**Personal Data Analysis**

Self-Reported Mood Prediction using Lifestyle Tracking Data

19479166

Diarmuid Brady

CA4016 Application Domains

7th December 2022

Dr. Cathal Gurrin

# 1 Acknowledgements

This endeavour would not have been possible without Professor Alan Smeaton, who set up the FLOURISH module at Dublin City University. Without him, I would not have received my Fitbit Versa 3 smartwatch, and the data I gathered and the analysis I performed would not have been possible. I am very grateful for his contribution.

I want to express my deepest appreciation to Luke Gahan and Dr Elie Gottlieb and their respective Data Science and Applied Sleep Science teams at SleepScore Labs. While working at SleepScore Labs, I learned invaluable data science skills that significantly influenced the decisions I made during this project. I also had the opportunity to develop my knowledge in the sleep science domain which provided me with a greater understanding of the data I was working with, allowing me to dig deeper into the analysis.

# 2 Introduction

Self-reflection is a valuable tool when trying to improve oneself. In today's world, wearables such as smartwatches and lifestyle-tracking applications such as habit or mood trackers are primary examples of technology that can aid and facilitate self-reflection. A lot is within our control to manage and improve our sleep, exercise, and mood. One approach I take to managing my health is digital self-tracking. Digital self-tracking devices provide quantified outputs of bodily performances that users can then draw upon to inform future behaviours (Clark, Southerton, and Driller 2022). I practice digital self-tracking by collecting data about myself to monitor different areas of my life and learn what factors are essential for maintaining a healthy and balanced lifestyle. Over the past two months, from September 30th to November 25th, I monitored my lifestyle using three digital self-tracking tools; the SleepScore app, the Fitbit Versa 3 smartwatch, and the CBT Thought Diary app. In this paper, firstly, I outline the data gathering, cleaning, and feature engineering stages. Secondly, I describe the data analysis stage and what new insights I gained into my health and lifestyle. Thirdly, I report on three machine learning experiments I carried out to evaluate the predictive power of the data gathered. Furthermore, I discuss the potential business and user needs for this data and the ethical considerations of using this data. Finally, I reflect on the crucial personal insights I have gained from engaging in this reflective process and how I can build upon this work in the future.

# 3 Personal Data Overview

In this section, I describe the three sources of data I am using in this project.

## 3.1 Fitbit

Fitbit is a health and fitness company that enables users to monitor their health and fitness by providing a selection of fitness trackers. During the tracking period, I wore my Fitbit Versa 3 smartwatch almost daily, with few exceptions. I exported both my activity and sleep data from the Fitbit dashboard. Each row represented one day. The activity data included:

- Date: date of record
- Calories Burned: daily calories burned
- Steps: daily steps counted
- Distance: total daily distance travelled
- Floors: daily floors climbed
- Minutes Sedentary: time spent sedentary in minutes
- Minutes Lightly Active: time spent lightly active in minutes
- Minutes Fairly Active: time spent fairly active in minutes
- Minutes Very Active: time spent very active in minutes
- Activity Calories: calories burned during activity.

For sleep data, each row referred to a single sleep record, and variables include

- Sleep onset ($Onset$): time of first falling asleep
- Sleep offset ($Offset$): time of final awakening
- Total Sleep Time ($TST$): total duration of sleep.
- Wakefulness After Sleep Onset ($WASO$): total duration awake between sleep onset and offset.

- Time In Bed ($TIB$): time spent in bed attempting to sleep.
- Number of Awakenings ($NOA$): number of brief and long awakenings during the sleep session.
- Rapid Eye Movement Sleep Duration ($REMSD$): time spent in the REM stage of sleep.
- Light Sleep Duration ($LSD$): time spent in Non-REM stages 1 and 2 of sleep.
- Deep Sleep Duration ($DSD$): time spent in Non-REM stage 3 of sleep (Shrivastava et al. 2014).

A healthy range for $TST$ is between 7-9 hours (Hirshkowitz et al. 2015). For Fitbit, $TIB$ begins at sleep onset and ends at sleep offset. $WASO$ up to 20 minutes is considered normal during the night (Suni 2022).

## 3.2 SleepScore

SleepScore Labs is a sleep improvement company that enables users to monitor and improve their sleep using a non-contact sleep tracker and sleep improvement guide through an app on the phone. At night, I logged several variables on the app before I set up the phone beside my bed to track my sleep. I requested my sleep data from SleepScore through email for the stated period. The data consists of both sleep data and daily log data, and each row represents one night of sleep. In addition, SleepScore provides additional sleep metrics along with those provided by Fitbit. However, the start and end time refers to the beginning and end of recording rather than sleep onset and offset. The sleep variables include

- Sleep Onset Latency ($SOL$): time it takes to get to sleep
- Average Respiration Rate ($ARR$): breathing rate throughout the night
- SleepScore ($SS$): a measure of overall sleep quality for a single night's sleep
- MindScore ($MS$): a measure of overall REM sleep quality
- BodyScore ($BS$): a measure of overall deep sleep quality (Labs 2017)
- Sleep Efficiency ($SE$): the percentage of time in bed spent asleep
- Sleep Maintenance Index ($SMI$): the percentage of time from sleep onset to offset spent asleep.
- Arousal Index (AI): the number of awakenings per hour of sleep time (NSF 2022).
- REM Percentage ($REMP$): the percentage of total sleep time spent in REM sleep
- Slow Wave Sleep Percentage ($SWSP$): the percentage of total sleep time spent in deep sleep
- GPS location: location at the time of tracking

A short $SOL$ indicates a greater sleep hunger while a very long $SOL$ indicates difficulty with entering sleep (Pacheco 2021). A lower $ARR$ indicates a deeper sleep. A normal range is between 12 and 16 Meadows 2021. $SS$ is based on $TST$, $SOL$, $LSD$, $DSD$, $REMSD$, and $WASO$. The difference between $SMI$ and $SE$ is that $SMI$ does not include $SOL$.

I also track several factors in the SleepScore app right before going to bed as a reflection on my day. I will refer to this as the daily log. These factors include sleepiness, mood, stress, and the number of alcoholic drinks and caffeine cups consumed.

### 3.3 CBT Thought Diary

CBT Thought Diary is an app that enables users to track, learn about, and improve their mood by providing guide journals, mental health assessments, and crash courses on Cognitive Behaviour and other related topics. I set a reminder to track my mood each day at four o'clock. I exported my data from the app into a text file, including each mood entry, associated emotions, and reasons for mood. $Mood$ exists on a five-point scale; Terrible, Bad, Okay, Good and Terrific. The reasons for a mood include exercise, family, friends, hobbies, love, night out, projects, school, self-care, sleep, and entertainment. *Hobbies* refers to fun activities, while $Projects$ are more serious endeavours often related to self-improvement. *Love* refers to anyone I am romantically involved. $NightOut$ refers to going to a bar or nightclub with my friends. $SelfCare$ refers to activities like relaxing and listening to music or reading a book. *Sleep* often refers to why I feel tired or why I have so much energy. *Entertainment* refers to watching movies and television or using social media.

# 4 Data Cleaning and Engineering

I carried out data cleaning and feature engineering in six stages; -Fitbit activity data cleaning

1. Fitbit sleep data cleaning
2. SleepScore sleep data cleaning
3. Fitbit and SleepScore feature engineering
4. SleepScore daily log data cleaning
5. CBT Thought Diary mood data cleaning

I worked with each dataset individually before merging them all. I manipulated the data with Python, stored the data in a Pandas DataFrame, and created visualisations using Seaborn and Matplotlib within a Jupyter Notebook. I followed the pep8 style guide for renaming columns and variables.

## 4.1 Fitbit Activity Data

The Fitbit activity data was very straightforward. I removed six days of data when I was not wearing the match. As a result, the data did not represent my actual activity levels, as illustrated in Figure 1. Fitbit reports numbers using commas, e.g., 1,000. I removed the commas and converted the relevant columns to numeric features, e.g., 1000. After data cleaning, I created a 7-day rolling mean, denoted by "R7DM", for each numeric column to capture my overall activity around that period.

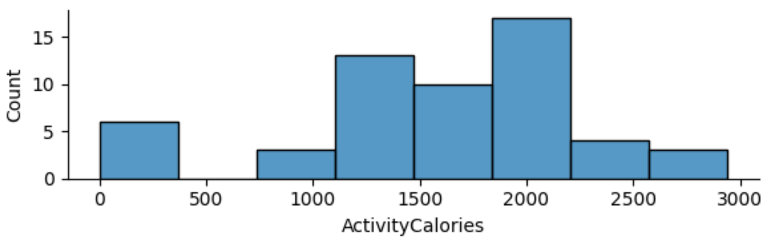


Figure 1: Active Calories Distribution with Inactive Days

## 4.2 Fitbit Sleep Data

Firstly, I created a single Date feature, then converted sleep onset and offset into hours past midnight. I chose hours over minutes because of its interpretability; for example, 0.5 means 00:30, and 16.25 means 16:15. For practical analysis, I changed the date to the following day. Mood refers to how I felt on the day, and sleep refers to how I slept the night before. One issue with bedtime is that when passing midnight, the hour jumps from 23 to 0. I transformed the bedtime by adding 24 hours to bedtimes after midnight and before noon (0 - 12), creating a new bedtime scale from 12 to 36 where 26 maps to 02:00 as $2 + 24 = 26$. Figure 2 illustrates how this transformation effectively captures the distribution of bedtimes. There are four peculiar records in Figure 2. Upon inspection, the very early and very late records were invalid. The early record around 19:00 was an error in which my watch had falsely detected sleep in my jacket pocket. The remaining two records referred to nights when I could not sleep; I had woken up at around 04:00. These two records had a twin record earlier in the night. Both records together represented the entire night of sleep. Therefore, I merged them, adding the time difference between the records to $WASO$ and $TIB$. To match the SleepScore data, I calculated additional features, including $AI$, $SWSP$, and $REMP$.

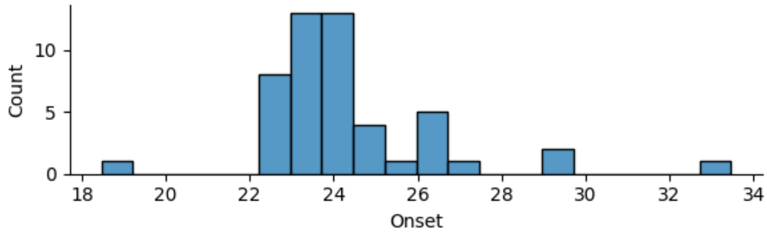


Figure 2: Fitbit Sleep Onset Distribution

## 4.3 SleepScore Data

I executed the cleaning procedure for the SleepScore data in a similar fashion to the Fitbit data. I first had to split the data into sleep and daily log. Since I manually turned on and turned off recording at night, there were fewer issues with bedtimes. However, there were signal quality issues. One night I was listening to a podcast, which corrupted half of the record because the app cannot function while the microphone is in use. I also found one record with an invalid $SMI$ value of 105. $SMI$ is measured on a scale from 0 to 100. Figure 3 shows the hypnogram for this record with the stages of wake (pink), REM (yellow), light (green), and deep (blue). Periods of black between two blocks represent no signal.

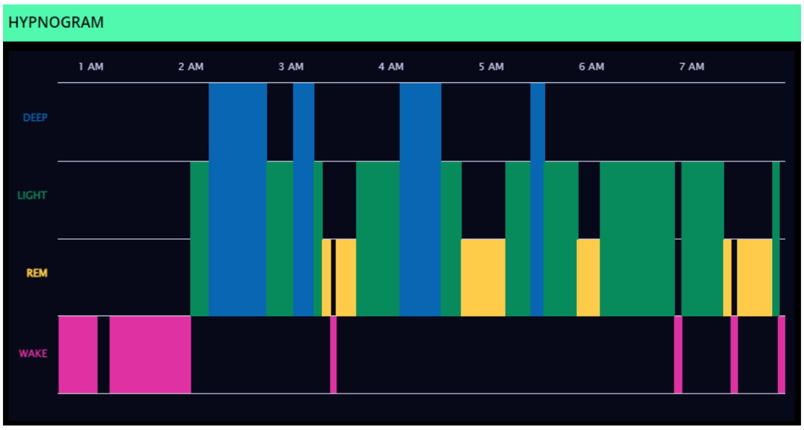


Figure 3: Extreme Sleep Maintenance Index Record

By reverse engineering the calculation for SMI, I discovered that SleepScore calculates SMI as $SMI = (TST/(TIB - SOL)) * 100$. The issue with this is that $TIB$ does not include the period of no signal (black space between two pink WAKE blocks), while $SOL$ does. I concluded that if $WASO$ is less than the period of no signal (black) during $SOL$, then $SMI$ will be greater than 100. In the next section, I recalculate $SMI$.

## 4.4 SleepScore and Fitbit Sleep Feature Engineering

I calculated many features for SleepScore and Fitbit together. NSF 2022 calculates sleep maintenance index as $SMI = (TST/TSDP) * 100$, where $TSDP$ (Total Sleep Duration Period) is the period between sleep onset and offset calculated as $TSDP = TST + WASO$. For SleepScore, I calculated the total time from sleep onset to offset, including the period of no signal. I corrected the $SMI$ error and ensured all values were between 0 - 100. While performing this task, I discovered that $TIB$ is calculated differently for Fitbit and SleepScore. For Fitbit, $TIB = TSDP$ but for SleepScore, $TIB = TSDP + SOL - NoSignalPeriod$. Mean Wake Duration ($MWD = WASO/NOA$) is the average duration of wakefulness after sleep onset. I calculated this feature to observe the relationship between the frequency and duration of wakeful periods. I then calculated the Sleep Midpoint and Sleep Regularity ($SR$). The midpoint is the mean of onset and offset. Sleep Regularity is a measure of day-to-day variability in sleep/wake patterns, and research shows it has significant effects on self-reported well-being (Phillips et al. 2017). Fischer, Klerman, and Phillips 2021 offers a selection of calculations for sleep regularity. Considering the data I possessed, I used the seven-day rolling standard deviation with a minimum of 4 days. To achieve an accurate calculation, I ensured that every row in the Pandas DataFrame represented each consecutive day from September 30th to November 25th. Therefore, null rows existed for days without a sleep record. Finally, I created the feature $IsWeekend$ to distinguish between the weekend and weekdays, where the weekend refers to the sleep records on Friday and Saturday nights.

## 4.5 SleepScore Daily Log

There was no cleaning required for the daily log. However, I created the feature $AtHome$ distinguishing the nights I slept at home versus at college using GPS location. I also calculated a seven-day rolling mean for $Sleepiness$ and $Stress$.

## 4.6 CBT Thought Diary Mood Data

The CBT Thought Diary mood data was all in a single text file. I had to parse each record using text processing. The only data I required for my analysis was my reported mood and the reason for the mood. I grouped some reasons, such as $Television/Movies$ and $SocialMedia$, into $Entertainment$. I created a one-hot encoding for all the reasons; each reason, e.g., $Friends$ and $Family$, became a binary feature. I transformed the mood feature into a binary feature of *Good* and *Bad* where *Good* consists of moods Good and Terrific, and *Bad* consists of moods Okay, Bad, and Terrible.

Once all the data was cleaned and engineered, I merged all the data into one DataFrame, and I was ready to move on to analysis.

# 5 Data Analysis

The quality of the questions asked determines the quality of the analysis. Therefore, I set out to answer the following questions:

- What are my current lifestyle habits?
- How has my lifestyle changed over the college semester?
- What lifestyle factors are associated with the good and bad days I experience?
- How does SleepScore sleep data compare to Fitbit sleep data?

In this analysis, I considered my lifestyle to be factors relating to my sleep, exercise, and mood. There were over eighty-five variables. Therefore, I needed to focus my analysis on a subset of critical factors.

## 5.1 What Are My Current Lifestyle Habits?

For sleep, I fall asleep at 23:55 and wake up at 07:31 on average. I sleep for 6 hours 58 minutes each night, just below the recommended 7 - 9 hours.

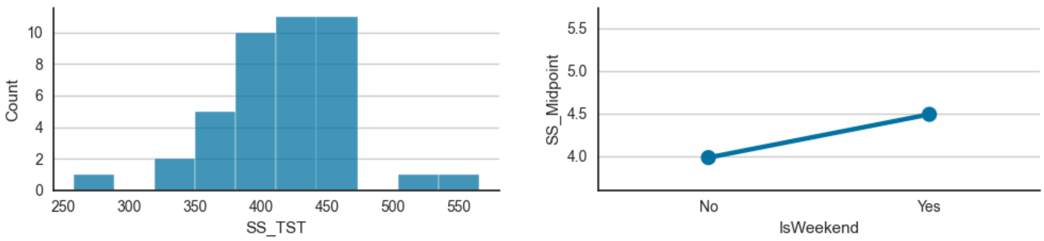


Figure 4: Total Sleep Time Distribution and Social Jetlag

Figure 4 shows low and high extreme values for total sleep time. The shorter sleep session relates to the night out at Halloween when I went to bed late. The two longer sessions (over 9 hours) were when I slept in on a Sunday after coming home from college. My mean sleep regularity is 24 minutes. In a week, the mean difference in midpoint is less than half an hour. In an ideal world, sleep regularity is zero. Social jetlag is the discrepancy in sleep/wake patterns between the weekdays and the weekend due to social obligations (Caliandro et al. 2021). In Figure 4, there is a 30-minute shift forward in sleep midpoint on the weekend, meaning that I go to bed and wake up later on the weekend compared to the weekdays. Fortunately, this is well below the 93-minute cut-off Sudy et al. set for young men aged 20-26 (Sűdy et al. 2019).

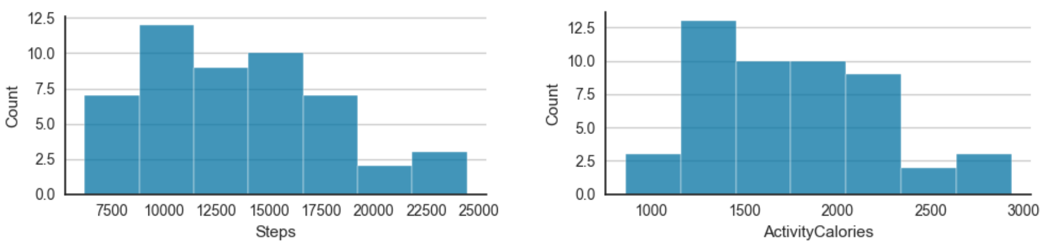


Figure 5: Daily Steps and Activity Calories Distribution

I clocked up over 13,000 steps a day for exercise, well above the recommended 10,000. When looking into the origins of 10,000 steps, I found that it came from a Japanese advertisement and had no scientific basis (Lieberman 2021)! However, Pozo Cruz et al. 2022 recently vindicated 10,000 for mitigating the risk of Dementia. I am highly active for 47 minutes and sedentary for 4 hours and 42 minutes on average daily. Due to activity, I burn 1769 calories daily, up as far as nearly 3000 calories one day! Figure 5 shows my daily calories and steps.

According to my daily log, I drink coffee and alcohol around once a week. However, as shown in Figure 6, I never drink more than one cup of coffee as it affects my ability to wind down in the evening, and I usually only have a few alcoholic drinks when I am out; on some night outs, I do not drink at all.

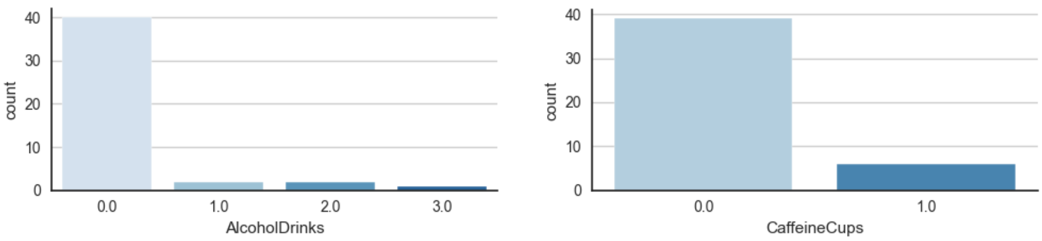


Figure 6: Alcohol and Caffeine Consumption

Finally, I feel good most days from my mood diary, with a mean score (0 - 4) of 2.4. Figure 7 compares the original *Mood* feature with the *Mood bin* feature.

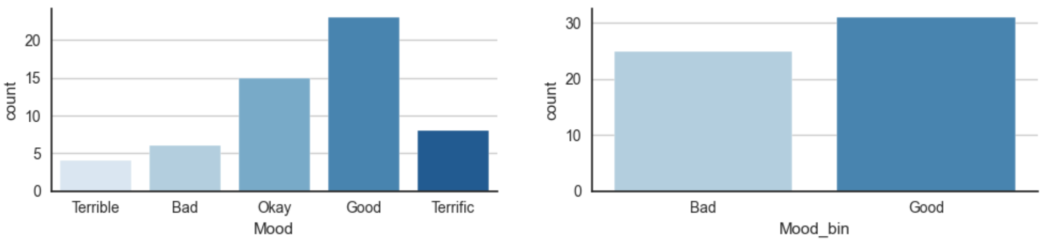


Figure 7: Comparison of 5-Point Scale and Binary Mood Features

## 5.2 How has my lifestyle changed over the semester?

Over my college semester, I became busier. I was active in many clubs, including boxing, gymnastics, and yoga, and societies, including DCU Enactus (a social enterprise society) and DCU Debate. I wanted to investigate how these extracurricular activities, alongside my study, impacted my lifestyle semester.

My activity increased over the semester with my steps and time spent actively increasing, as shown in Figure 8.

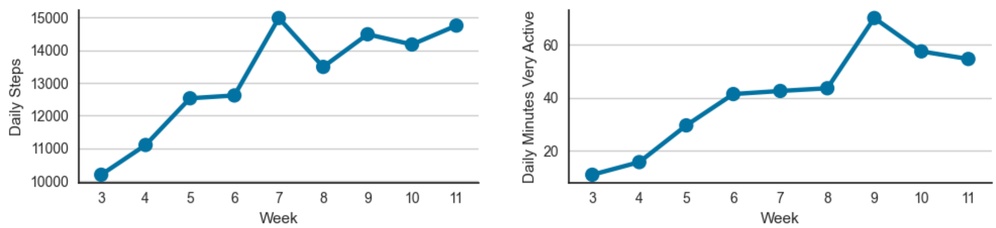


Figure 8: Daily Step and Minutes Very Active each week across the semester

Unfortunately, my sleep became increasingly irregular as each week progressed. For example, in Figure 9, my sleep regularity more than tripled from 12.6 minutes in week 3 to 39.6 minutes in week 10. My total sleep time dropped from 7 hours 36 minutes in week 3 to 6 hours 36 minutes in week 8. However, it rose again to 7 hours 13 minutes in week 10. Reflecting on this trend, I remember working later in the evenings on Enactus projects resulting in less sleep. As the weeks went by, I realized this was unsustainable, so I set a cut-off between 21:30 and 22:00 for any work on the laptop to allow myself to wind down properly before bed.

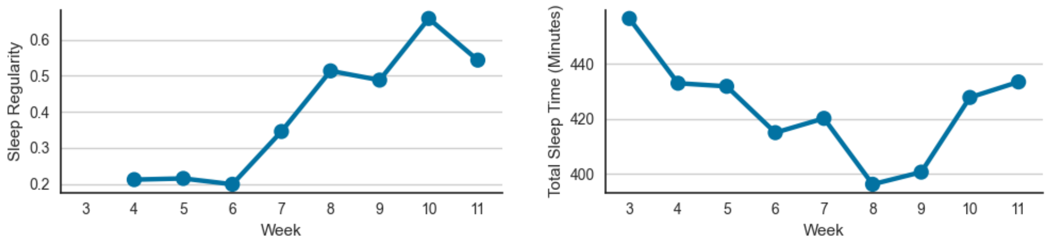


Figure 9: Total Sleep Time and Sleep Regularity each week across the semester

My mood stayed reasonably balanced and remained above the halfway threshold throughout the semester. Figure 10 illustrates how as each week went by, I was feeling well overall. Although, I was also reporting higher levels of sleepiness. The sleep graph in Figure 10 exemplifies how as my exercise, workload, and sleep debt increased, my feeling of sleepiness accumulated. Restoring my sleep by going to bed earlier helped alleviate the feelings of fatigue.

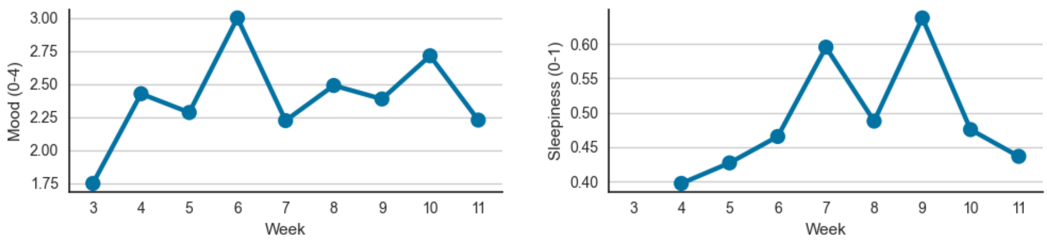


Figure 10: Mood and Sleepiness levels each week across the semester

## 5.3 What lifestyle factors are associated with the good and bad days I experience?

I compared various factors of my lifestyle on good and bad days. Unfortunately, due to the sparsity of samples for each factor, there are no significant differences, and I removed the error bars. I first examined the effects different types of people have on my life. Figure 11 illustrates the impact friends, family, and romantic interests have on my life. These are all reasons I reported alongside my mood in the CBT Thought Diary app. The y-axis represents the proportion of times a factor was the reason for a good or bad mood. Friends are associated with a high proportion of positive moods, my family also had a positive association, and finally, romantic interests had an equally positive and negative effect on my semester. Spending time with friends is one of the best things I can do for my mental health. Family is vital, and I should take out more time for them. For love, the results corroborate my subjective experience; it is as good as bad! I should manage my priorities carefully.

Figure 12 explores the consequences of a night out, alcohol, and caffeine on my mood. $NightOut$ is a reason I gave for my mood in the thought diary and reflects the day of the event, while $AlcoholDrinks$ and $CaffeineCups$ are factors I tracked referring to the previous day. $NightOut$ captures the anticipation and experience of an event. $AlcoholDrinks$ and $CaffeineCups$ capture


Figure 11: Association of Friends, Family, and Love with Good and Bad days

the effects of the beverages the following day. The proportions for each factor are low due to how infrequently I engage in the activities. Despite that, I never reported feeling bad on a night out, but the next day my mood was not as optimistic, a finding that I imagine many young people in college can attest to (McKinney and Coyle 2006). My mood may be associated with the context of an activity rather than the activity itself. For example, I drank coffee when I was fatigued and needed to be productive. Moreover, I used coffee to support days when I felt stressed about work. Moving forward, I will continue to drink alcohol, as it feels good on a night out, and caffeine helps when my energy is low, but I still need to get work done. However, I will continue to do so in moderation.

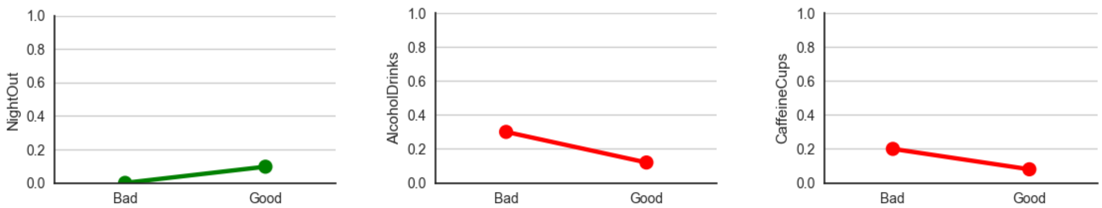


Figure 12: Association of a Night Out, Drinking Alcohol and Caffeine with Good and Bad days

Finally, I looked at exercise, extracurricular projects (Enactus), and school/college work. These are three reasons I gave in the thought diary app for my mood. Exercise is a gratifying part of my day as I almost always feel great during and after it. Projects and school, mainly, play a significant role in my life. They can be stressful at times, but there is great value in taking on challenging but eventually rewarding tasks.

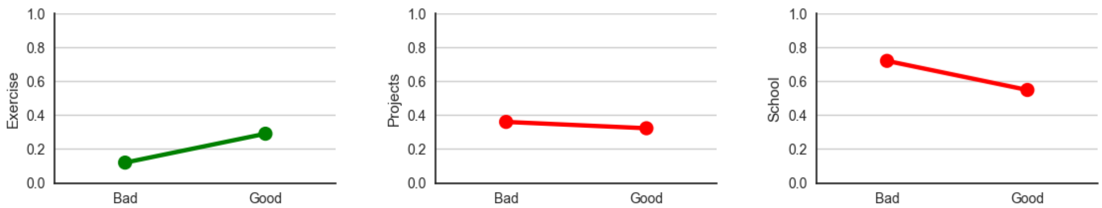


Figure 13: Association of Exercise, Projects and School with Good and Bad days

## 5.4 Comparative Analysis of Fitbit and SleepScore Data

Fitbit and SleepScore sleep tracking data have both are to be reliable, but they also have their limitations. For example, Fitbit overestimates Deep Sleep while SleepScore underestimates REM Sleep (Haghayegh et al. 2019). Therefore, since both technologies measure sleep differently and possess limitations, it is reasonable to hypothesise that my sleep data from Fitbit and SleepScore would differ. I tested this hypothesis by calculating the difference between Fitbit and SleepScore on a record level for each sleep variable. The number of awakenings was the only statistically significant difference. Fitbit calculates more awakenings than SleepScore with a mean difference of 22.18 awakenings and 95% confidence interval (CI) of (12.89, 31.46). Table 1 displays other non-significant results. The calculation used was $SleepScore - Fitbit$. Results greater than zero mean the SleepScore metric was more significant than the Fitbit.

Table 1: Non-Significant Differences Between SleepScore and Fitbit

| Var | TST | WASO | TIB | LSD | DSD | REMSD |
|---|---|---|---|---|---|---|
| Mean | 28.53 | 28.5 | 24.38 | -17.85 | 42.24 | 3.56 |
| Std | 34.25 | 15.4 | 44.77 | 32.58 | 24.01 | 25.14 |

I was surprised to see no significant differences in $TST$ and $WASO$ because, in the morning, I often noticed the difference in $TST$ when checking. Figure 14 presents the distribution of the differences between SleepScore and Fitbit for $TST$ and $WASO$. Fitbit calculated a longer $WASO$ than SleepScore across all nights I tracked.

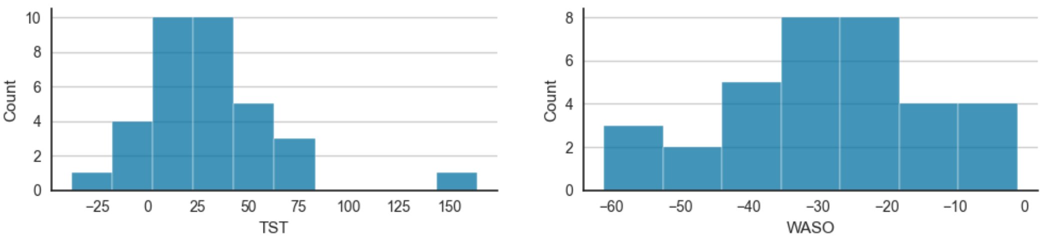


Figure 14: Total Sleep Time and WASO Difference Distribution

# 6 Data Modelling

I executed three tasks in the data modelling phases of the project:

1. Predicting SleepScore using only Fitbit data (Regression)
2. Predicting mood on a five-point scale (Regression)
3. Predicting mood as a binary variable (Classification)

I used the PyCaret AutoML Python Library to run machine learning algorithms and find the best model for each prediction task. I first created seven-day rolling mean and squared versions of all variables. Feature selection was a critical step due to the sheer quantity of features. I calculated Spearman's correlation for each feature to the target variable and selected features with a correlation greater than 0.25. Feature selection removed many features to mitigate the curse of dimensionality and ensure sufficient training samples for each feature. Next, I ran PyCaret's AutoML using these features and applied normalisation where necessary. Finally, I chose the top model and tuned it based on the chosen evaluation metric. For classification, I used accuracy. I chose this because it gave an overall evaluation of model performance, looking at both true positives and negatives as both classes were equally essential and reasonably balanced. For regression, I chose Mean Absolute Error ($MAE$) because I wanted to know how close the model predictions were to actual values, and $MAE$ provided interpretable results.

## 6.1 Predicting SleepScore using only Fitbit data (Regression)

Previously, we saw the differences between SleepScore and Fitbit data. Despite these differences, I investigated the predictive power of Fitbit sleep data for the SleepScore metric. As mentioned earlier, the SleepScore metric measures overall sleep quality on a scale of 0 to 100 for one night of sleep. Figure 15 shows a breakdown of SleepScore using several features; sleep duration ($TST$), time to get to sleep ($SOL$), light sleep ($LSD$), deep sleep ($DSD$), REM sleep ($REMSD$) and wake time ($WASO$). $TST$ (40%), $DSD$ (20%) and $REMSD$ (20%) place the greatest influence on SleepScore. Unfortunately, $SOL$ (5%) was unavailable to export from Fitbit, but I worked around it.

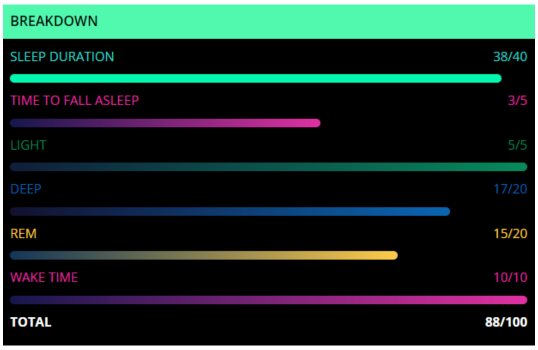


Figure 15: SleepScore Breakdown

The features selected can be seen in the heatmap in Figure 16. As expected, $TST$ has the most significant correlation with SleepScore.

Lasso Regression achieved the best model performance with $MAE$ of 2.23 and $R^2$ of 0.82 on the test set. The Figure 17 error plot shows how the predicted SleepScore was only off by 2.23 on average. SMI was most influential in the predictions, followed by $TST$ and $TSDP$. $SMI$ is calculated using $TST$ and $TSDP$ which simply consists of $TST$ and $WASO$. Intriguingly, the model made accurate predictions primarily using engineered features of $TST$ and $WASO$. Therefore, despite constituting 40% of the SleepScore, the model did not leverage deep or REM sleep.

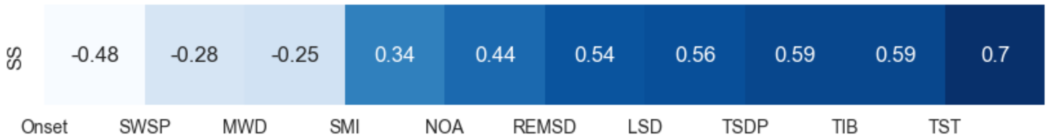


Figure 16: SleepScore Correlation Heatmap

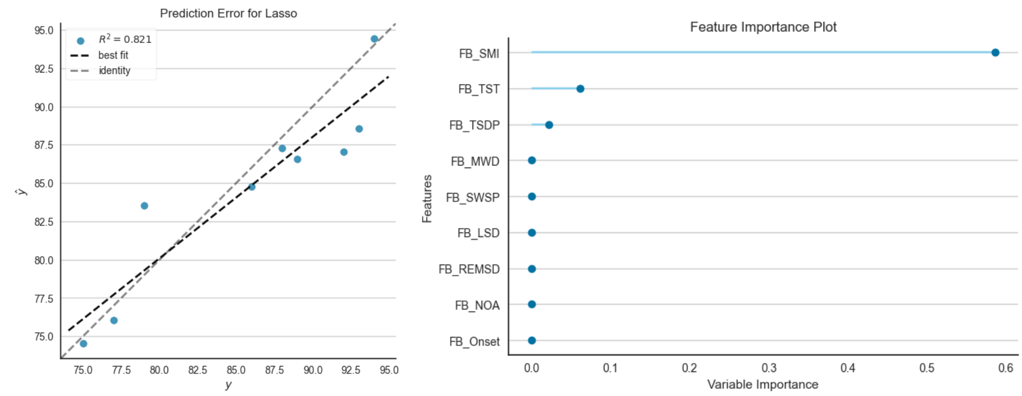


Figure 17: Lasso Regression Model Evaluation

## 6.2 Predicting mood on a five-point scale (Regression)

Before completing the classification task, I performed regression on the original $Mood$ feature, a more granular variable, and a challenging task. It was beneficial to investigate if the data could predict a more granular target variable to provide added context to the results of the subsequent classification task. Based on the analysis, I expected to see $Friends$ as a highly predictive variable. Therefore, I selected features by filtering out features with a correlation of less than 0.25 to either $Mood$ or $Mood_bin$. Figure 18 shows the heatmap for the features used in the regression. As hypothesised, $Friends$ is highly correlated with $Mood$. $Time$ is also a critical variable; this relates to the time of day I report, meaning reporting later in the day is associated with a more positive mood. Strangely, $REMP$, $REMSD$, and $MS$, which are all based on REM sleep, are negatively associated with $Mood$.


Figure 18: SleepScore Correlation Heatmap

Bayesian Ridge Regression achieved the highest performance with an $MAE$ of 1.1 on the test data. However, it received a very poor $R^2$ of -0.507. The mean would be more effective at predicting the target variable than the model itself. Figure 19 highlights how important it is to evaluate a model statistically and visually. The regression line of the predictions is far off the actual values. For feature importance, $Friends$ and $MS$ are the most influential. However, I did not consider them predictive based on the model's performance.

## 6.3 Predicting mood as a binary variable (Classification)

Due to the poor performance of the regression model, I resolved to simplify the problem and test the data on the binary target variable. I used the same features as before, with Figure 20 displaying

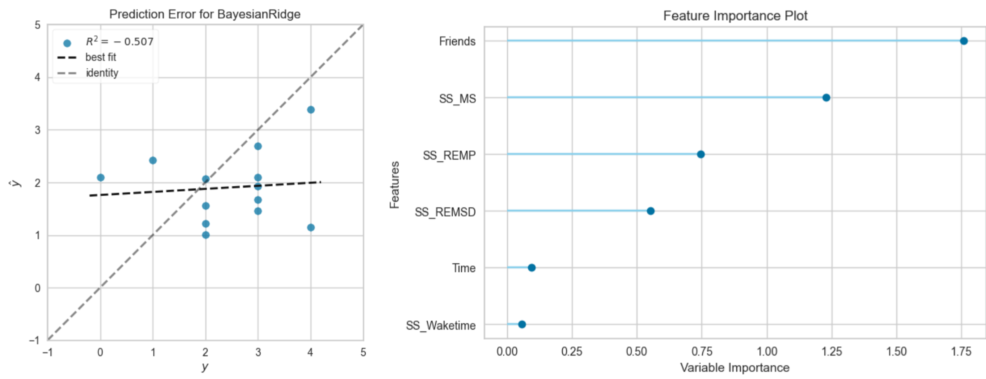


Figure 19: Bayesian Ridge Regression Model Evaluation

the correlation heatmap of $Mood_bin$.

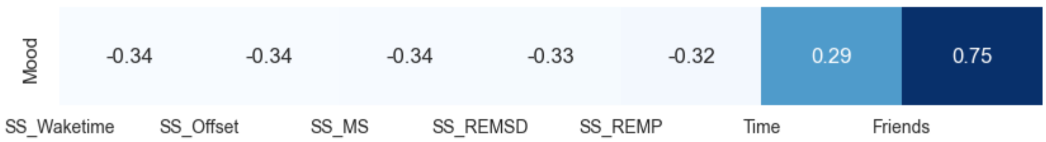


Figure 20: SleepScore Correlation Heatmap

Linear Discriminant Analysis ($LDA$) attained the highest performance. Figure 22 gives an overview of the results. Upon examining the results, I discovered the model's decisions were based solely on the feature $Friends$. If $Friends$ was 1, then the predicted $Mood$ was positive; if $Friends$ was 0, then the predicted $Mood$ was negative.

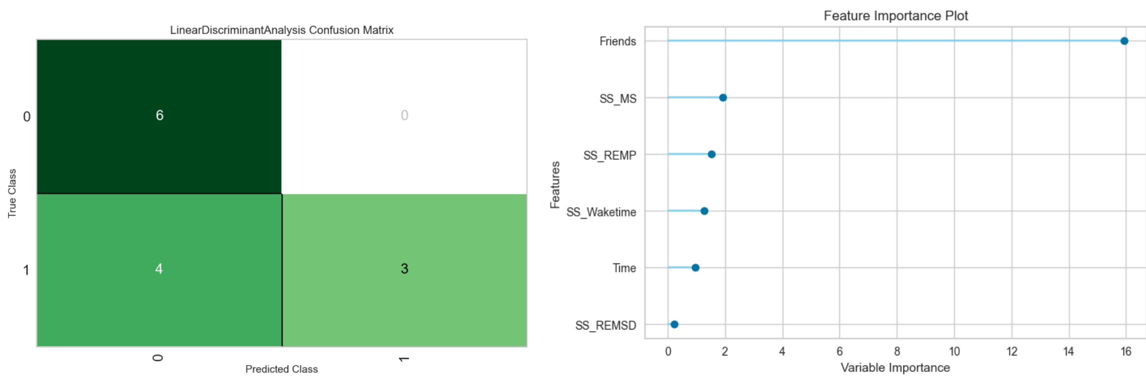


Figure 21: LDA Classification Model Evaluation

Figure 21 shows the Confusion Matrix. Many false negatives occurred in the model, leading to poor recall of the positive class. The model was inferior at detecting good days, or in other words, the model over-predicted bad days. The feature importance plot in Figure 21 aligned with the inference that it used $Friends$ for most of its decisions.

| | Model | Accuracy | AUC | Recall | Prec. | F1 | Kappa | MCC |
|---|---|---|---|---|---|---|---|---|
| 0 | Linear Discriminant Analysis | 0.6923 | 0.7143 | 0.4286 | 1.0000 | 0.6000 | 0.4091 | 0.5071 |

| | Time | Friends | SS_Waketime | SS_REMSD | SS_MS | SS_REMP | Mood_bin_num | Label | Score |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.725913 | 0 | -0.350465 | 0.703714 | 1.273915 | 0.402696 | 0 | 0 | 0.9952 |
| 1 | -0.514926 | 0 | -0.284582 | 0.662319 | 0.809800 | 0.668416 | 0 | 0 | 0.9959 |
| 2 | -0.610009 | 1 | 0.100761 | -0.952084 | -1.325127 | 0.574066 | 1 | 1 | 0.9984 |
| 3 | 0.317054 | 0 | 1.465815 | -1.324639 | -1.696419 | -1.285972 | 0 | 0 | 1.0000 |
| 4 | -0.120330 | 1 | -0.635325 | -1.117664 | -1.510773 | -0.985593 | 1 | 1 | 0.9999 |
| 5 | 1.291659 | 0 | -1.079915 | -1.986958 | -2.531826 | -2.154375 | 1 | 0 | 0.9996 |
| 6 | 0.597550 | 1 | -0.324397 | -0.289765 | -0.304075 | -0.419495 | 1 | 1 | 0.9999 |
| 7 | 0.231479 | 0 | -0.096414 | 0.289765 | 0.624154 | 0.058031 | 0 | 0 | 0.9970 |
| 8 | 0.150658 | 0 | 1.284281 | -0.165580 | 0.252862 | -0.142222 | 0 | 0 | 0.9996 |
| 9 | -0.472139 | 0 | 0.940174 | 0.331160 | 0.160040 | 0.083062 | 1 | 0 | 0.9994 |
| 10 | -0.329513 | 0 | -0.071767 | 1.448824 | 1.088269 | 1.338492 | 0 | 0 | 0.9987 |
| 11 | 0.588042 | 0 | -0.010149 | -0.289765 | 0.531331 | 0.366112 | 1 | 0 | 0.9989 |
| 12 | -0.533942 | 0 | -0.244769 | -0.041395 | 0.345685 | -0.063276 | 1 | 0 | 0.9942 |

Figure 22: LDA Classification Model Results

# 7 Business and User Need

There are a variety of sleep trackers available on the market currently. Each sleep tracker uses its algorithm to process and transform raw sensory data into sleep patterns. Therefore, merging data from several externally validated sleep trackers would be advantageous. Moreover, sleep is only one dimension of our life; while apps like Fitbit and SleepScore measure stress and overall mood, they lack the in-depth expertise and domain knowledge apps like CBT Thought Diary possess. Therefore, a service that assists users in integrating and examining several dimensions of their data would be successful. In this analysis, I merged various streams of my data to perform two main tasks 1) to predict the SleepScore metric using external Fitbit data and 2) to predict my mood. If performed on a larger scale, both of these tasks help 1) sleep-tracking companies integrate their data, producing a dataset where its sum is greater than the parts, and 2) educate users about what factors impact their experience using a data-driven approach.

# 8 Ethical Considerations

Any task that involves personal data undoubtedly poses ethical concerns. I had concerns about using my data. I performed all analysis tasks for this project using a Jupyter Notebook and saved them to my GitHub. Visually, the notebook reports all the code written and outputs generated. I had to consider what data I wanted publicly visible. The first step I took was to keep all data locally, ensuring no one else had access. Secondly, I was cautious about what data I was outputting in the notebook. CBT Thought Diary contains journal entries containing sensitive information about myself and others in my life. I needed to remove this before displaying any of the data. Thirdly, as part of my analysis, I used my geographical location from SleepScore to assess the days I was at home versus college. I am not too concerned about my college location as Dublin City University's location is not a mystery. However, I was uncomfortable with the coordinates to my home being easily accessible on my GitHub. Therefore, in all outputs, the coordinates have been hidden.

# 9 Key Personal Insights

This project has been a valuable learning experience, pooling data from various aspects of my life and exploring the trends and associations. Overall, I am happy with what my data has revealed about my lifestyle. I value my health, fitness, and emotional well-being. Based on the analysis I performed, I can confidently say I am a happy and healthy 21-year-old. I keep active every day with vigorous exercise dispersed throughout the week. At night, I get to bed at a reasonable time and sleep soundly with little difficulty. When my sleep hygiene falls to the waist-side, I usually notice and make the appropriate changes. I am fortunate to have an enjoyable social life filled with great people, whether they are my family, friends from home, college, or the various clubs and societies I am involved. I allow myself to go out, party and drink in a way that supports the other areas of my life rather than undermines them.

# 10 Future Work

Throughout this analysis, a constant challenge was that I had to contend with sparse data. I managed to track for close to sixty days in total. There were nights when I forgot to set up my phone for tracking or wear my watch resulting invalid records. The resulting DataFrame contained forty-two entire clean records. Suppose I revisit this project when I have gathered a more extensive dataset. I could answer more compelling questions, including the difference in lifestyle on and off the semester and how my lifestyle changes over Christmas or when I go on a holiday.

In the beginning, the quality of the questions asked determines the quality of the analysis. So rest assured, there are many fascinating questions yet to answer.

## References

Caliandro, Rocco, Astrid A. Streng, Linda W. M. van Kerkhof, Gijsbertus T. J. van der Horst, and Inês Chaves. 2021. "Social Jetlag and Related Risks for Human Health: A Timely Review." *Nutrients* 13, no. 12 (December): 4543. ISSN: 2072-6643, accessed December 6, 2022. https://doi.org/10.3390/nu13124543. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8707256/.

Clark, Marianne, Clare Southerton, and Matthew Driller. 2022. "Digital self-tracking, habits and the myth of discontinuance: It doesn't just 'stop'" [in en]. Publisher: SAGE Publications, *New Media & Society* (March): 14614448221083992. ISSN: 1461-4448, accessed December 6, 2022. https://doi.org/10.1177/14614448221083992. https://doi.org/10.1177/14614448221083992.

Fischer, Dorothee, Elizabeth B Klerman, and Andrew J K Phillips. 2021. "Measuring sleep regularity: theoretical properties and practical usage of existing metrics." *Sleep* 44, no. 10 (April): zsab103. ISSN: 0161-8105, accessed December 4, 2022. https://doi.org/10.1093/sleep/zsab103. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8503839/.

Haghayegh, Shahab, Sepideh Khoshnevis, Michael H Smolensky, Kenneth R Diller, and Richard J Castriotta. 2019. "Accuracy of Wristband Fitbit Models in Assessing Sleep: Systematic Review and Meta-Analysis." *Journal of Medical Internet Research* 21, no. 11 (November): e16273. ISSN: 1439-4456, accessed December 5, 2022. https://doi.org/10.2196/16273. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6908975/.

Hirshkowitz, Max, Kaitlyn Whiton, Steven M. Albert, Cathy Alessi, Oliviero Bruni, Lydia DonCarlos, Nancy Hazen, et al. 2015. "National Sleep Foundation's sleep time duration recommendations: methodology and results summary" [in eng]. *Sleep Health* 1, no. 1 (March): 40–43. ISSN: 2352-7226. https://doi.org/10.1016/j.sleh.2014.12.010.

Labs, SleepScore. 2017. *What Is A Sleep Score and How Is it Calculated — Sleepscore Labs,* October. Accessed December 6, 2022. https://www.sleepscore.com/blog/how-is-your-sleepscore-calculated/.

Lieberman, Daniel. 2021. *Exercised: Why Something We Never Evolved to Do Is Healthy and Rewarding.* December. ISBN: 978-0-525-43478-8.

McKinney, Adele, and Kieran Coyle. 2006. "Alcohol hangover effects on measures of affect the morning after a normal night's drinking" [in eng]. *Alcohol and Alcoholism (Oxford, Oxfordshire)* 41 (1): 54–60. ISSN: 0735-0414. https://doi.org/10.1093/alcalc/agh226.

Meadows, Austin. 2021. *Sleep Respiratory Rate* [in en], July. Accessed December 6, 2022. https://www.sleepfoundation.org/sleep-apnea/sleep-respiratory-rate.

NSF. 2022. *Definitions and Characteristics for Wearable Sleep Monitors,* September. https://www.thensf.org/wp-content/uploads/2022/10/ANSI-CTA-NSF-2052.1-A-FINAL.pdf.

Pacheco, Danielle. 2021. *Sleep Latency* [in en], August. Accessed December 6, 2022. https://www.sleepfoundation.org/how-sleep-works/sleep-latency.

Phillips, Andrew J. K., William M. Clerx, Conor S. O'Brien, Akane Sano, Laura K. Barger, Rosalind W. Picard, Steven W. Lockley, Elizabeth B. Klerman, and Charles A. Czeisler. 2017. "Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing" [in en]. Number: 1 Publisher: Nature Publishing Group, *Scientific Reports* 7, no. 1 (June): 3216. ISSN: 2045-2322, accessed December 7, 2022. https://doi.org/10.1038/s41598-017-03171-4. https://www.nature.com/articles/s41598-017-03171-4.

Pozo Cruz, Borja del, Matthew Ahmadi, Sharon L. Naismith, and Emmanuel Stamatakis. 2022. "Association of Daily Step Count and Intensity With Incident Dementia in 78430 Adults Living in the UK." *JAMA Neurology* 79, no. 10 (October): 1059–1063. ISSN: 2168-6149, accessed December 6, 2022. https://doi.org/10.1001/jamaneurol.2022.2672. https://doi.org/10.1001/jamaneurol.2022.2672.

Shrivastava, Deepak, Syung Jung, Mohsen Saadat, Roopa Sirohi, and Keri Crewson. 2014. "How to interpret the results of a sleep study." *Journal of Community Hospital Internal Medicine Perspectives* 4, no. 5 (November): 10.3402/jchimp.v4.24983. ISSN: 2000-9666, accessed December 6, 2022. https://doi.org/10.3402/jchimp.v4.24983. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4246141/.

Sűdy, Ágnes Réka, Krisztina Ella, Róbert Bódizs, and Krisztina Káldi. 2019. "Association of Social Jetlag With Sleep Quality and Autonomic Cardiac Control During Sleep in Young Healthy Men." *Frontiers in Neuroscience* 13. ISSN: 1662-453X, accessed December 4, 2022. https://www.frontiersin.org/articles/10.3389/fnins.2019.00950.

Suni, Eric. 2022. *Wakefulness After Sleep Onset* [in en], June. Accessed December 6, 2022. https://www.sleepfoundation.org/sleep-studies/wakefulness-after-sleep-onset.