Project One

Behavioral Impact of Social Media Usage on Productivity and Engagement

Team 11:
Cameron Magor
Angelina Wright
Daena Wallace

Table of Contents

Project Overview

Description:

This project investigates the behavioral impact of social media usage, with a focus on how different factors such as device type, location, profession, and engagement levels influence productivity and social media addiction. By analyzing data related to user habits on various devices, locations, and professions, we aim to uncover insights into how specific behaviors and patterns contribute to productivity loss and social media addiction.

Project Goals and Objectives:

Our primary goal is to analyze user data to identify patterns that could help explain the behavioral impact of social media usage. Specifically, we aim to understand how location, device type, profession, engagement levels and scroll rate contribute to addiction levels and productivity loss. Each question brings a different dimension to this analysis, helping us build a comprehensive understanding of the behavioral impacts of social media.

Project Ideation

Why This Topic Was Chosen:

As a team, we were interested in understanding the effects of social media on daily productivity, especially since social media has become an integral part of both personal and professional life. We wanted to explore the factors that contribute to social media addiction and productivity loss, as this could provide valuable insights for individuals looking to manage their time better and organizations aiming to enhance workplace productivity. Each team member contributed a unique perspective on the broader question of how social media impacts behavior, resulting in a multifaceted approach to the topic.

Data Collection and API Integration

Summary

Our data collection and integration process focused on identifying a dataset that matched our project goals of studying how behaviors impact productivity and engagement in various regions and setting up an API connection.

Initial Data Sources Considered:

- Dataset: "Mobile Device Usage and User Behavior"
- Location: Kaggle
- Issue: lacked the detailed geographic data we needed for our analysis

Final Dataset Selection:

- Dataset: "Time Wasters on Social Media"
- Location: Kaggle

API Integration:

- Set up Kaggle Account
- Created an Token to connect to API
- Installed Kagglehub to dev environment
- Added dependency 'import kagglehub' and 'kagglehub.login()' to Jupyter

Data Cleanup and Preparation

Data Cleaning and Standardization:

- Checked for null values and removed duplicate rows to ensure data accuracy.
- Corrected typos and standardized column names for clarity (e.g., "Barzil" to "Brazil" and "DeviceType" to "Device Type").
- Indexed column was removed when exporting the clean data to csv

Enhancing Data for Analysis:

- Created a "Time Spent" category column to group users by usage ranges (e.g., "0-50," "50-100") for insights into productivity loss.
- Grouped addiction levels (0-7) to improve visualization across locations.

Documentation and Clean Data Storage:

- Created a "data" folder to store the cleaned dataset for easy reference.
- Documented all data cleaning steps in the README file.

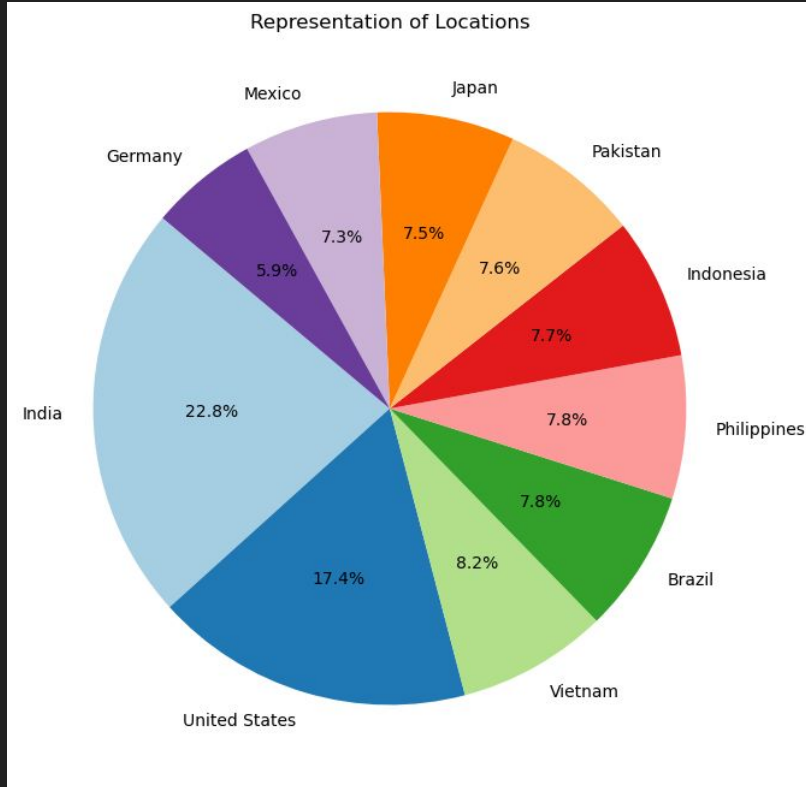# Are addiction levels higher in certain locations?

Objective:

To determine if self-assessed social media addiction levels vary by location, providing insights into regional patterns of social media usage and their impact on users.

Hypothesis:

Social media addiction is influenced by location, with certain regions, such as the United States, exhibiting potentially higher addiction levels compared to others.

Are addiction levels higher in certain locations?

Representation of Locations

The higher representation of India and the United States in the dataset may skew the overall outcome, making trends more reflective of behaviors in these countries rather than globally representative.
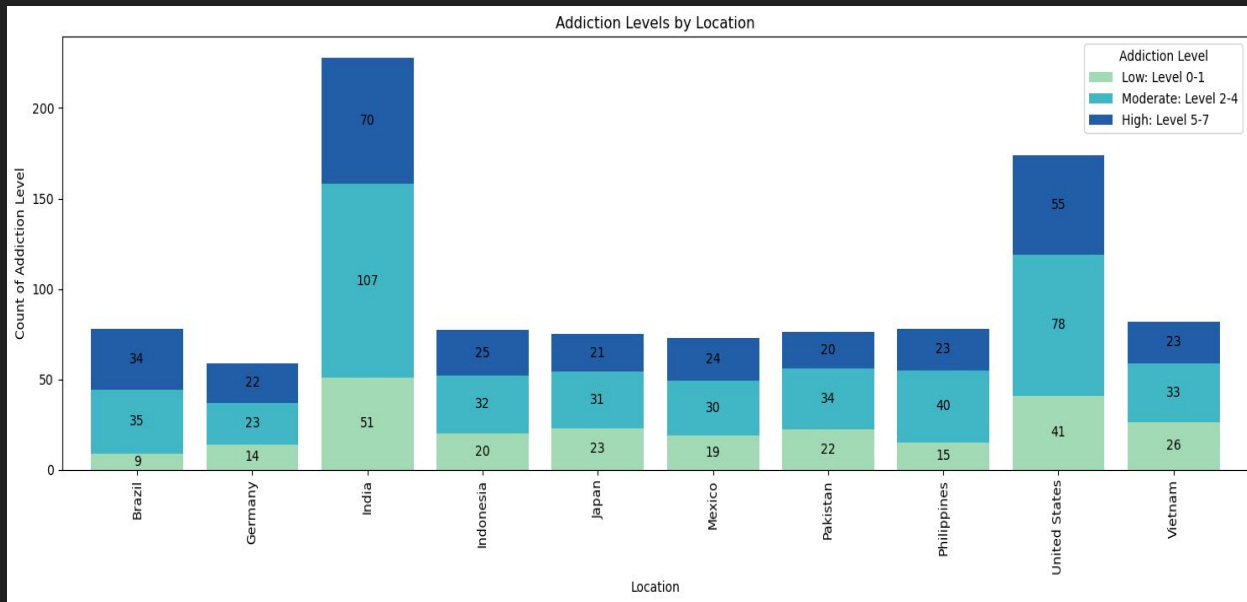
Key Insights

Location Representation

#1 - India: 22.8%
#2 - United States: 17.4%
#3 - Vietnam: 8.2%
#4 - Brazil & Philippines: 7.8%

Are addiction levels higher in certain locations?

India and the United States show that "High" addiction levels are likely influenced by their larger participant representation in the dataset.

Addiction Levels by Location

Key Insights

"High" Addiction Level
India: 70
United States: 55
Brazil: 34
Indonesia: 25

"Moderate" Addiction Level
India: 107
United States: 78
Philippines: 40
Brazil: 35

"Low" Addiction Level
India: 51
United States: 41
Vietnam: 26
Japan: 23

Are addiction levels higher in certain locations?

Major Findings

High Addiction: India and the United States have the highest levels of social media addiction, suggesting stronger dependency and frequent engagement.

Moderate Usage: India, the United States, the Philippines, and Philippines show moderate addiction levels, indicating regular yet balanced social media use.

Lower Addiction: Users in Brazil and Germany exhibit restrained social media use, which may have a less impact on productivity.

Skewed Data: Larger samples from India and the United States shape insights, making trends more reflective of these regions than others.

The data supports the hypothesis that social media addiction is influenced by location, with certain regions, including the United States, demonstrating higher self-reported addiction levels.

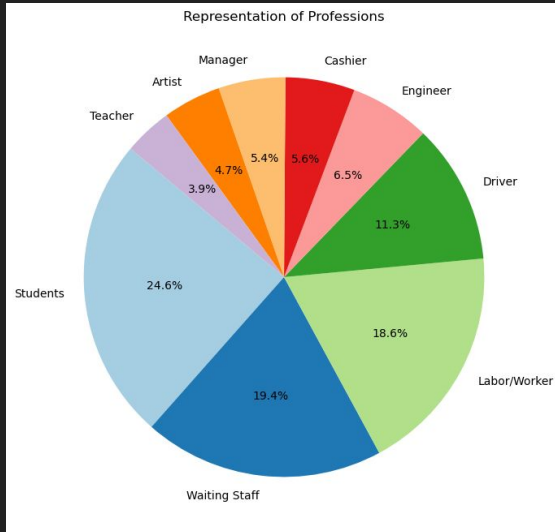# Is there a relationship between profession and device type?

Objective:

To explore whether different professions have a distinct preferences for device types used for social media, revealing patterns in device choice based on profession.

Hypothesis:

Profession influences device preference, with mobile-focused roles favoring smartphones and desk-based roles leaning toward computers and tablets.
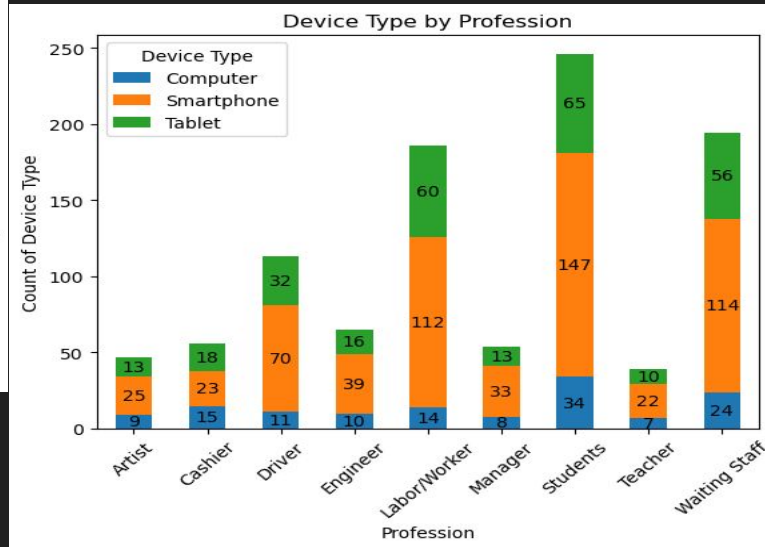
Is there a relationship between profession and device type?

Representation of Professions

TOP 4:

#1 - Students (24.6%)

#2 - Waiting Staff (19.4%)

#3 - Labor/Worker (18.6%)

#4- Driver (11.3%)

Device Type by Profession

Key Insights

Smart Phones

Widely used across professions with highest counts in Students (147), Waiting Staff (114) and Labor/Worker (112) suggesting these professions may use it for mobile flexibility.

Tablet

Highest use in Student (65) and Labor/Worker (60) suggesting relevance for education and ease of use in labor-intensive roles.

Computer

Most used by Students (34) and Waiting Staff (24) suggesting Students may use computers for educational purposes and Waiting Staff may use it for training.

Is there a relationship between profession and device type?

Major Findings

Smartphone Dominance in Mobile Professions: Professions that require mobility, such as Drivers (70) and Labor/Workers (112), show a strong preference for smartphones.

Tablet Use in Education: Students show a significant usage of tablets (65).

Computer Preference in Desk-Oriented or Administrative Tasks: Computers are not the dominant device in any of the professions, however, they are moderately used by Students (34) and Waiting Staff (24).

These findings support the hypothesis that social media behavior and device type are influenced by the nature of the profession, with mobile roles favoring smartphones for quick social media interactions, while educational roles adopt a blend of smartphones and tablets, likely for academic use and social media engagement.

Is there a relationship between scroll rate and engagement level?

Objective: The primary goal of this analysis is to investigate whether different types of devices (e.g., computers, smartphones, tablets) are associated with varying levels of scroll rates and engagement levels among users.

Hypothesis: Device type plays a significant role in scroll rate/engagement level, with certain devices (such as smartphones) potentially contributing to greater distraction compared to others.

Is there a relationship between scroll rate and engagement level(by device type)?
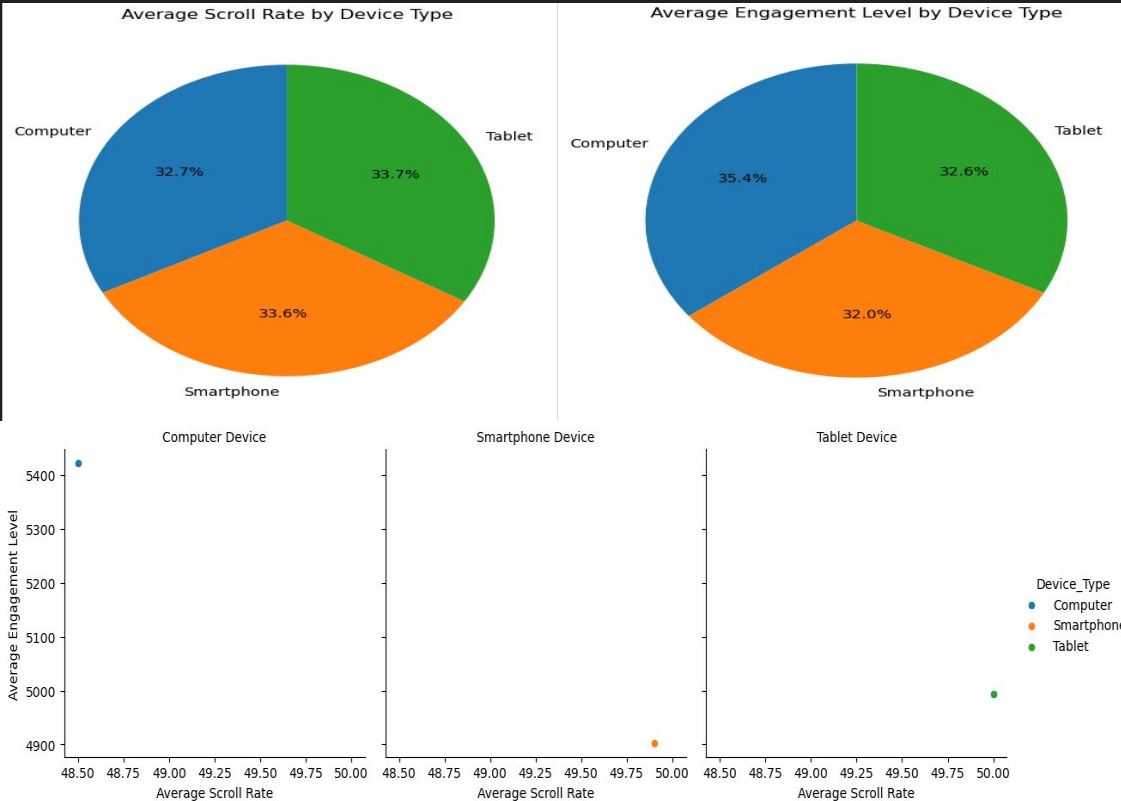
Key Insights

Device Comparison:

Despite having data from three separate devices, the average scroll rates and average engagements levels between them are almost homogenous.

User Behavior Insights:

Even though the average engagement levels are slightly higher for computers compared to other devices, the average scroll rates are slightly lower.

Target Audience:

Analyzing the data we observe that computer users have a slightly higher engagement level indicating that marketing efforts could focus more towards computers.

Scroll Rate Vs. Engagement Level Pt.2

Average Scroll Rate Vs. Engagement Level by Device Type

Relationship between Scroll Rate and Engagement Level
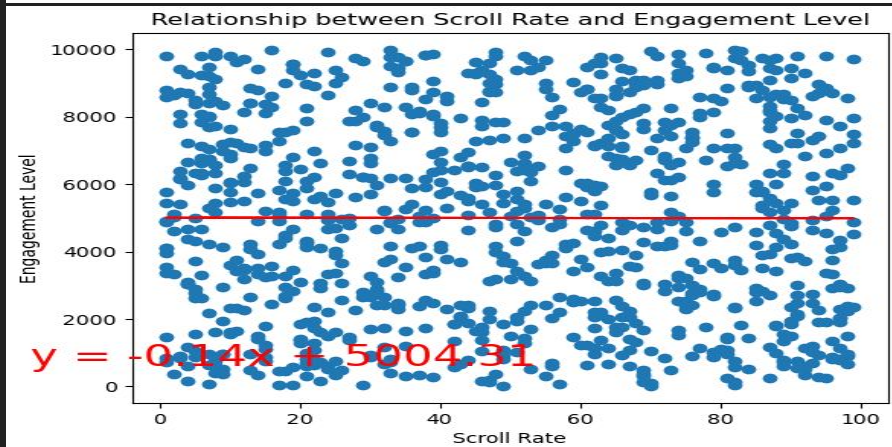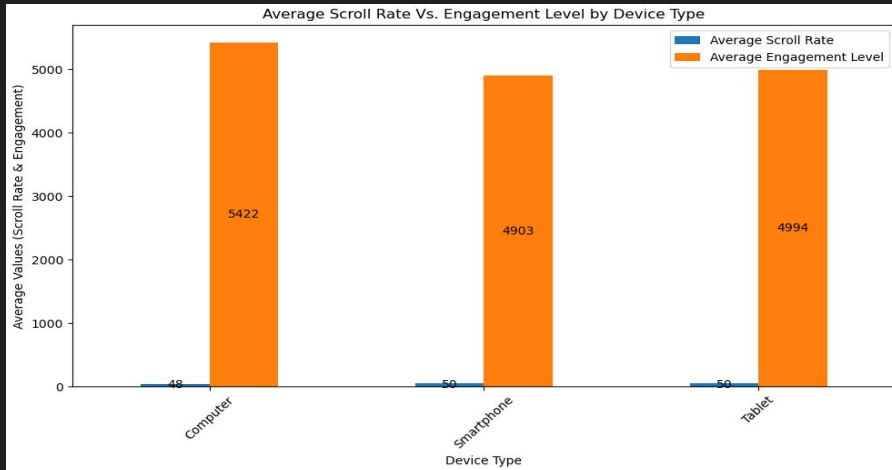
Correlation Coefficient: PearsonRResult(statistic=-0.0014412806711375228, pvalue=0.9636925819254766)

Average Scroll Rate
Computer: 48
Smartphone: 50
Tablet: 50

Average Engagement Level
Computer: 5422
Smartphone: 4903
Tablet: 4994

Regression Equation
Y = -0.14x + 5004.31

Correlation Coefficient
Statistic = -0.0014, pvalue = 0.964

Analysis:

Along with not seeing much of a relationship/correlation between the two variables, the correlation coefficient and line of regression both confirm that there is a very weak relationship and it is not statistically significant.

Is there a relationship between scroll rate and engagement level?

Major Findings

Device Type Impact: The analysis indicates that computers generally correlate with higher engagement levels than other devices. This is due to slightly lower scroll rates.

Profession-based Variability: Productivity loss/engagement levels on different devices varies by profession, with some jobs experiencing more distraction on certain device types.

# Does the type of device impact productivity loss?

Objective: The primary goal of this analysis is to investigate whether different types of devices (e.g., computers, smartphones, tablets) are associated with varying levels of productivity loss among users. With this study we are trying to understand the behavioral impact of technology, particularly in the context of social media usage.

Hypothesis: Device type plays a significant role in productivity loss, with certain devices (such as smartphones) potentially contributing to greater distraction and productivity loss compared to others.

Does the type of device impact productivity loss?

Distribution of Productivity Loss by Device Type
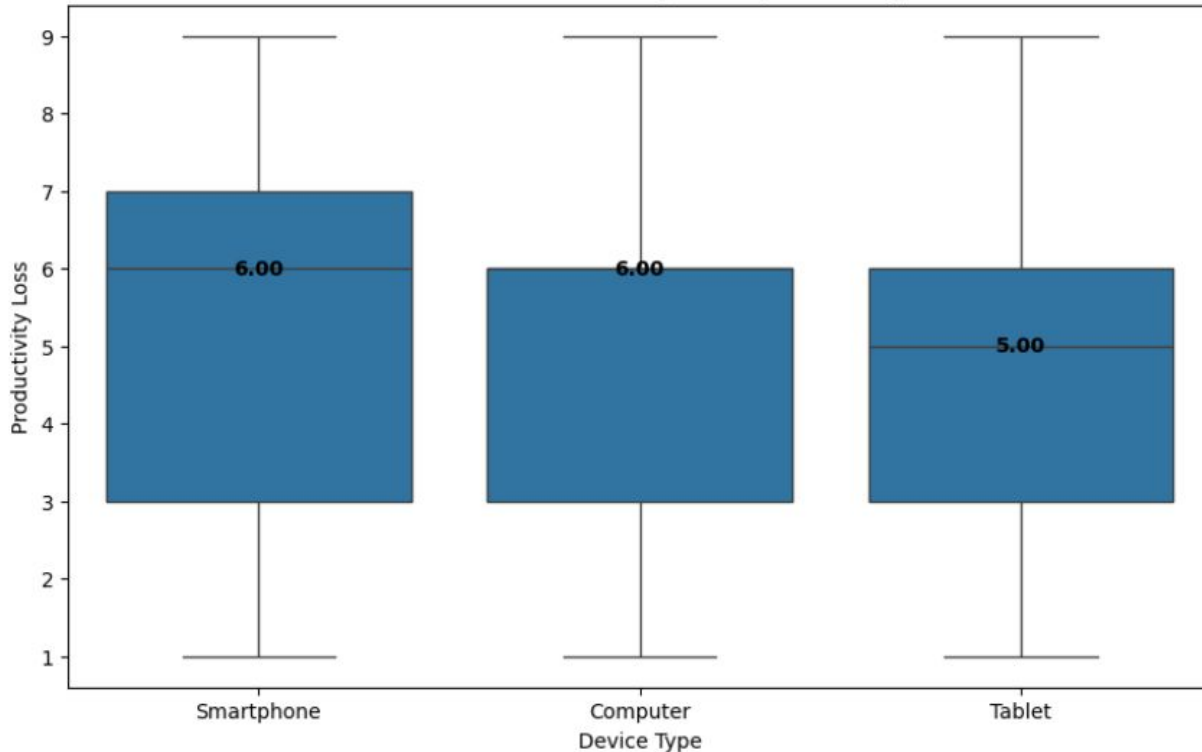
Key Insights

Smartphones:

Smartphones have the highest median productivity loss, with a broader range and more variability than computers and tablets.

Computers:

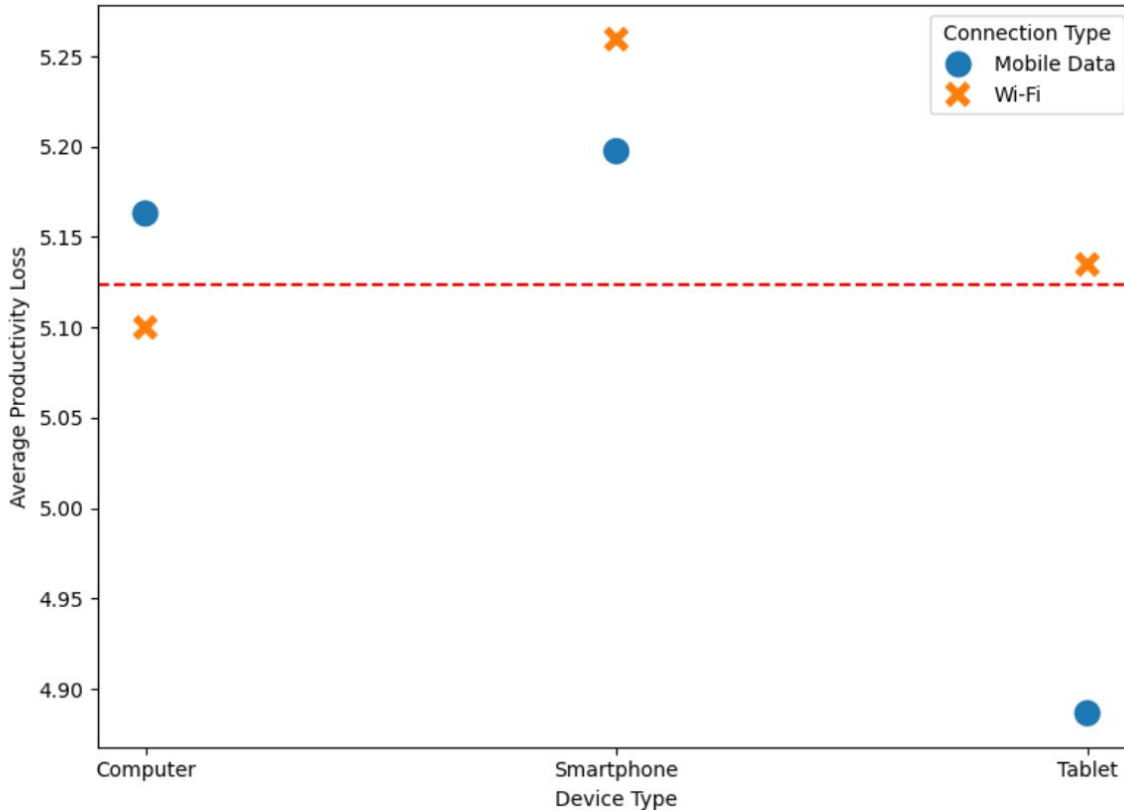Computers show the lowest media productivity loss and a slightly narrower distribution.

Tablets:

Tablets fall in between, with a median closer to computers but a wider spread than computers.

Does the type of device impact productivity loss?

Both device type and connection type might influence productivity loss, with Wi-Fi connections showing a tendency toward slightly higher productivity loss.

Average Productivity Loss by Device Type and Connection Type

Key Insights

Device Type:

Tablets have the highest productivity loss on WiFi, while computers shows the lowest on mobile data.

Connection Type:

Wi-Fi connections generally result in higher productivity loss than mobile data across all devices.

Average Line:

The red line marks the overall average productivity loss (~5.10). WiFi connected devices are mostly above this line, while mobile data connections are at or below it.

Does the type of device impact productivity loss?

Major Findings

Device Type Impact: The analysis indicates that smartphones generally correlate with higher productivity loss than other devices. This is probably due to other social media apps availability.

Connection Type Influence: The productivity loss was lower on mobile data compared to Wi-Fi for smartphones, hinting that users are more likely to be distracted by their phones while on mobile data.

Analysis and Methodology

Summary

Our analysis combined correlation analysis and summary statistics to explore patterns in the data and answer our research questions.

To ensure a cohesive approach, we focused on device type as a central theme, using a variety of visualizations to examine behavioral impacts of social media usage. Each question was represented with specific charts that highlighted different aspects of the data.

Key Challenges:

- Balancing high engagement values against lower scroll rate values required careful visualization choices to ensure clarity.
- Data imbalance, with certain regions (like India) having more participants, introduced potential skew in some of our analyses.
- Some data points lacked consistent measurement units, which posed additional formatting challenges.

Overall Summary Statement

Our analysis yielded several key findings regarding the impact of device type on productivity loss and how it relates to behavioral factors in social media usage.

Key Takeaways:

1. **Smartphones and Productivity Loss**: Smartphones show the highest and most varied productivity loss compared to other devices, making them a significant factor in productivity decline.

2. **Time Spent Consistency**: Productivity loss remains steady across different amounts of time spent on devices, suggesting that the type of device, rather than time spent, impacts productivity more.

3. **Profession-Based Variability**: Lower-wage professions, like cashiers and waiting staff, experience higher productivity loss, especially when using smartphones.

4. **Regional Differences**: Certain regions, such as India and Pakistan, show higher productivity loss. This could indicate cultural or environmental influences on social media usage and productivity.