Value of F: 2.680811

Means with the same letter are not significantly different.

Groups, Treatments and means

a F4 5.256

a F1 4.977

a F3 4.697

a F2 3.542

clarity std r Min Max

F1 5.471545 0.8440169 41 3.333333 7

F2 5.576923 1.2011094 13 3.833333 7

F3 5.362903 0.8485089 62 3.000000 7

F4 5.355263 0.9626439 38 3.333333 7

alpha: 0.05 ; Df Error: 150

Critical Value of F: 2.664907

Groups, Treatments and means

a F2 5.577

a F1 5.472

a F3 5.363

a F4 5.355

As we can see there are no significant differences between the factor values according to because each value ended up in the same group a.

**Conclusions:**

From part 1 we conclude that the factor System has a significant impact on politeness. From part 2 using the interaction plots and the two-way ANOVA results we conclude that C systems are significantly more clear than S systems. We also found that there is some interaction between the factors System and Comp\_Use\_Know with p-value close to 0.01. Because we used a significance level of 0.01 we could not conclude statistically that there is an interaction. However, because the p-value was close to 0.01 we decided to perform the interaction analysis and got the following conclusions:

When the system is C there are no differences among the different knowledge levels of computer use. When the system is S there is a large variance between the knowledge levels where the order from the highest to the lowest on the level of clarity is F4 (Expert), F1 (Medium), F3 (High), F2 (Basic) which is a bit surprising because we expect this order to be F1, F3, F1, F2. The level of clarity is not changed over the different systems for F4(Expert) which is reasonable. The post-hoc test we performed at the end didn't reveal any interaction between the factors values.