Correlation Between Stages of Sleep and Sleep Quality

Kaivan Khazeni, Sahith Jampala

Process Book

Table of Contents

Overview & Motivation

Basic Information:

- Title: Routines that Impact Sleep Quality
- Group Members:
    - Sahith Jampala:
        - uID: u1047503
        - Email: u1047503@utah.edu
    - Kaivan Khazeni
        - uID: u1135484
        - Email: u1135484@utah.edu

Background and Motivation:

- Inspiration behind the decision to dive into sleep data is based off of the leading fitness wearable, Whoop, which tracks numerous stats/data of the body ranging from Heart-Rate Variability, to Sleep Quality(REM, DEEP, SWS, LIGHT, ETC.)
- https://www.whoop.com
- Kaivan has had a whoop strap for the last year and has been interested in the health data field in terms of fitness and the body's overall readiness to take on strain versus recover.
- Since the strap and the company has had such an impact on Kaivan's mindset, Kaivan switched to Data Science as a major, and this would be the perfect opportunity to dive into that field(health/fitness data) with a visualization representation.
- Sleep has a profound impact on everyone's health, and the correlation between efficient sleep and better overall health is high. On the other end, those who do not sleep enough or sleep "well-enough" are not as healthy.

Overview:

- Primary Question : How do different durations of stages of sleep affect overall sleep quality?
- Want to learn:
    - What routines have the most/least impact on sleep?
    - How sleep-quality is measured?
    - How does sleep impact your overall health?
    - What is a better indicator of sleep, length or consistency?
    - Is there a correlation between higher qualities of sleep on certain days of the week?
- What we want to accomplish:
    - Showing correlation between stages and sleep

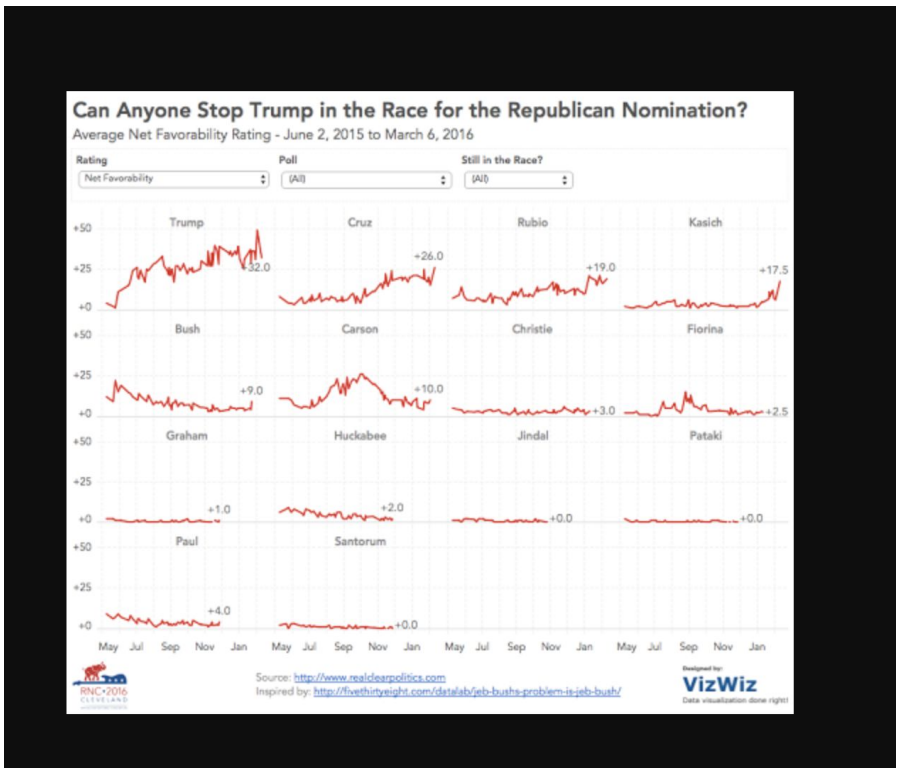
    - Benefit: Knowing what stages of sleep and the duration spent in them have good/bad impact on sleep; by knowing this we can determine what are ideal lengths of each stage to have while sleeping.
- Showing the different measurements of sleep in terms of stages of sleep
    - Benefit: The benefit of this would be to know how long each stage lasts, around what time in the night each stage happens, and the importance of each stage for your body.
- Showing that sleep does have a profound impact on bodies readiness the next day
- Highlighting the difference between length of sleep and the consistency of a sleep schedule (no matter the length)
    - Benefit: To show that consistency is more important than length of sleep; i.e. quality increases with consistency not length.

Related Work

WHOOP:
https://www.whoop.com/

- Whoop is a fitness wearable that gives the user their sleep data, strain data, and recovery scores.
- Whoop's sleep studies and visualizations were the original spark for our project idea. Whoop uses bed-time routines and other variables of the past day to determine sleep quality and recovery the next morning.

Example that Haihan gave us:



- Haihan advised us that our visualization at the time was not portraying enough information since we had access to 8 people (in the study) and only displayed one individual.
- The idea of showing all 8 people in some capacity (via this visualization provided) was going to be our new starting point and gave inspiration for the project. By showing all 8 people, the audience would get a better idea of how they compared to each other, in terms of sleep quality over a period of time.

Data

Source:

- https://datasets.simula.no/pmdata/
- This source contained the dataset of 16 individuals over a 5 month span
- This data was recorded using Fitbit, Google Forms, and PMSys for the several sets provided.
- Specifically, the dataset we were interested in was the sleep.json, which breaks down each night of sleep into the stages “”Light”, “Deep” and “REM, and the sleep_score.csv, which provides us with a sleep score (0-100 scale) and the minutes of sleep in each stage.
- This data was collected using the Fitbit Versa, which is another fitness wearable/tracker that tracks sleep and sleep quality(Similar to Whoop).
- The CSV file and .Json file were used simultaneously to form our final .csv used in JS

Scraping/Cleanup Method(s):
Because data was broken into separate CSV files for each person, we had to use Python and Microsoft Excel to join, scrape, and manipulate the data to match exactly what we wanted.

- Python: In python, we used Pandas’ data frames in order to visualize the data, as well as convert one of the json files into a CSV file. We did this so that we could then join the two datasets needed for each individual quite simply. Once we had the data frame exactly how we wanted it, we exported each individuals’ data as a CSV file which we then edited in Excel.
- Microsoft Excel: Once in Excel, all that was left was removing all the unnecessary columns in each individuals’ CSV file, converting the time columns into minutes from milliseconds, and then we were left with one final issue.
- Final Issue: The final issue that we ran into was that the stages of sleep and the duration of sleep in each stage were given in a dictionary object for each day for each person. We needed these values as separate columns, since CSV files cannot hold dictionary column types. In order to resolve this issue, we simply went back to python and wrote some code that pulled all the values out for each day for each person, and then converted each key in the dictionary to a column, and the value to be the duration of sleep in each stage.

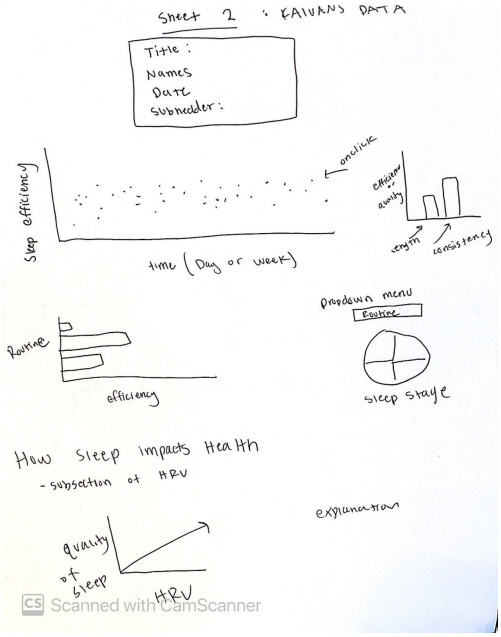
Exploratory Data Analysis

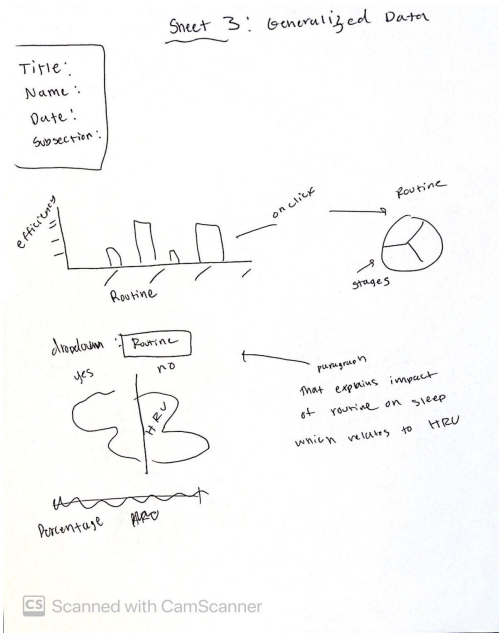
What visualizations did you use to initially look at your data? What insights did you gain? How did these insights inform your design?

- Originally, we were studying and observing the visualizations that Whoop provided Kaivan on his desktop and mobile phone. The visualitions range from bar-charts, to scatter plot, and give precise information that is linked together throughout the whole website/application.
- The correlation between sleep stages and quality of sleep led us to wanting to correlate the two or at least find cases where the stage of sleep helps increase quality of sleep.
- When we received our data, via the link provided in the 'Source' section, we were provided a .csv file and a .json file. This was in the realm of tabular data visualization which was exactly what we needed in order to scrape and manipulate the data found in order to fit the mold for our project.
- We then related the .csv file to the Whoop visualization and pondered how we could display this information we found using Whoop (and other fitness wearables) as inspiration for our own visualizations and graphs/charts.

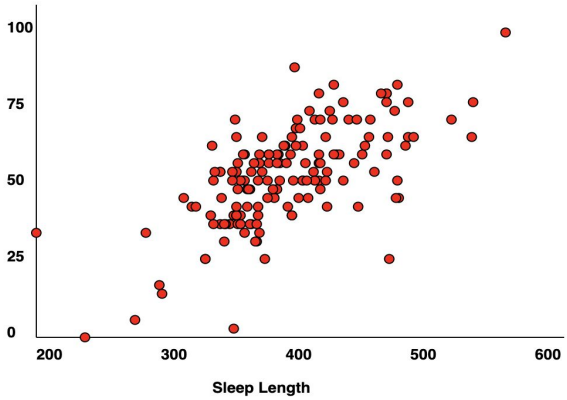
Design Evolution

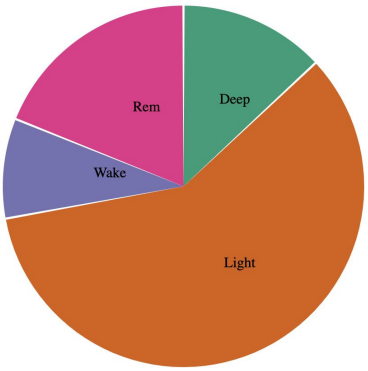
In the beginning, our goals were very different from this project, but because the desired data was not available, our design changed vastly. Here are a few examples of sketches that we were considering:





Because we were not able to access Kaivan's data from Whoop, we had to improvise and find general data online, but to our dismay, there were not many datasets that involved routines prior to sleep, and how that affected sleep quality. Instead, we were able to find a dataset that contained many different people and the time of sleep spent in each stage of sleep for about 150 days per person. Initially, we went with an implementation of a scatter plot of one persons' data, where each bubble represented a day and the quality of sleep on that day (y-axis). Clicking on one of these bubbles resulted in a piechart displaying to the right of the scatter plot, which broke down each stage of sleep in terms of the total duration of sleep for each night. Here is a picture of what this design looked like:





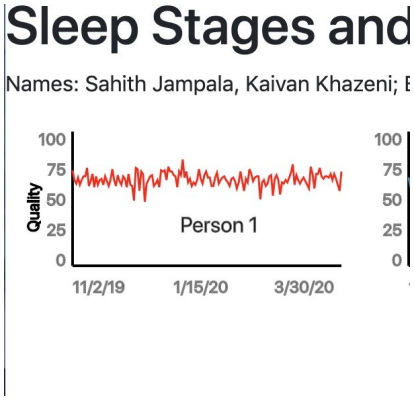
After meeting with Haihan, she recommended that we completely switch our design, as only displaying on individuals' data would not be very interesting. She also expressed that pie charts are not the best visualization tool to use in this scenario. From here, we were able to get to our final design strategy which is described with great detail in the next section, Implementation.

Implementation

Describe the intent and functionality of the interactive visualizations you implemented. Provide clear and well-referenced images showing the key design and interaction elements.

Interactivity is key in any visualization to keep the audience engaged and to add another dimension of information that is not seen on the page provided. Some features we implemented are as follow:

- Upon hover of the line-charts of the 8 people, the line turns red, highlighted the current person of interest

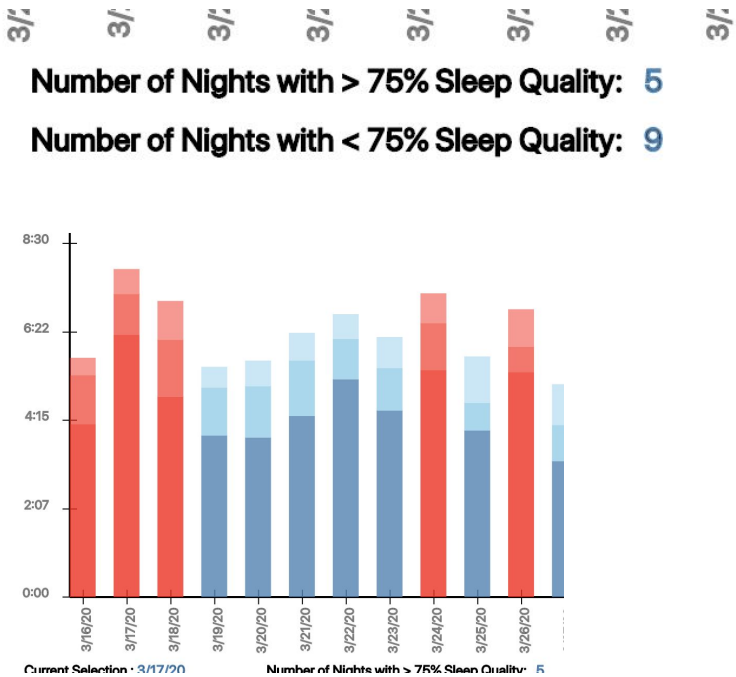


- When the highlighted line is clicked, the chart on the right will update to the person that is clicked.

- Two drop down options to select a new individual, and a starting date.

- A fun fact button to display hidden facts about sleep quality and day of the week/stage of sleep values.
- Upon mouse over of the individual bars, the current date, and minutes in each sleep stage are updated below the chart.

- Bottom right of the chart, the number of days below or above 75% sleep quality are listed and upon hover, the respective days will also highlight red.



Evaluation

What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?

What we gained from the visualizations chosen:

- We learned about how stages of sleep for a given individual can be very different on a night to night basis. We saw this through our bar-chart visualization when displayed a two week period of data and the person's sleep in hours(and stages in hour) per night.
- Another observation is the quality of sleep, which ranged through the week and were lower(or higher) on certain days of the week for the respective individuals
- Lastly, through the date selection option, we could see a trend of length of sleep through a month span, or even from the beginning of the study to the end(5 month period on average) and how quality of sleep(or length of sleep and sleep stages) changed.
- Our visualizations work great, not only do they display the data in a neat and precise matter, but they are fully functional via interactivity. Data is linked upon user selection which makes the transitions crisp.

Improvements:

- We could improve the user experience even more by creating custom date input from the user. This would imply the user had data that matched our .csv file, however if they had the format/template, they could provide the data and we would create the visualization for them
- Another improvement would be custom date selection, rather than a two week period with a custom start date, the user could select the start and end date on their own and it would be up to us to read the data and scale it appropriately.

Work Log

11/05/2020: Peer Feedback Review
Peers: Anne Senay, Jacob Flosiand, Kenny Ho

Notes Received:
(Note: our previous project plan was not attainable due to the lack of data available. The routine before bed we planned on using came from data from Whoop, however they were not able to supply Kaivan with his data in time for this project. We moved on to sleep stages and its effect on sleep quality, therefore the notes received are about the previous iteration of our proposal.)

- Choose two routines and compare them side by side
- Bar chart instead of column chart if using a lot of routines
- Clarify HRV Chart more; how generalized data will fit into the "blob chart"
- HRV Chart mainly to show which routine has the best HRV
- Possibly hover function on HRV chart where it shows the actual number of people in each percentage
- Possibly onclick function on HRV chart where it shows the details on yes/no; details shown through storytelling?
- Add good amount of storytelling components (with routine, "on average, 'routine' shows that …")
- Possibly add functionality for user to input their own info
- Ex. Body Figure, showing which areas of body user is taking care of
- Possibly reduce amount of visualizations, can be overwhelming, or maybe guide user through in detail, so that it is not

Thoughts on Feedback Session:

- Session was helpful in the sense that it helped us walk through the ideas with people that knew nothing about the topic. Basically like if a new user were to look at our visualization, we figured out what we need to explain/what we don't with the help of our peers.
- It was also helpful because they were able to provide unique ideas for the visualizations that we already have that we did not think about. They also provided alternate solutions to the problem we are facing with attaining data.

11/09/2020: Data Collection

- On this date, we were able to find data on sleep stages from a study. This data included the stages of sleep and the efficiency, however on another .csv file, the data had the quality of sleep and more information on the in and out of each sleep stage.
- We were able to use Python to manipulate and combine the two .csv files into one, and convert to Json.

11/10/2020: First commits on Github

- Our first step was to load the data into JS and read it
- After we were able to load the data and save to an object, the goal was to access a point of data and see how the Python conversions parsed from csv to Json and we fixed any errors we found
- Our main chart was created, which was a line chart showing the relation between sleep length and sleep quality.
- Our first implementation for interactivity was an onclick function that allows the user to click on a specific data point and to see how the stages compare to each other(Deep, Rem, Light, Awake) in a pie chart
- The axis labels are not final and the pie chart is not crisp however it is a step in the right direction

11/15/2020: Notes on Milestone Meeting with Mentor:

<u>Suggestions</u>

- Rather than showing a single person's data, show all 8 peoples' data that we have.
- Instead of Scatter Plot, do 8 different line-chart plots that have y-axis representing quality and x-axis representing the dates.
- This would make comparison of data more interesting than just looking at a single persons' data.
- Pick something other than piechart since Lex does not like piecharts.

Screenshot of example:

11/19/2020: Notes

- Changed single person's data to data of 8 different people.
- Able to access each person individually

11/23/2020: Notes

- Implemented the 8 overview graphs that show a line chart based on sleep quality from the start date to end date of each person.
- From here, we plan on implementing interactivity that allows the user to click on any persons' chart which will lead to our main plot of stages of sleep from each night.

11/24/2020: Notes

- Added functionality to each min-graph that changes the line chart's color on hover, and allows the user to click the line to show stacked vertical bar graphs.
- Created the svg and added the axes for the vertical bar graph.

11/25/2020: Notes

- Added the bars to the graph per each stage of sleep given for 14 days of data. Each bar represents the duration of sleep spent in this stage.
- All data working for the most part, but running into a problem of not all people having data on the same days, scaling issue for x-axis.
- Left off with mostly all charts populating correctly, other than person 3.

11/27/2020: Notes

- Today we added a legend to the vertical bar graphs to show which stage is represented by each color.
- Also added a dropdown selection that allows the user to select the start date of when the data begins and continue for the next 14 days of data available.

11/29/2020: Notes

- Added an information box, containing information about the specifics of each day in terms of rem, light, and deep sleep stages as well as total duration.
- On hover of a bar, the information instantly populates based on which bar is being hovered.
- Also added storytelling about the number of days from the given period that the person had > 75% sleep quality, and < 75% sleep quality.
- On hover of this information, the actual bars in each category will highlight to a different color.
- Added transitions to color switch.