Checking:

- **Normality**:

  Low-Carb, After 1 Month: W = 0.952, p = 0.695

  Low-Carb, After 2 Months: W = 0.975, p = 0.935

  Low-Fat, After 1 Month: W = 0.948, p = 0.646

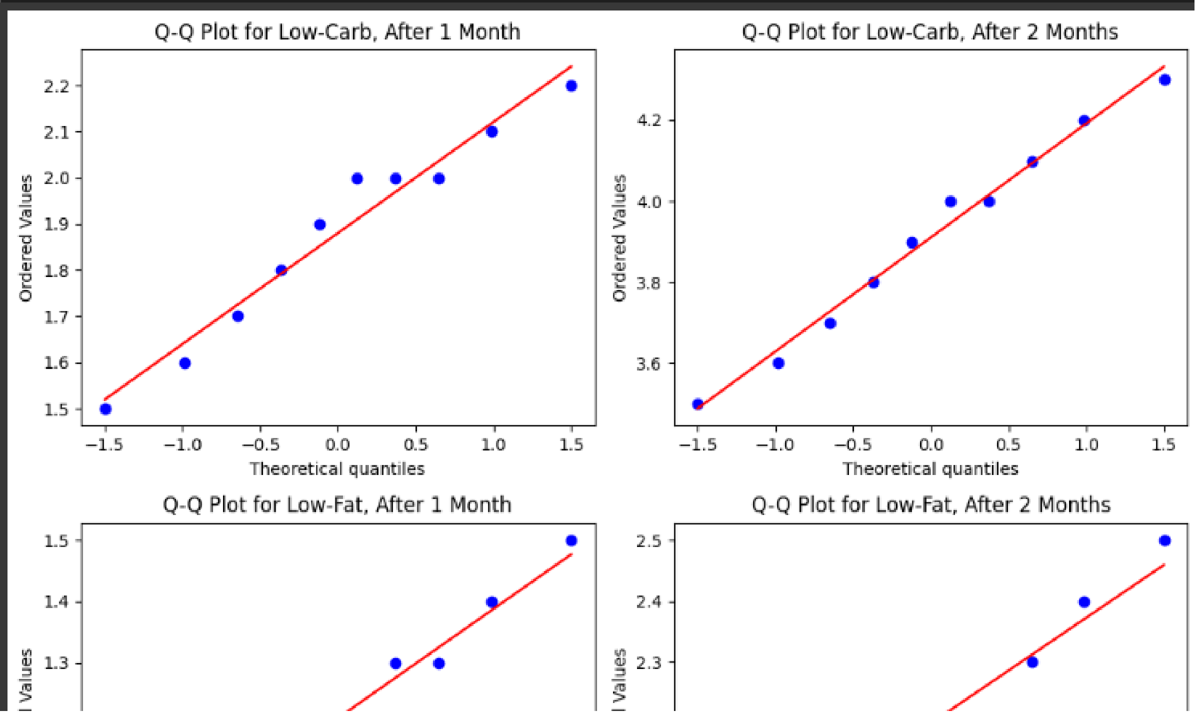
  Low-Fat, After 2 Months: W = 0.935, p = 0.494

  All *p*-values are above the threshold of 0.05, meaning that we **fail to reject the null hypothesis** in each case. This indicates that the data for each group at each time point follows a normal distribution.

```python
shapiro_low_carb_1 = stats.shapiro(low_carb_month_1)
shapiro_low_carb_2 = stats.shapiro(low_carb_month_2)
shapiro_low_fat_1 = stats.shapiro(low_fat_month_1)
shapiro_low_fat_2 = stats.shapiro(low_fat_month_2)


print("Shapiro-Wilk Test Results:")
print(f"Low-Carb, After 1 Month: W = {shapiro_low_carb_1.statistic}, p = {shapiro_low_carb_1.pvalue}")
print(f"Low-Carb, After 2 Months: W = {shapiro_low_carb_2.statistic}, p = {shapiro_low_carb_2.pvalue}")
print(f"Low-Fat, After 1 Month: W = {shapiro_low_fat_1.statistic}, p = {shapiro_low_fat_1.pvalue}")
print(f"Low-Fat, After 2 Months: W = {shapiro_low_fat_2.statistic}, p = {shapiro_low_fat_2.pvalue}")
```

```
Shapiro-Wilk Test Results:
Low-Carb, After 1 Month: W = 0.9522463889969321, p = 0.6951164516077839
Low-Carb, After 2 Months: W = 0.9752339025839641, p = 0.9346854448707635
Low-Fat, After 1 Month: W = 0.9480837918205034, p = 0.6458864004974137
Low-Fat, After 2 Months: W = 0.9345479260931479, p = 0.4940832001928637
```

```python
fig, axes = plt.subplots(2, 2, figsize=(10, 8))

stats.probplot(low_carb_month_1, dist="norm", plot=axes[0, 0])
axes[0, 0].set_title("Q-Q Plot for Low-Carb, After 1 Month")


stats.probplot(low_carb_month_2, dist="norm", plot=axes[0, 1])
axes[0, 1].set_title("Q-Q Plot for Low-Carb, After 2 Months")


stats.probplot(low_fat_month_1, dist="norm", plot=axes[1, 0])
axes[1, 0].set_title("Q-Q Plot for Low-Fat, After 1 Month")


stats.probplot(low_fat_month_2, dist="norm", plot=axes[1, 1])
axes[1, 1].set_title("Q-Q Plot for Low-Fat, After 2 Months")

plt.tight_layout()
plt.show()
```



- **Sphericity**:

  Mauchly’s Test of Sphericity:

  W: 0.094

  p-value: 5.896e-10 (extremely small)

  Since the p-value is much less than 0.05, sphericity is violated. This means that the assumption of equal variances of differences between time points does not hold.

  Greenhouse-Geisser Correction:

  The Greenhouse-Geisser epsilon (eps): 0.524.

The p-value with correction (p-GG-corr): 2.071e-12, which is still significant.

Interpretation:

The main effect of Time is significant: F (2, 38) = 221.51, p < .001 (both uncorrected and Greenhouse-Geisser corrected).

The Greenhouse-Geisser correction was applied because sphericity was violated.

The repeated measures ANOVA showed a significant effect of Time on weight loss, F (2, 38) = 221.51, p < .001, with a large effect size, $\eta^2 = 0.84$. Since Mauchly's test indicated a violation of the sphericity assumption, χ^2 (2) = 42.50, p < .001, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.52$). The Greenhouse-Geisser corrected p-value confirmed the significance of the Time effect, p < .001.

```
df_melt = pd.melt(df, id_vars=['Participant', 'Diet Type'],
                  value_vars=['Baseline', 'After 1 month', 'After 2 months'],
                  var_name='Time', value_name='Weight_Loss')
aov = pg.rm_anova(dv='Weight_Loss', within='Time', subject='Participant', data=df_melt, detailed=True)
mauchly = pg.sphericity(df_melt, dv='Weight_Loss', within='Time', subject='Participant')

print("Repeated Measures ANOVA:")
print(aov)

print("\nMauchly's Test of Sphericity:")
print(mauchly)
```

```
Repeated Measures ANOVA:
  Source         SS  DF         MS           F         p-unc     p-GG-corr  \
0   Time  93.334333   2  46.667167  221.512137  1.134267e-21  2.071801e-12
1  Error   8.005667  38   0.210675         NaN           NaN           NaN

        ng2       eps sphericity   W-spher       p-spher
0  0.835182  0.524741      False  0.094299  5.896243e-10
1       NaN       NaN        NaN       NaN           NaN

Mauchly's Test of Sphericity:
SpherResults(spher=False, W=0.09429927870847703, chi2=42.5030712889439, dof=2, pval=5.896242510895627e-10)
```

- **Homogeneity of Variance**:

The results from Levene's Test indicate the following:

W = 10.4740

p-value = 0.002

Interpretation: Since the p-value (0.002) is less than 0.05, this suggests that the variances between the two diet groups (Low-Carb and Low-Fat) are significantly different. In other words, the assumption of homogeneity of variance is violated.

```python
low_carb = df[['Baseline', 'After 1 month', 'After 2 months']].iloc[0:10].values.flatten()
low_fat = df[['Baseline', 'After 1 month', 'After 2 months']].iloc[10:20].values.flatten()

stat, p_value = stats.levene(low_carb, low_fat)

print(f"Levene’s Test for Homogeneity of Variance: W = {stat:.4f}, p = {p_value:.4e}")
```

```
Levene’s Test for Homogeneity of Variance: W = 10.4740, p = 2.0026e-03
```

```python
if p_value < 0.05:
    print("The variances are significantly different (p < 0.05). The assumption of homogeneity of variance is violate
else:
    print("The variances are approximately equal (p > 0.05). The assumption of homogeneity of variance is met.")
```

```
The variances are significantly different (p < 0.05). The assumption of homogeneity of variance is violated.
```

When the variances are not equal, the F-ratio used in ANOVA can become biased. This bias can lead to an inflated Type I error rate, meaning you might incorrectly reject the null hypothesis and conclude that there are significant differences between groups when there are none.

- **Independence**:

```python
duplicates = df[df.duplicated(['Participant'], keep=False)]
if duplicates.empty:
    print("All participants are independent (no duplicates found).")
else:
    print("Duplicate participants found:\n", duplicates)
```

```
All participants are independent (no duplicates found).
```

Based on the results of the two-way mixed model ANOVA, we can interpret the effects as follows:

1. Main Effect of Diet Type (Between-Subjects Effect)

$F(1, 18) = 174.707$, $p < 0.001$

Partial $\eta^2 = 0.907$ (indicating a large effect size)

Interpretation: The main effect of diet type is significant, suggesting that there is a statistically significant difference in weight loss between the Low-Carb and Low-Fat diet groups. Since the p-value is extremely low ($p < 0.001$), we can confidently reject the null hypothesis that there is no difference in weight loss between the two diet types. The large effect size ($\eta^2 = 0.907$) indicates that approximately 90.7% of the variability in weight loss can be attributed to the diet type, highlighting a strong practical significance of the diet intervention.

2. Main Effect of Time (Within-Subjects Effect)

$F(2, 36) = 2889.939$, $p < 0.001$

Partial $\eta^2 = 0.994$ (indicating a very large effect size)

Interpretation: The main effect of time is also significant, indicating that weight loss changed significantly over the three time points (Baseline, After 1 Month, After 2 Months), regardless of diet type. The extremely low p-value ($p < 0.001$) allows us to reject the null hypothesis that there is no change in weight loss over time. The very large effect size ($\eta^2 = 0.994$) suggests that about 99.4% of the variability in weight loss can be explained by the time factor, indicating a strong and meaningful change in weight loss across the different time points.

3. Interaction Effect (Diet Type × Time)

$F(2, 36) = -$

p-value: -

Interpretation: The interaction effect between diet type and time is not present in the results, as indicated by the empty Data Frame for this effect. This suggests that there is no significant interaction between diet type and time. In other words, the effect of diet type on weight loss does not depend on the time point; the differences in weight loss between the Low-Carb and Low-Fat diets are consistent across the different time measurements. Therefore, we do not reject the null hypothesis for the interaction effect, meaning the impact of the diet does not change over the time points assessed.

```python
diet_effect = aov.loc[aov['Source'] == 'Diet Type']
print("\nMain Effect of Diet Type:\n", diet_effect)


time_effect = aov.loc[aov['Source'] == 'Time']
print("\nMain Effect of Time:\n", time_effect)


interaction_effect = aov.loc[aov['Source'] == 'Diet Type:Time']
print("\nInteraction Effect (Diet Type x Time):\n", interaction_effect)
```

```
Main Effect of Diet Type:
       Source        SS  DF1  DF2        MS           F         p-unc  \
0  Diet Type  9.440667    1   18  9.440667  174.707334  1.048878e-10

   p-GG-corr       np2  eps sphericity  W-spher  p-spher
0        NaN  0.906594  NaN        NaN      NaN      NaN

Main Effect of Time:
   Source         SS  DF1  DF2         MS          F         p-unc  \
1   Time  93.334333    2   36  46.667167  2889.93922  1.779792e-40

      p-GG-corr      np2       eps sphericity   W-spher       p-spher
1  2.071801e-12  0.99381  0.524741      False  0.094299  5.896243e-10

Interaction Effect (Diet Type x Time):
 Empty DataFrame
Columns: [Source, SS, DF1, DF2, MS, F, p-unc, p-GG-corr, np2, eps, sphericity, W-spher, p-spher]
Index: []
```

https://github.com/yohannibuiltdiff/Applied-Multivariate