Experiment Report 2

University of Tsukuba
Experiment Design in Computer Sciences

Name:	Abdelrahman I. A. Abushbeka
Student Number:	202520711
Experiment Title:	The Relative Impact of Distance and Elevation on Caloric Expenditure in Cycling

1 Introduction: Background, Motivation, and a Question!

As an avid cycler, this is an experiment that seeks to answer a practical question I'm genuinely interested in when it comes to utilizing cycling as my main working out method, a question that I personally face quite often: "if I'm short on time and in need to burn a goal calories I set for myself, what matters more in this case, **distance of cycling, or elevation of cycling?** Which workout strategy is more sufficient, for example, a longer & flatter route or a shorter & steeper one? Does the terrain profile of a cycling workout significantly affect the efficiency of caloric expenditure?". Simply put, the core motivation for this experiment stems from a practical, real-world dilemma that I face quite often due to the number of responsibilities I have to deal with: how to achieve maximum fitness benefits through cycling within a limited timeframe.

The most important terminology for this experiment that I will use quite often is "caloric expenditure", which simply means the total number of calories that the body burns throughout the day (or throughout an exercise). It is a primary indicator of workout intensity and a key component of cardiovascular health maintenance. That is the focus of the experiment in a nutshell.

The primary objective is to quantify and compare the relative impact of these two fundamental workout components, distance and elevation. To achieve this, this study is designed as a controlled single-subject (me) experiment utilizing a Factorial Design. Our analysis is conducted using an **additive effects model**, which is a form of **Multiple Linear Regression** model (1), used to analyze the main effects of two continuous factors: Distance (km) and Elevation Up (m) on the response variable, Active Calories (Kcal). This statistical approach, which aligns with the principles of analyzing interaction effects, allows us to determine the independent caloric "value" of each kilometer cycled versus each meter climbed.

My goal behind such a type of exercise is to keep my heart health at check. I don't have specific workout out goals such as "weight loss" or "muscle strength"; I just want a well-rounded daily fitness regimen that keeps my cardiovascular and general health in check. From my research, I learned that burning ˜200 active calories per day (˜30 minutes of moderate aerobic activity or moderate cycling) supports heart health and general fitness, according to the American Heart Association (AHA) (2). That's why I chose 200 active calories as a goal to be burned for my cycling experiment. While this target is not used directly in the regression calculation, it serves as the motivational framework and will be used in the discussion part to translate the statistical findings into practical advice for me and for the reader.

I have to make a difference between "**Active Calories**" and just "**Calories**", as the human body burns calories by itself throughout the day, so those calories burned naturally are unaccounted for. From now on when I refer to calories I mean "Active Calories", which are the calories that we burn during an active exercise, in other words, calories burned with a conscious decision.

By applying Null Hypothesis Statistical Testing (NHST) and a multiple linear regression model, we will analyze which workout factor significantly impacts caloric expenditure.

2 Experiment Design

The design of this experiment is structured to provide a statistical answer to the central research question: which is more efficient for caloric expenditure in cycling, covering distance or gaining elevation? To investigate this, the experiment is designed as a controlled experiment, analyzing data from a single subject to minimize inter-individual variability.

The design moves beyond a simple comparison of ride "types" and instead employs a Factorial Design. The relationship between the factors and the response variable is analyzed using an additive effects model in the form of a Multiple Linear Regression (1). We utilize interaction effects analysis to learn the distinct and interactive contributions of both distance and elevation to caloric burn within each workout session, which allows us a precise quantification of the main effects of each factor.

2.1 The Variables

The experiment is defined by the following components:

- **Dependent Variable (Response):** Active Calories (Kcal). This is the primary continuous outcome we aim to predict.
- **Independent Variables (Factors):** These are the two primary components of a cycling workout whose influence I want to quantify.
  - Distance (km): A continuous variable representing the total distance covered.
  - Elevation Up (m): A continuous variable representing the total vertical meters gained.

2.2 The Model

This linear model allows us to translate our initial intuitive question about "Short & Steep" vs. "Long & Flat" into a precise, falsifiable statistical hypothesis comparing the two coefficients. To directly compare the efficiency of distance versus elevation, we will use an additive effects model. The statistical model used to test the relationship is:

$$\text{Active Calories} = \beta_0 + \beta_1(\text{Distance}) + \beta_2(\text{Elevation Up}) + \epsilon \tag{1}$$

In this statistical model:

- $\beta_0 \rightarrow$ intercept: The baseline caloric burn.
- $\beta_1 \rightarrow$ caloric efficiency of distance: The coefficient for Distance. It's the average additional Kcal burned for every kilometer cycled, holding elevation constant.
- $\beta_2 \rightarrow$ caloric efficiency of elevation: The coefficient for Elevation Up. It represents the average additional Kcal burned for every meter of elevation gained, holding distance constant (in reality holding distance constant during elevation gain with a bike is impossible, but let's assume that for the sake of the experiment).
- $\epsilon \rightarrow$ The residual error (unexplained variability).

2.3 The Hypotheses

My formulation of a falsifiable hypothesis started with the research question: Is there a statistically significant difference between the caloric burn efficiency of covering one kilometer of distance versus gaining one meter of elevation? Thus my null hypothesis and alternative hypothesis were the following:

- **Null Hypothesis (H_0):** There is no statistically significant difference in the caloric contribution of one kilometer of distance versus one meter of elevation gain. The effects of the two factors are statistically equivalent.
$$H_0 : \beta_1 = \beta_2 \tag{2}$$
- **Alternative Hypothesis (H_1):** There is a statistically significant difference in the caloric contribution of one kilometer of distance versus one meter of elevation gain.
$$H_1 : \beta_1 \neq \beta_2 \tag{3}$$

To test for potential complexities in this relationship, a second model including an interaction term (Distance * Elevation) was also fitted and evaluated as part of the analysis.

2.4 Data Protocol and Bias Management

The protocol established to gather data, ensuring its quality, and minimizing sources of error (bias) goes as the following:

- **Data Source and Equipment:** All data originates from a single subject (myself) using a consistent set of equipment (one bicycle, one Apple Watch Series 7). Data was automatically logged during "Outdoor Cycling" activities and exported from the Apple Health app as a JSON file.

- **Data Collection & Trial Definition:** Since this is a Factorial Design problem, the data was not collected following a single, repeated protocol. Rather, I collected data from my daily life, specifically my routes to university and to my part time job, trying out different routes with different variations in distance and elevation to fill out the appropriate data spectrum needed for a multiple regression model. A problem that I faced was gathering "short distance & high elevation" data points. So I established a protocol to gather this specific data type: I biked up and down a hill specified in Figure 1 multiple times, maintaining the same conditions for every ride. After I gathered this specific data type, I was able to continue with my statistical inference calculations.

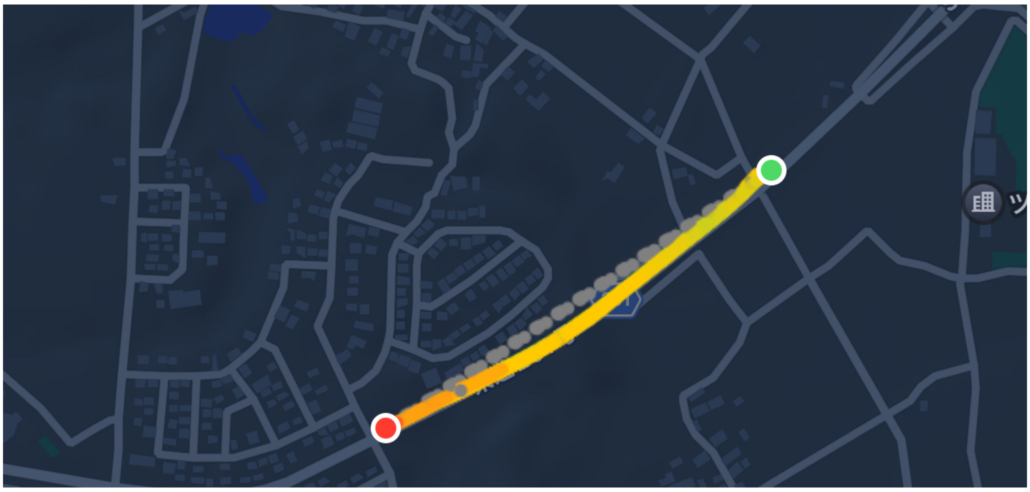

Figure 1: Map segment of the route used for collecting supplementary "Short & Steep" data points via hill repeats.

- **Data Filtering:** An experimental trial is defined as a single, complete cycling session. The raw data was pre-processed using a filtering protocol. A trial was included in the final dataset only if it met all of the following criteria: (1) Activity Type: Must be "Outdoor Cycling", not an indoor one. (2) Non-Zero Metrics: Must have a recorded Distance >0 km and Elevation Up >0 m. (3) Minimum Effort: Must have a Duration >5 minutes and Active Calories >10 Kcal to exclude non-representative sessions.

- **Bias and Nuisance Factor Control:** Instrumentation Bias was controlled by using the same calibrated devices. Inter-Subject Variability was entirely eliminated by using a single subject for all trials. Confounding Variables such as daily fatigue or weather are mitigated by analyzing a diverse dataset of rides collected over a two-month period, which helps average out random effects.

3 A Priori Analysis and Sample Size Determination

An a priori power analysis was conducted to establish a target sample size, to ensure the experiment possesses enough statistical power to confidently detect a meaningful relationship between the predictor variables and the response variable, and to minimize the risk of a Type II error.

The analysis was structured for a multiple linear regression model for these parameters:

- **Significance Level (α):** Set to 0.05, corresponding to a 95% confidence level.
- **Desired Statistical Power ($1 - \beta$):** Set to 0.80, to ensure an 80% probability of detecting a true effect if one exists.
- **Minimum Relevant Effect Size (f^2):** We targeted a medium effect size $f^2 = 0.15$.
- **Number of predictors:** Two (Distance and Elevation Up).

Using a statistical power calculator for a multiple regression F-test (3), the analysis concluded that the required sample size to achieve the desired power and significance is approximately $\mathbf{n = 68}$ valid workout sessions, out target for data collection.

4 Results and Analysis

Following the data protocol, the raw data from 74 workout sessions was filtered, resulting in a final analysis dataset of $n = 55$ valid "Outdoor Cycling" sessions. First we examine the descriptive statistics (Mean, Std. Dev., Min, Max), then post-hoc power analysis, checking model assumptions, and finally we perform the statistical inference.

4.1 Descriptive Statistics and Final Dataset

The final analysis dataset consists of $n = 55$ valid observations. The descriptive statistics for this dataset are presented in Table 1. A visual inspection via a scatter plot matrix (Figure 2) confirms positive correlations between the variables, which supports the use of a linear regression model.

Table 1: Descriptive Statistics of the Final Dataset (n=55)

Metric	**Mean**	**Std. Dev.**	**Min**	**Max**
Active Calories (Kcal)	527.89	372.96	115.19	1632.32
Distance (km)	5.68	3.78	0.90	15.70
Elevation Up (m)	17.47	20.19	2.00	107.78

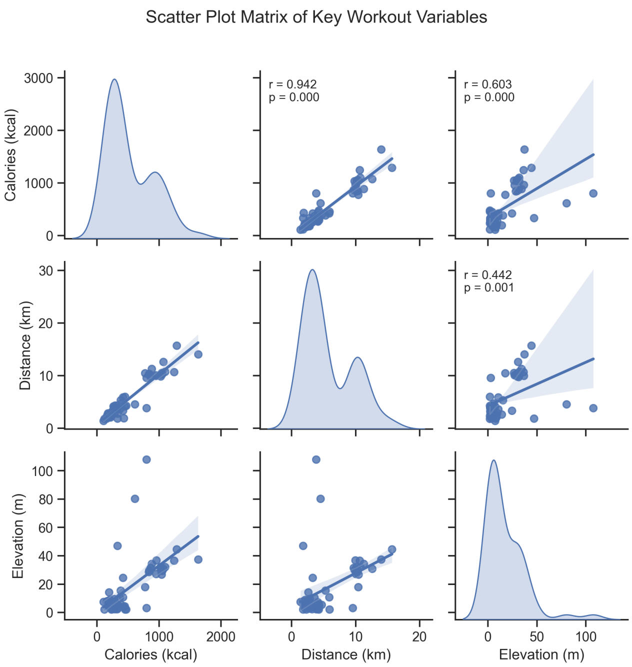

Figure 2: Scatter plot matrix showing relationships between key variables. The upper triangle displays the Pearson correlation coefficient (r) and its p-value.

4.2 Post-Hoc Statistical Power Analysis

The a priori analysis established a target sample size of $n \approx 68$. Our final sample size is $n = 55$. To determine the actual power of the experiment we perform a post-hoc power analysis. The model's summary R-squared value ($R^2 = 0.931$) was used to derive Cohen's f^2 via the formula $f^2 = R^2/(1 - R^2)$. This corresponds to a large effect size of $f^2 = 13.487$. Based on this, the post-hoc analysis: **Actual Power** $(1 - \beta)$ **>0.9999**.

4.3 Model Assumption Checking

The residuals of the model were analyzed as follows:

- **Normality:** A Shapiro-Wilk test yielded a p-value = 0.0134. As this is below $\alpha = 0.05$, we reject the null hypothesis of normality. This indicates a minor but significant deviation from a normal distribution, visually confirmed by the Q-Q plot in Figure 3.

- **Homoscedasticity:** A plot of residuals versus fitted values (Figure 4) showed a random scatter, satisfying the assumption of equal variance.
- **Independence:** This assumption is satisfied by the experimental design, where in each data point represents a distinct workout session.

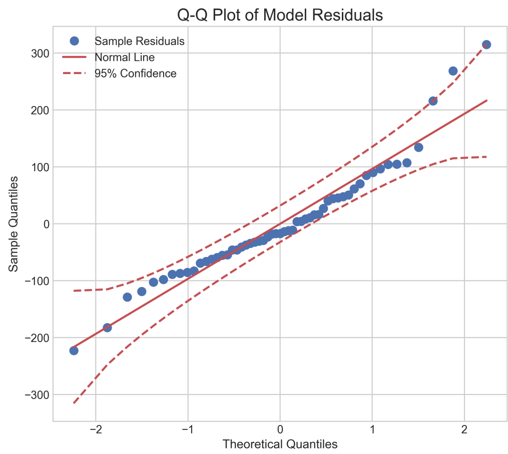

Figure 3: Q-Q plot of model residuals, showing slight deviation from normality at the tails.

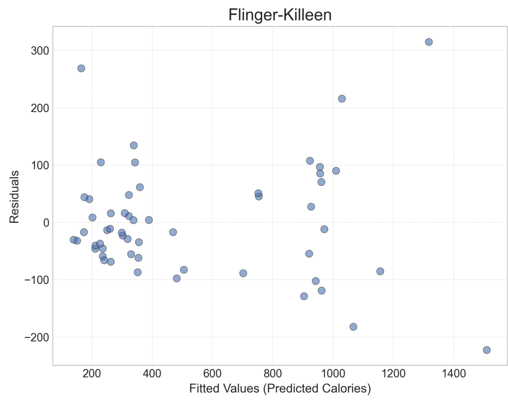

Figure 4: Plot of residuals versus fitted values, showing no discernible pattern, which supports the assumption of homoscedasticity.

4.4 Statistical Inference: Model Analysis and Hypothesis Testing

A two-stage modeling process was employed. First, a main effects model (‘Calories $\sim$ Distance + Elevation‘) was fitted. It has an R-squared of 0.928. A formal F-test on the primary hypothesis, $H_0 : \beta_1 = \beta_2$, resulted in a p-value of 7.34×10^{-24}.

Second, to test for non-additive effects, a full factorial model (‘Calories $\sim$ Distance * Elevation‘) was fitted. The p-value for the ‘Distance:Elevation‘ interaction term was 0.489, which is sad, it’s much higher than 0.05, higher than I expected.

Model Selection and Interpretation: As the interaction term was not statistically significant ($p > 0.05$), the simpler **main effects model is selected as the definitive model for interpretation**. The rejection of the null hypothesis in this model is therefore the primary finding. The data shows that one kilometer of distance (coefficient $\beta_1 \approx 82.8$) contributes approximately **19.2 times more** (82.8/4.3) to the total caloric expenditure than one meter of elevation gain (coefficient $\beta_2 \approx 4.3$).

5 Discussion

”is it more efficient to pursue distance or elevation to maximize caloric expenditure?”, our primary objective was to answer that question. Through a controlled factorial design analyzed with a multiple linear regression model, the study resulted in a clear statistical answer, the most **surprising finding** is the huge disparity in the caloric efficiency of the two factors. The analysis rejected the null hypothesis ($H_0 : \beta_1 = \beta_2$) of equal effect ($p = 7.34 \times 10^{-24}$), indicating for this subject, the contribution of distance to caloric burn vastly outweighs that of elevation. The model's coefficients quantify this relationship, that cycling one kilometer is approximately 19.2 times more calorically demanding than climbing one meter. This result changed my intuition that intense steep hill climbs are the most potent form of workout.

The **unsurprising finding** is that both factors are statistically independent predictors of calories burned. The main effects model demonstrated a fit, with an R-squared of 0.928. meanwhile the interaction model is non-significant ($p = 0.489$). This means the primary conclusion is strong, suggesting that the ∼19-to-1 efficiency ratio holds true regardless of whether a ride is flat or hilly; **the effects are simply additive**.

5.1 Limitations of the Experiment

The most significant limitation is the single-subject design. While it limits internal validity by controlling for all inter-individual differences, it limits the external validity (generalization) of the findings. The 19 to 1 efficiency ratio is personalized and may differ for athletes of different weights, fitness levels, or cycling styles.

6 Conclusion

The results provide a clear takeaway: while both factors contribute to a workout's intensity, **distance is the overwhelmingly dominant driver of caloric expenditure**. For a cyclist with limited time and a specific caloric goal such as the 200 Kcal benchmark for daily cardiovascular health, the most efficient strategy is to prioritize covering more ground. To burn 200 calories, this subject would need to either cycle approximately **2.4 kilometers** (200/82.8) on a flat route or climb **46.5 meters** (200/4.3) vertically. The superior efficiency is for distance-based effort.

References

[1] Bevans, R. (2020). *Multiple Linear Regression — An Introduction.* Scribbr. Retrieved from https://www.scribbr.com/statistics/multiple-linear-regression/

[2] American Heart Association. (2021). *AHA Recommendations for Physical Activity in Adults.* Retrieved from https://www.heart.org/en/healthy-living/fitness/fitness-basics/aha-recs-for-physical-activity-in-adults

[3] Soper, D.S. (2024). *A-priori Sample Size Calculator for Multiple Regression* [Software]. Available from https://www.danielsoper.com/statcalc/calculator.aspx?id=1