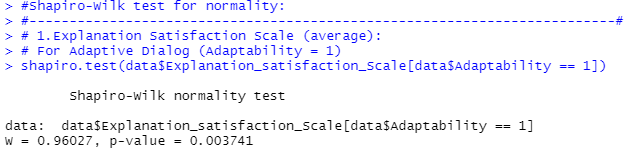
**Sawyer project- Understandable Robot Using ChatGPT**

**ALL Groups-statistical analysis Results-scenario 1+2-Subjective measures:**

**1.Explanation Satisfaction Scale (average):**

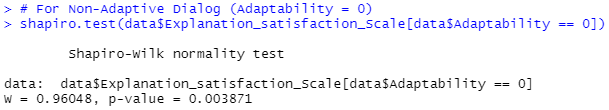
**Shapiro-Wilk test for normality:**

**#For Adaptive Dialog (Adaptability = 1):**

****

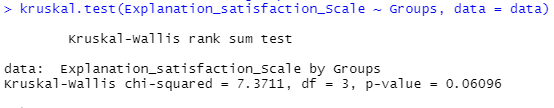
**Conclusion**: p-value>0.05 , The adaptive Dialog does not follow a normal distribution.

**For Non-Adaptive Dialog (Adaptability = 0):**



**Conclusion**: p-value>0.05 , The non-adaptive Dialog does not follow a normal distribution.

**Kruskal-Wallis Test: This test determines if there are significant differences between the four groups.**

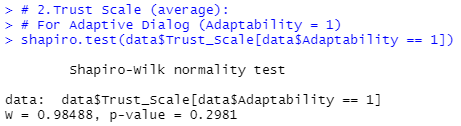


**Conclusion**: Since the p-value is greater than 0.05, there is **no statistically significant difference** in the medians of the Explanation satisfaction Scale across the four groups at the 5% significance level. Since the Kruskal-Wallis test shows no significant difference, we will not conduct further pairwise Mann-Whitney U tests.

**2. Trust Scale (average):**

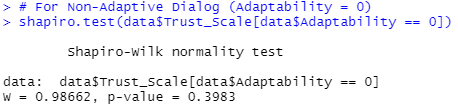
**Shapiro-Wilk test for normality:**

**For Adaptive Dialog (Adaptability = 1):**



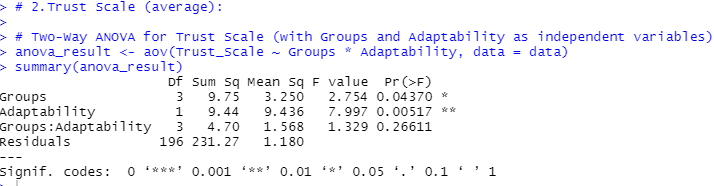
**Conclusion**: p-value>0.05 , The adaptive Dialog follows a normal distribution.

**For Non-Adaptive Dialog (Adaptability = 0):**



**Conclusion**: p-value>0.05 , The non-adaptive Dialog follows a normal distribution.

**Two-Way ANOVA for Trust Scale (with Groups and Adaptability as independent variables):**

****

**Main Effects:**

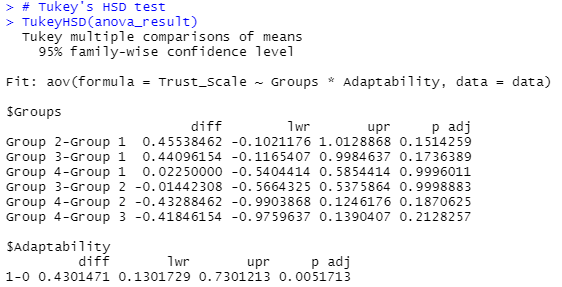
1. **Groups**: There is a significant main effect of **Groups** on the Trust Scale, with a p-value of **0.04370** (less than 0.05), indicating that there are differences in the Trust Scale between the groups.
2. **Adaptability**: There is a significant main effect of **Adaptability** on the Trust Scale, with a p-value of **0.00517** (less than 0.01), suggesting that adaptability influences the Trust Scale.

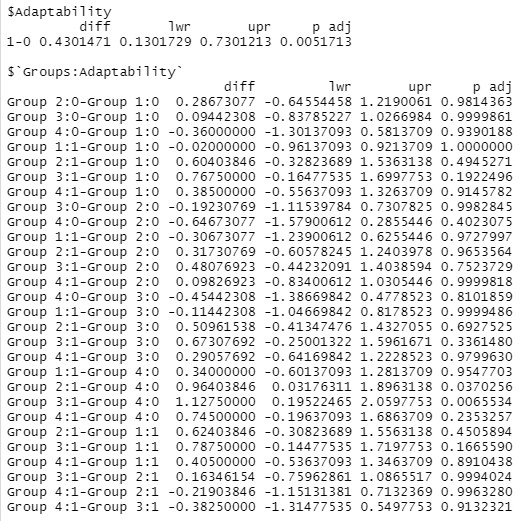
**Interaction Effect:**

1. **Groups: Adaptability**: The interaction effect between **Groups** and **Adaptability** is not significant, with a p-value of **0.26611** (greater than 0.05). This means that the effect of Groups on the Trust Scale does not depend on the level of Adaptability.

**Conclusion:**

* Both **Groups** and **Adaptability** have significant main effects on the Trust Scale.
* There is no significant interaction between the two factors (Groups and Adaptability), suggesting that their effects on Trust Scale are independent.

**Tukey's HSD test:**:



**Key points:**

**1. Main Effect of Groups:**

* The comparisons between **Groups** (Group 1, Group 2, Group 3, Group 4) show no statistically significant differences in the Trust Scale, as the p-values are all greater than 0.05 (e.g., for Group 2 vs Group 1: p = 0.151).
* The most notable difference is between **Group 2 and Group 4** (p = 0.037), where Group 2 shows a higher Trust Scale than Group 4, with a significant p-value below 0.05.

**2. Main Effect of Adaptability:**

* There is a significant difference between **Adaptability levels** (Adaptive vs. Non-Adaptive Dialog). The p-value for the comparison between Adaptability 1 (Adaptive) and Adaptability 0 (Non-Adaptive) is **0.005**, indicating that the Trust Scale is significantly different between these two levels of Adaptability.

**3. Interaction Effect of Groups and Adaptability:**

* For the interaction between **Groups and Adaptability**, the p-values are mostly above 0.05, indicating that the interaction effect is not statistically significant. For example, the comparison of **Group 2 with Adaptability 0 vs. Group 1 with Adaptability 0** has a p-value of **0.981**, suggesting no significant interaction between the two variables for those specific groups and Adaptability levels.
* However, significant differences are found for **Group 2 vs Group 4** when considering the interaction with Adaptability (p = 0.037 for **Group 2:1 - Group 4:0**), where Group 2 with Adaptability 1 differs significantly from Group 4 with Adaptability 0.

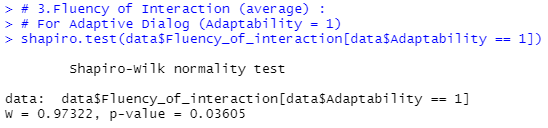
**Conclusions:**

* There are **no significant differences** in the Trust Scale across the groups when Adaptability is not considered.
* **Adaptability** (Adaptive vs. Non-Adaptive Dialog) has a significant impact on the Trust Scale, with Adaptive Dialog (1) showing higher scores than Non-Adaptive Dialog (0).
* The interaction effect between **Groups and Adaptability** is **mostly non-significant**, but there are some cases (like Group 2 vs Group 4) where the interaction might play a role.

**3. Fluency of Interaction (average):**

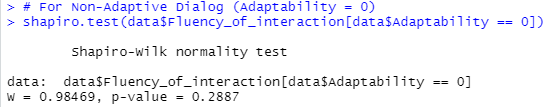
**Shapiro-Wilk test for normality:**

**For Adaptive Dialog (Adaptability = 1):**



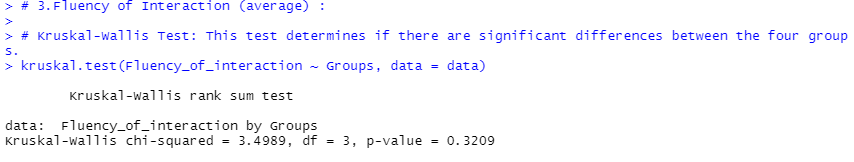
**Conclusion**: p-value>0.05 , The adaptive Dialog does not follow a normal distribution.

**For Non-Adaptive Dialog (Adaptability = 0):**



**Conclusion**: p-value>0.05 , The non-adaptive Dialog follows a normal distribution.

**Kruskal-Wallis Test: This test determines if there are significant differences between the four groups.**

****

**Conclusion**: The results of the **Kruskal-Wallis test** for the **Fluency of Interaction** show that the p-value is **0.3209**, which is greater than 0.05. This suggests that there is no statistically significant difference in the **Fluency of Interaction** across the four groups.

Since the Kruskal-Wallis test shows no significant difference, we will not conduct further pairwise Mann-Whitney U tests.