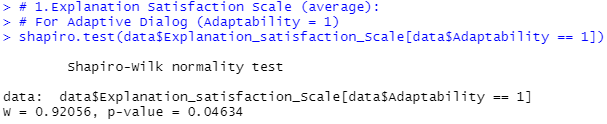
**Sawyer project- Understandable Robot Using ChatGPT**

**Group2-statistical analysis Results-scenario 1+2-Subjective measures:**

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

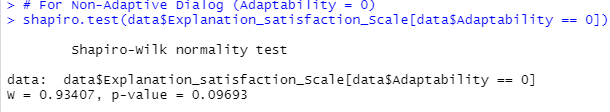
**Shapiro-Wilk test for normality:**

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

****

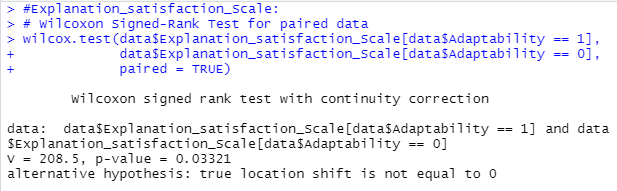
**Conclusion:** The p-value is less than 0.05, indicating that the data for **the Adaptive Dialog** condition **does not** **follow a normal distribution.**

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

****

The p-value is greater than 0.05, suggesting that the data for the **Non-Adaptive Dialog** condition **follows a normal distribution**.

**Wilcoxon Signed-Rank Test for paired data for Explanation satisfaction Scale:**



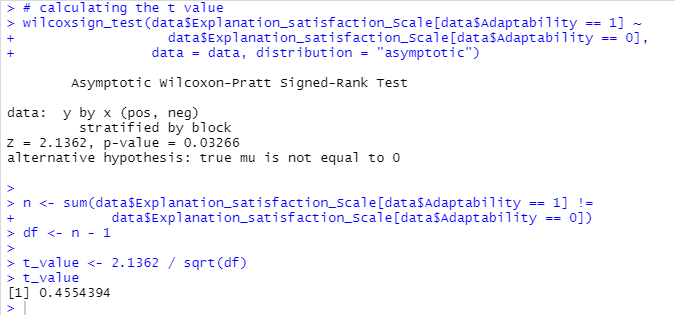
**Conclusion**: **Results:**

V = 208.5: The test statistic (V) represents the sum of ranks for the positive differences between the paired conditions.

p-value = 0.03321: The p-value is less than 0.05, indicating that the difference in Explanation Satisfaction Scale between the two conditions is statistically significant.

Alternative hypothesis: The test suggests that there is a true location shift (difference) between the two conditions.

Calculating the t value:

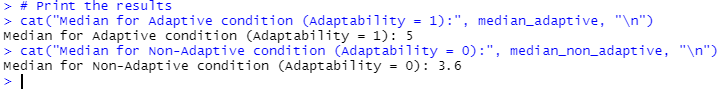


t value **≈** 0.455

**Conclusion:**

There is a statistically significant difference in Explanation Satisfaction Scale between Adaptive and Non-Adaptive Dialog conditions.

**Medians for Explanation satisfaction:**



**Adaptive Condition (Adaptability = 1):** Median = **5**

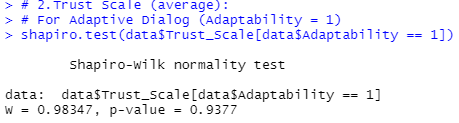
**Non-Adaptive Condition (Adaptability = 0):** Median = **3.6**

The higher median value in the Adaptive condition suggests that participants reported higher Explanation Satisfaction when the dialog was adaptive (Adaptability = 1).

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

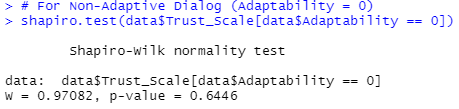
**Shapiro-Wilk test for normality:**

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

****

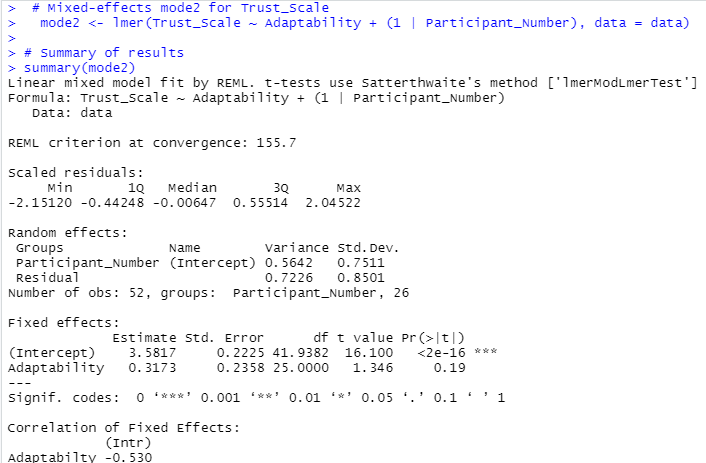
**Conclusion**: p-value>0.05 , The adaptive Dialog follows a normal distribution.

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



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

**# Mixed-effects model for Trust Scale:**

****

**Anova**:

תמונה שמכילה טקסט, גופן, צילום מסך, קו

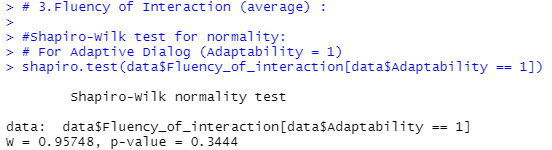
התיאור נוצר באופן אוטומטי

**Conclusion:** The p-value for **Adaptability** is **0.1904**, which is above the significance threshold of **0.05**. This means the **effect of Adaptability** on **Trust Scale** is **not statistically significant**.

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

**Shapiro-Wilk test for normality:**

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



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

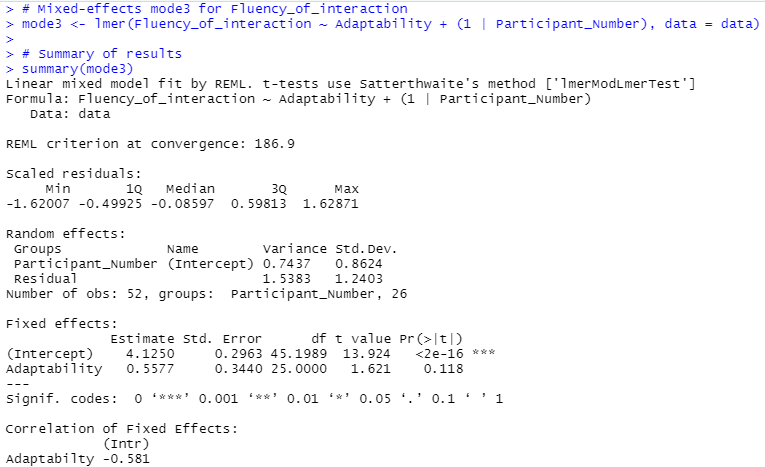
**For Non-Adaptive Dialog (Adaptability = 0):**

תמונה שמכילה טקסט, גופן, צילום מסך, קו

התיאור נוצר באופן אוטומטי

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

**Mixed-effects mode3 for Fluency of interaction**:



**Anova:**

תמונה שמכילה טקסט, גופן, צילום מסך, לבן

התיאור נוצר באופן אוטומטי

**Conclusion:** The p-value for **Adaptability** is **0.1175**, which is above the significance threshold of **0.05**. This means the **effect of Adaptability** on **Fluency of interaction** is **not statistically significant**.