CS210 INTRODUCTION TO

DATA SCIENCE PROJECT:

DATA ANALYSIS OF SLEEP

AND PHYSICAL ACTIVITY

HABITS

ALPER ÇAKAR

27884

**1-Motivation**

This project inspired by curiosity to comprehend my sleep and exercise habits. The importance of getting enough sleep and exercise cannot be emphasized in our day to day existence. These two factors are crucial in defining someone’s quality of life as they affect both physical endurance and mental health. I had two reasons to conduct this analysis. Firstly, I wanted more in depth data driven information about my well- being. I tried to find patterns and relationships between my sleep and physical activity habits. By this way, I wanted to have an idea about how well I slept and exercised. Secondly, I wanted to see what I can improve in terms of my habits that directly influence my health. Therefore, it can be concluded that motivation of this project was to analyze data and draw conclusions from this data about my sleep and exercise patterns.

**2-Data Source**

Data source of this project was health application on my mobile phone that tracks various of metrics. This data was transformed to excel by me manually. Data consists of 6 columns of information: days, steps taken, sleep duration, max stride length, min stride length and walking asymmetry. This data includes 90 days of information between 11th of October and 8th of January. First 30 days are referred as first month, second 30 days are referred as next month and last 30 days are referred as last month.

**3-Data Analysis**

Various of techniques was used during this project to conduct broad data analysis. Descriptive statistics like mean, std, minimum, maximum was found to understand the data. Moreover, visuals like scatter plot, box plots and histograms was used to determine the relationships, distributions and patterns in the data. Regression and correlation analysis was also conducted to see the relationships between different measurements more clearly

**4-Findings**

First analysis conducted was examination of sleep duration. Firstly descriptive statististics were found to see if there is any difference between the three months. Results were:

First Month:

Total Sleep Duration: 9905

Average Sleep Duration: 330.1666666666667

Std Sleep Duration: 91.45833169690779

Next Month:

Total Sleep Duration: 9907

Average Sleep Duration: 330.23333333333335

Std Sleep Duration: 71.48555838399294

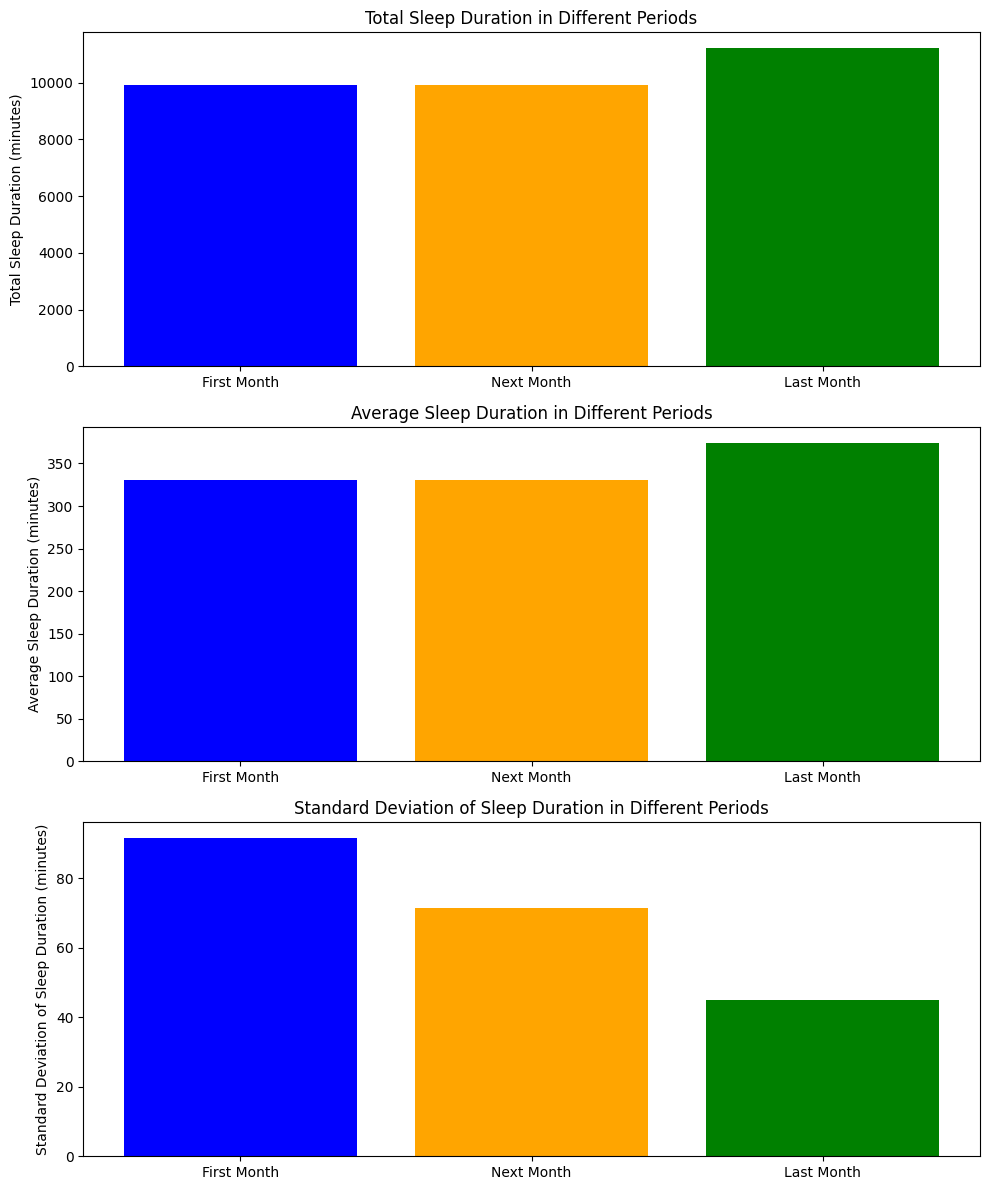
Last Month:

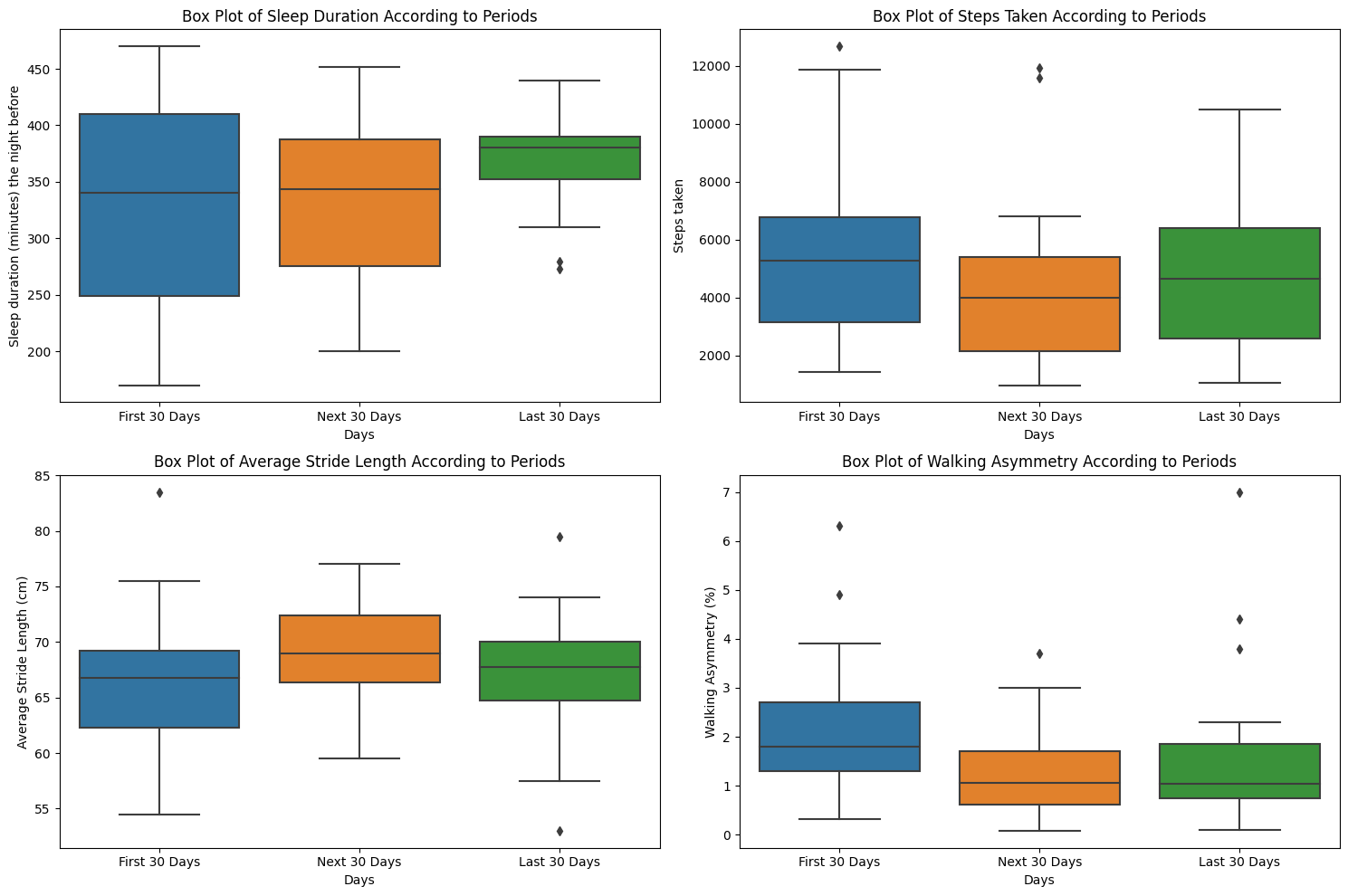
Total Sleep Duration: 11214

Average Sleep Duration: 373.8

Std Sleep Duration: 44.8856017156873

As it can be interpreted from these results, there wasn’t much difference between the months in terms of sleep duration. First two months are almost identical and the difference of last month from other two is less than 1 hour. Despite that standard deviation seems to be declining and last month has higher sleep duration which may indicate that my sleeping habits become healthier, during all three months sleep duration show the sign of unhealthy sleeping habit as experts suggest minimum 420 minutes of sleep a day. Related histograms and boxplot can be seen below:





Next analysis was examination of steps taken. Same as sleep duration, firstly descritive statistic analysis was conducted and after that boxplot and histograms were analyzed. Results can be seen below:

First Month:

Total Steps: 167411

Average Steps: 5580.366666666667

Std Steps: 3133.860101660827

Next Month:

Total Steps: 123413

Average Steps: 4113.766666666666

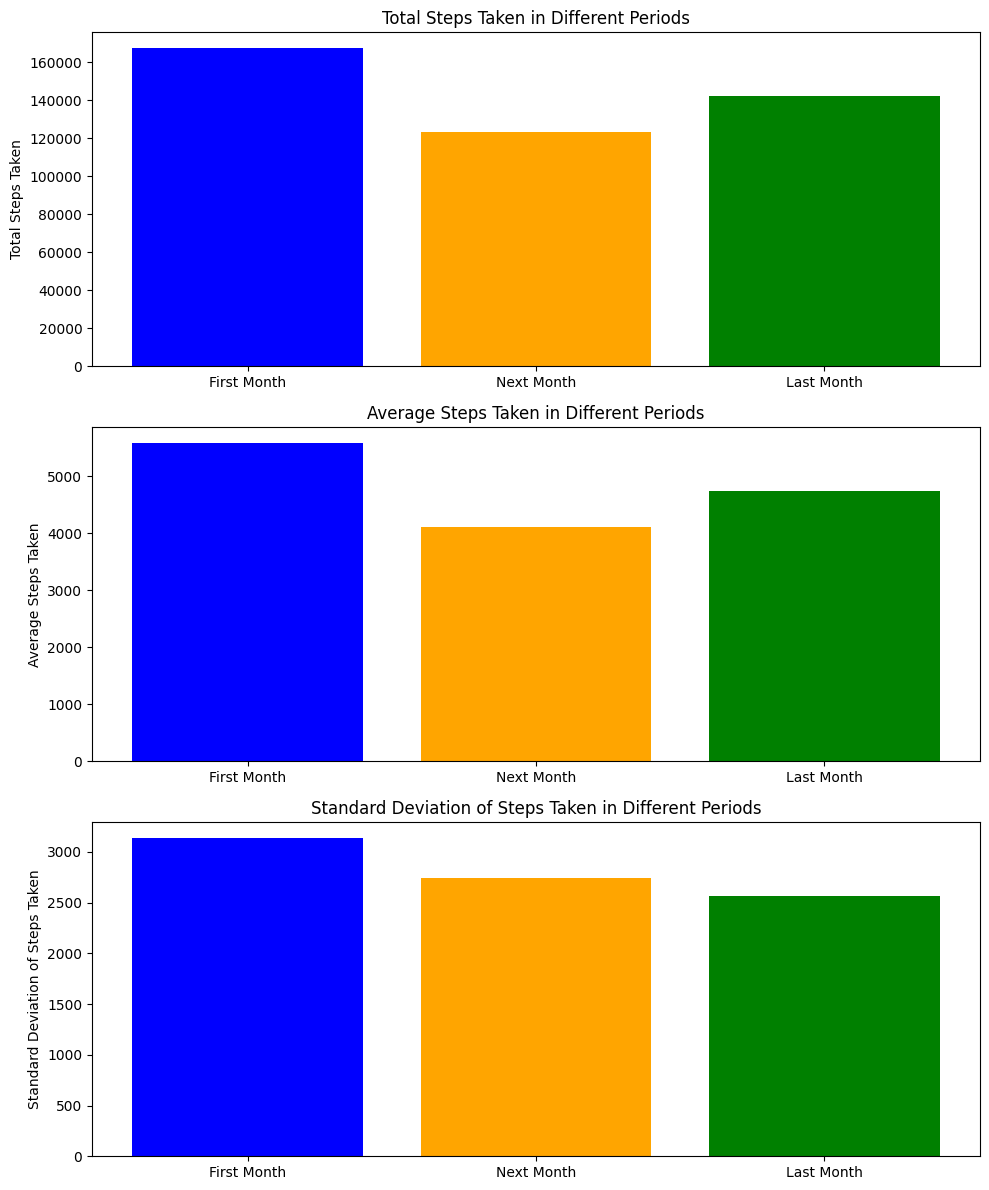
Std Steps: 2742.4013851862137

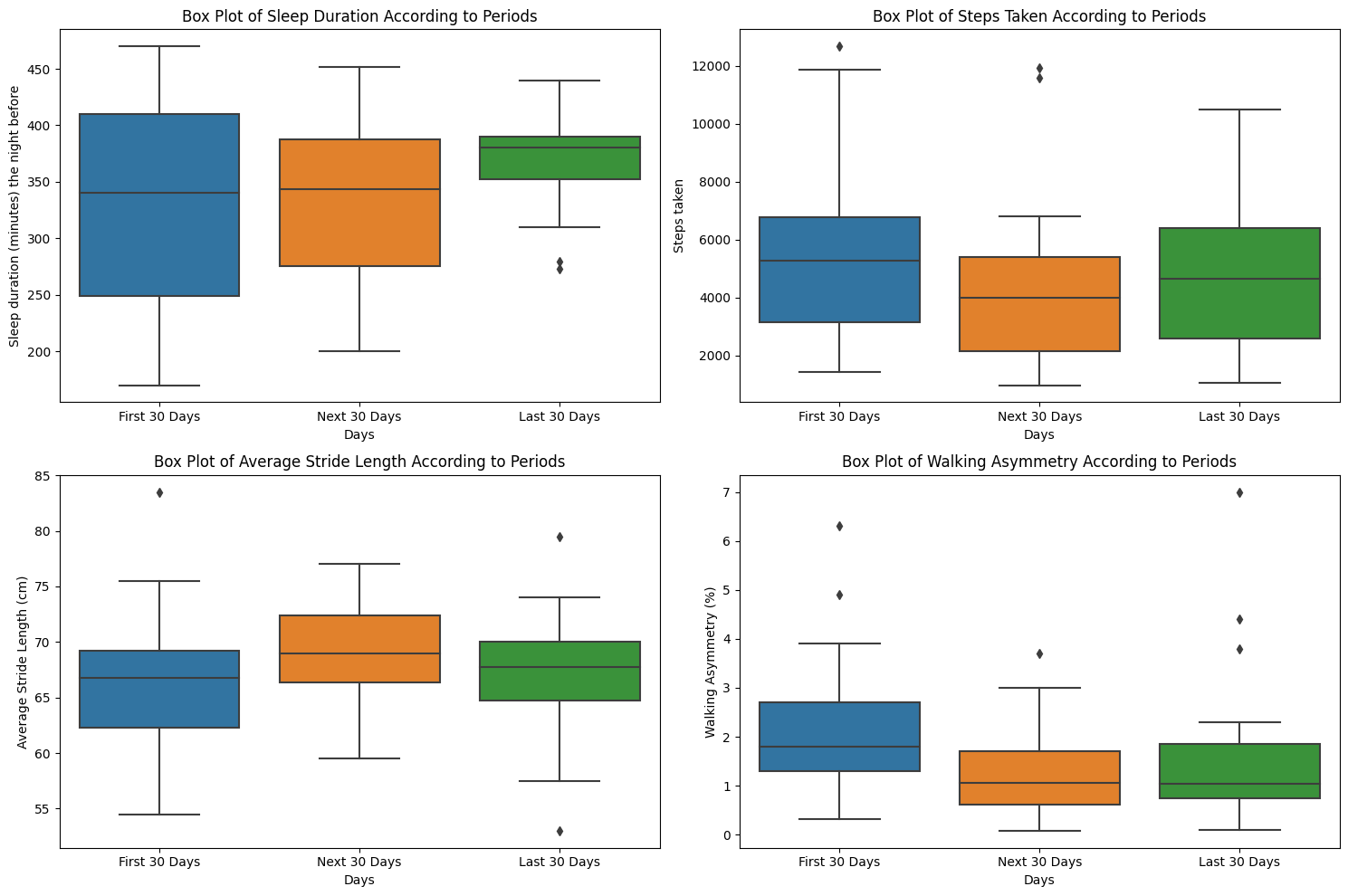
Last Month:

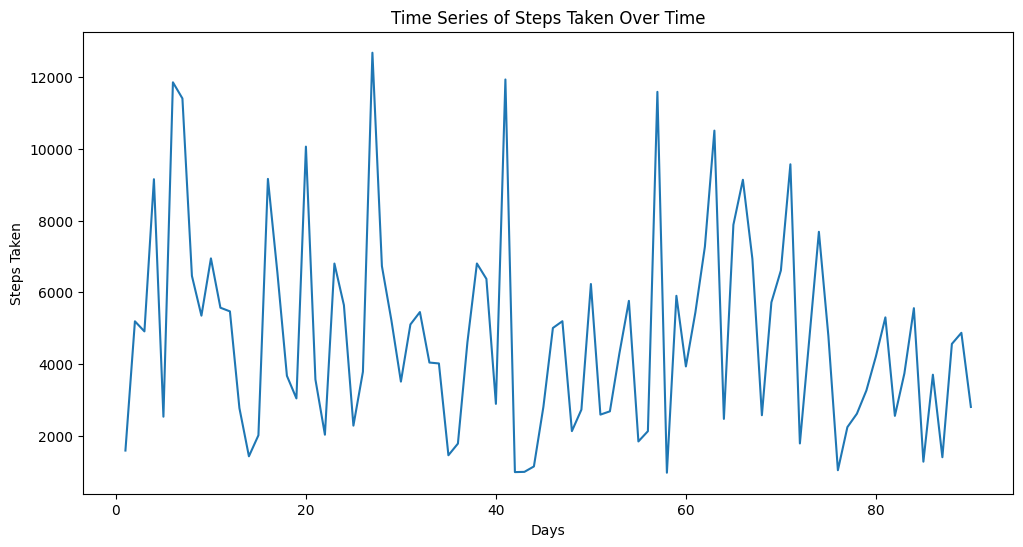
Total Steps: 142277

Average Steps: 4742.566666666667

Std Steps: 2565.982163016637



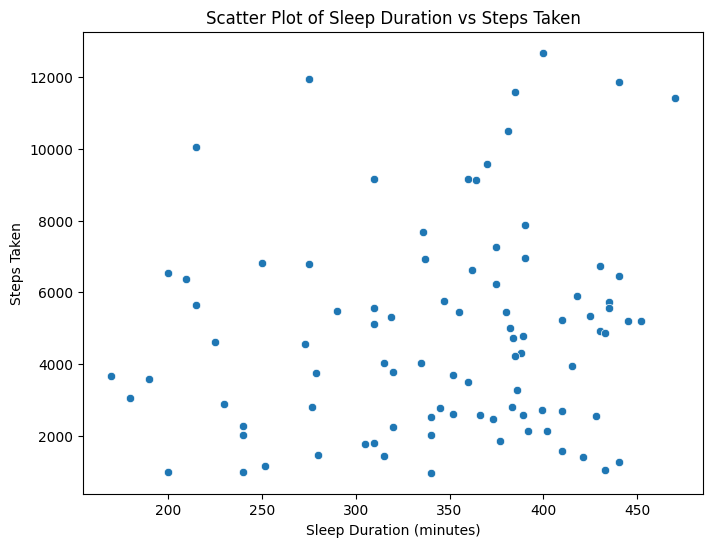




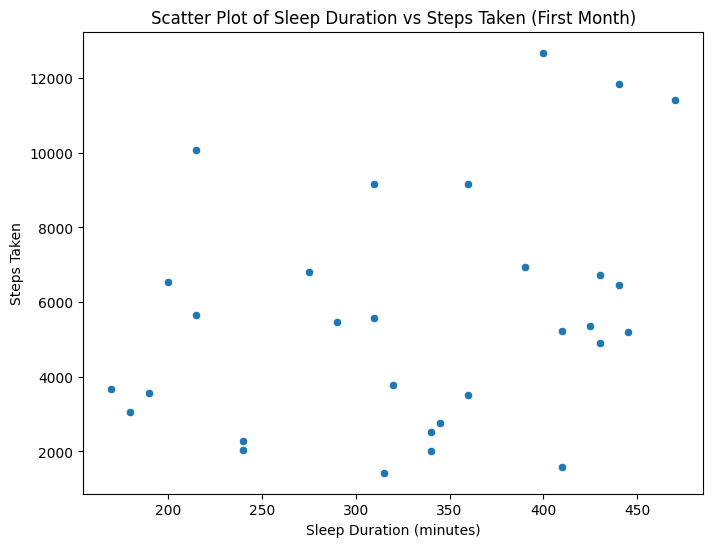
From the results, we can conclude that first month is significantly different than the second month and slightly different than third month. During the first month, my physical activity seems to be almost healthy as experts suggest 6000 steps a day. However other 2 months, my exercise level is not sufficient.

Next, another conducted analysis was the relationship between sleep duration and steps taken. To see if there is a correlation, scatter plot was printed and correlation coefficient was calculated for 3 months separately and overall. Results can be seen below:

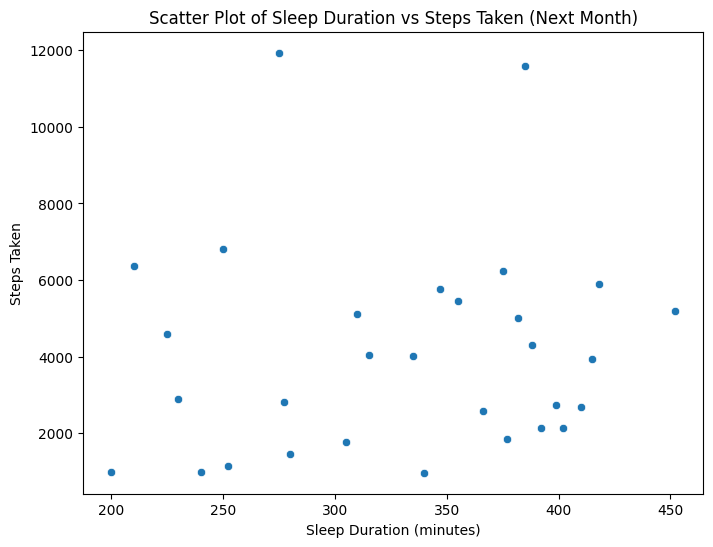
Correlation Coefficient between Sleep Duration and Steps Taken: 0.14373885180718465



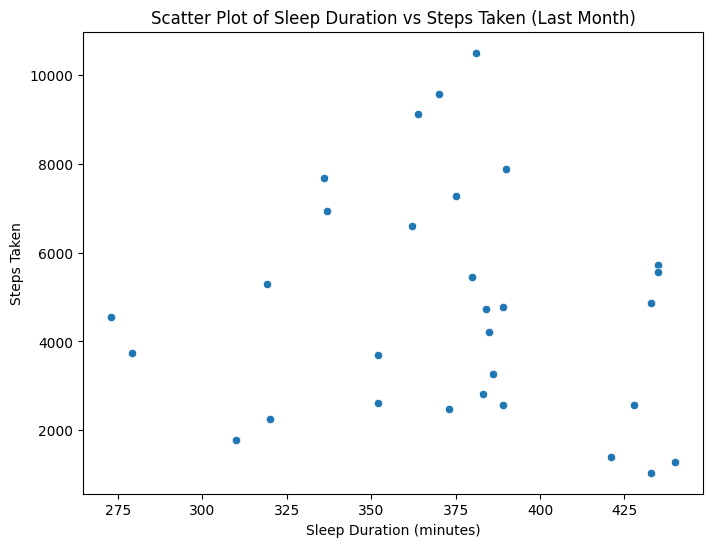
Correlation Coefficient between Sleep Duration and Steps Taken (First Month): 0.31680659280426776



Correlation Coefficient between Sleep Duration and Steps Taken (Next Month): 0.09566622979508294

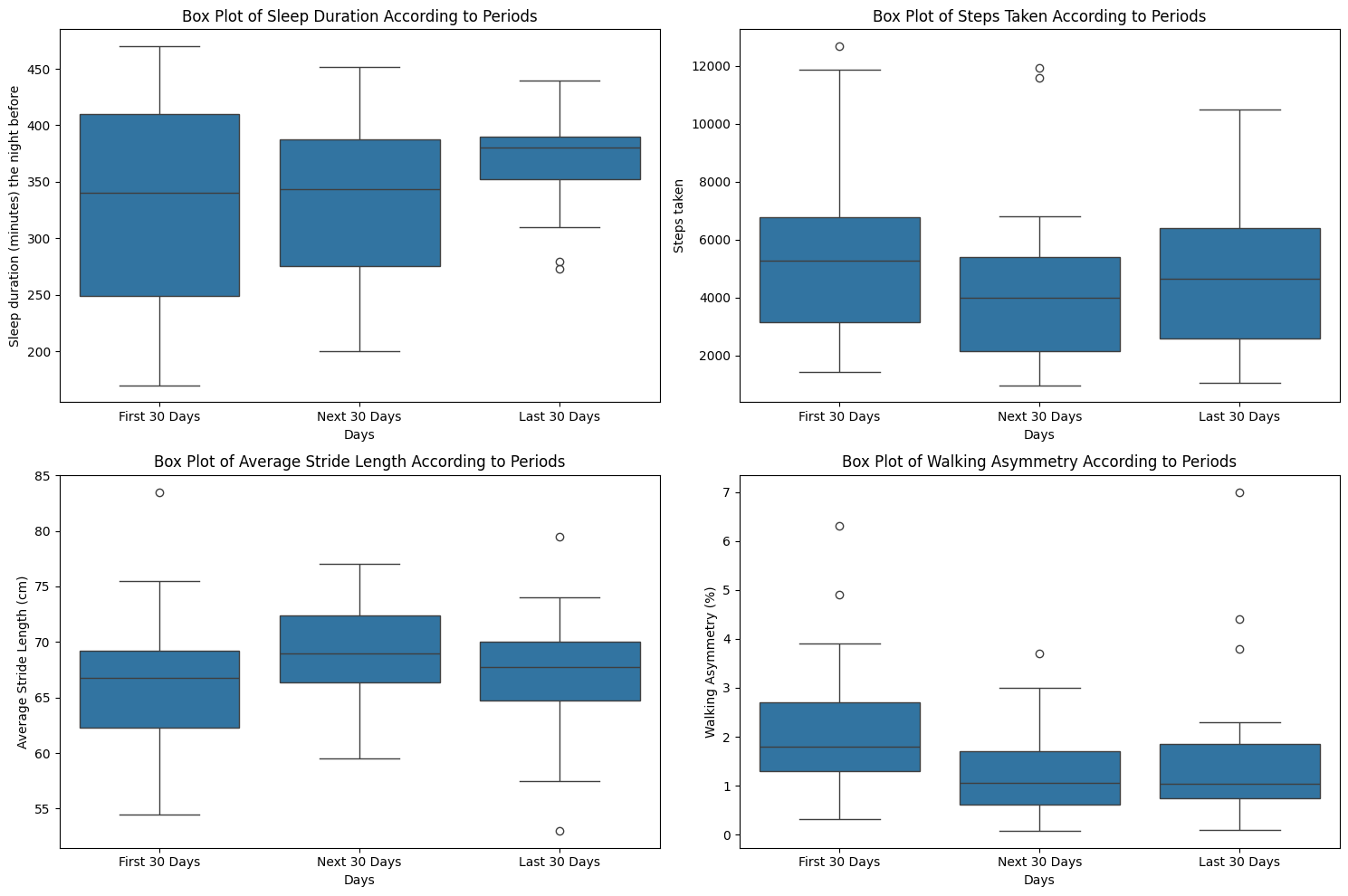


Correlation Coefficient between Sleep Duration and Steps Taken (Last Month): -0.11394128606167643

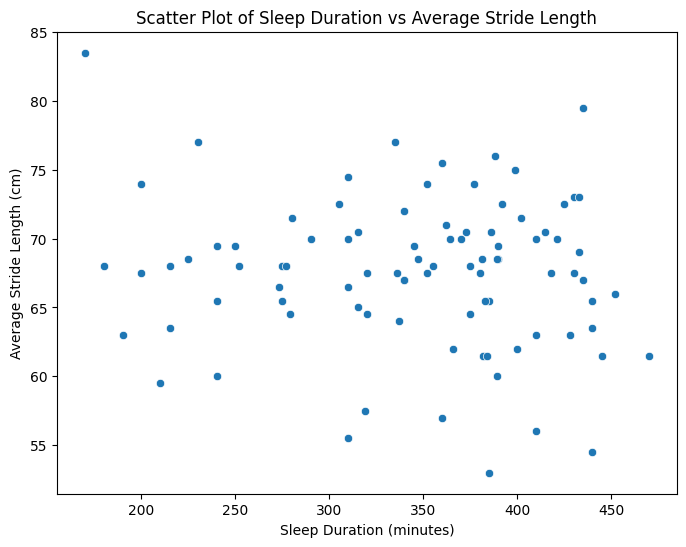


We can conclude that there is no correlation between sleep duration and steps taken as the correlation coefficient values are very low

After these examinations, average stride length was calculated by taking the average of mix and min stride length values. Firstly, box plot was printed to have an idea about the distribution. Then, correlation analysis was conducted to see if there is a relationship between sleep duration and average stride length. Results are below:



Correlation Coefficient between Sleep Duration and Average Stride Length: -0.10468392104759058



Both from the scatter plot and correlation coefficient value we can state that there is no correlation between the two measurements. After that, walking asymmetry data was examined. Because that walking asymmetry value for every individual can be very different, according to some experts what is more important is if there is a increasing trend. This may be an indicator for different diseases as experts state. So, to examine the trend of walking asymmetry, time series of walking asymmery was shown and descriptive statistics of three months seperately was calculated. Also scatter plot was examined to see if there is a relationship between sleep duration and walking asymmetry. Results are shown below:

Statistics for Walking Asymmetry - First 30 Days:

count 30.000000

mean 2.084000

std 1.328537

min 0.320000

25% 1.300000

50% 1.800000

75% 2.700000

max 6.300000

Name: Walking Asymmetry (%), dtype: float64

Statistics for Walking Asymmetry - Next 30 Days:

count 30.000000

mean 1.219667

std 0.851305

min 0.080000

25% 0.620000

50% 1.055000

75% 1.700000

max 3.700000

Name: Walking Asymmetry (%), dtype: float64

Statistics for Walking Asymmetry - Last 30 Days:

count 30.000000

mean 1.497000

std 1.408668

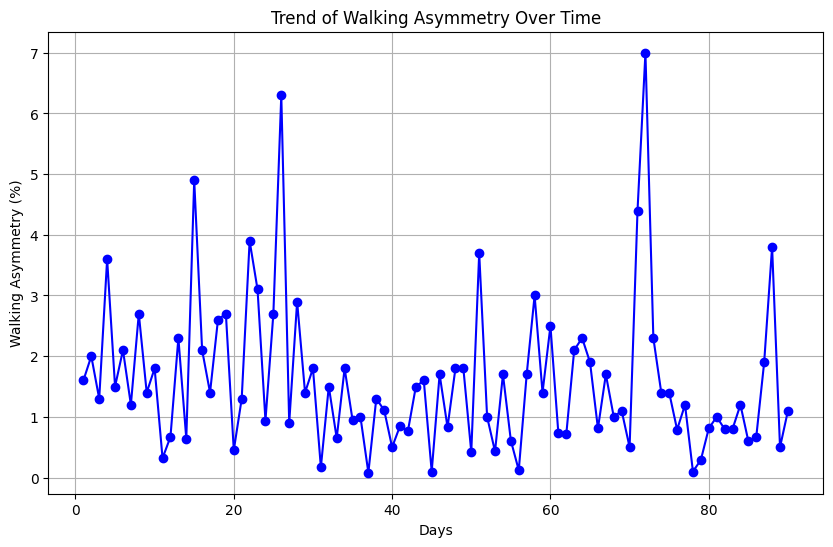
min 0.100000

25% 0.745000

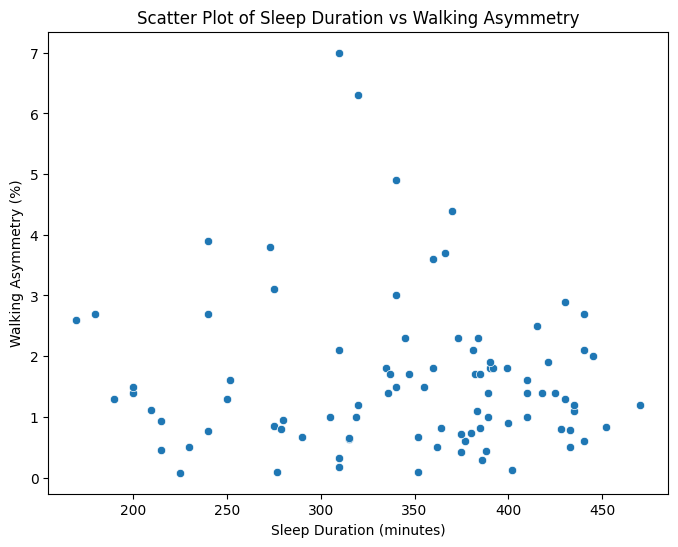
50% 1.050000

75% 1.850000

max 7.000000



Correlation Coefficient between Sleep Duration and Walking Asymmetry: -0.05964786654164678



Statistics show that walking asymmetry significantly decreased from first month to second month, and it slightly increased from second month to third. Most of that slight increase probably resulted from the value around 7%, walking asymmetry during day 72. Scatter plot and correlation coefficient showed no sign of relationship. Also when we consider the change in walking asymmetry, there is no sign of a disease statistically.

Another analysis was conducted to see if it is possible to guess the sleep duration based on other measurements. Linear regression was used for this analysis. Data was randomly split into two sets, 0.2 test data and 0.8 training data. As a result of this examination, mean squared error was calculated 5077.30970724438. Considering sqaure root of it is 71.25, we can say that model gives a moderate guess, not very close but also not very wrong. Lastly, days were seperated into weekdays and weekends, to see if there is a difference between average of measurements. Results are:

Day Type Steps Taken Sleep Duration The Night Before

Weekday 45.0625 4784.0 349.104167

Weekend 46.0000 4844.5 339.738095

Max Stride Length (cm) Min Stride Length (cm) \

Day Type

Weekday 83.416667 51.812500

Weekend 82.119048 52.785714

Walking Asymmetry (%) Average Stride Length (cm)

Day Type

Weekday 1.782292 67.614583

Weekend 1.392143 67.452381

We can state that there is no significant difference between values as measurements are very similar.

Consequently, we can conclude that sleep duration has no effect on my physical activity. My step count doesn’t seem far away from the target 6000 but my sleep habits seem to be the one measurements that needs to change as they seem to be very low compared to right amount of sleep. Also walking asymmetry data suggest normal values and no sign of a disease. I can say that as a result of this project, I was able to be more familiar with my sleeping and exercise habits.

**5-Limitations and Future Work**

One of the limitations of this project was that iphone’s reliability while measuring these data. Some people report that iphone’s location on our body while we are walking can change the walking asymmetry or stride length. Another limitation is that these measurements do not include the times that I lift weights or play basketball as a physical activity because I do not have my phone on me at these times. Also during a school term, my activity level is mostly based on my work load as sometimes I skip exercising to study or do homework so some of the results can be misleading. Future work might be that these measurements can be recorded for a longer time than 3 months, and also at a time that my work load doesn’t change much like during holiday maybe. This way project can have more reliable results.

**6-Resources**

<https://tr.mashable.com/saglik-zindelik/12160/saglikli-bir-yasam-icin-gunde-kac-adim-atmamiz-gerekiyor-10-bin-degil#:~:text=Bir%20grup%20ara%C5%9Ft%C4%B1rmac%C4%B1%2C%20d%C3%B6rt%20k%C4%B1tada,bin%20ad%C4%B1m%C4%B1n%20kafi%20olabilece%C4%9Fini%20g%C3%B6steriyor>.

<https://www.indyturk.com/node/485546/haber/osman-m%C3%BCft%C3%BCo%C4%9Flu-g%C3%BCnde-ka%C3%A7-ad%C4%B1m>

<https://www.apple.com/healthcare/docs/site/Measuring_Walking_Quality_Through_iPhone_Mobility_Metrics.pdf>

<https://www.wellandgood.com/iphone-gait-analysis/#:~:text=The%20average%20walking%20asymmetry%20for,about%20it%2C%22%20Glick%20says>.

<https://www.bbc.com/turkce/haberler/2015/02/150213_uyku_vakti_gch>