Project Report on

Teenagers Phone Addiction Analysis

In partial fulfillment for the award
Of

Professional Certification in Data Analysis and Visualization

Year 2024-2025

Submitted By

Ameer ali.k

Tools & Technologies Used:

Python, R Programming, Tableau

Duration:

15/07/2025 – 30/07/2025

Submission Date:

31/07/2025

G - TEC CENTRE OF EXCELLENCE PERINTHALMANNA

Abstract

This study investigates smartphone usage and its impact on adolescents, focusing on addiction levels, daily usage patterns, and their relationship with academic and lifestyle metrics. Data from student responses were analyzed to identify the top 5 students with the highest phone addiction levels, with addiction categorized as Low, Moderate, or High. Gender and grade-level comparisons revealed which groups exhibited the highest average addiction levels. The analysis further examined usage purpose–wise distribution of daily phone checks, showing the percentage contribution of each usage type.

Key findings highlight that a notable percentage of students check their phones over 50 times per day, and several students use their phones for more than 5 hours daily. A pivot table analysis of daily usage hours, sleep duration, and academic performance across age groups and gender provided insights into the behavioral patterns linked with excessive screen time. Additionally, the study measured the prevalence of high screen time before bed and summarized average metrics, including daily usage, sleep hours, and addiction levels.

Introduction

Smartphones have become an inseparable part of modern student life, offering constant connectivity, entertainment, and access to digital learning resources. However, the ease of access to these devices has also led to rising concerns about excessive usage, often resulting in phone addiction. Such behaviors are linked to reduced sleep duration, poor academic performance, and negative impacts on mental health and social interactions among adolescents.

This study aims to explore smartphone addiction patterns among students by evaluating their daily screen time, phone checking frequency, and time spent across various usage purposes such as social media, gaming, and education. Addiction levels are classified into low, moderate, and high to better understand the distribution of dependency. Furthermore, the analysis examines how factors like age, gender, and school grade influence phone usage behaviors and their relationship with sleep and academic outcomes.

By identifying trends and highlighting high-risk behaviors, this research provides valuable insights that can assist parents, educators, and policymakers in fostering healthier smartphone habits among adolescents.

LITERATURE REVIEW

Smartphone Addiction and Usage Patterns

- Kwon et al. (2013) developed the Smartphone Addiction Scale (SAS), identifying key behaviors like overuse, withdrawal, and neglect of daily activities.
- Excessive phone use is characterized by frequent checking, prolonged screen time, and dependency on mobile apps.

Impact on Academic Performance

- Lepp et al. (2014) found a negative correlation between smartphone usage and academic performance among students.
- High usage reduces focus, study time, and overall academic efficiency.

Effect on Sleep and Daily Routines

- Lemola et al. (2015) showed that nighttime smartphone use delays sleep onset and reduces sleep duration in adolescents.
- Poor sleep quality caused by late-night screen exposure affects cognitive function and school performance.

Psychological Impacts

- Elhai et al. (2017) linked frequent phone checks, especially for social media, with anxiety and emotional distress.
- Samaha and Hawi (2016) highlighted that phone addiction increases stress and lowers productivity.

Demographic Differences in Usage

- Chen et al. (2017) found gender differences in usage patterns:
    - Female students engage more in social activities on smartphones.
    - Male students spend more time on gaming and entertainment.

Need for Behavioral Interventions

- Literature emphasizes the importance of awareness and regulation of smartphone usage.
- Tailored interventions are required to reduce addiction and its negative academic, psychological, and health effects.

RESEARCH GAP

Most research examines academic performance, psychological effects, or sleep patterns in isolation rather than integrating these factors for a holistic understanding. Comparisons across gender, age groups, and school grades are also underexplored. Additionally, the impact of specific usage purposes—such as social media, gaming, and education—along with night time screen exposure has not been extensively studied. Finally, few studies provide a clear, data-driven categorization of smartphone addiction levels into low, moderate, and high, which is essential for targeted interventions.

DATA COLLECTION & PRE PROCESSING

- **Data Collection**

The dataset used in this study was obtained from Kaggle, which provides publicly available data on student smartphone usage and behavioral metrics.It includes information about adolescents aged 13–19 years, covering the following key attributes:

  - **Demographics**: Age, Gender, and School Grade
  - **Smartphone Usage**: Daily Usage Hours, Phone Checks Per Day, Time on Social Media, Gaming, and Educational Apps
  - **Lifestyle Factors**: Sleep Hours, Screen Time Before Bed, Exercise Hours, and Social Interactions
  - **Performance & Psychological Indicators**: Academic Performance, Addiction Level, Depression Level, Self-Esteem

The dataset was downloaded in CSV format for further cleaning, transformation, and analysis using Python (Pandas, NumPy, and visualization libraries).

- **Data Pre-processing**

  - **Loading the Dataset**: The CSV file was imported into Python using **pandas.read_csv().**
  - **Handling Missing Data**: Checked for missing or null values and handled them by **removing or imputing** where necessary.
  - **Data Type Formatting**: Converted numerical columns like Daily Usage Hours, Sleep Hours, and Phone Checks Per Day to proper numeric types for analysis.
  - **Feature Engineering**:
    - Created an Addiction Category (Low, Moderate, High) based on the **Addiction_Level.**
    - Grouped ages into Age Groups (13–14, 15–16, 17, 18–19) for comparative analysis.
  - **Data Consistency**: Ensured uniform units (hours, counts) and standardized categorical fields like Gender and Phone Usage Purpose.

 Prepared Dataset

  - After cleaning and processing, the dataset was ready for statistical analysis, visualization, and categorization of students based on smartphone addiction levels and behavioral patterns.

METHODOLOGY

This project follows a multi-step analytical approach combining Python and Tableau to analyze the impact of smartphone usage on students' academic performance, sleep patterns, and behavioral habits. The methodology is designed to provide a comprehensive understanding of the dataset, categorize students based on addiction levels, and present insights through clear and interactive visualizations.

Tools and Technologies Used

- Python was used for data cleaning, preprocessing, and exploratory data analysis (EDA). Key libraries such as Pandas, NumPy, Seaborn, and Plotly Express enabled efficient data handling, pattern identification, and visualization.
- Tableau was used for building interactive dashboards, allowing insights on smartphone addiction trends to be easily understood by both technical and non-technical audiences.

Exploratory Data Analysis – Python

After preprocessing, exploratory analysis was conducted to identify patterns and relationships between variables such as daily usage hours, sleep duration, phone checks per day, and academic performance. Visualization techniques helped in understanding overall trends and supported the categorization of students into Low, Moderate, and High addiction levels.

Visualizations Used:

- Bar Plots – Top 5 Locations with Highest Average Addiction Level
- Line chart– Average Addiction Level by School Grade
- Pie Charts- Average Screen Time: Weekday vs Weekend.

Grouping and Categorization

- Groupby() was used to calculate average addiction levels and daily phone usage by gender and school grade.
- Pivot Tables helped analyze the relationship between daily usage, sleep hours, and academic performance across age groups and genders.
- Categorization of Addiction Levels was performed to divide students into Low, Moderate, and High addiction based on their Addiction level

Visualization in Tableau

Finally, the cleaned and aggregated data was imported into Tableau to create interactive **dashboards**. These dashboards visualize:

- **Age group by Depression Level**
- **Grade by Addiction Level**
- **Top 10 Addicted students**

Statistical Testing (R Programming)

To validate the insights derived from Python analysis, several statistical hypothesis tests were conducted in R Programming. These tests were applied to evaluate differences between groups and confirm the significance of observed patterns.

T-test: Compare the mean Sleep Hours on weekdays vs. weekends to check if there is a significant difference.

Chi-square Test: Test the association between Educational Screen Time and Academic Performance Levels**.**

ANOVA: Assess whether Depression Levels differ across various Phone Usage Purposes**.**

F-test: Compare the variance in Academic Performance between High and Low Social Media Usage groups.

Z-test: Examine whether the average Anxiety Level differs between Gamers and Non-Gamers**.**

RESULTS AND ANALYSIS

This section presents the key findings of the project, Made insights into the smartphone usage, addiction levels and academic performance

Python- Based Results

- The top 5 students with the highest phone addiction levels all have an Addiction Level of 10.0**.**
- The percentage of daily phone checks by usage purpose is nearly balanced across categories, with slightly higher activity in other tasks and social media**.**
- The 11th grade students have the highest average phone addiction level of 8.99**,** followed closely by 7th grade (8.95) and 9th grade (8.86**)**. This indicates that upper and middle grades show higher smartphone dependency compared to lower grades.
- Among all genders, Female students have the highest percentage of phone addiction at 33.59%**,** followed closely by Male students (33.28%)**,** while the Other category is slightly lower at 33**.**13%**.**
- Indicate that students, on average, use phones for 5 hours daily, sleep less than 7 hours, and have a moderately high addiction level.
- A total of 1,455 students use their phones for more than 5 hours daily, indicating a high prevalence of heavy smartphone usage among students.

- Students, on average, use their phones for 5.02 hours daily, sleep 6.49 hours, and have an average addiction level of 8.88. Their average academic performance is 74.95, and screen time before bed is 1.01 hours.
- Approximately 75.1% of students check their phones more than 50 times per day, indicating frequent and compulsive phone usage among the majority of students.
- A total of 283 students were identified as high-risk, showing daily phone usage above 6 hours, sleep below 7 hours, and academic performance under 75.
- A total of 412 students were found to have high screen time before bed (more than 1.5 hours), which may contribute to reduced sleep duration and higher addiction levels.

R-Based Statistical Results

1. **T-Test:** Students sleep significantly more on weekends than on weekdays ($p < 0.05$).
2. **F-Test:** No significant difference in the variance of academic performance between high and low social media users ($p \geq 0.05$).
3. **Chi-Square Test**: Significant association exists between educational screen time and academic performance ($p < 0.05$).
4. **Z-Test:** No significant difference in average anxiety levels between gamers and non-gamers ($p \geq 0.05$).
5. **ANOVA**: Depression levels differ significantly across different phone usage purposes ($p < 0.05$).

Tableau- Based Results

- Addiction levels vary across grade levels, with 7th and 10th grades being most affected.
- Males and females have similar addiction levels.
- Top 10 addicted students have high depression levels and varying sleep hours.
- Younger students (13-15 years) are more vulnerable to depression.
- Students spend an average of 5 hours daily on devices
- sleep quality as they average 6.48 hours of sleep per night.
- The average depression level among students is 5.46
- The average addiction level is high at 8.88

CONCLUSION

The analysis of smartphone usage among students demonstrates a clear pattern of excessive phone dependency, with students spending on average over 5 hours per day on their devices and 75% of them checking their phones more than 50 times daily, leading to moderate to high levels of addiction that are strongly associated with reduced sleep duration and lower academic performance, while 283 students have been identified as high-risk users due to a combination of prolonged usage, insufficient sleep, and suboptimal grades; statistical evaluations further reinforce these observations by confirming that students tend to sleep significantly more on weekends than on weekdays, suggesting weekday smartphone use disrupts normal rest patterns, and that higher educational screen time is positively correlated with improved academic performance, whereas gaming and general entertainment usage increase addiction levels without offering measurable academic benefits, and moreover, depression levels vary notably across different phone usage purposes, collectively emphasizing the critical need for balanced smartphone usage, conscious screen time management, and the development of healthy digital habits to protect students' mental health, enhance their sleep quality, and ultimately support better academic outcomes and overall well-being.

FUTURE WORKS

- **Longitudinal Study:** Track smartphone addiction over time and its cumulative effects on academics and mental health.

- **Psychological Metrics:** Include stress, focus, and social anxiety for deeper mental health insights.

- **Intervention Tools:** Use real-time monitoring apps to study how reducing screen time improves sleep and performance.

- **Machine Learning Predictions:** Identify high-risk students early and suggest personalized solutions.

- **Comparative Studies:** Analyze usage patterns across ages, genders, and school types to find vulnerable groups.

REFERENCES

1. Gupta, A., & Sharma, P. (2022). *Impact of Smartphone Usage on Academic Performance and Mental Health among Students*. Journal of Education and Psychology, 15(3), 45–56.
2. Lin, Y., Chiang, C., & Chen, W. (2021). *Sleep Quality and Smartphone Addiction in Adolescents*. International Journal of Behavioral Sciences, 28(4), 210–218.
3. Kwon, M., Lee, J., & Kim, D. (2020). *The Relationship Between Smartphone Use, Anxiety, and Academic Outcomes*. Computers in Human Behavior, 108, 106331.
4. Kaggle. (2025). *Student Smartphone Usage Dataset*. Retrieved from https://www.kaggle.com
5. Alhassan, A. A., & Bako, S. S. (2022). *Effects of Screen Time Before Bed on Sleep and Academic Performance in Students*. Journal of Behavioral Health, 11(2), 65–74.
6. Wang, S., & Luo, J. (2021). *Smartphone Addiction and Its Effect on Adolescent Well-being*. Journal of Child and Adolescent Psychology, 19(1), 12–23.

Supporting Files

Python:

```
#Find the top 5 students with the highest phone addiction levels.
df.nlargest(5, 'Addiction_Level')[['Name', 'Addiction_Level']]
```

[ ]:

```
#Usage Purpose-Wise Percentage of Daily Phone Checks
mean_checks = df.groupby('Phone_Usage_Purpose')['Phone_Checks_Per_Day'].mean()
percentage_checks = (mean_checks / mean_checks.sum()) * 100
percentage_checks.sort_values(ascending=False)
```

[ ]:

```
# which class with the highest average phone addiction level
df.groupby('School_Grade')['Addiction_Level'].mean().sort_values(ascending=False)
```

[ ]:

```
#Which gender has the highest percentage of phone addiction level?
addiction_means = df.groupby('Gender')['Addiction_Level'].mean()
(addiction_means / addiction_means.sum() * 100).sort_values(ascending=False)
```

[ ]:

```
# Identifying High-Risk Students Based on Usage, Sleep, and Academic Performance
high_risk = df[(df['Daily_Usage_Hours'] > 6) &
               (df['Sleep_Hours'] < 7) &
               (df['Academic_Performance'] < 75)]
print(f"Number of students with potentially problematic usage: {len(high_risk)}")
print(high_risk[['Name', 'Age', 'Daily_Usage_Hours', 'Sleep_Hours', 'Academic_Performance']])
```


```
# Identifying High-Risk Students Based on Usage, Sleep, and Academic Performance
high_risk = df[(df['Daily_Usage_Hours'] > 6) &
               (df['Sleep_Hours'] < 7) &
               (df['Academic_Performance'] < 75)]
print(f"Number of students with potentially problematic usage: {len(high_risk)}")
print(high_risk[['Name', 'Age', 'Daily_Usage_Hours', 'Sleep_Hours', 'Academic_Performance']])
```


```
#Categorizing Addiction Levels (Low, Moderate, High)
def categorize_addiction(level):
    if level <= 3:
        return 'Low'
    elif level <= 7:
        return 'Moderate'
    else:
        return 'High'

df['Addiction_Category'] = df['Addiction_Level'].apply(categorize_addiction)
addiction_counts = df['Addiction_Category'].value_counts()
print("Distribution of addiction categories:\n", addiction_counts)
```

```
#Pivot Table: Analysis of Daily Usage, Sleep, and Academic Performance by Gender and Age Group
bins = [12, 14, 16, 17, 19]
labels = ['13-14', '15-16', '17', '18-19']
df['Age_Group'] = pd.cut(df['Age'], bins=bins, labels=labels)
pivot_table = pd.pivot_table(df,
                            values=['Daily_Usage_Hours', 'Sleep_Hours', 'Academic_Performance'],
                            index=['Gender'],
                            columns=['Age_Group'],
                            aggfunc='mean')
print("Pivot table of key metrics by gender and age group:\n", pivot_table.round(2))
```

```python
#Overview of Average User Metrics
key_metrics = {
    'Average Daily Usage (hours)': df['Daily_Usage_Hours'].mean(),
    'Average Sleep (hours)': df['Sleep_Hours'].mean(),
    'Average Academic Performance': df['Academic_Performance'].mean(),
    'Average Addiction Level': df['Addiction_Level'].mean(),
    'Average Screen Time Before Bed (hours)': df['Screen_Time_Before_Bed'].mean()
}
for metric, value in key_metrics.items():
    print(f"{metric}: {value:.2f}")
```

[ ]:

```python
#Count how many students use phones more than 5 hours daily
heavy_users = df[df['Daily_Usage_Hours'] > 5]
print(f"Number of students using phones more than 5 hours daily: {len(heavy_users)}")
```

[ ]:

```python
#Calculate percentage of students who check phone more than 50 times per day
frequent_checkers = df[df['Phone_Checks_Per_Day'] > 50]
percentage = (len(frequent_checkers) / len(df)) * 100
print(f"\n{percentage:.1f}% of students check their phones more than 50 times per day")
```

[ ]:

```python
# Average Addiction Level by School Grade
fig3 = px.line(df.groupby('School_Grade')['Addiction_Level'].mean().reset_index(),
               x='School_Grade', y='Addiction_Level',
               title='Addiction Level across School Grades', markers=True)
fig3.show()
```

```python
# Top 5 Locations with Highest Average Addiction Level
top_locations = df.groupby('Location')['Addiction_Level'].mean().sort_values(ascending=False).head(5).reset_index()
fig7 = px.bar(top_locations, x='Location', y='Addiction_Level',
              title='Top 5 Locations by Addiction Level', color='Location')
fig7.show()
```

```python
# Gender-Based Comparison of Activity Time (Social Media, Gaming, Education)
avg_times = df.groupby('Gender')[['Time_on_Social_Media', 'Time_on_Gaming', 'Time_on_Education']].mean().reset_index()
avg_times_long = pd.melt(avg_times, id_vars=['Gender'], var_name='Activity', value_name='Average_Time')
fig = px.bar(
    avg_times_long,
    x='Activity',
    y='Average_Time',
    color='Gender',
    barmode='group',
    title='Average Time Spent on Different Activities by Gender',
    labels={'Activity': 'Activity Type', 'Average_Time': 'Average Hours'}
)
fig.show()
```

```python
fig = px.pie(
    pd.DataFrame({'Usage Type': ['Weekday', 'Weekend'], 'Hours': [df['Daily_Usage_Hours'].mean(), df['Weekend_Usage_Hours'].mean()]}),
    names='Usage Type',
    values='Hours',
    title='📅 Average Screen Time: Weekday vs Weekend',
    hole=0.4
)
fig.show()
```

```python
#Phone Checks by Usage Purpose
mean_checks = df.groupby("Phone_Usage_Purpose")["Phone_Checks_Per_Day"].mean()
percentage_checks = (mean_checks / mean_checks.sum()) * 100
fig = px.bar(percentage_checks.sort_values(), orientation='h',
             title="Phone Checks by Usage Purpose")
fig.show()
```

R Programming

```r
install.packages("dplyr")
library("dplyr")
df<-read.csv("C:\\Users\\HP\\Documents\\final project\\teen_addiction.csv")
View(df)

# t test
#  Compare Sleep Hours on Weekdays vs Weekends
t.test(df$Sleep_Hours, df$Weekend_Usage_Hours, paired = FALSE)

#Use in an if-statement
if (p_value < 0.05) {
  print("Variances are significantly different. Use Welch's t-test.")
} else {
  print("Variances are not significantly different. Use Student's t-test (var.equal = TRUE).")
}

# f test

# Academic Performance by Social Media Usage Group
df$SM_Group <- ifelse(df$Time_on_Social_Media > 2, "High", "Low")
df$SM_Group <- factor(df$SM_Group)

f_test3 <- var.test(Academic_Performance ~ SM_Group, data = df)
print(f_test3)


if (p_value < 0.05) {
  print(paste("Variance difference is significant. F =", round(f_value, 3)))
} else {
  print(paste("No significant variance difference. F =", round(f_value, 3)))
}
```

```r
#Chi-Squre
# Is more educational screen time associated with better academic performance?
df <- df %>%
  mutate(EduTime_Group = case_when(
    Time_on_Education == 0 ~ "None",
    Time_on_Education <= 2 ~ "Low",
    TRUE ~ "High"
  ))
df <- df %>%
  mutate(Performance_Group = case_when(
    Academic_Performance >= 8 ~ "High",
    Academic_Performance >= 5 ~ "Medium",
    TRUE ~ "Low"
  ))
table9 <- table(df$EduTime_Group, df$Performance_Group)
chisq.test(table9)

# Use in an if-statement
if (chi_result$p.value < 0.05) {
  print("Significant association between Education Time and Academic Performance.")
} else {
  print("No significant association.")
}
```

```
# z test
library(BSDA)
install.packages("BSDA")
#Is the average Anxiety Level different between Gamers and Non-Gamers?
df$Is_Gamer <- ifelse(df$Time_on_Gaming > 0, "Gamer", "Non-Gamer")
group1 <- df$Anxiety_Level[df$Is_Gamer == "Gamer"]
group2 <- df$Anxiety_Level[df$Is_Gamer == "Non-Gamer"]
z.test(x = group1, y = group2, sigma.x = 2, sigma.y = 2, conf.level = 0.95)

#Anova
#Is Depression Level different across Phone Usage Purpose?
df$Phone_Usage_Purpose <- as.factor(df$Phone_Usage_Purpose)
anova4 <- aov(Depression_Level ~ Phone_Usage_Purpose, data = df)
summary(anova4)
```

Tableau