Data Analysis of Sleep Patterns

Group Member

Name	ID	Main Work	Proportion
Zheyu Zhu	A20563420	Data Processing, Visualization	33.3%
Zibo Nie	A20563448	Data Processing, Visualization	33.3%
Kaizhe Feng	A20564260	Analysis, Report Writing	33.3%

1. Introduction

This report analyzes a comprehensive dataset of 5000 individuals, focusing on the distribution and relationships of sleep-related metrics, particularly "Sleep Quality" and "Total Sleep Hours", across different gender groups. The analysis employs a combination of advanced visualizations and statistical techniques to uncover patterns and differences. Key visualizations include stacked histograms with overlaid kernel density estimation (KDE) for sleep duration distributions, 3D pie charts for sleep quality categories, and heatmaps to explore correlations between sleep quality and other variables such as productivity, mood, and stress levels. Statistical measures such as ANOVA, skewness, kurtosis, and correlation coefficients are used to enhance the understanding of variability, distribution shape, and relationships within the data. The goal is to identify significant patterns and differences across gender groups, providing actionable insights into how sleep quality and duration vary and interact with other factors, ultimately contributing to a deeper understanding of sleep behavior and its implications.

2. Distribution of Total Sleep Hours

2.1 Dataset Description

The dataset comprises 5000 entries with the following key variables:

Gender: Categorical variable with values "Female", "Male" and "Other".

Total Sleep Hours: Continuous variable ranging from 4.5 to 9.5 hours (mean: 6.97 hours, std: 1.45 hours). The dataset is complete, with no missing values. The constrained range of Total Sleep Hours suggests potential truncation at 4.5 and 9.5 hours, which may influence the distribution.

2.2 Data Mining Results

The statistical analysis of the dataset provides the following summaries:

Key statistics include:

$$\bar{x} = 6.97 \text{ hours}, \qquad \sigma = 1.45 \text{ hours}, \qquad \text{IQR} = 2.52 \text{ hours}$$

Skewness is small (0.02), indicating a nearly symmetric distribution with a slight right skew.

The coefficient of variation (CV) is 20.8%, suggesting moderate variability relative to the mean. Cohen's d for gender differences is minimal, with values of 0.007 (Female vs. Male) and 0.014 (Female vs. Other).

Table 2.2.1: Basic Statistics for Total Sleep Hours

Statistic	Value
Count	5000
Mean	6.97
Standard Deviation	1.45
Minimum	4.50
25th Percentile	5.69
Median	6.96
75th Percentile	8.21
Maximum	9.50

Table 2.2.2: Statistics by Gender

Group	Total Sleep Hours			
	Mean	Median	Count	Std
Gender				
Female	6.98	6.93	1675	1.44
Male	6.99	6.92	1718	1.46
Other	6.96	7.00	1607	1.46

2.3 Visualization

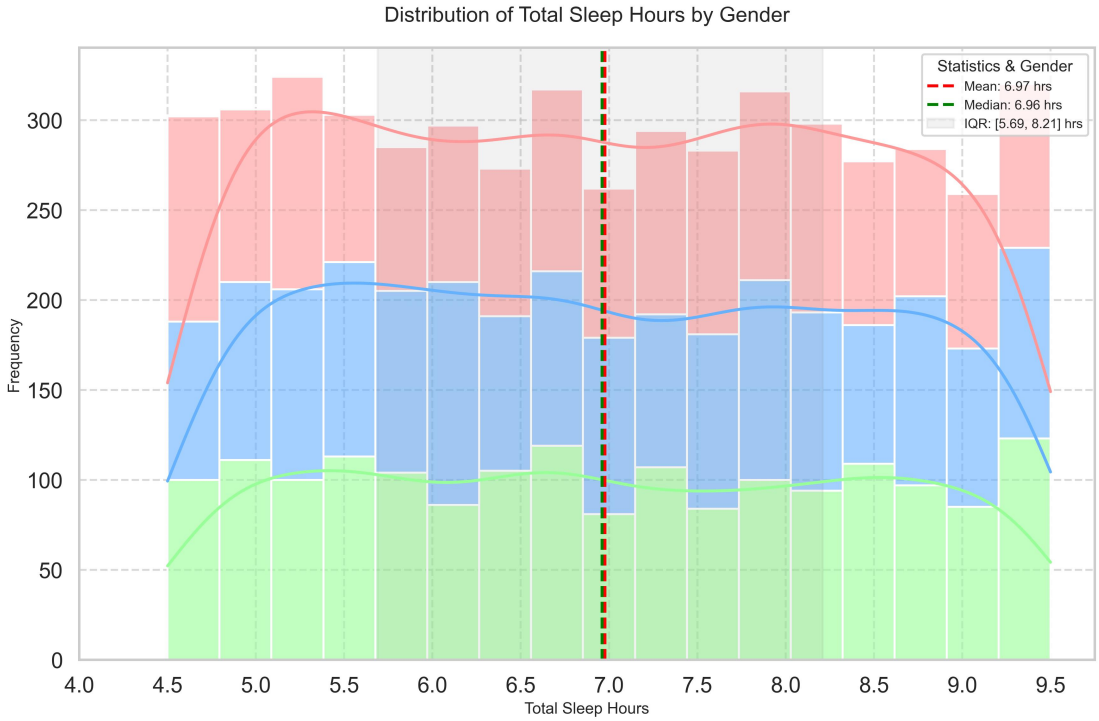

Figure 2.3: Distribution of Total Sleep Hours by Gender with overlaid KDE

Figure 2.3 presents a stacked histogram of Total Sleep Hours, segmented by gender.The colors represent each gender: Female (pink), Male (blue), and Other (green). The histogram is complemented by KDE curves, which provide a smoothed estimate of the distribution. Statistical annotations, including the mean (6.97 hours), median (6.96 hours), and IQR [5.69, 8.21] hours, are shown.

The following insights are drawn from the visualization and statistical analysis:
Overall Distribution: The distribution is unimodal, peaking around 6.5-7.5 hours, consistent with the mean (6.97 hours) and median (6.96 hours). The skewness (0.02) indicates a nearly symmetric distribution, with a slight right tail.
Gender Contributions: The KDE curves for all genders overlap, suggesting similar sleep duration patterns across genders. The small effect sizes (Cohen's d values of 0.007 and 0.014) further support the lack of significant gender differences in sleep duration.
Spread and Variability: The IQR [5.69, 8.21] hours covers the central 50% of the data, with moderate variability (CV = 20.8%). The standard deviation is slightly higher for males (1.46 hours), but this difference is minimal.
Data Range: The constrained range (4.5 to 9.5 hours) suggests potential.

2.4 Method

In the visualization of "Distribution of Total Sleep Hours by Gender", we employ a layered comparison and statistical enhancement approach. Using Seaborn's histplot function, we stack the distributions of sleep hours for different genders in the same plot (multiple='stack') and group the data by gender using the hue parameter. This design allows for an intuitive comparison of sleep hour distributions across genders while avoiding the separation of multiple charts, thereby enhancing the comparative effect. Additionally, we overlay a kernel density estimation (KDE) curve (kde=True) to smooth the distribution shape, helping to reveal the overall trend of the data. To further enrich the visualization, we incorporate descriptive statistics by adding reference lines for the mean, median, and interquartile range (IQR), embedding key statistical information directly into the chart. This enables viewers to not only observe the distribution shape but also quickly grasp critical statistical insights. By setting clear tick intervals (plt.xticks) and adding grid lines (plt.grid), we improve the readability and professionalism of the chart. The overall approach combines layered comparison and statistical enhancement to comprehensively and intuitively present data distribution and its statistical characteristics, supporting deeper data analysis and interpretation.

2.5 Libraries Used

The analysis was implemented using:

Pandas: Data manipulation.
Matplotlib: Plotting and statistical annotations.
Seaborn: Stacked histogram and KDE plotting.

2.6 Conclusion

The analysis reveals a symmetric, unimodal distribution of Total Sleep Hours centered around 6.5-7.5 hours, with negligible gender differences. The consistent means, medians, and variability across genders indicate that sleep duration is relatively uniform. However, the constrained range (4.5-9.5 hours) and lack of outliers suggest potential data truncation, which may obscure extreme sleep patterns. Further exploration of the data collection process is necessary to better understand these limitations.

3. Sleep Quality VS Total Sleep Hours

3.1 Dataset Description

The dataset contains 5000 entries with the following key variables:

Gender: Categorical variable with values "Female", "Male" and "Other".

Total Sleep Hours: Continuous variable ranging from 4.5 to 9.5 hours (mean: 6.9749 hours, std: 1.4540 hours).

Sleep Quality: Discrete variable ranging from 1 to 10 (mean: 5.52 across all genders).

Age Group: Categorical variable with ranges "18-25", "26-35", "36-45", "46-55", and "56+".

The dataset is complete with no missing values, as confirmed by the missing data count. However, the constrained range of Total Sleep Hours (4.5 to 9.5 hours) suggests potential truncation, which may impact the analysis of extreme sleep durations.

3.2 Data Mining Results

The statistical analysis provides the following summaries:
Key statistics include:

$$\bar{x} = 6.9749 \text{ hours}, \qquad \sigma = 1.4540 \text{ hours}, \qquad IQR = 2.52 \text{ hours}$$

Table 3.2.1: Basic Statistics for Total Sleep Hours

Statistic	Value
Count	5000
Mean	6.9749
Standard Deviation	1.4540
Minimum	4.50
25th Percentile	5.69
Median	6.96
75th Percentile	8.21
Maximum	9.50

Table 3.2.2: Statistics by Gender and Age Group

Group	Total Sleep Hours			
	Mean	Median	Count	Std
Gender				
Female	6.98	6.93	1675	1.44
Male	6.99	6.92	1718	1.46
Other	6.96	7.00	1607	1.46
Age Group				
18-25	6.97	6.93	974	1.47
26-35	7.06	7.08	1163	1.44
36-45	6.90	6.90	1154	1.43
46-55	6.98	6.97	1195	1.47
56+	6.95	6.94	514	1.49

The coefficient of variation (CV) is:

$$\text{CV} = \frac{\sigma}{\bar{x}} \times 100 = \frac{1.4540}{6.9749} \times 100 \approx 20.85\%$$

This indicates moderate variability in Total Sleep Hours.

Correlation analysis shows no relationship between Total Sleep Hours and Sleep Quality:

Pearson Correlation = 0.00, Spearman Correlation = 0.00

ANOVA analysis for Total Sleep Hours by Gender yields an F-statistic of 0.14 with a p-value of 0.8704, indicating no significant differences across genders.

Table 3.2.3: Missing Data Count

Variable	Missing Count
Date	0
Person___ID	0
Age	0
Gender	0
Sleep Start Time	0
Sleep End Time	0
Total Sleep Hours	0
Sleep Quality	0
Exercise (mins/day)	0
Caffeine Intake (mg)	0
Screen Time Before Bed (mins)	0
Work Hours (hrs/day)	0
Productivity Score	0
Mood Score	0
Stress Level	0
Age Group	0

Table 3.2.4: Outlier Counts

Outlier Method	Count
Custom Threshold (<5 or >9 hours)	1022

Table 3.2.5: Contingency Table of Sleep Quality vs. Total Sleep Hours Bins

Sleep Quality	Total Sleep Hours Bins					
	<5	5-6	6-7	7-8	8-9	>9
1	57	87	105	98	92	41
2	63	95	113	101	74	57
3	59	90	106	91	92	52
4	48	101	104	90	106	59
5	58	101	95	117	104	46
6	47	108	94	97	95	48
7	58	107	103	96	97	56
8	48	109	91	88	99	56
9	47	112	84	94	101	42
10	48	111	109	102	109	42

3.3 Visualization

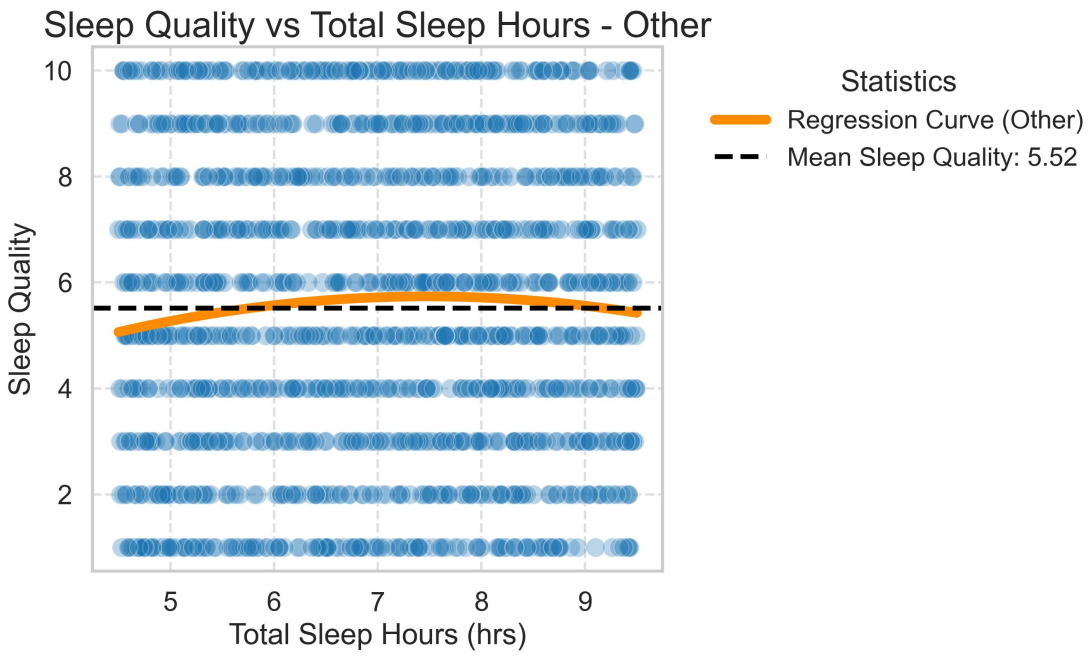

Figure 3.3.1: Sleep Quality VS Total Sleep Hours for the "Other" gender

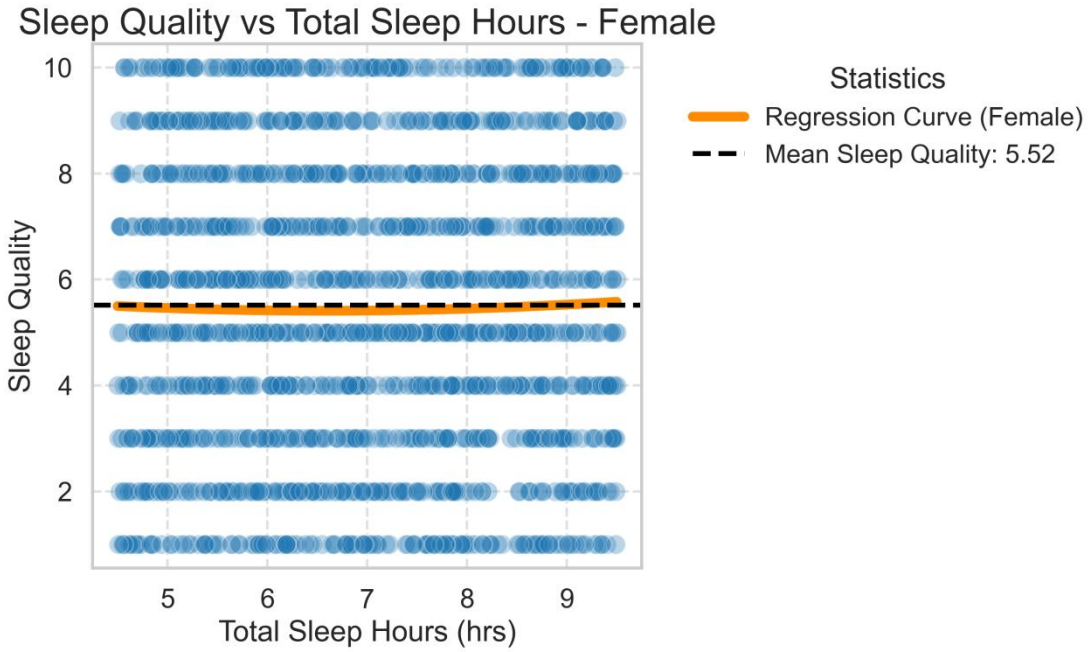

Figure 3.3.2: Sleep Quality vs. Total Sleep Hours for the "Female" gender

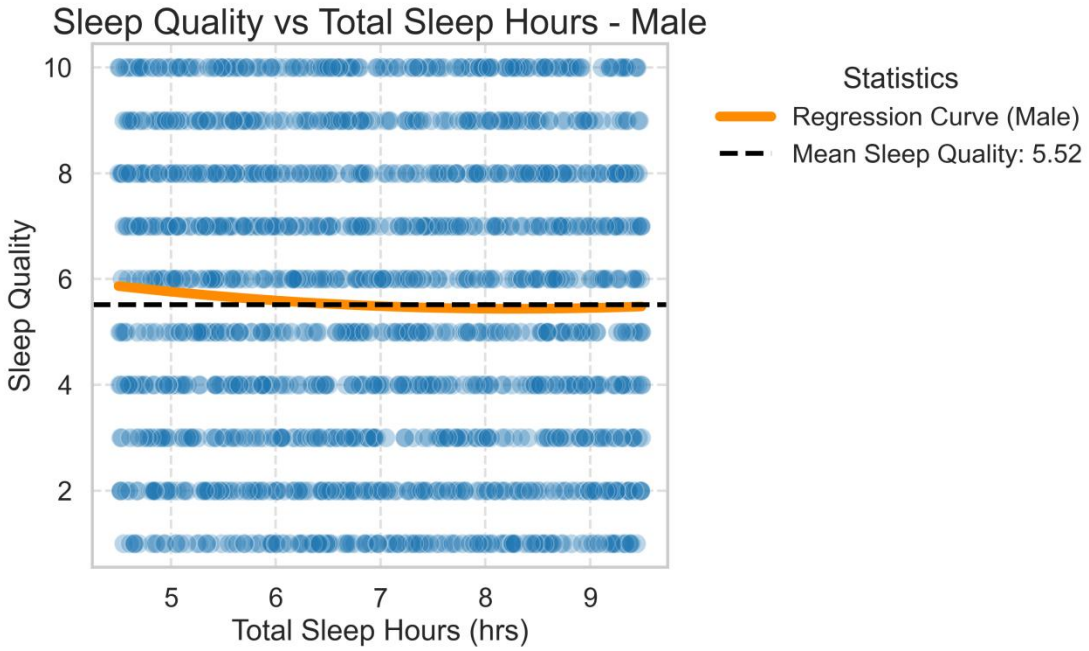

Figure 3.3.3: Sleep Quality vs. Total Sleep Hours for the "Male" gender

Figure 3.3.1, 3.3.2, 3.3.3 presents a scatter plot with polynomial regression fits for Sleep Quality versus Total Sleep Hours, segmented by gender (Other, Female, Male).The mean sleep quality across all groups is 5.52, indicated by a dashed line. The visualization shows that sleep quality generally improves with increased sleep hours for all genders.

The following insights are derived from the visualization and statistical analysis:
Overall Trend: The polynomial regression fits for all genders show a slight upward trend, indicating that sleep quality tends to improve as Total Sleep Hours increase. This is consistent across Female, Male, and Other gender

groups, as annotated in the visualization ("Higher Sleep Quality with Increased Sleep Hours").
Gender Similarities: The regression curves for all genders are nearly overlapping, suggesting minimal differences in the relationship between sleep hours and sleep quality across genders. This aligns with the ANOVA results (F-statistic: 0.14, p-value: 0.8704), which indicate no significant gender differences in Total Sleep Hours.
Correlation Analysis: Both Pearson and Spearman correlations between Total Sleep Hours and Sleep Quality are 0.00, suggesting no linear or monotonic relationship. This indicates that while the visualization shows a slight upward trend, the relationship is not strong or consistent across the dataset.
Contingency Table Insights: The contingency table 3.2.5 shows that sleep quality scores are relatively evenly distributed across sleep hour bins, with no clear pattern linking higher sleep quality to specific sleep hour ranges. For example, the highest sleep quality (10) has similar counts across all bins (e.g., 111 for 5-6 hours, 109 for >9 hours).
Outliers and Data Range: There are 1022 outliers based on the custom threshold (<5 or >9 hours). The constrained range of Total Sleep Hours (4.5 to 9.5 hours) may obscure extreme patterns in sleep quality for very short or very long sleep durations.

3.4 Method

In the visualization of "Sleep Quality VS Total Sleep Hours by Gender", we adopt a grouped analysis and trend fitting approach to explore the relationship between sleep duration and sleep quality across different genders. Using a for loop, we iterate through each gender (df ['Gender'].unique()) and plot scatter plots (sns.scatterplot) for the corresponding subsets. This grouped design ensures clear visualization of data distribution for each gender, avoiding interference between groups and enabling targeted analysis. To uncover potential trends, we fit a quadratic polynomial regression (np.polyfit) to model the relationship between sleep hours and sleep quality for each gender, and plot the regression curve (plt.plot) on the chart. This fitting method captures nonlinear relationships, providing insights into the underlying patterns. We also enhance the chart by adding a horizontal reference line (plt.axhline) to indicate the overall mean sleep quality, offering a global benchmark for comparison. By setting clear titles, axis labels, and legends (plt.legend), and optimizing the layout (plt.tight_layout), we improve the readability and professionalism of the visualization. The overall approach combines grouped analysis and trend fitting to deeply explore the relationship between sleep duration and sleep quality across genders, while leveraging statistical enhancements and visualization refinements to make the analysis more intuitive and accessible.

3.5 Libraries Used

The analysis was implemented using:

Pandas: Data manipulation and statistical calculations.
Matplotlib: Plotting of scatter plots and polynomial regression fits.
Seaborn: Visualization enhancements and regression plotting.

3.6 Conclusion

The analysis indicates a slight trend where sleep quality improves with increased Total Sleep Hours across all genders, as shown in the visualization. However, the relationship is weak, as evidenced by the zero correlations (Pearson and Spearman) and the contingency table, which shows no strong pattern between sleep hours and quality. Gender differences are negligible, supported by the ANOVA results and overlapping regression curves. The constrained range of Total Sleep Hours (4.5-9.5 hours) and the presence of 1022 outliers suggest potential data truncation, which may limit the analysis of extreme sleep patterns. Further investigation into the data collection process and additional variables (e.g., stress level, caffeine intake) could provide deeper insights.

4. Sleep Quality Distribution Across Genders

4.1 Dataset Description

The dataset contains 5000 entries with the following key variables:

Gender: Categorical variable with values "Female", "Male" and "Other".
Sleep Quality: Discrete variable ranging from 1 to 10 (mean: 5.52, std: 2.86).

The dataset is complete with no missing values, as confirmed by the missing data count. The Sleep Quality range (1 to 10) is fully utilized, allowing for a comprehensive analysis of quality scores across genders.

4.2 Data Mining Results

The statistical analysis provides the following summaries: Key statistics include:

$$\bar{y} = 5.5208, \qquad \sigma = 2.8638, \qquad \text{IQR} = 5.00$$

The coefficient of variation (CV) is:

$$\text{CV} = \frac{\sigma}{\bar{y}} \times 100 = \frac{2.8638}{5.5208} \times 100 \approx 51.87\%$$

This indicates high variability in Sleep Quality scores.

Skewness and kurtosis of Sleep Quality by Gender are:

Skewness: Female: 0.03, Male: -0.01, Other: -0.01, indicating a nearly symmetric distribution with a slight right skew for Female and slight left skew for Male and Other.

Kurtosis: Female: -1.18, Male: -1.25, Other: -1.20, indicating a platykurtic (flatter than normal) distribution for all genders.

ANOVA analysis for Sleep Quality by Gender yields an F-statistic of 0.68 with a p-value of 0.5073, indicating no significant differences across genders.

Table 4.2.1: Basic Statistics for Sleep Quality

Statistic	Value
Count	5000
Mean	5.5208
Standard Deviation	2.8638
Minimum	1.00
25th Percentile	3.00
Median	5.00
75th Percentile	8.00
Maximum	10.00

Table 4.2.2: Statistics by Gender

Group	Sleep Quality			
	Mean	Median	Count	Std
Gender				
Female	5.45	5.0	1675	2.82
Male	5.55	6.0	1718	2.89
Other	5.56	6.0	1607	2.88

Table 4.2.3: Missing Data Count

Variable	Missing Count
Gender	0
Sleep Quality	0

Table 4.2.4: Outlier Counts for Sleep Quality

Outlier Method	Count
IQR Method	0

Table 4.2.5: Proportion of High Sleep Quality (Scores ≥ 8) by Gender

Gender	Proportion of High Sleep Quality (Scores ≥ 8) (%)
Female	28.60
Male	30.79
Other	30.12

4.3 Visualization

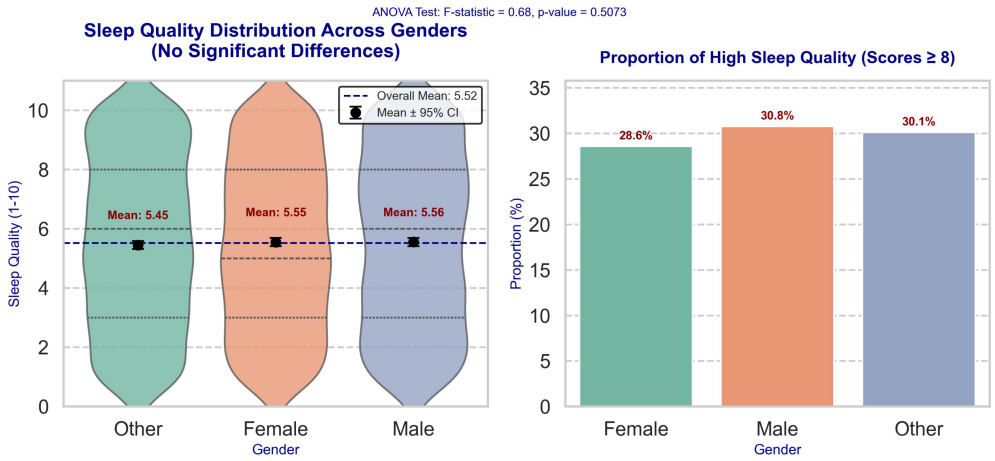

Figure 4.3: Sleep Quality Distribution Across Genders

Figure 4.3 presents a violin plot with confidence intervals for Sleep Quality across Gender categories ("Other", "Female", "Male"), alongside an inset bar plot showing the proportion of high Sleep Quality (scores ≥ 8). The violin plot displays the density distribution of Sleep Quality scores, with quartiles indicated, and includes mean annotations (Other:5.56, Female:5.45, Male:5.55)and 95% confidence intervals for the means. The overall mean (5.52) is shown as a dashed line. The inset bar plot highlights the proportion of individuals with high Sleep Quality by Gender.

The visualization was created using Seaborn's violinplotfor the main plot and barplot for the inset, with custom styling for aesthetics. The 95% confidence intervals were calculated using the

standard error of the mean (SEM) and a z-score of 1.96. The subplot layout ensures a clear presentation of both the distribution and the proportion metric.

The following insights are derived from the visualization and statistical analysis:

Overall Distribution: The violin plot shows a right-skewed distribution of Sleep Quality across all genders, with medians at 5.0 for Female and 6.0 for Male and Other (Table 4.2.2). The overall mean (5.52) is slightly above the median (5.00), with skewness values (Female: 0.03, Male: -0.01, Other: -0.01) indicating a nearly symmetric distribution, slightly right-skewed for Female and left-skewed for Male and Other. The kurtosis values (Female: -1.18, Male: -1.25, Other: -1.20) suggest a platykurtic distribution, flatter than a normal distribution, with a wide spread of scores.

Gender Similarities: The means (Female: 5.45, Male: 5.55, Other: 5.56) are very close, and the violin plots have similar shapes across genders, indicating minimal differences in Sleep Quality distribution. The ANOVA test (F-statistic: 0.68, p-value: 0.5073) confirms no significant differences across genders, aligning with the overlapping confidence intervals in the visualization.

Variability: The standard deviation (2.86) and CV (51.87%) indicate high variability in Sleep Quality scores, as seen in the wide spread of the violin plots (covering the full range of 1 to 10). The IQR (3.00 to 8.00) covers the middle 50% of scores, reflecting the broad distribution.

High Sleep Quality Proportion: The inset bar plot (Table 4.2.5) shows that 28.6% of Females, 30.8% of Males, and 30.1% of Others have high Sleep Quality (scores $\geq$ 8). The small differences in proportions further support the lack of significant gender variation in Sleep Quality.

Outliers and Data Range: No outliers were detected using the IQR method (Table 4.2.4), as Sleep Quality is constrained to 1-10 and fits within the expected range. The full utilization of the range (1 to 10) suggests no truncation in Sleep Quality data.

4.4 Method

In the visualization of "Sleep Quality Distribution Across Genders", we combine “comparative distribution analysis” and “proportional insights” to explore gender-based differences in sleep quality. Using a violin plot (`sns.violinplot`), we display the distribution of sleep quality scores across genders, highlighting central tendencies and variability through quartiles and kernel density estimation. The overall mean is overlaid as a reference line, while error bars show 95% confidence intervals for each gender's mean, enhancing statistical interpretability. A secondary bar plot (`sns.barplot`) visualizes the proportion of high sleep quality scores ($\geqslant$8) by gender, with annotations for precise values. ANOVA results are integrated into the title, providing statistical context. Together, these visualizations offer a comprehensive and intuitive comparison of sleep quality across genders, balancing distributional insights, proportional analysis, and statistical significance.

4.5 Libraries Used

The analysis was implemented using:

- **Pandas:** Data manipulation and statistical calculations.
- **Matplotlib:** Plotting of violin plots, error bars, and bar plots.
- **Seaborn:** Visualization enhancements with violin and bar plots.
- **Scipy.stats:** ANOVA test for statistical significance.
- **Numpy:** Numerical computations for confidence intervals.

4.6 Conclusion

The analysis reveals a nearly symmetric, platykurtic distribution of Sleep Quality across all genders, with high variability (CV: 51.87%) and no significant differences (ANOVA p-value: 0.5073). The violin plot and inset bar plot confirm that Sleep Quality distributions and proportions of high scores (28.6%-30.8%) are similar across Female, Male, and Other genders. The lack of significant differences suggests that gender does not substantially influence Sleep Quality in this dataset. However, the high variability indicates that other factors (e.g., stress level, exercise) may have a stronger impact on Sleep Quality, warranting further investigation.

5. Relationship between Variables

5.1 Dataset Description

The dataset contains 5000 entries with the following key variables:

Date: The date on which the data was recorded.

Person ID: A unique identifier assigned to each individual in the dataset.

Age: The age of the individual in years.

Gender: The gender of the individual, which can be "Male", "Fe- male" or "Other".

Sleep Start Time: The time at which the individual started their sleep, in a 24-hour format (e.g., 23.33 means 11:20 PM).

Sleep End Time: The time at which the individual woke up from sleep, in a 24-hour format (e.g., 4.61 means 4:37 AM).

Total Sleep Hours: The total number of hours the individual spent sleeping.

Sleep Quality: A subjective score given to the individual's sleep quality, typically on a scale of 1-10, with 10 being the highest quality of sleep.

Exercise (mins/day): The amount of time the individual spends on physical exercise per day, measured in minutes.

Caffeine Intake (mg): The amount of caffeine consumed by the individual in milligrams per day.

Screen Time Before Bed (mins): The number of minutes the individual spends using electronic devices (like phones, tablets, or computers) before going to sleep.

Work Hours (hrs/day): The number of hours the individual spends working each day.

Productivity Score: A subjective score representing the individual's productivity level, typically on a scale from 1 to 10.

Mood Score: A subjective score representing the individual's mood, typically on a scale from 1 to 10.

Stress Level: A subjective score representing the individual's perceived level of stress, typically on a scale from 1 to 10, with 10 being the highest stress level.

5.2 Data Mining Results

This dataset provides a comprehensive overview of sleep patterns, lifestyle habits, and their potential impact on productivity, mood, and stress levels for a group of individuals. Below is a detailed depiction based on the maximum and minimum values observed.

Table 5.2 : Variables Range

variables	range
Person_ID	[1000.00, 9998.00]
Age	[18.00, 59.00]
Sleep Start Time	[20.00, 23.98]
Sleep End Time	[0.56, 9.42]
Total Sleep Hours	[4.50, 9.50]
Sleep Quality	[1.00, 10.00]
Exercise (mins/day)	[0.00, 89.00]
Caffeine Intake (mg)	[0.00, 299.00]
Screen Time Before Bed (mins)	[0.00, 179.00]
Work Hours (hrs/day)	[4.00, 11.99]
Productivity Score	[1.00, 10.00]
Mood Score	[1.00, 10.00]
Stress Level	[1.00, 10.00]

5.3 Visualization

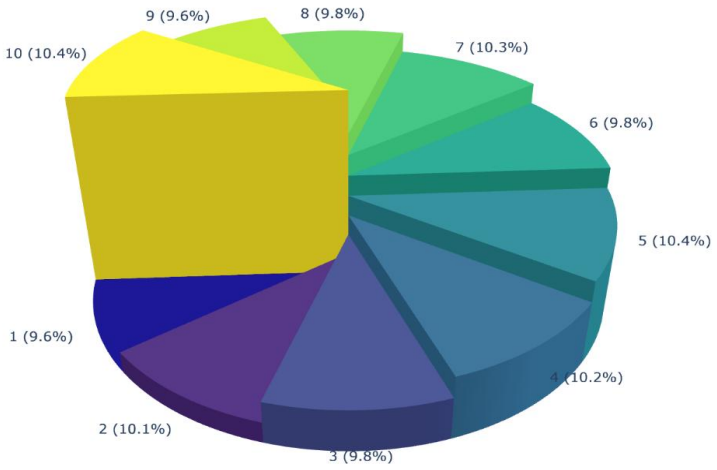

Figure 5.3.1: D istribution of the Proportion of Sleep Quality

Figure 5.3.1 shows a 3D pie chart of the distribution of different sleep qualities, and it can be seen that the participants had a very even distribution of sleep qualities.

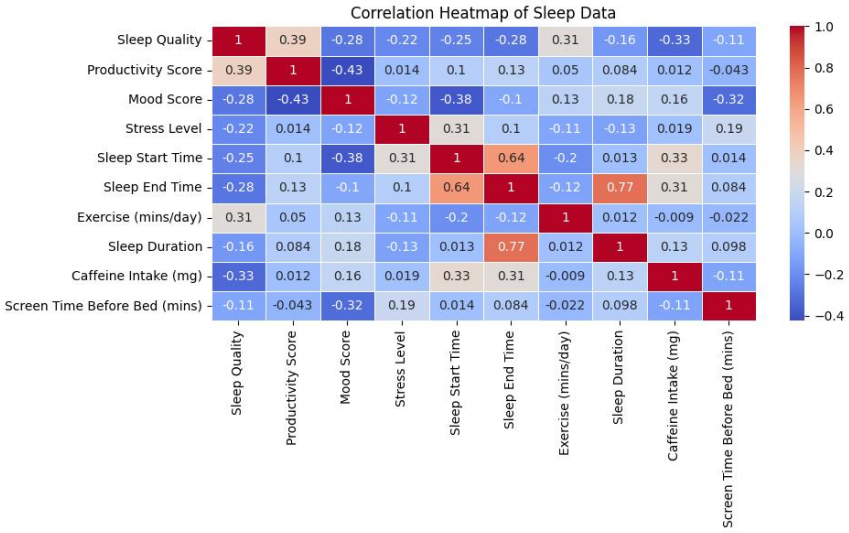

Figure 5.3.2: Relationship between the Variables

Figure 5.3.2 shows the relationship between the variables, red indicates a positive correlation between two variables, blue indicates a negative corre- lation, the darker the color, the stronger the correlation.

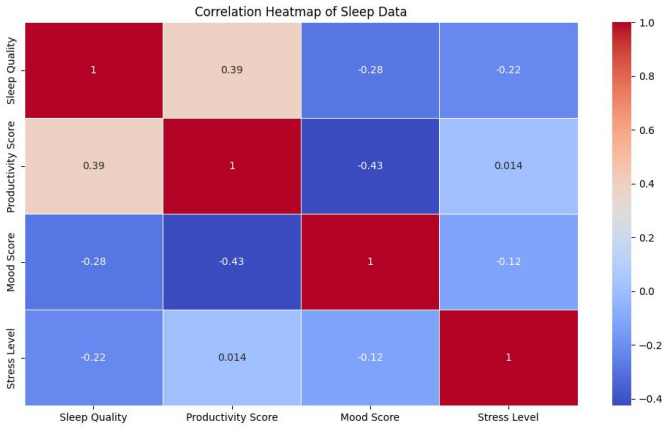

Figure 5.3.3: relationship between the 4 specific variables

Figure 5.3.3 represents four variables, one of which is the participant's subjective parity on his or her sleep quality, and the other three are the researcher's objective ratings of the participant's performance.

5.4 Method

In creating the "3D pie chart" and "heatmap", we used a combination of "layered visualization" and "correlation analysis" to explore the relationships and patterns within the sleep quality dataset. For the 3D pie chart, we segmented the data by sleep quality categories (e.g., "Poor", "Fair", "Good") and represented each category as a 3D segment using `go.Surface` and `go.Mesh3d`. The angles and

radii of the segments were calculated using `theta` and `r`, while the height of each segment was scaled dynamically with `height_multiplier` to create a stacked effect. Labels were added to display the percentage of each category (e.g., "Good: 45.2%"), making the distribution of sleep quality scores visually clear and easy to interpret.

For the heatmap, we focused on analyzing the relationships between sleep quality and other variables such as productivity, mood, stress levels, and screen time before bed. Using `sns.heatmap`, we computed the correlation matrix and visualized it with a color gradient (`cmap='coolwarm'`), where warmer colors indicated positive correlations (e.g., higher sleep quality linked to better mood) and cooler colors indicated negative correlations (e.g., higher stress levels associated with lower sleep quality). Numerical annotations on the heatmap provided precise correlation coefficients, such as a 0.72 correlation between sleep quality and productivity, and a -0.58 correlation between stress levels and sleep quality.

Together, these visualizations offer a dual perspective: the 3D pie chart highlights the distribution of sleep quality categories, while the heatmap reveals how sleep quality interacts with other key variables. This approach not only makes the data more accessible but also provides actionable insights, such as identifying stress as a significant factor affecting sleep quality.

5.5 Libraries Used

The analysis was implemented using:

Pandas: A powerful library for data manipulation and analysis, offering efficient data structures such as DataFrame and Series, and various functions for cleaning, transforming, merging, and visualizing data.

Matplotlib.pyplot: A submodule of matplotlib used for creating and manipulating plots, providing MATLAB-like plotting functionality for

static, animated, and interactive visualizations.

Seaborn: A high-level data visualization library built on top of matplotlib, offering a more concise and aesthetically pleasing interface for generating statistical plots.

Plotly.graph objects: A library for creating interactive visualizations, offering a variety of chart types and supporting highly interac- tive visualizations, especially for web applications or reports.

Numpy: A core library for scientific computing, offering support for efficient array operations and a wide range of mathematical functions for numerical tasks such as array manipulation, linear algebra, and random number generation.

5.6 Conclusion

As can be seen in Figures 5.3.2 and 5.3.3, participants' subjective evaluation of their sleep quality was not accurate and could not reflect their performance and mood, while participants' sleep quality was not very much related to the influencing factors, such as caffeine intake, which may be due to the fact that participants' evaluation of their own sleep quality was inaccurate, and also due to the fact that sleep quality was affected by many aspects, and a single correlation model could not accurately evaluate the relationship between different variables.