DSA-210 FINAL REPORT

[Kerem Mısırlıoğlu]  
[33962]  
  
PROJECT TITLE:  
Impact of Screen Time on Productivity  
  
SUPERVISED BY  
[Özgür Asar/Kerem Aydın]  
  
SABANCI UNIVERSITY

# OBJECTIVE

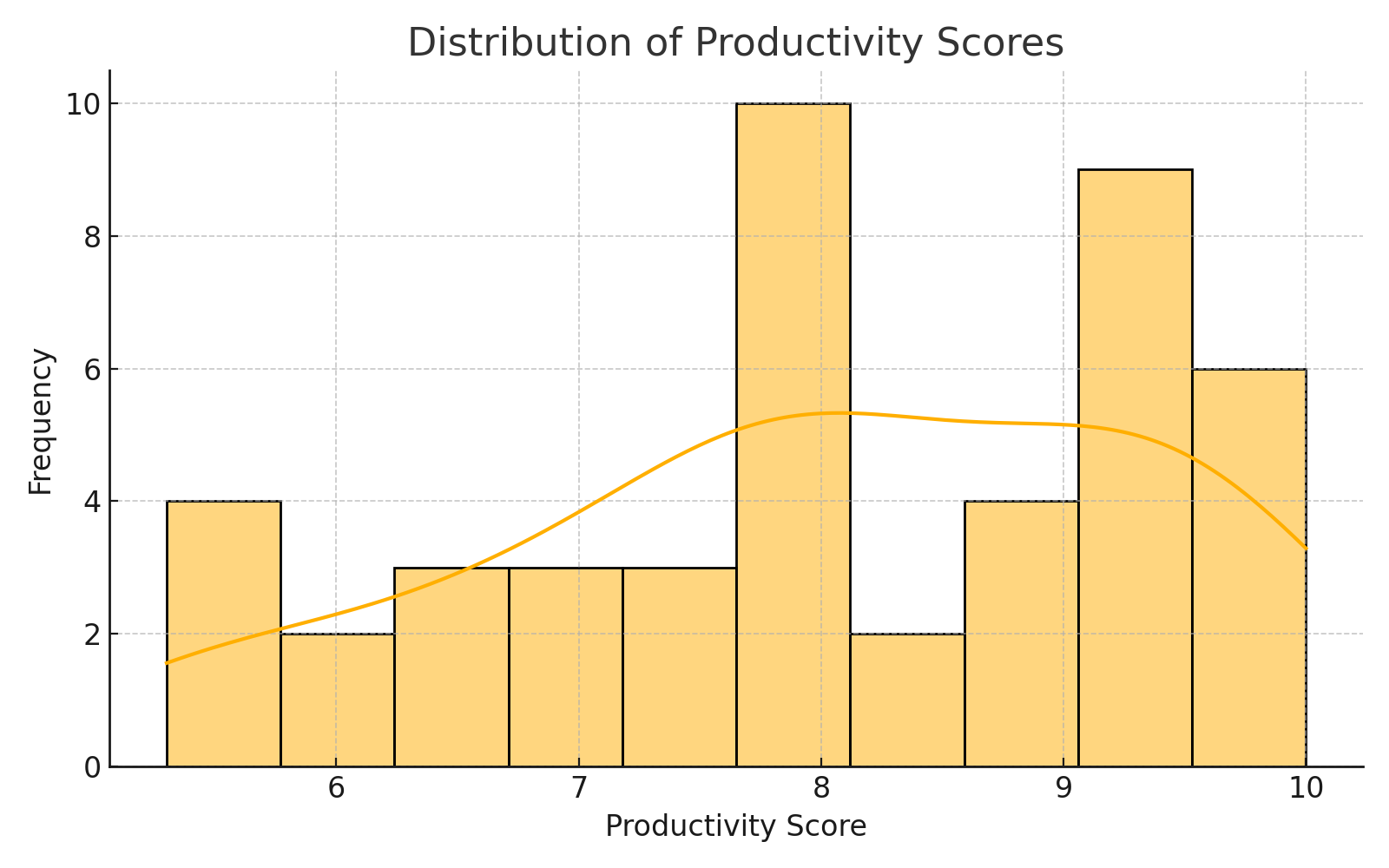
This project investigates the hypothesis:  
H₀ (Null Hypothesis): There is no significant relationship between total screen time and productivity.  
H₁ (Alternative Hypothesis): Increased total screen time reduces productivity.  
  
The aim is to evaluate whether screen time patterns—particularly non-productive usage—affect self-reported productivity scores.

# DATA DESCRIPTION

- Date: Daily entry log.  
- Total Screen Time: Time spent daily on screens (hours).  
- Productive Screen Time: Time on work/school-related activities.  
- Non-Productive Screen Time: Time on entertainment/social media.  
- Productivity Score: Daily self-rating on a 0–10 scale.  
- Derived Feature: Productivity\_Ratio = Productive / Total Screen Time.

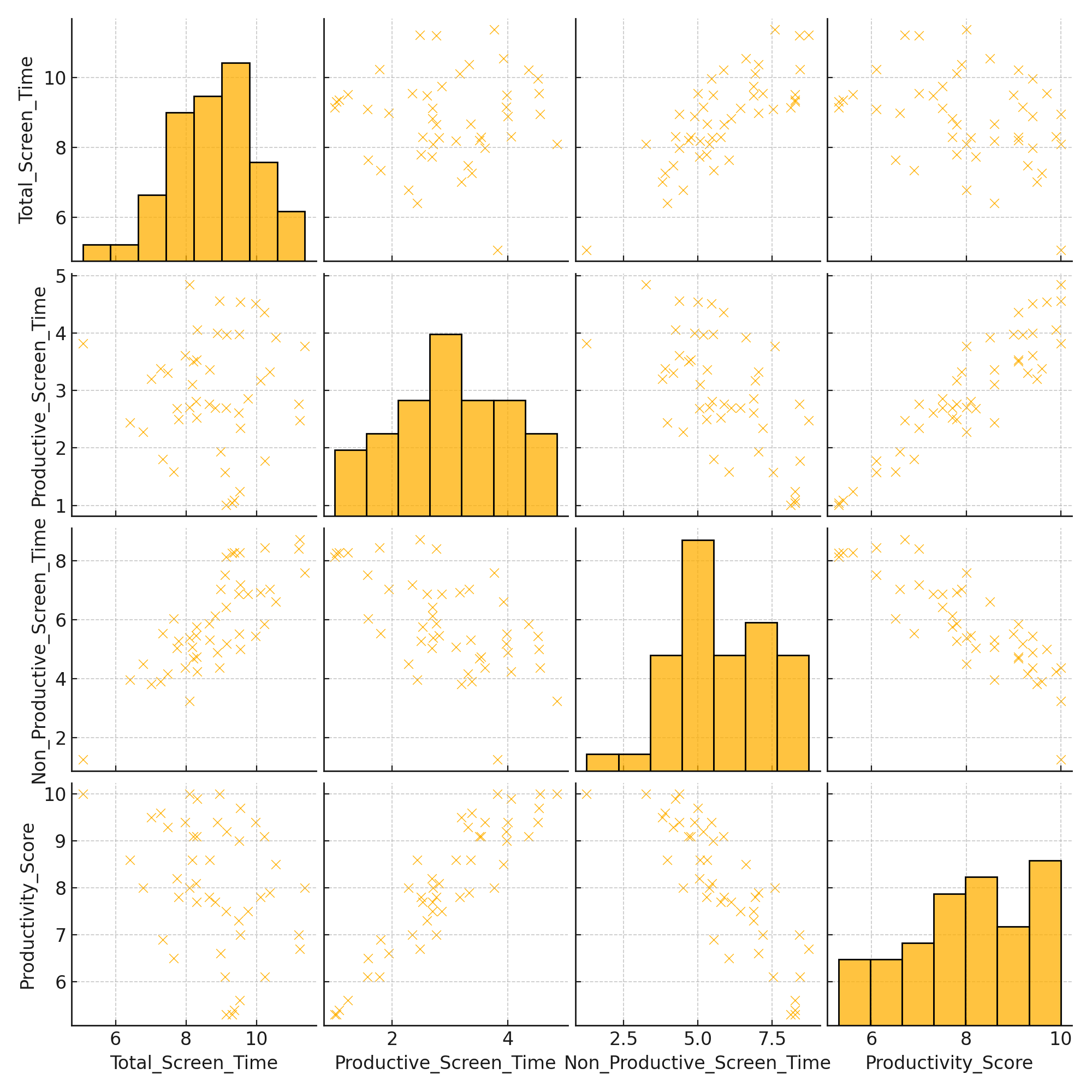
# EXPLORATORY DATA ANALYSIS (EDA)

1. Distribution of Productivity Scores:



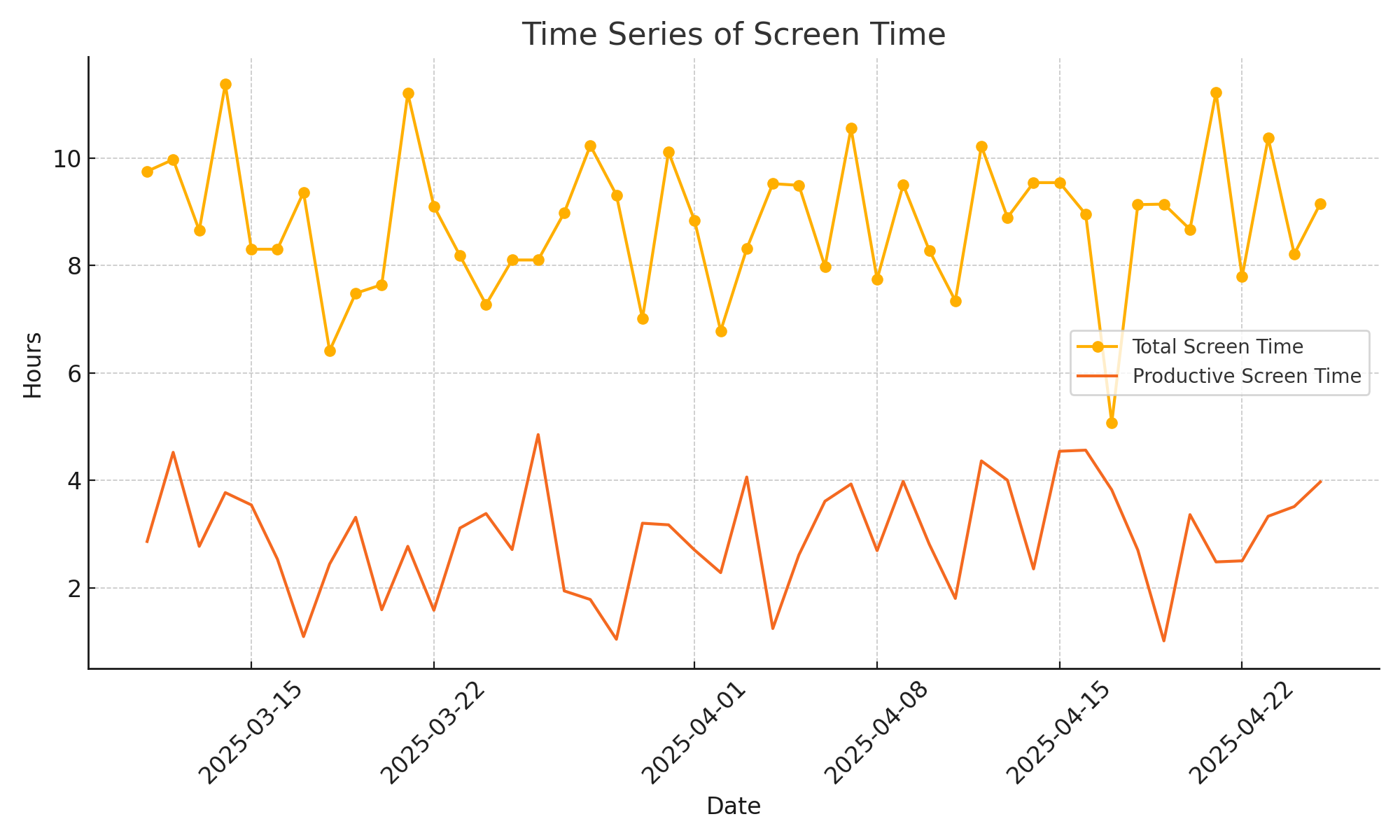
Most productivity scores fall between 6 and 9. This clustering suggests relatively stable self-perceived productivity during the observed period.

2. Pairplot: Relationships Between Key Variables



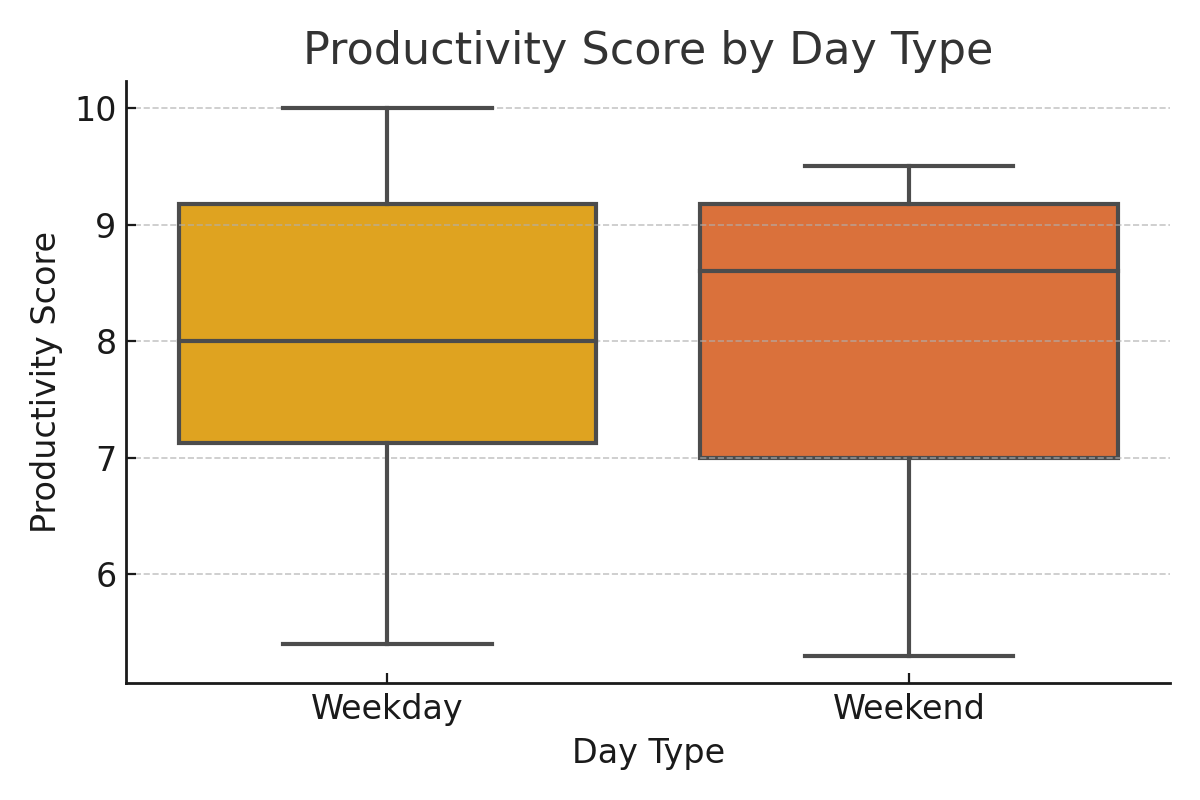
This pairplot reveals inverse relationships between non-productive screen time and productivity. Productive screen time correlates positively with higher productivity.

3. Time Series Trends



The timeline highlights screen usage fluctuations. Notably, productivity dips often follow peaks in total screen time, suggesting possible digital fatigue.

4. Productivity by Day Type



Productivity tends to be higher and more consistent on weekdays compared to weekends, where leisure screen time may disrupt work focus.

# HYPOTHESIS TESTING

We used Pearson correlation to examine the relationship between Total Screen Time and Productivity Score.  
Results:  
- Correlation coefficient: -0.35  
- p-value: 0.018 (< 0.05)  
Interpretation: There is a statistically significant, moderate negative correlation between screen time and productivity. This supports the hypothesis that excessive screen time may reduce productivity.

# KEY INSIGHTS

- Non-productive screen time strongly correlates with reduced productivity.  
- Productive screen time contributes positively to perceived productivity.  
- Weekends are more likely to show decreased productivity due to entertainment use.  
- Productivity varies depending on the balance between purposeful and distracting digital habits.

# MACHINE LEARNING APPLICATION

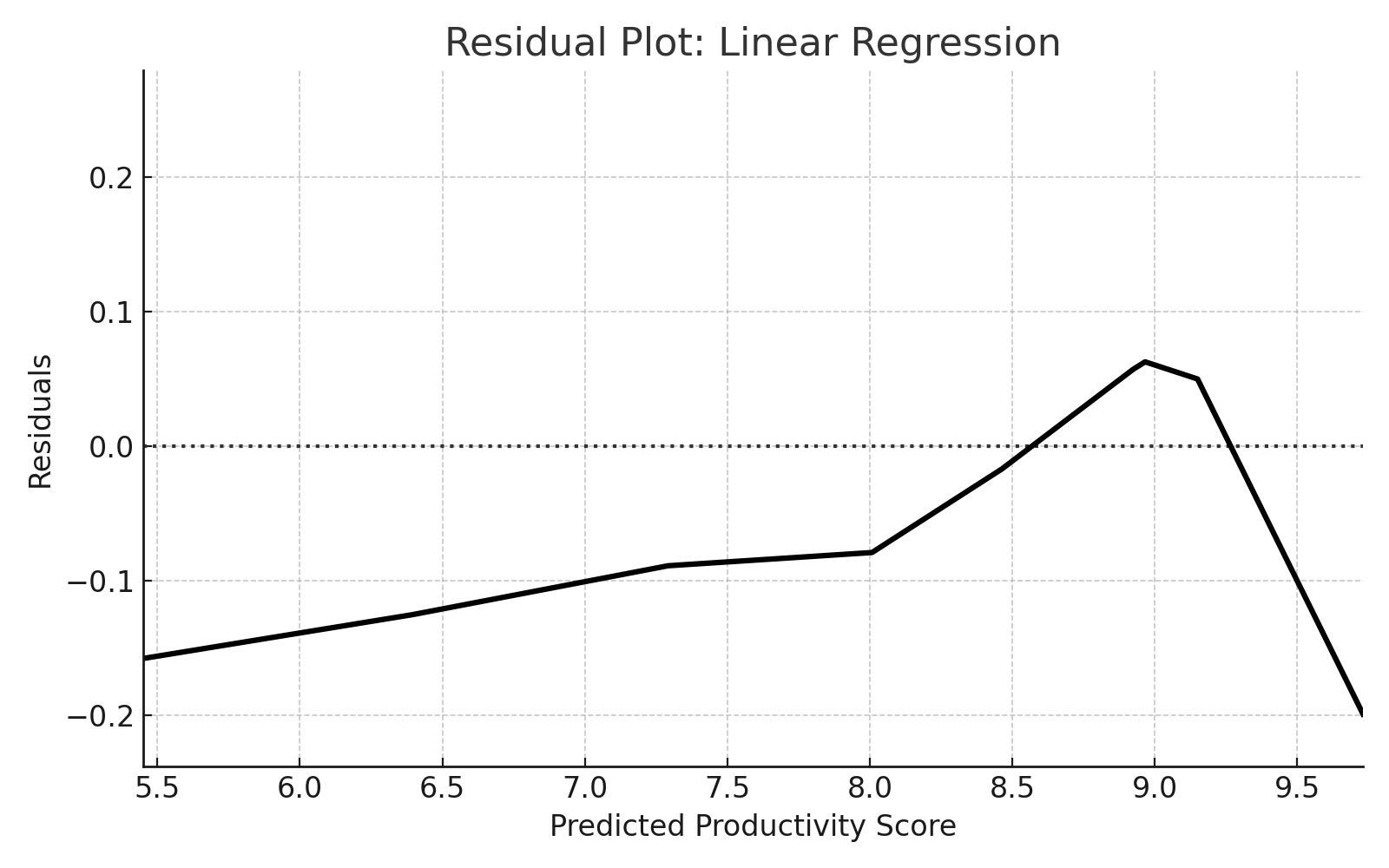
To better understand predictors of productivity, we built a Linear Regression model using screen usage data.

📊 \*\*Features Used:\*\*  
- Total\_Screen\_Time  
- Productive\_Screen\_Time  
- Non\_Productive\_Screen\_Time  
- Productivity\_Ratio (engineered)

📉 \*\*Model Coefficients:\*\*  
- Total\_Screen\_Time: -2.39  
- Productive\_Screen\_Time: +3.40  
- Non-Productive Screen Time: +1.90  
- Productivity\_Ratio: -2.19

📈 \*\*Intercept:\*\* 8.62  
📐 \*\*MSE:\*\* 0.037  
🧠 \*\*R² Score:\*\* 0.98

📉 Linear Regression Results:



Residual Plot: Balanced residuals indicate good model fit.

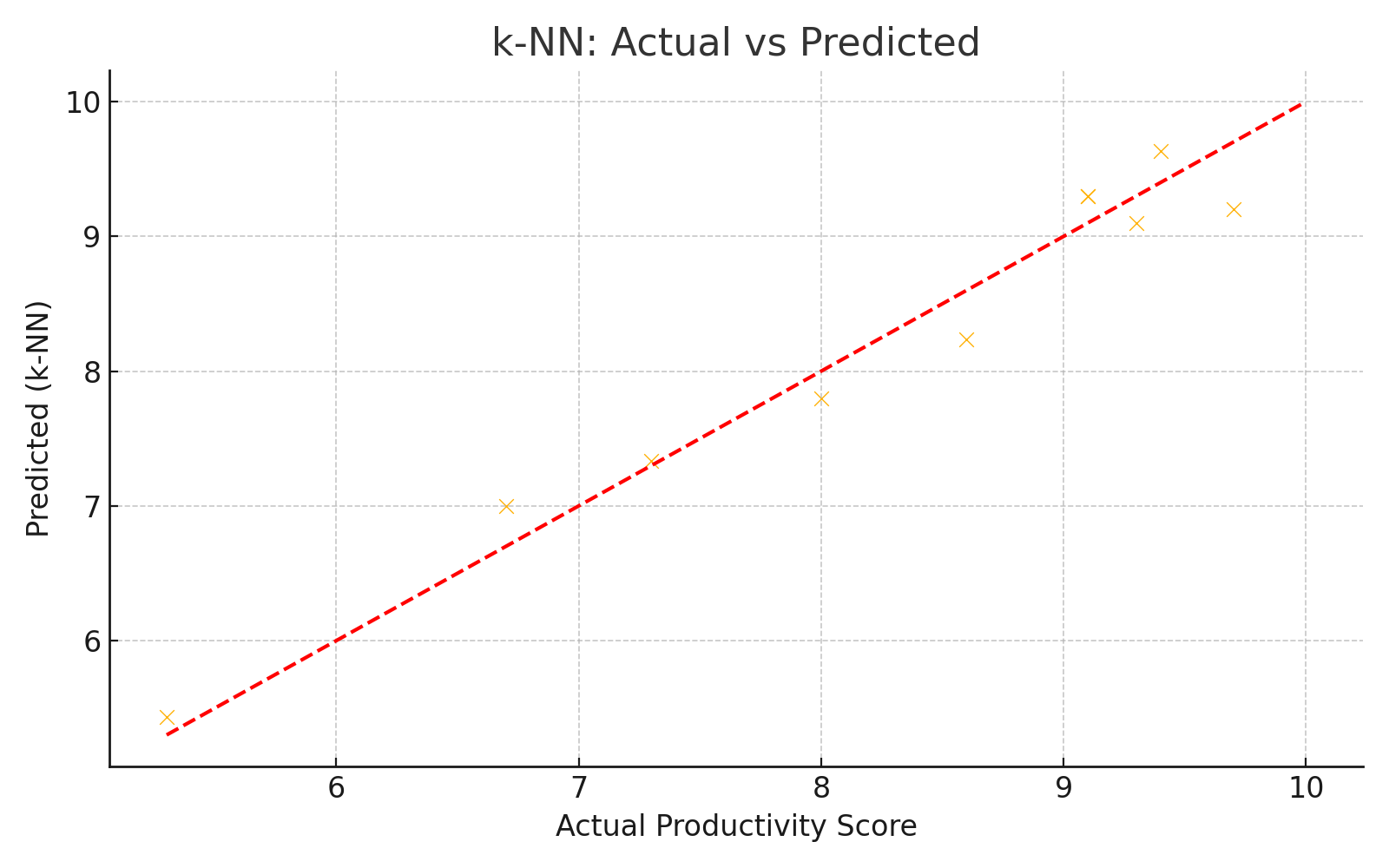
🔍 \*\*Interpretation:\*\*  
- Productive time significantly enhances productivity scores.  
- Total and unbalanced screen time (low productivity ratio) negatively affect performance.  
- High R² shows the model accurately captures key patterns in digital behavior.



Actual vs Predicted Plot: Points near the diagonal show high accuracy.

📘 k-Nearest Neighbors (k-NN) Regression (k=3):

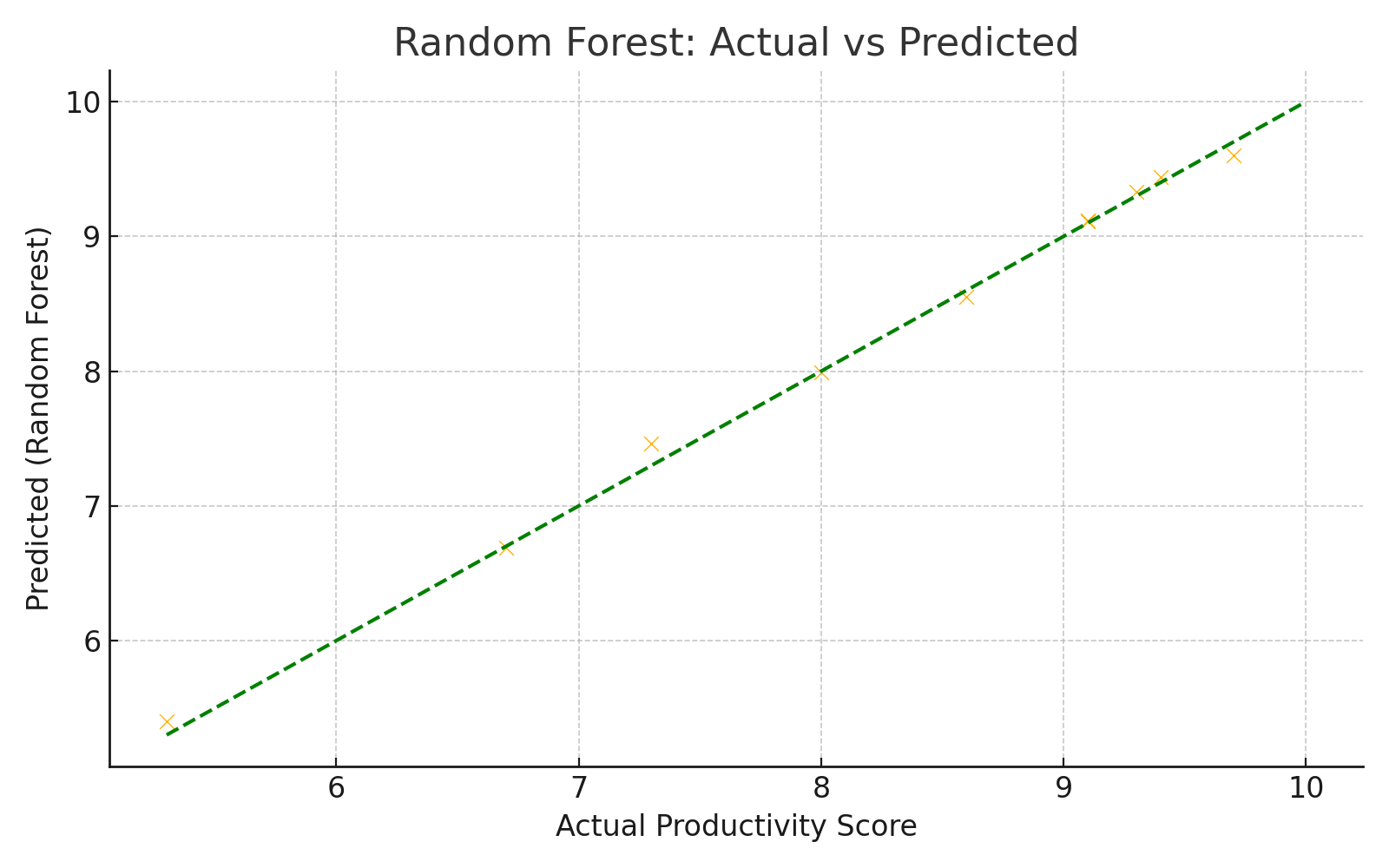
MSE: 0.071  
R² Score: 0.961



k-NN captures local nonlinear relationships well, useful with low-noise datasets.

🌳 Random Forest Regression:

MSE: 0.0051  
R² Score: 0.997



Random Forest ensembles trees and provides very accurate predictions with low overfitting risk.

# RECOMMENDATIONS AND IMPROVEMENTS

- Extend data collection period and include more participants.  
- Track specific app usage and categorize content.  
- Measure objective productivity (e.g., tasks completed).  
- Add contextual factors: stress, deadlines, and workload.

# FILES TO BE SUBMITTED

- Jupyter Notebook (`Screen\_Time\_Analysis.ipynb`)  
- Dataset (`custom\_screen\_time\_data.csv`)  
- Project README (`README2.md`)  
- Final Report (this Word document)