**Triathlon Multiple Linear Regression Worksheet**

**Introduction:**

The Triathlon Multiple Linear Regression worksheet focuses on analyzing the relationships and predictive capabilities of multiple variables in the context of triathlon performance. Specifically, we will explore the prediction of run times using swim times and bike times as predictors. By employing simple regression models, checking conditions with Residuals vs. Fits plots, and conducting hypothesis tests and confidence intervals, we aim to identify significant predictors and understand their contextual implications. The dataset used for this analysis comprises the 2022 Canadian finishers of the Lake Placid Ironman.

**Learning Objectives:**

1. Understand the concept of multiple linear regression and its application in analyzing sports data.

2. Recognize the importance of checking assumptions, such as assessing residuals vs. fits plots, for validating the regression models.

3. Gain insights into the significance of predictors through hypothesis tests and confidence intervals.

4. Comprehend the interpretation and contextual implications of the findings from multiple linear regression analyses in the triathlon setting.

**Methods:**

To successfully complete this worksheet, students should have prior knowledge of the following statistical concepts:

1. Familiarity with simple linear regression, including interpretation of slopes, intercepts, and residuals.

2. Understanding of hypothesis testing and confidence intervals, particularly in the context of regression analysis.

3. Knowledge of residual analysis, including the interpretation of residual plots.

4. Basic understanding of the triathlon sport and its components (swimming, biking, and running) to better grasp the contextual implications of the statistical analyses.