Caffeine and Sleep Study

Motivation:

It’s often that me and my friends are complaining about how we haven’t slept as due to our excessive caffeine consumption. My motivation is to clarify whether we are right about blaming caffeine and if so how could we prevent accidental overconsumption of caffeine. In this study the aim is to find a relation between caffeine consumption and sleep quality. Main objective is to locate a relation so that we can train a model and predict sleep quality based on a certain level of caffeine consumption.

Data Source:

Data used for sleep analysis is directly fetched from apple health. Dataset for caffeine intake is from an IOS app called HiCoffee, as apple health’s caffeine metrics does not provide hourly caffeine consumption info. I've logged my caffeine consumption for 2 months and fetched the raw data from the app. All other steps of cleaning, organizing etc. is done in Jupyter Notebook (provided in Github repo) based on raw data and %100 reproduceable.

Data Analysis

> Format of Apple Health Data: Apple Health creates multiple sleep object instances for each night based on movement detection with a start and end date (Such As:). Based on that they can target total number of awakening sleep stage by the duration of each sleep instance etc. However, this makes a quite complicated raw data.

Data Cleaning:

- Filtered out only sleep metrics among all health data.
- Clustered all data under date index to gather all sleep instances of a particular night together.
- Cleaned duplicate rows of sleep instances.
- Cleaned overlapping sleep instance data.
- Shift sleep data frame by -1 as the date created is the morning of awakening but not the night of the start of the sleep session.
- Removed all missing rows of caffeine data

- Cluster caffeine consumption on a daily basis.
- Divide the day into four time zones and structure my consumption in those four intervals, which are ['08.00-12.00', '12.00-16.00', '16.00-20.00', '20.00-24.00'].
- Caffeine in the body has a halving time, which means that my caffeine intake in the first period will also affect my caffeine levels in later intervals, thus created another caffeine data frame considering halving effect.
- Merged two datasets and removed each row which lacks either caffeine data or sleep data. (There were some as some days I forgot to wear apple watch and some days I haven't logged my caffeine consumption.)

Sleep DF:

creationDate	bed_time	awake_time	sleep_counts	time_in_bed	REM_Total	Deep_Sleep_Total	Core_Sleep_Total
2023-10-24	2023-10-23 23:31:05	2023-10-24 08:10:05	12	519.0	102.5	51.0	339.0
2023-10-25	2023-10-25 00:05:58	2023-10-25 08:44:28	6	518.5	87.5	49.0	348.0
2023-10-26	2023-10-25 23:45:44	2023-10-26 08:09:14	10	503.5	81.0	45.5	350.5
2023-10-27	2023-10-27 00:15:04	2023-10-27 07:54:34	9	459.5	86.0	67.5	275.0
2023-10-28	2023-10-28 00:55:30	2023-10-28 08:55:00	12	479.5	91.0	77.0	291.5

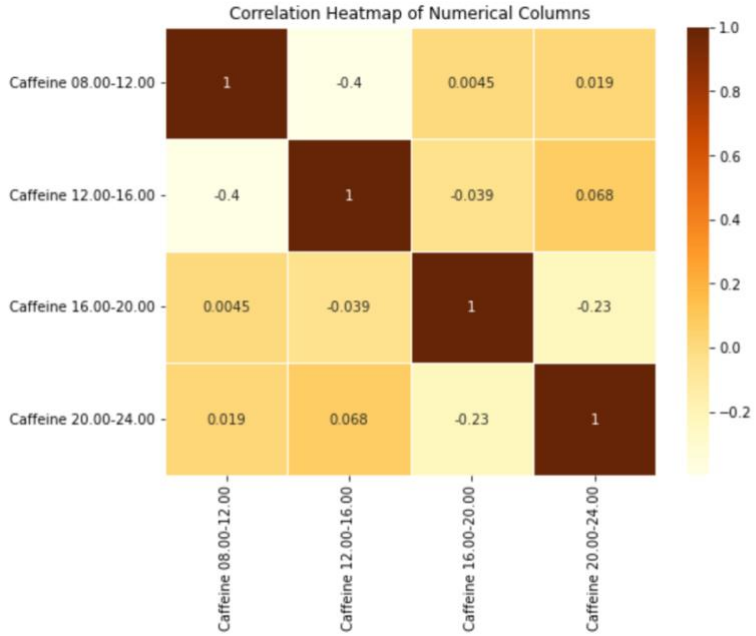
Caffeine DF:

Date	Caffeine (mg)	Caffeine 08.00-12.00	Caffeine 12.00-16.00	Caffeine 16.00-20.00	Caffeine 20.00-24.00
2023-10-21	75	0.0	0.0	75.00	37.500
2023-10-22	167	0.0	60.0	137.00	68.500
2023-10-23	81	15.0	73.5	36.75	18.375
2023-10-24	88	46.0	65.0	32.50	16.250
2023-10-25	34	2.0	33.0	16.50	8.250

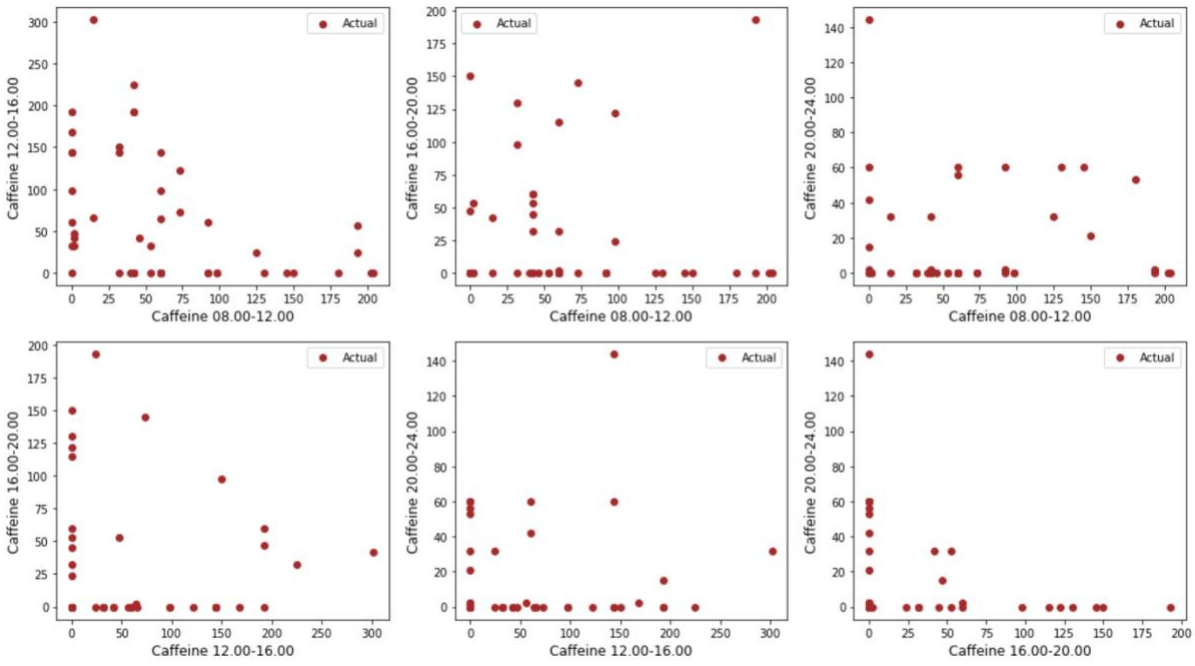
Caffeine Consumption Correlation analysis:

- Investigated whether my consumption of caffeine in a certain interval has an effect on a later interval.

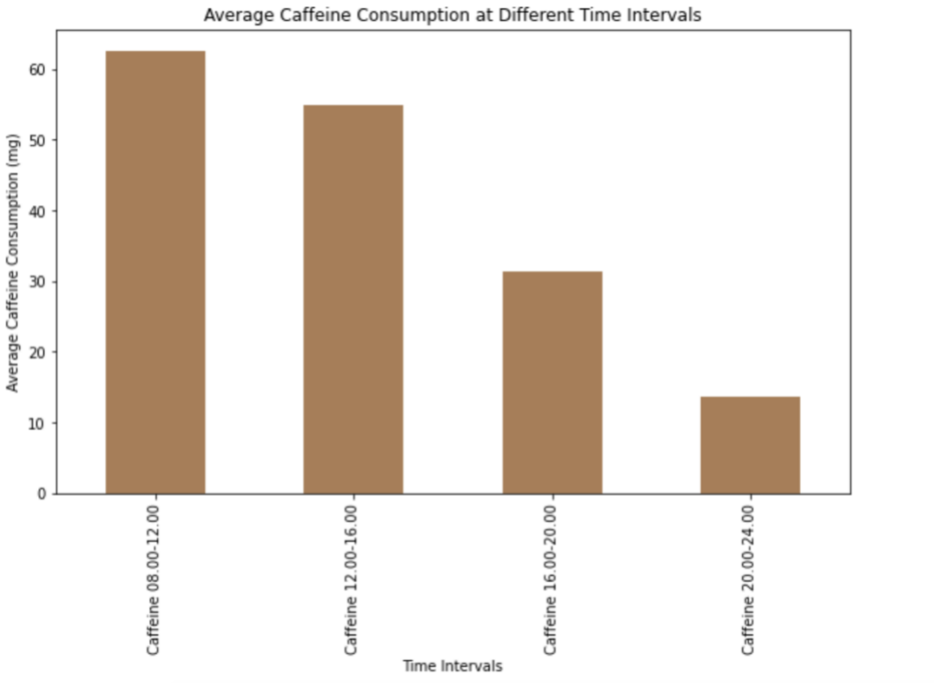
Correlations Heat Map:



Individual Correlations of each time interval:



Average Caffeine Consumption:



Review of Caffeine Consumption Correlation analysis:

1. My caffein Consumption between 08.00-12.00 correlates negatively with my consumption between 12.00-16.00. Which might mean I either consume caffein in morning or compansate my lack of consumption in the first period later.
2. My caffein Consumption between 16.00-20.00 correlates negatively with my consumption between 20.00-24.00. Which might mean I compensate my lack of consumption in one of those intervals. If I drink coffee during 16.00-20.00, I tend to consume less coffee before sleeping.
3. Other than that, there isn't any significant correlation.

Sleep Analysis:

For sleep scoring there are guidelines provided by authorities such as AASM (American Association of Sleep Medicine). However, their methodologies are much more complex including EEG and EKGs so based on their motivation Create an algorithm to score sleep based on Apple data.

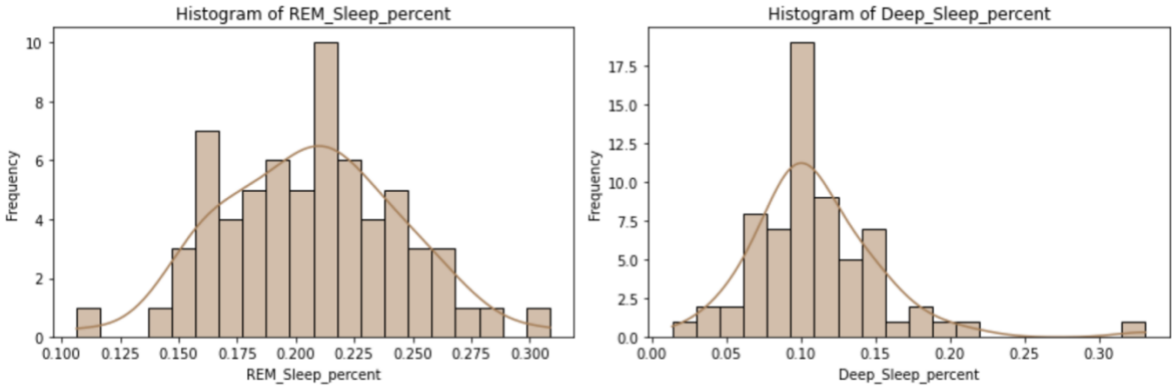
1. Add helpful data to our dataset which are REM sleep duration, Deep sleep duration, Core Sleep Duration.
2. Based on AASM lets derive sleep efficiency metric such as (time on sleep / time in bed)
3. Add percentages of each sleep stage.

Derived Data Frame:

creationDate	bed_time	awake_time	sleep_counts	time_in_bed	REM_Total	Deep_Sleep_Total	Core_Sleep_Total	REM_Sleep_percent	Deep_Sleep_percent	Core_Sleep_percent	Sleep_Efficiency
2023-10-24	2023-10-23 23:31:05	2023-10-24 08:10:05	12	519.0	102.5	51.0	339.0	0.197495	0.098266	0.653179	0.948940
2023-10-25	2023-10-25 00:05:58	2023-10-25 08:44:28	6	518.5	87.5	49.0	348.0	0.168756	0.094503	0.671167	0.934426
2023-10-26	2023-10-25 23:45:44	2023-10-26 08:09:14	10	503.5	81.0	45.5	350.5	0.160874	0.090367	0.696127	0.947368
2023-10-27	2023-10-27 00:15:04	2023-10-27 07:54:34	9	459.5	86.0	67.5	275.0	0.187160	0.146899	0.598477	0.932535
2023-10-28	2023-10-28 00:55:30	2023-10-28 08:55:00	12	479.5	91.0	77.0	291.5	0.189781	0.160584	0.607925	0.958290

Analyze Numeric Values of Sleep Data Frame:

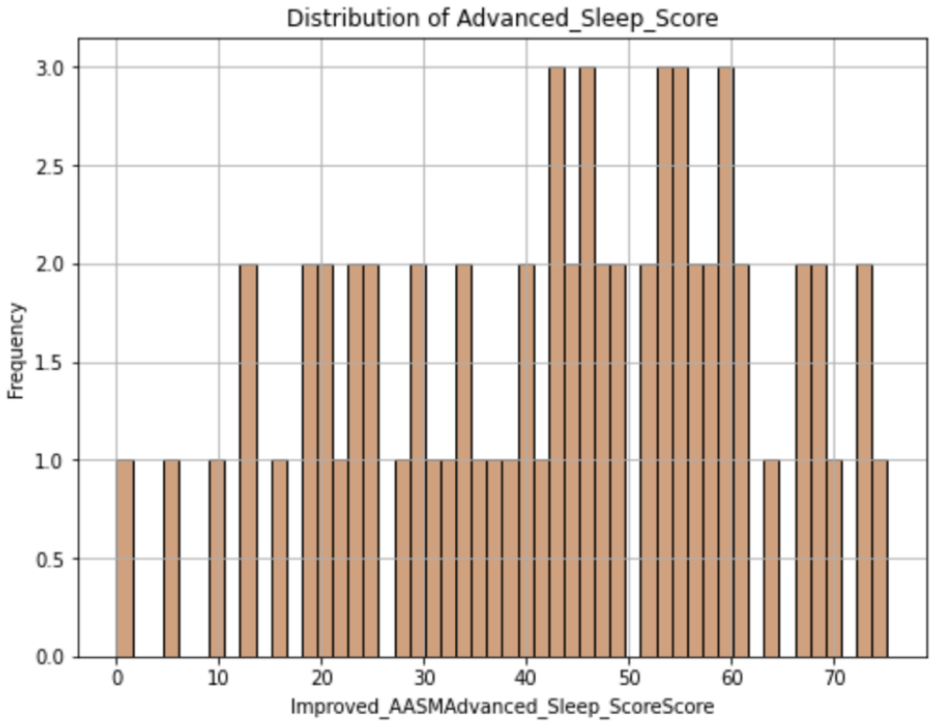
	Mean	Median	Standard Deviation	Variance	Minimum	Maximum
sleep_counts	9.348485	9.000000	2.820215	7.953613	3.000000	19.000000
time_in_bed	482.242424	487.250000	53.697675	2883.440326	285.500000	644.000000
REM_Total	99.909091	97.500000	22.997811	528.899301	40.500000	157.000000
Deep_Sleep_Total	52.227273	50.750000	17.669657	312.216783	7.000000	108.000000
Core_Sleep_Total	309.909091	312.750000	44.889686	2015.083916	144.500000	417.500000
REM_Sleep_percent	0.206350	0.208479	0.037829	0.001431	0.106472	0.308642
Deep_Sleep_percent	0.110640	0.102577	0.045655	0.002084	0.013725	0.330998
Core_Sleep_percent	0.641381	0.638158	0.051707	0.002674	0.506130	0.731452
Sleep_Efficiency	0.958370	0.967711	0.033124	0.001097	0.824635	0.994965



Based on that:

I created a model which scores sleep as they are under the distribution of the given parameters to score each sleep among each other. That approach also gave me a chance to classify the sleep sessions as good or bad.

Result of the scoring algorithm:



- I also added a categorical classifier which classifies sleep as good if the score is above mean and bad if the score is below mean

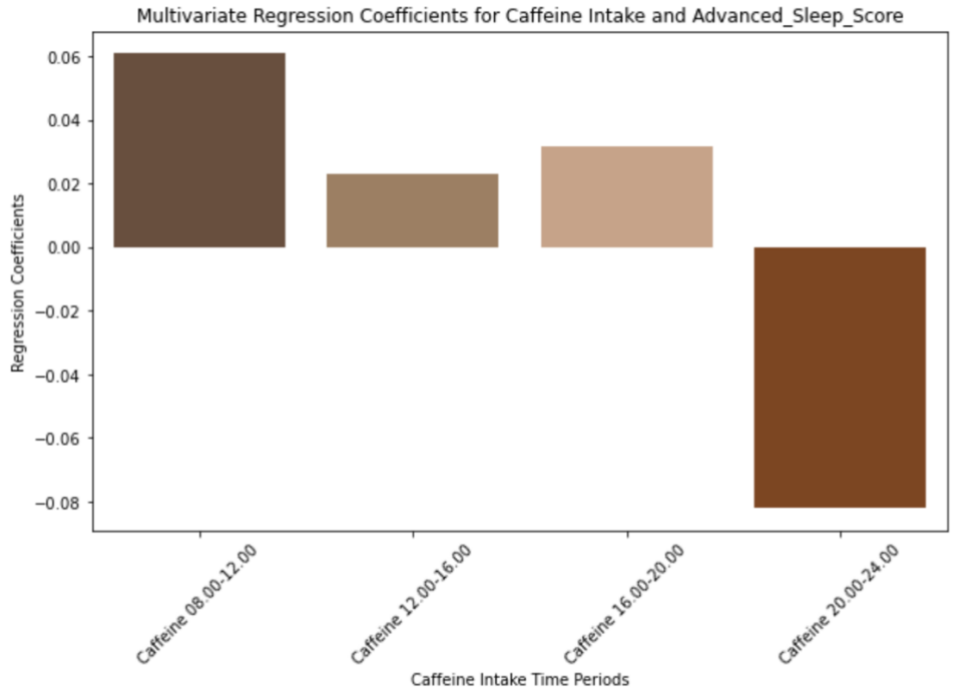
Train Models:

Multi Variable Linear Regression:

Method:

- Fit a linear regression with caffein consumption intervals ['Caffeine 08.00-12.00','Caffeine 12.00-16.00','Caffeine 16.00-20.00','Caffeine 20.00-24.00'] as independent variables and sleep score as dependent variable

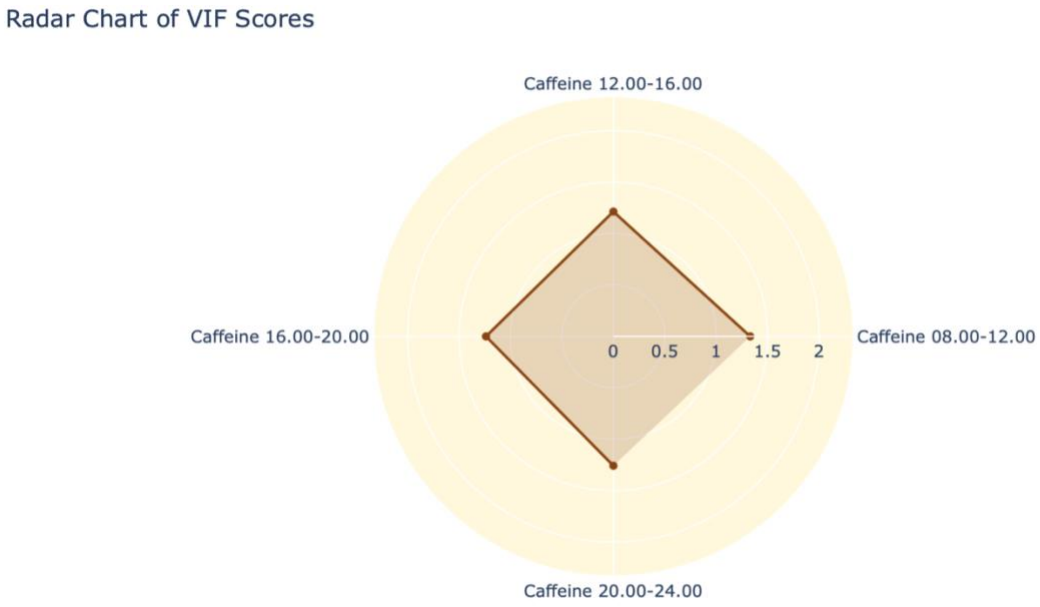
Multivariate Regression Coefficients for Caffeine Intake and Advanced_Sleep_Score:



Test The Model:

- R-squared: 0.088
- MAE: 14.926
- Minor Conclusion: There isn't a significant regression.

Analyze VIF Scores:

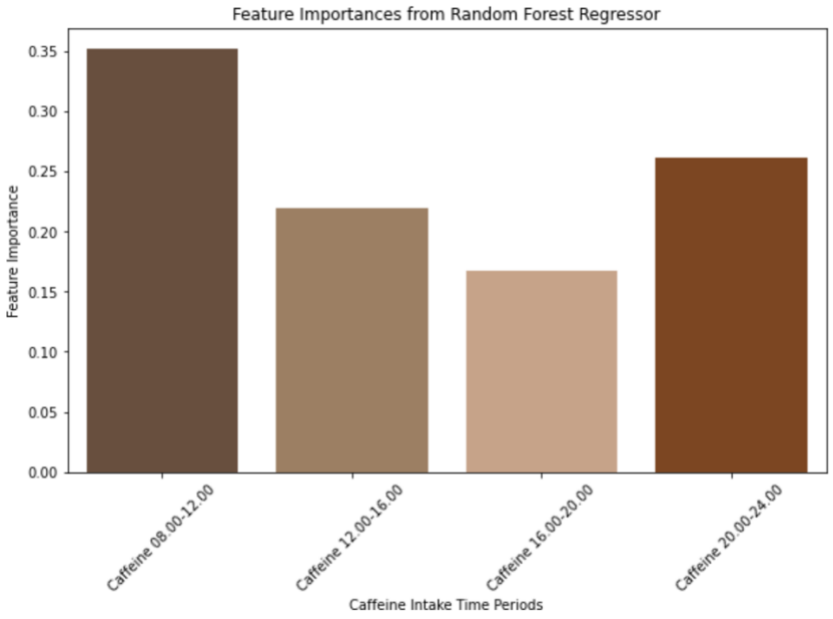


Random Forest Regression:

Method:

- Fit a random forest regression model using train test split as using 0.8 of the data to train and 0.2 of the data to test where caffein consumption intervals ['Caffeine 08.00-12.00', 'Caffeine 12.00-16.00', 'Caffeine 16.00-20.00', 'Caffeine 20.00-24.00'] as independent variables and sleep score as dependent variable.

Feature Importance of Random Forest Regressor:



Test The Model:

- R-squared: -0.64
- MAE: 20.35

R-squared being less than 0 means its more accurate to estimate the mean for each time. Which means the model fails to predict.

- Minor Conclusion: There is no significant linear regression.

Random Forest Classifier:

Method:

- Fit a Random Forest Classifier using train test split as using 0.8 of the data to train and 0.2 of the data to test where caffein consumption intervals ['Caffeine 08.00-12.00', 'Caffeine 12.00-16.00', 'Caffeine 16.00-20.00', 'Caffeine 20.00-24.00'] as independent variables and Weel_Sleep categorical parameter as dependent variable (Weel_Sleep is split as 1 and -1 based on the logic of sleep score above mean or below mean).
- Cross Validate
- Optimize Parameters
- Leave one out cross validate based on optimized parameters

Test The Model:

- Mean Accuracy: 0.44
- 3-Fold Cross-Validation Accuracy Scores: 0.38, 0.29, 0.64

Tune hyperparameters and cross validate with hyperparameters:

- Best Hyperparameters: {'max_depth': 5, 'min_samples_split': 3, 'n_estimators': 25}

Test with leave one out as data frame has not many entries:

- Mean Accuracy: 0.83

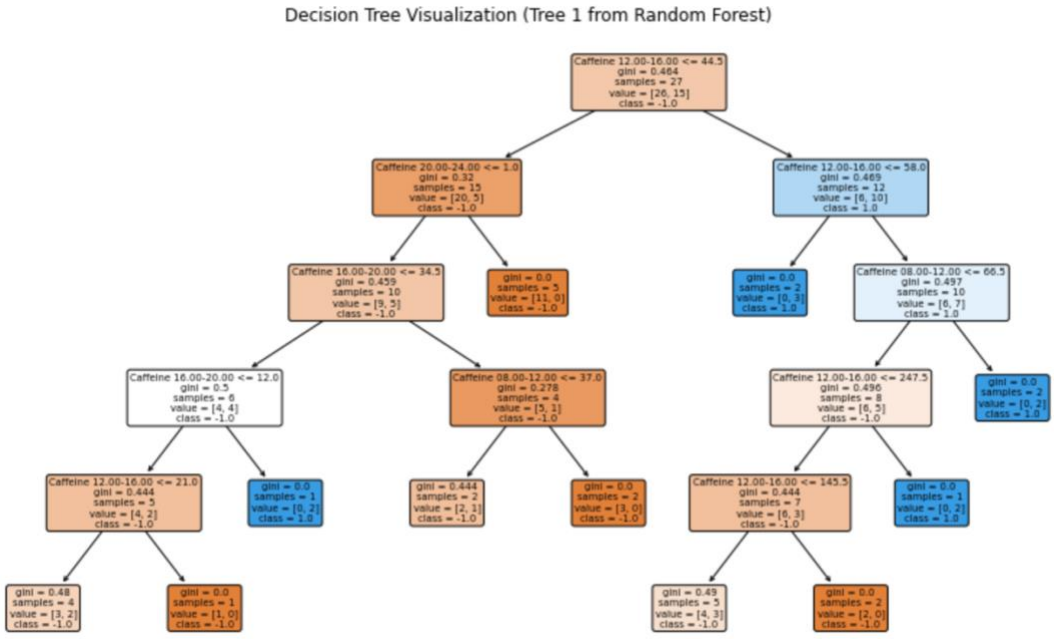
Test for Base Accuracy

- Based on the dataset there is 33 bad sleep instances and 33 good sleep instances which means we have %50 base accuracy. And our model has 0.83 accuracy. So, use p-value test to find out if we have an accurate model with 0.05 significance.

- Null Hypothesis (H0): The mean accuracy of the Random Forest Classifier is equal to the level of random chance (50% for binary classification).
- Alternative Hypothesis (H1): The mean accuracy of the Random Forest Classifier is significantly different from the level of random chance.

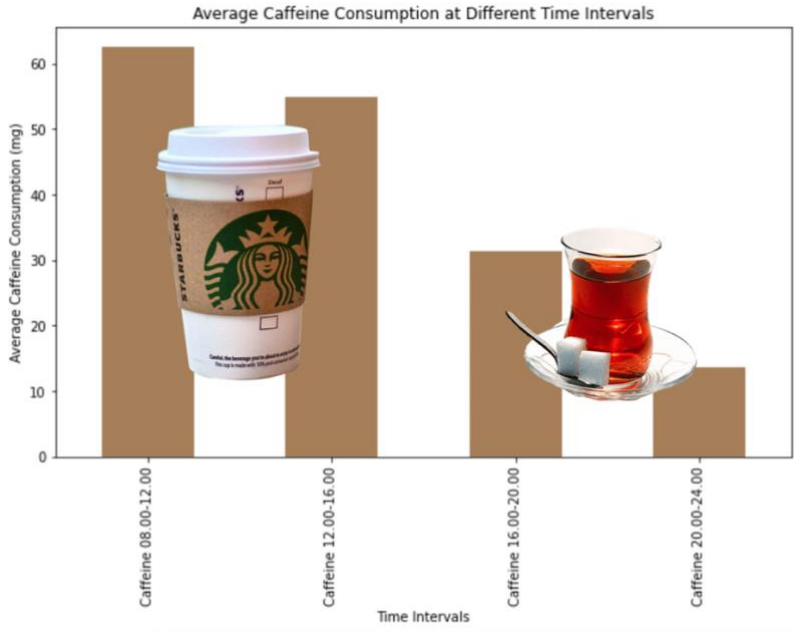
- T-statistic: 6.17132654441729
- P-value: 1.104494396483327e-07
- Reject the null hypothesis: There is a significant difference from chance level.

Example Decision Tree of the Classifier:



Conclusion:

- There isn't a simple significant regression between caffeine consumption and sleep quality. Sleep quality is a lot more complex to model than only one dependence of caffeine consumption. There might be other parameters such as stress level, alcohol consumption, painkillers etc. Based on that we can observe it is not possible to predict a significant sleep quality metric.
- However, things change as we aim to classify wither we sleep well or not based on caffein consumption, my trained model of Random Forest Classifier has 83 percent accuracy. This means that as a practice application I can use this model is to predict if it is okay to drink one more cup of coffee at a certain time of the day.
- A simpler outcome is that based on the average consumption of caffeine and my subjective opinion about my sleep that I'm happy with my overall sleep. If I won't go out to extremes and stick with my average coffee consumption distribution, I will most probably be okay with my sleep. Which indicates a tall Starbucks drink in the first half of the day, to wake up :) and a short drink or a cup of tea as a Turkish tradition.



Limitations:

- Lack of track of alcohol consumption. AASM defines alcohol as a major predictor of sleep quality.
- Ignorance of stress level.
- Ignorance of daily activity. (I was wearing apple watch only during nights, so I haven't got any significant activity tracking)

- Lack of indicating the time when I get into bed as Apple Sleep Mode activates automatically after an hour.

Future Work:

- Find the caffein halving time of a person based on their metabolism and implement a new data frame of caffeine level based on this halving time

Resources:

IBER, C., ANCOLI-ISRAEL, S., CHESSON JR., A. L., QUAN, S. F., (2007). The AASM Manual for Scoring of Sleep and Associated Events. Retrieved from https://www.sleep.pitt.edu/wp-content/uploads/2020/03/The-AASM-Manual-for-Scoring-of-Sleep-and-Associated-Events-2007-.pdf