Mobile Usage & Behaviour Analysis

Project Report

Author: Satyam Saurabh

Tools Used: Python (Pandas, NumPy, Matplotlib, Seaborn), Power BI, Excel
Dataset Size: 7,500+ user records

February 11, 2026

Contents

1 Executive Summary

This project analyzes mobile usage patterns from 7,500+ users across 5 device models to identify digital dependency risks and behavioral trends. The analysis includes comprehensive data exploration, visualization, and interactive dashboard creation to provide actionable insights on digital well-being, OS preferences, and battery optimization.

Key Highlights:

- Analyzed 7,500 user records across multiple dimensions
- Developed a 5-level user behavior classification system
- Created 10 comprehensive visualizations using Python
- Built an interactive 4-page Power BI dashboard
- Identified 6.03% of users as high-risk for digital dependency

2 Project Objectives

Primary Objectives:

- Analyze mobile usage patterns across different demographics (age groups, gender)
- Identify digital dependency risks among users, especially younger demographics
- Evaluate device and OS performance in terms of battery drain and data consumption
- Develop a behavior classification framework to categorize users from Minimal to Dangerous usage levels
- Create interactive visualizations to communicate insights effectively

Business Value:

- Help app developers design healthier usage features
- Assist device manufacturers in battery optimization
- Provide insights for digital well-being initiatives
- Support data-driven decision making

3 Dataset Description

Dataset Overview:

- Total Records: 7,500 users
- Number of Features: 14 columns
- Data Quality: No missing values, no duplicates

Key Features:

Feature	Description	Data Type
user_id	Unique identifier for each user	Integer
age	User's age	Integer
age_group	Categorized age groups (18-25, 26-35, 36-45, 46-60)	String
gender	User's gender (Male/Female)	String
device_model	Smartphone model (5 models)	String
os	Operating system (Android/iOS)	String
avg_screen_time_hrs	Daily screen time in hours	Float
daily_data_gb	Daily data consumption in GB	Float
app_count	Number of installed apps	Integer
battery_drain_pct	Daily battery consumption percentage	Float
charging_freq	Daily charging frequency	Integer
primary_use	Primary usage category	String
usage_score	Overall usage score	Float
behaviour_category	Pre-defined behavior category	String

Demographic Distribution:

- Age Groups: 26-35 (39.9%), 36-45 (32.4%), 18-25 (16.8%), 46-60 (10.9%)
- Gender: Male (54.5%), Female (45.5%)
- Device Models: iPhone 14 (30.8%), Samsung Galaxy S24 (24.5%), Xiaomi 13 (19.7%), Google Pixel 8 (15.0%), OnePlus 11 (9.9%)
- OS: Android (69.2%), iOS (30.8%)

4 Methodology

Step 1: Data Collection and Preparation

Loaded dataset using Pandas; Verified data types and structure; Checked for missing values and duplicates; Validated data consistency.

Step 2: Exploratory Data Analysis (EDA)

Generated statistical summaries; Analyzed distributions across categorical variables; Identified patterns and trends; Calculated correlation between numerical features.

Step 3: Feature Engineering

Created `Risk_Indicator` column based on screen time (> 8 hrs) and app count (> 70 apps).

Developed `custom_behaviour_class` using normalized scoring:

$Score = (\text{screen\_time}/\text{max\_screen\_time}) + (\text{app\_count}/\text{max\_app\_count})$

Classification: Minimal (< 0.4), Light ($0.4 - 0.8$), Moderate ($0.8 - 1.2$), Heavy ($1.2 - 1.6$), Dangerous (> 1.6).

Step 4: Data Visualization

Created 10 comprehensive visualizations using Matplotlib and Seaborn; Designed visualizations for different analysis dimensions; Saved high-resolution charts (300 DPI) for reporting.

Step 5: Dashboard Development

Built 4-page interactive Power BI dashboard; Implemented cross-page filtering using slicers; Created DAX measures for dynamic calculations; Applied consistent color scheme and formatting.

5 Data Analysis Process

5.1 Screen Time Analysis

Objective: Understand screen time patterns across demographics.
Methodology: Calculated average screen time by age group, gender, and device model; Compared usage patterns across different segments.
Key Findings:

- Average screen time: 6.41 hours/day
- Highest screen time: Age group 18-25 (6.87 hrs/day)
- Gender difference: Minimal (Female: 6.43 hrs, Male: 6.38 hrs)
- Device with highest screen time: Google Pixel 8 (6.59 hrs)

5.2 Data Consumption Analysis

Objective: Analyze daily data usage patterns.
Methodology: Grouped data consumption by age group and OS; Identified high-consumption user segments.
Key Findings:

- Average daily data usage: 2.44 GB/day
- Highest consumption: Age group 18-25 (3.39 GB/day)
- OS comparison: Android (2.45 GB) vs iOS (2.41 GB) - nearly equal

5.3 Battery Drain Analysis

Objective: Evaluate battery performance across devices.
Methodology: Calculated average battery drain by device model; Analyzed charging frequency by behavior category.
Key Findings:

- Average battery drain: 77.65%
- Highest drain: Google Pixel 8 (80.82%)
- Charging frequency increases with usage intensity
- Dangerous users charge 1.34 times/day vs Minimal users at 0.15 times/day

5.4 Digital Dependency Risk Analysis

Objective: Identify users at risk of excessive digital dependency.
Methodology: Set thresholds: Screen time > 8 hrs AND app count > 70; Calculated risk distribution by age group; Analyzed younger users (18-34) separately.
Key Findings:

- 6.03% of users identified as high-risk (452 users)
- Age group 26-35 has most high-risk users (189 users)
- Younger users (18-34) show higher dependency with 6.61 hrs average screen time

5.5 Behavior Classification

Objective: Categorize users into 5 behavioral levels.
Methodology: Created custom classification framework based on app count and screen time; Normalized scores for fair comparison; Applied thresholds for 5 categories.
Classification Results:

- Light: 72.5% (5,441 users)
- Moderate: 16.8% (1,258 users)
- Minimal: 9.8% (732 users)
- Heavy: 0.9% (67 users)
- Dangerous: 0.03% (2 users)

6 Key Findings

Finding 1: Digital Dependency Concerns

- 6.03% of users show high digital dependency (452 users)
- Younger users (18-34) are more prone to excessive usage
- Clear correlation between age and screen time
- Recommendation: Implement digital well-being features targeting younger demographics

Finding 2: Device Performance Insights

- Google Pixel 8 shows highest battery drain (80.82%)
- Minimal variation in screen time across devices
- Device model has less impact than user behavior on battery drain
- Recommendation: Focus on user education rather than device-specific optimizations

Finding 3: OS Market Dynamics

- Android dominates with 69.2% market share
- Minimal difference in data consumption between OS
- OS preference remains consistent across age groups
- Recommendation: Maintain cross-platform development strategies

Finding 4: Usage Patterns

- Primary usage: Work (36.1%), Social (34.2%), Mixed (17.2%), Gaming (12.5%)
- Work-related usage dominates across age groups
- Younger users show more diverse usage patterns
- Recommendation: Tailor features based on primary use cases

Finding 5: Behavioral Trends

- Majority of users (72.5%) show "Light" usage patterns
- Only 0.03% fall into "Dangerous" category
- Strong correlation (r=0.891) between screen time and usage score
- Recommendation: Use behavior classification for personalized interventions

7 Power BI Dashboard

7.1 Dashboard Structure

Page 1: Executive Summary

- Purpose: High-level overview of key metrics
- Components: 4 KPI Cards (Total Users, Avg Screen Time, Avg Data Usage, High Risk %); 5-Level Behavior Classification (Column Chart); OS Market Share (Donut Chart); Screen Time by Age Group (Bar Chart).
- Key Insights Display: Quick snapshot of overall dataset; Immediate visibility of risk metrics; Behavior distribution at a glance.

Page 2: Demographic Insights

- Purpose: Detailed demographic analysis

- Components: Screen Time by Gender (Column Chart); Data Consumption by Age Group (Line Chart); OS Preference by Age Group (Stacked Bar Chart); App Count by Age Group (Column Chart); Primary Usage Distribution (Pie Chart).
- Key Insights Display: Gender-based patterns; Age-wise consumption trends; Usage category preferences.

Page 3: Digital Well-being & Risk Analysis

- Purpose: Focus on digital dependency and risks
- Components: 3 KPI Cards (High Risk Users, Young Users Screen Time, Dangerous Users); High-Risk Users by Age Group (Bar Chart); Digital Dependency Risk Pattern (Scatter Plot); Usage Score Distribution (Histogram); Risk Indicator Table (Top 100 users).
- Key Insights Display: Risk distribution across demographics; Detailed risk patterns; Individual user risk scores.

Page 4: Device & Battery Optimization

- Purpose: Device performance and battery insights
- Components: Screen Time by Device Model (Bar Chart); Battery Drain by Device (Bar Chart); Charging Frequency by Behavior (Column Chart); Device Market Share (Treemap).
- Key Insights Display: Device-specific performance metrics; Battery optimization opportunities; Market distribution.

7.2 Interactive Features

Slicers (Available on all pages):

- Age Group: Multi-select dropdown
- Gender: Button-style selector
- OS: Button-style selector
- Device Model: Multi-select dropdown
- Behavior Class: Multi-select list

Slicer Benefits: Cross-page filtering; Dynamic data exploration; Personalized analysis views.

8 Technical Implementation

8.1 Python Analysis

Libraries Used: Pandas (Data manipulation and analysis), NumPy (Numerical computations), Matplotlib (Static visualizations), Seaborn (Statistical data visualization).

Code Structure:

```
# Step 1: Data Loading
df = pd.read_csv('user_behavior_dataset.csv')

# Step 2: Data Cleaning
# - Check for missing values
# - Remove duplicates
# - Verify data types

# Step 3: Feature Engineering
# - Create Risk_Indicator
# - Develop custom_behaviour_class

# Step 4: Statistical Analysis
# - Descriptive statistics
```

```
# - Correlation analysis
# - Group-wise aggregations

# Step 5: Visualization
# - 10 comprehensive charts
# - High-resolution outputs (300 DPI)
```

8.2 Power BI Implementation

Data Model: Single table structure; Calculated columns for custom metrics; DAX measures for dynamic calculations.

Interactivity: Cross-filtering enabled between visuals; Synchronized slicers across pages; Drill-through capabilities.

9 Conclusions

Major Insights

- Digital Well-being Concern: 6.03% of users show concerning usage patterns. Younger demographics (18-34) require targeted interventions. Behavior classification successfully identifies at-risk users.
- Device & OS Insights: Device choice has minimal impact on usage patterns. OS market share stable across demographics. Battery optimization more dependent on user behavior than device.
- Usage Patterns: Work and social media dominate usage. Screen time averages 6.41 hours daily. Strong correlation between app count and screen time.
- Demographic Trends: Age is stronger predictor of usage than gender. Younger users show higher data consumption. Usage patterns stabilize with age.

Project Impact

- For App Developers: Design features promoting healthier usage; Implement time management tools; Create age-specific experiences.
- For Device Manufacturers: Focus on battery optimization; Develop smart charging features; Enhance power management systems.
- For Digital Well-being Initiatives: Target younger demographics; Use behavior classification for interventions; Monitor high-risk indicators.

10 Future Scope

Potential Enhancements

- Predictive Analytics: Machine learning models to predict usage trends; Early warning system for digital dependency; Personalized recommendations.
- Time-Series Analysis: Track usage changes over time; Identify seasonal patterns; Monitor intervention effectiveness.
- Advanced Segmentation: Create more granular user segments; Develop persona-based analysis; Geographic analysis if location data available.
- Real-Time Dashboard: Live data integration; Automated alerts for high-risk patterns; Dynamic threshold adjustments.
- Extended Analysis: Social media platform breakdown; App category-wise usage; Notification impact analysis.

Technical Improvements

- Automation: Automated data pipeline; Scheduled dashboard refresh; Email reports for stakeholders.
- Advanced Visualizations: 3D visualizations; Network analysis; Geospatial mapping (if location data available).
- Integration: API connections to live data sources; Mobile app integration; Cloud deployment.

Contact Information

Project Author: Satyam Saurabh
Email: satyam2610saurabh@gmail.com
LinkedIn: `https://www.linkedin.com/in/satyam-saurabh`
GitHub: `https://github.com/i-satyamsaurabh`
Portfolio: `https://i-satyamsaurabh.github.io/My-Portfolio/`