Fitbit Sleep Analysis

A Quantified Self Study

Background

There are many motivations behind self-tracking, but with the rise of newer technologies coupled with recent trends in physical/mental improvement, it is natural to want to study yourself. For me, this obsession started with trying to figure out if I had some gastrointestinal issues in 2017. At first, I told myself perhaps it was my age since I am no longer in my early 20s. As I researched further, I found the Nourish Balance Thrive Podcast. *Voila!* My passion of self improvement through nutrition and sleep was ignited. This podcast spoke to me because the host, Christopher Kelly, was a data scientist turned health pioneer. I also worked in the field of data science and needed answers about my health. In the past year, I have played around with my sleep but not intensely tracking it. Since the beginning of my data science immersive with General Assembly, I started tracking my sleep with a Fitbit. With that said, the purpose of this analysis, is two fold. I want to be able to see if a consistent week of good sleep will lead to more good sleep, and I would like to look at my sleep data not only on a daily view but track my sleep over a long period.

Problem Statement & Data

What questions are we trying to answer?

1. Does consistent good sleep lead to more good sleep?
2. What other factors can impact sleep?
3. What else can we explore?

There were 3 main data files used for this analysis: granular sleep data, sleep header data, and caffeine intake data.

Granular sleep data was minute level measures of time asleep, awake, and restless. The sleep header data gave aggregate level information about sleep. The caffeine intake data had 23 days of caffeine intake in mg.

High Level Statistics

	Mean	StdDev	Min	Max
TimeInBed (Mins)	408.04	123.68	61.00	696.00
Awake (Mins)	2.19	2.35	0.00	9.00
Restless (Mins)	17.73	10.58	0.11	54.00
Asleep (Mins)	388.10	117.48	56.00	677.00
Caffeine Intake (mg)	147.32	113.11	0.00	330.00

Let’s play with some Visualizations!

Interesting Inferences

Saturday is my most volatile day in terms of sleep and caffeine consumption.

As I sleep throughout the night, it seems that I have more minutes of sleep between the hours of 12am and 3am.

I have more consistent sleep Mondays - Wednesdays and Friday nights.

The later I go to bed, the more consistent my minutes asleep are as well.

~450%

The Standard Deviation of minutes of sleep on Thursdays and Saturdays are 450% larger than Tuesdays, Wednesdays, and Fridays!?!?

But what about modeling?

We tried using a linear regression for intake of caffeine to minutes of sleep, but that was inconclusive. When we tried to model minutes of sleep the next day, we had some signal but still some target leakage concerns. That also holds true for the RNN.

Conclusion & Next Steps

In terms of modeling, the data I collected did not seem to be sufficient and complete enough to give much signal if any at all. In terms of statistical inference it seems that we can answer the three questions posed in the first few slides. In addition, the demo code shows that anyone can easily use the code and recreate their own inferences about their sleep! The next steps should be collect more data. Perhaps we can use a user that has 1 years worth of data and very little days missing. I also recommending incorporating additional health data such as steps per day or food consumption.

Questions???