EDA: Body Composition and Workout Routines Analysis for Gym Members to Optimise Physical Fitness

This report delves into an exploratory data analysis of gym members body composition and workout routines. The goal is to identify key factors influencing physical fitness and provide recommendations for optimising the body composition of all gym members.

Introduction to the study and its objectives

1 Study Scope

The study aims to understand the relationship between body composition, workout routines, and physical fitness outcomes for gym members.

2 Data Analysis

This involves exploring data collected on members body composition metrics and workout routines to identify key factors influencing fitness.

3 Optimising Workout Programmes

The analysis will provide valuable insights for developing personalised workout routines to enhance gym members' overall body composition.

Data collection process and key variables

Source: Gym Members Exercise Dataset on Kaggle

Variable	Description
Age	Member's age in years
Gender	Male or Female
Height	Member's height in meters
Weight	Member's weight in kilograms
Body Fat Percentage	Percentage of body mass that is fat
Body Mass Index (BMI)	It is calculated by dividing a person's weight in kilograms by the square of their height in meters
Workout Frequency	Number of gym visits per week
Calories Burned	Total calories burned during each session
Workout Type	Type of workout performed

Exploratory analysis of body composition metrics

Body Mass Index (BMI)

The BMI categories are standard for both men and women:

- **Underweight**: BMI < 18.5
- **Normal weight**: BMI 18.5–24.9
- **Overweight**: BMI 25–29.9
- **Obesity**: BMI ≥ 30

$$BMI = \frac{\text{Weight (kg)}}{\text{Height (m)} \times \text{Height (m)}}$$

Body Fat Percentage

Healthy Body Fat % for Men by Age:

- **Ages 20-39**: 8-19%
- **Ages 40-59**: 11-21%
- **Ages 60-79**: 13-24%

Healthy Body Fat % for Women by Age:

- **Ages 20-39**: 21-32%
- **Ages 40-59**: 23-33%
- **Ages 60-79**: 24-35%

Data Cleaning: BMI and Body Fat Percentage

Incongruent Record Filtering

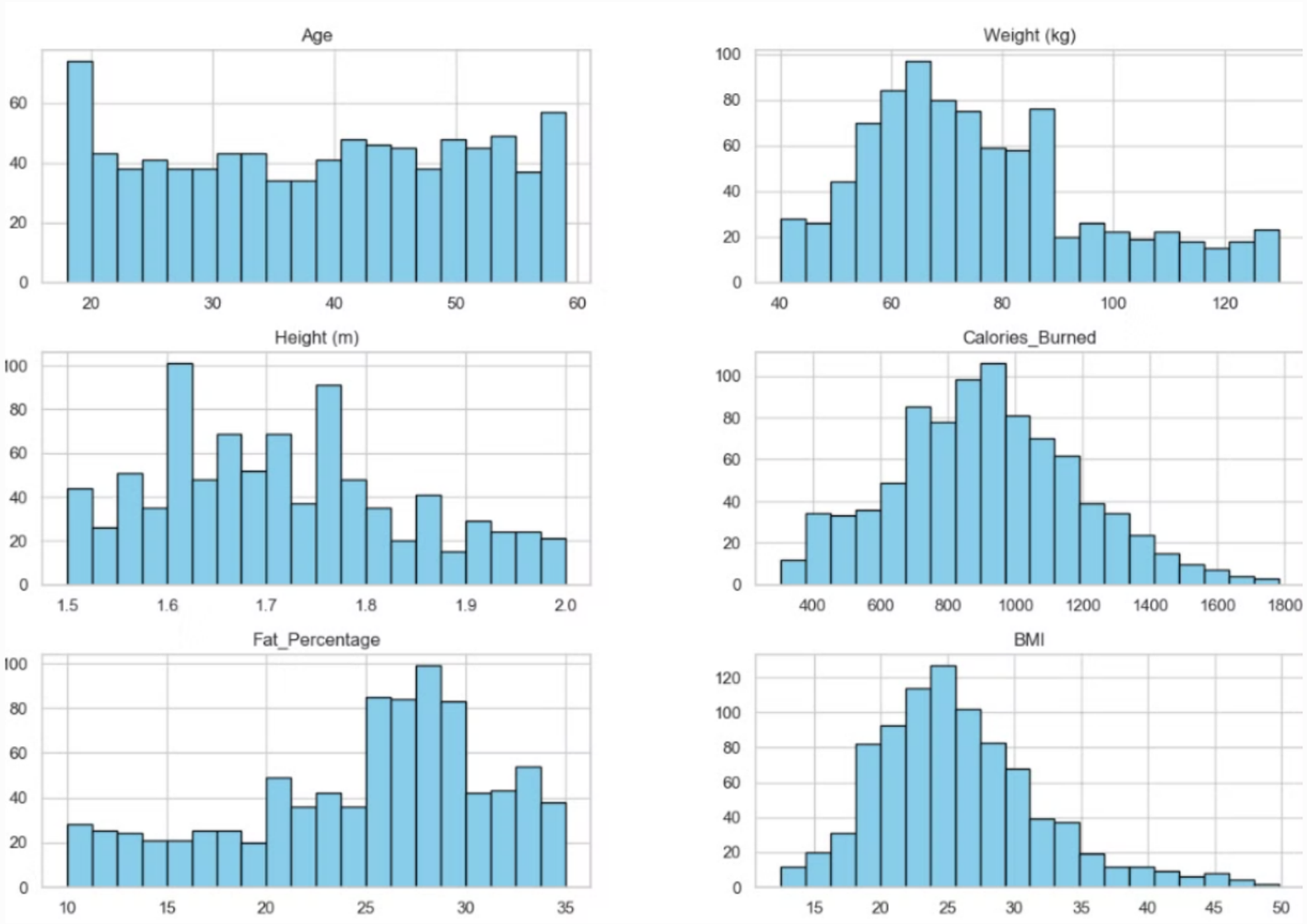
First of all, creation of a code that identifies and removes inconsistent records from the dataset based on:

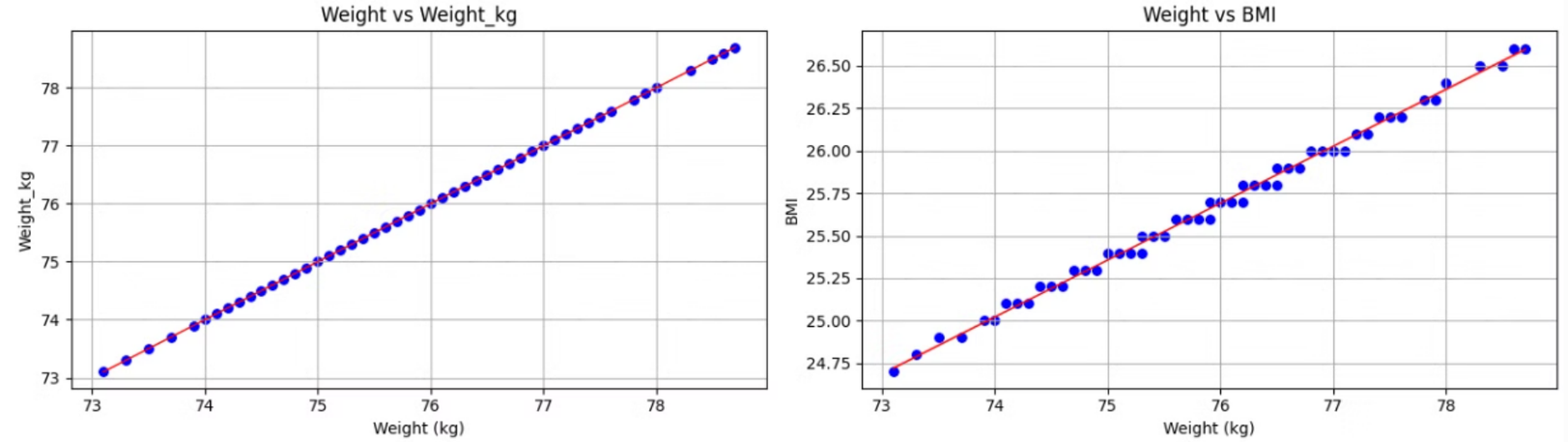
1. **BMI and Fat Percentage Mismatch**: Eliminates **"Underweight"** records with **high fat** or **"Obesity"** records with **low fat**.
2. **Unrealistic Values**: Filters out entries with **extreme fat percentage** (<3% or >50%), **weight** (<30 kg or >200 kg), and **height** (<1.2 m or >2.5 m).

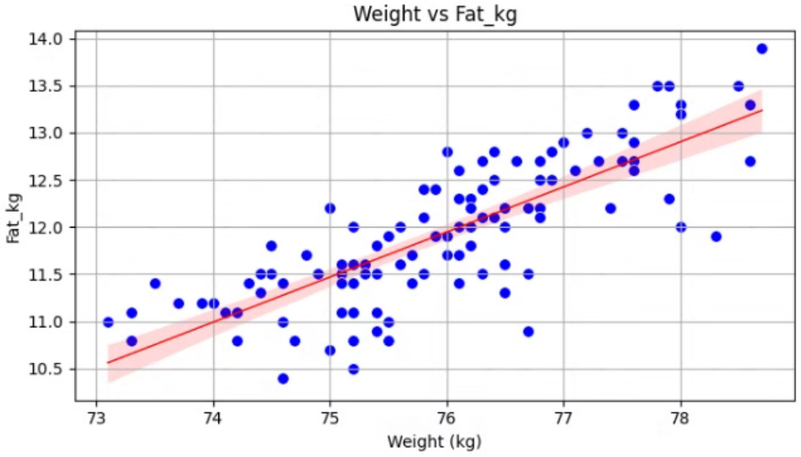
This ensures the dataset is accurate and reliable for analysis.

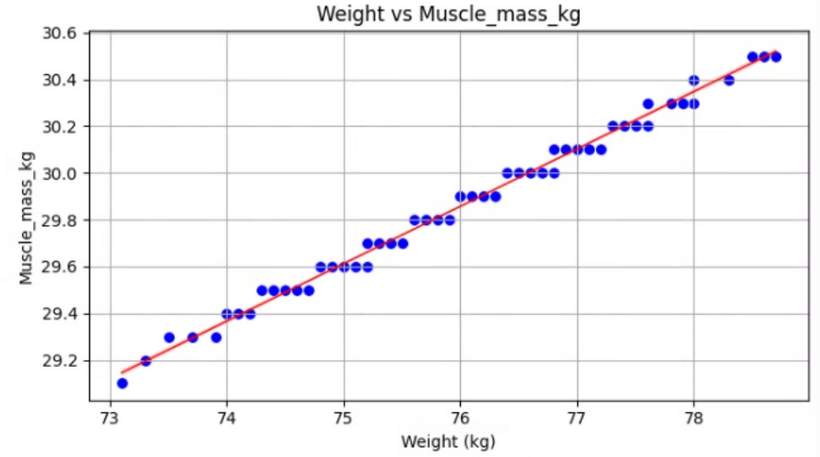
Data Distribution and Removal

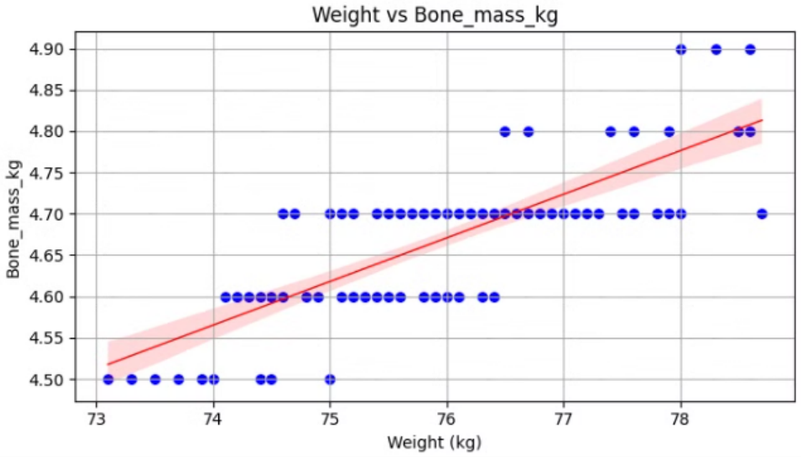
The initial dataset contained 973 imputs for each single gym member with it's metrics. Total records removed: 93; Underweight with High Fat: 92; Obesity with Low Fat: 1 , leaving the data distribution aprox. 50% males and females as in origin. Leaving the dataset like this with the main variables we are working on as shown below but **Gender** and **Workout_Frequency (days/week)**

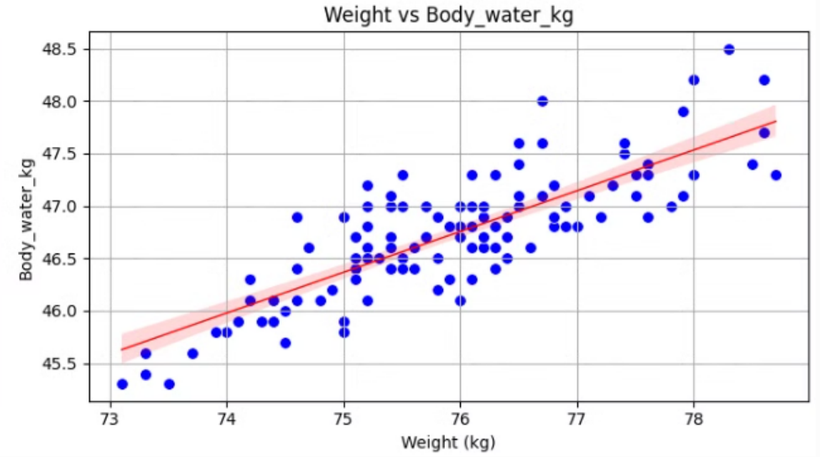








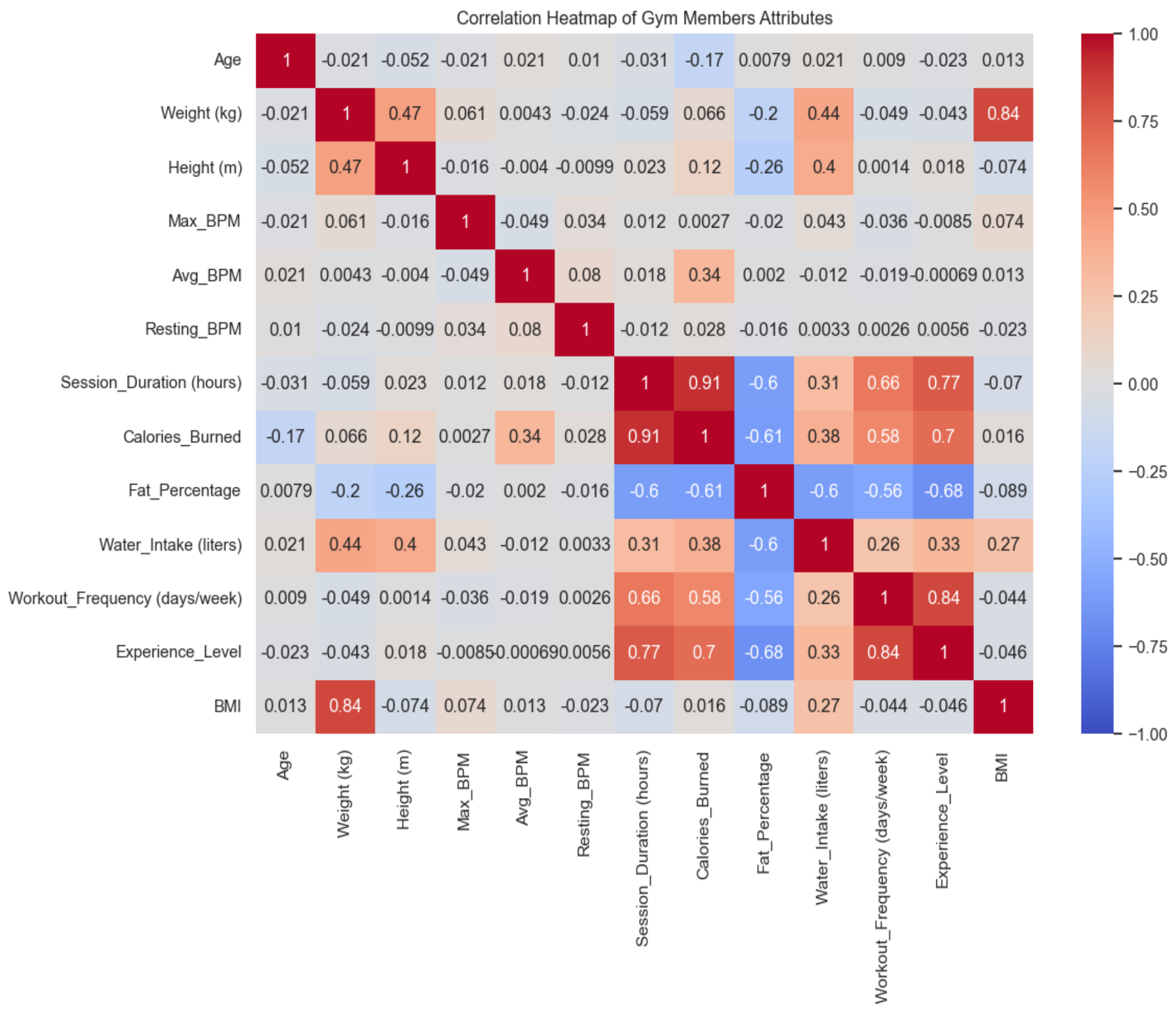




Member Body Composition Analysis Over 4 Months

This analysis explores the body composition changes over time for a single gym member just to check the relevance and relation of BMI, Weight and Fat as on our dataset we only have single records for each gym member.

Source: [Own registers via Garmin Index S2 scale](https://connect.garmin.com/modern/weight)



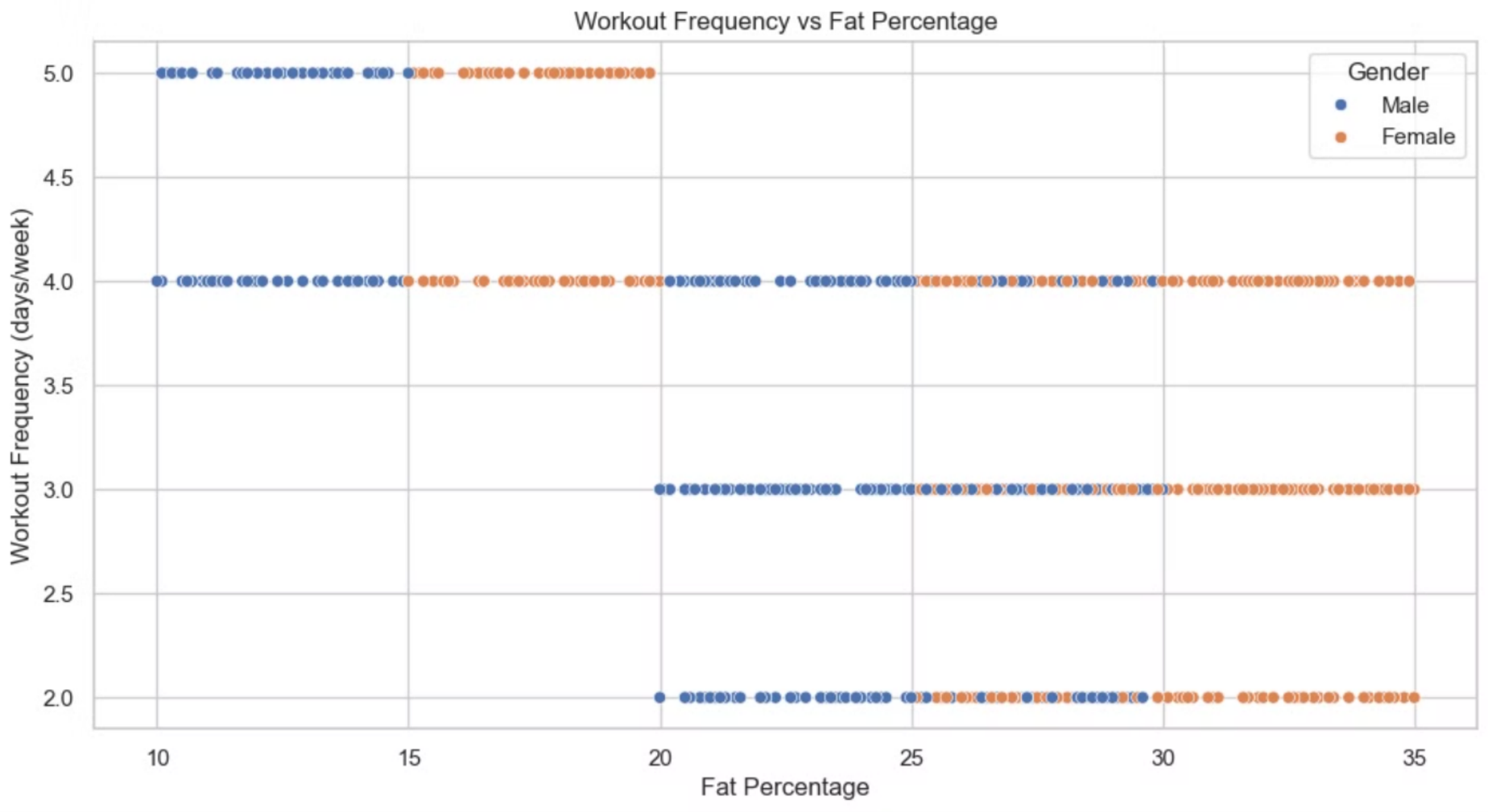
Correlation Matrix for All Gym Members

This correlation matrix analyses the relationship between different body composition metrics for all gym members.

- Positive correlations:
  - BMI – Weight
  - Experience_Level
    - Workout_Frequency
    - Session_Duration
    - Calories_Burned
- Negative correlations:
  - Fat_Percentage
    - Workout_Frequency
    - Session_Duration
    - Calories_Burned
    - Experience_Level

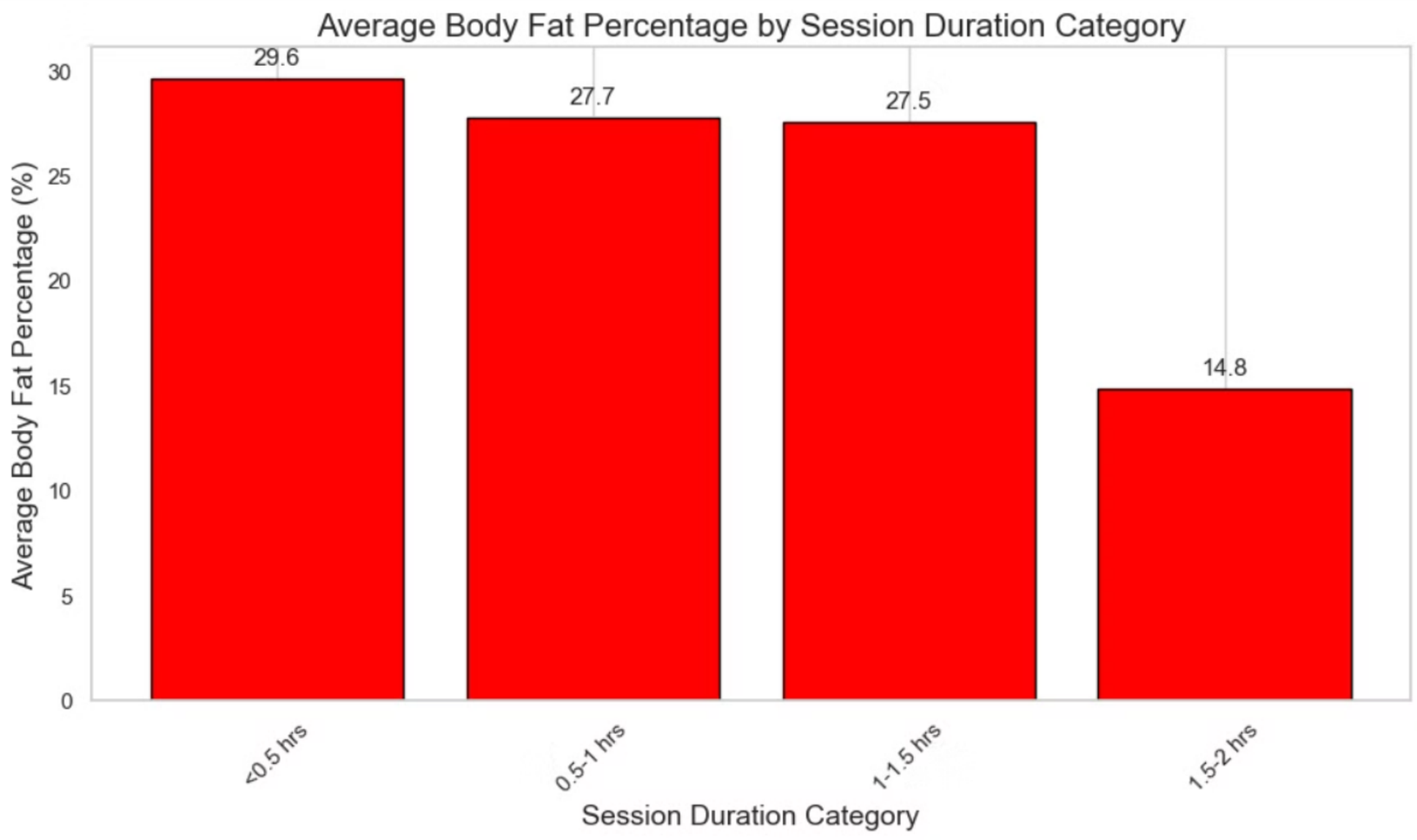
Workout Frequency and Body Fat Percentage

The average body fat percentage tends to decrease with more frequent workouts.



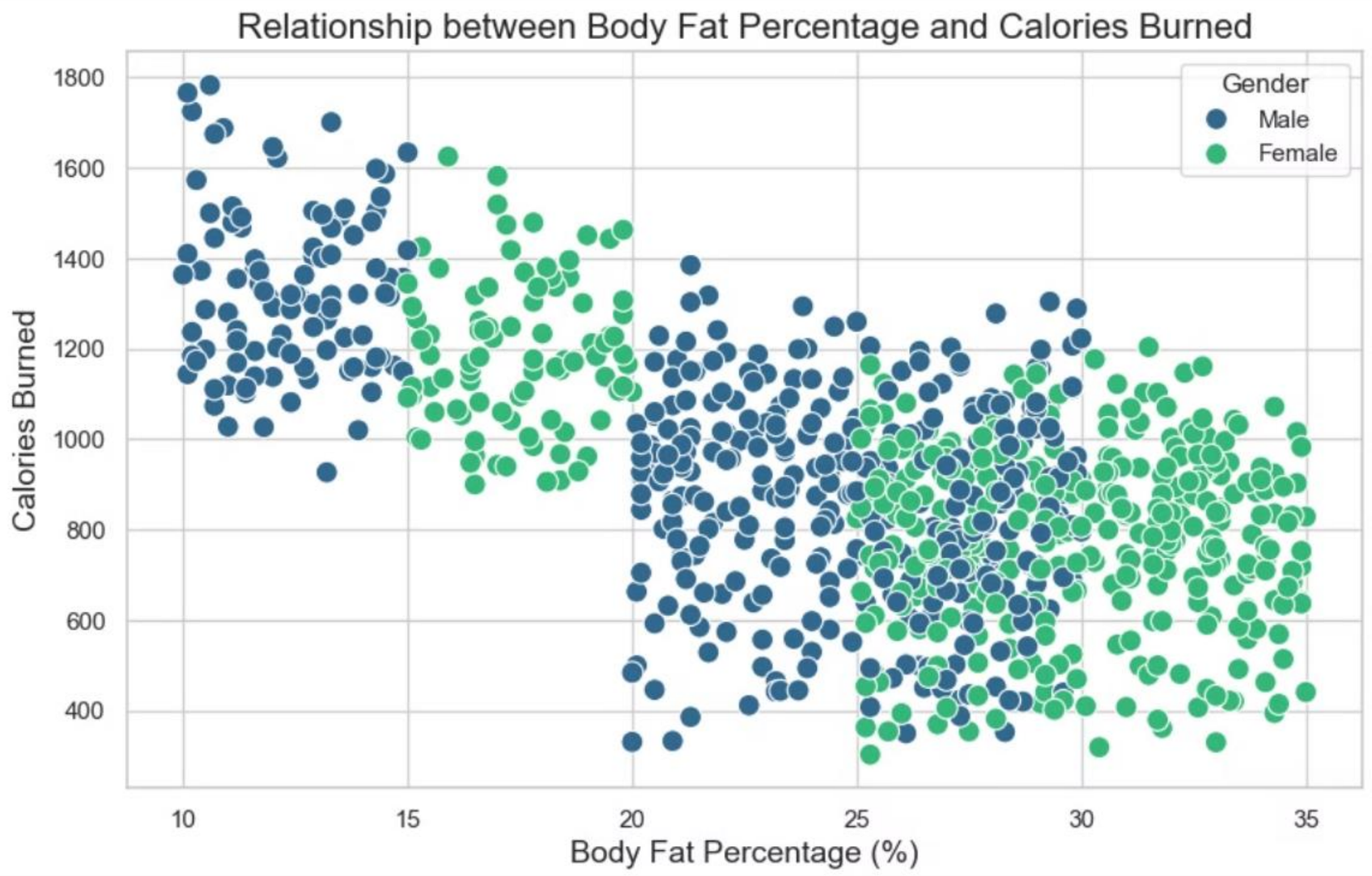
Relationship between Workout Duration and Body Fat Percentage

Members who participate in longer sessions tend to have lower body fat percentage, suggesting a correlation between workout duration and fat reduction.



Calories Burned and Body Fat Percentage

This allows us to see how the two are connected and to identify that the intesity of the workout is crucial in lowering the body fat.



Best workouts to gain muscle and lose fat

The most experienced members of the gym shows us that a combination of Strength and HIIT training are the best options to decrease fat and rise muscle

Identifying key factors influencing physical fitness

1

Workout Frequency

Consistent workout frequency is a significant factor influencing physical fitness, as it provides regular opportunities for calorie expenditure and lowering fat mass. 4-5 days/week

2

Workout Duration

Longer workout sessions allow for greater calorie expenditure and fat loss, contributing to improved fitness levels. 1,5 - 2 hours/session

3

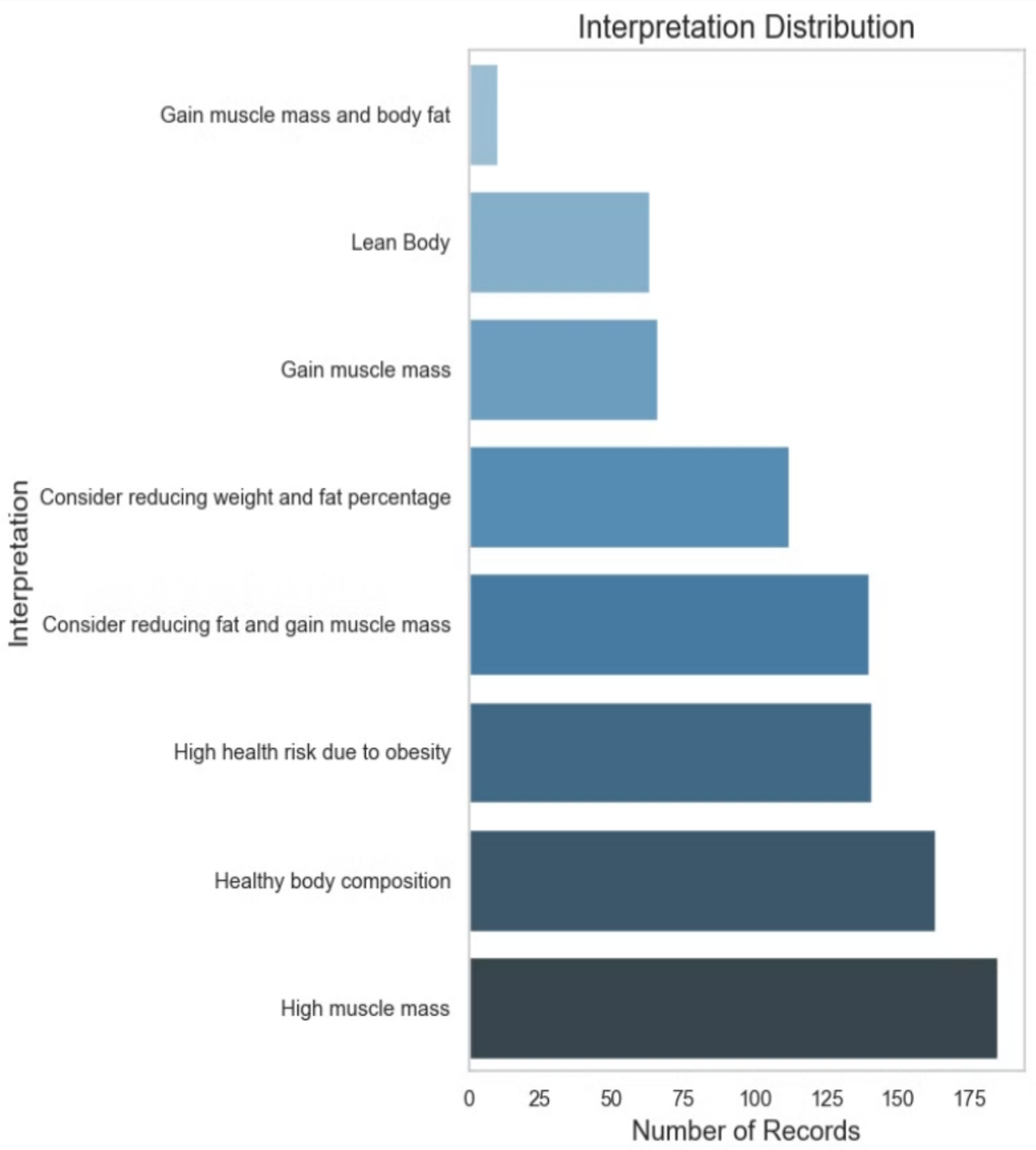
Workout Intensity

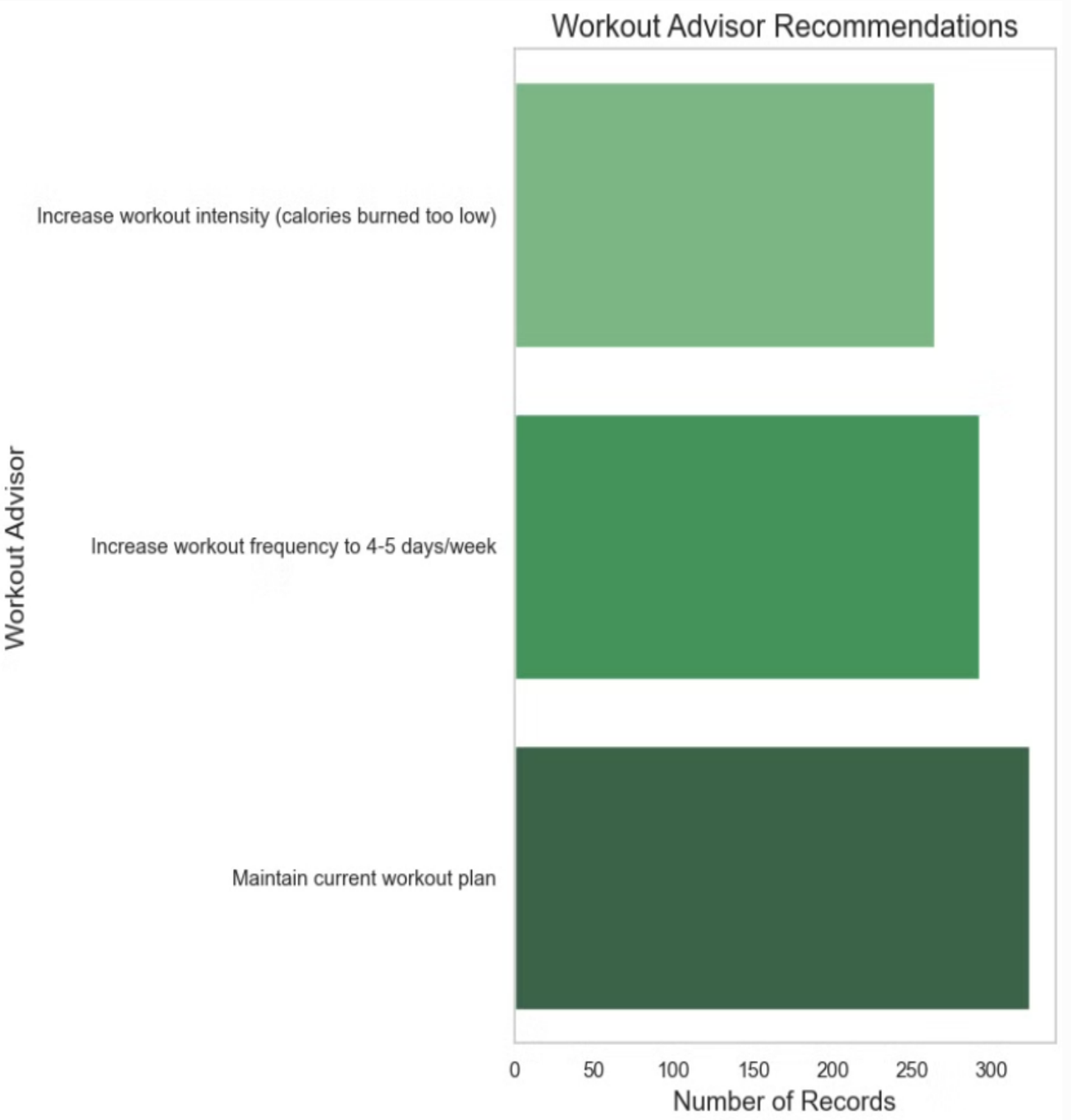
Higher intensity workouts lead to greater calorie burn and metabolic benefits, resulting in more significant fitness improvements like fat loss. 915 kcal/session at least

Body composition interpreter and workout advisor for optimising workout routines

```python
# Interpretation function based only on BMI and fat percentage
def interpret_body_composition(row):
    bmi_status = row['BMI_Status']
    fat_status = row['Fat_Status']

    # Interpretation based on BMI and fat percentage
    if (bmi_status == "Overweight" or bmi_status == "Obesity") and (fat_status == 'Healthy' or fat_status == 'Low Fat (Consider increasing fat)'):
        return "High muscle mass"
    elif bmi_status == "Underweight" and "High Fat" in fat_status:
        return "Gain muscle mass and reduce body fat"
    elif bmi_status == "Normal weight" and fat_status == "Healthy":
        return "Healthy body composition"
    elif bmi_status == "Normal weight" and fat_status == "Low Fat (Consider increasing fat)":
        return "Lean Body"
    elif bmi_status == "Normal weight" and fat_status == "High Fat (Consider reducing fat)":
        return "Consider reducing fat and gain muscle mass"
    elif bmi_status == "Overweight":
        return "Consider reducing weight and fat percentage"
    elif bmi_status == "Obesity":
        return "High health risk due to obesity"
    elif bmi_status == "Underweight" and fat_status == "Healthy":
        return "Gain muscle mass"
    elif bmi_status == "Underweight" and fat_status == "Low Fat (Consider increasing fat)":
        return "Gain muscle mass and body fat"
    else:
        return "General improvement suggested"

# Add the Interpretation column to the DataFrame
df_gym['Interpretation'] = df_gym.apply(interpret_body_composition, axis=1)

# Select relevant columns for display
columns_to_display = [
    'Age', 'Gender', 'Weight (kg)', 'Height (m)', 'BMI', 'BMI_Status',
    'Fat_Percentage', 'Fat_Status', 'Interpretation'
]

# Create a summary DataFrame
df_gym_summary = df_gym[columns_to_display]

# Display the updated DataFrame
df_gym_summary.head()

# Export the DataFrame to a CSV file (optional)
# df_gym_summary.to_csv('gym_body_composition_analysis.csv', index=False)

# print("CSV file exported successfully: gym_body_composition_analysis.csv")
```

	Age	Gender	Weight (kg)	Height (m)	BMI	BMI_Status	Fat_Percentage	Fat_Status	Workout_Frequency (days/week)	Session_Duration (hours)	Calories_Burned	Workout_Type	Interpretation	Workout_Advisor
0	56	Male	88.3	1.71	30.20	Obesity	12.6	Healthy	4	1.69	1313.0	Yoga	High muscle mass	Maintain current workout plan
1	46	Female	74.9	1.53	32.00	Obesity	33.9	High Fat (Consider reducing fat)	4	1.30	883.0	HIIT	High health risk due to obesity	Increase workout intensity (calories burned to...
2	32	Female	68.1	1.66	24.71	Normal weight	33.4	High Fat (Consider reducing fat)	4	1.11	677.0	Cardio	Consider reducing fat and gain muscle mass	Increase workout intensity (calories burned to...
5	56	Female	58.0	1.68	20.55	Normal weight	15.5	Low Fat (Consider increasing fat)	5	1.59	1116.0	HIIT	Lean Body	Maintain current workout plan
6	36	Male	70.3	1.72	23.76	Normal weight	21.3	High Fat (Consider reducing fat)	3	1.49	1385.0	Cardio	Consider reducing fat and gain muscle mass	Increase workout frequency to 4-5 days/week
7	40	Female	69.7	1.51	30.57	Obesity	30.6	Healthy	3	1.27	895.0	Cardio	High muscle mass	Increase workout intensity (calories burned to...
8	28	Male	121.7	1.94	32.34	Obesity	28.9	High Fat (Consider reducing fat)	4	1.03	719.0	Strength	High health risk due to obesity	Increase workout intensity (calories burned to...
9	28	Male	101.8	1.84	30.07	Obesity	29.7	High Fat (Consider reducing fat)	3	1.08	808.0	Cardio	High health risk due to obesity	Increase workout frequency to 4-5 days/week
10	41	Male	120.8	1.67	43.31	Obesity	20.5	Healthy	2	0.82	593.0	HIIT	High muscle mass	Increase workout intensity (calories burned to...
12	57	Male	112.5	1.61	43.40	Obesity	22.1	High Fat (Consider reducing fat)	3	1.24	1013.0	Cardio	High health risk due to obesity	Increase workout frequency to 4-5 days/week
13	41	Male	94.5	2.00	23.62	Normal weight	27.6	High Fat (Consider reducing fat)	3	1.18	794.0	HIIT	Consider reducing fat and gain muscle mass	Increase workout frequency to 4-5 days/week
14	20	Male	117.7	1.81	35.93	Obesity	26.4	High Fat (Consider reducing fat)	3	1.35	1195.0	Yoga	High health risk due to obesity	Increase workout frequency to 4-5 days/week
15	39	Female	42.5	1.75	13.88	Underweight	26.2	Healthy	2	1.13	740.0	Strength	Gain muscle mass	Increase workout intensity (calories burned to...
16	19	Female	64.0	1.53	27.34	Overweight	29.8	Healthy	3	1.33	1111.0	HIIT	High muscle mass	Maintain current workout plan
17	41	Female	43.8	1.77	13.98	Underweight	31.9	Healthy	3	1.19	884.0	Cardio	Gain muscle mass	Increase workout intensity (calories burned to...
18	47	Female	66.8	1.75	21.81	Normal weight	32.8	Healthy	3	1.13	742.0	Strength	Healthy body composition	Increase workout intensity (calories burned to...
19	55	Female	75.2	1.67	26.96	Overweight	25.2	Healthy	2	1.37	1030.0	HIIT	High muscle mass	Maintain current workout plan
20	19	Male	89.0	1.77	28.41	Overweight	28.9	High Fat (Consider reducing fat)	4	1.50	1048.0	Strength	Consider reducing weight and fat percentage	Maintain current workout plan
21	38	Male	71.9	1.77	22.95	Normal weight	25.7	High Fat (Consider reducing fat)	2	1.12	875.0	Cardio	Consider reducing fat and gain muscle mass	Increase workout frequency to 4-5 days/week
22	50	Female	71.0	1.68	25.16	Overweight	33.1	High Fat (Consider reducing fat)	2	1.17	848.0	Yoga	Consider reducing weight and fat percentage	Increase workout frequency to 4-5 days/week
23	29	Male	120.9	1.78	38.16	Obesity	28.1	High Fat (Consider reducing fat)	2	0.78	721.0	Yoga	High health risk due to obesity	Increase workout frequency to 4-5 days/week
24	39	Female	64.3	1.69	22.51	Normal weight	26.9	Healthy	4	1.25	925.0	HIIT	Healthy body composition	Maintain current workout plan
25	42	Female	63.7	1.71	21.78	Normal weight	26.1	Healthy	3	1.42	1080.0	Yoga	Healthy body composition	Maintain current workout plan
26	44	Male	65.2	1.80	20.12	Normal weight	27.2	High Fat (Consider reducing fat)	2	0.73	502.0	HIIT	Consider reducing fat and gain muscle mass	Increase workout frequency to 4-5 days/week
28	45	Male	84.9	1.86	24.54	Normal weight	14.2	Healthy	5	1.64	1104.0	HIIT	Healthy body composition	Maintain current workout plan
29	33	Female	78.0	1.68	27.64	Overweight	32.3	High Fat (Consider reducing fat)	3	1.29	871.0	HIIT	Consider reducing weight and fat percentage	Increase workout frequency to 4-5 days/week
30	32	Male	108.2	1.80	33.40	Obesity	28.4	High Fat (Consider reducing fat)	3	1.27	964.0	HIIT	High health risk due to obesity	Increase workout frequency to 4-5 days/week
31	20	Female	65.4	1.52	28.31	Overweight	28.0	Healthy	4	1.03	654.0	Yoga	High muscle mass	Increase workout intensity (calories burned to...
32	54	Female	50.2	1.61	19.37	Normal weight	28.2	Healthy	2	1.48	1046.0	HIIT	Healthy body composition	Maintain current workout plan
33	24	Female	58.9	1.51	25.83	Overweight	31.7	Healthy	2	1.04	816.0	Cardio	High muscle mass	Increase workout intensity (calories burned to...
34	38	Male	81.4	1.71	27.84	Overweight	10.2	Healthy	5	1.52	1237.0	HIIT	High muscle mass	Maintain current workout plan
35	26	Male	127.6	1.73	42.63	Obesity	27.3	High Fat (Consider reducing fat)	3	1.32	1162.0	Strength	High health risk due to obesity	Increase workout frequency to 4-5 days/week
36	56	Female	59.3	1.56	24.37	Normal weight	33.4	High Fat (Consider reducing fat)	4	1.26	879.0	Cardio	Consider reducing fat and gain muscle mass	Increase workout intensity (calories burned to...
37	35	Male	96.9	1.71	33.14	Obesity	24.2	High Fat (Consider reducing fat)	3	1.34	1069.0	Strength	High health risk due to obesity	Increase workout frequency to 4-5 days/week
38	21	Male	62.6	1.81	19.11	Normal weight	27.6	High Fat (Consider reducing fat)	4	1.30	1072.0	Strength	Consider reducing fat and gain muscle mass	Maintain current workout plan
40	31	Female	48.8	1.51	21.40	Normal weight	28.2	Healthy	2	1.48	969.0	Cardio	Healthy body composition	Maintain current workout plan
42	43	Male	113.2	1.83	33.80	Obesity	29.9	High Fat (Consider reducing fat)	4	1.45	962.0	Yoga	High health risk due to obesity	Maintain current workout plan
43	19	Female	60.5	1.59	23.93	Normal weight	30.1	Healthy	2	1.14	730.0	Cardio	Healthy body composition	Increase workout intensity (calories burned to...
44	37	Male	124.2	1.76	40.10	Obesity	29.3	High Fat (Consider reducing fat)	2	1.50	1304.0	HIIT	High health risk due to obesity	Increase workout frequency to 4-5 days/week
47	25	Male	88.1	1.95	23.17	Normal weight	22.7	High Fat (Consider reducing fat)	2	1.41	1016.0	Cardio	Consider reducing fat and gain muscle mass	Increase workout frequency to 4-5 days/week





Use cases and development

1 GYM Owners

- Data-Driven Program Design
- Service Optimization
- Increase customer engagement

2 GYM Users

- Easily know your physical condition
- Adapt their training routine

3 Continued Research

Additional research with larger samples and longitudinal data may predict evolution, maintain user motivation, and increase customer retention.

Thank you for your attention

Source Code and Data

View the code and data used for this analysis on Github.

https://github.com/MikelTelo/eda_for_gym_owners_members

LinkedIn Profile

Connect with Mikel Telo on LinkedIn to discuss the analysis in more detail.

https://www.linkedin.com/in/mikel-telo-a96a371a/