**Project 6 Report - Relationship between Lifestyle Factors and Sleep Efficiency**

**ALY6000 71135 Introduction to Analytics**

**Professor: Dr. Arash Shamseddini**

**Team 7 Member: Honglin Wang, Jun Ma & Zihui Yao**

**Due Date: 10/27/2024**

Introduction

Sleep is an essential part of human health. It helps repair body, regenerate tissues, build bone, build muscle, and strengthen the immune system. Some important variables such as sleep duration, sleep efficiency, REM sleep percentage[1], deep sleep percentage, and light sleep percentage are highly related with sleep quality and health. Sleep efficiency is closely linked to both age and gender, with younger individuals typically exhibiting higher sleep efficiency than older adults. Furthermore, women often maintain slightly better sleep quality than men, but age-related sleep disturbances appear more pronounced in men[2]​. In addition, some factors such as alcohol consumption, smoking, and exercise frequency may also have impact on sleep quality.

In this report, a dataset containing 452 observations, and 14 variables is used for analyzing the relationship between some variables related to sleep quality. The variables in the dataset record the age, gender, bedtime, wakeup time, sleep duration, sleep efficiency, REM sleep percentage, deep sleep percentage, light sleep percentage, awakenings, caffeine consumption, alcohol consumption, smoking status, and exercise frequency.

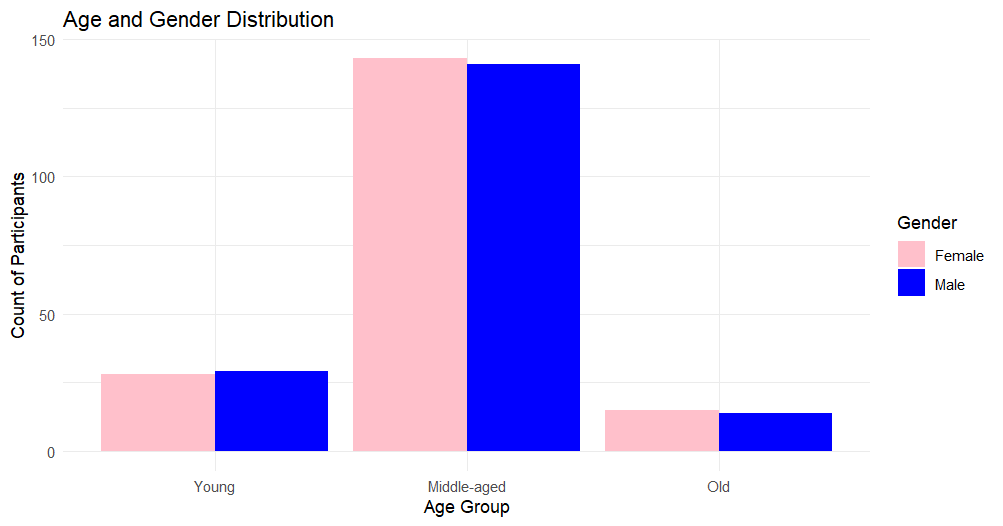
The objective of this study is to have a deeper understanding of the potential factors affecting sleep quality by conducting data analysis and statistical analysis on the collected dataset.

Methodology

The dataset is in csv format, and is loaded into R environment using `read.csv` function. After checking the missing values in this dataset using function `colSums` to calculate the number of missing values in each column, four columns (Awakenings, Caffeine Consumption, Alcohol Consumption, and Exercise Frequency) have a few missing values. Since the number of missing values is not too large, those observations containing missing values are directly removed using function `drop\_na()` from `tidyr` package. After removing missing values, outliers in column “Sleep Duration” and “Caffeine Consumption” were identify using IQR method. This method calculates the first quantile (Q1) and third quantile (Q3) of a variable, and then use Q3-Q1 to calculate IQR. Then, the lower bound is set to Q1-1.5\*IQR and the upper bound is set to Q3+1.5\*IQR. Those observations with value outside of the lower bound and upper bound were removed from the dataset. Finally, the cleaned data is used for creating visualizations using `ggplot` function from R package `ggplot2`. In addition, the `cut` function is used to create categorical variable from numerical variable by setting several cut points.

Result

In this project, several statistical analysis methods were conducted to analyze the dataset, and some meaningful plots were created for visualizing the dataset from different aspects. These plots including frequency distribution of categorical variable or numerical variable with only a few unique values, histogram displaying distribution of continuous variables, scatter plots with linear regression line for displaying the relationship between two numerical variables, and box plots for displaying the distribution of numerical variables grouped by a categorical variable.

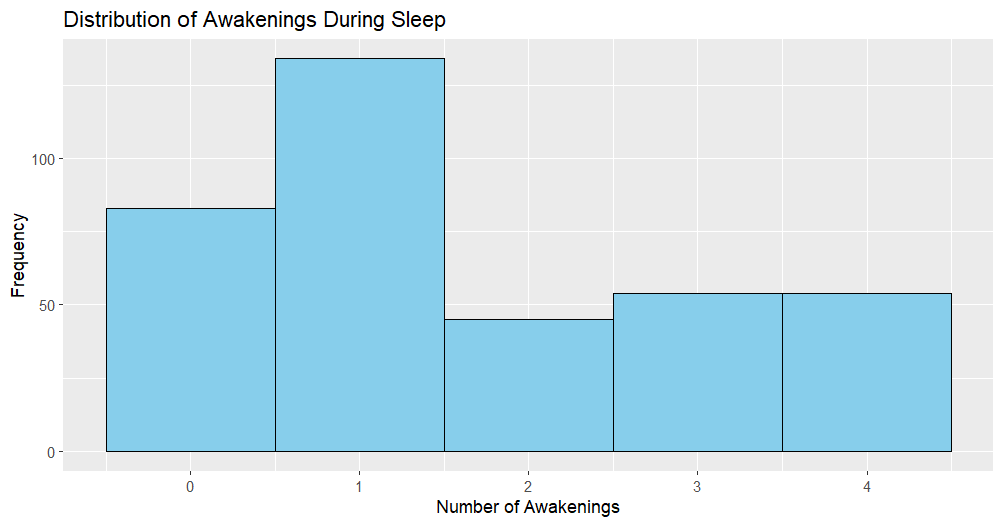
 A graph of a number of bars

Description automatically generated

*3.1 Distribution of Age Group (Categorical) and Gender 3.2 Distribution of Age (Continuous) by Gender*

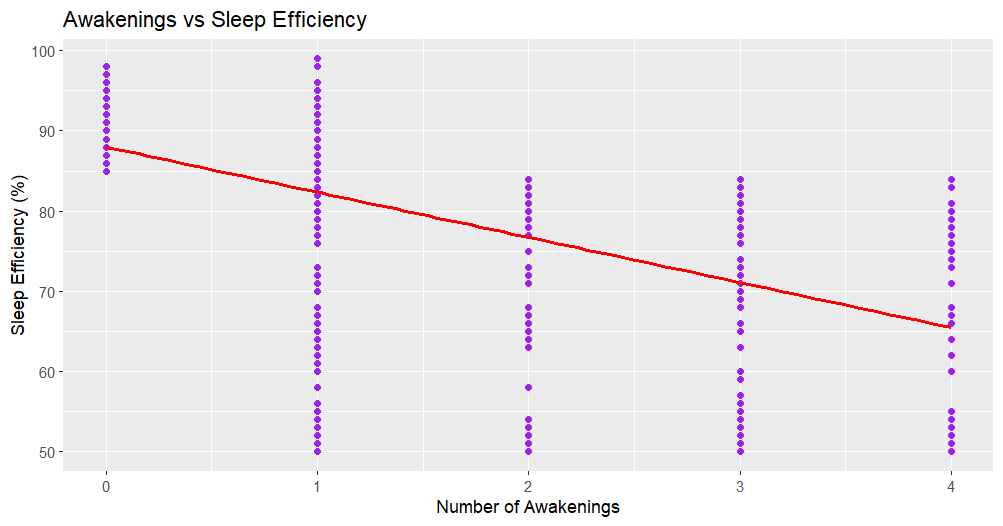
Figure 3.1 shows the distribution of age groups and gender, with age split into Young, Middle-aged, and Old. The large proportion of middle-aged individuals aligns with global trends, where the working-age population (15-64) comprises about 65% of the total population[3]. However, the gender distribution is inconsistent with global patterns, where males typically outnumber females at birth, and women predominate in the 65+ age group due to higher life expectancy[3]. Therefore, this set of experimental data is consistent with the global age distribution, but the gender ratio is not.

Figure 3.2 displays the distribution of age colored by gender. It can be observed that, the mean age of male participants is higher than female participant obviously. In addition, the age distribution of the male participants has two peaks, which are around 30 years and 55 years. This demonstrates that the age distribution of participants selected for analysis is not uniform.



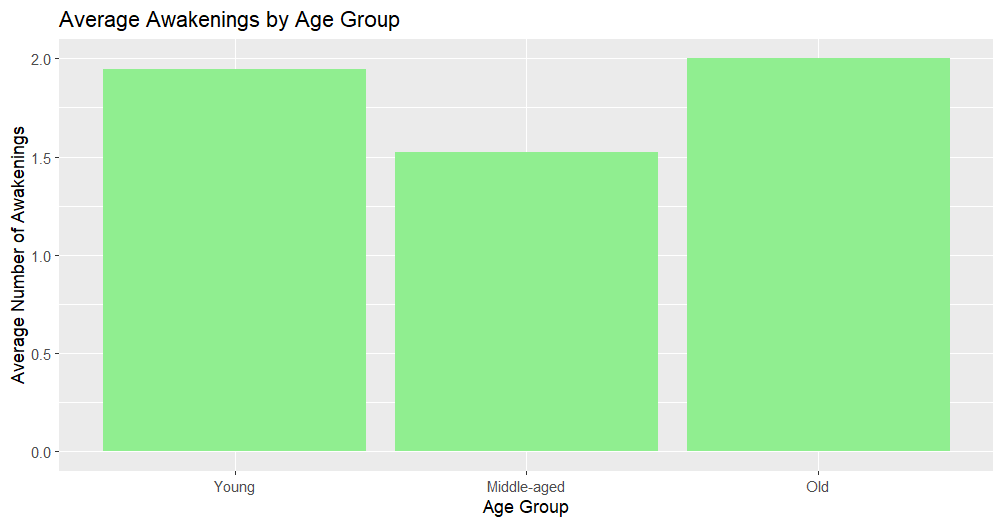
*3.3 Distribution of Awakenings During Sleep*

Figure 3.3 displays the distribution of number of awakenings during sleep. Obviously, most participants have one awakening during sleep. However, there are also a large percent of participants have more than two awakenings during sleep.



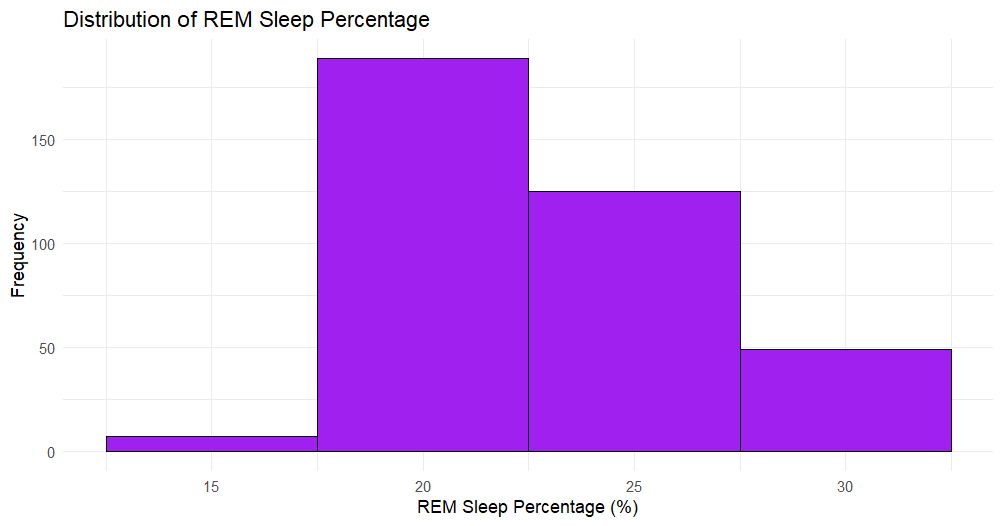
*3.4 Awakenings vs Sleep Efficiency*

Sleep efficiency is the ration between the time a person spends asleep, and the total time dedicated to sleep. It is an important indicator. A sleep efficiency of 80% or more is considered normal[4]. Figure 3.4 displays the relationship between sleep efficiency and awakenings. According to the red linear regression line, there is an apparent negative relationship between the two variables. Such results indicate the sleep efficiency significantly decreases as number of awakenings increases.



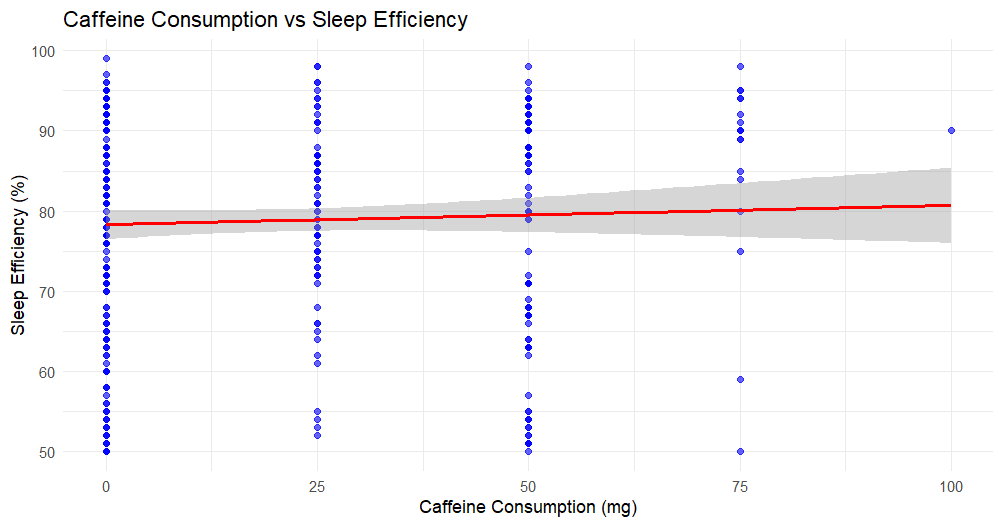
*3.5 Average Awakenings by Age Group*

Figure 3.5 displays the average number of awakenings by age group, and we can observe that young participants and old participants have higher number of awakenings (close to 2), and middle-aged participants have lower number of awakenings (close to 1.5).



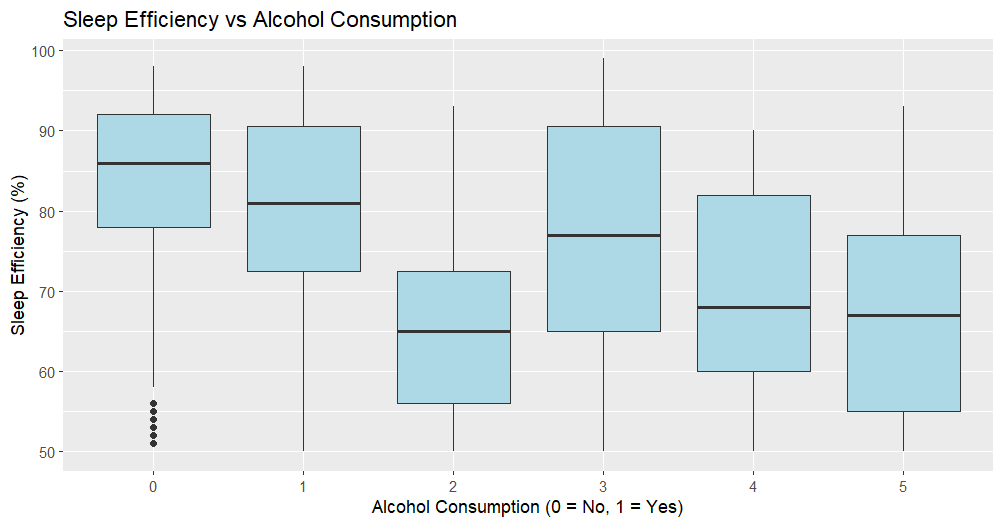
*3.6 Distribution of REM Sleep Percentage*

Figure 3.6 presents the distribution of REM sleep percentages among participants. Most participants exhibit a REM sleep percentage between 20% and 25%, which aligns with typical REM sleep patterns. A small number of participants have a lower REM percentage, around 15%. REM sleep, or Rapid Eye Movement sleep, plays a crucial role in cognitive functions such as memory consolidation. Normally, REM sleep constitutes about 20% to 25% of total sleep time in adults, during which brain activity is similar to when awake, and dreams commonly occur[5].



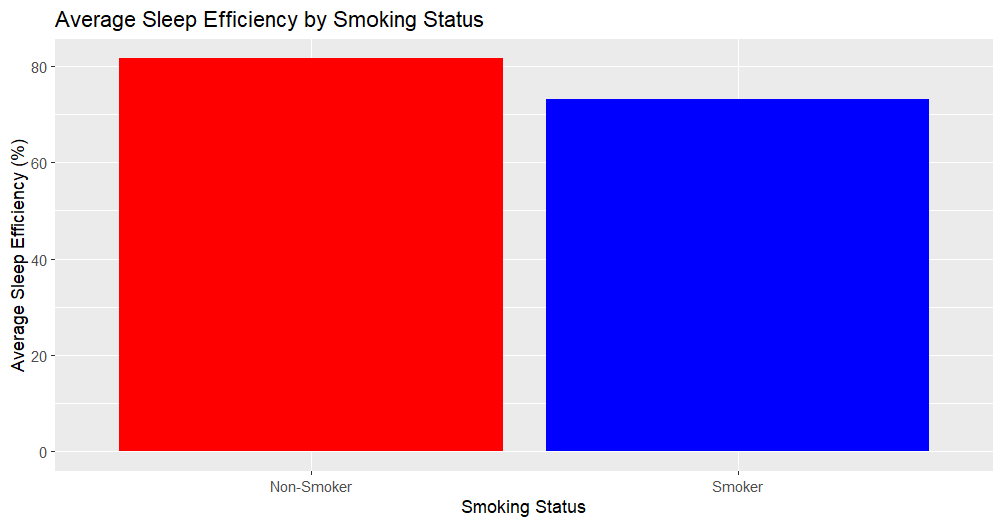
*3.7 Caffeine Consumption vs Sleep Efficiency*

In Figure 3.7, the plotted data shows no strong correlation between caffeine consumption and sleep efficiency, which is surprising considering the general consensus that caffeine disrupts sleep quality. The red regression line remains almost flat, suggesting minimal impact of caffeine intake on sleep efficiency across the sample. However, previous research highlights that caffeine can reduce sleep efficiency by approximately 7% when consumed within proximity to bedtime, particularly affecting total sleep duration, onset latency, and increasing light sleep[6]. This discrepancy may suggest that individual differences or caffeine tolerance could be factors in the varying results.



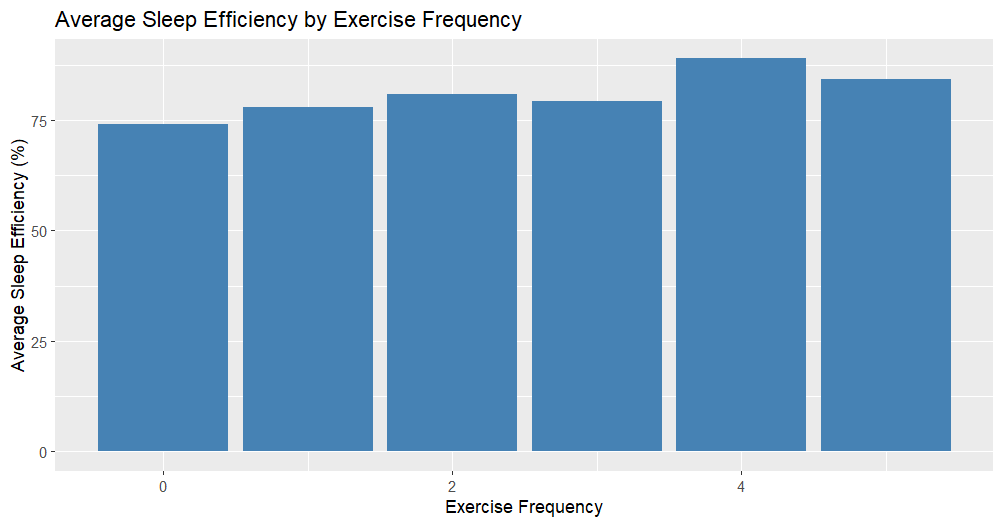
*3.8 Sleep Efficiency vs Alcohol Consumption*

Figure 3.8 displays a boxplot illustrating the relationship between alcohol consumption and sleep efficiency. As observed, individuals who do not consume alcohol (category 0) have the highest sleep efficiency, with a median around 90%. As alcohol consumption increases, there is a noticeable decline in sleep efficiency, especially for categories 2 and 5, where the median drops closer to 70%. This aligns with research indicating that alcohol disrupts sleep, particularly by reducing REM sleep, which is essential for restful sleep and cognitive function[7]. Excessive drinking before sleep is not advised, as it leads to poorer sleep efficiency and fragmented sleep patterns.



*3.9 Average Sleep Efficiency by Smoking Status*

Figure 3.9 compares the average sleep efficiency between smokers and non-smokers. We can see non-smokers have relatively higher sleep efficiency (higher than 85%), and those smokers have lower sleep efficiency (lower than 75%). This pattern supports research showing that smoking negatively impacts sleep quality, reducing overall sleep efficiency due to nicotine's stimulating effects on the body and its disruption of normal sleep patterns[8].



*3.10 Average Sleep Efficiency by Exercise Frequency*

Figure 3.10 explores the relationship between average sleep efficiency and exercise frequency. As shown in the figure, sleep efficiency increases with higher exercise frequency, with participants who exercise more frequently (2 to 4 times per week) exhibiting improved sleep efficiency. This is consistent with findings in sleep research, which show that regular physical activity enhances sleep quality, particularly by reducing the time taken to fall asleep and increasing sleep efficiency[9]. Exercise has been shown to have positive effects on both sleep onset and duration, further highlighting its role in improving overall health and well-being through better sleep.

Conclusion

The analysis of the sleep dataset shows that sleep quality is influenced by several lifestyle factors. Key findings indicate that sleep efficiency decreases with more awakenings and alcohol consumption, while increased exercise frequency is linked to better sleep. Interestingly, caffeine consumption shows minimal correlation with sleep efficiency, possibly due to individual differences in tolerance or dependence, which may reduce its typical disruptive effects on sleep quality, especially when caffeine is consumed close to bedtime[10]. Additionally, young and old participants tend to experience more awakenings on average compared to middle-aged participants.

In summary, this study suggests that maintaining a healthy lifestyle can significantly improve sleep quality, contributing to overall health and well-being. It advocates for the adoption of healthy habits, such as regular exercise and moderate alcohol consumption, to enhance sleep efficiency and, by extension, overall well-being. Additionally, for individuals who smoke and experience sleep issues, attempting to quit smoking or reducing their cigarette intake[11] is a crucial step towards maintaining good health and achieving a higher quality of life. Lastly, as age advances, sleep efficiency tends to diminish. For the elderly, engaging in regular physical activity[12] becomes particularly important to improve sleep quality.

Reference

1. H. R. Colten, B. M. Altevogt, and I. of M. (US) C. on S. M. and Research, “Sleep Physiology,” in *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*, National Academies Press (US), 2006. <https://www.ncbi.nlm.nih.gov/books/NBK19956/>
2. Sophie Desjardins, Sylvie Lapierre, Carol Hudon, Alain Desgagné, Factors involved in sleep efficiency: a population-based study of community-dwelling elderly persons, *Sleep*, Volume 42, Issue 5, May 2019, zsz038, <https://doi.org/10.1093/sleep/zsz038>
3. Our World in Data. (2022). *Gender ratio*. Our World in Data. <https://ourworldindata.org/gender-ratio>
4. Sleep WA. (n.d.). *Interpreting a sleep study*. Sleep WA. <https://www.sleepwa.com.au/interpreting-a-sleep-study/>
5. Cleveland Clinic. (n.d.). Sleep basics. Cleveland Clinic. <https://my.clevelandclinic.org/health/body/12148-sleep-basics>
6. Gardiner, C., Weakley, J., Burke, L. M., Roach, G. D., Sargent, C., Maniar, N., Townshend, A., & Halson, S. L. (2023). The effect of caffeine on subsequent sleep: A systematic review and meta-analysis. *Sleep Medicine Reviews*, 69, 101764. <https://doi.org/10.1016/j.smrv.2023.101764>
7. Piedmont Healthcare. (n.d.). *How does alcohol affect your sleep?* Piedmont Healthcare. <https://www.piedmont.org/living-real-change/how-does-alcohol-affect-your-sleep>
8. Frontiers in Psychiatry. (2024). *The impact of smoking on sleep quality and efficiency*. <https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2024.1407741/full>
9. Tandfonline. (2019). *The relationship between physical activity and sleep quality*. <https://www.tandfonline.com/doi/full/10.1080/21679169.2019.1623314#abstract>
10. Kerpershoek, E., Antypa, N., Van Den Berg, M., & Van Der Hulst, M. (2018). Caffeine intake and sleep quality in adults. *Journal of Sleep Research*, 27(5), e12625. <https://doi.org/10.1111/jsr.12625>
11. Lohsoonthorn, V., Khidir, H., Casillas, G., Lertmaharit, S., Tadesse, M. G., Pensuksan, W. C., Rattananupong, T., Sonkprasert, T., & Williams, M. A. (2022). Sleep quality and sleep patterns in relation to consumption of energy drinks, caffeinated beverages, and other stimulants among Thai college students. *Sleep Epidemiology*, 2, 100009. <https://doi.org/10.1016/j.sleepe.2022.100009>
12. Del Pozo Cruz, B., Mañas, A., Martín-García, M., Marín-Puyalto, J., García-García, F. J., Rodríguez-Mañas, L., & Guadalupe-Grau, A. (2020). Frailty and multimorbidity patterns in the elderly and their relation with physical activity and sitting time: An isotemporal substitution approach. *International Journal of Behavioral Nutrition and Physical Activity*, 17, 111. <https://doi.org/10.1186/s12966-020-0913-3>