Introduction to Data Science
(CS PROF ELECT 2)

[Group 5]

Analyzing Smartphone Usage Patterns Dataset

Exploring screen time, app usage, and ussr behavior trends across different demographics.

Group Members

Canturias, Christian Gahuman

Salazar, Christian Sapalleda

Dondoyano, Gyel Dimagnaong

Jamolod, Janah Marie Pangasian

Montecalvo, Demetrio III Pansoy

1. Introduction

As mobile technology advances rapidly, it is important for technology companies, app developers, and even healthcare professionals to know how people utilize their smartphones. This project examines a dataset of smartphone use to identify patterns in screen time, app usage, and behavioral variation by demographic groups.

Research Questions

1. How much time do people spend on their phones daily on average?
2. Does more apps mean more screen time?
3. Are there differences in smartphone usage based on age or gender?
4. Is there a difference in screen time between Android users and iOS users?
5. Which demographic group shows the highest/lowest screen time?

Objectives

1. Identify general smartphone usage trends.
2. Explore how usage varies by age, gender, and device OS.
3. Examine the relationship between screen time and app usage.

2. Data Description

Database Source

https://.kaggle.com/datasets/valakhorasani/mobile-device-usage-and-user-behavior-dataset

Observations

700 unique users with various demographics, smartphone models, operating systems, number of installed apps, etc.

Variables

- User ID
- Age
- Gender
- Device Model

- **Operating System** (Android or iOS)
- **Screen On Time Daily** (Hours/day)
- **App Usage Time Daily** (Minutes/day)
- **Number of Apps Installed**

3. Data Cleaning & Preprocessing

- We checked for missing values in the columns.
  - No missing data was found, all data exists and is ready for analysis.
  - No further action needed.
- We checked for outliers in the dataset.
  - The highest screen time recorded was 12 hours, which is not far from the second highest (11.9 hours) or others and does not appear to be an extreme outlier.

- The highest app usage time recorded was 598 minutes (almost 10 hours), which was not far from the second highest (597 minutes) or others and does not appear to be an extreme outlier.
- We checked for inconsistent variables.
  - App Usage Time was given in minutes a day.
  - To make things consistent and uniform, we created a copy of App Usage Time and divided all values by 60, to create App Usage Time (hours/day) instead of App Usage Time (minutes/day).

App Usage Time (minutes)	App Usage Time (Hours)
393	6.56

- Normalized usage time values.
  - While the screen time and app usage values were consistent, all values were verified to be within a reasonable range to ensure they align with typical smartphone usage patterns.
- Removed Device Model, Battery Drain, and Data Usage columns.
  - These variables were not used in the analysis because it holds a large set of disparate values that don't give valuable groupings or trends that can contribute to answering our research questions. It wouldn't add significant information to trends of user behavior and was hence excluded during preprocessing.
- Converted categorical data.
  - To enable numerical analysis and visualizations (correlation, regression, and boxplots), categorical variables were re-coded to numeric codes using Minitab's Text to Numeric tool. This was done because statistical functions in Minitab need numerical input.
  - Each new numeric column was created separately to preserve the original text-based data for reference.
  - Gender was coded as 0 for Female and 1 for Male.

Gender (Original)	Gender Code
Female	0
Male	1

  - Operating System was coded as 0 for Android and 1 for iOS.

Operating System (Original)	Operating System Code
Android	0
iOS	1

- Grouped ages by general age groups.
  - To simplify the wide range of age values and make comparisons easier, ages were grouped into four categories: 18–24, 25–34, 35–44, and 45+. This allowed for clearer visualization and identification of behavioral trends across different life stages. The grouped data preserved individual visibility while reducing clutter in graphs.

4. Exploratory Data Analysis (EDA)

- Summary Statistics
  - Screen On Time (Hours/Day)

Statistic	Value
N (Sample Size)	700
N* (Missing Values)	0
Mean	5.27 hours
Standard Error	0.116
Standard Deviation	3.07 hours
Minimum	1 hour
Q1	2.5 hours
Median	4.9 hours
Q3	7.4 hours
Maximum	12 hours

  - App Usage Time (Hours/Day)

Statistic	Value
N (Sample Size)	700
N* (Missing Values)	0
Mean	4.52 hours
Standard Error	0.112
Standard Deviation	2.953
Minimum	0.5
Q1	1.9 hours
Median	3.8 hours
Q3	7.3 hours
Maximum	10 hours

  - Apps Installed

Statistic	Value
N (Sample Size)	700
N* (Missing Values)	0
Mean	51
Standard Error	1.02

Standard Deviation	26.9
Minimum	10
Q1	26
Median	49
Q3	74
Maximum	99

- ○ Age

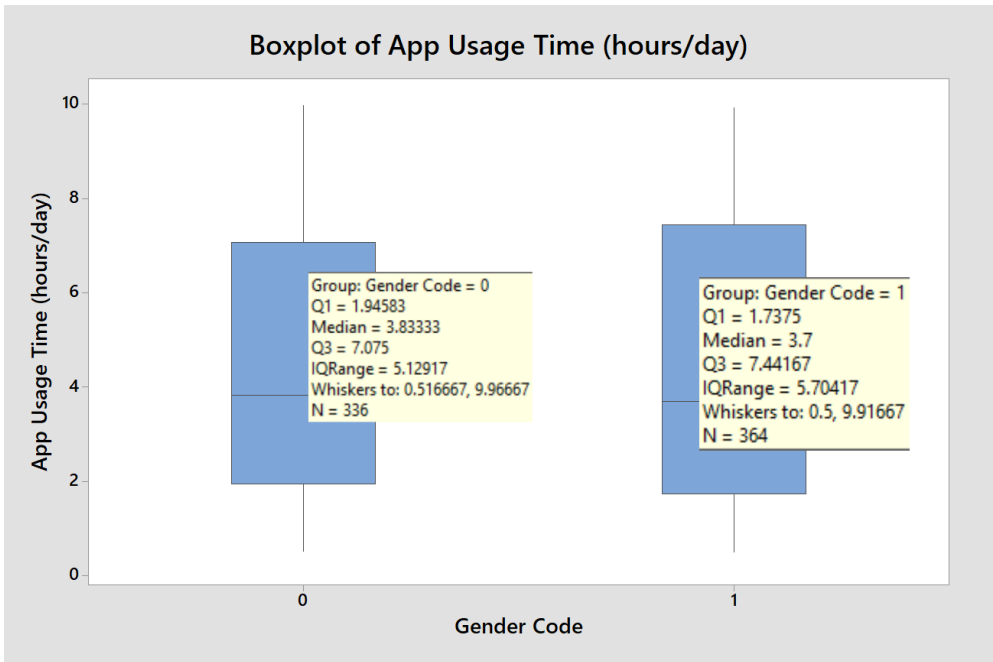| Statistic	Value
N (Sample Size)	700
N* (Missing Values)	0
Mean	38.5
Standard Error	0.5
Standard Deviation	12
Minimum	18
Q1	28
Median	38
Q3	49
Maximum	59

- **Visualizations**
  - ○ Histogram of Screen On Time Daily in General
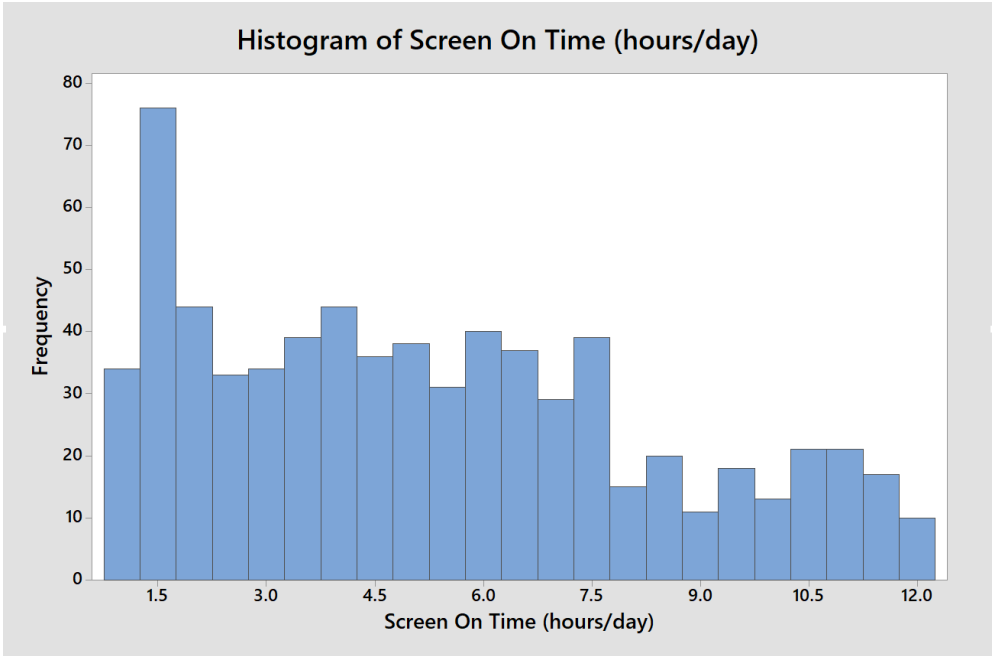
The histogram of Screen On Time (hours/day) reveals a right-skewed distribution, indicating that most users spend relatively few hours with their phone screens on each day. The highest frequency is observed between 1 and 2 hours/day, while a
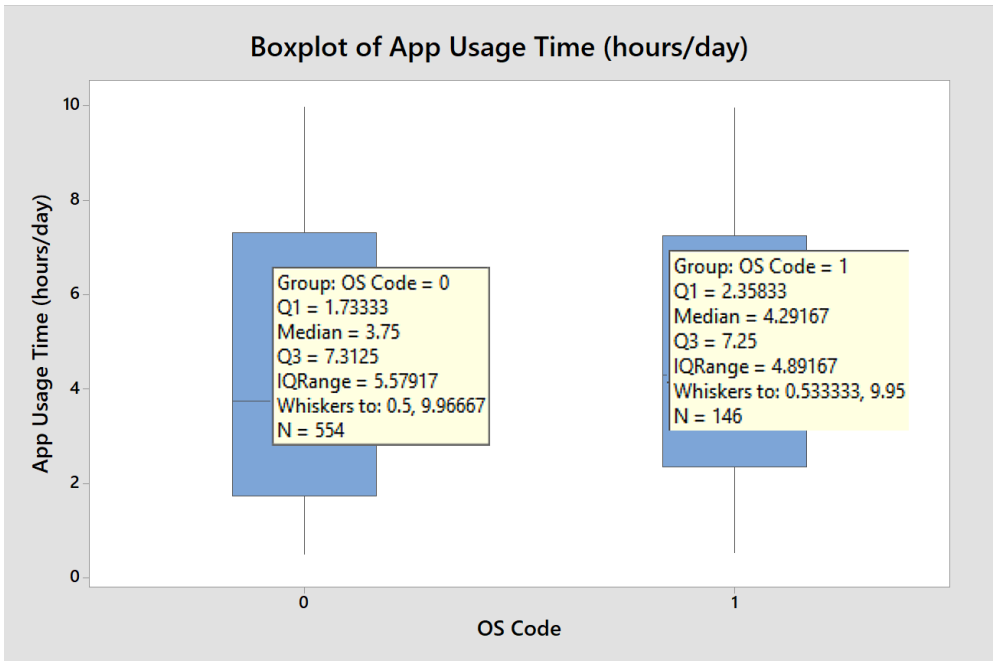
majority of users fall within the 1 to 5 hours/day range. A smaller portion of users exhibit significantly higher screen times, reaching up to 12 hours/day, suggesting a group of heavy or continuous phone users. This variation highlights diverse smartphone usage behaviors across the dataset.

- ○ Boxplot App Usage by Gender
  (Female = 0, Male = 1)

Boxplot of App Usage Time (hours/day)

Group: Gender Code = 0
Q1 = 1.94583
Median = 3.83333
Q3 = 7.075
IQRange = 5.12917
Whiskers to: 0.516667, 9.96667
N = 336

Group: Gender Code = 1
Q1 = 1.7375
Median = 3.7
Q3 = 7.44167
IQRange = 5.70417
Whiskers to: 0.5, 9.91667
N = 364

Both groups show similar distributions, with slight differences in central tendency and spread. Female (Code 0) has a median app usage time of approximately 3.83 hours/day, while Male (Code 1) has a slightly lower median of 3.7 hours/day. The interquartile range (IQR) is broader for Male (5.70) than for Female (5.13), suggesting more variability in app usage among users in Male. Both groups have a wide range of usage, extending up to nearly 10 hours/day, with a few low-usage outliers.

- ○ Boxplot App Usage by Operating System
  (Android = 0, iOS = 1)

Boxplot of App Usage Time (hours/day)

Group: OS Code = 0
Q1 = 1.73333
Median = 3.75
Q3 = 7.3125
IQRange = 5.57917
Whiskers to: 0.5, 9.96667
N = 554

Group: OS Code = 1
Q1 = 2.35833
Median = 4.29167
Q3 = 7.25
IQRange = 4.89167
Whiskers to: 0.533333, 9.95
N = 146

Median usage of apps is greater among iOS users (4.29

hours/day) than among Android users (3.75 hours/day), implying iOS users spend somewhat more time utilizing apps per day. Android users exhibit a larger interquartile range (IQR = 5.6) than iOS users (IQR = 4.89), showing greater variation in daily app use among Android users. The whiskers range from 0.5 to 9.96 hours for Android and from 0.53 to 9.95 hours for iOS, with very similar overall ranges. Sample sizes are not equal: Android (n = 554) and iOS (n = 146), which can impact comparison validity. Both groups overall have similarly comparable ranges, but iOS users have slightly more median usage with slightly less spread.

○ Scatterplot Number of Apps Installed vs. App Usage Time

Scatterplot of App Usage Time (hours/day) vs Apps Installed
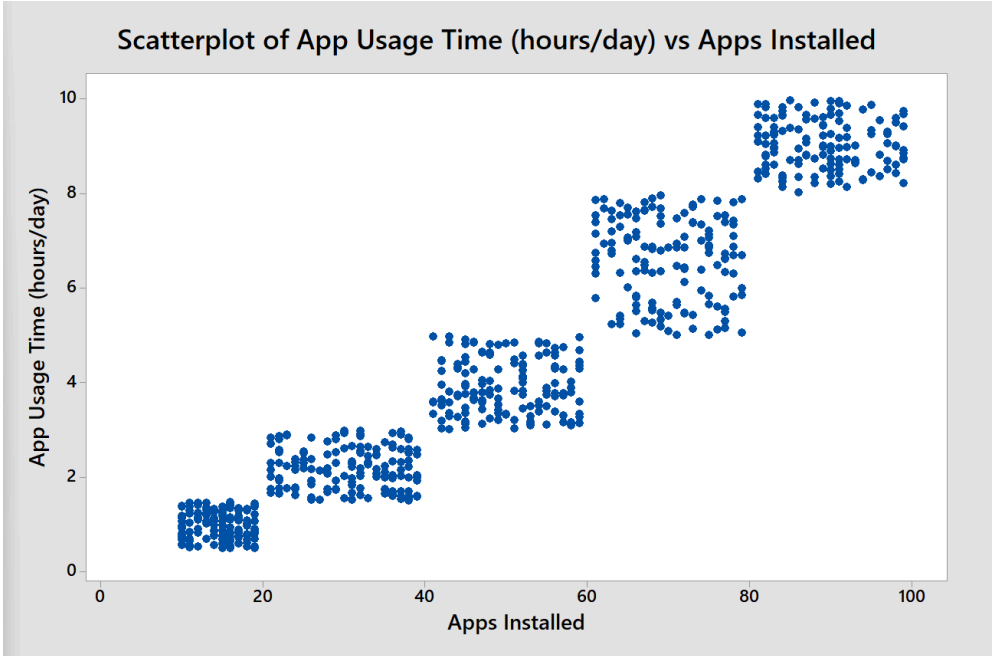
There is a definite positive correlation. As the number of apps installed grows, so does the amount of time spent using apps per day. Users with fewer apps like 10–20 use their phones for 1–2 hours/day. Users with lots of apps like 80–100 use their phones for 8–10 hours/day. The points cluster in distinct groups. This indicates that users naturally group into categories by app number and usage time. Each cluster appears somewhat flat, meaning that within a cluster users are somewhat different in usage but not tremendously. This plot implies a close relationship between the number of apps an individual installs and the amount of time they use their phone. Individuals with more apps probably have more reasons or circumstances to use their phones, resulting in greater daily usage.

○ Boxplot of App Usage Time per Age Group
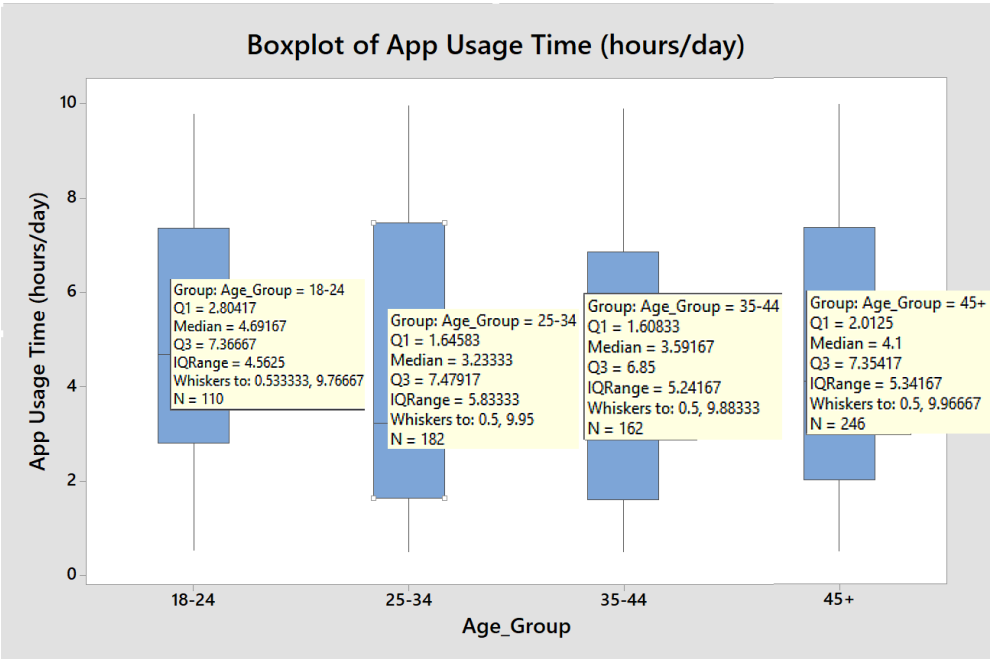
The boxplot shows the daily app usage time per age group as 18–24, 25–34, 35–44, and 45+. The highest median app usage of around 4.69 hours a day is for the 18–24 age group, showing they use apps more frequently than the other groups. The 25–34 year group, however, has the lowest median usage at approximately 3.23 hours per day but the broadest interquartile range (IQR), indicating higher variability in use among this age group. The 45+ age group also indicates high usage on the higher side, with a few people using apps almost for 10 hours a day, and a median of approximately 4.1 hours. All cohorts exhibit a reduced whisker at around 0.5 hours, reflecting light use by some users. Generally, although younger adults use apps more regularly and heavily, extensive use is observed in all age groups.

5. Key Findings & Insights

- General Screen Time Usage
  - The average daily screen-on time is 5.27 hours, with values ranging from 1 to 12 hours.
- Screen Time by OS
  - iOS users have a slightly higher median screen time (5.05 hrs) compared to Android users (4.8 hrs).
  - However, Android users show greater variability, with a wider range of usage and more outliers.
  - Sample size is imbalanced: Android (554) vs. iOS (146).
- Age Group Trends
  - Younger users (18–24) tend to have higher average screen time and app usage.
  - Older users (45+) show lower median screen time with tighter IQRs, indicating more consistent and moderate use.

6. Conclusion & Recommendations

Our breakdown of smartphone use patterns yields interesting information about usage behavior. Users spent an average of **5.27 hours** of screen time and displayed comparable app usage times. Younger age ranges **(18–34)** had increased screen time as well as app usage, reflecting higher levels of device activity than older users. **iOS users** edged **Android users** marginally in terms of screen time, although Android users displayed more varied usage patterns.

Notably, users with a higher number of installed apps also had longer screen and app usage times, possibly indicating a relationship between app diversity and device dependency.

To App Developers & Marketers,
Target efforts at younger age groups, especially 18–34-year-old users, since they are the most active and inclined to use a range of apps.

To UX/UI Designers,
Bear in mind that iOS users tend to spend slightly more regular time on their phones, improving user experience for extended session stays.

To Digital Wellness Advocates,
Interventions to decrease screen dependency should specifically focus on high-app-count users and younger populations, as these users exhibit indicators of greater smartphone entrenchment.

Future Work Recommendations
Match app usage time with particular app categories, if possible. Examine gender-specific behavior, if there are meaningful patterns. Explore the effect of device model or OS version, which was ignored here because of categorical complexity.

---

7. References

Dataset Source

Mobile Device Usage and User Behavior Dataset
https://.kaggle.com/datasets/valakhorasani/mobile-device-usage-and-user-behavior-dataset

Software Used
- Minitab Statistical Software – for data analysis, summary statistics, and visualizations.
- Google Docs – for report writing.
- Microsoft PowerPoint – for presentation preparation.

Statistical Methods & Visualization Techniques
- Basic descriptive statistics (mean, median, standard deviation, quartiles)
- Boxplots, bar charts, histograms
- Age group binning for categorical analysis
- Style and Structure Reference