Applied Statistics with R | AiQuest | Project 1

Title: Weight Loss Program Analysis

Deadline 16 March 2025

Background

You are a researcher at Titan Research Institute. Recently, you have been presented with a problem. Obesity and overweight conditions have become a major public health challenge worldwide, contributing to an increased risk of chronic diseases such as diabetes, hypertension, cardiovascular diseases, and metabolic disorders. As lifestyle-related factors, including diet, physical activity, and sleep patterns, play a crucial role in weight management, structured weight loss programs have been developed to guide individuals toward healthier living.

Despite numerous diet and exercise programs available, the effectiveness of these programs varies significantly across individuals. This variation can be attributed to multiple factors, such as personal adherence to dietary guidelines, engagement in physical exercise, sleep quality, and demographic characteristics like age and gender. Therefore, a data-driven approach is necessary to better understand the relationship between these factors and weight loss outcomes.

In a recent experiment conducted by your institute, 35 participants voluntarily enrolled in a 12-week weight loss program. The **primary goal** of this study is to assess the effectiveness of lifestyle modifications, particularly diet adherence, exercise intensity, and sleep quality on weight loss outcomes in individuals. For each participant, data was collected on:

1. **Demographics:**
    a. Age (in years)
    b. Gender (Male/Female)
    c. Baseline Weight (Weight before lifestyle modifications - in kilograms)
    d. After Weight (Weight after 12 weeks of weight loss program)

2. **Lifestyle/Behavioral Factors:**
    a. Diet Adherence: A score indicating how closely the participant followed the prescribed diet (on a scale from 6 to 10)
    b. Exercise Intensity: A score representing the intensity of the exercise regimen (on a scale from 5 to 10)
    c. Sleep Quality: A score reflecting the quality of sleep during the program (on a scale from 5 to 10)

Additional Instruction:
Please create a new variable, **Weight Loss (kg)**, by subtracting the *After Weight* from the *Baseline Weight*. This will work as your outcome / dependent variable.

Project Objectives

1. **Data Exploration:**
    - Summarize the dataset by calculating descriptive statistics (mean, standard deviation, etc.) for all variables.
    - Visualize the relationships (where appropriate) among the variables using appropriate plots (e.g., histograms, scatter plots, boxplots).

2. **Statistical Analysis:**
    - Perform a regression analysis to determine how demographics and lifestyle factors predict Weight Loss.
    - Check the assumptions of the regression model (normality, homoscedasticity, linearity, etc.) and report any issues if exists.

3. **Interpretation and Reporting:**
    - Interpret the regression coefficients to explain the effect of each predictor on Weight Loss.
    - Discuss which factors are statistically significant and their practical implications in the context of the weight loss program. Provide a clear and concise conclusion summarizing your findings.

Submission Guidelines

- Submit your R markdown code and knitted HTML file by 16 March 2025.
- Your knitted HTML report should include:
    - Title of the study, your name, and date of submission.
    - Code for data import, exploratory analysis, and visualization.
    - Code for the regression analysis.
    - Inline comments (using #) where necessary.
- Ensure that your code is well-organized, and figures have proper labeling.
- Your report (the knitted HTML file) should include:
    - The purpose of the study.
    - Results/key findings supported by enough statistical evidence.
    - A discussion on the implications of your findings and any limitations of your analysis.
- Cite any external sources or references used in your analysis (any website / books / tutorials).

Evaluation Criteria

- Clarity and depth of exploratory analysis and visualizations.
- Correct application and interpretation of regression analysis.
- Quality of written communication and overall presentation of the project.
- Creativity and critical thinking in discussing the findings.

^_^ Good luck with your project, and ensure that your analysis is both rigorous and clearly presented!

- Md. Ahsanul Islam

Originality is the ability to be able to hide your source.
- Benjamin Franklin