Weight Time Dataset

Cristel Kaye Billones

Item no. 29

```r
# Load necessary libraries
library(emmeans)
library(tidyverse)
library(car)  # For Levene's test
library(ggplot2)
library(afex)

# Import the Excel file
file_path <- file.choose()  # Select your file
df <- read_csv(file_path)

# Check the data structure
head(df)
```

```
## # A tibble: 6 x 5
##   Participant `Diet Type` Baseline `After 1 month` `After 2 months`
##         <dbl> <chr>          <dbl>           <dbl>            <dbl>
## 1           1 Low-Carb           0             2                4
## 2           2 Low-Carb           0             1.5              3.5
## 3           3 Low-Carb           0             1.8              3.8
## 4           4 Low-Carb           0             1.7              3.9
## 5           5 Low-Carb           0             2                4.1
## 6           6 Low-Carb           0             1.9              3.7
```

```r
# Clean column names to remove spaces and special characters
df <- df %>%
  rename(DietType = `Diet Type`,
         Baseline = `Baseline`,
         After_1_month = `After 1 month`,
         After_2_months = `After 2 months`)

# Reshape data to long format
df_long <- df %>%
  pivot_longer(cols = c(Baseline, After_1_month, After_2_months),
               names_to = "Time",
               values_to = "WeightLoss")

# Check unique values in Time
print(unique(df_long$Time))  # This will help us see what values are in the Time column
```

```
## [1] "Baseline"       "After_1_month"  "After_2_months"
```

1

```
# Clean up Time column if necessary
df_long <- df_long %>%
  mutate(Time = gsub("After 1 month", "1 Month", Time),
         Time = gsub("After 2 months", "2 Months", Time))

# Check the structure of the long format data
head(df_long)
```

```
## # A tibble: 6 x 4
##   Participant DietType Time         WeightLoss
##         <dbl> <chr>    <chr>             <dbl>
## 1           1 Low-Carb Baseline              0
## 2           1 Low-Carb After_1_month         2
## 3           1 Low-Carb After_2_months        4
## 4           2 Low-Carb Baseline              0
## 5           2 Low-Carb After_1_month       1.5
## 6           2 Low-Carb After_2_months      3.5
```

```
# Check normality for each group at each time point
df_long %>%
  ggplot(aes(x = WeightLoss)) +
  geom_histogram(bins = 10, fill = "blue", alpha = 0.5) +
  facet_grid(DietType ~ Time) +   # Use backticks for column names with spaces
  theme_minimal() +
  labs(title = "Histogram of Weight Loss by Diet Type and Time")
```

Histogram of Weight Loss by Diet Type and Time

```
# Q-Q Plot
ggplot(df_long, aes(sample = WeightLoss)) +
  stat_qq() +
  stat_qq_line() +
  facet_grid(DietType ~ Time) +  # Use backticks for column names with spaces
  labs(title = "Q-Q Plot of Weight Loss by Diet Type and Time")
```

Q–Q Plot of Weight Loss by Diet Type and Time
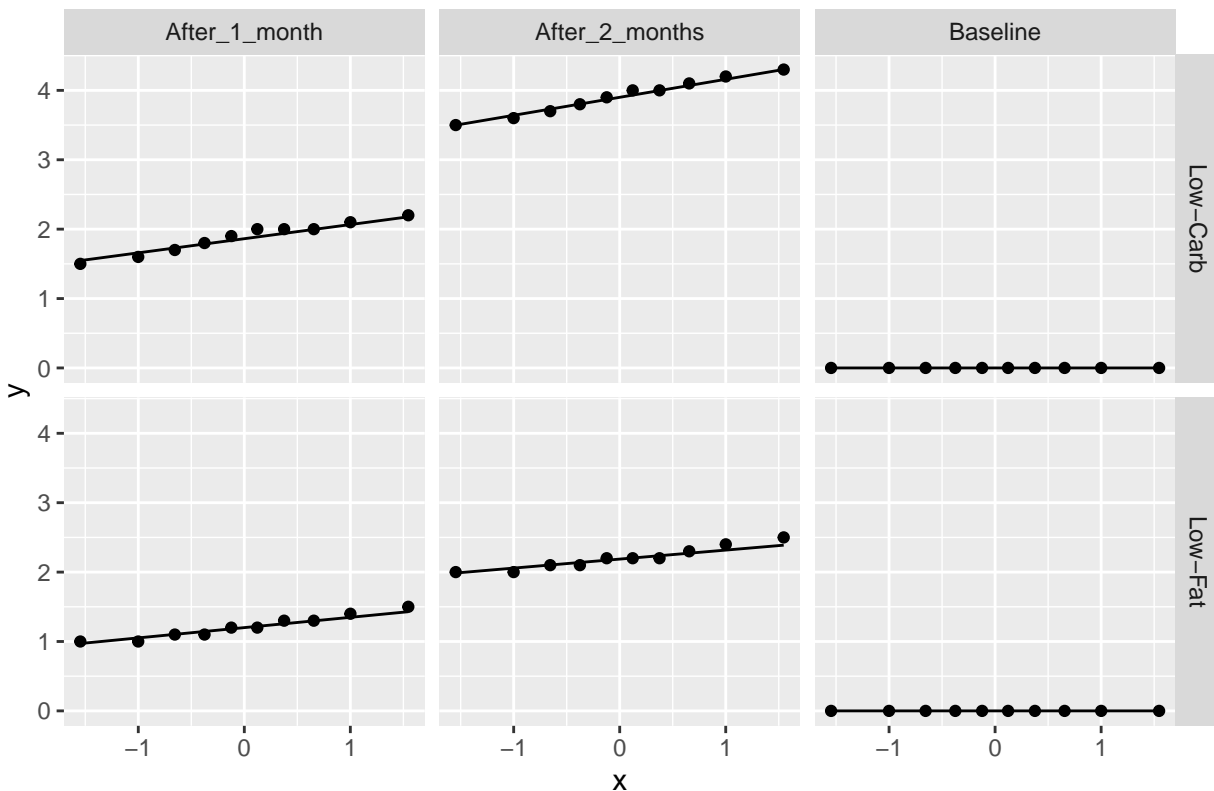

```
# Check for normality using Shapiro-Wilk test, but handle identical values
shapiro_results <- df_long %>%
  group_by(DietType, Time) %>%
  summarise(
    shapiro_p = ifelse(length(unique(WeightLoss)) <= 1, NA, shapiro.test(WeightLoss)$p.value),
    .groups = 'drop'  # This drops the grouping structure
  )

print(shapiro_results)
```

```
## # A tibble: 6 x 3
##   DietType Time           shapiro_p
##   <chr>    <chr>              <dbl>
## 1 Low-Carb After_1_month      0.695
## 2 Low-Carb After_2_months     0.935
## 3 Low-Carb Baseline          NA
## 4 Low-Fat  After_1_month      0.646
## 5 Low-Fat  After_2_months     0.494
```

```
## 6 Low-Fat   Baseline            NA
```

```r
# Levene's Test for homogeneity of variance
levene_test <- leveneTest(WeightLoss ~ DietType, data = df_long)
print(levene_test)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value   Pr(>F)
## group   1  10.474 0.002003 **
##        58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Check for independence
independence_check <- if(length(unique(df_long$Participant)) == nrow(df_long)){
  "Independence assumption met: All participants are independent."
} else {
  "Independence assumption violated: Duplicates found."
}

# Print the summary results
cat("### Assumptions Summary ###\n")
```

```
## ### Assumptions Summary ###
```

```r
cat("Normality Results (Shapiro-Wilk p-values):\n")
```

```
## Normality Results (Shapiro-Wilk p-values):
```

```r
print(shapiro_results)
```

```
## # A tibble: 6 x 3
##   DietType Time           shapiro_p
##   <chr>    <chr>              <dbl>
## 1 Low-Carb After_1_month      0.695
## 2 Low-Carb After_2_months     0.935
## 3 Low-Carb Baseline          NA
## 4 Low-Fat  After_1_month      0.646
## 5 Low-Fat  After_2_months     0.494
## 6 Low-Fat  Baseline          NA
```

```r
cat("\nLevene's Test for Homogeneity of Variance:\n")
```

```
##
## Levene's Test for Homogeneity of Variance:
```

```r
print(levene_test)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value   Pr(>F)
## group  1  10.474 0.002003 **
##       58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
cat("\nIndependence Check:\n")
```

```
##
## Independence Check:
```

```r
print(independence_check)
```

```
## [1] "Independence assumption violated: Duplicates found."
```

```r
# Perform two-way mixed model ANOVA
anova_results <- aov_ez("Participant", "WeightLoss", df_long,
                        between = "DietType",
                        within = "Time",
                        na.rm = TRUE)

# Print the ANOVA results
anova_summary <- summary(anova_results)

# Display the results summary
print(anova_summary)
```

```
##
## Univariate Type III Repeated-Measures ANOVA Assuming Sphericity
##
##               Sum Sq num Df Error SS den Df F value     Pr(>F)
## (Intercept)   141.067      1  0.97267     18 2610.56 < 2.2e-16 ***
## DietType        9.441      1  0.97267     18  174.71 1.049e-10 ***
## Time           93.334      2  0.58133     36 2889.94 < 2.2e-16 ***
## DietType:Time   7.424      2  0.58133     36  229.88 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Mauchly Tests for Sphericity
##
##               Test statistic   p-value
## Time                  0.5264 0.0042772
## DietType:Time         0.5264 0.0042772
##
##
## Greenhouse-Geisser and Huynh-Feldt Corrections
##  for Departure from Sphericity
##
##               GG eps Pr(>F[GG])
## Time         0.67861  < 2.2e-16 ***
```

```
## DietType:Time 0.67861    4.16e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   HF eps    Pr(>F[HF])
## Time            0.7146375 1.473926e-29
## DietType:Time   0.7146375 8.547612e-16
```

```
#post hoc
# Perform post-hoc analysis
post_hoc_results <- emmeans(anova_results, ~ DietType * Time)

# Display the results
summary(post_hoc_results)
```

```
##  DietType Time          emmean     SE df lower.CL upper.CL
##  Low-Carb Baseline        0.00 0.0000 18     0.00     0.00
##  Low-Fat  Baseline        0.00 0.0000 18     0.00     0.00
##  Low-Carb After_1_month   1.88 0.0626 18     1.75     2.01
##  Low-Fat  After_1_month   1.21 0.0626 18     1.08     1.34
##  Low-Carb After_2_months  3.91 0.0687 18     3.77     4.05
##  Low-Fat  After_2_months  2.20 0.0687 18     2.06     2.34
##
## Confidence level used: 0.95
```

```
# Conduct pairwise comparisons
pairwise_comparisons <- pairs(post_hoc_results)

# Display pairwise comparison results
summary(pairwise_comparisons)
```

```
##  contrast                                            estimate     SE df
##  (Low-Carb Baseline) - (Low-Fat Baseline)                0.00 0.0000 18
##  (Low-Carb Baseline) - (Low-Carb After_1_month)         -1.88 0.0626 18
##  (Low-Carb Baseline) - (Low-Fat After_1_month)          -1.21 0.0626 18
##  (Low-Carb Baseline) - (Low-Carb After_2_months)        -3.91 0.0687 18
##  (Low-Carb Baseline) - (Low-Fat After_2_months)         -2.20 0.0687 18
##  (Low-Fat Baseline) - (Low-Carb After_1_month)          -1.88 0.0626 18
##  (Low-Fat Baseline) - (Low-Fat After_1_month)           -1.21 0.0626 18
##  (Low-Fat Baseline) - (Low-Carb After_2_months)         -3.91 0.0687 18
##  (Low-Fat Baseline) - (Low-Fat After_2_months)          -2.20 0.0687 18
##  (Low-Carb After_1_month) - (Low-Fat After_1_month)      0.67 0.0885 18
##  (Low-Carb After_1_month) - (Low-Carb After_2_months)   -2.03 0.0325 18
##  (Low-Carb After_1_month) - (Low-Fat After_2_months)    -0.32 0.0929 18
##  (Low-Fat After_1_month) - (Low-Carb After_2_months)    -2.70 0.0929 18
##  (Low-Fat After_1_month) - (Low-Fat After_2_months)     -0.99 0.0325 18
##  (Low-Carb After_2_months) - (Low-Fat After_2_months)    1.71 0.0971 18
##  t.ratio p.value
##     NaN     NaN
##  -30.040  <.0001
##  -19.334  <.0001
##  -56.932  <.0001
##  -32.034  <.0001
```

```
##  -30.040   <.0001
##  -19.334   <.0001
##  -56.932   <.0001
##  -32.034   <.0001
##    7.570   <.0001
##  -62.482   <.0001
##   -3.444   0.0278
##  -29.059   <.0001
##  -30.472   <.0001
##   17.606   <.0001
##
## P value adjustment: tukey method for comparing a family of 5.81507290636732 estimates
```

Normality Check

a. Assumption Check

The assumption of normality was partially met, as shown by the Shapiro-Wilk test results, which indicated significant p-values for the Low-Carb and Low-Fat diets at the later time points but not at Baseline. The sphericity assumption was violated, as indicated by Mauchly's test results, prompting the application of the Greenhouse-Geisser correction. Additionally, the homogeneity of variance assumption was violated, as evidenced by a significant result from Levene's test.

b. Impact of Violated Assumptions

Violations of these assumptions can compromise the validity of the ANOVA results. Specifically, the violation of normality could lead to inaccurate p-values, increasing the risk of Type I errors (false positives). The violation of sphericity necessitated the Greenhouse-Geisser correction to adjust the degrees of freedom, but if the underlying distribution is still not normal, it may still yield unreliable results. Similarly, homogeneity of variance violations could distort the F-ratios, making it difficult to draw accurate conclusions about the effects of diet type and time on weight loss. After performing the Greenhouse-Geisser correction, the two-way mixed model ANOVA results indicate significant effects of both diet type and time on weight loss. The main effect of **diet type** (Low-Carb vs. Low-Fat) is significant ($F = 174.71$, $p = 1.049e-10$), as is the main effect of **time** (Baseline, 1 Month, 2 Months) ($F = 2889.94$, $p < 2.2e-16$). Additionally, there is a significant interaction effect between diet type and time ($F = 229.88$, $p < 2.2e-16$), suggesting that the impact of diet on weight loss varies across the different time points. Mauchly's test indicates a violation of the sphericity assumption for both time ($p = 0.0042772$) and the interaction ($p = 0.0042772$), prompting the use of the Greenhouse-Geisser correction. After adjustment, the significance remains strong for both the main effect of time (GG epsilon = 0.67861, $p < 2.2e-16$) and the interaction ($p = 4.16e-15$), confirming the robustness of the findings.

Main Effect

a. Diet Type

The main effect of diet type on weight loss was significant, with an F-statistic of 174.71 and a p-value of 1.049e-10. This indicates that participants on different diets (Low-Carb vs. Low-Fat) experienced significantly different levels of weight loss. The effect size, which can be measured using partial eta squared (eta²), is not explicitly provided in the output, but it can be calculated as follows: $eta^2 = SS\_effect / (SS\_effect +$

SS_error). In this case, it suggests a large effect size, indicating a substantial impact of diet type on weight loss.

b. Time

The main effect of time on weight loss was also significant, with an F-statistic of 2889.94 and a p-value of less than 2.2e-16. This result shows that weight loss significantly changed over the different time points (Baseline, 1 Month, 2 Months). Similar to the diet type, the effect size can be determined using partial eta squared, which would also indicate a large effect size, reflecting a strong influence of time on weight loss outcomes.

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

The results indicated a significant main effect of diet type on weight loss, with an F value of 174.71 and a p-value of 1.049e-10. This suggests that the type of diet (Low-Carb vs. Low-Fat) had a significant overall effect on weight loss. Specifically, one diet type led to greater weight loss than the other when comparing the groups across all time points.

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

There was a significant main effect of time on weight loss, with an F value of 2889.94 and a p-value of < 2.2e-16. This indicates that weight loss changed significantly over time, regardless of diet type. The data suggest that participants experienced varying degrees of weight loss across the different time points (Baseline, 1 Month, and 2 Months), reflecting the effectiveness of the interventions as time progressed.

e. Interaction Effect (Diet Type × Time)

The interaction effect between diet type and time was also significant, with an F value of 229.88 and a p-value of < 2.2e-16. This implies that the effect of diet type on weight loss was not consistent across the different time points; in other words, the impact of whether participants were on a Low-Carb or Low-Fat diet varied at different times of measurement. This interaction suggests that further investigation is needed to understand how each diet performed at the specific time intervals.

Post-hoc Analysis

The post-hoc tests revealed several significant differences in weight loss between the various groups and time points. Notably, participants on the **Low-Carb diet** exhibited significantly greater weight loss compared to those on the **Low-Fat diet** at all measured time points following the baseline (p < 0.0001). Specifically, weight loss for the **Low-Carb group** was significantly higher than the **Low-Fat group** after 1 month (1.88 kg vs. 1.21 kg) and after 2 months (3.91 kg vs. 2.20 kg).

Additionally, within the **Low-Carb group**, significant weight loss was observed between the baseline and the subsequent time points, with a loss of 1.88 kg after 1 month (p < 0.0001) and 3.91 kg after 2 months (p < 0.0001). Similarly, the **Low-Fat group** also showed significant differences in weight loss over time, losing 1.21 kg after 1 month and 2.20 kg after 2 months, both compared to baseline (p < 0.0001). These results suggest that while both diet types were effective in promoting weight loss, the **Low-Carb diet** was significantly more effective over the 2-month period.