Decoding Dreams: A Data-Driven Dive into Sleep Efficiency

Coby Alvarez, Elias Schablowski, Shivani Kotian, Daniel Hernandez, Erick Vergara
California State University, Los Angeles
Email: { calvar108, skotian, eschabl, dherna197, evergar9 }@calstatela.edu

Motivation

The average person spends about 26 years sleeping in their life which equates to 9,490 days or 227,760 hours. That's one third of our entire lives spent asleep in bed! Yet for most people sleep is a mystery and they don't know the first thing about what is actually happening when they sleep. It's hard to improve something that you don't fundamentally understand. The bottom line is - The better you sleep, the longer you live. Sleep is a major part of our everyday lives, it greatly affects our productivity and our health throughout our lives.

Poor sleep quality has been linked to a multitude of ailments from dementia[1], to attention lapses, reduced cognition, delayed reactions, and mood shifts. Our motivation for analyzing this dataset is to find the reasons that might affect a good night's sleep and what changes can people make to get better sleep.

Approach

Our approach to analyzing and visualizing the intricacies of sleep patterns involves a comprehensive exploration of the dataset titled "Sleep Efficiency," which includes metrics such as age, gender, bedtime and wakeup times, sleep duration and efficiency, REM and deep sleep percentages, awakenings, and lifestyle factors like caffeine and alcohol consumption, smoking status, and exercise frequency. The envisioned visualizations aim to uncover hidden patterns and correlations within the data, facilitating a deeper understanding of how various factors contribute to sleep quality.

The process begins with data preprocessing to ensure accuracy and usability, followed by the creation of a series of targeted visualizations. Each visualization serves a specific purpose:

A histogram that will offer insights into the overall sleep efficiency among participants, highlighting common efficiency levels and outliers. By grouping participants into age ranges and visualizing average sleep durations, we can observe potential trends and deviations, suggesting how age may influence sleep. A scatter plot will help identify correlations between how long people sleep and the quality of that sleep, potentially revealing optimal sleep durations for high efficiency. Box plots will compare sleep efficiency across different levels of caffeine and alcohol consumption, revealing their impact on sleep quality. A line or bar chart examining the

[1] https://doi.org/10.1016/j.smrv.2017.06.010

relationship between exercise frequency and sleep efficiency could indicate if regular physical activity contributes to better sleep. A scatter plot exploring the link between the frequency of awakenings and sleep efficiency might uncover if less interrupted sleep correlates with higher efficiency. Through bar charts or box plots, we'll examine if there are significant differences in sleep patterns between genders, considering duration, efficiency, and sleep stages. Analyzing sleep efficiency or duration based on smoking status will highlight how this lifestyle choice may affect sleep.

Furthermore, we can use scatter plots to analyze deep sleep vs light sleep percentage, and categorize participants, as well as linking certain activities to deep sleepers and light sleepers.

Milestones

1 The objective of our visualization will focus on finding the patterns that come with levels of sleep and their correlation with lifestyle. It will show how things like food or caffeine have an impact on the levels of sleep people get. We will begin with a quantitative data analysis on comparisons and outcomes using categorical formats to clarify the data. Starting with a basic bar graph and showing comparisons with exercise frequency, the analysis goes deeper into how certain foods and other factors contribute to sleep efficiency.

2 Design of the visualizations will be focused on our audience accessibility and preattentive learning from visualizations by simplifying complex visuals, increasing colorblind accessibility, and continuously improving charts through iterations. In design the sleep relationships data can be negatively affected by misleading factors. One factor that can affect design of visualizations are outliers. Ignoring outliers and void data may significantly help to filter out and view data clearer. They may portray participants who opt out of giving information skewing data. Therefore, designing multiple chart interpretations of the same quantitative data can provide a higher-level depiction of data. Going through the project in a coding-like level by level and iterative fashion overall improves the quality of visualizations.

3 Results of visualizations may provide neutral, negative, or positive findings of factors affecting sleep effectivity. Finding solid evidence that sleep efficiency reduces, improves, or doesn’t get affected by our data columns is the most important purpose of the project. Data visualization results transform information into knowledge that is easier to process generating new areas of focus, insights, and questions. Future advice regarding sleep could visually contribute to the understanding of living healthier and longer lives.

Extensions

Beyond the scope of this project, we envision incorporating phone usage before sleeping. This way we can explore how using technology impacts sleep efficiency. Another addition would be to compare how sleep impacts one's mental health. This way we can make charts showing how lower levels of sleep increase chances of certain conditions. Furthermore, we can compare healthy habits to bad habits to distinguish what has a higher impact in helping sleep efficiency. To help distinguish this the habits we can try to add age groups, different populations, and their work to help make the result more accurate and better.

Tools

List Tools that you plan to use for this project (e.g., Tableau for data visualization, Excel for Data recording and storage, and Python for cleaning your data).

- Tableau Desktop - Data Visualization and Analysis
- Microsoft Excel - Data Storage

References

List references that you cited in this proposal. One of the references is the dataset that you plan to use.

https://www.kaggle.com/datasets/equilibriumm/sleep-efficiency
Redirecting, doi.org/10.1016/j.smrv.2017.06.010. Accessed 21 Mar. 2024.