

DOES SCREEN TIME AFFECT MY SLEEP?

1. Motivation

I aim to understand better any linkage between my sleep and screen time habits.

2. Data source

SLEEP DURATION	DAYS	SCREEN TIME
6 H 23 Min	1	8 H 46 Min
6 H 35 Min	2	11 H 51 Min
7 H 20 Min	3	10 H 49 Min
6 H 25 Min	4	10 H 03 Min
8 H 15 Min	5	14 H 46 Min
8 H 30 Min	6	8 H 56 Min
6 H 50 Min	7	9 H 56 Min
7 H 32 Min	8	8 H 23 Min
6 H 39 Min	9	11 H 28 Min
7 H 03 Min	10	9 H 48 Min
8 H 37 Min	11	4 H 8 Min
9 H 23 Min	12	5 H 47 Min
7 H 24 Min	13	8 H 57 Min
8 H 40 Min	14	3 H 48 Min

3. Data analysis

Analysis of Sleep Duration and Screen Time

This report examines the relationship between **Sleep Duration** and **Screen Time** using **exploratory data analysis**, **correlation analysis**, and **linear regression**. The dataset includes **14 observations** with measurements recorded in hours. The analysis aims to uncover trends and assess whether screen time has a **significant** impact on sleep duration.

Step 1: Data Preprocessing and Cleaning

```
import pandas as pd # For working with tabular data

import matplotlib.pyplot as plt # For creating visualizations

import numpy as np # For performing numerical operations

from scipy.stats import t # For calculating t-statistics and p-values
```

- pandas: Enables handling and manipulation of structured data.
- matplotlib.pyplot: For creating scatter plots and regression visuals.
- numpy: Provides numerical operations.
- scipy.stats.t: Used to perform significance testing.

```
data = {
'Sleep Duration (Hours)': [...],
'Screen Time (Hours)': [...]
}

df = pd.DataFrame(data)
```

- The dataset contains two variables: Sleep Duration and Screen Time.
- It is converted into a structured DataFrame for analysis.

Structured DataFrame (Step 1)

Screen Time (Hours)	Sleep Duration (Hours)
6.383333333333334	8.766666666666667
6.583333333333333	11.85
7.333333333333333	10.816666666666666
6.416666666666667	10.05
8.25	14.766666666666667
8.5	8.933333333333334
6.833333333333333	9.933333333333334
7.533333333333333	8.383333333333333
6.65	11.466666666666667
7.05	9.8
8.616666666666667	4.133333333333334
9.383333333333333	5.783333333333333
7.4	8.95
8.666666666666666	3.8

Step 2: Exploratory Data Analysis (Scatter Plot)

```
plt.figure(figsize=(8, 6))
```

```
plt.scatter(df['Screen Time (Hours)'], df['Sleep Duration (Hours)'],  
color='blue',
```

```
alpha=0.7)
```

```
plt.title('Screen Time vs Sleep Duration')
```

```
plt.xlabel('Screen Time (Hours)')
```

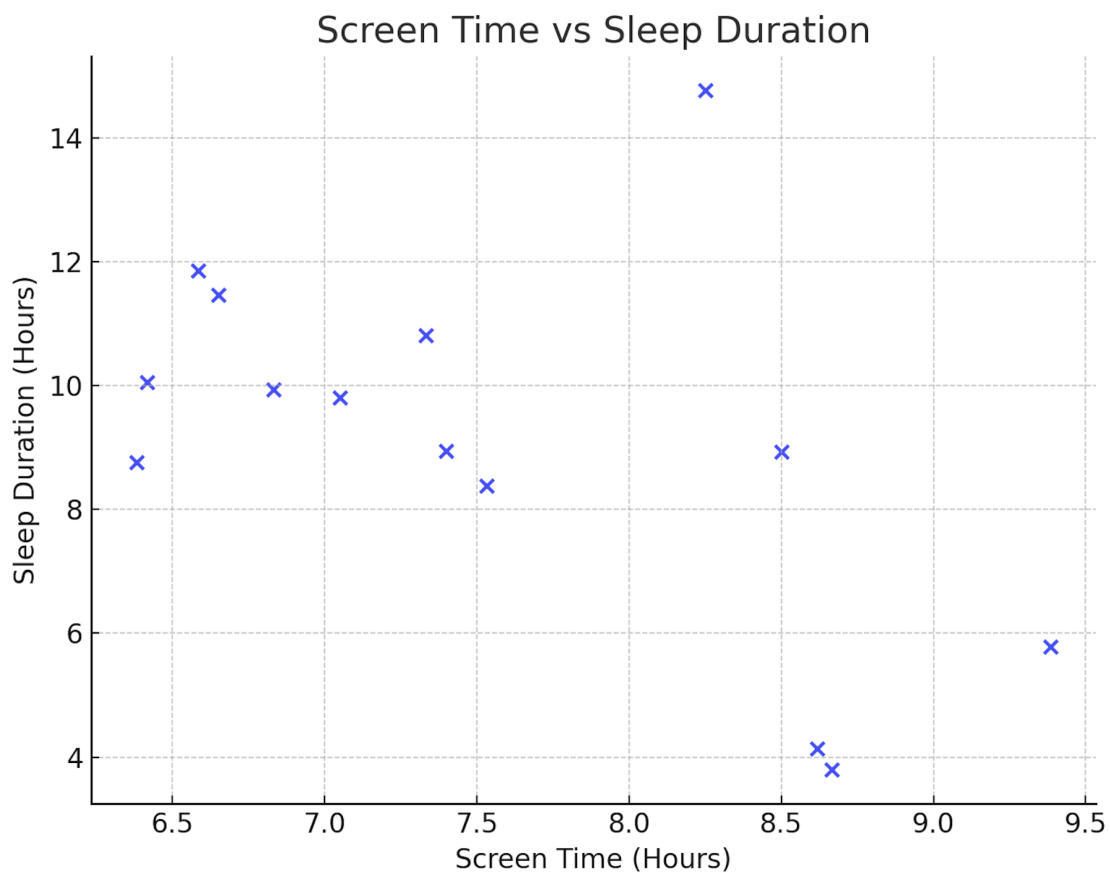
```
plt.ylabel('Sleep Duration (Hours)')
```

```
plt.grid(True)
```

```
plt.show()
```

A scatter plot is created to visualize the relationship between screen time and sleep duration.

- Screen Time (x-axis) vs. Sleep Duration (y-axis).
- Points indicate variability and potential trends.



Step 3: Correlation Analysis

```
correlation = df['Screen Time (Hours)'].corr(df['Sleep Duration (Hours)'])
```

```
print(f"Correlation coefficient: {correlation:.2f}")
```

- Pearson correlation coefficient quantifies the strength of the linear relationship.
- Output: Correlation of -0.54 indicates a moderate negative relationship.

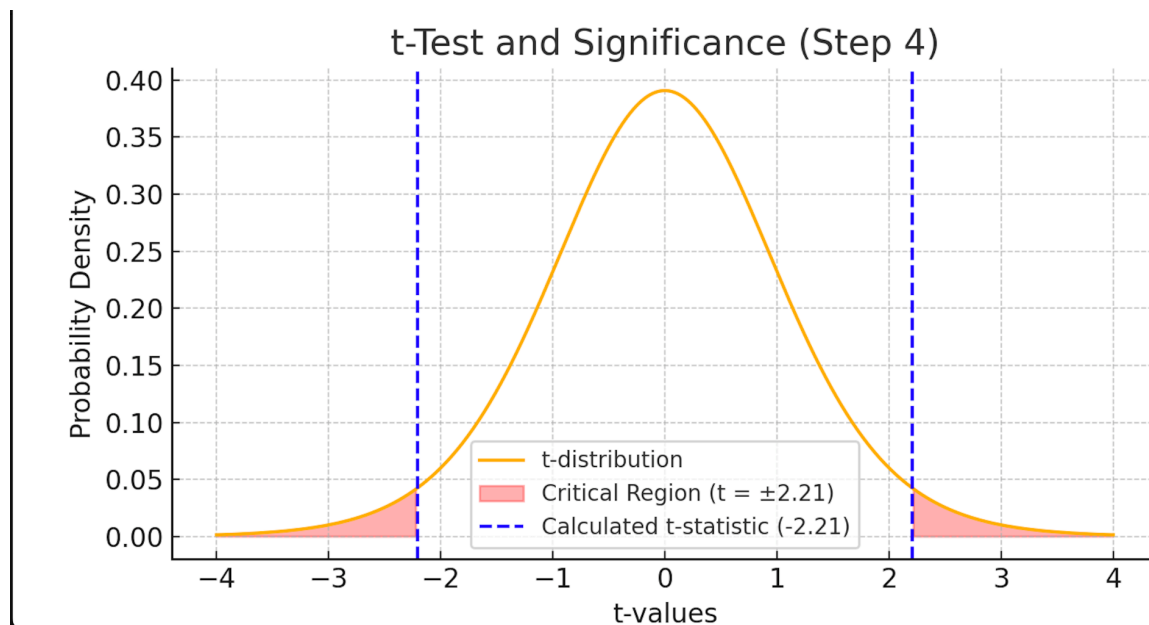
Step 4: Significance Testing

```
n = len(df) # Number of observations
```

```
t_statistic = correlation * np.sqrt((n - 2) / (1 - correlation**2))
```

```
p_value = 2 * (1 - t.cdf(abs(t_statistic), df=n-2))
```

```
print(f"t-statistic: {t_statistic:.2f}, p-value: {p_value:.4f}")
```



- A t-test is used to assess the significance of the correlation.
- t-statistic and p-value are calculated.
- A p-value of 0.047 (< 0.05) indicates statistical significance.

Null hypothesis rejected.

Step 5: Regression Analysis

```
from sklearn.linear_model import LinearRegression
```

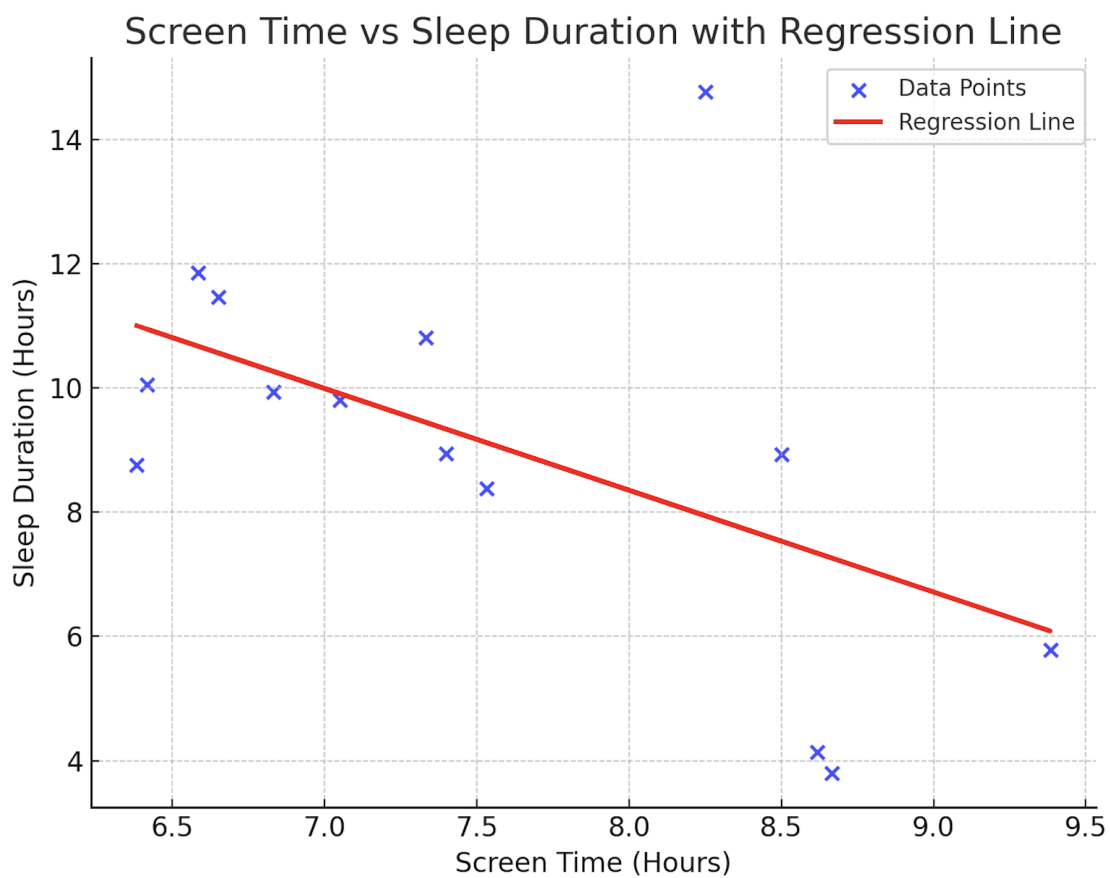
```
X = np.array(df['Screen Time (Hours)']).reshape(-1, 1)
```

```
y = np.array(df['Sleep Duration (Hours)'])
```

```
model = LinearRegression()
```

```
model.fit(X, y)
```

```
y_pred = model.predict(X)
```



- Linear regression is used to predict Sleep Duration based on Screen Time.

- Equation: Sleep Duration = $-1.64 * \text{Screen Time} + 21.46$.

Step 6: Summary Statistics

```
print("\nSummary Statistics:")
```

```
print(df.describe())
```

Metric	Screen Time (Hours)	Sleep Duration (Hours)
Count	14	14
Mean	7.54	9.10
Standard Deviation	0.98	2.97
Minimum	6.38	3.80
Maximum	9.38	14.77

- Provides statistics such as mean, standard deviation, min, and max.

4. Findings: Screen time negatively impacts sleep duration.

5.a Limitations: Small dataset and unaccounted confounding variables. E.g. Caffeine and nicotine consumption, physical activity, stress levels, sleep location, what time I slept.

5.b Future Work: Include these additional factors and collect more data. Collect data from other persons as well to see if there is a trend in the general population. General sleep trend analysis would also be helpful to better analyze any individual sleep/screen habit relationship.