**Sleep Data Analysis - DSA210 Term Project**

**Motivation**

In today’s fast-paced, technology-driven world, sleep is often neglected due to excessive screen time, academic pressures, and varying physical activity levels. As a student, I noticed significant changes in my sleep patterns, especially during midterm exam periods and on days with high TikTok usage. This project aims to investigate the impact of TikTok activity (daily login count and message count), step count, and midterm exam days on sleep quality, using data collected from my Apple Watch and iPhone Health app.

**Data Sources**

The project relies on two types of data collection methods: automated and manual. Automated data includes sleep quality metrics, daily step counts, and TikTok usage data, such as login and message counts. These were recorded using my Apple Watch and iPhone Health app. Sleep metrics include total sleep hours, REM sleep duration, and an overall sleep quality score. Manual data involves recording midterm exam dates in a JSON file to analyze the effect of academic stress on sleep patterns.

**Dataset Overview**

The dataset consists of the following variables:

• **Date**: The day the data was recorded.

• **TikTok\_Login\_Count**: Number of times TikTok was accessed.

• **Sleep\_Quality**: A numerical measure of overall sleep quality (0–100).

• **Step\_Count**: The number of steps taken each day.

• **TikTok\_Message\_Count**: Total number of messages sent on TikTok.

• **Exam\_Day**: A binary variable indicating whether a day was a midterm exam day (1 for exam day, 0 otherwise).

**Methodology**

The project followed a structured methodology, beginning with data preprocessing to clean and normalize the dataset. Exploratory Data Analysis (EDA) was conducted to identify trends, patterns, and anomalies. Statistical analyses, including Pearson and Spearman correlations, were performed to quantify the relationships between variables. A t-test was used to evaluate the impact of exam days on sleep quality. Visualization techniques such as time-series plots, scatter plots, bar charts, and heatmaps were used to illustrate key findings and support the analysis.

**Analysis and Findings**

The analysis revealed significant insights into the relationship between digital activity, physical activity, academic stress, and sleep quality. TikTok usage was found to negatively impact sleep quality, with higher login and message counts correlating with reduced sleep hours and poorer REM sleep duration. Step counts, on the other hand, positively influenced sleep quality, highlighting the importance of physical activity in promoting restorative sleep. Exam days showed a clear negative impact on sleep quality, likely due to heightened stress and irregular schedules.

Statistical tests supported these observations. The Pearson correlation coefficient between sleep quality and TikTok login count was -0.71, indicating a strong negative relationship. A positive correlation of 0.26 was observed between step count and sleep quality. The t-test for exam days versus non-exam days showed a statistically significant difference, further emphasizing the adverse effect of academic stress on sleep.

Visualizations enhanced the understanding of these findings. Time-series plots captured daily fluctuations in sleep quality, while heatmaps illustrated correlations between variables. Scatter plots highlighted the negative relationship between TikTok usage and sleep quality, while bar charts compared sleep quality on exam and non-exam days.

**Limitations**

The project has some limitations. The dataset is limited to my personal records, which may not generalize to a broader population. Important external factors such as diet, caffeine intake, and stress levels were not included, although they likely influence sleep quality. Additionally, the accuracy of Apple Watch and iPhone sensors could introduce measurement errors.

**Future Work**

To address these limitations, future work will focus on expanding the data collection period to improve the robustness of the analysis. Additional variables, such as diet, caffeine consumption, and stress levels, will be incorporated to provide a more holistic view of the factors affecting sleep quality. Advanced techniques, including machine learning algorithms, will be employed to predict sleep quality based on multiple variables, such as TikTok activity, step count, and midterm schedules. These improvements will make the findings more generalizable and actionable.

**Conclusion**

This project demonstrates the interconnectedness of digital activity, physical activity, academic stress, and sleep quality. The findings confirm that excessive TikTok usage and midterm exam stress negatively impact sleep, while physical activity promotes better sleep quality. By identifying these patterns, this project provides actionable insights for improving sleep habits through better time management, reduced screen time, and increased physical activity. Future efforts will aim to refine these insights and explore predictive models to help individuals optimize their sleep and overall well-being.