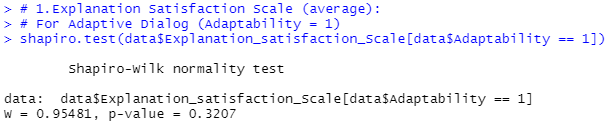
**Sawyer project- Understandable Robot Using ChatGPT**

**Group4-statistical analysis Results-scenario 1+2-Subjective measures:**

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

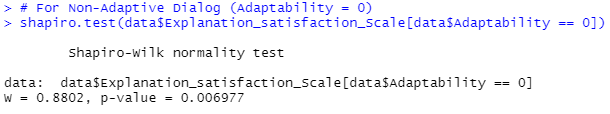
**Shapiro-Wilk test for normality:**

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



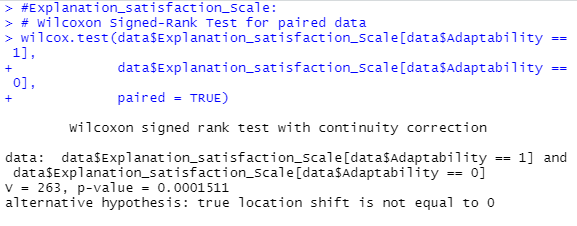
**Conclusion :**The p-value is greater than 0.05, suggesting that the data for the **Adaptive Dialog** condition **follows a normal distribution**.

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



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

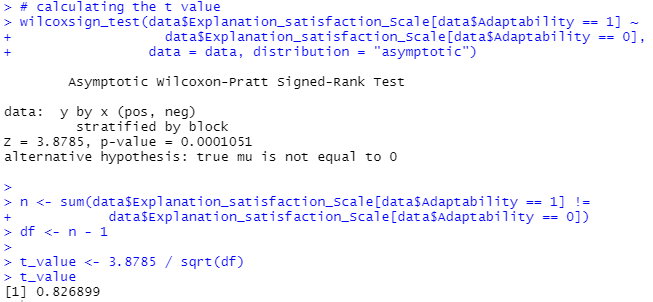
**Wilcoxon Signed-Rank Test for paired data for Explanation satisfaction Scale:**



Key results:

1. **Test Statistic (V)**:  
   The test statistic V=263 represents the sum of signed ranks for the differences between paired observations.
2. **P-Value**:  
   The p-value p=0.0001511 indicates strong statistical significance (p<0.001). This suggests that there is a significant difference in **Explanation Satisfaction** between the **Adaptable** and **Non-Adaptable Dialog** conditions.
3. **Alternative Hypothesis**:  
   The alternative hypothesis (H1states that the median difference between the paired conditions is not zero (i.e., there is a shift in satisfaction levels).

**Calculating t value:**

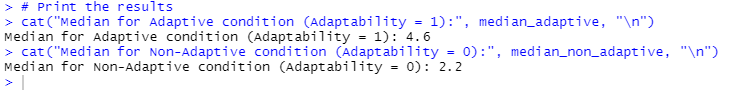


t value = 0.827

**Conclusion:**

Participants rated their **Explanation Satisfaction** significantly higher in the **Adaptive Dialog** condition compared to the **Non-Adaptive Dialog** condition.

**Medians for Explanation satisfaction:**



**Adaptive Condition (Adaptability = 1):** Median = **4.6**

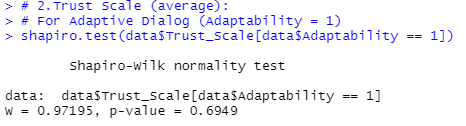
**Non-Adaptive Condition (Adaptability = 0):** Median = **2.2**

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):**

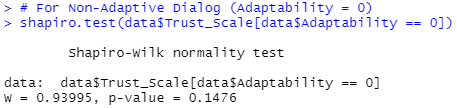
**Shapiro-Wilk test for normality:**

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



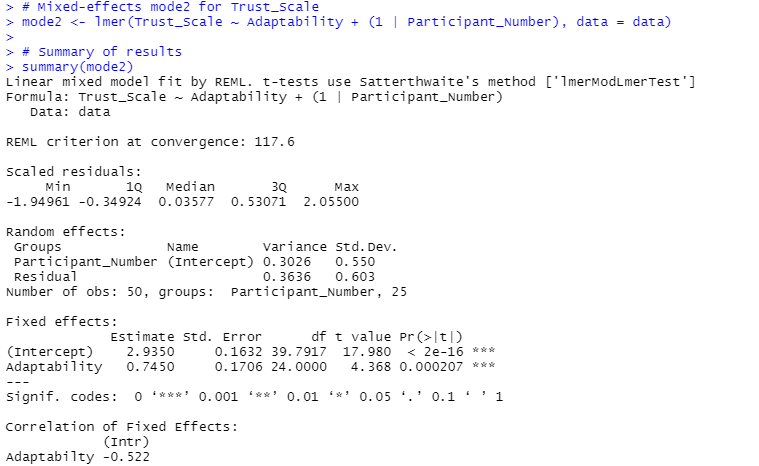
**Conclusion**: p-value>0.05 , The adaptive Dialog follows a normal distribution.

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

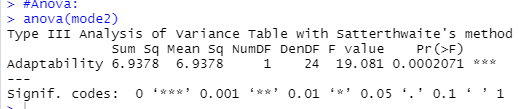


**Conclusion**: p-value>0.05 , The adaptive Dialog follows a normal distribution.

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

****

**Anova:**



**Key Findings for Trust Scale (Updated Model):**

**Fixed Effects:**

* **Intercept (Baseline Trust Scale):**
  + Estimate: **2.935** (significant, p<0.001), representing the average Trust Scale score in the **Non-Adaptive Dialog** condition (Adaptability=0\text{Adaptability} = 0Adaptability=0).
* **Effect of Adaptability:**
  + Estimate: **0.745** (significant, p=0.0002), indicating that the **Trust Scale** increases by **0.745 points** in the **Adaptive Dialog** condition compared to the **Non-Adaptive Dialog** condition.

**Random Effects:**

* **Participant-Level Variability:**
  + Variance: **0.3026**, reflecting differences in trust levels across participants.
* **Residual Variance:**
  + Variance: **0.3636**, representing unexplained variability in trust scores.

**ANOVA Results:**

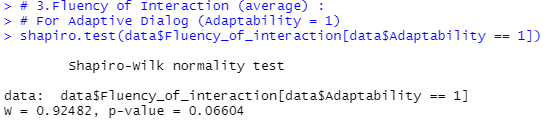
* F(1,24)=19.08,p=0.0002:
  + The effect of **Adaptability** on the **Trust Scale** is highly statistically significant.

**Conclusion:**  
Adaptive Dialog significantly improves trust compared to Non-Adaptive Dialog, with an estimated increase of **0.745 points**. The model highlights both the fixed effect of dialog type and variability across participants.

**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.

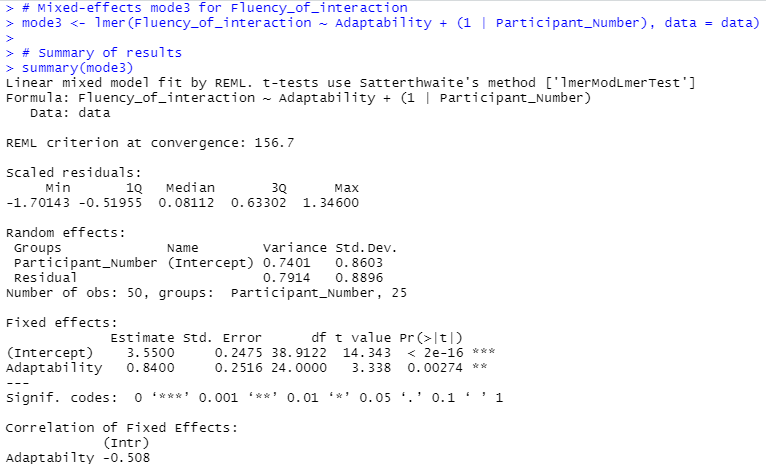
**For Non-Adaptive Dialog (Adaptability = 0):**

תמונה שמכילה טקסט, גופן, צילום מסך, לבן

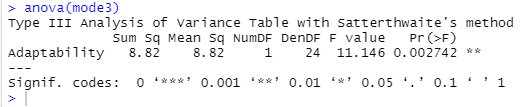
התיאור נוצר באופן אוטומטי

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

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



**Anova:**



**Key Findings for Fluency of Interaction (Mixed-Effects Model):**

**Fixed Effects:**

* **Intercept (Baseline Fluency of Interaction):**
  + Estimate: **3.55** (significant, p<0.001), representing the average fluency score in the **Non-Adaptive Dialog** condition (Adaptability=0\text{Adaptability} = 0).
* **Effect of Adaptability:**
  + Estimate: **0.84** (significant, p=0.0027), indicating that **Fluency of Interaction** improves by **0.84 points** in the **Adaptive Dialog** condition compared to the **Non-Adaptive Dialog** condition.

**Random Effects:**

* **Participant-Level Variability:**
  + Variance: **0.7401**, reflecting differences in fluency across participants.
* **Residual Variance:**
  + Variance: **0.7914**, representing unexplained variability in fluency scores.

**ANOVA Results:**

* F(1,24)=11.15 ,p=0.0027:
  + The effect of **Adaptability** on **Fluency of Interaction** is statistically significant.

**Conclusion:**  
Adaptive Dialog significantly enhances fluency compared to Non-Adaptive Dialog, with an estimated improvement of **0.84 points**. The model accounts for individual variability, confirming the importance of adaptability in interaction fluency.