This is the main submission document. ***Save and rename this document filename with your registered full name as Prefix before submission.*** *Submit both this word document and PDF format.*

Class	1
Full Name	CHEN WEICHEN
University Email	N2301599J@e.ntu.edu.au

** : Delete and replace as appropriate.*

Declaration of Academic Integrity

By submitting this assignment for assessment, I declare that this submission is my own work, unless otherwise quoted, cited, referenced or credited. I have read and understood the Instructions to CBA.PDF provided and the Academic Integrity Policy.

I am aware that failure to act in accordance with the University's Academic Integrity Policy may lead to the imposition of penalties which may include the requirement to revise and resubmit an assignment, receiving a lower grade, or receiving an F grade for the assignment; suspension from the University or termination of my candidature.

I consent to the University copying and distributing any or all of my work in any form and using third parties to verify whether my work contains plagiarised material, and for quality assurance purposes.

Please insert an "X" within the square bracket below to indicate your selection.

[X] I have read and accept the above.

Table of Contents

Answer to Q1:

Figure 1.0 indicates that there is a minor negative link between the percentage of deep sleep and the number of daily steps; that is, as the number of daily steps grows, the percentage of deep sleep slightly drops. The data points were mostly centered in the midpoint range (about 4000 to 8000 steps) and around 60% of the deep sleep percentage, with few outliers below the 40% deep sleep percentage. Nevertheless, this association was not significant. This implies that while step count is a substantial role, other factors including age, food, stress levels, and kind of daily activity might also have a big impact on deep sleep. Therefore, relying just on step count may not be sufficient to evaluate or enhance sleep quality; instead, a number of elements and healthy behaviors must be taken into account in order to improve sleep quality overall.

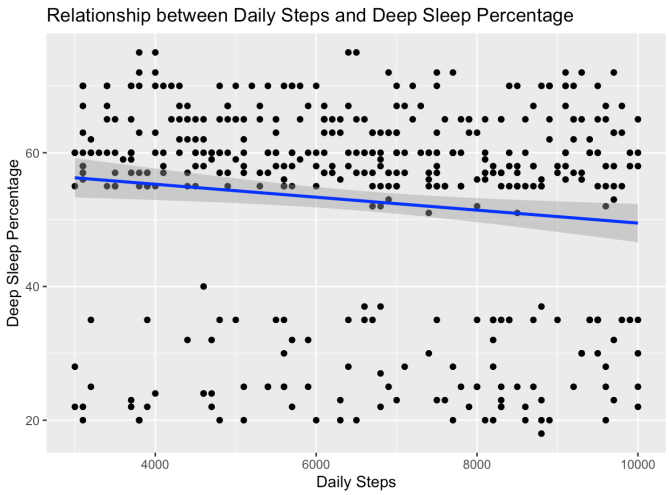

Figure 1.0

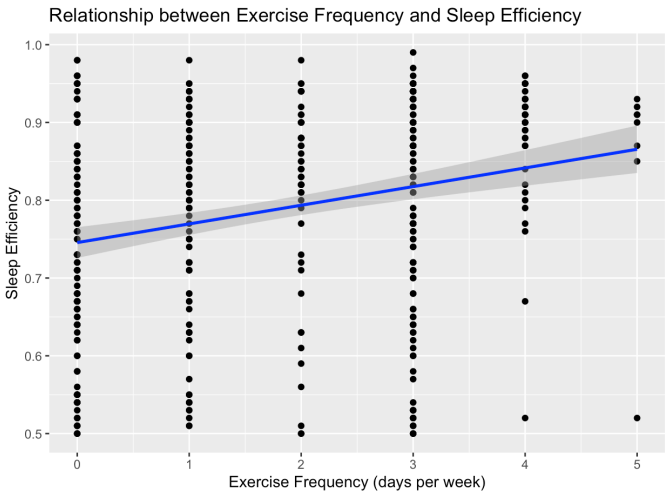

Figure 1.1

From figure 1.1 We can see from the graph that there isn't a meaningful linear correlation between the amount of sleep and how well we sleep. Regardless of the length of sleep—six, seven, or eight hours—the total sleep efficiency stays at 0.8. This shows that just lengthening or reducing sleep time does not, for the majority of people, appreciably alter sleep efficiency. The data distribution revealed that the majority of persons had sleep efficiencies ranging from 0.75 to 0.85, which suggests that overall sleep quality was largely consistent. Individual variations persist, though, indicating that optimizing sleep quality would necessitate taking into account additional variables in addition to sleep length, such as medical issues and lifestyle choices.

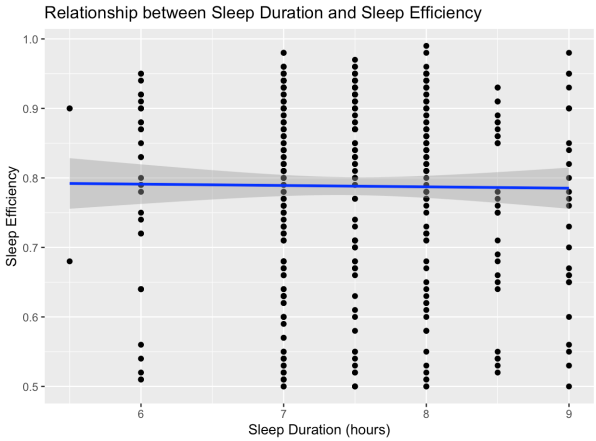

Figure 1.2

Based on the figure 1.2 , we discovered a strong positive relationship between the frequency of weekly exercise and the effectiveness of sleep. Sleep efficiency rises considerably along with the frequency of weekly activity. According to the statistics, the majority of people sleep between 0.75 and 0.85, with those who exercise more frequently—four to five days a week—sleeping at the highest efficiency. A bar graph that illustrates how average sleep efficiency rises with increased exercise frequency lends more credence to this conclusion. It is thus advised to maintain a frequency of exercise at least three to four days a week for maximum sleep efficiency and general health. Regular exercise not only helps to enhance physical health but also greatly improves sleep quality.

Answer to Q2:

Research Question

How do lifestyle factors (exercise frequency, caffeine consumption, alcohol consumption, and daily steps) predict sleep efficiency in adults?

Target Variables:

Sleep Efficiency

Sleep efficiency, or the ratio of time in bed to time sleeping, is a critical measure of the quality of sleep. Examining the variables that predict sleep efficiency might reveal important information about how lifestyle choices affect both the quality of sleep and general health.

Answer to Q3:

	Model	RMSE	MAE	R_squared
1	Linear Regression	0.1221005	0.1018471	0.2161053

Figure 2.0

	Model	RMSE	MAE	R_squared
1	Pruned CART	0.1180559	0.09778667	0.2657113

Figure 2.1

In the current study, we evaluated how well the linear regression model and the pruned CART model predicted sleep efficiency. We discovered that the pruned CART model performed better than the linear regression model on all three indicators after computing and comparing the RMSE, MAE, and R-squared metrics. In particular, the Pruned CART model's low RMSE and MAE show that its prediction error is small. Its R-squared value is also greater, suggesting a superior capacity for data interpretation. Therefore, it can be concluded that the Pruned CART model performs best in predicting sleep efficiency

Answer to Q4:

How do lifestyle factors (exercise frequency, caffeine consumption, alcohol consumption, and daily steps) predict sleep efficiency in adults?

Based on the results of the CART model and linear regression model after pruning, we found that age, exercise frequency, and daily step count were the most important factors in predicting sleep efficiency in adults. Specifically, age had the greatest impact on sleep efficiency, suggesting that sleep patterns change significantly as we age. Regular exercise and higher daily steps are strongly associated with better sleep quality, emphasizing the importance of maintaining an active lifestyle. Although alcohol and caffeine consumption also have an impact on sleep efficiency, the effect is relatively small. Moderate alcohol consumption and reduced caffeine intake, especially before bedtime, can still help improve sleep efficiency.

Analysis of linear regression models further supported these findings, showing that alcohol consumption had a significant negative impact on sleep efficiency, while exercise frequency and age had a significant positive effect on sleep efficiency. Although the negative effect of daily steps on sleep efficiency is small, its statistical significance suggests that moderate attention is required. Overall, these results highlight the critical role of maintaining a healthy lifestyle in improving sleep quality. Combining the results of these two models, we can more fully understand and optimize the lifestyle factors that influence sleep efficiency.

Answer to Q5:

Factor	Importance Score
Age	0.3744810
Alcohol consumption	0.2644895
Caffeine consumption	0.1204813
Daily Steps	0.2879988
Exercise frequency	0.3528406

Based on the results of the CART model after pruning, we found that age, exercise frequency, and daily step count were the most important factors in predicting sleep efficiency in adults. Specifically, the importance of age score was 0.3744810, indicating that sleep patterns changed significantly with age. The importance of exercise frequency score was 0.3528406, indicating that people who exercise regularly generally have better sleep quality. With an importance score of 0.2879988, higher daily steps are strongly associated with higher sleep efficiency, emphasizing the importance of maintaining an active lifestyle. Although the importance scores of alcohol and caffeine consumption were 0.2644895 and 0.1204813, respectively, indicating that they had little effect on sleep efficiency, moderate alcohol consumption and reduced caffeine intake, especially before bedtime, still helped to improve sleep efficiency. These results highlight the important role of a healthy lifestyle in improving sleep quality.

```
                         Estimate   Std. Error   t value     Pr(>|t|)
(Intercept)             7.532832e-01 3.382494e-02 22.270050 2.261533e-65
`Alcohol consumption`  -2.728471e-02 4.396146e-03 -6.206506 1.811466e-09
Age                     1.463444e-03 5.339025e-04  2.741033 6.494542e-03
`Exercise frequency`    2.678485e-02 4.748634e-03  5.640538 3.939523e-08
`Caffeine consumption`  4.798638e-04 2.841622e-04  1.688697 9.232361e-02
`Daily Steps`          -7.765095e-06 3.344985e-06 -2.321414 2.093882e-02
```

According to the results of the linear regression model, age, exercise frequency, and alcohol intake are important factors in predicting sleep efficiency. The regression coefficient of alcohol intake was -0.02728, indicating that sleep efficiency was significantly reduced with each additional unit of alcohol intake ($P < 0.001$). The regression coefficient for age was 0.001463, indicating that sleep efficiency increased with each additional year ($p < 0.01$). The regression coefficient of exercise frequency was 0.02678, indicating that increasing exercise frequency significantly improved sleep efficiency ($P < 0.001$). The regression coefficient for daily steps was -7.765e-06, although there was a slight negative effect ($p < 0.05$), but the effect was small. The regression coefficient of caffeine intake was 0.0004799, and the effect was not significant ($P \approx 0.09$). Overall, maintaining an active lifestyle (such as increasing exercise frequency) and moderate alcohol consumption are essential for improving sleep efficiency, although daily step count and caffeine intake have a small impact.

Answer to Q6:

The model developed in this study can be used to provide personalized health recommendations aimed at improving sleep efficiency. By analyzing an individual's lifestyle data, such as age, exercise frequency, daily steps, caffeine intake, and alcohol intake, the model can predict their sleep efficiency and provide customized recommendations to improve sleep quality.

Health & Fitness Apps:

Integration: The model can be integrated into health and fitness apps that track the user's activities and habits.

Data collection: These apps already collect daily steps, exercise habits, and even dietary habits such as caffeine and alcohol intake.

Recommendations: With models, apps can analyze this data and provide personalized recommendations. For example, if the model predicts inefficient sleep due to high caffeine intake or lack of exercise, the app can notify the user and provide specific recommendations for improvement, such as reducing caffeine intake or increasing physical activity.