**Analysis of the Relationship Between Screen Tiem and Well-being Among Adolescents**

**Group Reports**

**Course:** HIT140 Foundations of Data Science (2024)

**Submission Date:** 18/10/2024

**Group:** 41

**Group Members:**

1. Zenith Sharma (s371125)
2. Shem Ramudan (s372783)
3. Sagun Pandey (s374592)
4. Bipin Poudel (s374247)

**Table of Contents**

1. Introduction
2. Data Overview and Exploration
3. Feature Engineering
4. Data Visualization
5. Model Development and Linear Regression Analysis
6. Discussion and Limitations
7. Conclusion
8. Individual Contributions

**List of Figures**

1. Correlation Heatmap of Screen Time and Well-being Indicators

2. Computer Screen Time During the Week & Weekend

**1. Introduction**

In today's tech-driven world, there’s a growing concern about how too much screen time might affect the mental health of young people, especially teenagers. Whether gaming, scrolling on smartphones, binge-watching shows, or spending hours working on computers, screens have become a big part of their everyday routines. However, studies suggest that excessive screen exposure could be linked to higher stress, anxiety, and lower self-esteem.

This project explores the relationship between screen time and well-being through data science. Our goal is to understand how different types of screen time, whether during weekdays or weekends, impact adolescents' self-reported well-being. We plan to predict well-being scores by analyzing screen time data alongside demographic factors such as gender and socioeconomic background.

We will be working with three key datasets. One contains demographic information, the second captures the time spent on various devices, and the third provides well-being indicators. Using data visualization and linear regression, we aim to explore how screen time affects the well-being of adolescents. Specifically, we want to understand the impact that different amounts and types of screen use have on their overall mental health.

**2. Data Overview and Exploration**

**2.1 Description of Datasets**

The analysis leverages three datasets that each capture a different aspect of the study,

**1. Demographics (dataset1. csv):** Contains base demographic information for 120,000 adolescents and associated descriptions. Key variables include:

* **ID:** This is an identifier of the individual participant.
* **Gender:** Binary variable (1 = male, 0 = female)
* **Minority Status:** If it is minority (1) or not a minority (0).
* **Deprived:** Takes the value of 1 when a participant lived on an area with high deprivation indices, otherwise it takes the value of 0.

**2. Screen Time (dataset2. zip | csv):** Daily screen time for about 113,000 respondents measuring computer use, video gaming device-usage and smart phone and TV usage on both weekend days and weekdays. The variables include:

* C\_we, C\_wk: Computer usage weekends and weekdays.
* Table 3 G\_we, G\_wk: Video game time use weekend and weekday.
* S\_we, S\_wk: Smartphone (Weekend/Weekday)
* T\_we, T\_wk: Television watching on weekends and weekdays.

**3. Indicators of well-being csv):** 102,580 rows of self-reported well-being scores covering optimism, relaxation, energy, and confidence. The scale is 1 to 5 (where each indicator can be given a rating of one, two, three, four or five).

* 1 = None of the time
* 5 = All of the time

**2.2 Data Cleaning Process**

To prepare the datasets in such a way that they can be fed to analysis methods, we are required to do following cleaning steps:

* **Missing Data**: Missing values in the screen time and well-being outcomes were listwise deleted to maintain consistency across variables.
* **Merge the Datasets** — We merge both the datasets based on ID (as it is useful) We matched this information to each respondent, providing data on all their demographics and screen time for us to also have well-being scores.
* **Outlier Detection:** To detect outlier, we checked some screen time variables and kept those outliers if the value is seemingly acceptable for teenagers.
* **Standardization:** No transformation was necessary anymore since all the numerical features were already standardized.

**3. Feature Engineering**

**3.1 New Feature Creation**

We changed it in order to more accurately model the effects of screen time and add some better normalization that we do for per-app data usage information.

* **Gross Screen Time**: Sum of all screen time variables on weekdays and weekends for the 4 platforms (computers, smartphones, gaming consoles & TV) This provided a complete estimate of overall screen time for that participant.
* **Device-Specific Usage:** We computed overall screen time per device type each device category (e.g., total gaming by adding up weekday and weekend hours of that individual-specific devices Usage)
* **Screen Time Ratio:** We calculated a ratio of weekend screen time compared to weekday screen times as an additional measure indicating how differently participants may have behaved during weekdays versus on weekends.

**3.2 Objective of feature creation**

It will depend on the type of device, and day/ time when used. We generated a complete picture of the habits of everyone by aggregating screen times across devices. The hypothesis is that high amounts of screen use, especially on weekend days, are associated with stronger negative effects in well-being than moderate to balanced weekday-screen-patterns.

**4. Data Visualization**

**4.1 Correlation Analysis**

We created a correlation **heatmap** to understand the relationship between screen time and well-being. These heats maps allowed us to observe patterns of associations between screen time and the positive well-being indicators, such as Optimism (Optm), feeling relaxed (Relx), and feeling close to others (Clsep).

Several salient results were highlighted in the following heatmap:

* **Gaming on Weekends (G\_we)** had a modest positive relationship with decreased Engs and Optm, meaning that greater time spent gaming during weekends was associated with lower engagement energy and optimism of the adolescents.
* **Smartphone Usage** had a strong negative correlation with almost all well-being indicators **(S\_we, S\_wk)** The finding indicates that too much smartphone use--especially the day before school --is linked with lower well-being in adolescents.
* **TV Watching** was negatively correlated with most well-being indicators at weaker levels as compared to Complete PA (e.g. belief that tomorrow will be a better day, feeling of inspiration) suggesting that too much TV time could also be associated with lower feeling of confidence and relaxation.

**4.2 Distribution of Screen Time**

For even more detail on screen time patterns, we also made a histogram of the distribution between computer screen times for weekends and weekdays. In the plot below, we can see that number of hours on computers were higher among adolescents in weekends than weekdays because they spent reduced time at school and probably at some other activities (more free time).

**5. Model Development and Linear Regression Analysis**

**5.1 Model Selection**

We decided to perform a linear regression because it is relatively easy and allows relationships between dependent variables (screen time) and independent variable (well-being). The response variables were well-being factors (e.g., optimism, relaxation, confidence), and the predictors were:

* + Total Screen Time on all devices
  + Weekends and weekdays screen-time
  + Demographic features like gender, minority status and deprivation

**5.2 Model Training**

We split the dataset 80-20%, using 80% to training and test our models on another or remaining %20 data. The regression model was made with scikit-learn module in Python.

**5.3 Results**

The Model provided the following insights:

* Smartphone usage on weekends (S\_we) significantly negatively predicted all most of the well-being indicators.
* During weekends, the more time that young adolescents used smartphones overall — compared to other activities, the less they self-reported optimism about their futures and feeling relaxed or confident daily.
* Gaming time had a moderate negative impact on energy levels, suggesting that long hours of gaming may lead to fatigue or lower energy.
* Demographics: Depending on the age group, females reported slightly lower well-being compared to males but rarely did this measure have a significant effect.

The fit of the model, as measured by its overall **R-squared** value (ranging from 0.15 to 0.22), explained approximately between up to offering about one-fifth variance in well-being scores accounted for screen time variables 15–22% This is not a high explanatory power, but it is also standard for social science models, as mental health has many contributing factors.

**6. Discussion and Limitations**

**6.1 Key Findings**

The findings showed that greater screen time - especially on a smartphone, playing video games or using social media and the Internet - was associated with a lower level of well-being in young people. Smartphone usage at the weekend was particularly harmful, with almost half of respondents reporting that it made them less positive about their life (46%) or reduced levels of relaxation and feeling confident. Gameplay Time Affected Energy Levels- Too Much Gameplay Guarantees Fatigue And Low vitality Television viewing appeared less important and was weakly negatively correlated with psychological well-being scales such as confidence or relaxation.

Well-being was affected to a lesser extent by demographic variables, such as gender and deprivation status. While some females reported slightly lower well-being scores, this difference was smaller than the effect of screen time.

**6.2 Limitations**

This analysis has opened a lot of gaps despite the insights provided. For one, the well-being data is reported by individuals themselves; this can lead to bias or errors in responses. Data on the emotions of the participants might not have been accurately reported or were misunderstood by either side, making it less reliable.

Second, this is a correlational study, and it therefore does not establish causality between screen time and well-being. So, there is an association, but it is unclear whether reduced well-being directly results from screen time or if other factors are at play. Moreover, the explanatory power of our model was low (R-squared ranged from 0.15 to 0.22), implying that a lot remains unexplained and not included as covariates in crew factor affect most likely how well-being amongst adolescents is shaped by their caregivers' work arrangements.

**7. Conclusion**

This study looked at how screen time affects the well-being of teenagers using data science methods. We found that too much screen time, especially on smartphones and over the weekends, was linked to lower well-being. Excessive time spent on smartphones and gaming was tied to feeling less optimistic, less energetic, and less relaxed. In contrast, screen time for educational purposes seemed to have less of a negative impact.

However, our model explained only 15-22% of the variation in well-being, suggesting that other factors, such as social interactions, exercise, and individual traits, also play important roles. Demographic factors like gender and deprivation had minimal effects, although girls reported slightly lower well-being on average.

Overall, the findings support the idea that managing screen time is important for mental health. Still, future research should include more data and explore other factors to get a clearer picture of these relationships.

**8. Individual Contributions**

**1.** **Sagun Pandey (Project Lead and Data Analyst):** I took charge of organizing the project, keeping the team on track with deadlines and managing meetings. I led the data analysis, combining datasets, finding correlations, and creating visualizations. I also ensured the results were accurate and helped interpret the data.

**2. Shem Ramudan (Report Writing and Literature Review):** Shem focused on writing the report, making sure it met academic standards. Mr. Ramudan did the literature review, added references, and polished the language to ensure everything flowed smoothly.

**3.** **Zenith Sharma (Coding and Visualization):** Zenith was responsible for writing the Python code to analyze the data and create visualizations like histograms and heatmaps. Mr. Sharma also worked with the team to troubleshoot and optimize the code.

**4. Bipin Paudel (Research and Data Interpretation):** Bipin helped make sense of the results by researching similar studies and suggesting the implications of the findings. Mr. Paudel also contributed to discussing the limitations and future directions of the research.