Hypotheses

1. **H0 :** The number of students using dating apps are only 30% of the population. (YouGov, 2017) (Q5)

**HA :** The number of students using dating apps are less than 30% of the population. (this one is left-tailed is it? Need to change to 2 tailed?)

1. **H0 :** There is no significant difference between usage of dating   
    apps between men and women among undergraduate students in   
    UM.

**HA :** There is significant difference between usage of dating apps between men and women among undergraduate students in UM.

1. **H0 :** There is no significant difference between heterosexuals and   
    non heterosexuals in using dating apps.

**HA :** There is significant difference between heterosexuals and non heterosexuals in using dating apps.

1. **H0 :** There is no significant difference between gender on their   
    opinion towards the effectiveness of dating apps.

**HA :** There is significant difference between gender on their opinion   
 towards the effectiveness of dating apps.

1. **H0 :** There is no significant difference between gender on the   
    safeness of dating apps.

**HA :** There is significant difference between gender on the safeness   
 of dating apps.

**Hypotheses created**

1. **Hypothesis 1**

**H0 :** The number of students using dating apps are equal to 29% of the population. (YouGov, 2017) (Q5)

**HA :** The number of students using dating apps are less than 29% of the population.

1. **Hypothesis 2**

**H0 :** There is no significant difference on the usage of dating apps between gender among undergraduate students in UM.

**HA :** There is significant difference on usage of dating apps between gender among undergraduate students in UM.

1. **Hypothesis 3**

**H0 :** There is no significant difference between heterosexuals and non-heterosexuals in using dating apps.

**HA :** There is significant difference between heterosexuals and non-heterosexuals in using dating apps.

1. **Hypothesis 4**

**H0 :** There is no significant difference between gender on their opinion towards the effectiveness of dating apps.

**HA :** There is significant difference between gender on their opinion towards the effectiveness of dating apps.

1. **Hypothesis 5**

**H0 :** There is no significant difference between gender on the safeness of dating apps.

**HA :** There is significant difference between gender on the safeness of dating apps.

**Assumptions of the statistical tests used**

**Assumptions of Chi-Square Independence test**

* The two variables should be measured at ordinal or nominal level (categorical data).
* The two variables should consist of two or more categorical, independent groups.
* Expected frequencies of more than 5 for each group in each variable.

<https://www.spss-tutorials.com/spss-chi-square-independence-test/#assumptions>

<https://statistics.laerd.com/spss-tutorials/chi-square-test-for-association-using-spss-statistics.php>

**Assumptions made for Sign test**

* The dependent variable should be measured on a continuous(interval or ratio) or ordinal level.
* The independent variable should consist of two categorical, “related group” or “matched pairs”.
* The paired observations for each participant need to be independent.
* The difference scores (differences between the paired observations) are from a continuous distribution.

<https://statistics.laerd.com/spss-tutorials/sign-test-using-spss-statistics.php>

**Assumptions made for Wilcoxon Signed Rank test**

* Dependent variable should be measured at the ordinal or continuous level.
* Independent variable should consist of two categorical groups which are related or matched pairs.
* Data distribution of the two related groups should be symmetrical.
* (if data distribution is not similar, run Sign test)

<https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php>

**Assumptions made for Mann-Whitney U test**

* Dependent variable should be measured at the ordinal or continuous level.
* Independent variable should consist of two categorical, independent groups
* Independence of observations, there is no relationship between the observations in each group or between the group themselves.
* Data distribution for both groups of the independent variable should have same or similar shape.

**Assumptions made for Kruskal-Wallis H test**

* Dependent variable should be measured at the ordinal or continuous level.
* Independent variable should consist of two or more categorical, independent groups.
* Independence of observations, there is no relationship between the observations in each group or between the group themselves.
* Data distribution for both groups of the independent variable should have same or similar shape.

**Result and discussion for each hypothesis**

**Result for Hypothesis 1**

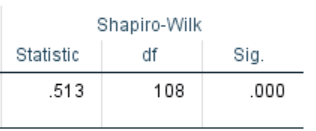
****

Figure .2 Normality test

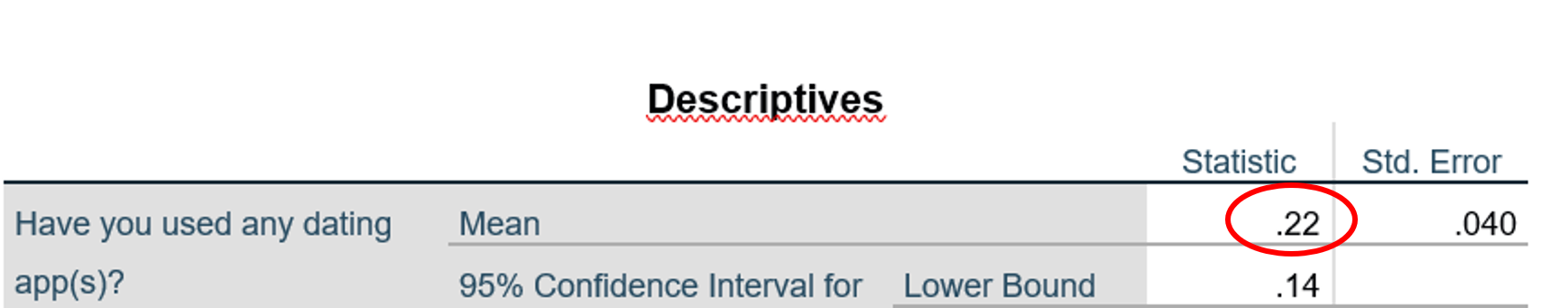


Figure 1.1 Mean from descriptive statistic

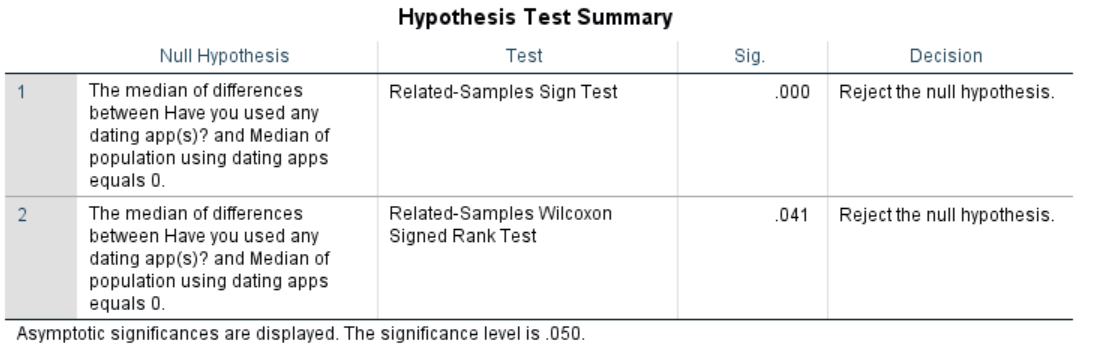


Figure .3 Hypothesis Test Summary

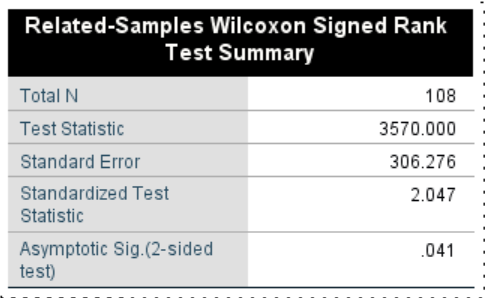
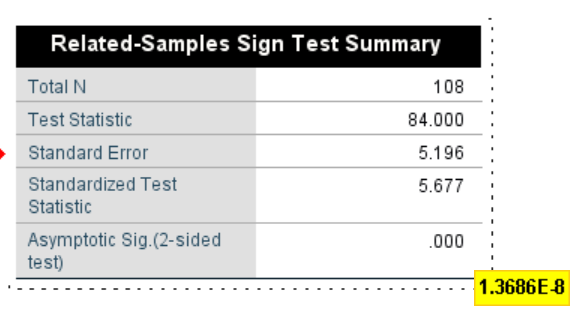


Figure 1.4 One-sample Sign test Summary Figure 1.5 One-sample Wilcoxon Signed Rank Test Summary

Normality test(Shapiro-Wilk test): p-value(2.7985E-17) is lower than 0.05, the data significantly deviate from a normal distribution.

One-Sample Wilcoxon Signed Rank Test: The standardized test value(z-vlaue) = 2.047. Asymptotic significance value(p-value) = 0.041 (less than 0.05).

Sign Test (One Sample Sign Test): The standardized test value(z-vlaue) = 5.677. Asymptotic significance value(p-value) = 0.000 (1.3686E-8, less than 0.05).

The p-value for both tests are less than the α value(0.05). Thus, the null hypothesis is rejected. The number of students using dating apps are not equal to 29% of the population.

From the descriptive statistic, the mean shows that only 22% of the students used dating apps. Although it is not statistical, the alternative hypothesis of less than 29% of students using dating apps is accepted.

**Result for Hypothesis 2**

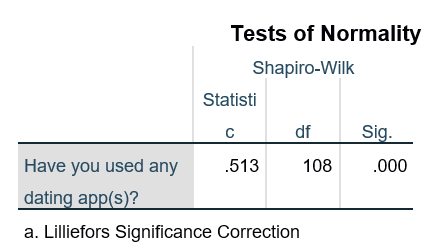
****

Figure 2.1 Normality test

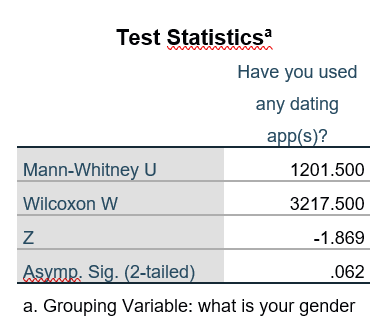


Figure 2.3 Mann-Whitney U test statistics

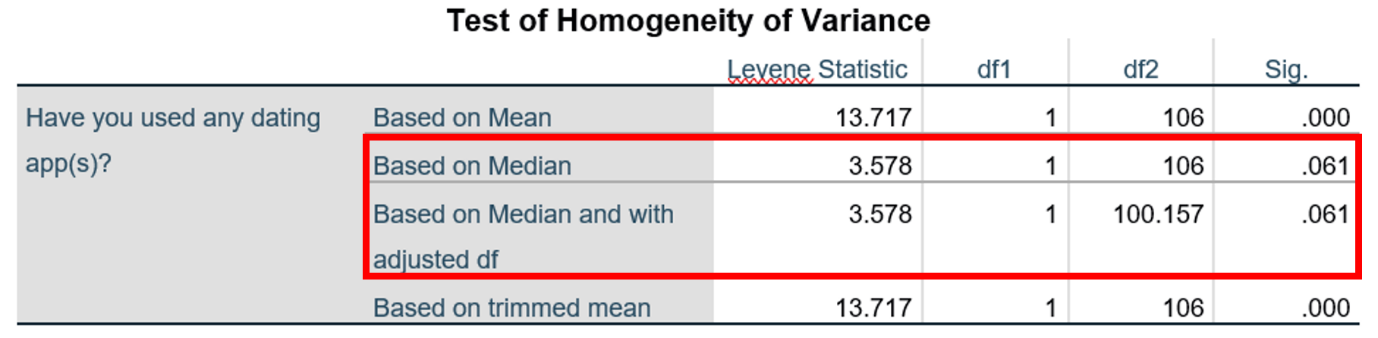


Figure 3.2 Non-parametric version of test of Homogeneity of Variance

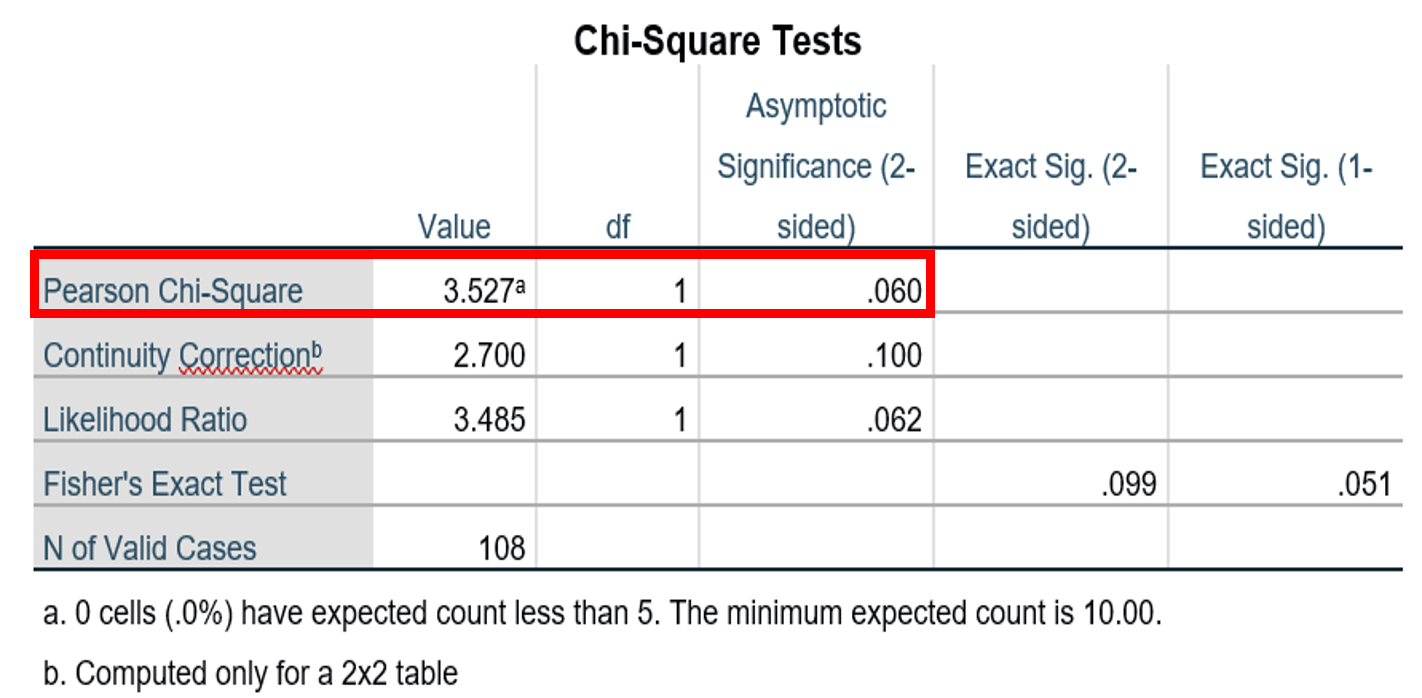


Figure 2.5 Chi-Square test summary

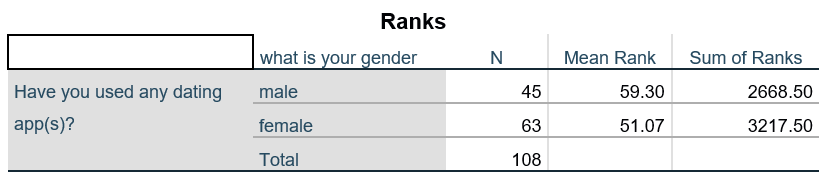


Figure 2.4 Mann-Whitney U test mean ranks

Normality test: The p-value (2.7985E-17) based on Shapiro-Wilk test is lower than 0.05. The bar chart (Appendix 2.1.1) also shows there is obvious difference between the gender on their use on dating apps. The data failed to pass the normality test.

Test on homogeneity of variance: We obtained the values of Levene statistic for “based on median” and “based on median and with adjusted df” are both the same(3.578). The p-value for both are 0.061, which is greater than 0.05, which implied that the variances for heterosexual and non-heterosexual are homogenous.

Based on Mann-Whitney test, the mean rank of male students is higher than female students’ (59.30 > 51.07). This shows that more male students in UM are using dating apps. However, the p-value is 0.062(greater than 0.05), indicating that there is no statistically significant difference. The null hypothesis is not rejected in this case.

Since the two variables tested(“Gender” and “Have you used dating apps”) can be treated as two categorical variables, Chi-square Independence test is used to test the association between the two variables. The bar chart (Appendix 2.1.1) and the contingency table (Appendix 2.1.2) provided some information and more direct view about the association between the two variables. The Pearson Chi-Square p-value (Asymptotic significance(2-sided)) is 0.60 (greater than 0.05), which is not statistically significant. We concluded that there is association found between gender and usage of dating apps. (X2(1) > = 3.527, p = 0.060). The usage of dating apps is dependent on the gender.

**Result for hypothesis 3**

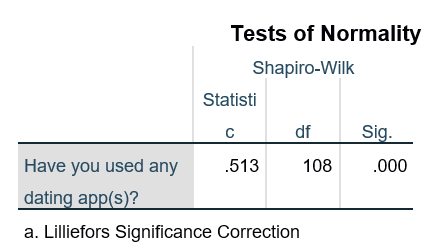
****

Figure 3.1 Normality test

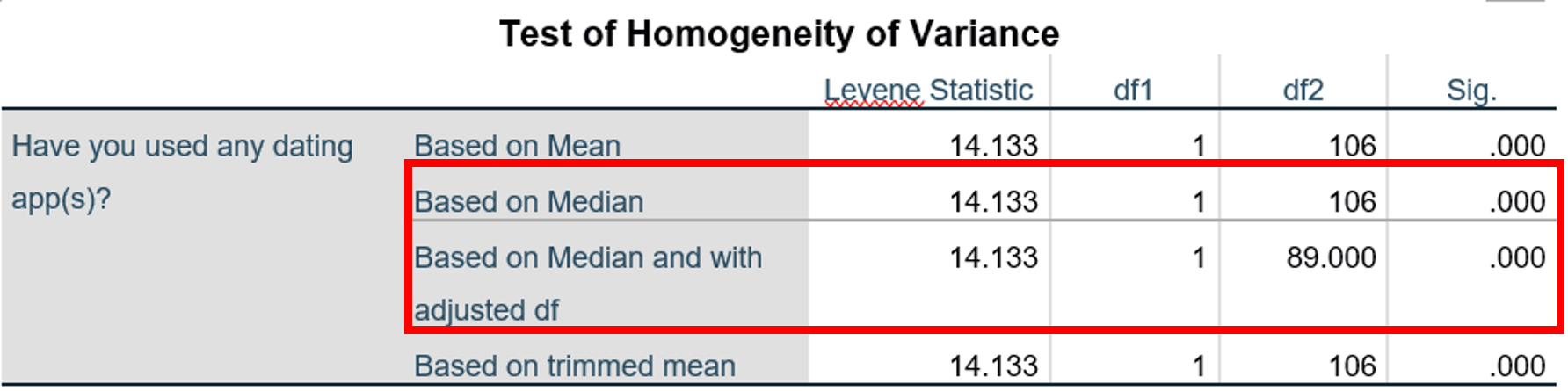


Figure 4.2 Non-parametric version of test of Homogeneity of Variance

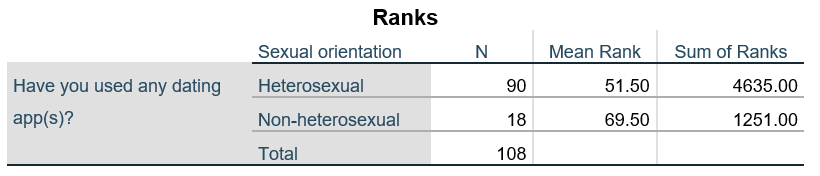


Figure 3.3 Mann-Whitney U test Ranks

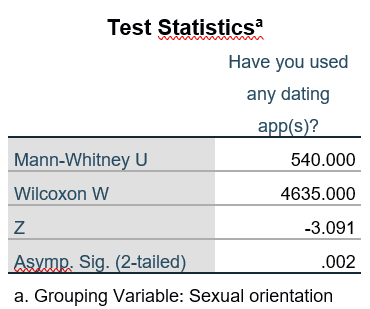


Figure 3.4 Mann-Whitney U test statistics

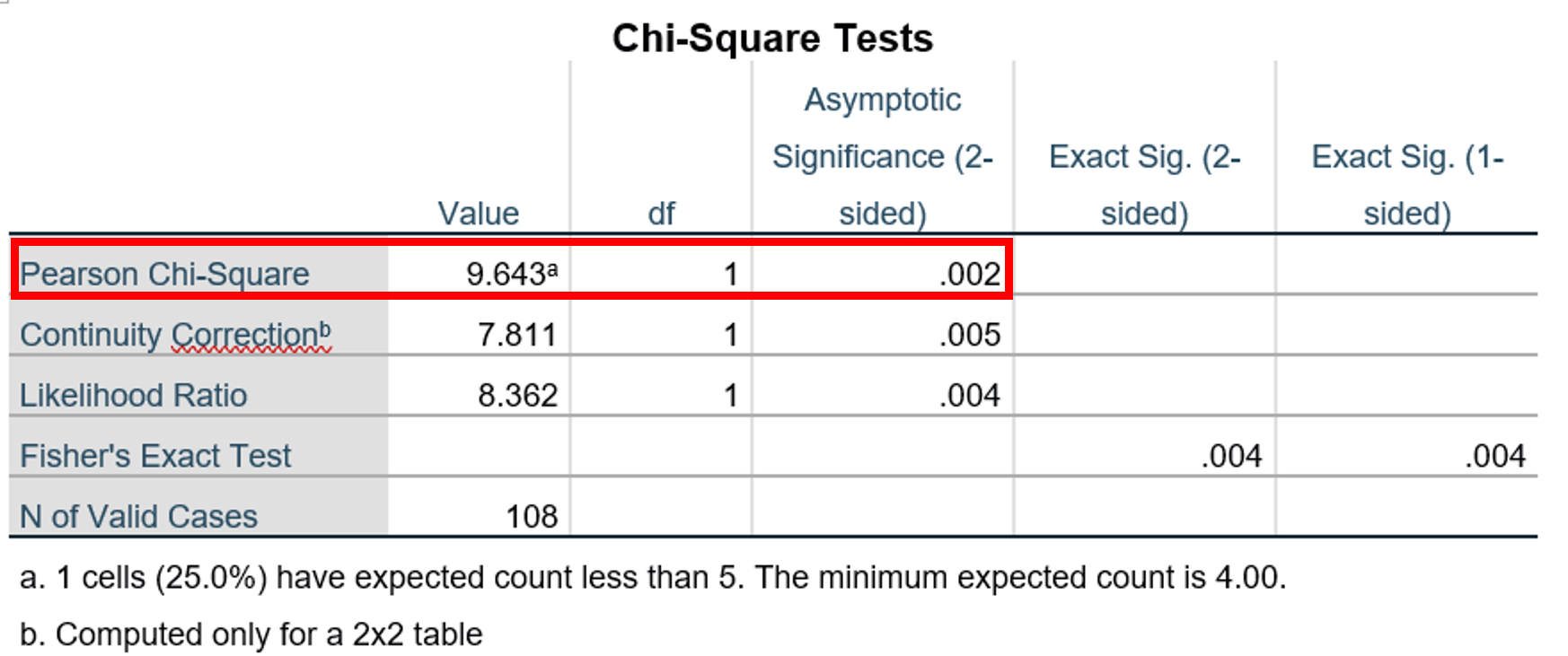


Figure 3.4 Chi-Square test summary

Normality test: Using the result of normality test in hypothesis 2, the data(usage of dating apps) failed to exhibit normal distribution as the p-value is lower than 0.05. Another normality test with sexual orientation involved (Appendix 3.1.1) also showed the same result. The histograms (Appendix 3.1.2) showed the distribution of heterosexual/non-heterosexual on their usage of dating apps.

Test on homogeneity of variance: We obtained the values of Levene statistic for “based on median” and “based on median and with adjusted df” are both the same(14.133). The p-value of the former(0.000279) and the latter(0.000304) smaller than 0.05, which implied that the variances for heterosexual and non-heterosexual are not homogenous. Implicitly we assume that the data distribution for two groups are not homogenous.

Mann-Whitney U test: From the mean rank, there was an obvious, numerically difference between heterosexual and non-heterosexual who use the dating app. From the mean rank we do not know if it is statistically significant because it is not normally distributed. The z-value(-3.091) is far from 0, and the p-value(0.002) is less than 0.05, thus we reject the null hypothesis, as there is significant difference between non-heterosexual and heterosexual in the usage of dating apps. The reason is there is very unequal number of heterosexual and non-heterosexual respondents.

Chi-square Independence test is used to test the association between the two variables. The bar chart contingency table (Appendix 3.1.3) provided some information and more direct view about the association between the two variables. The Pearson Chi-Square p-value (Asymptotic significance(2-sided)) is 0.002 (lower than 0.05), which is statistically significant. We concluded that there is no association was found between sexual orientation and usage of dating apps. (X2(1) > = 9.643, p = 0.002). The usage of dating apps is independent on the sexual orientation.

**Result for hypothesis 4**

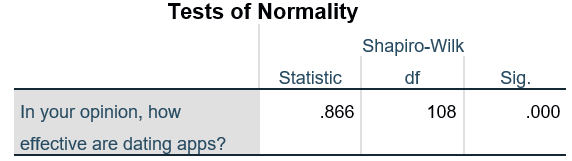


Figure 4.1 Normality test

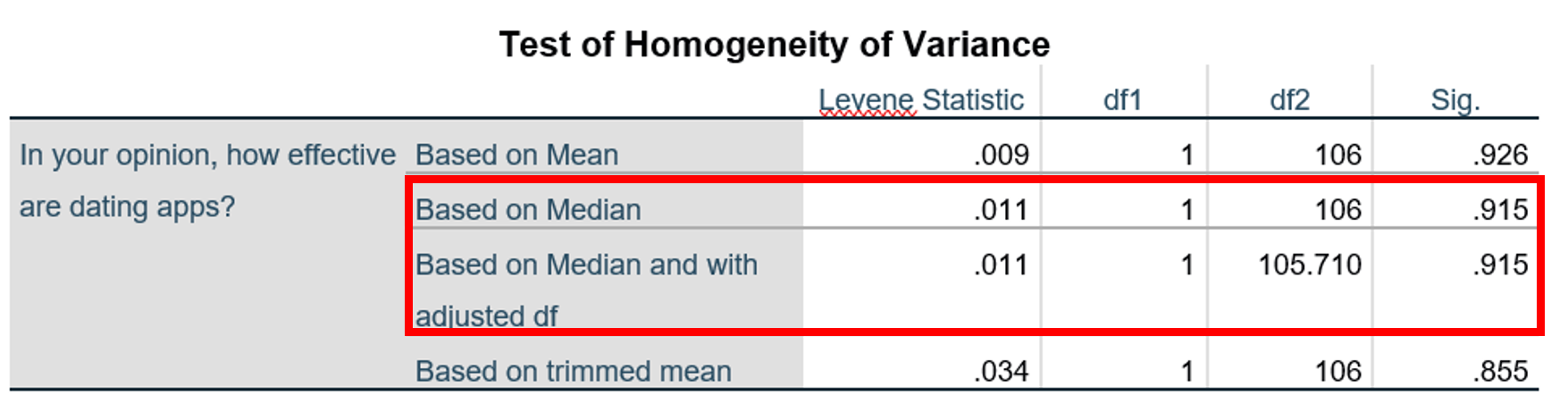


Figure 4.2 Non-parametric version of test of Homogeneity of Variance

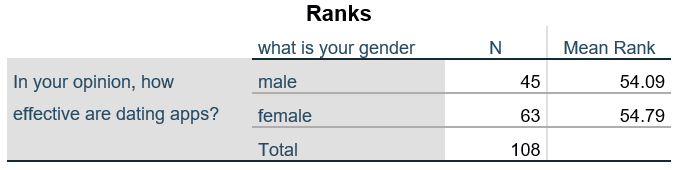


Figure 4.3 Kruskal-Wallis H test Mean Ranks Table

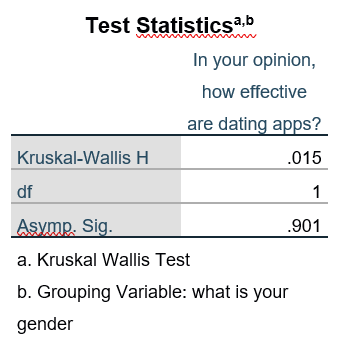


Figure 4.4 Kruskal-Wallis H test summary

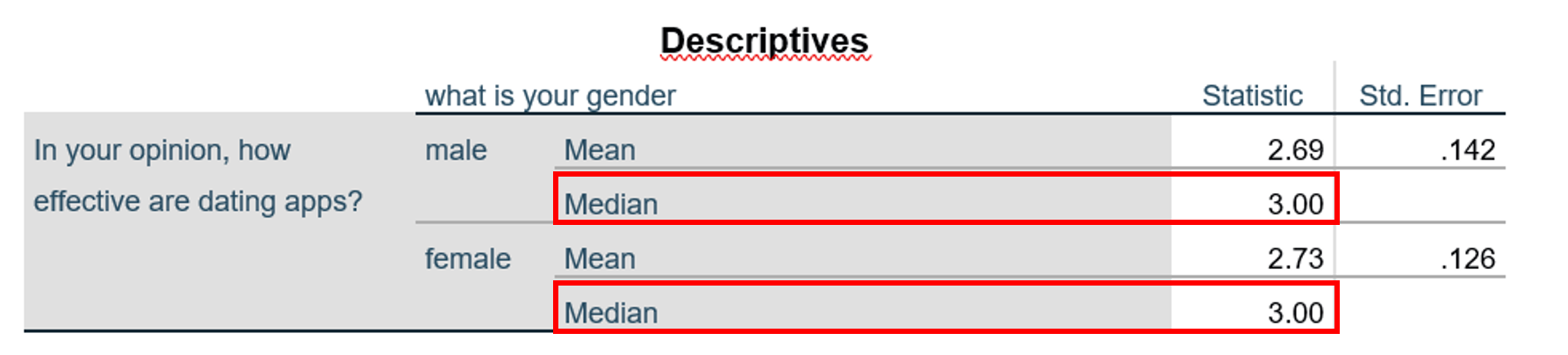


Figure 4.5 Median

Normality test: The p-value(1.8462E-8) of Shapiro-Wilk test is less than 0.05, this statistically significant result indicates that the data are not normally distributed. By examining on the bar chart from descriptive statistic, we can see that the data are not distributed normally (Appendix 4.1.1).

Test on homogeneity of variance: The Levene Statistic values for both “based on median” and “based on median and with adjusted df” are the same, 0.011. As well as the p-values, which are both 0.915, we can assume that there is no statistically significant difference between the variances of two groups of data since it failed to reject the null hypothesis.

Kruskal-Wallis H test: Although there is a difference between number of male and female students, we focused on the mean rank. The mean ranks for male(54.09) and female(54.78) are similar, which numerically showed that they have similar rating on the effectiveness of dating apps. The H value is 0.015, and the Asymptotic significance value(p-value) is 0.901, which is far greater than 0.05. Therefore, the null hypothesis is not rejected. T There is high possibility that there is no statistically significant difference between gender on the effectiveness of dating apps. The median is 3.00 for both male and female, most of the respondents think the dating apps are moderately effective in looking for partner.

**Result for hypothesis 5**

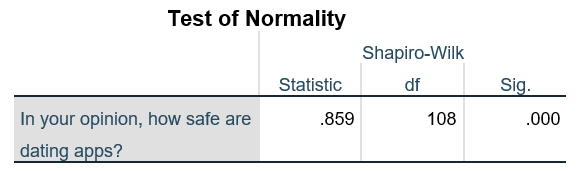
****

Figure 5.1 Normality test

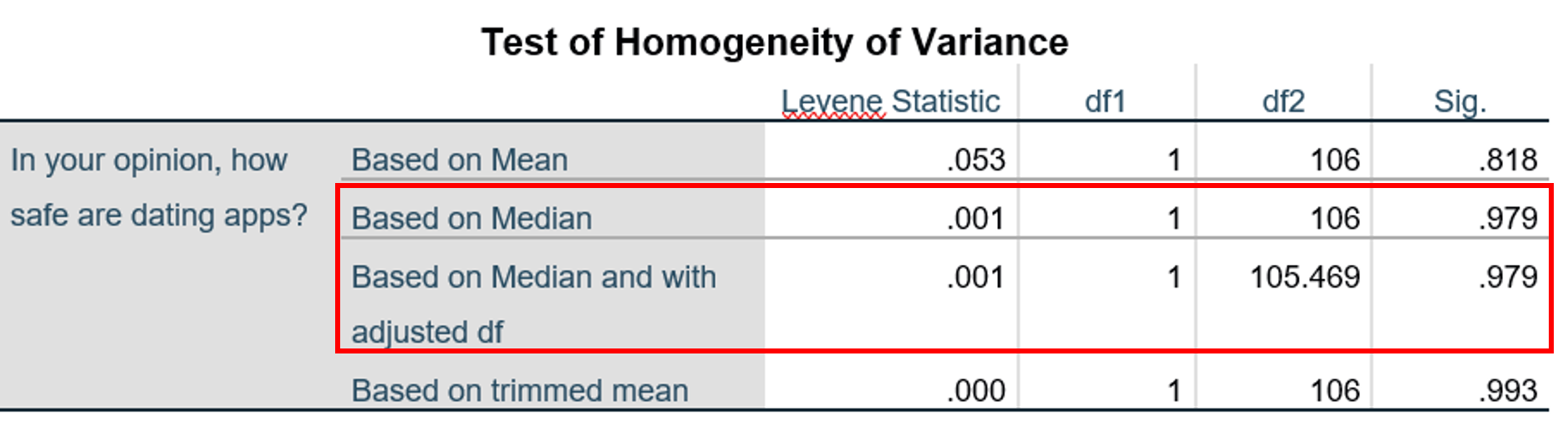


Figure 5.2 Non-parametric version of test of Homogeneity of Variance

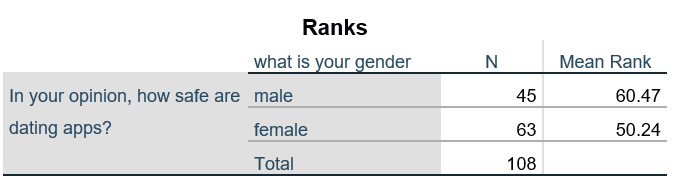


Figure5.3 Kruskal-Wallis H test Mean Ranks Table

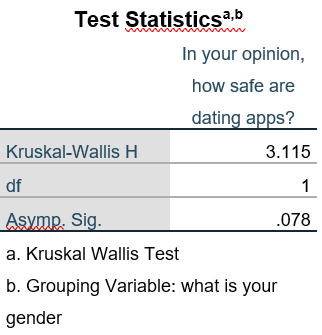


Figure 5.4 Kruskal-Wallis H test summary

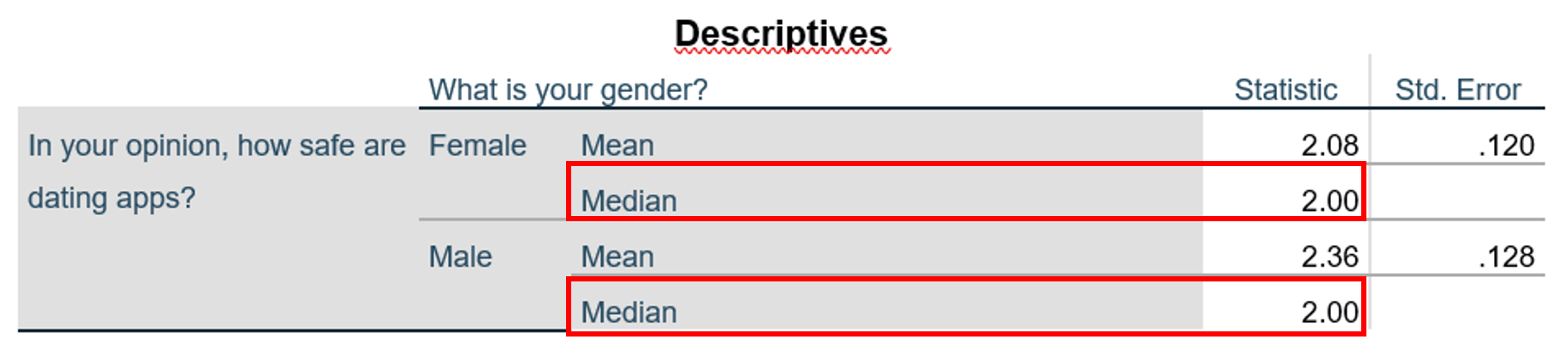


Figure 5.5 Median

Normality test: The p-value(9.8161E-9) is lower than 0.05. The data for these two groups are not distributed normally.

Test on homogeneity of variance: The Levene statistic values and the p-values from the “based on median and with adjusted df” are 0.001 and 0.979 respectively, as the p-value is greater than 0.05, the variances are roughly homogenous. The assumption of similar data distributions has been confirmed.

Kruskal-Willis test: the mean rank of both male and female shows significant difference. Male students tend to give higher rating on the safety usage of dating apps. However, the p-value equals to 0.078, which basically means there is a 7.8% chance of finding our sample results if gender does not have any effect in the opinion on the safeness of dating apps. The p-value is still higher than 0.05, showing no statistically significant difference. The null hypothesis was not rejected despite of the difference shown in the mean ranks. Similar to the median, both male and female think that the dating apps are unsafe.

Conclusion (not done yet)