Exploring Exercise Patterns and Their Relationships with Physiological Metrics: A Visual Analytics Approach

Raisson Leal Silva and Pedro Luís Azevedo Costa
Faculty of Electrical and Computer Engineering at the State University of Campinas
University of Campinas (UNICAMP)
Emails: raissonls@hotmail.com, pdcost@icloud.com

***Abstract*—This project investigates the relationship between exercise patterns and physiological metrics of gym members using visual analytics and statistical methods. Leveraging a simulated dataset, we address three key research questions: (1) Which type of exercise burns more calories: aerobic or anaerobic? (2) Does BMI influence workout type choice? (3) Is weekly workout frequency associated with higher calories burned per session? Exploratory data analysis, hypothesis testing [1], and regression modeling [2] reveal actionable insights to support personalized fitness optimization.**

I. Introduction

The intersection of exercise patterns and physiological metrics offers valuable opportunities for optimizing fitness strategies. This study explores three pivotal questions:

1) Which type of exercise burns more calories: aerobic or anaerobic?
2) Does BMI influence workout type choice?
3) Is weekly workout frequency associated with higher calories burned per session?

Using a comprehensive simulated dataset, we employ exploratory data analysis (EDA), hypothesis testing [1], and regression modeling [2] to uncover data-driven insights. The ultimate goal is to inform personalized workout recommendations that enhance fitness outcomes.

II. Methods

A. Dataset

The Gym Members Exercise Dataset [3] includes 973 samples with detailed attributes encompassing:

- **Demographics:** Age, gender, weight, and height.
- **Workout Details:** Type, frequency, duration, and calories burned.
- **Physiology:** BMI, heart rate metrics, body fat percentage, and water intake.

B. Analytical Approach

The analysis followed these steps for each research question:

1) Research Question 1: Which Type of Exercise Burns More Calories?: To evaluate whether aerobic exercises burn more calories than anaerobic exercises, we conducted a **Permutation Test**, a non-parametric approach ideal for comparing groups without assuming a normal distribution [1]:

- **Test Statistic:** The observed difference in mean calories burned between aerobic and anaerobic exercises.
- **Hypotheses:**
  - H_0: No difference in calories burned between aerobic and anaerobic exercises.
  - H_1: Aerobic exercises burn more calories than anaerobic exercises.
- **Permutation Process:** Pooling calorie data, shuffling labels, and calculating test statistics for 2,000 permutations.
- **p-value:** The proportion of permuted test statistics greater than or equal to the observed value.

2) Research Question 2: Does BMI Influence Workout Type Choice?: To investigate the relationship between BMI and workout type:

- **Visual Analysis:** Box plots, density plots, histograms, and Q-Q plots were used to assess distribution trends and normality.
- **Hypothesis Testing:** An ANOVA test was performed to evaluate significant differences in BMI across workout types [4]. Levene's test confirmed the assumption of homogeneity of variances [5].

3) Research Question 3: Is Weekly Workout Frequency Associated with Calories Burned?: To explore the association between workout frequency and calories burned:

- **Visual Analysis:** Scatter plots, box plots, and density plots were created to identify trends.
- **Correlation Analysis:** Pearson correlation coefficient was calculated to assess the relationship [6].
- **Linear Regression:** A regression model was fitted to quantify the impact of workout frequency on calories burned [2].

Fig. 1: Calories burned distribution for aerobic and anaerobic exercises.

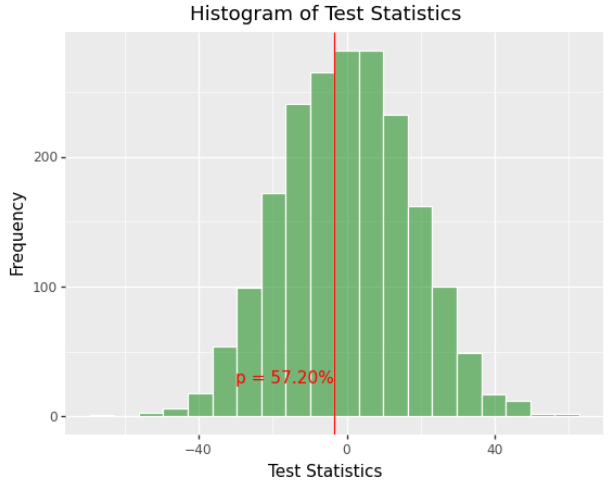

Fig. 2: Histogram of permuted test statistics with observed value highlighted.

III. Results

A. Research Question 1: Which Type of Exercise Burns More Calories?

The permutation test revealed no statistically significant difference in calories burned between aerobic and anaerobic exercises ($p = 0.572$). Both exercise types exhibited similar distributions of calories burned (Fig. 1). The histogram of permuted test statistics with the observed value highlighted (Fig. 2) confirms this result.

B. Research Question 2: Does BMI Influence Workout Type Choice?

ANOVA analysis returned a p-value of 0.38, indicating no statistically significant differences in BMI across workout types [4]. Visual analysis corroborated this result with overlapping BMI distributions for all workout categories (Fig. 3). The

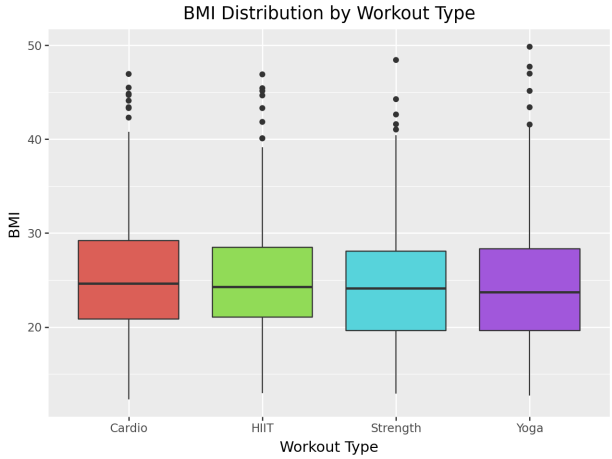

Fig. 3: Box plot of BMI distribution by workout type.

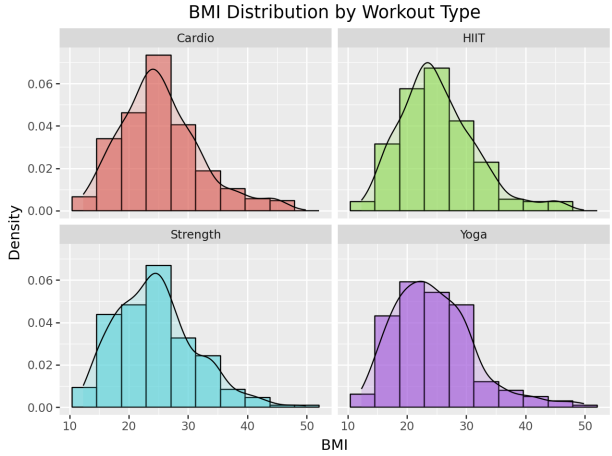

Fig. 4: Histogram of BMI distribution by workout type.

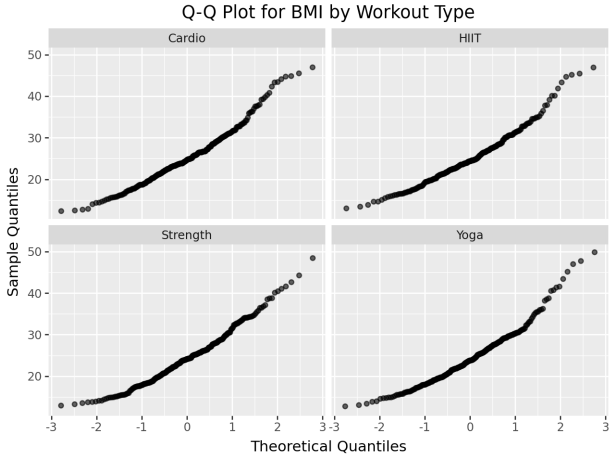

Fig. 5: Q-Q plot of BMI distribution by workout type.

histograms (Fig. 4) and Q-Q plots (Fig. 5) further confirmed that BMI distributions approximate normality.

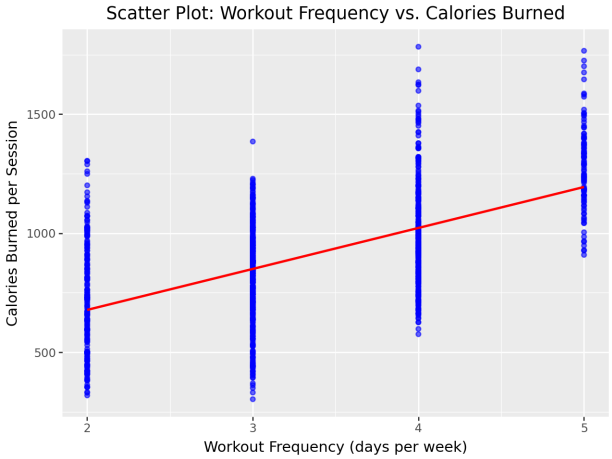

Fig. 6: Scatter plot: Workout frequency vs. calories burned.

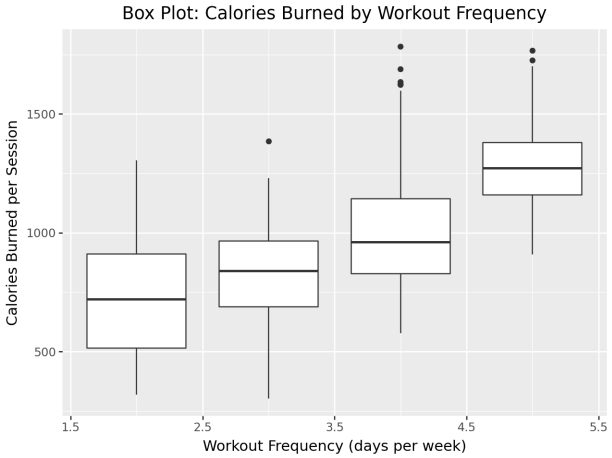

Fig. 7: Box plot: Calories burned by workout frequency.

C. Research Question 3: Is Weekly Workout Frequency Associated with Calories Burned?

Linear regression analysis revealed a moderate positive relationship between workout frequency and calories burned ($r = 0.58, p < 0.001$) [2]. Each additional workout day was associated with an increase of 172 calories burned per session, with the model explaining 33.2% of the variance ($R^2 = 0.332$). The scatter plot (Fig. 6), box plot (Fig. 7), and residuals plot (Fig. 8) illustrate these findings.

IV. Discussion

Key insights from the analysis include:

- **Exercise Type and Calories Burned:** Both aerobic and anaerobic exercises are equally effective in calorie expenditure, suggesting that factors like intensity and duration are more critical determinants.
- **BMI and Workout Choice:** BMI does not significantly influence workout type selection. Individual goals and preferences likely play a larger role in workout decisions.
- **Workout Frequency and Calories Burned:** Consistent training significantly improves calorie burn per session, emphasizing the importance of workout frequency for achieving fitness goals.

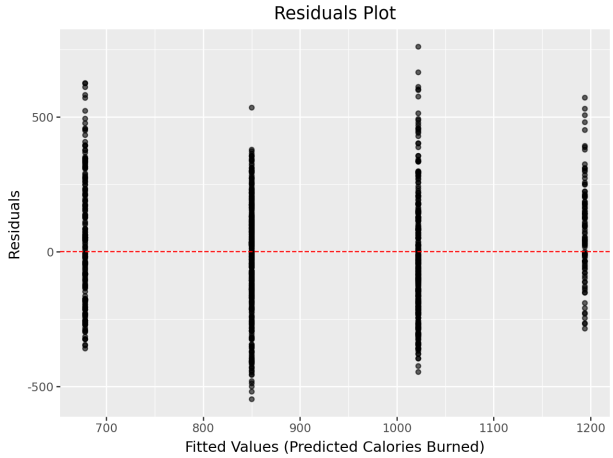

Fig. 8: Residuals plot: Fitted values vs. residuals.

V. Conclusion

This study demonstrates that while exercise type and BMI have limited influence on calorie burn, workout frequency is a critical factor. These findings can inform personalized training strategies and highlight the importance of consistent exercise. Future research could examine additional variables such as exercise intensity and psychological factors to further enhance insights.

Acknowledgments

The authors extend their heartfelt gratitude to Profa. Wu, Shin, the instructor of the IA376M course at UNICAMP, for her invaluable guidance and mentorship throughout this project. We also thank our peers for their insightful feedback and collaborative spirit, which greatly enriched our learning experience.

References

[1] P. Good, *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Springer Science & Business Media, 2013.

[2] X. Su, X. Yan, and C.-L. Tsai, "Linear regression," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 3, pp. 275–294, 2012.

[3] V. Khorasani, "Gym members exercise dataset," https://www.kaggle.com/datasets/valakhorasani/gym-members-exercise-dataset, 2024, accessed: 2024-11-10.

[4] L. St, S. Wold *et al.*, "Analysis of variance (anova)," *Chemometrics and intelligent laboratory systems*, vol. 6, no. 4, pp. 259–272, 1989.

[5] J. L. Gastwirth, Y. R. Gel, and W. Miao, "The impact of levene's test of equality of variances on statistical theory and practice," *Statistical Science*, vol. 24, no. 3, pp. 343–360, 2009.

[6] I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," *Noise reduction in speech processing*, pp. 1–4, 2009.