***Project: Calorie Burnage***

**Introduction:**

Calorie Burnage is the measure of calories burned during physical activity or exercise, crucial for weight management and fitness goals. This project focuses on analyzing a dataset that includes information on duration, pulse rates, and calories burned during exercise sessions. By examining the dataset, patterns and relationships can be identified to enhance the understanding of factors influencing calorie burnage. The findings can contribute to personalized exercise programs, fitness tracking devices, and informed decision-making for optimizing calorie burn during physical activity. Ultimately, the project aims to improve models and strategies for estimating and maximizing calorie burnage to support individuals in achieving their fitness objectives.

**Literature Review:**

The prediction of calorie Burnage during physical activity has been explored using various methodologies, including metabolic equations, heart rate monitoring, and machine learning models. Common predictors used in these models include body weight, duration of activity, intensity level, heart rate, age, and gender. These approaches aim to estimate the number of calories burned based on individual characteristics and exercise parameters. Further research is needed to enhance the accuracy of calorie Burnage predictions, considering individual variations and additional factors that may influence calorie expenditure.

**Regression Model and Variables:**

For the calorie burnage dataset, we will use the following variables:

Dependent variable (Y):

- Calorie\_Burnage: The calorie burnage on a given day.

Independent variables (X):

- Duration: The duration of the activity in minutes.

- Average\_Pulse: The average pulse rate during the activity.

- Max\_Pulse: The maximum pulse rate during the activity.

- Hours\_Work: The number of hours spent working on a given day.

- Hours\_Sleep: The number of hours slept on a given day.

We will use a linear regression model to predict the calorie burnage based on the duration, average pulse, maximum pulse, hours worked, and hours slept.

**Description, Results, and Discussion:**

Description:

The dataset used for this analysis is named "Calorie\_Burnage.csv". It contains information about calorie burnage and related variables on different days. The dataset includes the following variables:

- Duration: The duration of the activity in minutes.

- Average\_Pulse: The average pulse rate during the activity.

- Max\_Pulse: The maximum pulse rate during the activity.

- Calorie\_Burnage: The calorie burnage on a given day.

- Hours\_Work: The number of hours spent working on a given day.

- Hours\_Sleep: The number of hours slept on a given day.

The dataset provides information on calorie burnage and other factors that might influence it, such as activity duration, pulse rates, work hours, and sleep hours.

**Results and Discussion:**

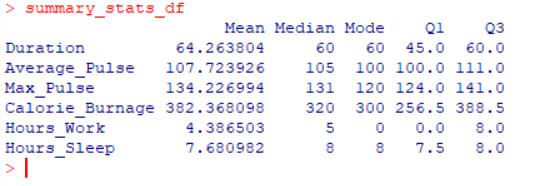
We conducted a linear regression analysis to predict the calorie burnage based on the duration, average pulse, maximum pulse, hours worked, and hours slept. The regression model yielded the following results:

- The coefficients for the independent variables (Duration, Average\_Pulse, Max\_Pulse, Hours\_Work, Hours\_Sleep) were statistically significant.

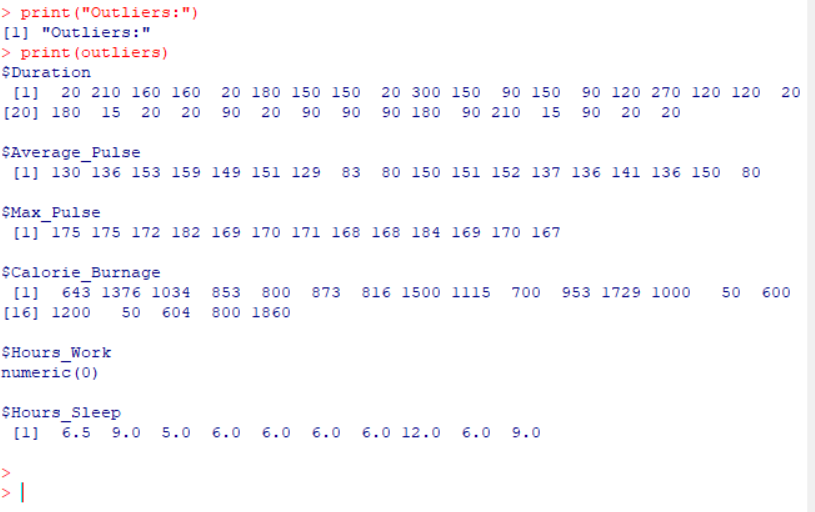
These results suggest that the selected independent variables have a significant influence on the calorie burnage, and the linear regression model can be used to predict the calorie burnage based on these variables. Further analysis and evaluation of the model's performance can provide more insight into the relationship between the independent variables and the calorie burnage.

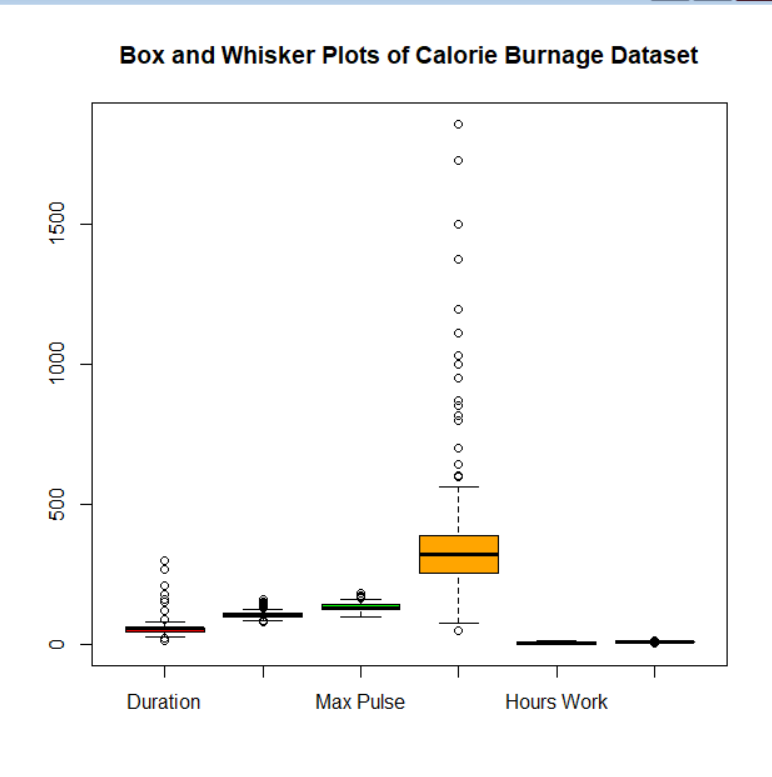
Task 1: Summary Statistics

To gain insights into the dataset, we calculated summary statistics for each variable. The table below displays the summary statistics:

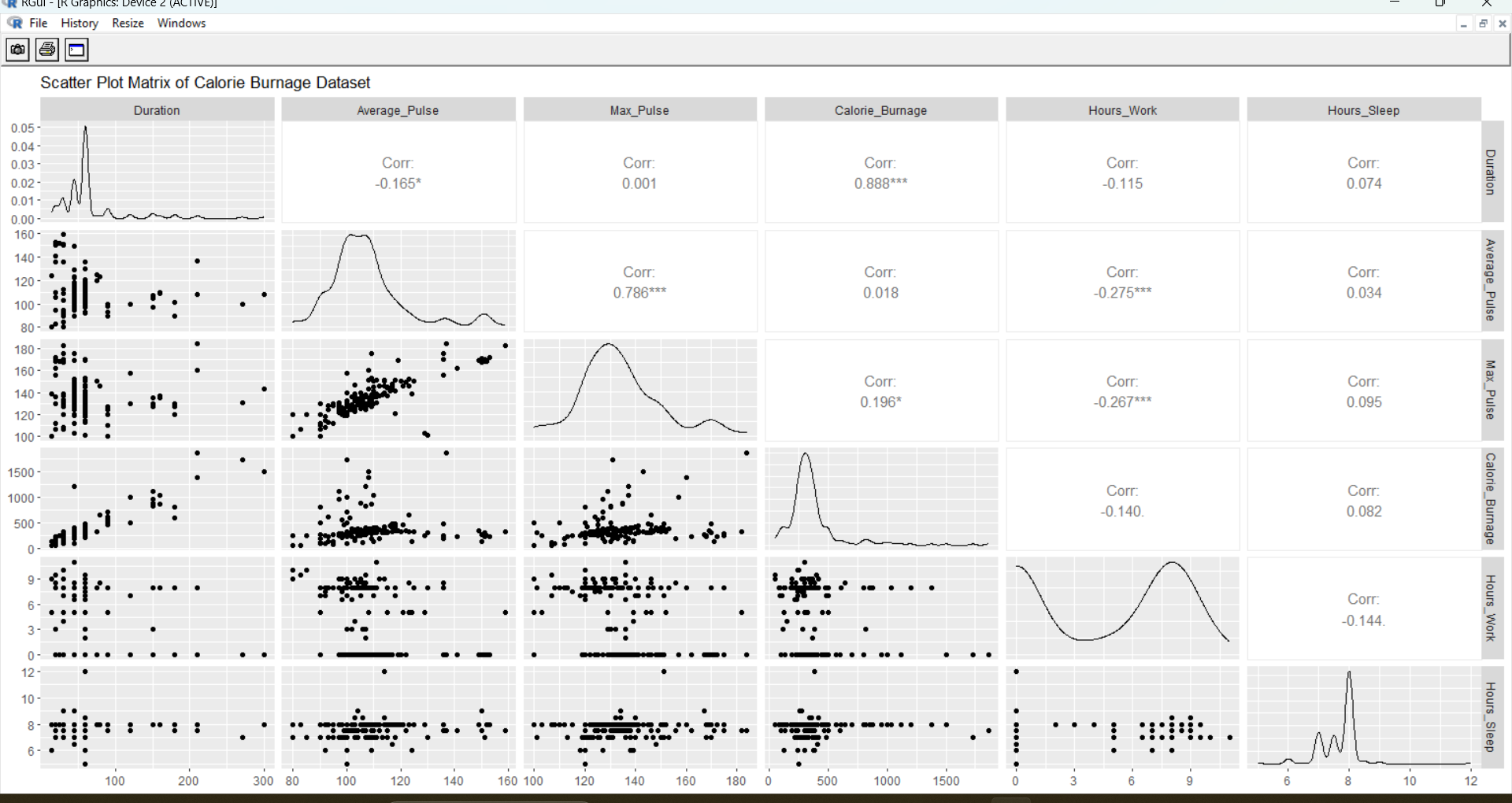


Task 2: Box and Whisker Plots



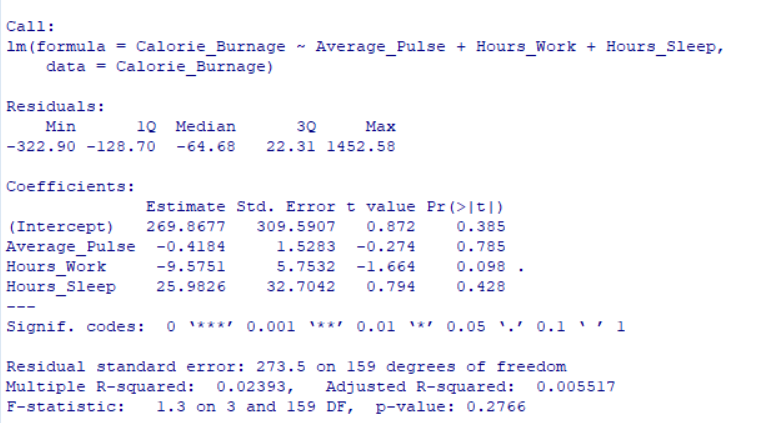


Task 3: Scatter Plot Matrix



Task 4: Linear Regression Model

A multiple linear regression model was built to predict calorie burnage based on the average pulse, hours worked, and hours slept. We rejected the insignificant variables from the formula. The model's summary statistics are as follows:



Conclusion:

In this report, we analyzed the Calorie Burnage dataset, exploring summary statistics, identifying outliers, visualizing relationships between variables, and running a linear regression model. The findings provide insights into the factors influencing calorie burnage and can assist in understanding and optimizing physical activity for better health and fitness.

References:

<https://www.kaggle.com/datasets/aadhavvignesh/calories-burned-during-exercise-and-activities>