**Discovering Deeper Health Insights Through My Smartwatch Data**

Many people use smart gadgets—like smartwatches—to track aspects of their lifestyle such as activity, heart rate, stress, and sleep. While the accompanying apps offer helpful dashboards, they often don't reveal the full story hidden in the thousands of data points collected daily. In this article, I dive into my own data to uncover meaningful patterns and insights.

**My Data Source**

In 2022, I purchased the [**Garmin**](https://www.bestbuy.ca/en-ca/search?path=brandName%253AGARMIN&search=garmin+watches) Venu 2 Plus smartwatch, drawn by its accurate sensors and impressive battery life. Over two years of consistent use, the device has collected comprehensive data on my daily heart rate, sleep, stress, activity levels, respiration, Pulse Oximetry (Pulse Ox), and more. So far, it has tracked over **6.5 million steps** and more than **5,100 kilometers** of walking.

As a data scientist, I wanted to go beyond the app’s visualizations. I focused my analysis on two key questions:

1. How can I improve the quality and consistency of my sleep?
2. What factors influence how many calories I burn during different activities?

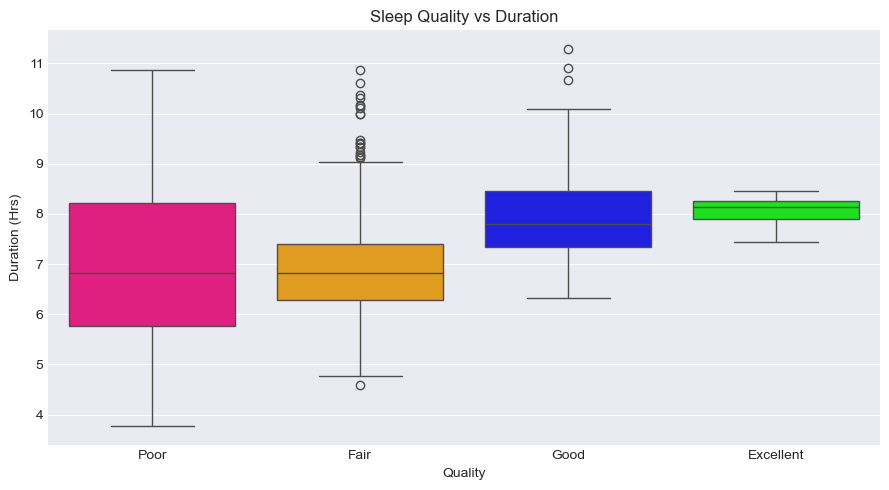
Using Python and libraries like Pandas, Matplotlib, Seaborn, Scikit-learn, and XGBoost, I explored and modeled my health data to find answers.

**Exploring Sleep Patterns**

My smartwatch categorizes sleep quality into four groups: *Excellent*, *Good*, *Fair*, and *Poor*. To understand what contributes to better sleep, I began by comparing these labels to key sleep metrics.

**Sleep Duration vs. Quality**

A clear pattern emerged: longer sleep was associated with better quality. On average, nights rated as *Good* or *Excellent* lasted around 8 hours, while *Poor* nights were notably shorter.



**Bedtime vs. Quality**

There was a weaker, yet noticeable trend linking earlier bedtimes to better sleep. *Fair* or *Poor* nights were more common following later bedtimes.

**Wake Time vs. Quality**

Interestingly, wake-up time didn’t show a consistent relationship with sleep quality in my data.

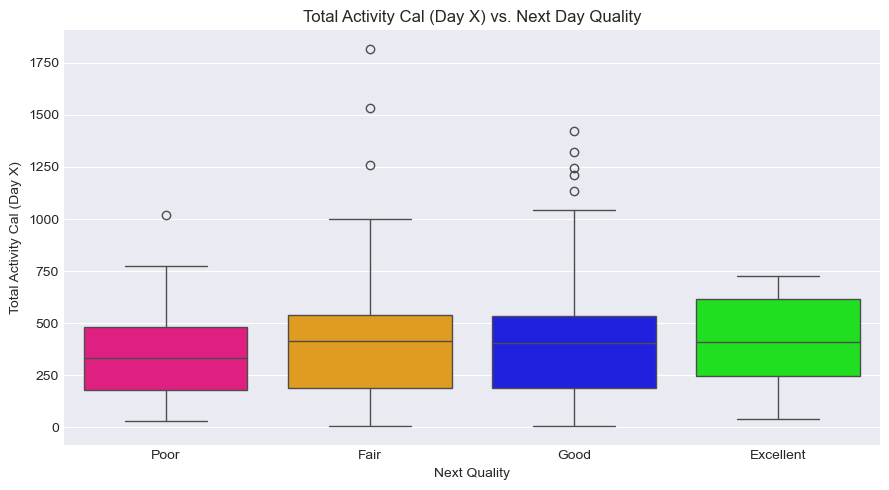
**The Role of Stress, Heart Rate, and Activity**

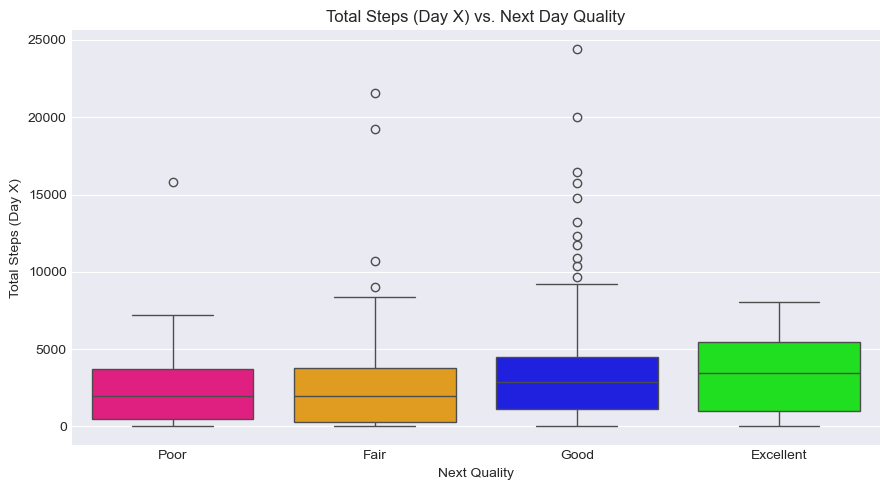
Recognizing that restorative sleep is vital, I investigated how stress and resting heart rate (RHR) affect sleep. A lower RHR during sleep was strongly associated with better quality, prompting me to explore how my daily stress and activity levels influenced RHR.



By integrating my activity data, I observed that days involving higher maximum heart rate, more calories burned, and exceeding approximately 4,500 steps were more often followed by *Good* or *Excellent* sleep. While correlation isn't causation, the data hints that consistent, moderately intense physical activity may support better sleep.



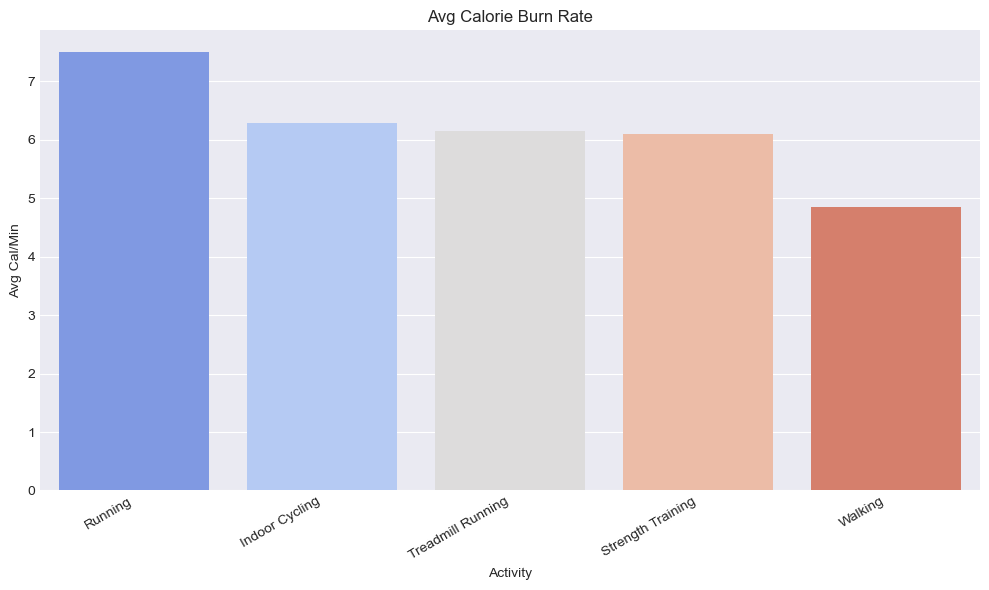




**Understanding Activity Intensity and Calorie Burn**

Next, I analyzed calorie burn by activity type. By grouping workouts and calculating calories burned per minute, I estimated the relative intensity of each activity:

1. **Running** – 7.50 Calories/Minute (Highest)
2. **Indoor Cycling** – 6.28
3. **Treadmill Running** – 6.15
4. **Strength Training** – 6.10
5. **Walking** – 4.85 (Lowest)

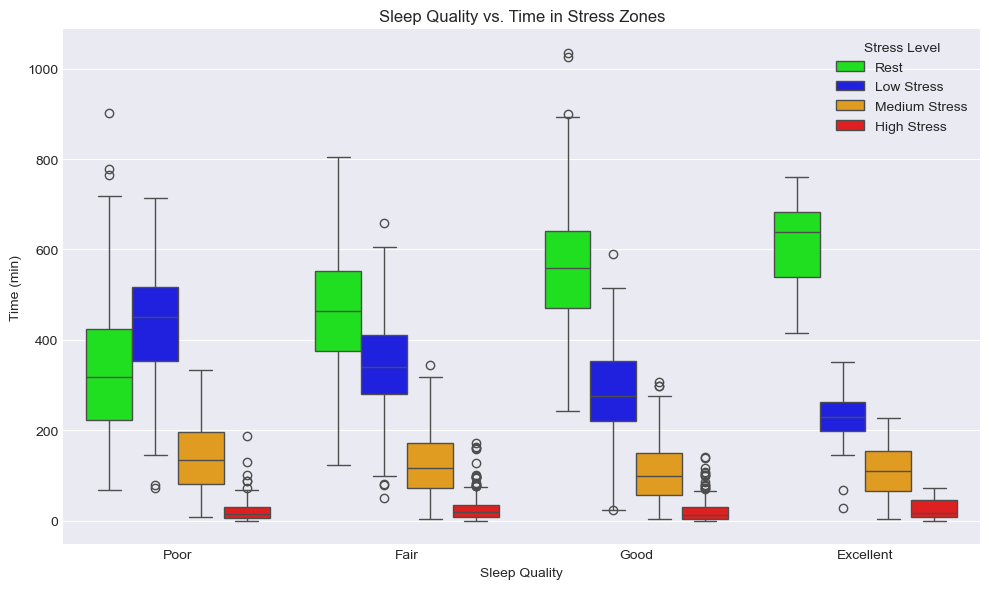


These results align with expectations. Further analysis (e.g., *Calories vs. Avg Heart Rate* plots) confirmed that higher heart rates during activity led to more calories burned. Activities like running and strength training produced higher average heart rates and calorie burn rates compared to walking. To match these effects, walking would require either a faster pace or significantly longer duration.

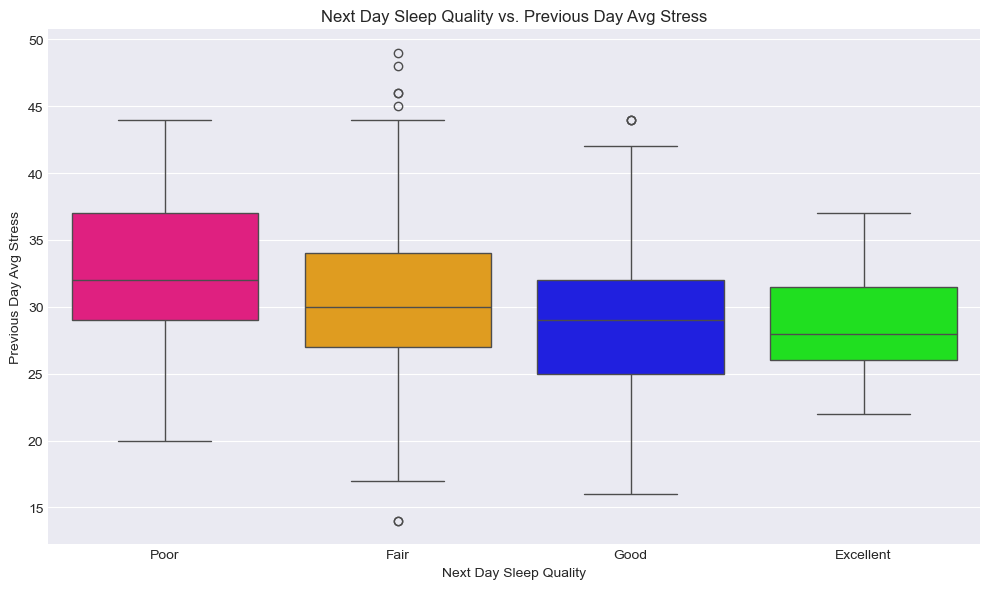
**Stress Patterns and Sleep**

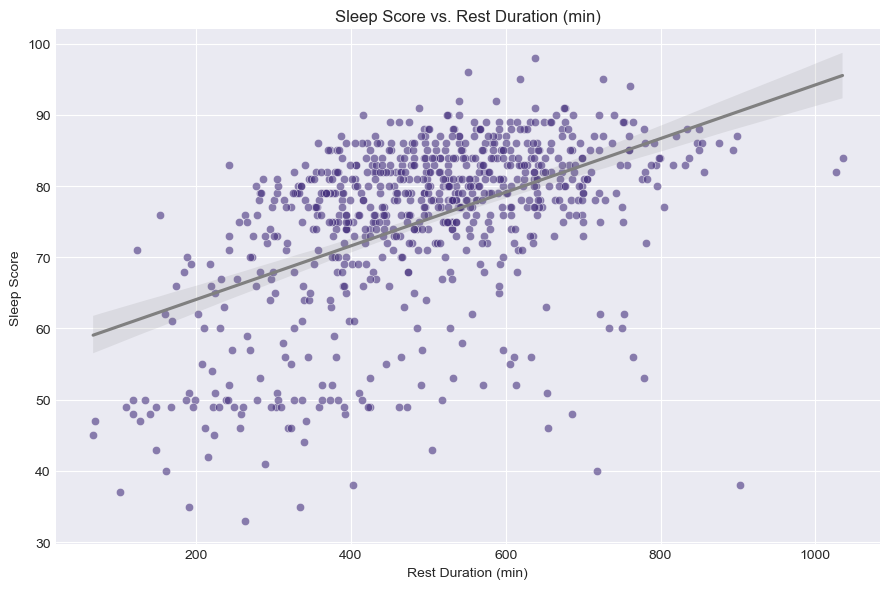
Returning to stress, I looked at how daily stress impacted sleep the following night. I found a consistent trend: higher stress levels—particularly time spent in *Medium* or *High* stress—were followed by lower sleep scores and shorter durations.

Interestingly, I didn't find a strong same-day link between physical activity and stress scores. This suggests that external factors, such as work or life events, likely play a larger role in daily stress fluctuations. Nevertheless, good sleep and regular exercise may still help manage how stress is perceived and handled.

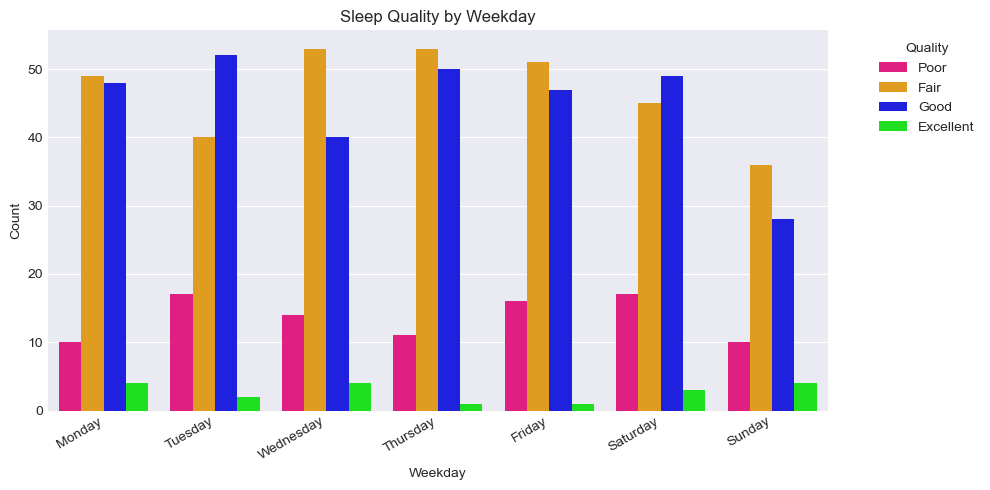
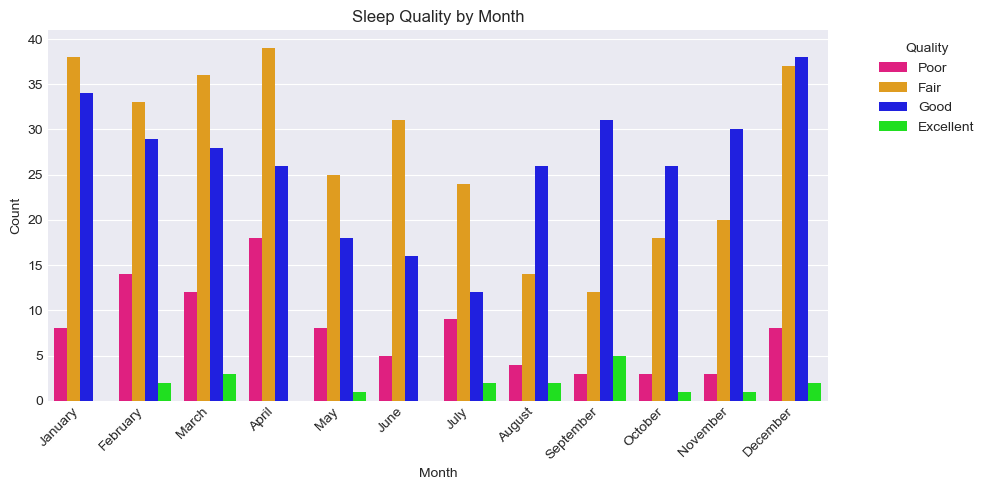


Days where more time was logged in the *Rest* state were associated with improved sleep quality and duration the next night—another signal that recovery and stress management support better sleep.





**Calendar Trends**

I also examined trends by day of the week and by month. Weekday analysis didn't reveal strong differences in sleep quality. However, monthly trends suggested a slight improvement in sleep during the late summer months (August and September), possibly influenced by seasonal lifestyle factors. 

**Predicting Sleep Quality with Machine Learning**

In the final stage of analysis, I applied machine learning techniques to see if I could predict my sleep quality category (Poor, Fair, Good, or Excellent) based on the various activity, stress, and physiological metrics from the preceding day. For this, I tested several algorithms, including Random Forest, XGBoost, and Logistic Regression, using a pipeline that handled data imputation (filling missing values), scaling (standardizing feature ranges), and class imbalance (using SMOTE to oversample rarer categories). I used 39 different metrics as input features for the models, including heart rate details from specific activities, stress durations, total steps, aggregated activity metrics, and previous day's metrics.

The results showed that the **Random Forest model achieved the highest accuracy** on the test data, correctly predicting the sleep quality category about **55.6%** of the time. While this indicates some predictive power, particularly for the common 'Fair' (61% F1-score) and 'Good' (67% F1-score) categories, the model (like the others tested) struggled significantly with predicting the rarer 'Poor' (38% F1-score) and especially the 'Excellent' (0% F1-score) sleep categories, which is a common challenge with imbalanced datasets.

To understand what factors the model found most influential in making its predictions, I examined the feature importances from the Random Forest model. As of **April 20, 2025**, the Top 3 most impactful features associated with my sleep quality according to the model were:

1. **Low Stress** (Duration in low stress state)
2. **Rest** (Duration in rest state)
3. **Avg Stress** (Overall average stress score for the day)

This feature importance ranking strongly suggests that **stress-related metrics were key indicators** associated with my sleep quality according to the model. Factors like how much time I spent in low stress or rest states, my overall average stress, and even the previous day's rest duration, were highly influential. Other factors appearing in the top 15 included activity heart rates (Max HR during Strength Training, overall Max HR), respiration rate, heart rate variability, previous day's stress, Pulse Ox, and total activity time/calories. This aligns with the relationships observed during the earlier exploratory data analysis, emphasizing the potential impact of daily stress and recovery on sleep outcomes.

**Conclusion**

To achieve better sleep quality, several key factors were identified through the analysis. Increasing total sleep duration and going to bed earlier are strongly associated with improved rest and recovery. Maintaining overall physical health through regular daily activity and achieving a higher heart rate during exercise also contribute positively to sleep quality. While daily stress may not always be fully controllable, the data suggests that our actions—such as staying active and managing rest—can mitigate its negative effects. In short, healthy routines during the day lead to better nights.