Deliverable 2: Project Proposal

Johnathan Nguyen Foundations for Data Analytics

Introduction

Background Information:

Caffeine is something that has become part of the daily life amount students and people in the workforce. By tracking the amount of caffeine a person intakes, it can help give data that provides insights into health, productivity, and a person's sleep quality. This would also allow for personal relevance because I am part of that student population who drinks caffeine daily in order to stay focused.

Purpose of the Project:

The purpose of the project is to analyze caffeine intakes and their patterns using the dataset that was given. The goal is to find patterns using the average intake, frequency, beverage types, and timings to get a better picture of the population's consumption and how it may impact their health, such as sleep.

Significance:

This project is interesting because it can help lead to a better understanding on the impacts of caffeine consumption on focus, sleep distribution, and whether age, gender, and time of the day could affect these. This could help contribute to more research on caffeine's impact.

Objectives:

Primary Objectives:

The primary goal of this project is to use caffeine intake to identify the patterns in consumption by using drinks, type of day, and the amount consumed, and how it will impact the outcomes of people's sleep distribution and focus.

Secondary Objectives:

- Visualize intake patterns
- Explore whether demographics influence caffeine consumption
- Whether people's beverage types are linked to higher or lower caffeine amounts
- Check whether age or gender influences the difference in caffeine consumption habits or how they digest it (how it impacts their focus or sleep)

Methedoloy:

Approach:

I will start by cleaning the data and organizing it so that it can be consistent and a lot easier to read. This would help with the next step of getting statistics such as averages and distributions so that I could identify patterns in the data. This would also help me with analyzing the correlations between caffeine intake, focus levels, and sleep quality to see if I could find trends within the data.

Tools and Resources:

- Excel
- R (tidyverse, dplyr)
- Caffeine intake dataset (CSV file)

Timelines:

- Week 1: Import and Clean the Dataset
- Week 2: Get statistics and grab information
- Week 3: Find correlations and trends between caffeine intake, focus, and sleep
- Week 4: Finalize conclusions and prepare final report

Expected Outcomes:

Deliverables:

- Report

Success Criteria:

To consider this a successful project, I would want the data to be cleaned and analyzed in a way that shows patterns between caffeine intake, focus levels, and sleep quality. I would also consider this as successful if I could produce at least a few tables or visual representations of the data and patterns so that it is easy to understand.

Challenges and Limitations:

Potential Challenges:

One challenge is that many of the variables have a true or false value, such as beverage type, gender, and time of day, which would mean I need to do more cleaning so that it would be easier to analyze. Another challenge will be making sure everything is meaningful because some of the datasets have decimal numbers. For example, why does age have decimals from 0-1?

Limitations:

The limitations of this dataset include a lack of outside factors, such as stress, workload, or their lifestyles, which could impact caffeine intake and their sleep. It is also a somewhat small dataset with only 500 samples, which may limit the results and how it can be applied to a bigger population.

Conclusion:

Summary of the Proposal:

The project will be using the caffeine intake dataset and analyzing it to find patterns in consumption, beverage choice, and timing, and how these relate to outcomes like focus and sleep quality. Once I clean the data, I will be able to find correlations and trends that will help show the analysis of students and a broader population.

Call to Action:

I am committed to completing this project so that I can find results using the data, so I can practice analyzing it and presenting it in a proposal. It will also allow me to move forward with the analysis and produce insights that would be useful to me personally and maybe to others who are also looking at caffeine intakes and their impacts on focus and sleep.