Nafisat Ibrahim
Tuesday, May 07th 2024

**Project**

## Task 1: Create Side-by-side boxplots

**Code:**

```
library(ggplot2)
data <- read.csv("WalkData4650.csv")
data_long <- reshape(data, varying = c("Usual_Walk", "Teabag_Walk", "Putey_Walk"),
                     v.names = "EE", timevar = "WalkingStyle",
                     times = c("Usual Walk", "Teabag Walk", "Putey Walk"),
                     direction = "long")
ggplot(data_long, aes(x = WalkingStyle, y = EE, fill = WalkingStyle)) +
  geom_boxplot() +
  labs(title = "Boxplot of Energy Expenditure by Walking Style",
       x = "Walking Style",
       y = "Energy Expenditure (kcal/kg/min)") +
  theme_minimal()
```

**Output:**


**Interpretation:**

- The "Teabag Walk" has the highest median EE and also shows the greatest variability in energy expenditure among the three walking styles. This suggests that the Teabag Walk, likely being more physically exaggerated, requires more energy.

- The "Usual Walk" shows the least energy expenditure and less variability, indicating it is the most efficient walking style.
- The "Putey Walk" has a lower median EE compared to the Teabag Walk but is slightly higher than the Usual Walk, with moderate variability.

## Task 2: Analysis of Variance (ANOVA)

**Code:**

anova_result <- aov(EE ~ WalkingStyle, data = data_long)
summary(anova_result)

**Output:**

```
> anova_result <- aov(EE ~ WalkingStyle, data = data_long)
> summary(anova_result)
             Df Sum Sq Mean Sq F value Pr(>F)    
WalkingStyle  2 0.7528  0.3764     672 <2e-16 ***
Residuals    36 0.0202  0.0006                   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Interpretation:**

- The results from the Analysis of Variance (ANOVA) on the energy expenditure (EE) across different walking styles are highly significant, with an F-value of 672 and a p-value less than 0.0001. This outcome allows us to reject the null hypothesis, which stated that the mean energy expenditures for all walking styles are equal.

The significant p-value indicates that there are indeed differences in the average energy expenditures among the walking styles tested—Usual Walk, Teabag Walk, and Putey Walk.

## Task 3: Multiple Linear Regression

**Code:**

teabag_data <- data_long[data_long$WalkingStyle == "Teabag Walk", ]
model <- lm(EE ~ BMI + Gender, data = teabag_data)
summary(model)

**Output:**

```
> teabag_data <- data_long[data_long$WalkingStyle == "Teabag Walk", ]
> model <- lm(EE ~ BMI + Gender, data = teabag_data)
> summary(model)

Call:
lm(formula = EE ~ BMI + Gender, data = teabag_data)

Residuals:
      Min        1Q    Median        3Q       Max
-0.032312 -0.013699  0.003413  0.007208  0.035670

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.170981   0.038727   4.415  0.00130 **
BMI          0.007524   0.001601   4.701  0.00084 ***
Gender      -0.019904   0.012811  -1.554  0.15133
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02039 on 10 degrees of freedom
Multiple R-squared:  0.6933,	Adjusted R-squared:  0.632
F-statistic:  11.3 on 2 and 10 DF,  p-value: 0.002712
```

**Interpretation:**

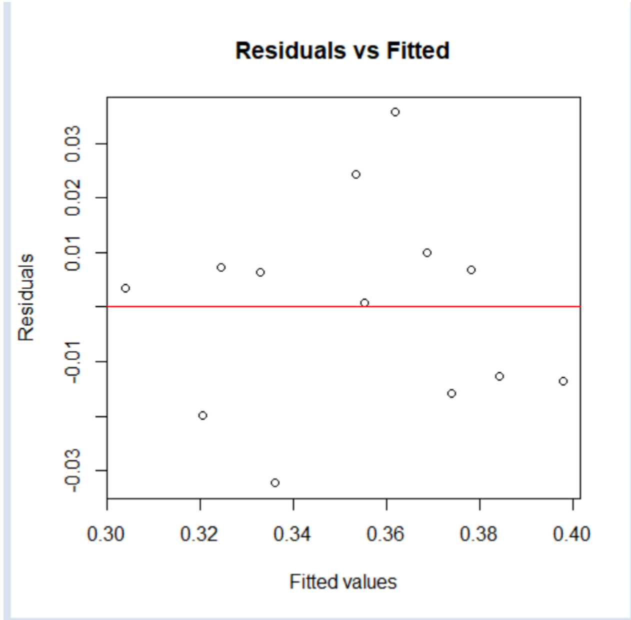
- Body Mass Index (BMI) significantly impacts energy expenditure during the Teabag Walk, with each unit increase in BMI resulting in a corresponding increase in energy expenditure of 0.007524 kcal/kg/min. This highlights the importance of body composition in determining the energy cost of exaggerated walking styles.
- Gender does not significantly influence energy expenditure for the Teabag Walk when controlling for BMI.
- The model explains approximately 69.33% of the variance in energy expenditure, as indicated by the R-squared value, and is statistically significant with an F-statistic of 11.3 (p-value = 0.002712). This strong model fit suggests that BMI is a robust predictor of energy expenditure for this specific walking style.

## Task 4: Residual Analysis

**Code:**

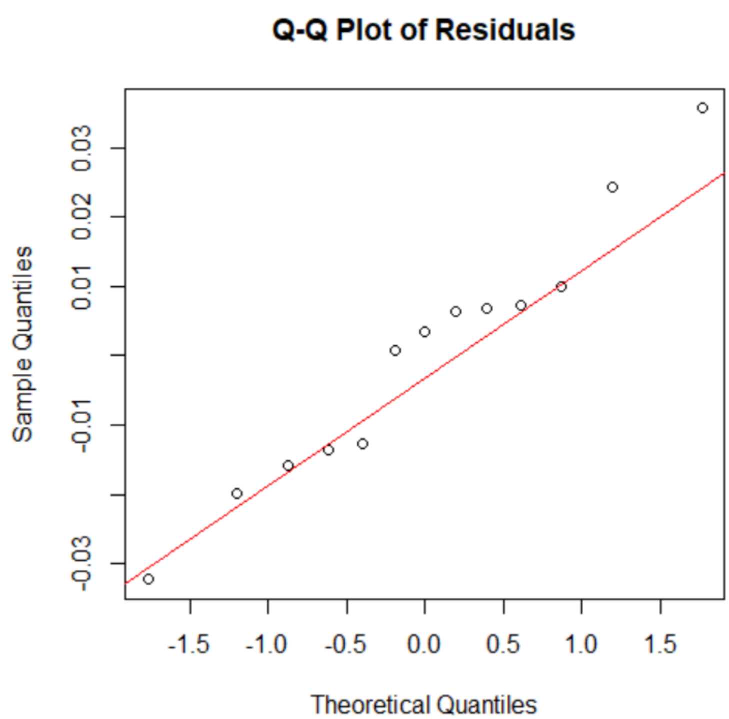
Residual plot

```
plot(model$residuals ~ fitted(model),
    main = "Residuals vs Fitted",
    xlab = "Fitted values",
    ylab = "Residuals")
abline(h = 0, col = "red")
```




Q-Q plot of residuals

qqnorm(model$residuals, main = "Q-Q Plot of Residuals")
qqline(model$residuals, col = "red")




**Interpretation:**

- The regression model for predicting energy expenditure based on BMI and gender in the Teabag Walk is reliable.

**Task 5: Report**


# Energetic Silly Walks: Analyzing the Impact of Walking Styles on Energy Expenditure

Exploring how walking styles affect energy levels.

**Nafisat Ibrahim**
**Tuesday, May 7th 2024**

## Abstract

This report provides a comprehensive statistical analysis of the energy expenditure associated with different walking styles, particularly those exemplified in Monty Python's "Ministry of Silly Walks." Utilizing data from an experimental study, we analyzed the energy costs of three distinct walking styles: Usual Walk, Teabag Walk, and Putey Walk. The analyses employed included boxplots for visual comparison, ANOVA for testing differences in means, and multiple linear regression to assess the influence of Body Mass Index (BMI) and gender on energy expenditure.

## Introduction

Inspired by Monty Python’s famous "Ministry of Silly Walks" sketch, this study takes a closer look at how different funny walking styles can affect the amount of energy we use. Those silly walks raise interesting questions about the physical effects of unusual movements. This report aims to measure the energy spent on these exaggerated walking styles.

## Methods

In this study, we analyzed a dataset inspired by the ‘Quantifying the benefits of inefficient walking’ article from the British Medical Journal. Data were constructed based on the experimental setups described in the article, focusing on adults who performed three different walking trials (Gaesser et al., 2022). The variables analyzed included energy expenditure (EE) in kcal/kg/min, walking styles (Usual Walk, Teabag Walk, Putey Walk), Body Mass Index (BMI), and gender.

Statistical analyses were performed using R. Initially, boxplots were created to visually compare EE across the three walking styles. Subsequently, an Analysis of Variance (ANOVA) was conducted to test for significant differences in EE among these styles. To further explore the effects of BMI and gender on EE, multiple linear regression was applied. Plots, including residual plots and Q-Q plots, were used to assess the appropriateness of the regression model.

## Results

First of all, the boxplots for energy expenditure by walking style showed distinct variations between the groups. The Teabag Walk had the highest median energy expenditure, significantly surpassing the Usual Walk and Putey Walk. This was visually evident from the longer upper quartile, suggesting a higher range of energy used during the Teabag Walk (see Figure 1). The lack of outliers in the plots indicates that the data were consistently distributed within each walking style.

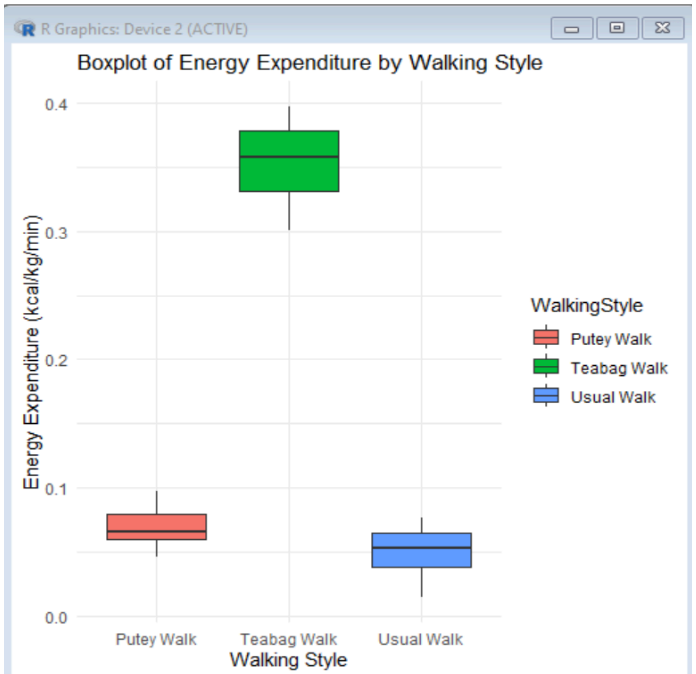



Figure 1: Boxplot of Energy Expenditure by Walking Style

Secondly, the Analysis of Variance demonstrated a significant difference in energy expenditures among the walking styles, with an F-statistic of 672 and a p-value less than 0.0001, strongly suggesting that not all walking styles consume energy equivalently.

Then, the multiple linear regression analysis focused on the Teabag Walk style, assessing the impact of BMI and gender on energy expenditure. The model demonstrated that BMI is a significant predictor, with each unit increase in BMI associated with an increase in energy expenditure by approximately 0.0075 kcal/kg/min ($p = 0.00084$). Conversely, gender did not significantly impact energy expenditure ($p = 0.151$), suggesting no notable difference between males and females in this context. The overall fit of the model was robust, with an R-squared of 0.6933, indicating that about 69.33% of the variability in energy expenditure is explained by the model.

Finally, evaluations of the regression model included a residuals vs. fitted values plot and a Q-Q plot. The residuals vs. fitted values plot showed a random dispersion of points around the horizontal line at zero (see Figure 2). And, the Q-Q plot supported the assumption of normality in the residuals, with points largely aligning along the theoretical line, except for minor deviations at the tails (see Figure 3).

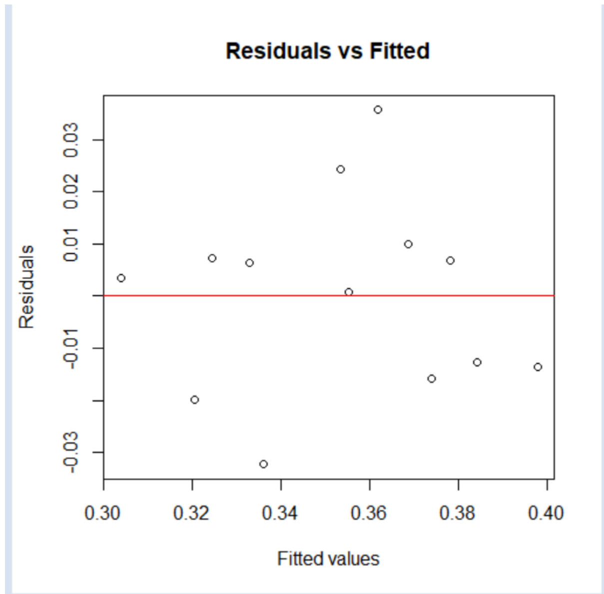



Figure 2: Residual vs Fitted

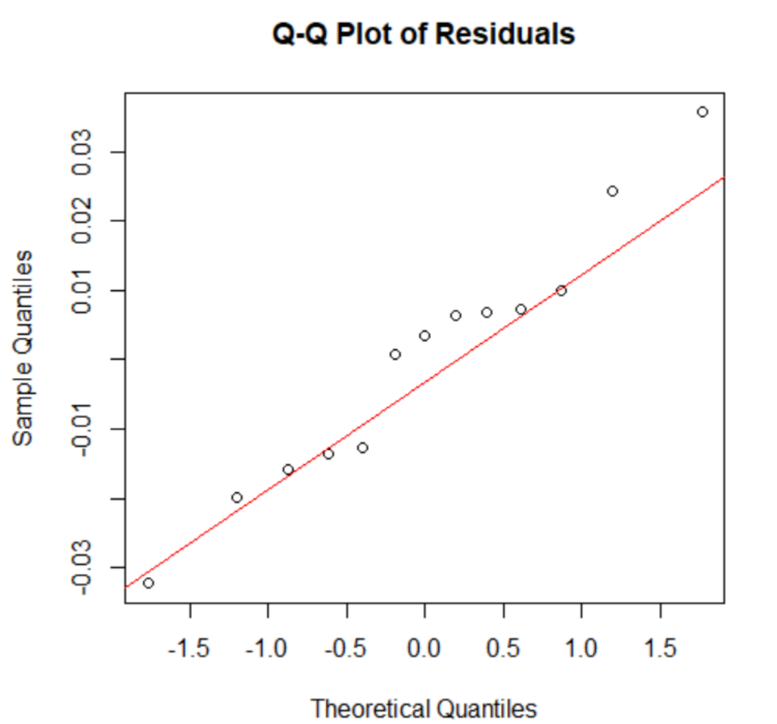



Figure 3: Q-Q Plot of Residuals

## Discussion

This study shows that walking in a silly, exaggerated way, like the Teabag Walk from Monty Python's sketch, really does use more energy than normal walking. This is pretty important for understanding how unusual movements can be tiring and what that might mean for people's health or fitness. It's also interesting to see that someone's body size and shape can change how much energy they use during these kinds of activities.

These findings are useful for anyone whose job involves moving in unique ways—like actors or performers—and could even help physical therapists plan better rehab exercises. However, since the study used a small and specific set of data, we have to be careful about saying this is true for everyone.

Going forward, we could dig deeper into how other physical traits affect energy use in weird walking styles. More studies with different kinds of people and in various places would help confirm these results and make them more useful for everyone.

## Conclusion

In this study, we learned a lot about how different ways of walking can use up different amounts of energy. The silly, exaggerated walks like the Teabag Walk use more energy than normal walking, and how much energy you use can depend a lot on your body size, as shown by the BMI results. However, whether you're a man or a woman doesn't seem to make a big difference in this case.

## Reference

1. Gaesser, G. A., Poole, D. C., & Angadi, S. S. (2022). Quantifying the benefits of inefficient walking: Monty Python inspired laboratory based experimental study. BMJ, e072833. https://doi.org/10.1136/bmj-2022-072833