Descriptive Statistics & Visualization

Descriptive Statistics

Setup

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 108 entries, 0 to 107
Data columns (total 21 columns):
 #   Column                           Non-Null Count  Dtype  
---  ------                           --------------  -----  
 0   Timestamp                        108 non-null    object 
 1   Age Group                        108 non-null    object 
 2   Gender                           108 non-null    object 
 3   Education Level                  108 non-null    object 
 4   Occupation                       108 non-null    object 
 5   Exercise Days/Week               108 non-null    object 
 6   Device Usage (hrs/day)           108 non-null    object 
 7   Screen Time Before Sleep         108 non-null    object 
 8   Height (cm)                      83 non-null     float64
 9   Weight (kg)                      92 non-null     float64
 10  Bedtime                          108 non-null    object 
 11  Wake-up Time                     108 non-null    object 
 12  Sleep Onset Time                 108 non-null    object 
 13  Nap Duration                     108 non-null    object 
 14  Sleep Duration (hrs/24hr)        107 non-null    object 
 15  Sleep Quality                    108 non-null    int64  
 16  Sleep Disturbances               108 non-null    object 
 17  Sleep Medication                 108 non-null    object 
 18  Language                         108 non-null    object 
 19  BMI                              80 non-null     float64
 20  Calculated Night Sleep Duration  105 non-null    float64
dtypes: float64(4), int64(1), object(16)
memory usage: 17.8+ KB
```

Some random samples from the dataset:

	Age Group	Gender	Education Level	Occupation	Exercise Days/Week	Device Usage (hrs/day)	Screen Time Before Sleep	Heig (cı
13	25-34	Female	Master's	Student	0 Days	7+ Hours	2+ Hours	162
40	25-34	Male	Doctorate	Professional/Office Worker	0 Days	7+ Hours	2+ Hours	165
81	25-34	Male	Doctorate/Prof.	Student	3-4 Days	1-3 Hours	30-60 Minutes	N
36	35-44	Male	Doctorate	Professional/Office Worker	3-4 Days	4-6 Hours	30-60 Minutes	164
61	35-44	Female	Master's	Professional/Office Worker	1-2 Days	7+ Hours	2+ Hours	175

Overall descriptive stats

	Height (cm)	Weight (kg)	Sleep Quality	BMI	Calculated Night Sleep Duration
count	83.000000	92.000000	108.000000	80.000000	105.000000
mean	165.305542	67.415217	3.444444	24.552500	7.036952
std	8.321679	12.798085	0.824092	4.245503	1.368431
min	150.000000	43.000000	2.000000	17.500000	1.670000
25%	160.000000	59.800000	3.000000	21.500000	6.500000
50%	167.000000	68.000000	3.000000	23.550000	7.000000
75%	171.000000	75.000000	4.000000	26.600000	8.000000
max	185.000000	100.000000	5.000000	39.400000	9.750000

- **Sleep Quality:** On average, respondents rated their sleep quality around 3 on a scale, indicating moderate sleep quality.
- **BMI:** The average Body Mass Index (BMI) is around 23.55, with a range extending from 16.5 to 39.4.
- **Calculated Night Sleep Duration:** The average night sleep duration is around 7 hours, with a wide range from 1.67 hours to almost 9.75 hours.

	Age Group	Gender	Education Level	Occupation	Exercise Days/Week	Device Usage (hrs/day)	Screen Time Before Sleep	Bedtime
count	108	108	108	108	108	108	108	108
unique	5	3	5	7	4	4	4	18
top	25-34	Male	Master's	Student	1-2 Days	7+ Hours	30-60 Minutes	23:00
freq	72	67	47	47	43	43	45	24

- **Age Group:** The most common age group among respondents is 25-34.
- **Gender:** A slightly higher number of male respondents compared to females.
- **Education Level:** The majority of respondents have a Master's degree.
- **Occupation:** Many respondents are students.
- **Exercise Days/Week:** '1-2 Days' is the most common response for exercise frequency.
- **Device Usage (hrs/day):** A large portion of respondents use devices for '7+ Hours' per day.
- **Screen Time Before Sleep:** '30-60 Minutes' is the most common duration for screen time before sleep.
- **Sleep Disturbances:** 'Rarely' is the most frequent response, indicating that most respondents rarely experience sleep disturbances.
- **Sleep Medication:** The majority of respondents do not use sleep medication.
- **Language:** English is the most common language among respondents.

Sleep related fields

	Sleep Duration (hrs/24hr)	Sleep Quality	Sleep Disturbances	Sleep Medication	Calculated Night Sleep Duration
count	107	108.000000	108	108	105.000000
unique	3	NaN	5	2	NaN
top	6+ Hours	NaN	Rarely	No	NaN
freq	64	NaN	48	105	NaN
mean	NaN	3.444444	NaN	NaN	7.036952
std	NaN	0.824092	NaN	NaN	1.368431
min	NaN	2.000000	NaN	NaN	1.670000
25%	NaN	3.000000	NaN	NaN	6.500000
50%	NaN	3.000000	NaN	NaN	7.000000
75%	NaN	4.000000	NaN	NaN	8.000000
max	NaN	5.000000	NaN	NaN	9.750000

Sleep Duration:

The most common reported sleep duration is '6+ Hours'. This suggests that a majority of the respondents are getting the minimum recommended amount of sleep for adults, which is usually around 7-9 hours. However, without more specific data on those who sleep '6+ Hours' (e.g., whether they are closer to 6 or 9 hours), it's hard to assess the adequacy of sleep duration precisely.

Sleep Quality:

The average sleep quality score is 3.44 out of 5. This indicates a moderate level of sleep quality among the respondents. A score closer to 5 would suggest better sleep quality, so there's room for improvement. The presence of some variability (standard deviation of 0.82) suggests differing sleep quality experiences among respondents.

Sleep Disturbances:

'Rarely' being the most common response for sleep disturbances is a positive sign, suggesting that most respondents do not frequently experience sleep disturbances.

Sleep Medication:

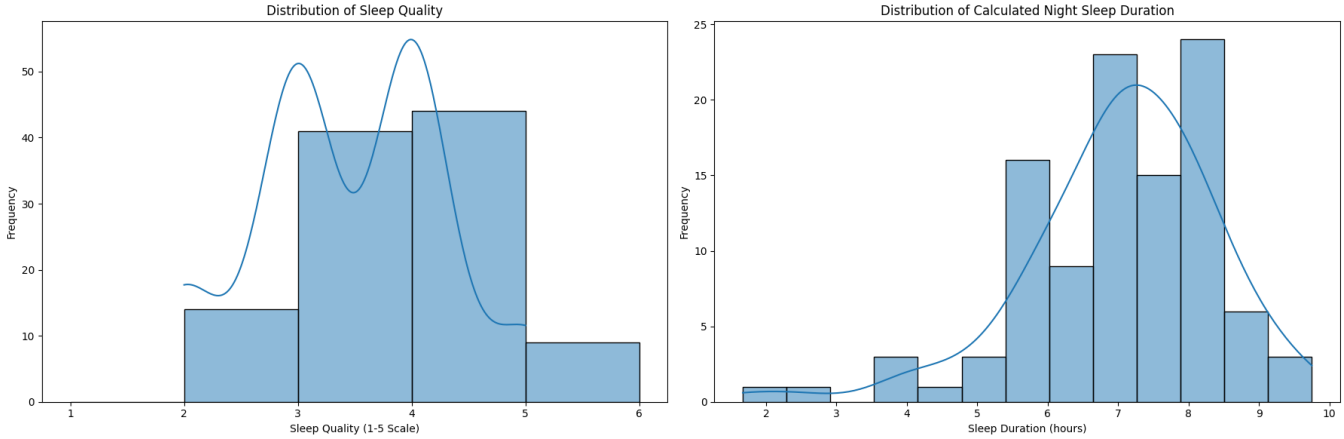
The fact that the vast majority of respondents do not use sleep medication (105 out of 108) could indicate that sleep issues are not severe enough to require medication, or there might be a preference for non-medical approaches to sleep improvement.

Calculated Night Sleep Duration:

The calculated average night sleep duration is approximately 7 hours, which aligns with general sleep recommendations. However, the range (minimum of 1.67 hours and a maximum of 9.75 hours) indicates significant variability among respondents. Overall, these statistics suggest a relatively positive picture in terms of sleep quantity (with most respondents getting 6 or more hours of sleep) and a moderate level of sleep quality. However, the variability in sleep quality and duration indicates that experiences vary significantly among individuals. This variability could be explored further to understand what factors (like lifestyle, diet, or exercise) might be influencing sleep patterns.

Visualization

Overview



Distribution of Sleep Quality:

The sleep quality scores are distributed mainly between 2 and 4, with the majority of respondents reporting a sleep quality of 3. There's a smaller number of respondents with sleep quality scores of 5, indicating excellent sleep quality.

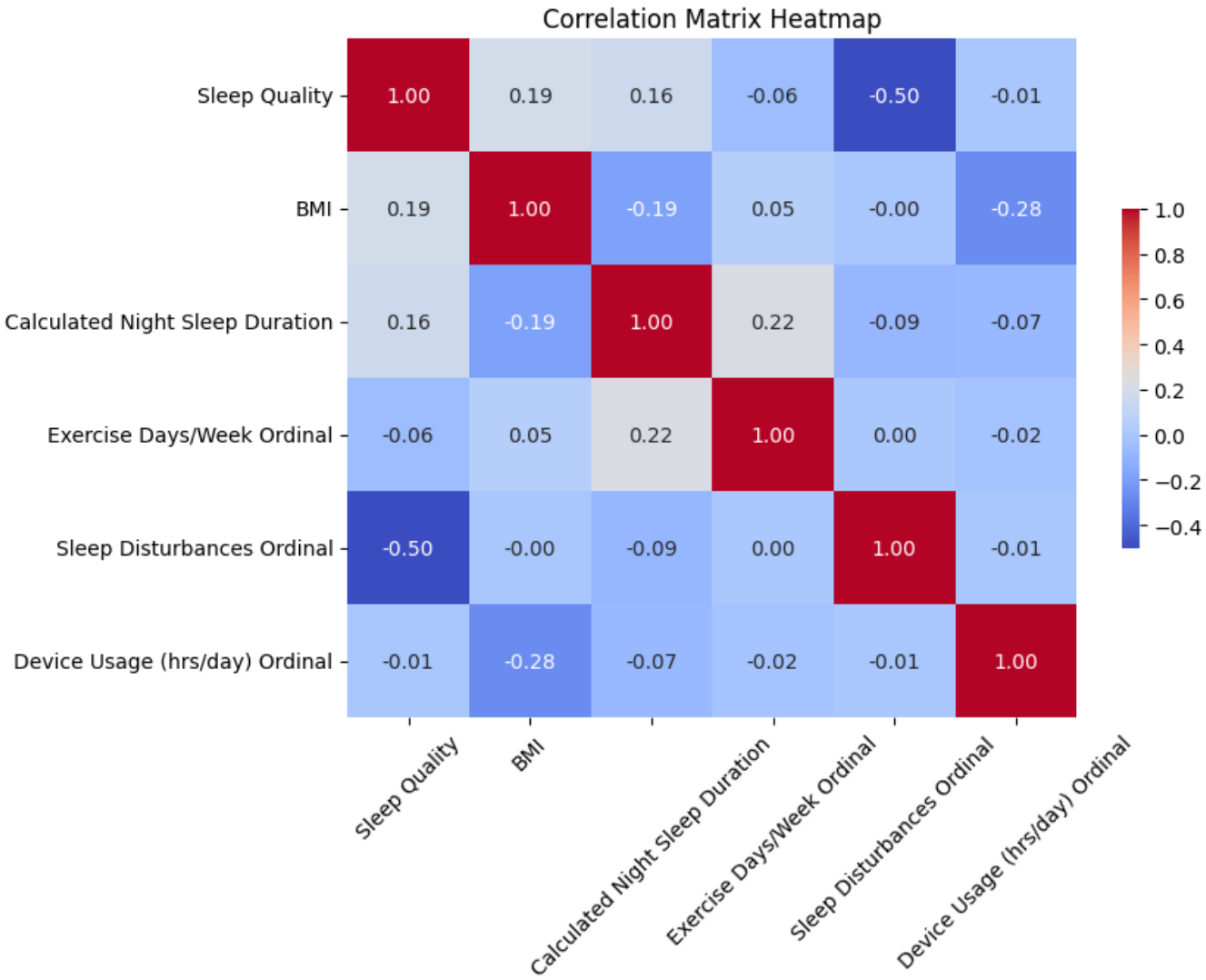
Distribution of Calculated Night Sleep Duration:

The histogram shows a fairly normal distribution centered around 7 hours, which aligns with general sleep recommendations. There are fewer instances of very short (<5 hours) or very long (>9 hours) sleep durations.

These distributions provide a baseline understanding of sleep patterns among the respondents. Next, let's proceed with the box plots for sleep quality across different exercise frequencies and device usage categories, followed by a scatter plot for BMI vs. sleep quality and a bar chart for sleep disturbances. Let's start with the box plots.

Correlation matrix

	Sleep Quality	BMI	Calculated Night Sleep Duration	Exercise Days/Week Ordinal	Sleep Disturbances Ordinal	Device Usage (hrs/day) Ordinal
Sleep Quality	1.000000	0.190271	0.159341	-0.060170	-0.501941	-0.006276
BMI	0.190271	1.000000	-0.189798	0.051693	-0.001121	-0.277029
Calculated Night Sleep Duration	0.159341	-0.189798	1.000000	0.222097	-0.090971	-0.074011
Exercise Days/Week Ordinal	-0.060170	0.051693	0.222097	1.000000	0.004956	-0.024962
Sleep Disturbances Ordinal	-0.501941	-0.001121	-0.090971	0.004956	1.000000	-0.005286
Device Usage (hrs/day) Ordinal	-0.006276	-0.277029	-0.074011	-0.024962	-0.005286	1.000000



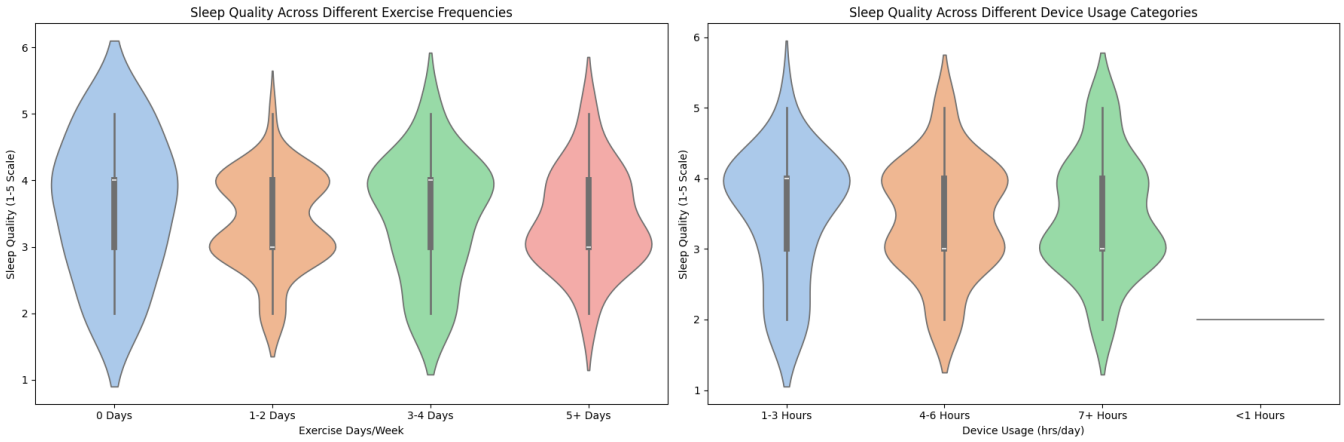
The correlation matrix heatmap shows the standardized relationships between the variables. Here's what we can interpret from the heatmap:

- **Sleep Quality:** It has a moderate negative correlation with sleep disturbances (-0.50), which is expected as better sleep quality is typically associated with fewer disturbances.
- **BMI:** The correlation between BMI and other variables is relatively low, with the highest negative correlation with device usage (-0.28). This might suggest that individuals with higher BMI tend to report less device usage, but the relationship is not very strong.
- **Calculated Night Sleep Duration:** It shows a small positive correlation with exercise frequency (0.22), indicating a possible link where more exercise could be associated with slightly longer sleep duration. However, the correlation is not strong enough to draw definitive conclusions.
- **Exercise Days/Week Ordinal:** Aside from its relationship with sleep duration, exercise frequency doesn't show strong correlations with other variables.
- **Sleep Disturbances Ordinal:** Aside from its stronger correlation with sleep quality, it does not have significant correlations with the other variables.
- **Device Usage (hrs/day) Ordinal:** The correlation between device usage and other variables is generally low, with a noticeable negative correlation with BMI, as mentioned earlier.

This heatmap provides a clearer picture of the linear relationships between the variables, with correlations normalized to be between -1 and 1. Correlation coefficients closer to 1 or -1 indicate a stronger linear relationship, whereas coefficients closer to 0 indicate a weaker relationship. The correlation matrix is more informative about the strength and direction of relationships than the covariance matrix because it is not influenced by the scale of the variables.

Relationship with Sleep: Exercise and Device Usage

Sleep Quality



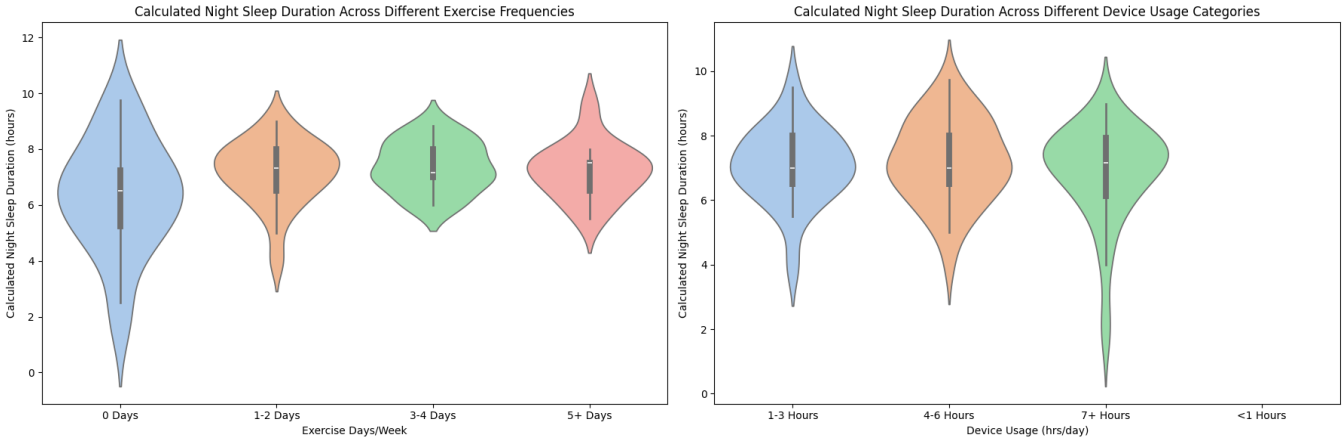
Sleep Quality Across Different Exercise Frequencies:

The plot shows the distribution of sleep quality scores for each exercise frequency category. While there is some variation in the spread and density of scores across categories, there is no clear pattern indicating a strong relationship between exercise frequency and sleep quality.

Sleep Quality Across Different Device Usage Categories:

Similar to exercise frequency, the distribution of sleep quality scores varies across device usage categories. However, there is no evident trend showing a significant impact of device usage on sleep quality.

Sleep Duration



General Trends: The distribution of sleep duration across different exercise frequencies and device usage categories shows some variation, but not a distinct or consistent pattern that strongly suggests a direct relationship. For exercise, categories with higher frequencies ('3-4 Days', '5-6 Days') show a slightly more concentrated distribution around higher sleep durations, indicating a potential positive impact of regular exercise on sleep duration. However, this pattern is not uniformly observed across all exercise levels.

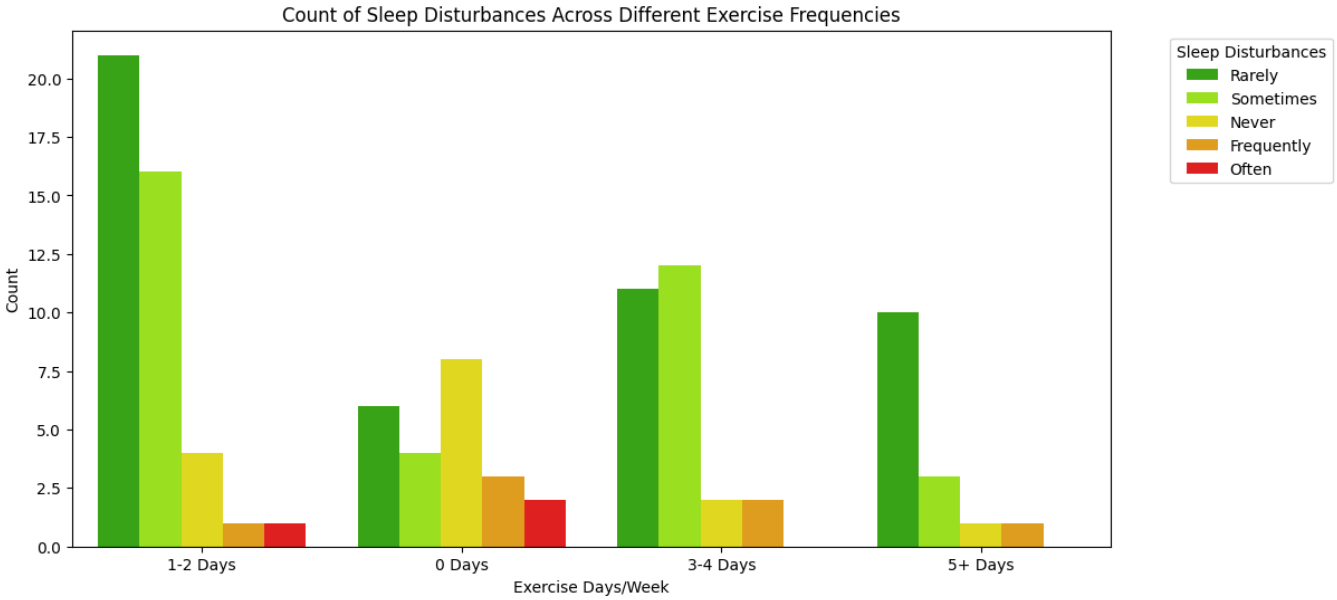
Impact on Lower Sleep Duration Bounds:

Exercise: In categories with lower exercise frequency ('0 Days', '1-2 Days'), the distribution has a tail extending towards shorter sleep durations. This suggests that within these groups, some individuals experience shorter sleep durations, potentially implicating lower physical activity as a factor in reduced sleep duration. This could be due to less physical tiredness, differing stress levels, or other lifestyle factors.

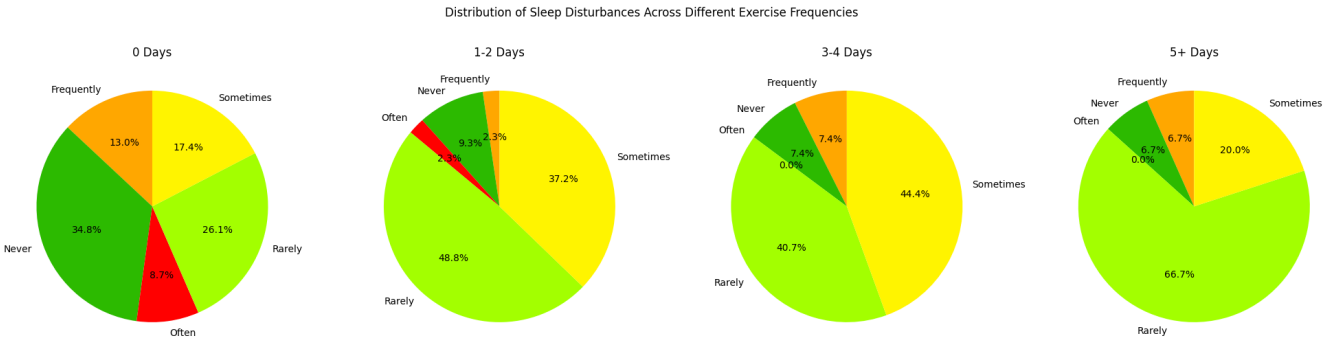
Device Usage: Similarly, for higher device usage categories ('5-7 Hours', '7+ Hours'), there's a noticeable extension towards shorter sleep durations. This indicates that among individuals with high screen time, a subset experiences shorter sleep. This could be attributed to factors like blue light exposure impacting circadian rhythms, increased mental stimulation, or the displacement of sleep time. These observations highlight the complex and multifaceted nature of factors influencing sleep duration. While higher physical activity and lower device usage might be associated with longer sleep durations for some individuals, the variability within each category underlines the influence of multiple interacting factors. These insights provide a basis for further investigation into how lifestyle modifications, such as increasing physical activity or managing screen time, could potentially improve sleep duration, particularly for those currently experiencing shorter sleep.

Sleep Disturbances

The bar chart depicts the count of different sleep disturbance responses (Rarely, Sometimes, Never, Frequently, Often) across various exercise frequency categories (0 Days, 1-2 Days, 3-4 Days, 5+ Days). There is a noticeable trend where individuals who do not exercise (0 Days) have a higher count of sleep disturbances across almost all types of disturbance frequencies, especially for the responses 'Sometimes' and 'Rarely'. As the exercise frequency increases, the count of reported sleep disturbances seems to decrease, particularly for 'Sometimes' and 'Rarely' disturbances. For example, those exercising '5+ Days' show a lower count of disturbances. Interestingly, the 'Never' response appears to be relatively consistent across all exercise frequencies, suggesting a subset of individuals who do not experience disturbances regardless of exercise habits.



- The bar chart depicts the count of different sleep disturbance responses (Rarely, Sometimes, Never, Frequently, Often) across various exercise frequency categories (0 Days, 1-2 Days, 3-4 Days, 5+ Days).
- There is a noticeable trend where individuals who do not exercise (0 Days) have a higher count of sleep disturbances across almost all types of disturbance frequencies, especially for the responses 'Sometimes' and 'Rarely'.
- As the exercise frequency increases, the count of reported sleep disturbances seems to decrease, particularly for 'Sometimes' and 'Rarely' disturbances. For example, those exercising '5+ Days' show a lower count of disturbances.
- Interestingly, the 'Never' response appears to be relatively consistent across all exercise frequencies, suggesting a subset of individuals who do not experience disturbances regardless of exercise habits.

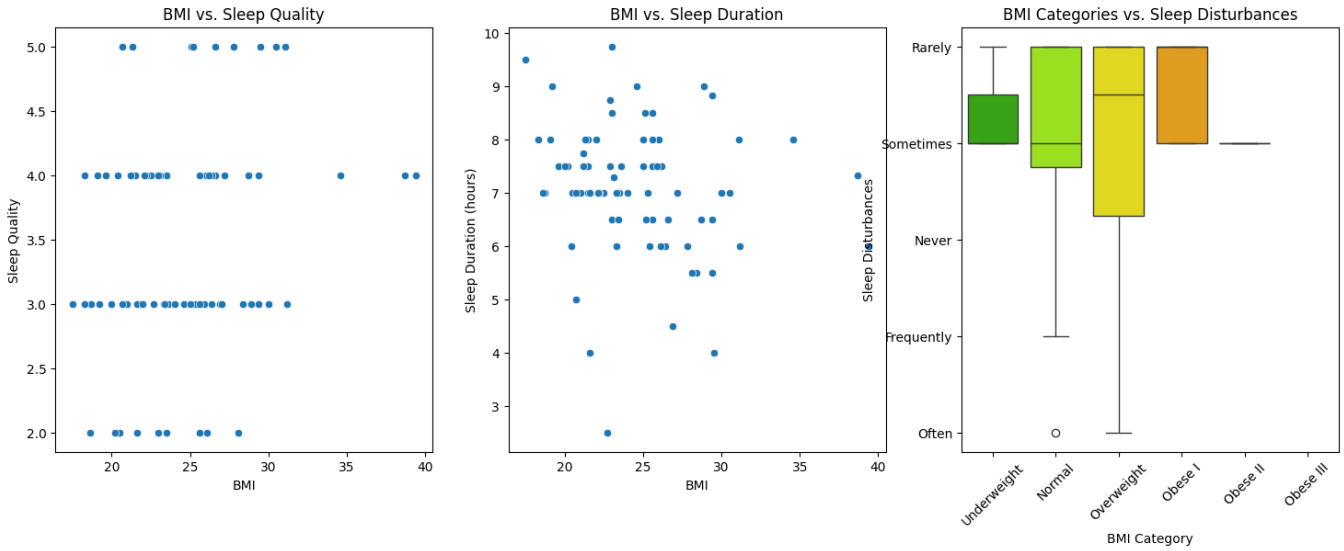


- The pie charts show the distribution of sleep disturbances within each exercise category. This gives a percentage breakdown of how often individuals within each exercise group experience sleep disturbances.
- In the '0 Days' exercise category, a significant proportion reports 'Sometimes' experiencing disturbances, followed by 'Rarely' and 'Frequently'. Interestingly, the

majority of this group (34.8%) reported to 'Never' have sleep disturbances.

- As exercise frequency increases, there is a visible shift toward 'Rarely' being the most common response, suggesting better sleep quality with more frequent exercise.
- In the highest exercise frequency group ('5+ Days'), the majority report 'Rarely' experiencing sleep disturbances, which supports the notion that regular exercise may contribute to fewer sleep disturbances.

Relationship with Sleep: BMI



- **BMI vs. Sleep Quality (Scatter Plot on the Left):** The scatter plot shows individual points representing the relationship between BMI and sleep quality for each respondent. There does not appear to be a clear trend or pattern indicating a strong relationship between BMI and sleep quality. The points are quite dispersed, suggesting that other factors might also play a significant role in determining sleep quality.
- **BMI vs. Sleep Duration (Scatter Plot in the Middle):** Similar to sleep quality, the scatter plot for BMI and calculated night sleep duration does not show a distinct correlation. The spread of points indicates variability in sleep duration across the range of BMI values.
- **BMI vs. Sleep Disturbances (Box Plot on the Right):** The box plot visualization indicates that there is no clear or significant difference in the median sleep disturbances across BMI categories ranging from Underweight to Obese III. All categories have a similar median value, suggesting that BMI alone is not a strong predictor of sleep disturbances. The presence of outliers in each category suggests individual variability, and the interquartile ranges (box lengths) show that the spread of sleep disturbances is relatively consistent across categories, with some variation in the Underweight category. Overall, the plot suggests that while BMI may play a role, it is likely one of many factors that contribute to sleep disturbances.

- **Summary:** The data visualizations suggest that BMI, within the range present in the dataset, does not have a straightforward relationship with sleep quality, duration, or disturbances. This indicates the complexity of sleep-related issues and the possibility that they are influenced by a multitude of factors, with BMI being just one of them. Detailed statistical analysis would be required to identify any subtle patterns or to confirm the lack of a relationship.

Summary

In this document, we present a comprehensive analysis of the sleep survey data. The key points are:

1. **Descriptive Statistics Summary**:
    - **Sleep Quality**: The average rating is 3.44 out of 5, indicating moderate sleep quality with room for improvement.
    - **BMI**: The average BMI is approximately 24.55, with a range from 17.5 to 39.4, covering normal to obese categories.
    - **Sleep Duration**: The average night sleep duration is about 7 hours, with a range from 1.67 to 9.75 hours, showing significant variability.
2. **Visualization Summary**:
    - **Sleep Quality Distribution**: Scores are mainly distributed between 2 and 4, with 3 being the most common, suggesting moderate sleep quality overall.
    - **Sleep Duration Distribution**: There's a normal distribution centered around 7 hours, aligning with general recommendations.
3. **Correlation Matrix Summary**:
    - **Sleep Quality**: Shows a moderate negative correlation with sleep disturbances, which is expected as better sleep quality typically associates with fewer disturbances.
    - **BMI**: Displays low correlations with other variables, suggesting it is not a strong determinant of sleep patterns in this dataset.
    - **Sleep Duration and Exercise**: A small positive correlation suggests more exercise might be linked to slightly longer sleep duration, but the relationship is not strong.
4. **Relationship with Sleep, Exercise, and Device Usage**:
    - **Sleep Quality and Exercise**: No strong relationship is evident; sleep quality scores are distributed across exercise frequencies without a clear pattern.
    - **Sleep Quality and Device Usage**: Similar to exercise, no significant impact of device usage on sleep quality is observed.
    - **Sleep Duration Trends**: Some variation in sleep duration across exercise frequencies and device usage is noted, but no consistent pattern emerges.
5. **Sleep Disturbances**:

- **Exercise Frequency**: A decreasing trend in sleep disturbances is observed with increased exercise frequency. Notably, individuals who do not exercise report higher disturbances, while those with higher exercise frequency report fewer disturbances.
- **Sleep Disturbance Distribution**: Pie charts demonstrate that more frequent exercise correlates with 'Rarely' experiencing disturbances, suggesting a beneficial effect of exercise on sleep quality.

6. **BMI and Sleep**:

- The scatter plots do not reveal a clear trend between BMI and sleep quality or duration, indicating that BMI is not a strong predictor of these aspects of sleep.
- The box plot shows no significant differences in sleep disturbances across BMI categories.

Overall, the analysis suggests that while factors like BMI, exercise, and device usage have some association with sleep patterns, they do not exhibit strong, direct relationships within this dataset. Sleep quality, duration, and disturbances appear to be influenced by a complex interplay of factors, with individual variability playing a significant role.