

Progress Report #1

Group 9

Eric Wu

Dongyeon Kang

Sunwoo Kim

Introduction

In our pursuit of creating a dynamic and informative website, our primary goal is to construct a functional platform that fully utilizes the potential of data visualization. Central to our project is the utilization of the "Sleep health and lifestyle" dataset, obtained from Kaggle.com. Our goal is twofold: firstly, to transform this dataset into an engaging and accessible resource, making it interesting and user-friendly so that it can be used easily. Secondly, we also aim individuals to be able to examine the data from various perspectives and grasp the connections between sleep and lifestyle. This progress report marks the initial stages of our project, highlighting our ongoing efforts to build a user-friendly website that promises to be an invaluable tool.

Data

The dataset "Sleep health and lifestyle" is a classification dataset from Kaggle.com and is used to predict the type of sleep disorder based on all other factors, with an emphasis on sleep and exercise routines. Some of the variables included are sleep duration, quality of sleep, physical activity level, stress level, BMI categories, blood pressure, heart rate, etc. We have downloaded the CSV version of the dataset from Kaggle.com.

Task and Visualization Tools

Before creating visualizations, our team will clean the dataset to leave the necessary variables for this project. We will create interactive data exploration features to allow users to filter and select specific subsets of the dataset based on certain criteria. Histograms will be used to visualize the distribution of continuous variables such as sleep duration and heart rate. Bar charts will be implemented to compare categorical variables like BMI categories or types of sleep disorders. These visualizations are valuable for healthcare professionals or researchers. Moreover, box plots will be useful for identifying outliers and understanding overall data spread. Last but not least, we will try to implement other kinds of data visualizations based on the course materials. These visualization tools are essential for showcasing the prevalence of health conditions and lifestyle factors.

Collaboration Plan

After a detailed discussion between the three group members, our group decided to split all the responsibilities three ways because everyone was comfortable with coding, presentation, and writing tasks. We will be describing our responsibilities using the responsibility table. In addition, it's understandable that someone will do more of the writing/coding/presenting on some tasks, which will be discussed in group meetings. The reason for not assigning roles is to make sure every group member is involved in tasks related to class learning. Moreover, the designated meeting time of the group is every Tuesday and Thursday from 2:00 to 3:00 P.M. in Westgate, and we decided to have at least one meeting per week, and emergencies will be arranged accordingly. Lastly, the group created a group chat to make sure there was an easy way of communication between the group members.

Reference

Tharmalingam, L. (2023, September 18). Sleep health and lifestyle dataset. Kaggle.
<https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset>

Group Activity Table

Team Member	Responsibilities
Eric Wu	Data and Collaboration Plan Section
Dongyeon Kang	Task and Visualization Tools Section
Sunwoo Kim	Introduction Section

Group Activity Log

Date	Activity	Attendance
9/19/2023	Discuss the format and details regarding progress report #1.	All team member except Sifei Zhou