Digital Dreams - Unveiling the Impact of Screen Time on Sleep Quality

Carly Chick (cfc91, section 6), Meryl Gideon (mcg212, section 7)
Rutgers University - New Brunswick

1 Introduction

In a time that is ruled by the use of technology, screen time (either from your phone, tv, etc) has become an important part of our daily lives. Although these technologies offer many options for entertainment, there is a concern about their effects on human health, particularly sleep. Sleep plays an important role in physical restoration, mental clarity, and emotional regulation. However, disruptions to healthy sleep patterns have been linked to many health issues, including cardiovascular disease, cognitive decline, and increased stress levels.

Emerging evidence suggests that screen exposure, especially during evening hours, may negatively influence sleep quality by suppressing melatonin production and changing circadian rhythms. The blue light emitted by screens acts like daylight, signaling to the brain that it is not time to sleep, potentially leading to delayed sleep onset and reduced overall sleep duration. This is concerning given the widespread use of devices right before bedtime.

This study's goal is to examine the relationship between screen time and sleep quality by analyzing a publicly available data set. By exploring patterns across different demographics, this research aims to encourage more informed screen time habits by making graphs that reveal relationships between screen time and sleep quality and by creating and analyzing a linear regression model on a chosen target variable and predictors.

2 Data Preprocessing and Cleaning

The dataset used in this study was in CSV format and was obtained from Kaggle, titled "Social Media Usage & Sleep Data (Singapore)". This data set has 500 rows and 20 columns. This data set includes information on individuals' social media usage habits alongside self-reported sleep patterns and demographic information such as age and gender. We want to stress that the data in this data set was collected in Singapore, so the specific statistics we get from the linear regression model that we implemented and the relationships between certain variables that we find when plotting graphs might not apply to other similar data sets from other countries.

To prepare the data for analysis, several pre-processing steps were taken. First, unnecessary columns that were not important to the research question were removed to simplify the data set and focus on variables that directly related to screen usage and sleep quality. A check for null values confirmed that the data set contained no missing data, eliminating the need for row deletion.

To help compare across different age ranges, a new column labeled "Age Groups" was created. This column categorizes individuals into age brackets, allowing for clearer visualization of trends related to screen time and sleep quality across demographic groups. These steps ensured a clean and structured data set, usable for further exploratory and statistical analysis.

3 Exploratory Data Analysis

To explore the relationships between the variables in the data set, we plotted many types of graphs, namely scatter plots, bar plots, a heatmap, and a line plot. All of these graphs were plotted using the seaborn package. To begin we will discuss the scatter plots that were plotted.

The first scatter plot that we plotted explores the relationship between pre-sleep screen time use (measured in minutes) and sleep quality rating for different age groups. Sleep quality is measured on a scale from 1 to 5, 5 being the best and 1 being the worst. We plotted a separate scatter plot for each age group, 18-24 year olds, 25-34 year olds, 35-60 year olds, and 61+ year olds, so a total of four subplots were made. Looking at all the subplots, it seems that if the pre-sleep screen time use is limited, then sleep quality increases (meaning more people rated their sleep quality between a 3 and 5), no matter what age you are. This suggests that pre-sleep screen time use and the sleep quality rating are negatively correlated with each other. A picture of the graphs that were generated can be seen below:

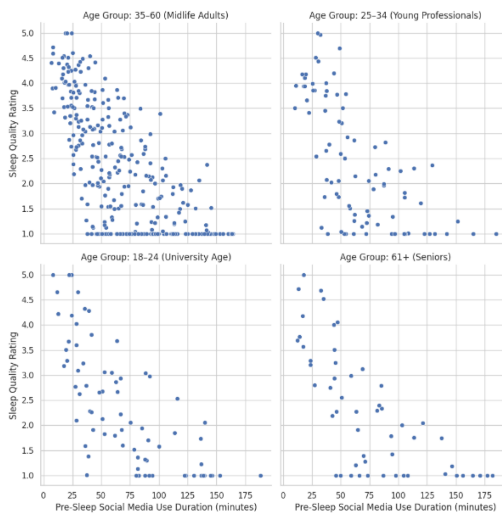

Figure 1: Sleep Quality Rating vs. Pre-Sleep Social Media Use Duration (minutes) across four age groups

The second scatter plot that we plotted explores the relationship between pre-sleep screen time use and sleep quality rating for different genders. There are three labels for gender in our data set: male, female, and other. We plotted a separate scatter plot for each gender. Looking at all the subplots, it seems that if screen time use before bedtime is limited, then all genders are getting better sleep (meaning there are more sleep quality ratings between 3 and 5). This suggests that pre-sleep screen time use and sleep quality rating are negatively correlated with each other. A picture of the graphs that were generated can be seen below:

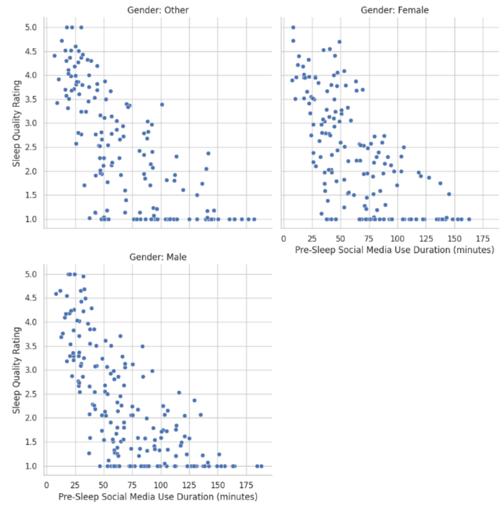

Figure 2: Sleep Quality Rating vs. Pre-Sleep Social Media Use Duration (minutes) across different genders

The third scatter plot that we generated simply explores the relationship between pre-sleep screen time use and sleep quality rating for the entire data set. This scatter plot nicely connects the two scatter plots we discussed above into one big scatter plot, making it easier to see the nature of the relationship between these two variables across the entire data set and not just within specific subgroups of the data set. As seen in the two scatter plots above, this scatter plot also shows that limiting phone use before going to sleep will cause sleep quality ratings to increase (once again, this means that most sleep quality ratings are between 3 and 5). Once again, pre-sleep screen time use and sleep quality rating are negatively correlated with each other. A picture of the graph that was generated can be seen below:

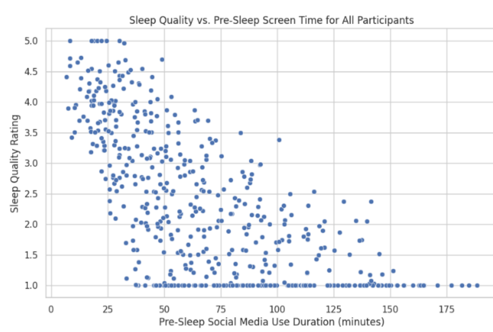

Figure 3: Sleep Quality Rating vs. Pre-Sleep Social Media Use Duration (minutes) across entire data set

Now, we will discuss the bar plots that were plotted. The first bar plot that we plotted explores the relationship between the type of social media content consumed and average total sleep time (measured in hours). We found that there were only three unique labels in the 'type of social media content consumed' column of our data set, namely social interaction, news, and entertainment. The average total sleep time for these three types of social media content were relatively the same, a little more than 5 hours, which is well below the 7 to 9 hours of sleep that many health professionals recommend. So, although we can not conclude that consuming a specific genre of social media content will aid in getting more sleep per night, we can conclude that consuming any type of social media content before bed will not aid the average person in getting more sleep per night. A picture of the plot that was generated can be seen below:

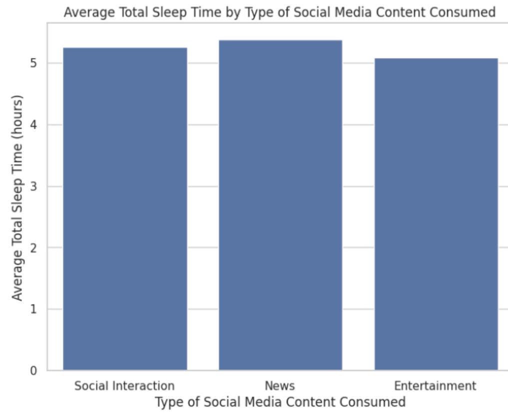

Figure 4: Average Total Sleep Time (hours) vs. Type of Social Media Content Consumed

The next bar plot that we plotted explores the relationship between the type of social media content that was consumed before bed and average sleep latency (measured in minutes). As before, the type of social media content was broken up into three categories: social interaction, news, and entertainment. Sleep latency is the time it takes for a person to transition from full wakefulness to sleep. From the graph, we notice that there is not a huge difference between the average sleep latencies for different types of social media content. They are all around the same value, around 23 minutes. For the average person, their sleep latency is between 10 to 20 minutes according to most sources. So, there is around a three minute difference minimum to a 13 minute difference maximum between the average sleep latency found in the data set and the average sleep latency of a healthy person. So, although we can not conclude that consuming a specific type of social media content will cause a person's sleep latency to increase or decrease dramatically, we can conclude that consuming any type of social media content before bedtime can have a significant impact on the average person's sleep latency, increasing it from 3 minutes minimum to 13 minutes maximum. A picture of the plot that was generated can be seen below:

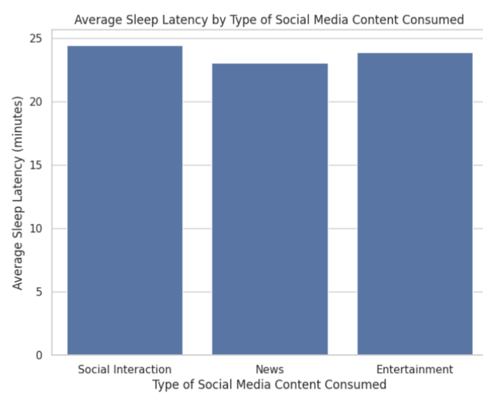

Figure 5: Average Sleep Latency (minutes) vs. Type of Social Media Content Consumed

Finally, we will discuss the heat map that models how average melatonin levels (measured in pg/mL) is affected by screen time (using the pre-sleep social media use duration (minutes) column from our data set) and blue light exposure before sleep (measured in minutes). Melatonin is the hormone that helps regulate your sleep-wake cycle. Blue light is a unique type of light that is most commonly emitted from phone screens. We decided to create four separate subplots based on the same age groups that were used in a previous section: 18-24 year olds, 25-34 year olds, 35-60 year olds, and 61+ year olds. This was done to find out whether or not the effect on average melatonin levels by screen time and blue light exposure would affect people of different ages differently. To interpret the heatmap, note that cells that are more red in color mean that the melatonin levels are at a healthy number, while cells that are more blue in color suggest that melatonin levels may be at an unhealthy number. Looking at all 4 heatmaps, it seems that low screen time and low blue light exposure results in higher, healthy melatonin levels on average across all age groups while high screen time and high blue light exposure results in lower, unhealthy melatonin levels on average across all age groups. This suggests that melatonin levels are negatively correlated with screen time and blue light exposure, and that this negative correlation is independent of age. A picture of the plot that was generated can be seen below:

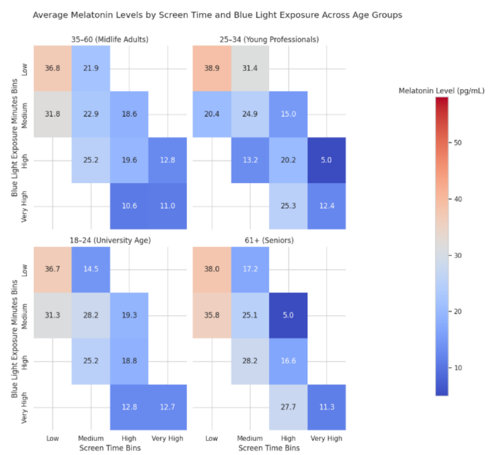

Figure 6: Average Melatonin Level (pg/mL) as measured by Pre-Sleep Social Media Use Duration (minutes) and Blue Light Exposure Before Sleep (minutes)

4 Data Modeling

To explore how screen time affects sleep quality and efficiency, we used a supervised learning approach with Linear Regression to model Sleep Efficiency based on key predictors. We chose Sleep Efficiency as our target variable because it is a more scientifically reliable measure than the self-reported Sleep Quality Rating, which can be subjective and statistically inconsistent. The purpose of this model was not just to make predictions but also to understand which features had the strongest and most meaningful influence on sleep efficiency through the interpretation of slope values (coefficients).

From the cleaned data set, we then selected the following four features as independent variables based on their relevance to the research question: Pre-Sleep Social Media Use Duration (minutes), Blue Light Exposure Before Sleep (minutes), Sleep Latency (minutes), and Melatonin Level (pg/mL).

These features were chosen because they reflect behaviors directly related to screen exposure and sleep. The dependent variable was Sleep Efficiency. We then split the data into training and testing sets using an 80/20 ratio for fair validation.

After training the Linear Regression model using scikit-learn, we examined the coefficients (slopes) to interpret the influence of each independent variable on Sleep Efficiency:

Pre-Sleep Social Media Use Duration: Slope = -0.0385 $\rightarrow$ This negative coefficient suggests that increased social media use before bed is associated with a slight decline in sleep efficiency, which supports the hypothesis that screen time can disrupt sleep.

Blue Light Exposure Before Sleep: Slope = +0.0140 $\rightarrow$ This coefficient was unexpectedly positive and very small, indicating a weak relationship with sleep efficiency. One possible explanation is that many users now enable blue light filters or "night mode" settings on their devices, which reduce the intensity of blue light and may limit its measurable effect on sleep. As a result, self-reported exposure might not fully capture the true impact.

Sleep Latency (Time to Fall Asleep): Slope = -0.1888 → This negative slope indicates that longer time to fall asleep is strongly associated with lower sleep efficiency, a relationship consistent with known sleep research.

Melatonin Level: Slope = +0.3053 → This was the most influential predictor in the model. Higher melatonin levels corresponded with better sleep efficiency, reinforcing melatonin's important role in regulating sleep cycles.

These slope values helped identify which factors should be most emphasized when considering how to improve sleep outcomes.

Originally, our model's R^2 score was around 0.78. To improve the fit, we used a Polynomial Regression pipeline (degree = 2) to capture possible nonlinear relationships. We also performed cross-validation (3-fold) to validate the model's consistency. The model's average R^2 score then improved to 0.823, suggesting a strong fit and good explanatory power.

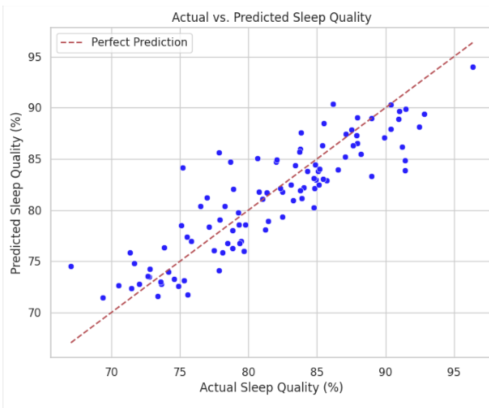

Figure 7: Final Linear Regression Model

5 Concluding Our Findings

This project's goal was to test the hypothesis that increased screen time before sleep negatively affects sleep quality and efficiency. Our analysis supported this through both visual and statistical methods.

Using scatter plots, we observed a consistent negative correlation between pre-sleep screen time and self-reported sleep quality across different age groups and genders. In almost all the cases, people who used screens less before bed reported higher sleep quality ratings. Bar plots revealed that the type of social media content consumed before bed had little effect on total sleep time, but all types were associated with average sleep durations below recommended levels. Also, we found that watching any kind of social media content before bed was associated with slightly elevated sleep latency. The heatmaps further illustrated that higher screen time and blue light exposure were linked to lower melatonin levels across all age groups, which is an important physiological sign of disrupted sleep.

To quantify these relationships, we built a linear regression model with sleep efficiency as the target. While direct measures like screen time and blue light exposure had weaker coefficients than expected, sleep latency and melatonin levels emerged as strong predictors. These results suggest that the impact of screen time may be more indirect through things like delayed sleep onset and hormonal disruption.

Overall, our visualizations and model findings reinforce the idea that heavy screen use before bedtime can reduce sleep quality and efficiency. The consistent trends seen in our graphs and the explanatory strength of our regression model (R^2 = 0.823) offer strong evidence in favor of our original hypothesis. Our study shows the value of combining exploratory data analysis with statistical modeling to better understand how modern behaviors affect health and underscores the importance of mindful screen use for better sleep.

6 References

Tools and Libraries Used: Python, Pandas, NumPy, Scikit-learn, matplotlib, seaborn, and Jupyter Notebook.

Data set from Kaggle: https://www.kaggle.com/datasets/globalmediadata/socialmediausage-sleepdata-sg