

# EFFECT OF SLEEP ON EFFICIENCY

- Sinem Sönmez 30854



# MY DATA SET

Screen Usage of Mac, Ipad,  
Iphone





# HYPOTHESIS OF THIS PROJECT

- "There is an optimal wake-up time for each individual that maximizes their productivity, and deviating from this time, whether waking up too early or too late, leads to reduced productivity throughout the day."

date	hour	minute
2024-09-30	11	30
2024-10-01	6	6
2024-10-02	8	1
2024-10-03	7	39
2024-10-04	11	40
2024-10-05	5	29
2024-10-06	6	21
2024-10-07	6	1
2024-10-08	5	26
2024-10-09	5	1
2024-10-10	9	44
2024-10-11	5	53
2024-10-12	6	12
2024-10-13	9	24
2024-10-14	6	1
2024-10-15	4	3

Firstly, I scraped my screen usage data for my phone, tablet, and laptop using a third party apps and scripts.

I tagged apps as good, bad, or neutral. To calculate efficient time, I divided good app usage to total usage

date	total_usage	good_app_usage	efficient_time
2024-09-30	1966	1481	0.7533062054933876
2024-10-01	13934	2519	0.1807808238840247
2024-10-02	6099	0	0.0
2024-10-03	12905	1577	0.12220069740410694
2024-10-04	3064	76	0.024804177545691905
2024-10-05	180	0	0.0
2024-10-06	3205	198	0.061778471138845555
2024-10-07	11059	7178	0.6490641106790849
2024-10-08	18407	3631	0.19726191122942358
2024-10-09	28035	2348	0.08375245229177813
2024-10-10	14997	8596	0.5731813029272521
2024-10-11	11244	540	0.048025613660619
2024-10-12	7023	0	0.0
2024-10-13	12228	238	0.019463526333006215
2024-10-14	35765	22216	0.6211659443590102
2024-10-15	32291	4705	0.1457062339351522
2024-10-16	8962	112	0.0124972104440973
2024-10-17	5711	920	0.16109262826125023

These are  
good app  
examples for  
mac, iPad and  
iPhone

net.wochuywdm.collanote496	10	2024-09-30 12:07:59	2024-09-30 12:08:09	2024-09-30 12:08:09	lpad	1
net.wochuywdm.collanote496	18	2024-09-30 12:13:24	2024-09-30 12:13:42	2024-09-30 12:13:42	lpad	1
net.wochuywdm.collanote496	1043	2024-10-01 06:43:27	2024-10-01 07:00:50	2024-10-01 07:00:50	lpad	1
net.wochuywdm.collanote496	10	2024-10-01 07:17:02	2024-10-01 07:17:12	2024-10-01 07:17:12	lpad	1
net.wochuywdm.collanote496	10	2024-10-01 07:46:13	2024-10-01 07:46:23	2024-10-01 07:46:23	lpad	1
net.wochuywdm.collanote496	10	2024-10-01 07:47:01	2024-10-01 07:47:11	2024-10-01 07:47:11	lpad	1

com.microsoft.skype.teams	1	2024-10-15 04:03:19	2024-10-15 04:03:20	2024-10-15 04:03:20	IPhone	1
com.microsoft.skype.teams	11	2024-10-22 13:51:54	2024-10-22 13:52:05	2024-10-22 13:52:05	IPhone	1
com.microsoft.skype.teams	8	2024-10-23 09:54:18	2024-10-23 09:54:26	2024-10-23 09:54:26	IPhone	1
com.microsoft.skype.teams	9	2024-11-12 11:27:16	2024-11-12 11:27:25	2024-11-12 11:27:25	IPhone	1
com.microsoft.skype.teams	4	2024-11-12 11:27:49	2024-11-12 11:27:53	2024-11-12 11:27:53	IPhone	1

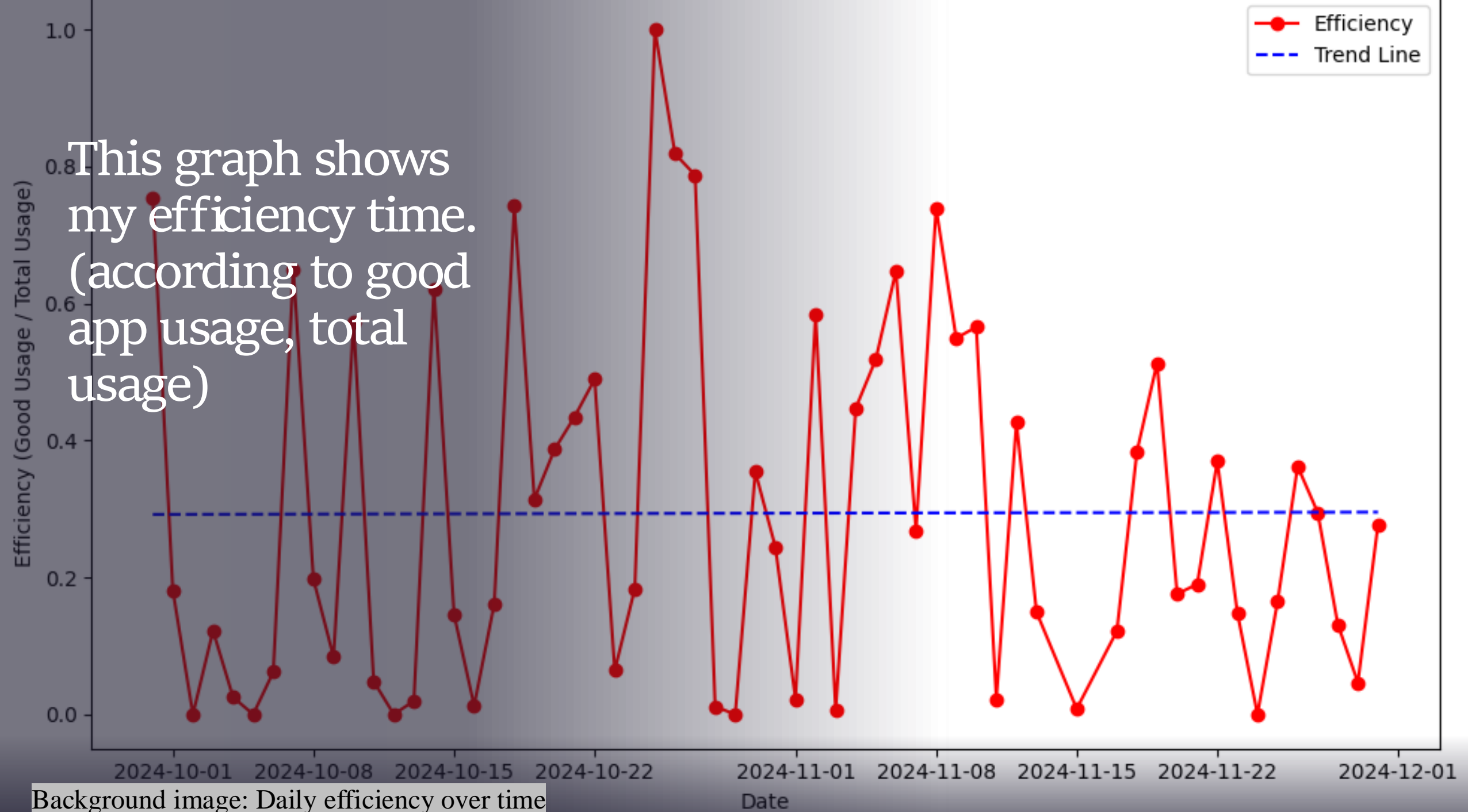
com.microsoft.Word	11	2024-11-08 12:48:27	2024-11-08 12:48:38	2024-11-08 12:48:38	Mac	1
com.microsoft.Word	1	2024-11-08 12:48:38	2024-11-08 12:48:39	2024-11-08 12:48:39	Mac	1
com.microsoft.Word	14	2024-11-08 12:49:34	2024-11-08 12:49:48	2024-11-08 12:49:48	Mac	1
com.microsoft.Word	1	2024-11-08 12:54:58	2024-11-08 12:54:59	2024-11-08 12:54:59	Mac	1
com.microsoft.Word	2	2024-11-08 12:55:14	2024-11-08 12:55:16	2024-11-08 12:55:16	Mac	1
com.microsoft.Word	3	2024-11-08 12:55:18	2024-11-08 12:55:21	2024-11-08 12:55:21	Mac	1



After reviewing the data I had, I saw that it was not enough and added another dataset.

- Apple health application

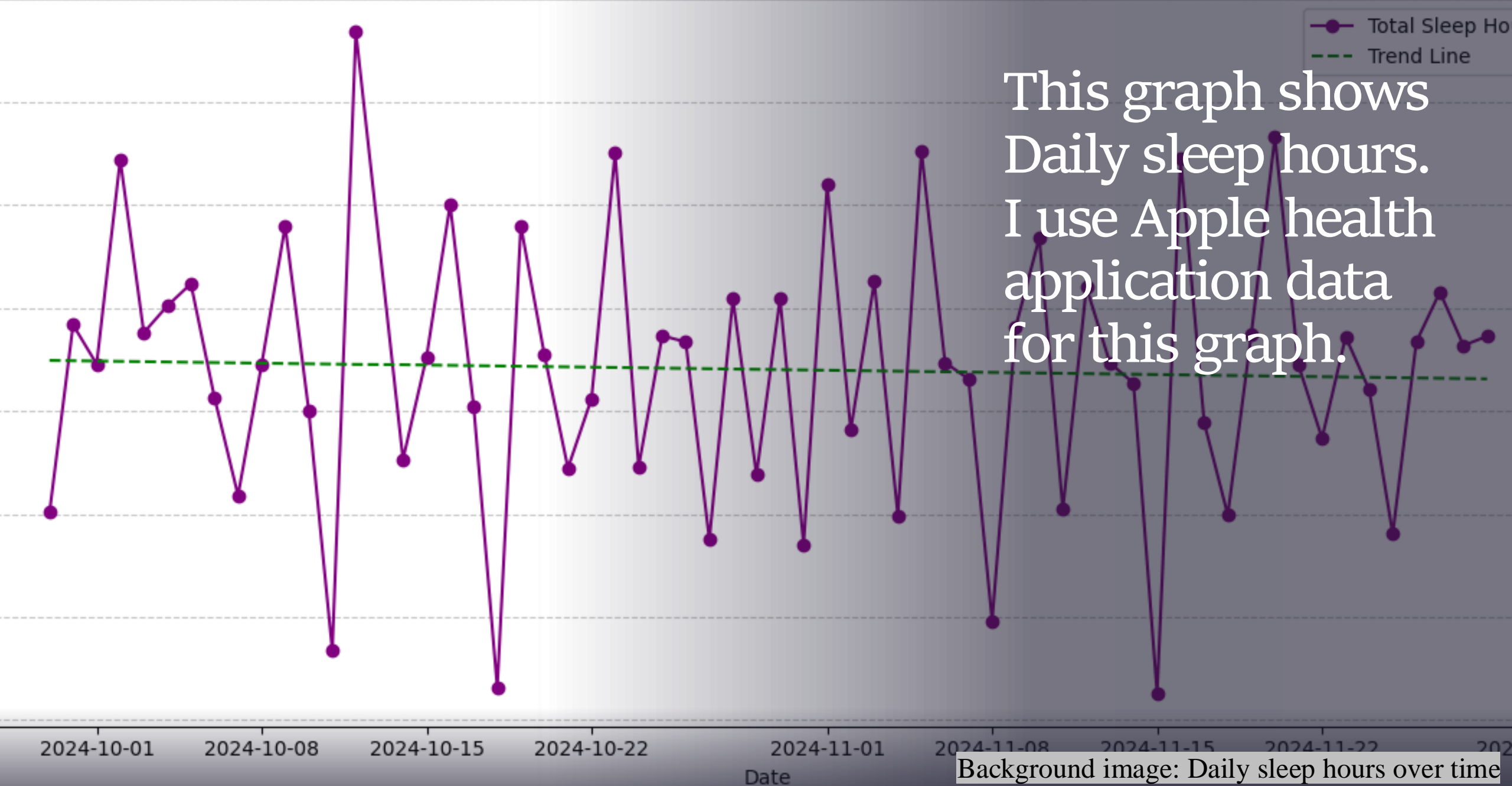
I use wake-up time and sleeping time for this project.



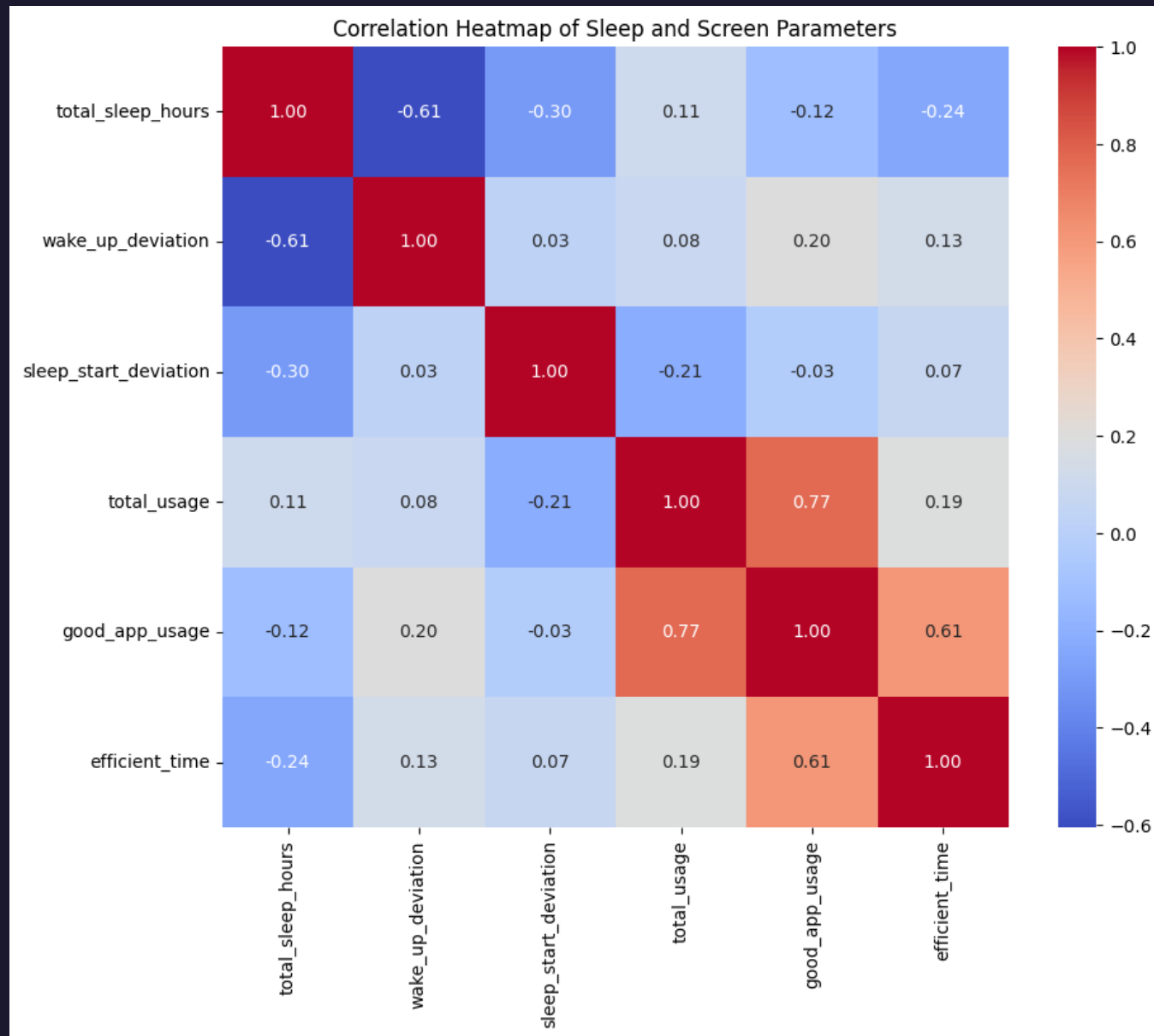
Background image: Daily efficiency over time



# Daily Sleep Hours Over Time

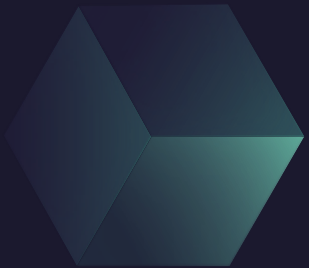


Background image: Daily sleep hours over time



I created this heatmap to observe any meaningful correlations between sleep data and screen usage data.





After observing the heatmap, I did not see any meaningful correlation. For further exploration, I use this data in random forest regression.

Before using random forest regression, I prepared the data by shuffling and splitting it into train and test parts.

```
-
3 df_shuffled = shuffle(df, random_state=42)
4
5 X = df[["wake_up_deviation", "sleep_start_deviation", "total_sleep_hours", "total_usage"]].astype(float)
6 y = df[["efficient_time"]].astype(float)
7
8 print(df.head())
9
10 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

# Random forest regression part

```
[ ] 1 regressor = RandomForestRegressor(random_state=42)
    2
    3 regressor.fit(X_train, y_train)
```



Finally, I used test data and a regression model to predict values corresponding to the test data. Then, compare these values to the actual data.



## Root Mean Squared Error (RMSE):

This value should ideally be as small as possible for a well-performing model.


## R-squared ( $R^2$ ):

A negative  $R^2$  typically means:

- The model is poorly trained.
- The model could be underfitting or inappropriate for the data.

```
1 y_pred = regressor.predict(X_test)
2
3 mse = mean_squared_error(y_test, y_pred)
4 rmse = np.sqrt(mse)
5 r2 = r2_score(y_test, y_pred)
6
7 print(f"Root Mean Squared Error: {rmse}")
8 print(f"R-squared: {r2}")
9
```

```
⇒ Root Mean Squared Error: 0.3367844487907777
   R-squared: -0.34802801942628525
```



In conclusion, the null hypothesis is rejected. The correlation between sleep pattern and day productivity was not seen. So, It can be said sleep pattern and day productivity do not correlate.



Thank you for  
your time.

