EDA Analysis: Calories burnt in Gym.

Mayuri Ande

Bellevue University

DSC530- Data Exploration And Analysis

Matthew Metzger

11/17/2025

EDA on Gym Membership Dataset

The Gym membership is analyzed to get the insights on the customers attending the gym and the calories burnt based on Age, sessions, BMI

**Key Findings:**

* A Strong positive correlation between session duration and calories burned is observed. This suggests that longer workout sessions tend to result in higher calories. But at the same time, the calories burnt will be affected based on the frequency of the workout. As the frequency of workout increases, it takes more time to burn the same amount calories compared to previous sessions.
* The correlation between age and BMI was found to be very weak, indicating that age alone is not a strong predictor of BMI.
* The distribution of workout frequency suggests the presence of two distinct groups: one group with lower workout frequency and another with higher frequency.
* The distribution of BMI is right-skewed, indicating a larger proportion of individuals with lower BMI values and a smaller proportion with higher BMI values in the given dataset.

**Missed opportunities:**

* The dataset does not show the membership duration and the progressive BMI information to determine the causation of the BMI/exercise and calories burnt.
* There might be several factors like health that are not accounted for by customers that might have an impact on the analysis.

**Assumptions:**

* There was assumption made the calories burnt are directly related to the session duration. But OLS regression results showed that as number of workouts increased the calories burnt will be decreased, showing it takes more time to burn same number of calories.

**Challenges faced:**

Finding the right dataset and choosing the 5 variables that would be included in analysis.

Understanding the correlation between variables and later plotting them. I would like to learn more about functions and various plots.