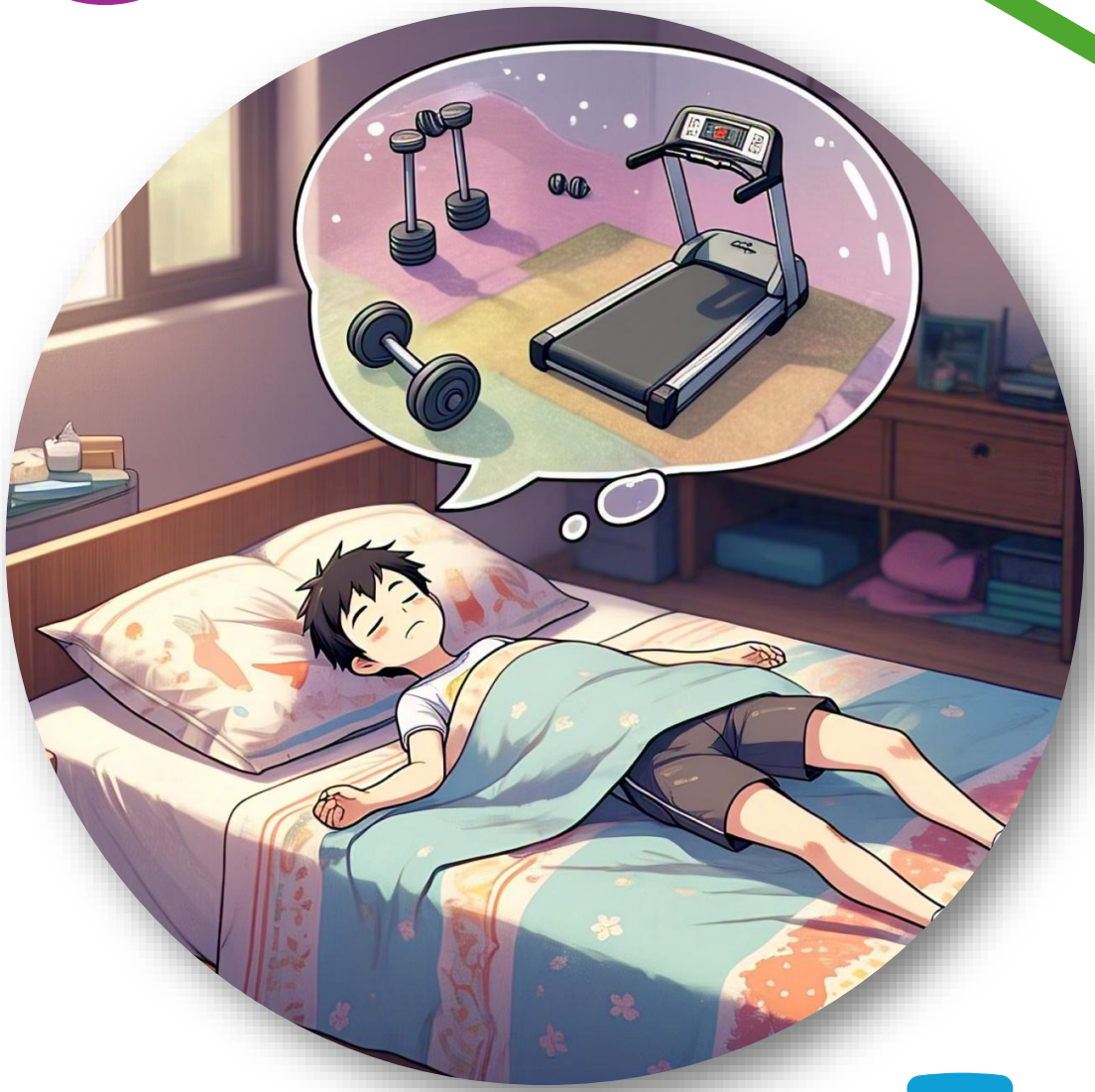


SLEEP DURATION AND WORKOUT QUALITY

FINAL PROJET 530-T301

DANIEL SOLIS TORO

BELLEVUE UNIVERSITY

How does sleep duration before a workout influence burned calories, workout duration, and post-workout mood?

Adequate sleep boosts energy and performance, whereas poor sleep can diminish effectiveness, increase fatigue, and contribute to negative post-workout emotions.

This study explores the extent of sleep duration's influence and examines potential variations based on age, gender, and workout type.

DATA SET: Workout & Fitness Tracker

Overview

This dataset contains 10,000+ records of workout and fitness-related data collected from various fitness apps and devices. It includes various metrics such as workout type, duration, calories burned, heart rate, steps taken, and other health-related parameters.

VARIABLES

User ID: A unique numerical identifier assigned to each user in the dataset.

Age: The age of the user, ranging from 18 to 60 years.

Gender: The gender of the user (Male, Female, or Other).

Height (cm): The height of the user in centimeters.

Weight (kg): The weight of the user in kilograms.

VARIABLES

Workout Type: The type of workout performed (e.g., Cardio, Strength, Yoga, HIIT, Cycling, Running).

Workout Duration (mins): The total workout duration in minutes.

Calories Burned: The number of calories burned during the workout session.

Heart Rate (bpm): The average heart rate (beats per minute) recorded during the workout.

Steps Taken: The total number of steps taken during the workout (applicable for walking/running workouts).

To analyze the research question, the following variables were selected:

Independent Variable:
- Sleep duration (hours of sleep before the workout).

Dependent Variables:
- Calories Burned: Total calories burned during the workout.
- Workout Duration (mins): Total time spent in a workout
- Post-Workout Mood: Energized, Neutral, Fatigued.

Variables that could influence the relationship:
- Age
- Gender

ANALYSIS OF THE MAIN VARIABLES

SLEEP DURATION

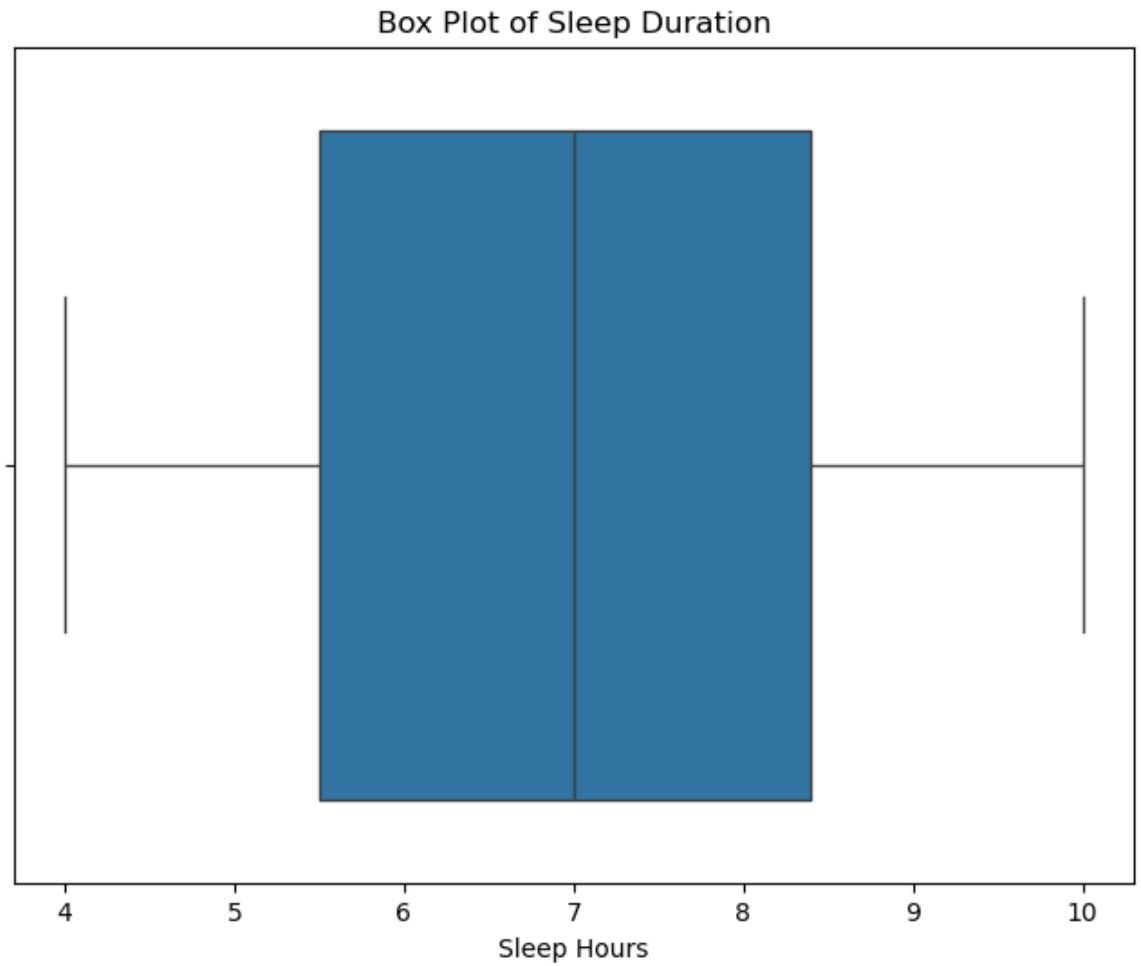
OUTLIERS CHECK

Reasoning:

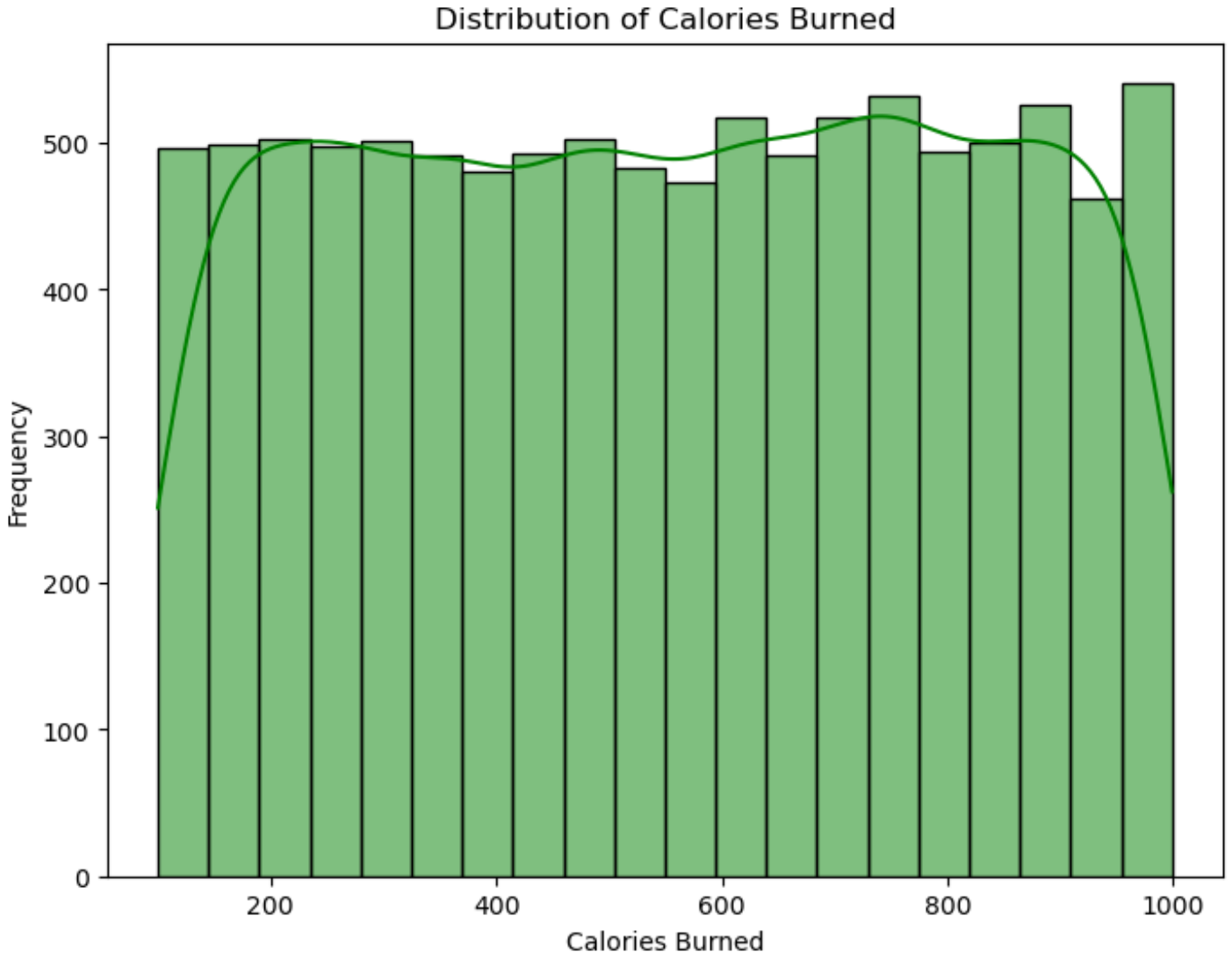
Extremely short sleep durations may indicate data entry errors or users with irregular sleep patterns (e.g., shift workers).

Extremely long sleep durations may be unrealistic or indicate errors (e.g., users forgetting to stop tracking sleep).

Output: None outliers were identified



CALORIES BURNED



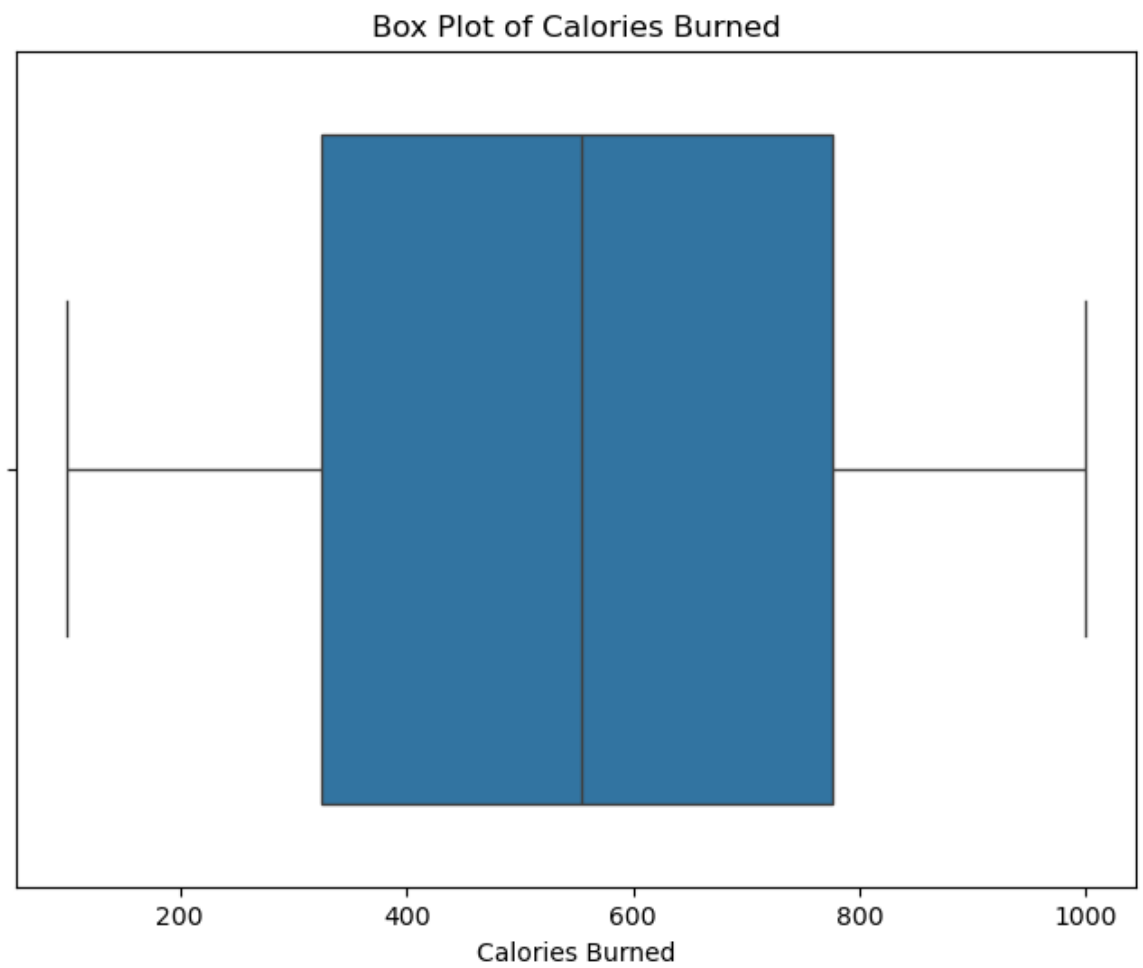
OUTLIERS CHECK

Reasoning:

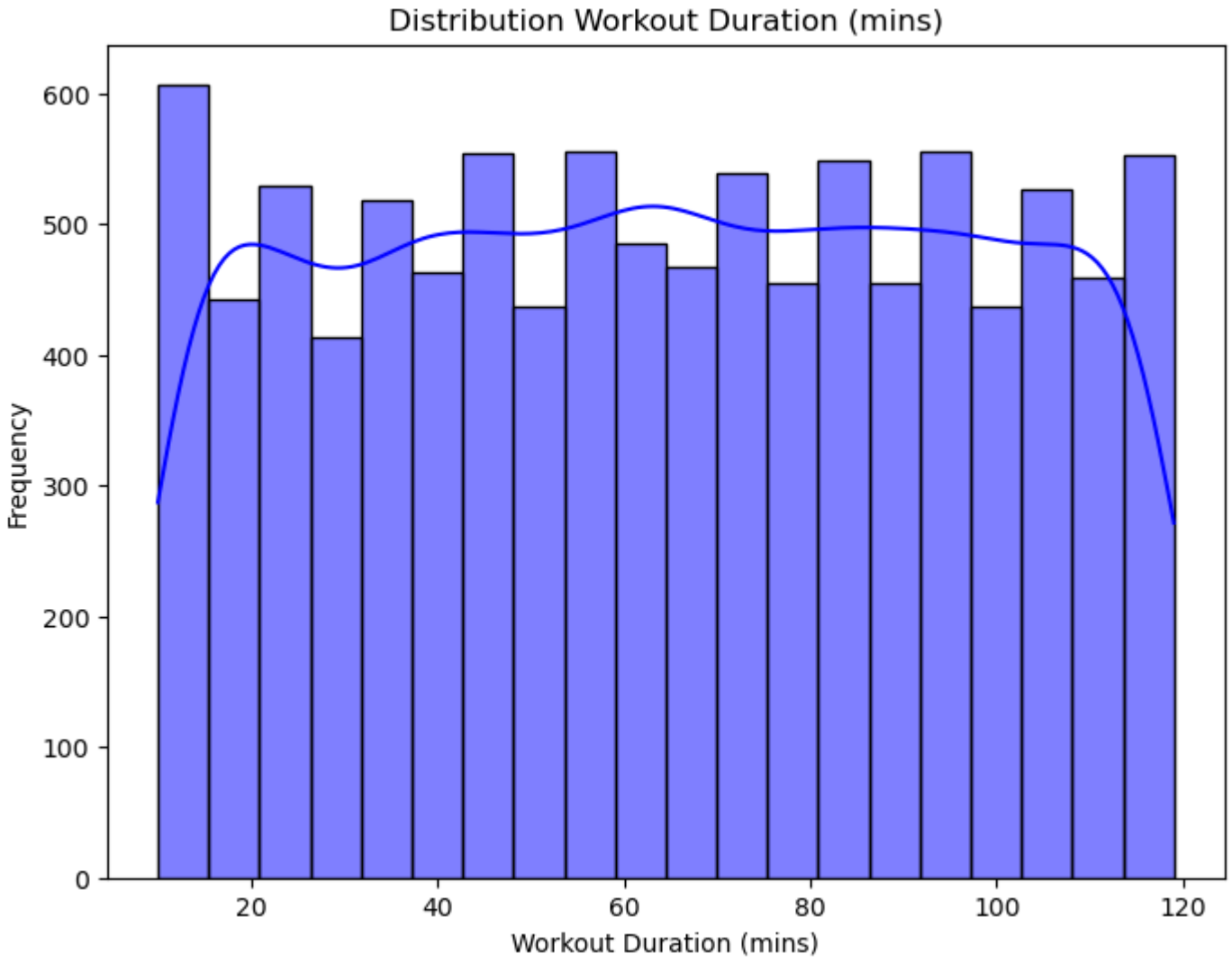
High-calorie burns may result from long or intense workouts, but extreme values could indicate errors in tracking or unrealistic data.

Low-calorie burns may represent incomplete or very short workouts.

Output: None outliers were identified



WORKOUT DURATION

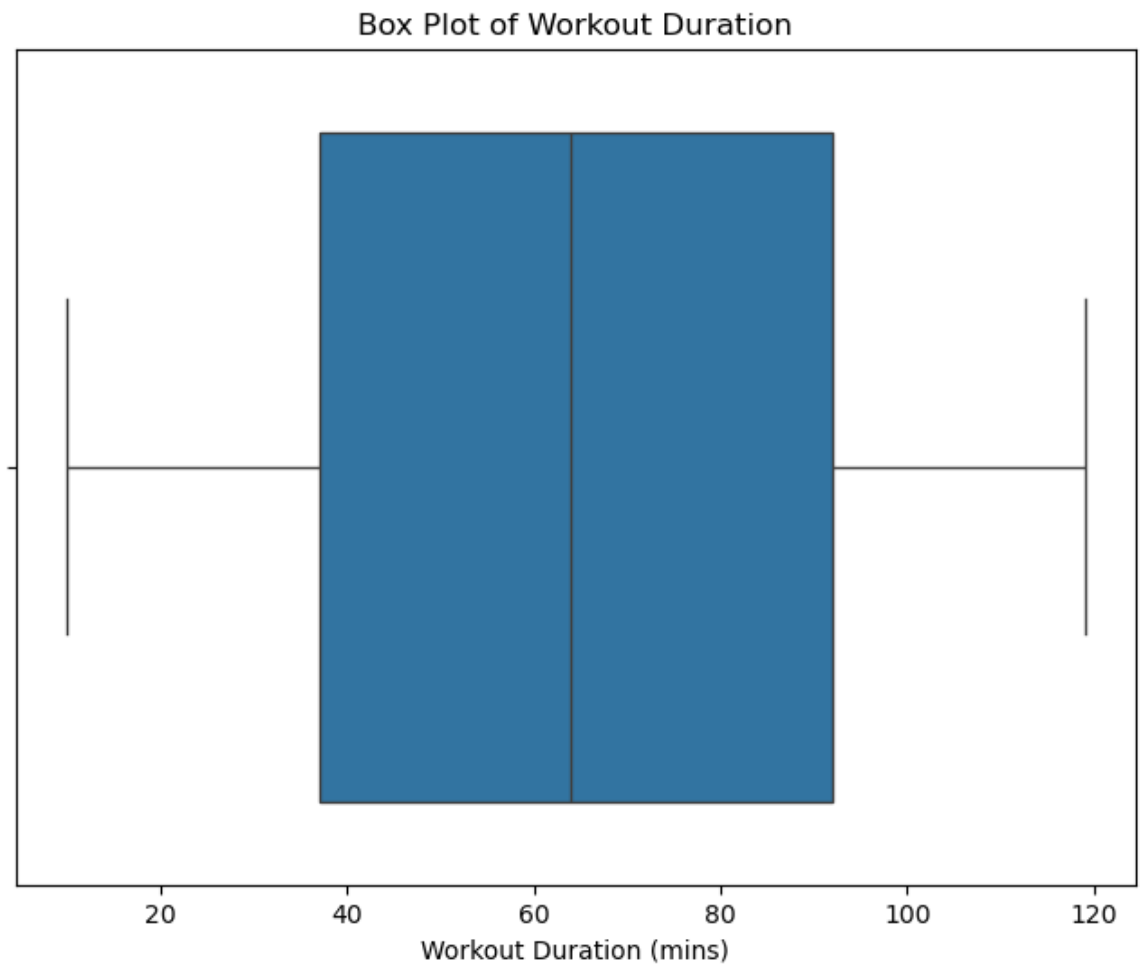


OUTLIERS CHECK

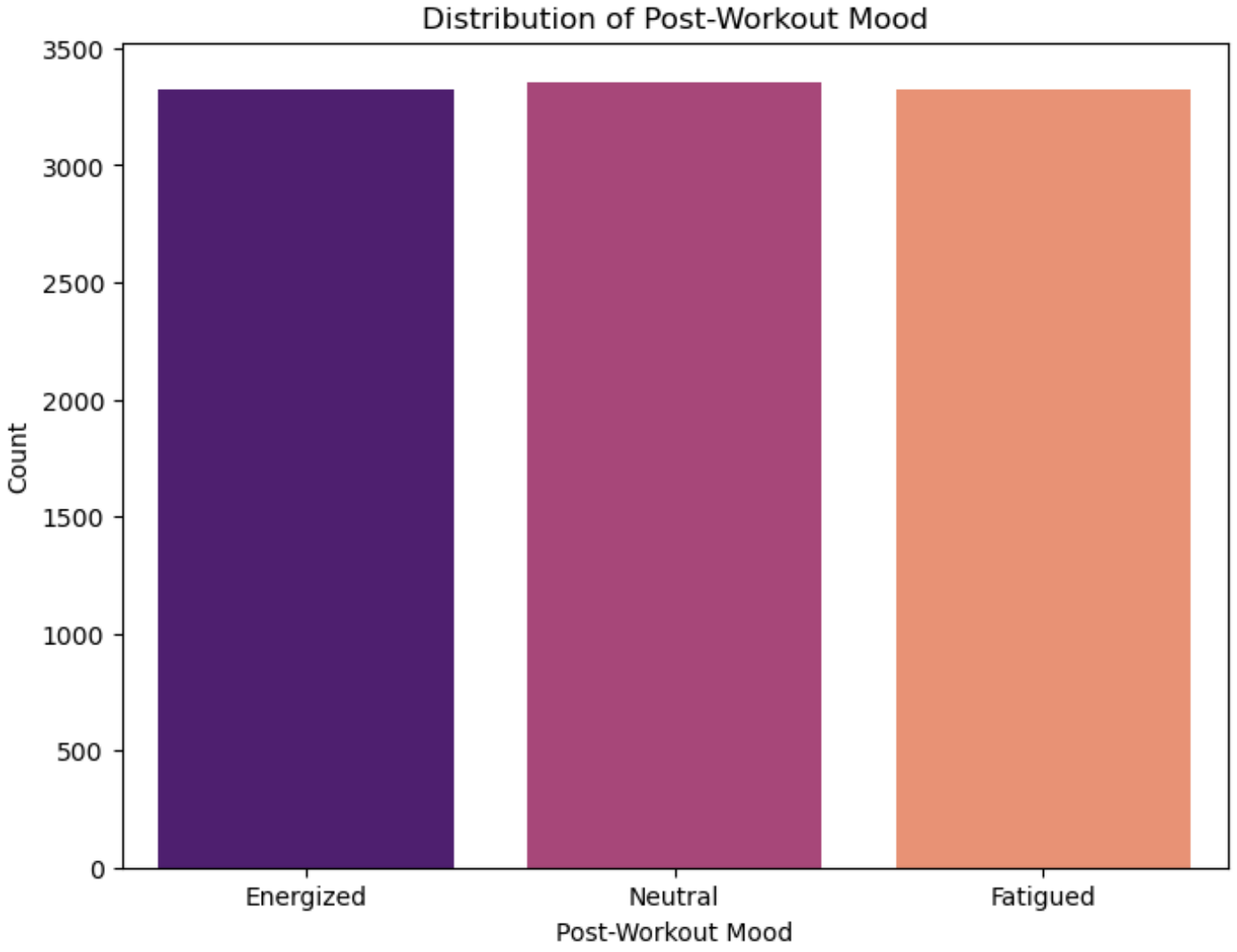
Reasoning:

Long durations may indicate users forgetting to stop tracking, while short durations may represent incomplete workouts.

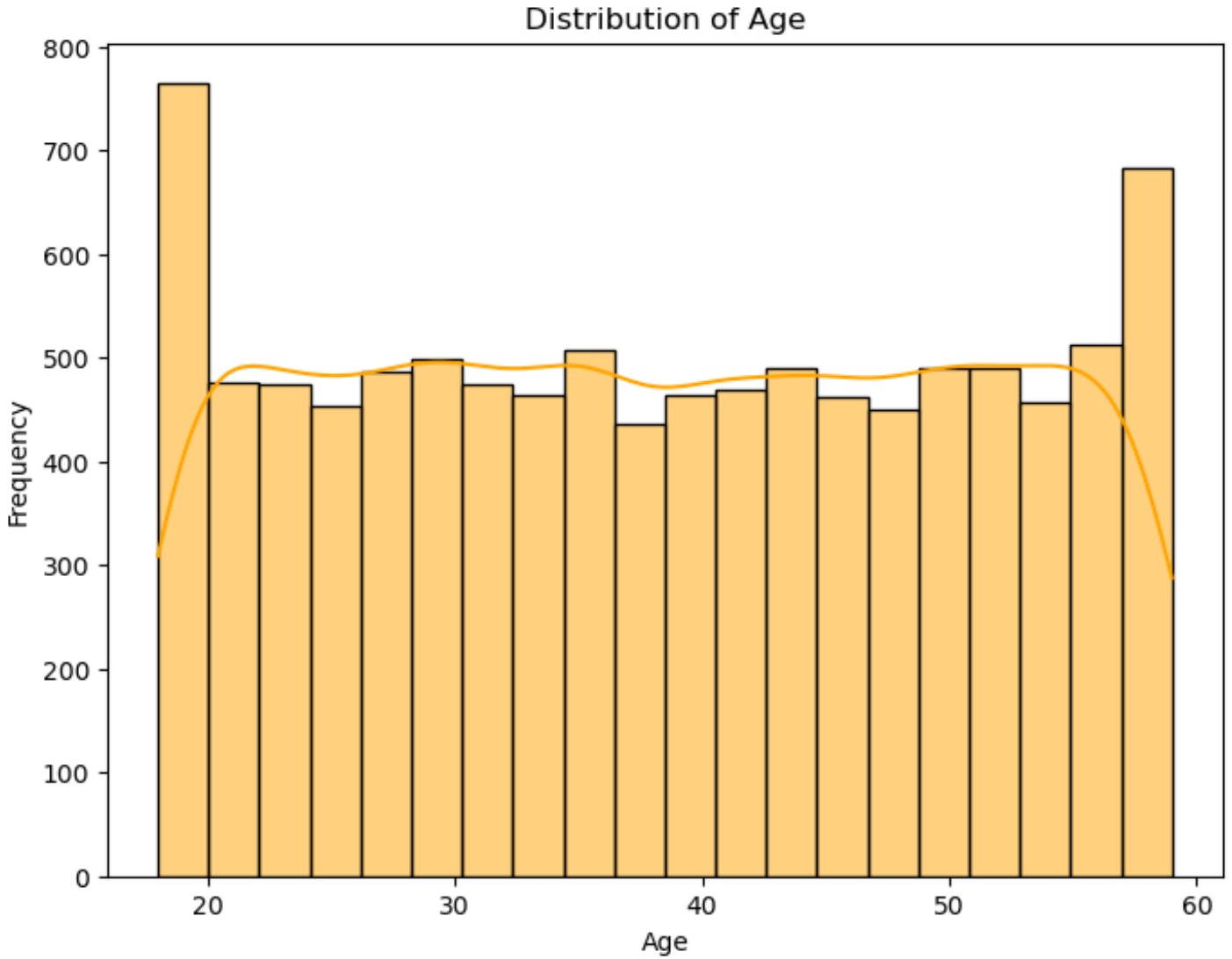
Output: None outliers were identified



POST-WORKOUT MOOD



AGE

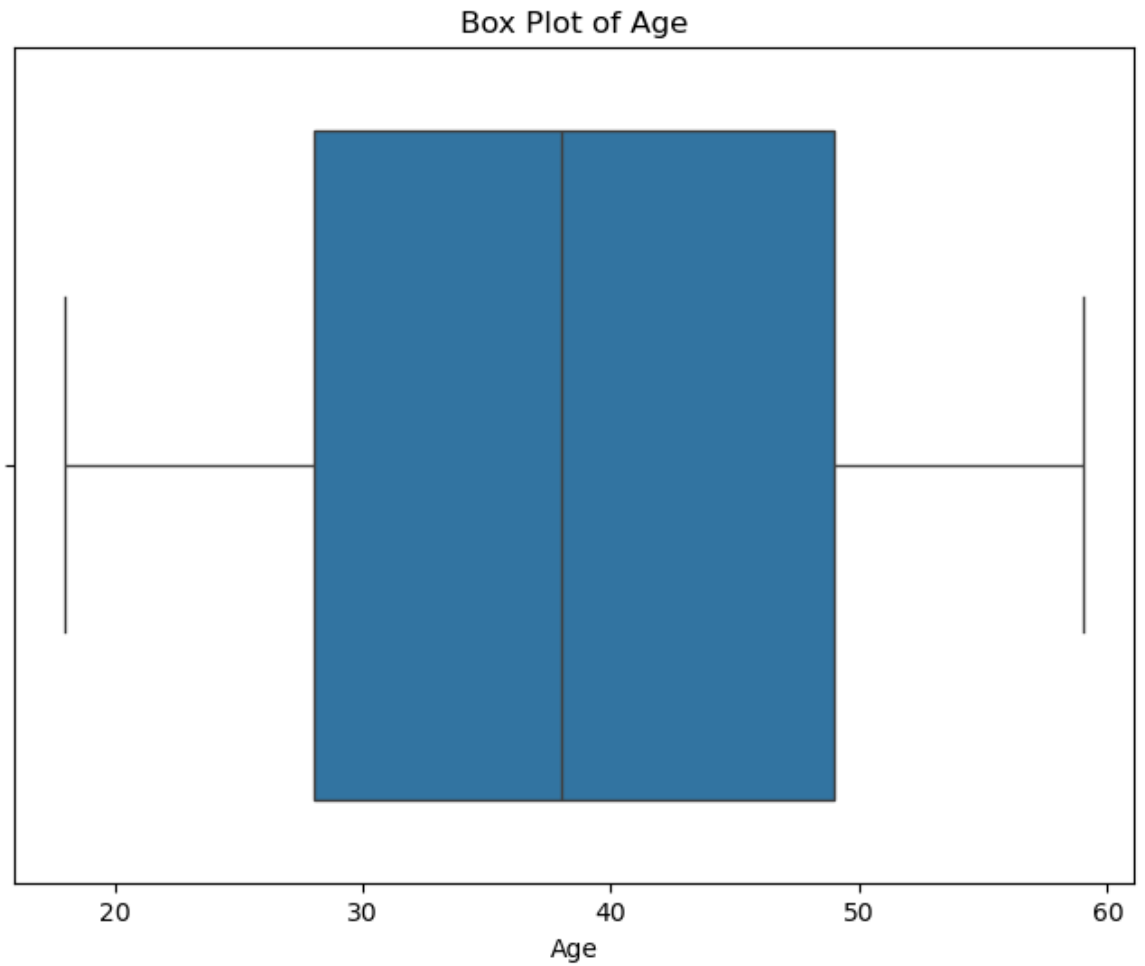

Distribution of Age
Frequency
800
700
600
500
400
300
200
100
0
20
30
40
50
60
Age

OUTLIERS CHECK

Reasoning:

These may be data entry errors or users outside the target demographic.

Output: None outliers were identified



GENDER

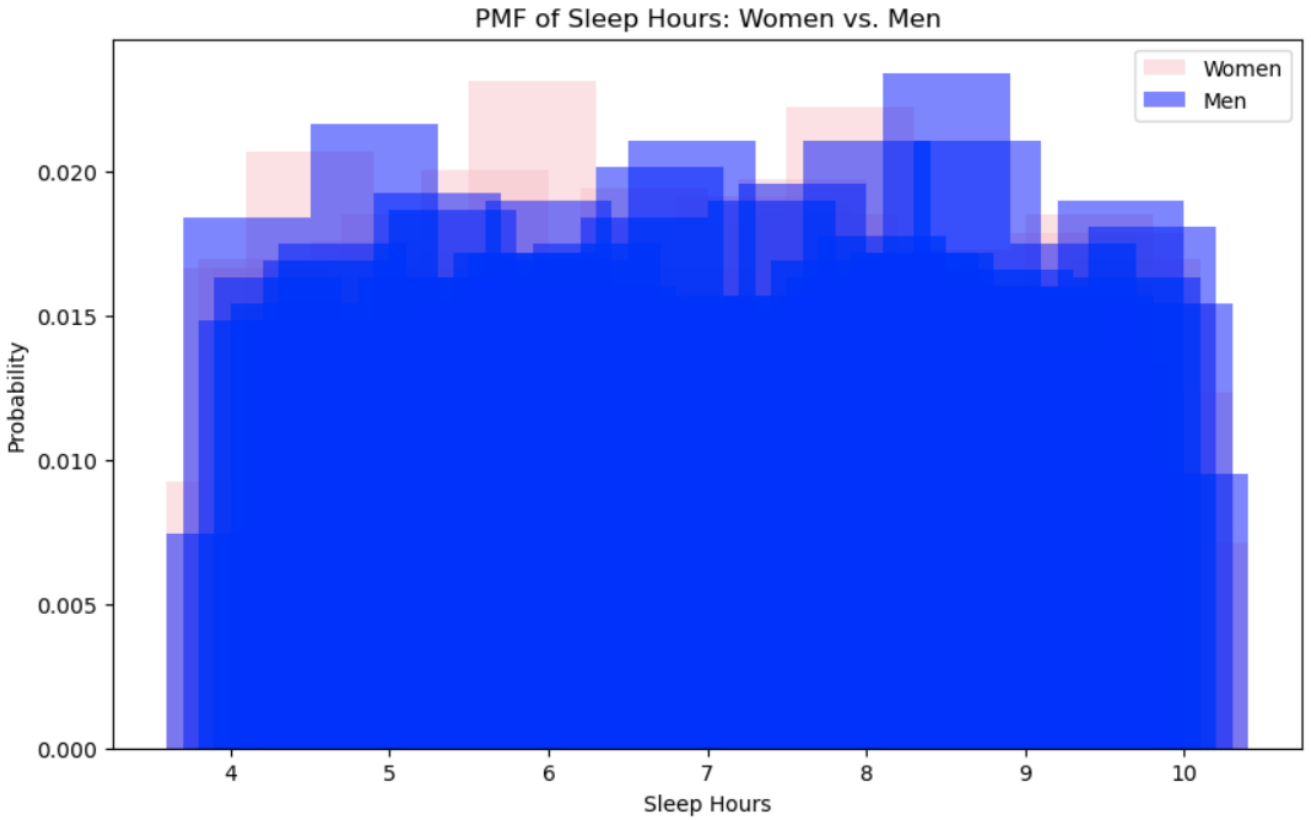

Distribution of Gender
Count
3500
3000
2500
2000
1500
1000
500
0
Male
Other
Female
Gender

DESCRIPTIVE ANALYSIS

Variable	Mean	Std Dev	Min	Max	Range	Q1	Q3	IQR	Skewness	Kurtosis
Sleep Hours	6.98	1.71	4.0	10.0	6.0	5.5	8.4	2.9	0.005072	-1.164787
Calories Burned	552.10	260.62	100.0	999.0	899.0	325.0	776.0	451.0	-0.015800	-1.205459
Age	38.39	12.17	18.0	59.0	41.0	28.0	49.0	21.0	0.006228	-1.213378
Workout Duration (mins)	64.46	31.81	10.0	119.0	109.0	37.0	92.0	55.0	-0.008161	-1.184800

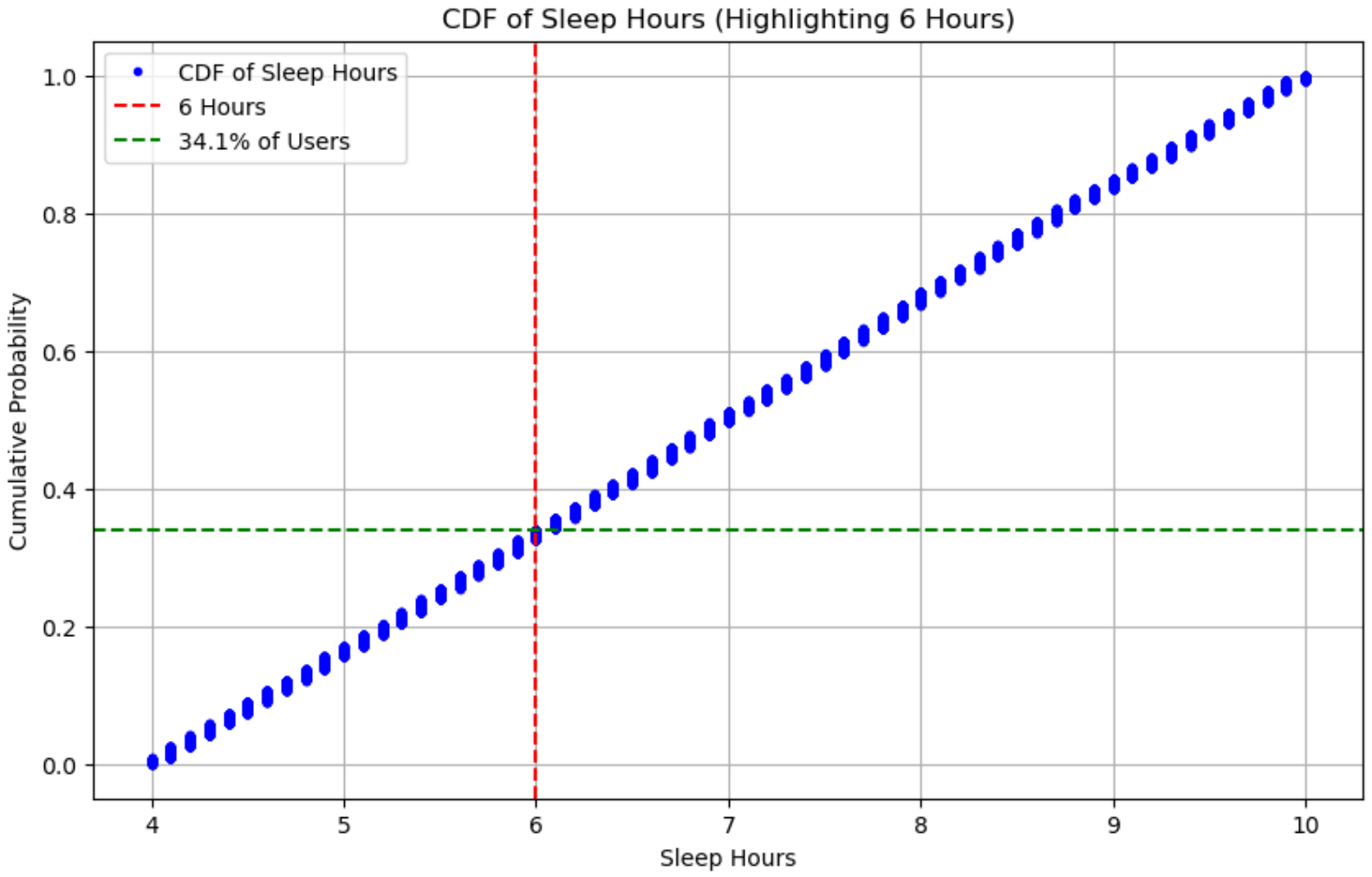
PMF of Sleep Hours Based on Gender

Filter the Data based on Gender- Scenario 1: Sleep Hours for Women.- Scenario 2: Sleep hours for Men.



The KL Divergence of 0.02 and the PMF plot suggest that sleep hours are very similar between women and men.

KL Divergence (Women vs. Men): 0.02



CDF OF SLEEP HOURS

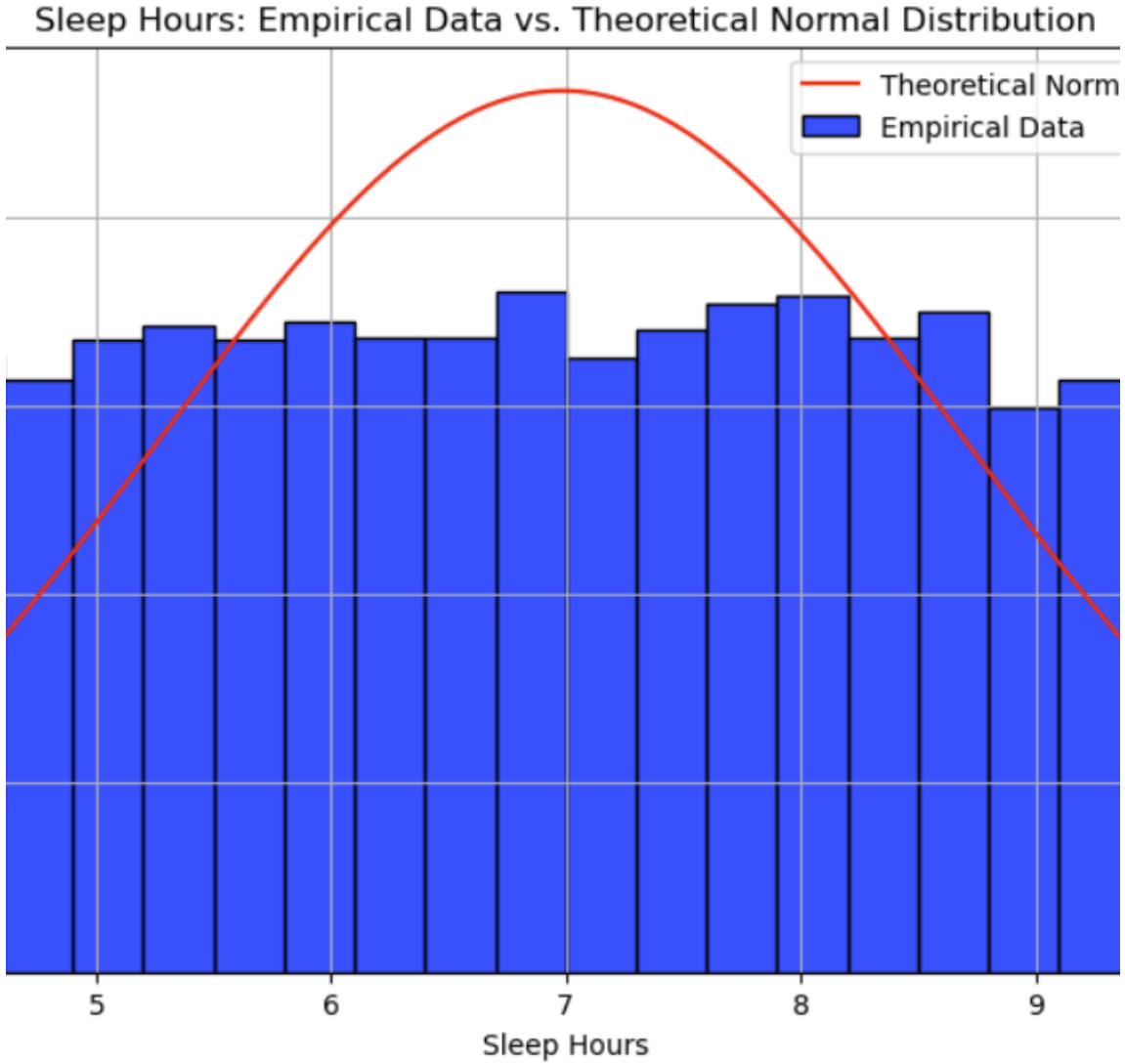
34.1% of users sleep 6 hours or less

Analytical Distribution Plot for Sleep Hours

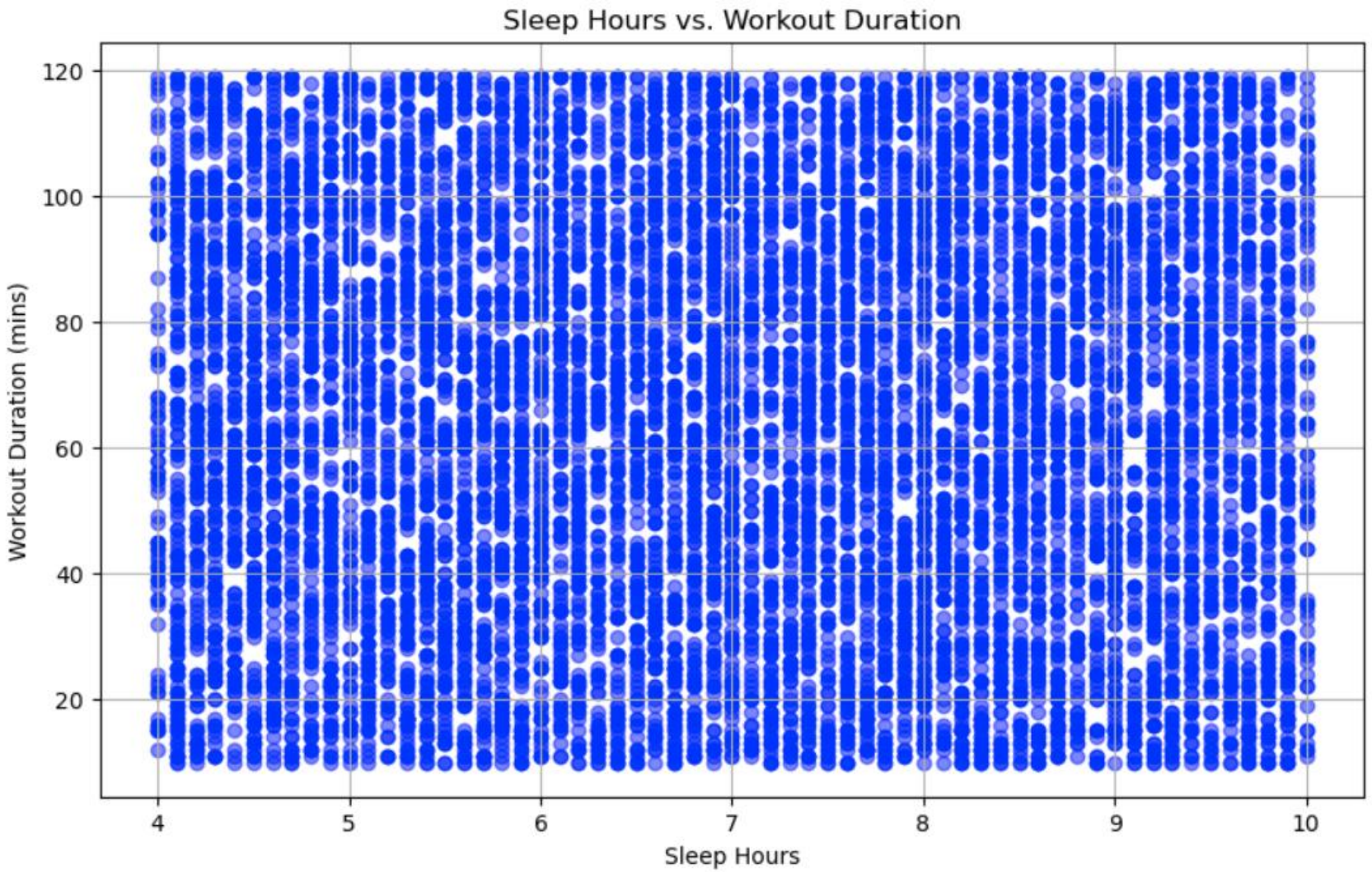
KS Statistic: 0.0646

P-value: 0.0000

The empirical distribution of Sleep Hours is **significantly different** from a Normal Distribution.



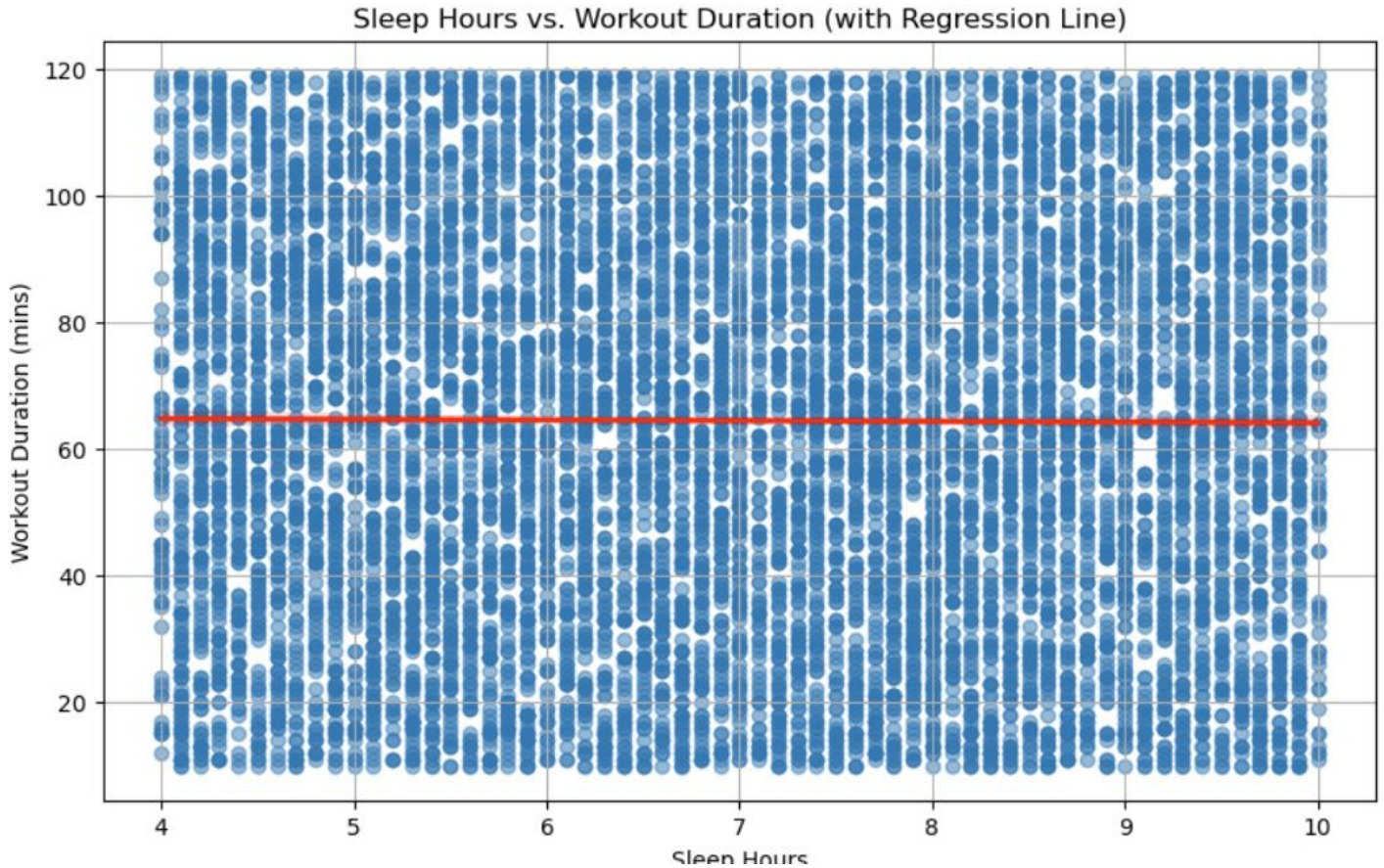
Analysis of the relationship between Sleep Hours and Workout Duration



SLEEP HOURS VS WORKOUT DURATION

- **Covariance:** -0.32
- **Pearson's Correlation**: -0.01
- **P-value:** 0.5538

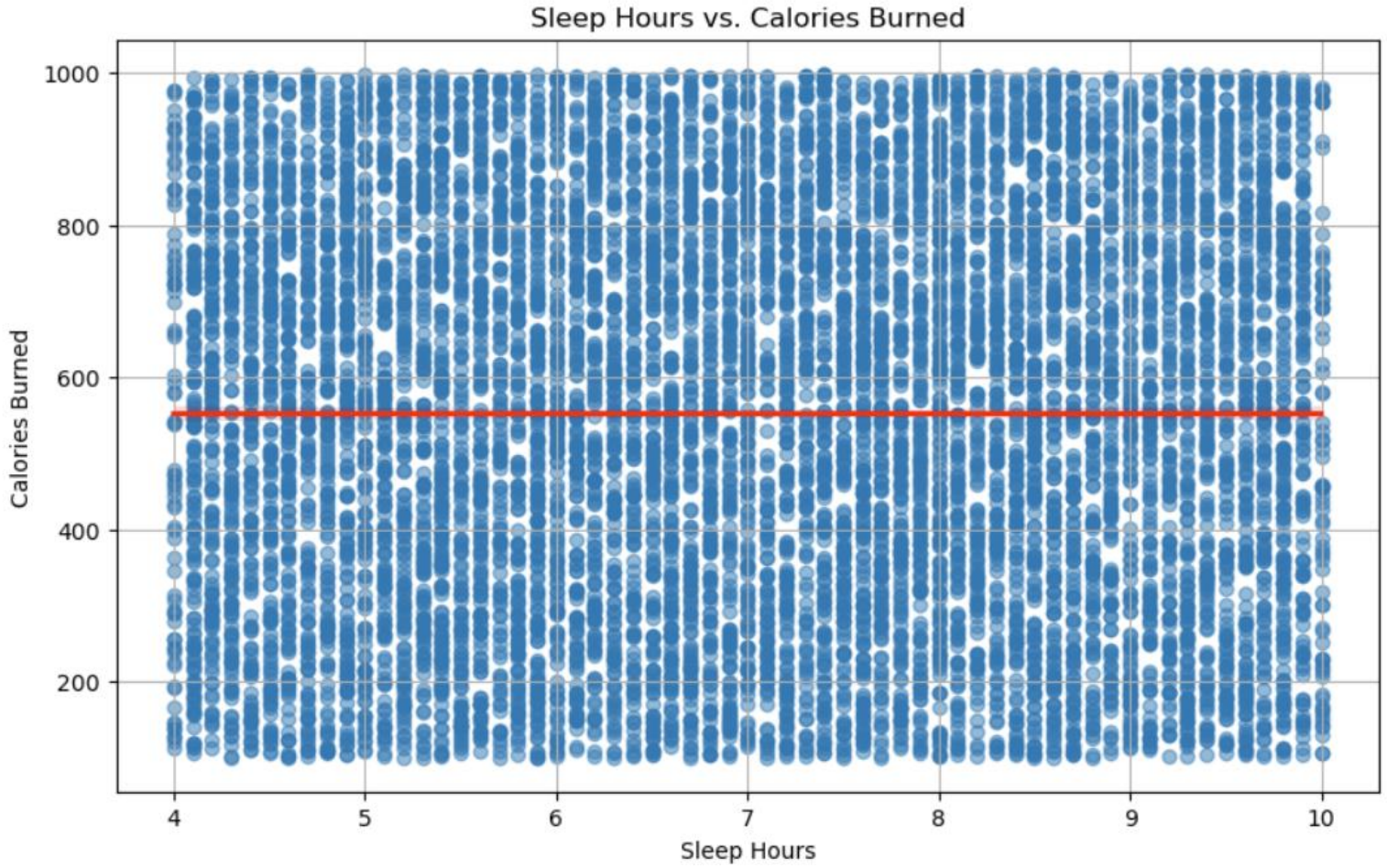
No statistically significant linear relationship between Sleep Hours and Workout Duration.



SLEEP HOURS VS CALORIES BURNED

- **Pearson's Correlation**: -0.00
- **P-value:** 0.9967

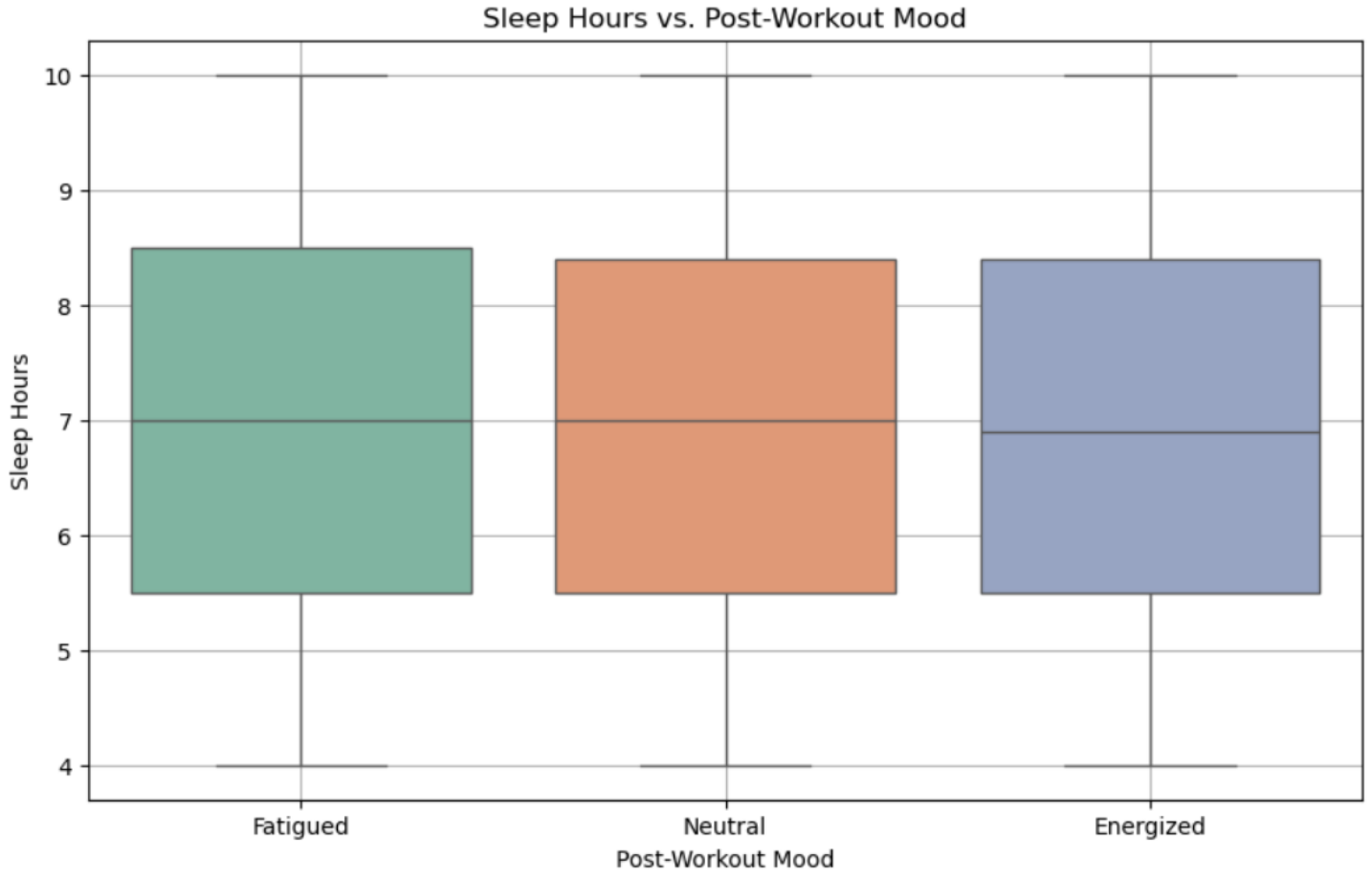
No statistically significant linear relationship between Sleep Hours and Calories Burned



SLEEP HOURS VS POST-WORKOUT MOOD

ANOVA F-statistic: 0.57

No significant difference in post-workout mood across different sleep hour groups.



REGRESSION MODELS

CALORIES BURNED MODEL

- R-squared: 0.000
- Intercept: 544.62
- Sleep_Hours coefficient: -0.0032
- Age coefficient: 0.1953

The model is **not statistically significant**, and there is no meaningful relationship between sleep, age, and calories burned in this dataset.

Calories Burned Model Summary:

OLS Regression Results

Dep. Variable:	Calories_Burned	R-squared:	0.000
Model:	OLS	Adj. R-squared:	-0.000
Method:	Least Squares	F-statistic:	0.4160
Date:	Thu, 27 Feb 2025	Prob (F-statistic):	0.660
Time:	14:47:07	Log-Likelihood:	-69819.
No. Observations:	10000	AIC:	1.396e+05
Df Residuals:	9997	BIC:	1.397e+05
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	544.6236	13.720	39.695	0.000	517.729	571.518
Sleep_Hours	-0.0032	1.525	-0.002	0.998	-2.993	2.987
Age	0.1953	0.214	0.912	0.362	-0.224	0.615

Omnibus:	9083.566	Durbin-Watson:	1.991
Prob(Omnibus):	0.000	Jarque-Bera (JB):	605.130
Skew:	-0.016	Prob(JB):	3.96e-132
Kurtosis:	1.795	Cond. No.	216.

WORKOUT DURATION MODEL

- R-squared: 0.000
- Intercept: 64.13
- Sleep_Hours coefficient: -0.1098
- Age coefficient: 0.0288

The model is **not statistically significant** and does not have any meaningful predictive power for workout duration based on sleep hours and age.

Workout Duration Model Summary:

OLS Regression Results

Dep. Variable:	Workout_Duration_mins	R-squared:	0.000
Model:	OLS	Adj. R-squared:	-0.000
Method:	Least Squares	F-statistic:	0.7816
Date:	Thu, 27 Feb 2025	Prob (F-statistic):	0.458
Time:	14:47:07	Log-Likelihood:	-48787.
No. Observations:	10000	AIC:	9.758e+04
Df Residuals:	9997	BIC:	9.760e+04
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	64.1263	1.675	38.290	0.000	60.843	67.409
Sleep_Hours	-0.1098	0.186	-0.590	0.555	-0.475	0.255
Age	0.0288	0.026	1.101	0.271	-0.022	0.080

Omnibus:	7616.132	Durbin-Watson:	1.996
Prob(Omnibus):	0.000	Jarque-Bera (JB):	585.162
Skew:	-0.007	Prob(JB):	8.58e-128
Kurtosis:	1.815	Cond. No.	216.

MOOD AFTER WORKOUT MODEL

- R-squared: 0.000
- Intercept: 1.0125
- Sleep_Hours coefficient: -0.0045
- Age coefficient: 0.0005

There is **no significant** relationship between sleep, age, and mood after the workout in this dataset.

Mood After Workout Model Summary:

OLS Regression Results

Dep. Variable:	Mood_After_Workout	R-squared:	0.000
Model:	OLS	Adj. R-squared:	-0.000
Method:	Least Squares	F-statistic:	0.7136
Date:	Thu, 27 Feb 2025	Prob (F-statistic):	0.490
Time:	14:47:07	Log-Likelihood:	-12144.
No. Observations:	10000	AIC:	2.429e+04
Df Residuals:	9997	BIC:	2.432e+04
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975
Intercept	1.0125	0.043	23.596	0.000	0.928	1.09
Sleep_Hours	-0.0045	0.005	-0.939	0.348	-0.014	0.00
Age	0.0005	0.001	0.737	0.461	-0.001	0.00

Omnibus:	62003.557	Durbin-Watson:	1.972
Prob(Omnibus):	0.000	Jarque-Bera (JB):	930.591
Skew:	-0.000	Prob(JB):	8.41e-203
Kurtosis:	1.506	Cond. No.	216.

Incorrect Assumptions:

The analysis assumed that sleep duration alone would be a significant predictor of workout quality. However, the results suggest that sleep duration, as measured in this dataset, does not have a meaningful impact on calories burned, workout duration, or post-workout mood. This indicates that other factors, such as sleep quality or overall health, may play a more critical role.

Challenges and Understanding:

One of the main challenges was interpreting the lack of significant relationships between sleep duration and the dependent variables. It was initially assumed that sleep would have a more pronounced effect on workout outcomes. Additionally, understanding the implications of the low R-squared values in the regression models was challenging, as they indicated that the models had little to no predictive power. This raised questions about whether the dataset was sufficient to answer the research question or if other variables needed to be considered.

Conclusion

While the analysis provided valuable insights into the dataset, it also highlighted the complexity of factors influencing workout quality and the need for more comprehensive data to fully understand the relationship between sleep and exercise.