Cs210 Project Report

Hanifi Talha Ozan

27818

Data Analaysis of my Apple Health Data

I started this project because I wanted to learn more about my health using the data from my Apple Health app. Nowadays, we have so much health information from our phones and watches, but we don't always use it to help us. I thought it would be a good idea to look closely at this information to see what it tells me about my daily activities, like walking and exercising. My main goal is to find out if this data can help me live a healthier life. For example, by looking at how many steps I take each day or how fast I walk, I can learn what I need to do to be more active and healthy. This project is important to me because it's about my health and using data in a smart way. I believe that the information from Apple Health can help me make better choices every day. It's like having a guide that shows me how to be healthier based on what the numbers say. The data for this project was sourced directly from my iPhone, utilizing the Apple Health app, which is a built-in feature of the iOS system. Apple Health comprehensively tracks a variety of health and fitness metrics, automatically collecting data from the phone's sensors and any connected wearable devices, like an Apple Watch.

Data Preparation:

I started this project by getting my health data ready for analysis. This data came from my iPhone's Apple Health app. The app keeps track of lots of health stuff every day. To make sense of this data, I first had to organize it. I looked at different kinds of data like how many steps I took, how many stairs I climbed, how long I listened to music on headphones, and more. Some of these, like steps and stairs, needed to be added up for each day. For others, like how fast I walked or how evenly I walked, I needed to find the average for each day. Using Python, I wrote code to do this. This code took the detailed data and made it into daily summaries. I checked to make sure it was doing this right by looking at the columns in each type of data. Once it was all set, I saved these daily summaries into new files. This way, I could look at them and find out more about my health habits. From doing this, I learned how to turn a lot of detailed data into something easier to understand. Now, I can see patterns like which days I walk more or how my walking changes over time. This is just the first step in my project, but it's important because it helps set everything else up. Next, I'll look more at these daily summaries to find out interesting things about my health and activity. After organizing my daily health data, I found that some days were missing. This happens when my phone doesn't record data for a day. To fix this, I added new rows for these missing days in my health datasets. For each missing day, I filled in the average value of all other days. This way, I could still learn about my health on days when my phone didn't record anything. For other information like which device I used or the unit of measurement, I filled in the missing spots with the details from the day just before. This makes sense because I usually use the same device and units don't change much day-to-day.

Filling Gaps in Data:

I also noticed that some days had zero values, which might not be right. To get a clearer picture, I changed these zeros to the average value of days that weren't zero. But I made sure that these zero days didn't affect the average calculation. By filling in these missing parts, I made my health data more complete. Now, I can trust that my analysis covers every day, not just the days my phone remembered to record. This helps me see the full picture of my health and activity patterns. Plus, changing the zero days means my averages are more accurate and show a better view of my regular activity.

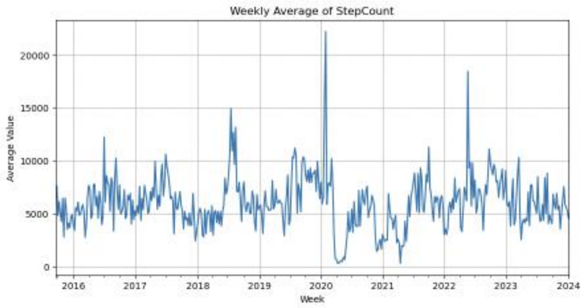
Analyzing Weekly Averages:

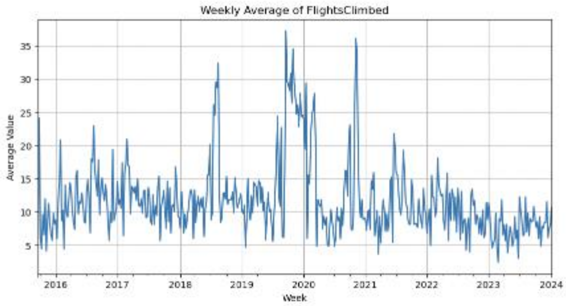
The next step in my health data project was to look at weekly trends. I wanted to see how my activities changed from week to week across several years. This is important to understand any long-term patterns in my behavior and health. For this, I used each of my health datasets, like the number of steps I took, how many stairs I climbed, and how much I used headphones. I changed the daily data into weekly averages. This means I combined each week's data into one average number to make it easier to see trends over time.

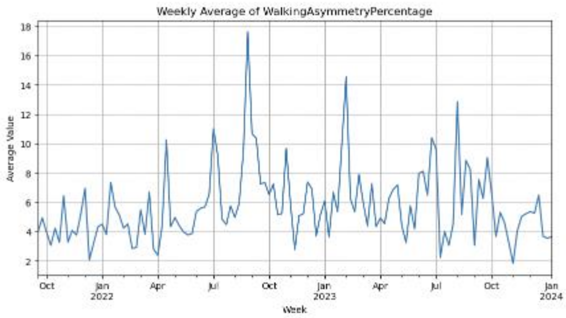
Visualizing Pandemic Impact:

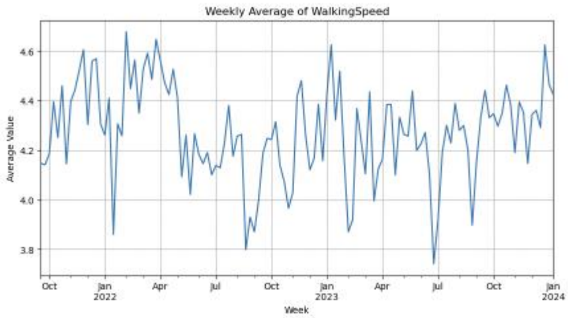
An interesting part of this analysis was seeing how the COVID-19 pandemic affected my activities. During the early months of the pandemic, we all had to stay at home. This change in routine showed up in my data, especially in the step count and flights climbed datasets. When I plotted these weekly averages, I could clearly see a change during the pandemic lockdown. My steps and stairs climbed went down because I was staying at home more. This visual change on the graph helped me understand just how much the pandemic changed my daily life.

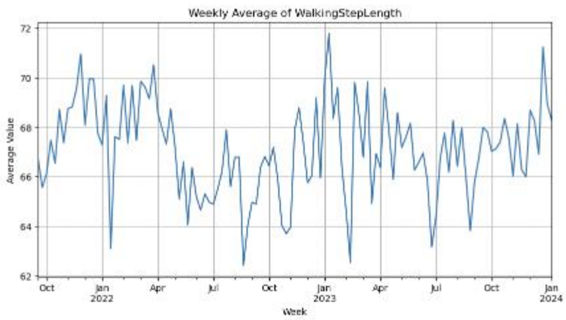
Week

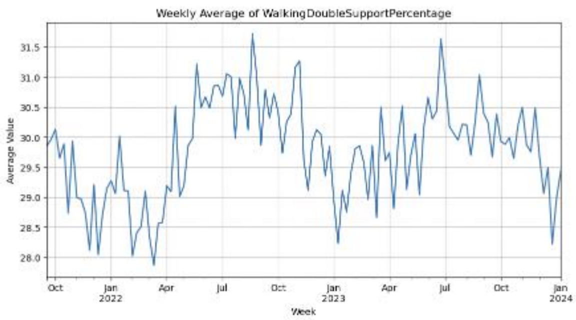

Weekly Average of StepCount
Average Value
0
5000
10000
15000
20000
2016
2017
2018
2019
2020
2021
2022
2023
2024
Week


Weekly Average of FlightsClimbed
Average Value
5
10
15
20
25
30
35
2016
2017
2018
2019
2020
2021
2022
2023
2024
Week


Weekly Average of WalkingAsymmetryPercentage
Average Value
2
4
6
8
10
12
14
16
18
Oct
Jan 2022
Apr
Jul
Oct
Jan 2023
Apr
Jul
Oct
Jan 2024
Week


Weekly Average of WalkingSpeed
Average Value
3.8
4.0
4.2
4.4
4.6
Oct
Jan 2022
Apr
Jul
Oct
Jan 2023
Apr
Jul
Oct
Jan 2024
Week


Weekly Average of WalkingStepLength
Average Value
62
64
66
68
70
72
Oct
Jan 2022
Apr
Jul
Oct
Jan 2023
Apr
Jul
Oct
Jan 2024
Week


Weekly Average of WalkingDoubleSupportPercentage
Average Value
28.0
28.5
29.0
29.5
30.0
30.5
31.0
31.5
Oct
Jan 2022
Apr
Jul
Oct
Jan 2023
Apr
Jul
Oct
Jan 2024
Week

Understanding Relationships Between Activities:

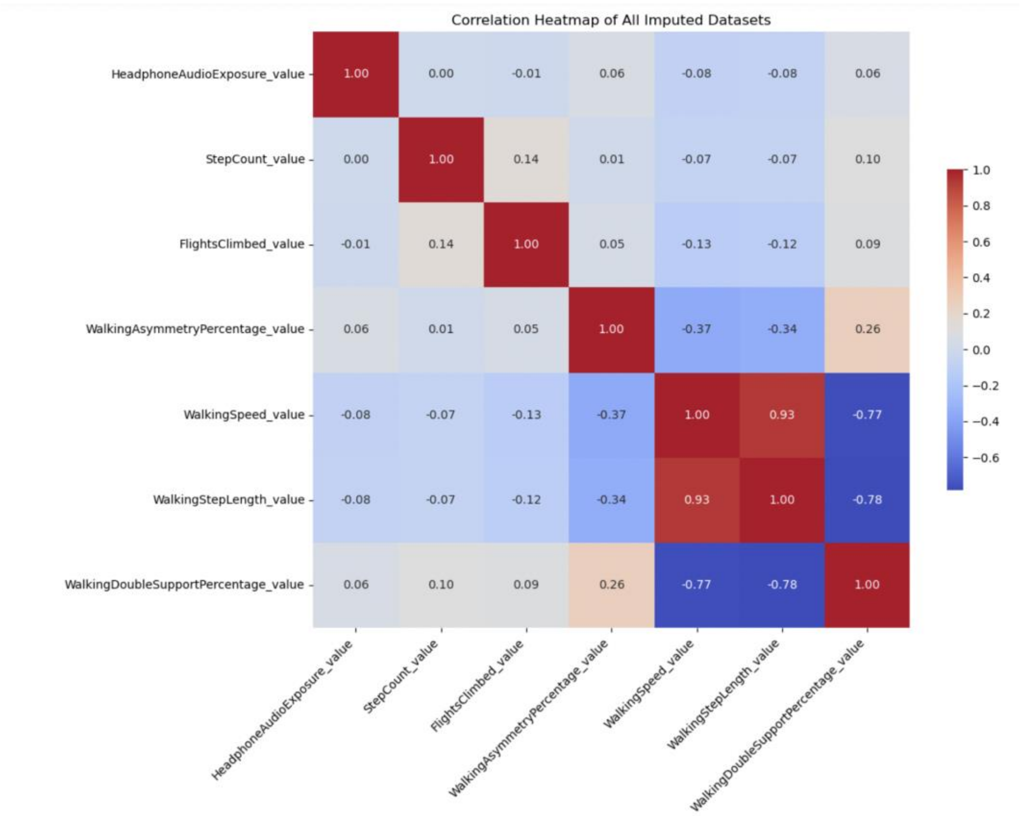
The fourth step in my health data analysis was looking at how different activities might be related to each other. I wanted to see if doing more of one activity could affect another. For example, if I walk faster, does that mean I take longer steps? Or does walking more steps mean I climb more stairs? To find out, I used a special kind of chart called a heatmap. This chart uses colors to show how strongly different things are connected. I made this heatmap for all my health activities, like step count, flights climbed, and walking speed.

Key Findings:

The heatmap showed some interesting connections:

1. Walking Speed and Step Length: The faster I walked, the longer my steps tended to be. This makes sense because longer steps usually mean a quicker pace. The heatmap showed a strong positive link between these two, which means they usually go up or down together.

2. Walking Double Support Percentage Correlation: There were two notable strong negative correlations here. First, as the walking double support percentage increased, both my walking speed and step length decreased. Walking double support percentage refers to the portion of the walking cycle where both feet are in contact with the ground. A higher percentage often indicates a more cautious walking pace, which is typical when stability is prioritized. The negative correlation indicates that on days when my walking was more cautious and stable (possibly due to fatigue, carrying loads, or walking on uneven surfaces), I tended to take shorter, slower steps.



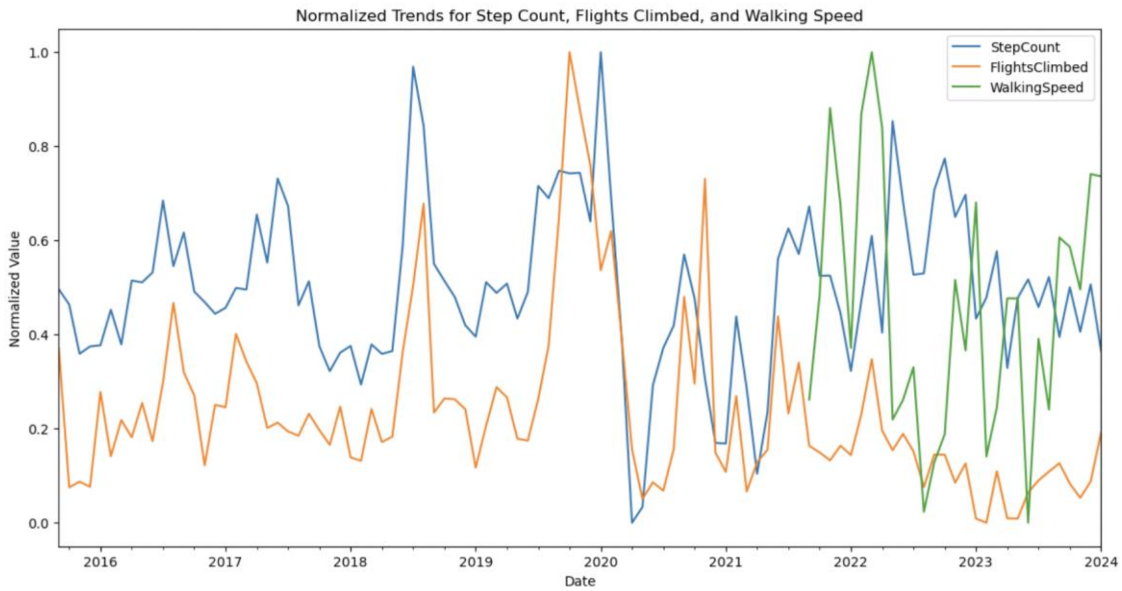
Trend Analysis Across Activities Normalization and Trend Comparison:

In the fifth phase of my data analysis, I aimed to compare the trends of my three main physical activities: step count, flights climbed, and walking speed. To accurately compare them, I normalized the data. Normalization is a technique that adjusts values measured on different scales to a common scale, often 0 to 1. This way, each dataset can be compared fairly without one overpowering another due to different ranges of values.

Plotting and Observations: After normalization, I plotted all three datasets together. The chart displayed clear patterns and trends over time. It was particularly revealing to see the similarity between step count and flights climbed, which, as confirmed by the correlation heatmap, have a strong positive correlation.

1. Step Count and Flights Climbed: Both these activities showed similar ups and downs in the graph. This trend suggests that on days when I walked more, I also tended to climb more stairs. The movements are closely linked, indicating that when I'm more active in one aspect, it's likely to be reflected in the other.

2. Walking Speed: While the walking speed also fluctuated, the pattern was distinct from the other two activities. This indicates that my walking speed may be influenced by different factors than just the number of steps or flights climbed.

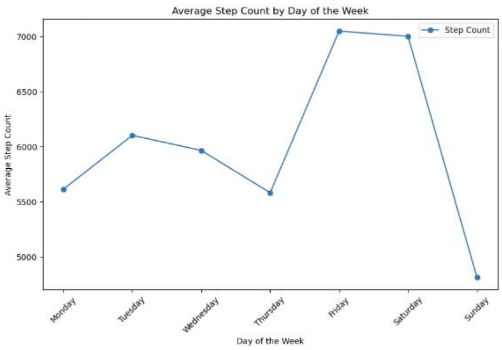


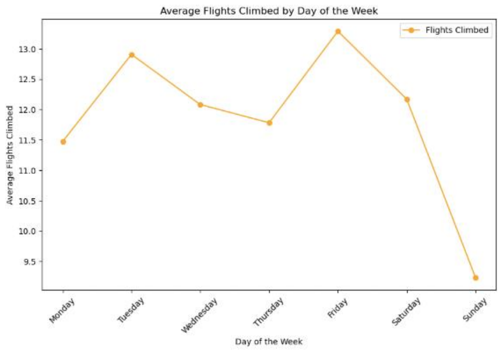
Assessing Daily Activity Patterns: In the sixth part of my data analysis, I examined my activity levels based on the days of the week. This approach provided a clear picture of which days I was most active and which days I tended to be less active. By doing this, I could identify trends and potentially optimize my weekly routine to enhance my overall fitness and well-being.

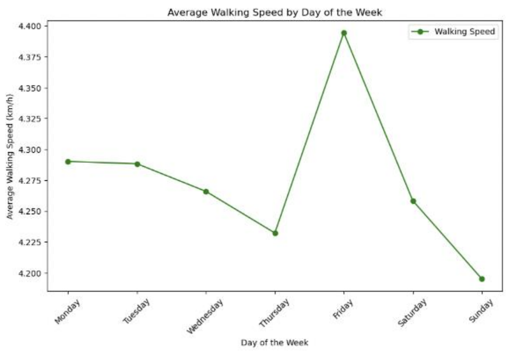
Active Days vs. Rest Days: The analysis revealed that Friday stood out as the day with the highest average activity, reflected in both my step count and flights climbed. This could be due to the end-of-week rush or perhaps more social activities that require moving around. On the other end of the spectrum, Sunday was the least active day, which might be attributed to it being a traditional rest day or used for more sedentary activities.

Strategic Exercise Planning: With this insight, I see an opportunity to balance my weekly activity levels. Implementing an exercise plan on Sundays could increase my overall activity and prevent the dip in physical movement that the data shows. Since Sundays are typically less busy, dedicating time to exercise could have a significant impact on my average activity, especially since any increase on this typically inactive day would raise the overall weekly average.

Visualizing the Trends: To properly visualize the data, I created separate plots for each activity type, which helped address any scale issues from combining different metrics. These plots clearly illustrated the average activity levels for step count, flights climbed, and walking speed across the days of the week.





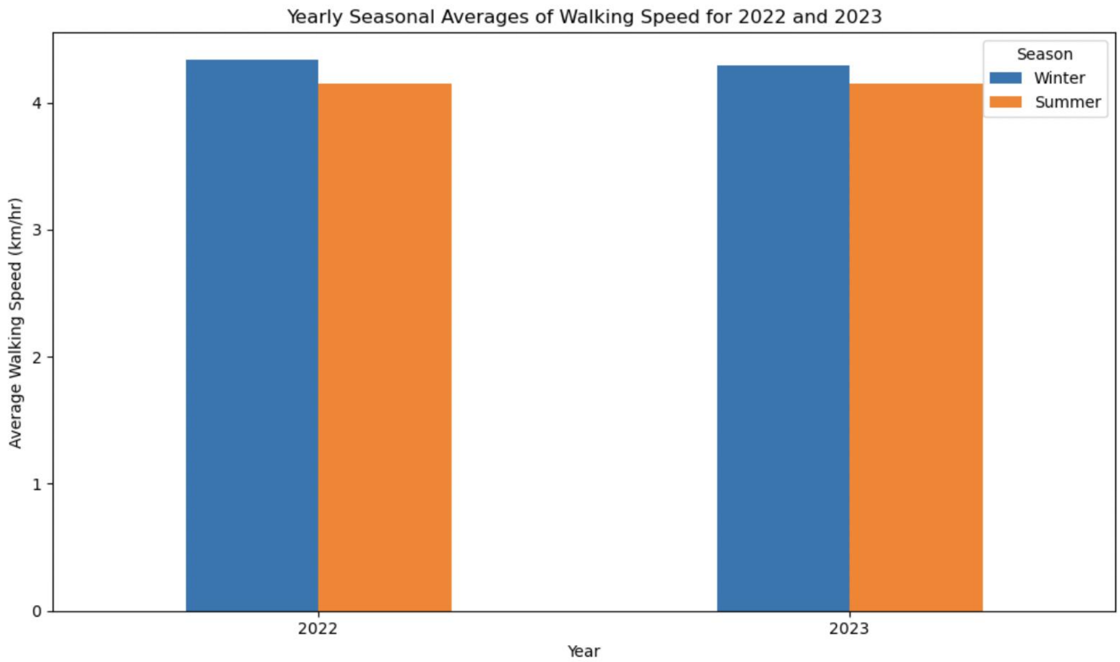


Walking Faster in Winter

Why I Did This Study: I was curious if I walked faster when it was cold outside. I thought maybe I speed up to stay warm.

What I Found Out: After looking at my walking speeds, I saw that I do walk a bit faster in winter than in summer. In winter, my average speed was about 4.3 km/h, and in summer, it was closer to 4.1 km/h.

Checking It Twice: I didn't just look at one year. I checked the winters and summers of both 2022 and 2023. I made a bar chart to show the speeds for each season, and both years showed the same thing: I walk faster when it's cold.

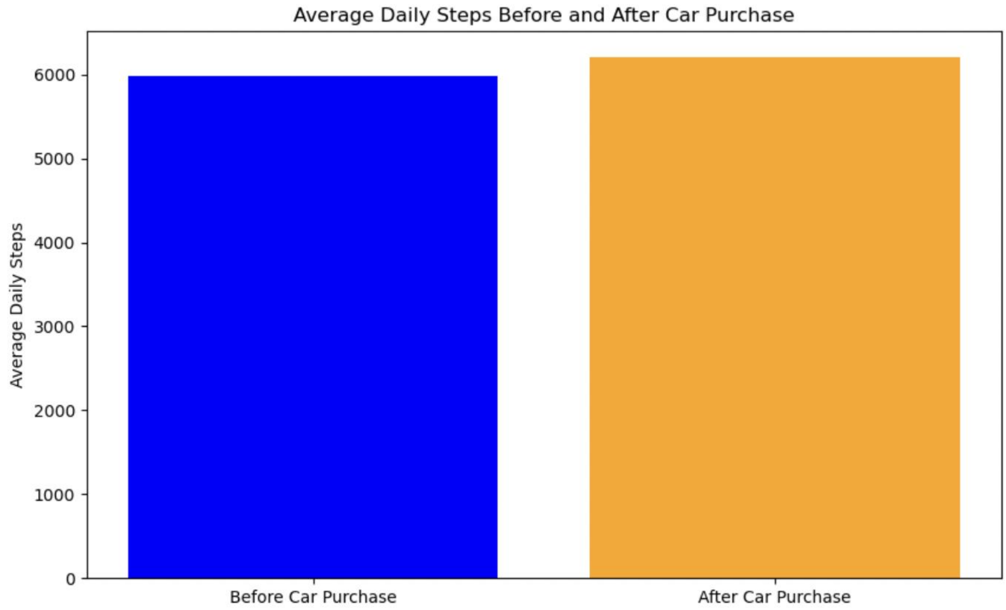


My Study on Walking More or Less with a Car

I was thinking about something that many people talk about – do we walk less when we have a car? After I got my car in October 2022, I wanted to see if this was true for me too. So, I decided to look at my own walking data from my iPhone to find out.

How I Looked at My Data: I used the walking and step count info that my iPhone Health app has been saving since 2015. First, I just compared how much I walked before I got my car to how much I walked after. But then I thought, maybe that's not fair because lots of things in my life have changed since 2015. So, I focused on just the year before I got my car and the time after that.

What I Found Out: It turns out, I was walking a bit more before I got the car. I used to walk around 6,604 steps every day, but after getting the car, it went down to about 6,202 steps. It's not a huge change, but it does seem like having a car might make me walk a little less.



My Study on Headphone Types and Audio Exposure

I was curious if different types of headphones would change how much noise I was exposed to. So, I took a look at my Sony WH-1000XM5 and Apple AirPods to see which one had higher dB levels when I used them.

How I Analyzed My Data: I used the data from my iPhone that tells me about my headphone audio exposure. I created a way to tell which data was from which headphones, and then I looked at the average dB for each one.

What I Found Out: It turns out my AirPods usually play music louder than my Sony headphones. The AirPods have an average dB of about 75.93, while the Sony ones are lower, at 74.09. I think this is because I tend to turn up the volume more when I use AirPods compared to the Sony headphones.

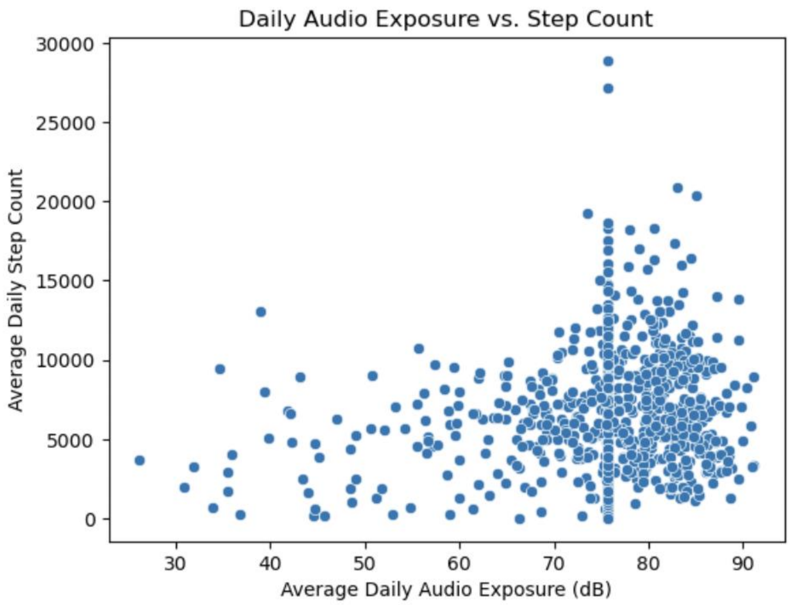
What I Think About This: This might be because of how the quality of headphones are like maybe the Sony headphones are better at noise-canceling, so I don't feel like I need to turn the volume up as much.

Exploring the Link Between Sound Exposure and Walking Habits

I got curious about whether the loudness of my headphones would affect how much I walk. It's an interesting thought, right? Like, do I end up walking more on days when my music's louder? I took my iPhone data that shows both my daily step count and how loud my headphones were. I lined up the dates to match and then checked out the average loudness and steps for each day.

Findings from the Data: It turns out there isn't much of a link between the two. The correlation between how loud my headphones are and how many steps I take is pretty low.

Interpreting the Results: This low correlation suggests that loud music doesn't really push me to walk more. So, even if my headphones are booming, it doesn't mean I'll be pacing more that day.




```
The correlation between daily audio exposure and step count is: 0.1546918287443016
```
