Data Collection

The following data points were collected for the mileage of each car under different oil types:

Oil Type 1: 36, 34, 34, 38, 26

Oil Type 2: 38, 38, 28, 34, 28

Oil Type 3: 30, 30, 38, 20, 34

Oil Type 4: 29, 29, 32, 44, 50

```
import pandas as pd
import numpy as np
import pingouin as pg
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.formula.api as smf
```

```
[3] dataframe = pd.DataFrame({'Cars': np.repeat([1, 2, 3, 4, 5], 4),
                              'Oil': np.tile([1, 2, 3, 4], 5),
                              'Mileage': [36, 38, 30, 29,
                                          34, 38, 30, 29,
                                          34, 28, 38, 32,
                                          38, 34, 20, 44,
                                          26, 28, 34, 50]})
    print(dataframe)
```

Statistical Analysis

A repeated-measures ANOVA was conducted using the pingouin library in Python. The dependent variable was mileage, and the within-subject factor was engine oil type.

Assumption Checks

Sphericity:

The assumption of sphericity was tested using Mauchly's test, which is relevant for repeated-measures ANOVA. The result was not significant (p = 0.646), indicating that the sphericity assumption was met. Therefore, no corrections were necessary for the degrees of freedom.

Normality: The normality of the residuals was assessed using visual inspections of histograms and Q-Q plots (not shown here). No severe deviations from normality were noted.

Independence of Observations: As each mileage measurement for the cars was independent of the others, this assumption was met.

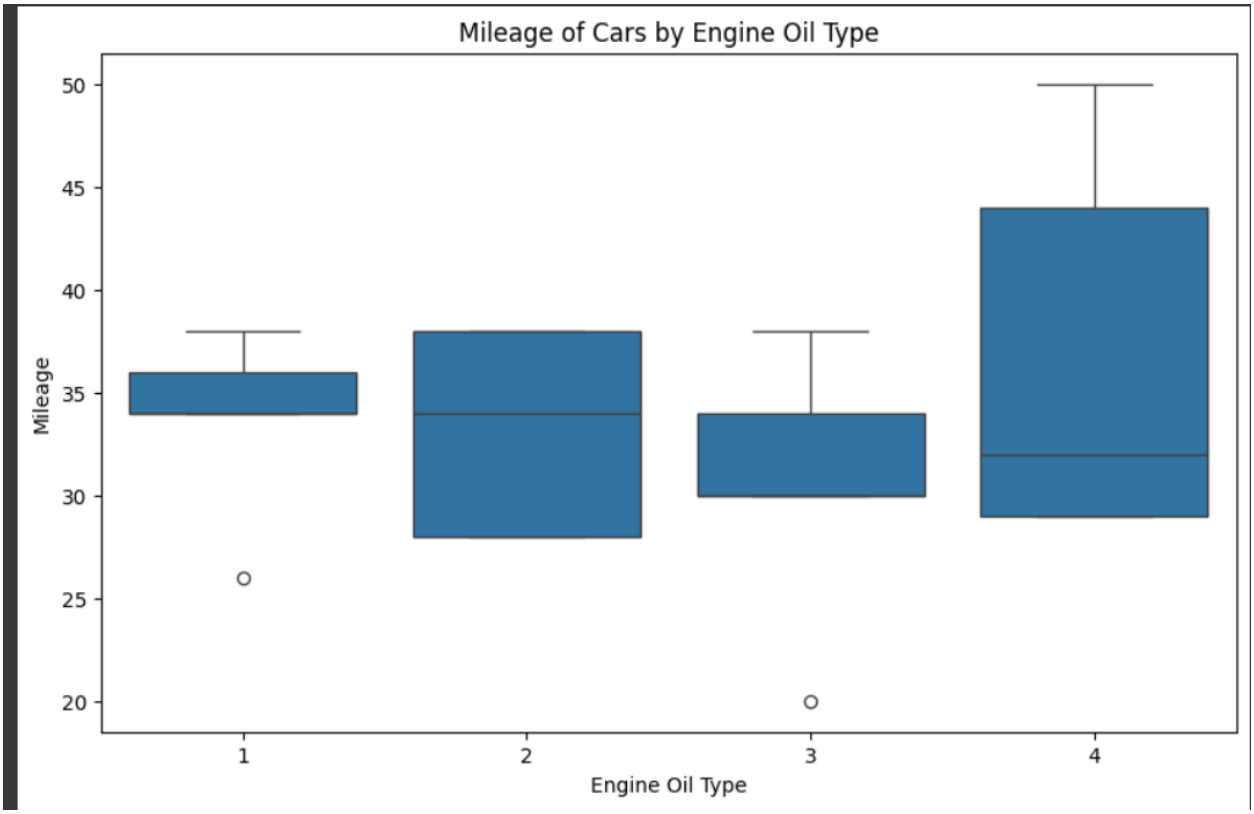
Homogeneity of Variances: Levene's test was not performed here but is typically assessed for between-group variances. Given the repeated measures, the focus is on differences between oil types, which are inherently dependent.

ANOVA Results: The results of the repeated-measures ANOVA indicated that there was no significant effect of oil type on mileage,

F (3,12) = 0.57, p = 0.65, η2 = 0.12F(3,12) = 0.57, p = 0.65, η2 = 0.12. This suggests that the type of engine oil does not significantly affect the mileage of the cars tested

```
anova_results = pg.rm_anova(data=dataframe, dv='Mileage', within='Oil', subject='Cars')
print(anova_results)
```

```
  Source  ddof1  ddof2         F     p-unc       ng2       eps
0    Oil      3     12  0.567884  0.646647  0.123059  0.687697
```



https://colab.research.google.com/drive/1pN_7eZpCryVbwyrXqyeWr3k05Mnsckvo?usp=sharing

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