Homework 4
Problem 3
Ayad Masud
733009045

I certify that I have personally done the coding, generated the figures and written the report without aid from anybody else, and that I have not plagiarized, self-plagiarized, or used AI-generated text. I certify that I have acknowledged any sources I used to complete this assignment. ARM.

1 Part 1: Data Collection

Data for this research project was collected using several different instruments and was relatively straightforward. Most importantly the heartrate and sleep time data were collected using a smart watch. This data was collected every day for around 3 months and was then put into a excel spreadsheet to track. Specifically, I would record the heartrate of the exercise that day and then the following morning after I have slept, I would record the sleep quality consequent of the exercise from the previous day. The sleep quality for the experiment was measured using the Karolinska Sleepiness Scale to rate the sleep quality out of 10. A copy of my dataset is attached in the zip file.

2 Part 2: Data Analytics

The research question I wanted to answer with this project is how my exercise habits affect my sleep quality. I wanted to know what type of exercise schedule I should create for myself to optimize my sleep. Getting into the analysis figure 1 is a correlation matrix between heartrate, sleep quality, and sleep time. We can see that the only meaningful correlation in the matrix is between sleep quality and heartrate, which has a positive correlation. So, when heartrate is elevated, or high sleep quality is better. There is a slight positive correlation between sleep time and sleep quality and almost zero correlation between sleep time and heartrate. Figure 2 shows a box plot for each heart rate zone I trained at and the average quality of sleep that resulted from each respective heartrate zone. Most of my exercise took place in a moderate range between 120-160 bpm. We can see an upward trend, as heartrate goes up the sleep quality also goes up. This matches the original hypothesis I made when I began this experiment. This makes sense because higher heartrate most of time means outputting more effort and energy during exercise, which in turn makes me very tired and sleepy for bed that night. Figure 3 is another visual representation to show the linear nature of heartrate and how that corresponds with sleep quality. We can see more clearly in this diagram with the clusters that within approximately 80-110 bpm I had the worst quality sleep, around 5.75 on the Karolinska sleepiness scale which means some signs of sleepiness. I had the best quality sleep at around 170-190 bpm or high intensity exercises. That had an average sleep score of 7.5 which is labeled as alert on the Karolinska sleepiness scale. I also conducted a hypothesis test with the following hypothesis. The null hypothesis is that heartrate has no significant effect on sleep quality and the alternative hypothesis is that heartrate does have a significant effect on sleep quality. After conducting the hypothesis test using python functions, the p-score came out to 0.000007 which is less than the 0.05 significant threshold. Therefore, we can say reject the null hypothesis and say that heartrate does have a significant effect on sleep quality. As a conclusion I now know that intense exercise will give me really good quality sleep, so I may start doing intense exercise the day before an exam, so I sleep well and wake up refreshed and full of energy for the exam.

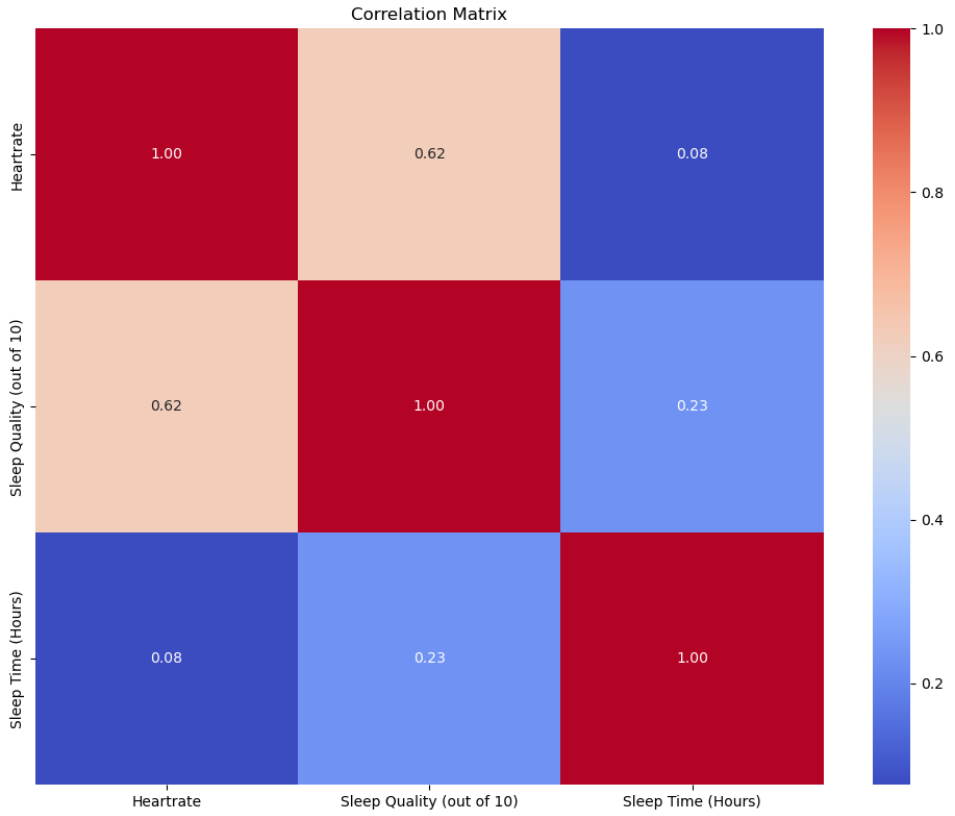

Figure 1. Correlation matrix between heartrate, sleep quality (out of 10), and sleep time (hours)

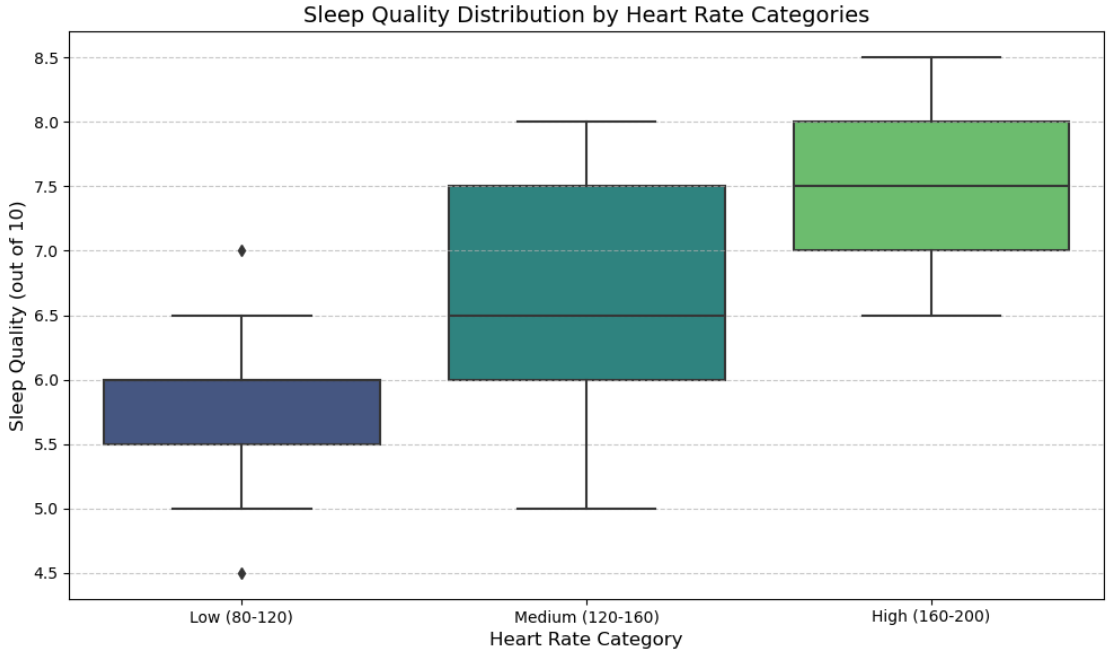

Figure 2. Sleep quality distribution by heart rate categories

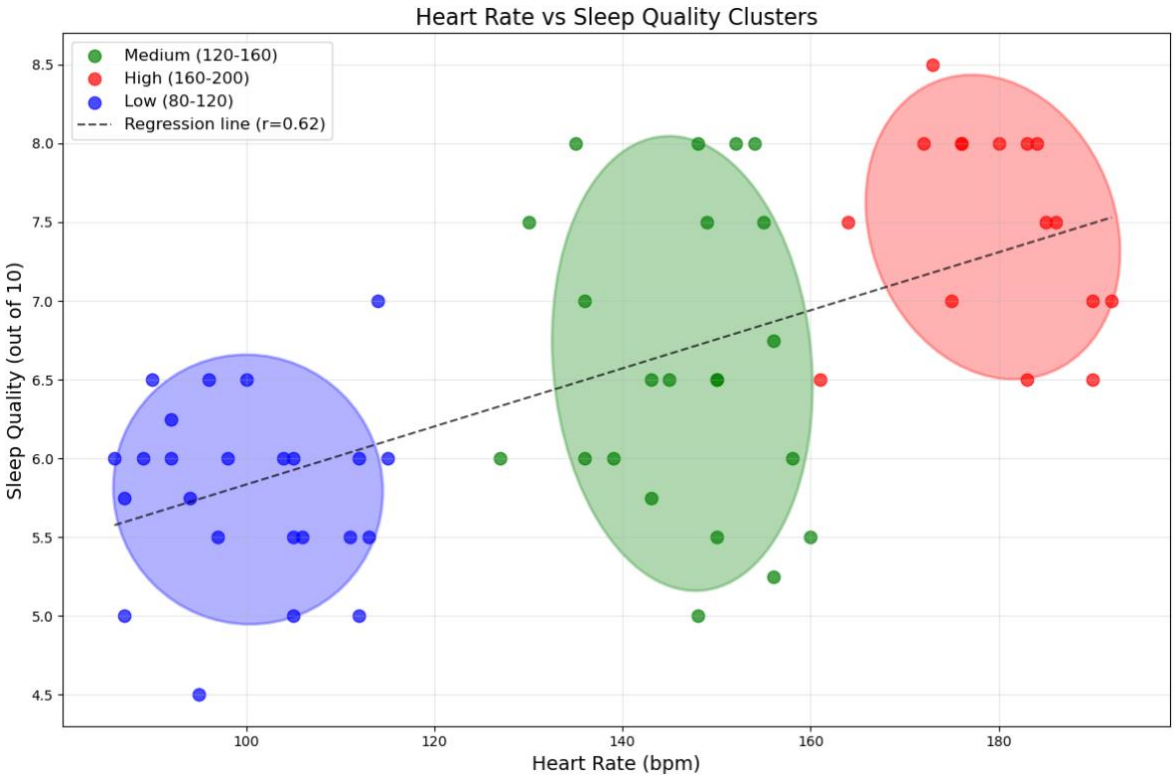

Figure 3. Heart rate vs sleep quality clusters.

3 Resources used to achieve this goal

Canvas: Homework template

Python Library: NumPy, pandas, seaborn, matplotlib, sci-kit learn

4 References

scikit-learn. “Sklearn.metrics.confusion_matrix — Scikit-Learn 0.21.3 Documentation.” *Scikit-Learn.org*, 2019, scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion_matrix.html.