

Milestone 2: Data, Methods, and Analysis Reporting Update

Project Title: Analyzing the Impact of Daily Screen Time on Mental Health

(This is an individual Project.)

1. Data Overview

The dataset used is the 'Mental Health and Technology Usage Dataset' obtained from Kaggle. It contains 10,000 entries and 14 columns representing user demographics, technology usage, and mental health indicators. All columns were confirmed to have non-null values, and data types were appropriate for analysis.

Key Variables: Screen_Time_Hours, Mental_Health_Status, Stress_Level, Sleep_Hours, Physical_Activity_Hours, Social_Media_Usage_Hours.

2. Methods and Exploratory Analysis

Exploratory Data Analysis (EDA) included generating summary statistics and multiple visualizations to understand data patterns. Histograms, boxplots, and heatmaps were used to assess distributions, group comparisons, and correlations among variables. A simple linear regression was conducted to explore the relationship between Screen_Time_Hours and Mental_Health_Status.

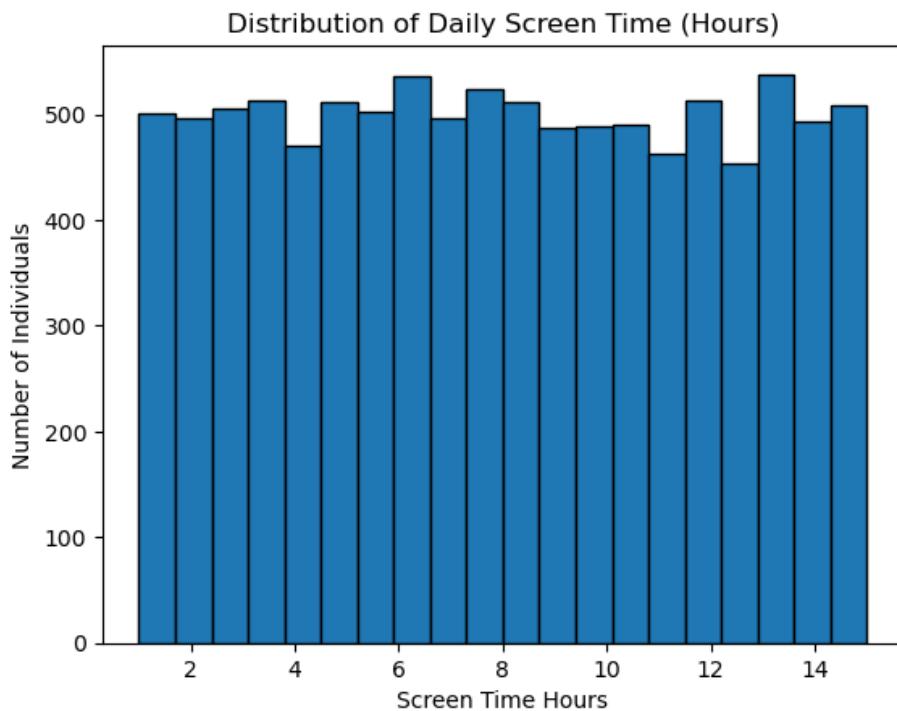


Figure 1: Distribution of Daily Screen Time (Hours)

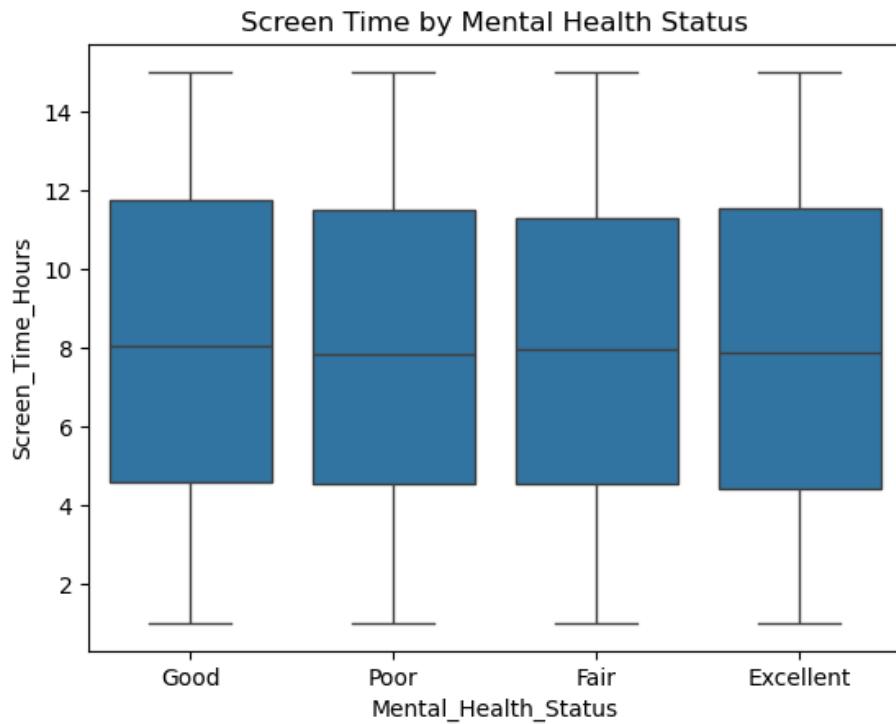


Figure 2: Screen Time by Mental Health Status

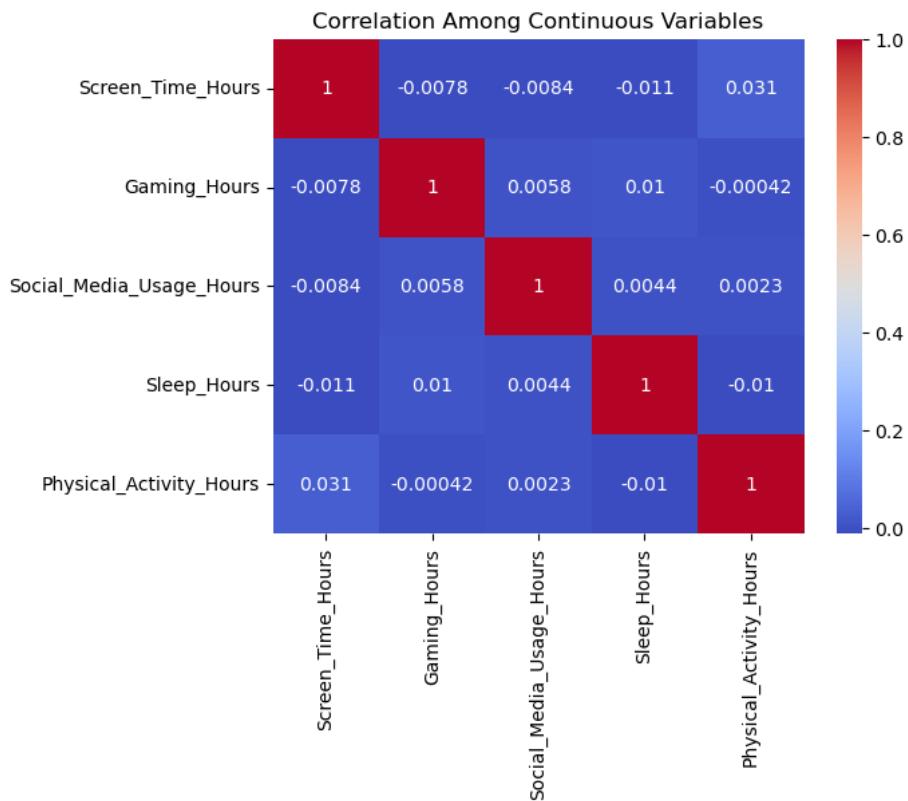


Figure 3: Correlation Among Continuous Variables

3. Key Observations

- Data quality was strong with no missing or duplicate values.
- The histogram shows screen time is evenly distributed among participants.
- Boxplots indicate minimal variation in screen time across mental health categories.
- Correlation heatmap reveals very weak relationships among continuous variables.
- Regression results:

Intercept = 2.50, Slope = 0.00, $R^2 = 0.000$, suggesting no linear relationship between screen time and reported mental health status in this dataset.

4. Early Insights

Initial findings show that screen time alone is not a strong predictor of mental health self-assessment. The low correlation and near-zero regression slope indicate other variables e.g., physical activity, sleep, or social interaction may play a more significant role. Future analyses will incorporate these factors into a multivariate model.