

FINAL REPORT

TERM PROJECT

DGM 6108

PROGRAMMING FOUNDATION FOR DIGITAL
MEDIA

NAME: RASIKA RAMESH PARATE

NORTHEASTERN UNIVERSITY

TABLE OF CONTENT

1. Introduction.....	3-4
2. Overview of Data Recorded & Process.....	4-5
3. Table of Data Recorded.....	5-7
4. Rationale for Visualization.....	7-9
5. Final Visualization.....	9-12
6. Analysis.....	12-14
7. Alternate Visualization.....	15-16

Relationship Between Sleep duration and Stress level.

Introduction

In my final term project that I have been working on since the project was introduced from step one for building the project that was writing the hypothesis the very first thought that strike in my mind was that, why not collect a data on a specific topic that I am trying to find and answer for and make a visualization on the basis of the same topic, so I took some time to think what all questions do I have and then I came to a point where I noticed that from a very long time I was struggling with my sleeping pattern and at that point of time I did not know that what are the factors that are mostly affecting my sleep duration.

Therefore, after doing some observation on my day-to-day routine I was able to find three factors that affects my sleep duration that was stress level, workload and the time spent on mobile phone talking to my family or friends, out of these factors I narrowed down to two major factor which are stress level and workload and added a new factor that is how all these factors affects my mood.

The hypothesis for the term project is As my sleep duration increases my stress level decreases. And the other factor that affects my sleep duration is the workload that I have and all this factors ultimately affects my mood.

The whole term project is built in a way where the entire course till the end will definitely gave me an answer to my hypothesis for what I was looking for since the start of the project. The very first step was hypothesis writing, then giving a draft idea about the project. The phase two was about making your data into objects and arrays. then presenting at least 3 days of data.

After all this the main part began from making sketches of your visualization, and the next task was to writing a code for any of your 1 visualization in draft but it cannot be a simple scatter or a bubble plot. I made a lollipop chart. As it as for the first time I did not do any experiments I wanted a basic lollipop to see how my output turns out and does it justify my hypothesis or not. For my second attempt I made a scatter plot.

Before I was introduced to JavaScript Object Notion (JSON) I was adding the data manually in the JavaScript file. Then the next step for the project was to translate the data collected into a JSON file. I named the variables such as “sleepDurtion”, “stressLevel”, “workLoad”, and “mood”. I also added other data that I collected in my JSON file which is time taken to fall asleep as “fallAsleep”

At first, I faced some issues while working with JSON file. When I was manually adding the data to in the JavaScript file, I was able to get the output but when I added the code to link the data from my JSON file the data was not been addressed. But after that I solved that issue just for trial purpose to see if my JSON file is working properly, I added only three data properties to see if the data is been plotted after some tries and errors, I got the third data as the output I wanted. Later I added the data that I actually wanted to get as the output. But the output when I go live was showing a blank canvas. After checking console there were some errors which got solved. After one of my data was been plotted on X axis the values on the same scale were not in ascending order instead, they were getting plotted in a way I added my data in a JSON

file. To change the vales in the desired order that I want I gave variables to the data of my X and Y axis such as “sleepDurationMin”, “sleepDurationMax” and “stressLevelMin”, “stressLevelMax” in short, I set a minimum and maximum values. The data was being plotted but it was out of the margin so later I changed my margin values and got the code fully working.

Overview of Data Recorded

For the first three days when I started to record the data, I made three different tables for sleep duration, stress level and workload. I noted the date, time taken to fall asleep, start and end time, mood, and duration in hours. For stress level I recorded the reason for stress and percentage of stress at the end of the day. For workload I recorded the start and end time when I used my laptop, its duration in minutes and frequency, as well as defined my workload on the scale of 1 to 5.

But later I changed the way I collected the data for 42 days in a different way. I recorded the in one table I started by recording the date, sleep duration in minutes, stress level in percentage, workload on the scale of 1 to 5, mood on scale 1 to 5 and time taken to fall asleep in minutes.

Data Collection Process: -

I recorded all the data in a Word document between 19th October 2022 till 14th December 2022. In total 42 days of data was collected. Sleep Duration data was recorded in a way that I used to note the time when I go to bed which as the start time and the end was when I wake up from sleep. The total hours of sleep I get per night was then converted into minutes. The minimum range for my sleep duration was 200 minutes and maximum was 700 minutes. Stress Level was calculated at the end of the day using a smart watch in percentage and the range was between 25% to 75% Workload data was recorded by how much work I had in a day on the scale of 1 to 5 where

1 = Low, 2 = Moderate-low, 3= Moderate-high, 4 = High and 5 = Very high.

The data for mood was recorded at the end of the day on the basis of what my mood is before going to bed on the scale of 1 to 5 where

1= Happy, 2 = Calm, 3 = Sad, 4 = Angry, 5 = Tensed.

Before I was introduced to JSON just for trail purpose, I saved my data in a CSV file named data.csv and added the link for the csv file in the JSON file. I had this alternate option ready for my backup incase if my JSON file didn't work. But after paying attention to the code and working on it I am successful in converting my data in JSON file and adding that to my JavaScript and make it work in a way I want to.

Table

Data Recorded Between 19th October to 14th December 2022

After collecting the three days of data in the beginning, I changed the complex data collection process to a simple form so that whenever I want to add a new data for a new data, I open my word document and add the particular recorded data. I found the way I recorded the data in my second table easier and more convenient for me. So instead of all the other data I collected for a separate factor I narrowed down to the main factors for my visualization.

Therefore, I sorted the first three days of data that is from 19th to 21st October and added it to the new table: -

Sr No.	Date	Sleep Duration(minutes)	Stress Level in % (percentage)	Workload Scale (1 to 5)	Mood (1 to 5)	Time taken to Fall Asleep(minutes) (00:05 to 00: 25)
1	10/19/22	590	45	4	3	5
2	10/20/22	460	57	2	2	15
3	10/21/22	385	62	5	5	25
4	10/22/22	230	74	4	5	13
5	10/23/22	423	57	3	4	3
6	10/24/22	629	70	5	4	18
7	10/25/22	290	59	3	2	1
8	10/26/22	535	67	5	4	7
9	10/27/22	446	38	4	3	18
10	10/28/22	371	58	3	2	14
11	10/30/22	500	32	2	2	20
12	10/31/22	310	51	3	5	14
13	11/02/22	525	30	1	2	10
14	11/03/22	396	64	4	4	14
15	11/04/22	320	74	5	5	04
16	11/05/22	432	40	4	3	23
17	11/06/22	544	30	2	1	04
18	11/07/22	342	56	4	5	11
19	11/08/22	470	48	3	4	12
20	11/09/22	240	61	2	3	24
21	11/11/22	610	35	1	4	5
22	11/12/22	250	70	5	5	15
23	11/13/22	315	28	1	2	12
24	11/15/22	515	55	4	3	9
25	11/16/22	342	38	2	1	12
26	11/17/22	485	49	1	1	3
27	11/19/22	625	29	2	3	16
28	11/20/22	205	65	5	4	25
29	11/22/22	160	68	3	2	10
30	11/23/22	690	36	1	1	15
31	11/25/22	650	42	2	2	8
32	11/26/22	360	28	4	3	17
33	12/01/22	475	46	4	2	10
34	12/02/22	695	36	2	1	7
35	12/07/22	400	50	3	2	18
36	12/08/22	275	69	5	4	25

37	12/09/22	513	44	4	2	10
38	12/10/22	620	35	3	1	4
39	12/11/22	608	33	1	2	04
40	12/12/22	260	72	5	4	23
41	12/13/22	577	59	4	2	12
42	12/14/22	453	38	5	3	07

Data Collected in Function (JavaScript) Manually

```
[{
  sleepDuration: 590,
  stress Level: 45,
  workload: 4,
  mood: 4,
  fall Asleep: 5
}, // [10/19/2022] date when the information was gathered
{
  sleepDuration: 460,
  stressLevel: 57,
  workLoad: 2,
  mood: 3,
  fallAsleep: 15
}, // [10/20/2022] date when the information was gathered
{
  sleepDuration: 385,
  stressLevel: 62,
  workLoad: 5,
  mood: 2,
  fallAsleep: 20
}] // [10/21/2022] date when the information was gathered
```

Example of data added manually

Data Collected in JavaScript Object Notion (JSON) File

```
[
{
  "sleepDuration": 590,
  "stressLevel": 45,
  "workLoad": 4,
  "mood":3,
  "fallasleep":5
},
{
  "sleepDuration": 460,
  "stressLevel": 57,
  "workLoad": 2,
  "mood":2,
```

```
"fallasleep":15
},
{
  "sleepDuration": 385,
  "stressLevel": 62,
  "workLoad": 5,
  "mood":5,
  "fallasleep":25
}
]
```

Example of data added in JSON file

Rationale For Visualizations

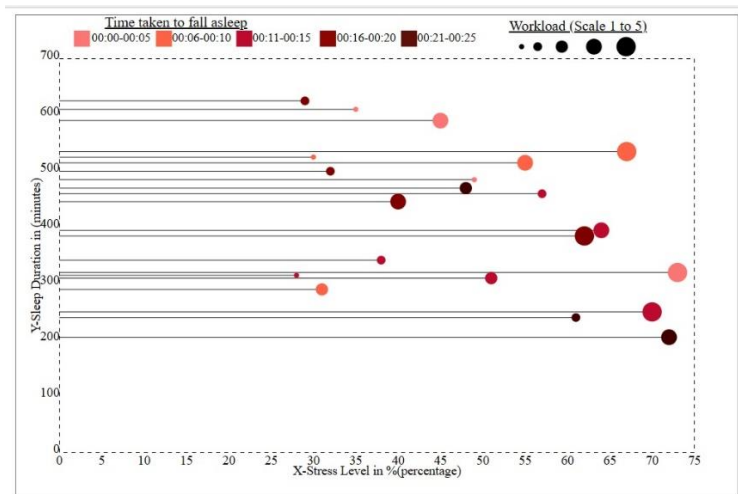
For both of my attempts I tried visualization such as basic lollipop chart and scatter plot. Apart from that I considered other visualizations formats such as Cleveland dot plot, and bar graph. But out of these four visualizations format the most suitable format for representing my data is a Lollipop Chart. For the kind of data, I have recorded, after creating a visualization and looking to the output I felt that the chart that I ultimately choose gave me a better clarification of the data and the output was justifying the hypothesis.

For all the visualization formats I had 3 fix data that are sleep duration, stress level, workload, and the 4th data that differed was mood or time taken to fall asleep. After I was introduced to how the color scheme should be selected for a particular type of data, I got to know that if the data is continuous then I can use color gradient for example if my 4th property in the visualization was time taken to fall asleep then the data was recorded in minutes that is from 00:05 minutes to 00:25 minutes. But for my final visualization format I used my 4th property as mood which was on the scale 1 to 5 and the color for the mood was determined on the basis of research done. I did the research regarding the meaning behind the colors, which particular mood represents which color and according to that I gave color to the property mood.

When I created a Lollipop chart as my final visualization in the beginning, I added only three data properties that are the more important factor in my hypothesis that I want to prove which are, sleep duration, stress level and workload. By simply creating a lollipop chart the output I got made more sense than any other visualization format. Later I added the fourth property in y final visualization that is mood.

Rationale For Final Visualization (Attempts)

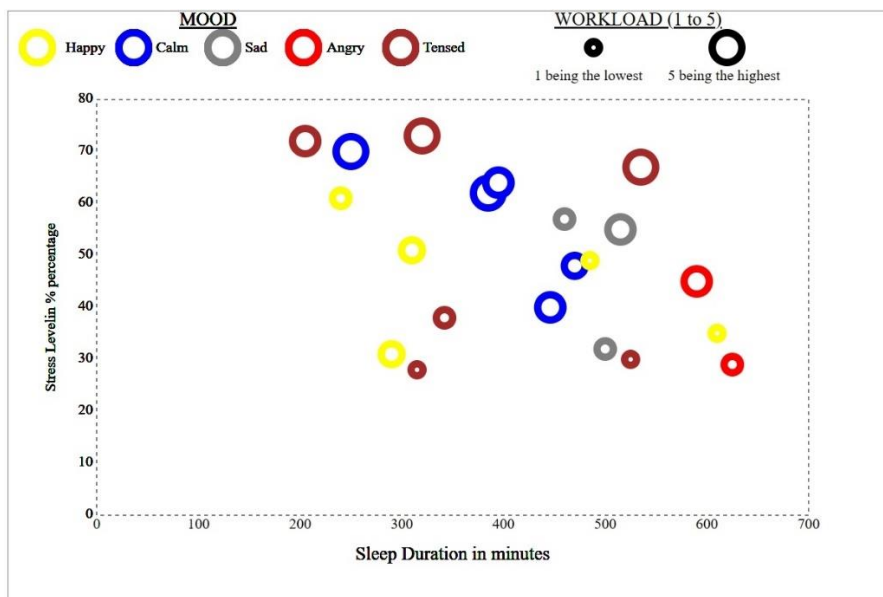
Attempt 1: Lollipop chart



For the attempt 1, I created a lollipop chart where I used 4 data. The X axis represents Stress Level in percentage and Y axis represents Sleep Duration in minutes. And the circles are drawn with the help of data Workload (1 to 5) where 1 is low and 5 is very- high. And the color to the circles is given by the data of Time taken to fall asleep, as the data that determines color is a continuous data and the color scale, I used color gradient which are shades of red color.

The points are being plotted on the basis of what stress level percent is and total sleep duration. Both the data meeting at one point creates a circle of different radius depending on the workload and then giving the color on the 4th property. For example, if my stress level is 72% and sleep duration is 215 minutes and the workload is 5 then the circle of radius 5 will be created at the point and the time taken to fall asleep is 00:11 to 00:15 minutes - strong red color.

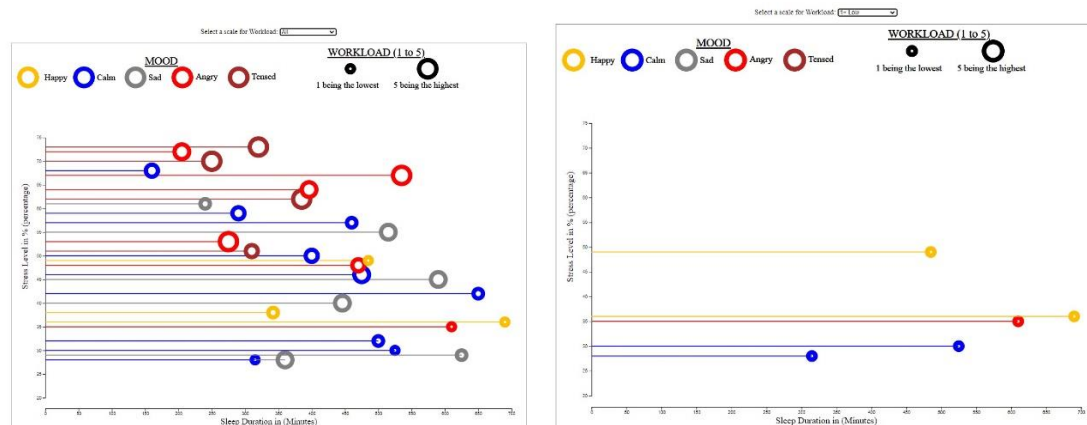
Attempt 2: Scatter plot for draft and made it for final Lollipop chart



For the second attempt I created a scatter plot in a dramatic way. I used 4 data properties for this visualization in which X axis represents Sleep Duration in minutes and Y axis represents Stress Level in percentage and different size of circles is defined by workload on and the color is defined by the property mood.

The changes made in attempt 2 are: I flipped the data of both the X and Y axis and added a new fourth data property that is mood. I added 5 different moods on the scale of 1 to 5.

Draft Visualization



Final Visualization

In this final visualization all the data is similar to the scatter plot. I added few more days of data and also added lines that connect each circle that is been plotted to give more clarification so that whenever someone looks at the visualization, they can follow the line that goes from Y axis to the stress level point and can understand what data I am trying to portray. To make it more appealing I added same color to the lines of the circle on the basis of mood.

I decided to go with Lollipop chart. I wanted to add an arc at the outer edge of the circle but the data that I wanted to add was time taken to fall asleep but the data was not that relevant to my hypothesis so I decided to add the major data which mostly helped me to reach to a conclusion.

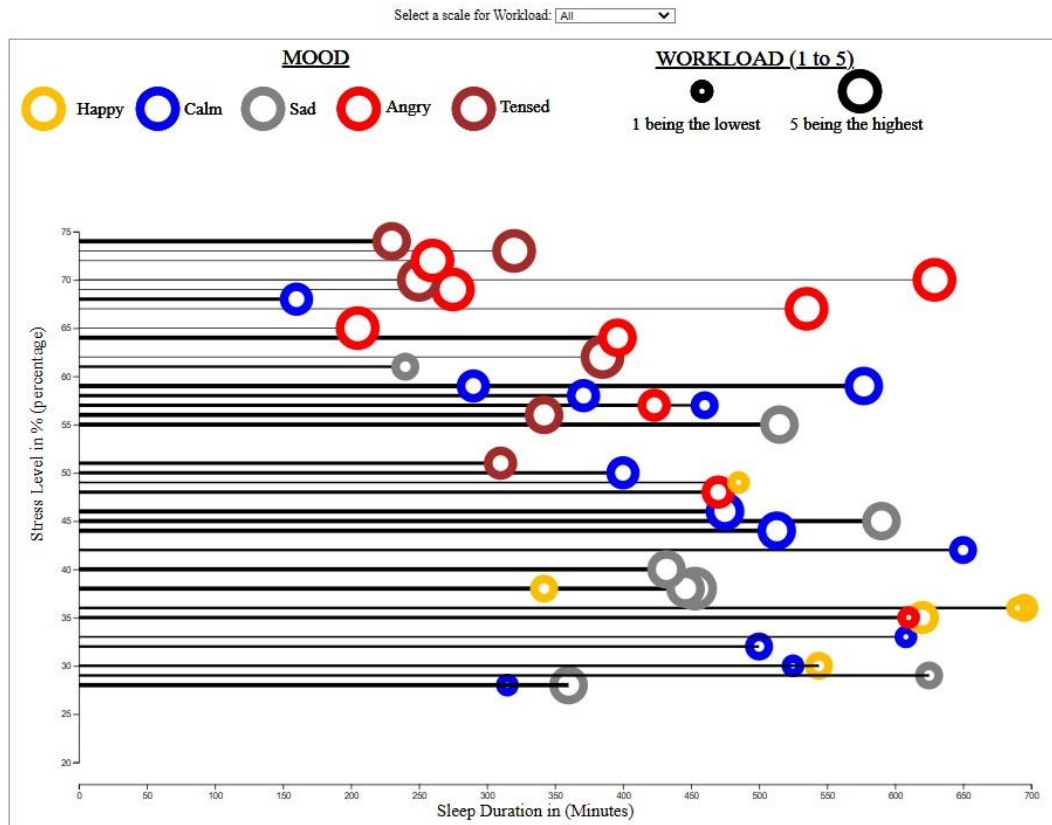


Fig (1)

The value Y axis in the final visualization starts from 20 instead of 0 because when I did research about a minimum stress level a person can have then the range was between 25 to 75 percent. If I create a scale from 0 then there will be a huge blank space as there will be no data plotted at that particular stress level.

After I select “All” option in a dropdown menu I get all the data plotted in the canvas but most of the data is been clustered and it is confusing for the viewer to understand. Therefore, to avoid this problem I added interactivity to my visualization. I also mentioned that what number of workloads determines.

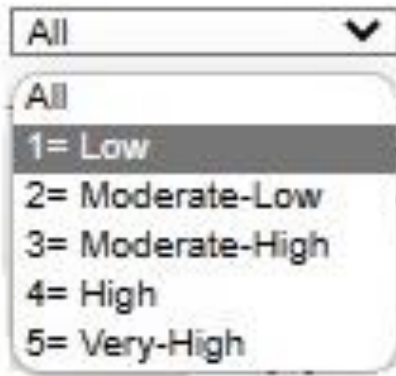


Fig (2)

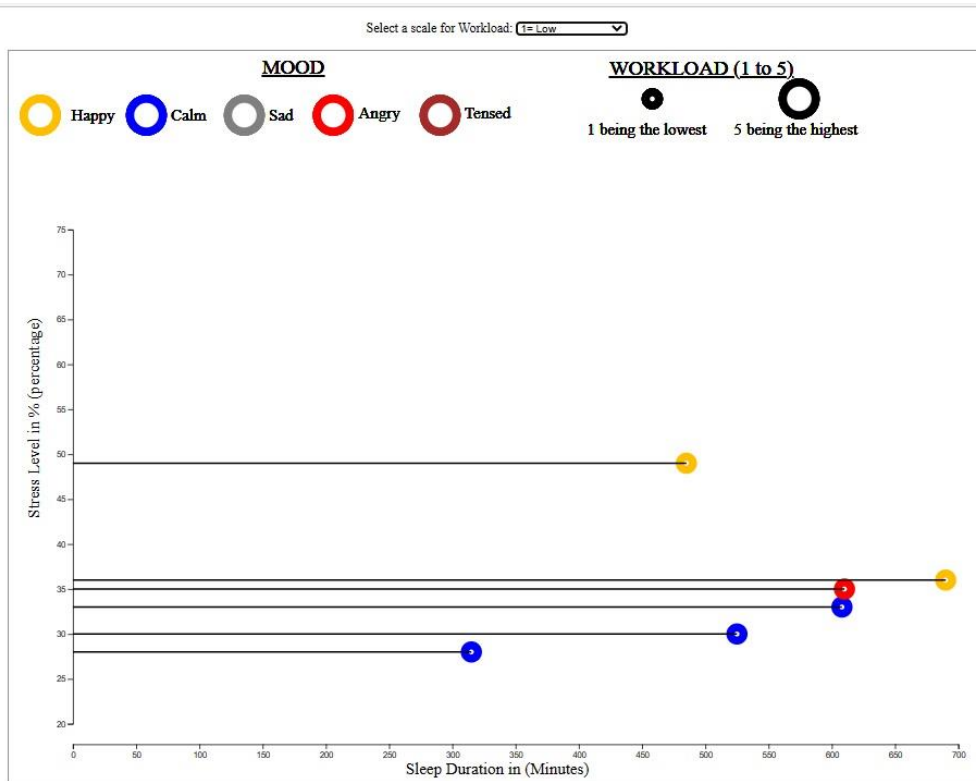


Fig (3)

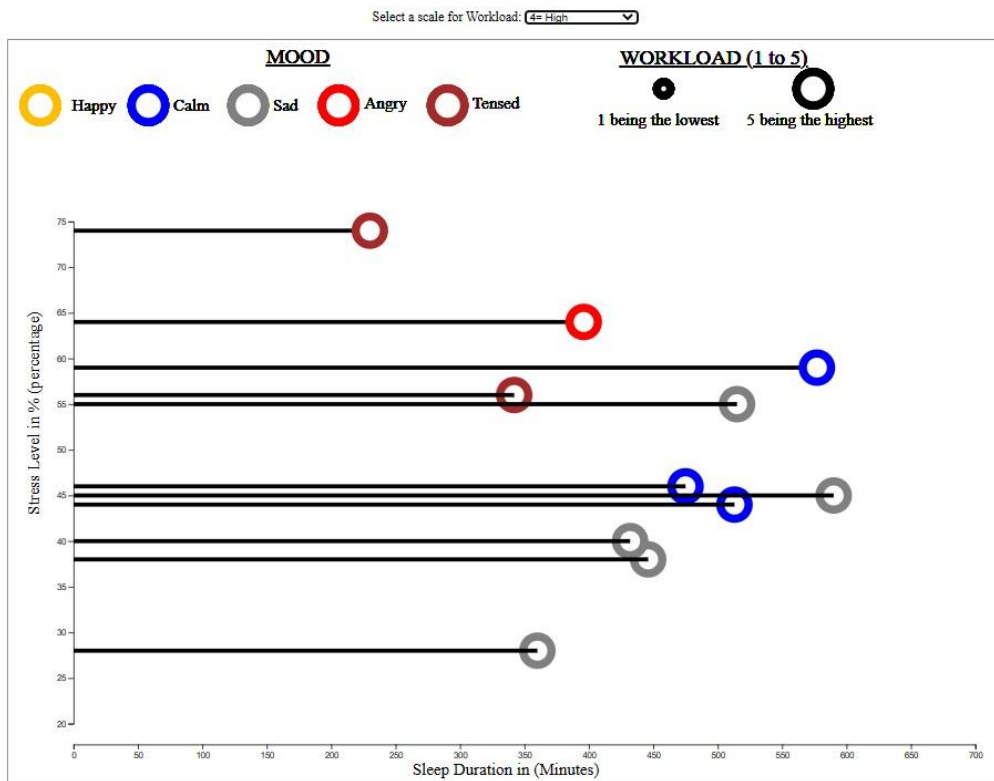


Fig (4)

In the final visualization the X axis represents Sleep Duration in minutes and Y axis represents Stress Level in percentage and the circles are created by Workload data and the color to them is defined by Mood. The lines that connect the circles changes the width according to the workload data only.

Instead of workload data to define the line I changed it to data of Fall Asleep but the data I feel is not supporting my visualization.

The images that are added above explains Fig (1) shows all the data that I have added in the JSON file which has 6 weeks of data. Later in Fig (2) the image that shows a dropdown option which adds interactivity to the final visualization. When you sort the data on the basis of workload. In Fig (3) the image shows the data of days when the workload was 1 which is low.

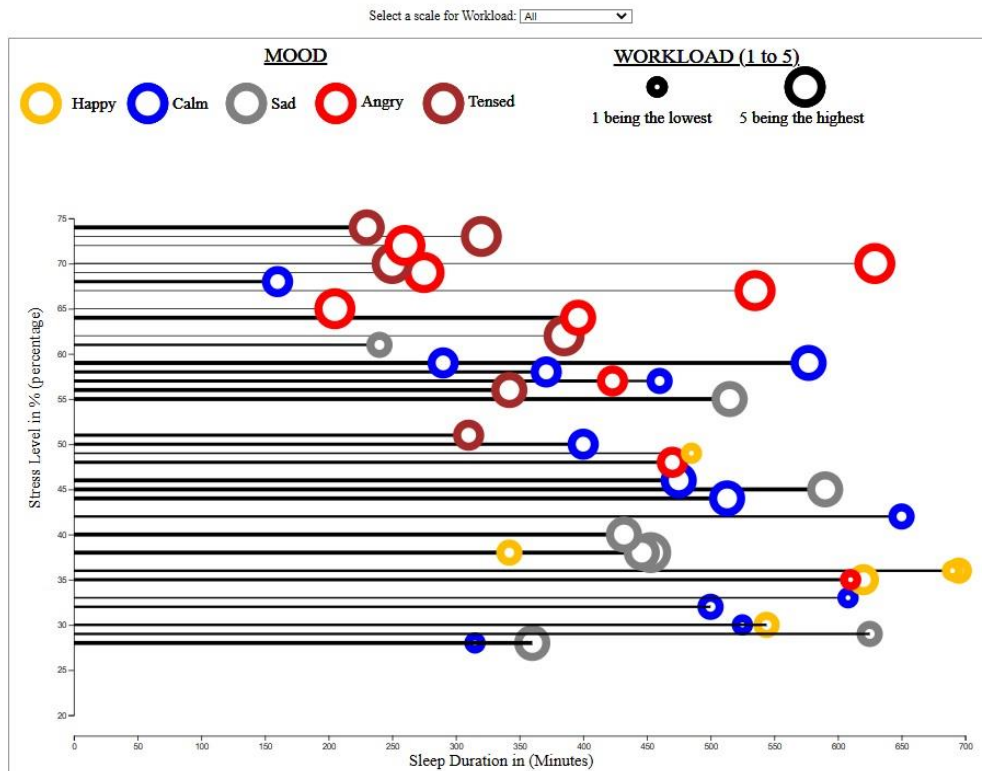
Workload data not only creates different size of circle but it also changes the stroke width of the line that starts from Y axis and connects to the circles that are plotted.

Analysis

Questions, Hypothesis Conclusion

Q1. What factors mostly affects my sleep duration?

Hypothesis: As my sleep duration decreases my stress level increases and other factor is workload.



As the above visualization shows the lollipop chart has data that affects my sleep duration. Stress level percentages are directly getting plotted as it represents the Y axis.

Conclusion

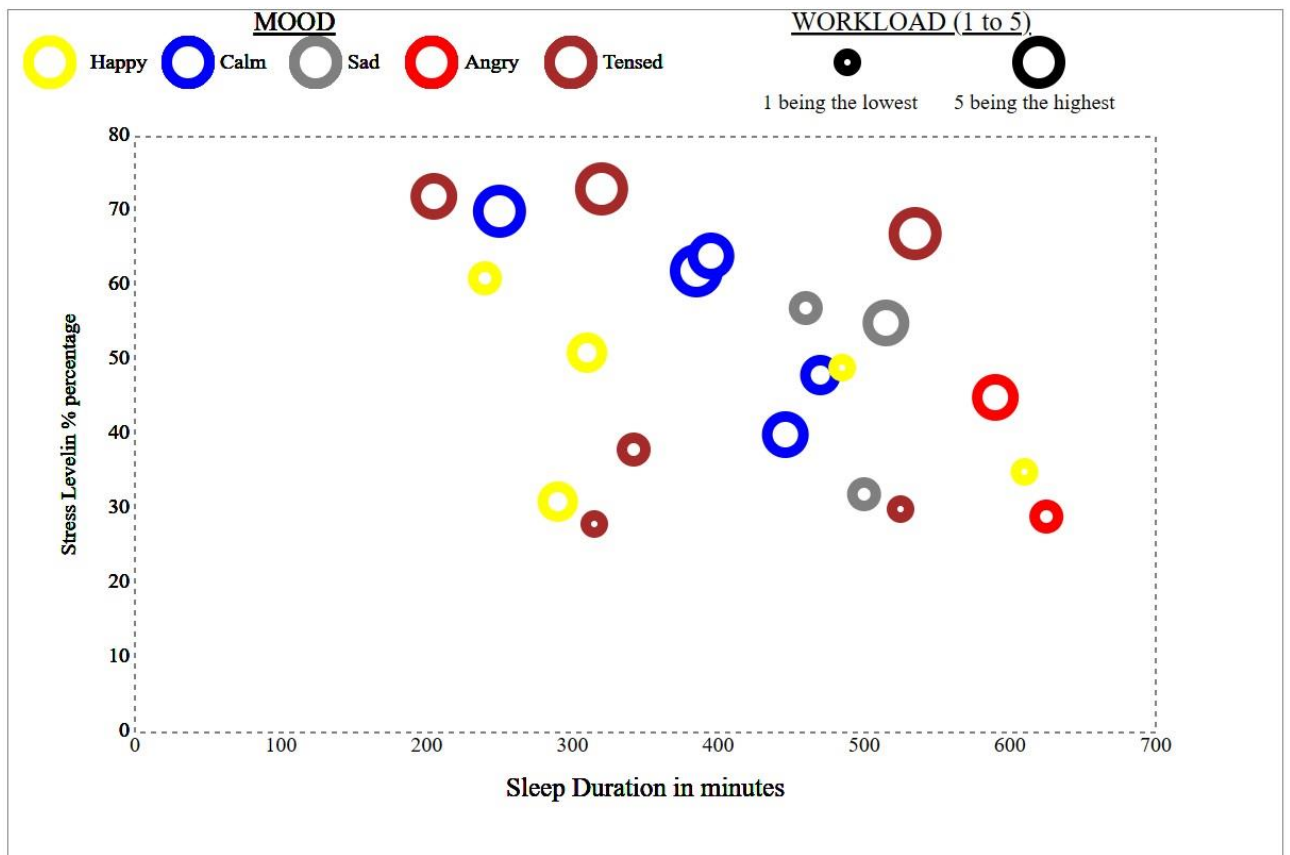
On days when I had enough time of sleep my stress level is normal but when I had less time of sleep then I have high stress.

At maximum points when the sleep duration is more the stress level is high.

If I get enough time of sleep per night then my stress level automatically gets back to normal.

Q2. How is Workload related to sleep and stress level?

Hypothesis: If my workload increases my sleep duration decreases and stress level increases.

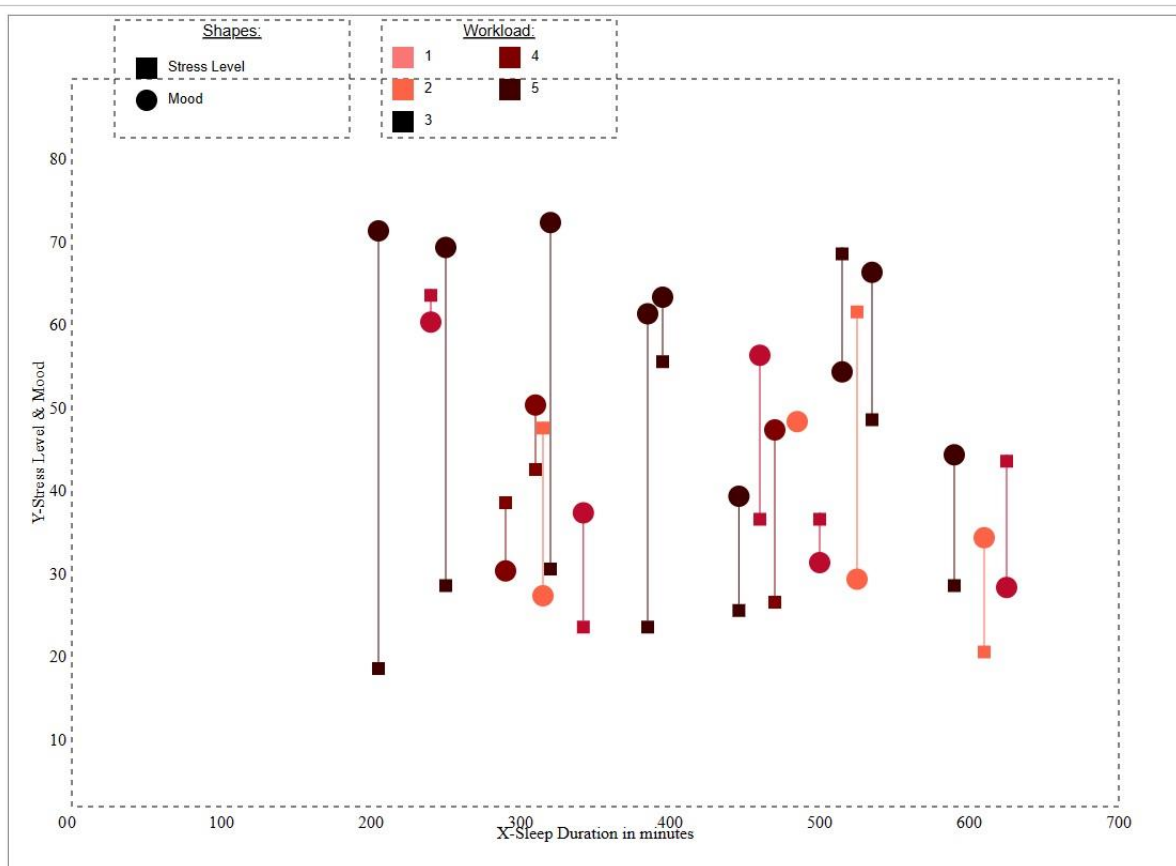


Conclusion

On days when I have workload between 3 to 5 that is from moderate-high to very-high then I have observed that I get less time of sleep per night as most of times I have to complete my work to meet the deadlines. And if I have more workload that means my stress level will go up and all these factors will definitely affect my mood. On the other side if I have less workload then my stress level will be low and my mood will be happy.

Alternate Visualizations

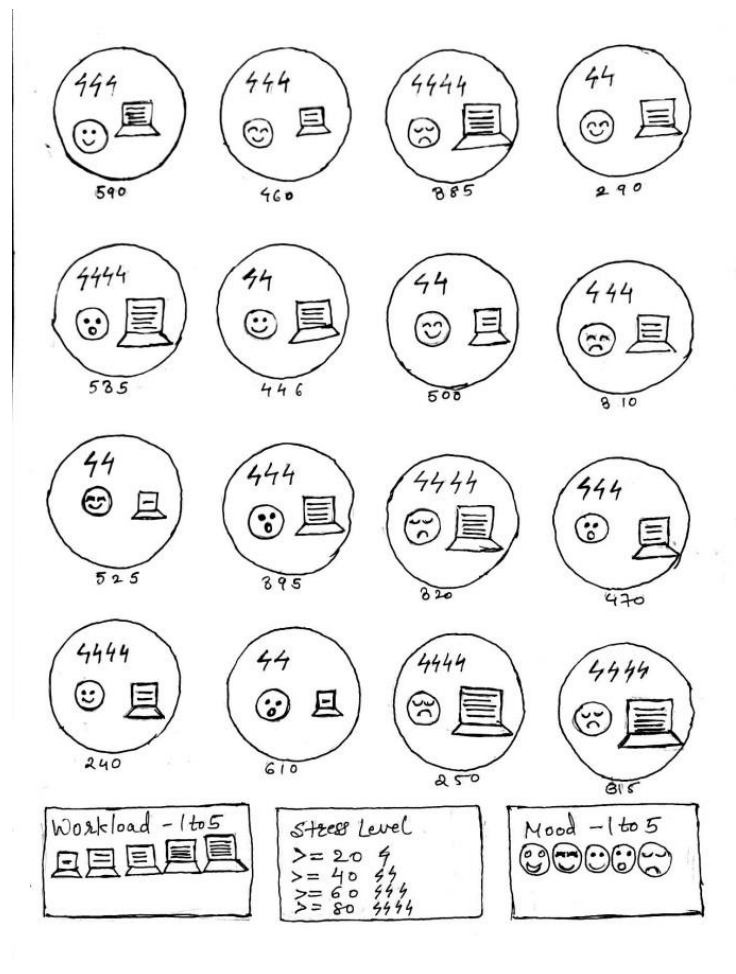
Visualization



Cleveland Dot Plot

This is the visualization that I created for my attempt 2 draft. X axis represents Sleep duration, Y axis represents Stress level and Mood, where I have added key of two different shapes for stress level it is represented by rectangle and circles for mood and the color was given by the property of workload. But when I actually look at the output to find the answer to my question the output was confusing as there are two data plotted on Y axis which is stress level and mood which I don't think goes hand in hand. If I had recorded stress level two times a day such as record the stress level at the start of the day and at the end of the day then I could have added the stress level on one axis to see what different output I get. And then instead of lollipop chart I could have used Cleveland dot plot for the final visualization.

Sketch



Creative Sketch

This is the creative sketch that I visualized but this sketch is looks more like an infographic and not something that will help me to reach a conclusion point.

Citation

Holtz, Yan. "Lollipop Chart." The D3 Graph Gallery, <https://d3-graph-gallery.com/lollipop>.