

EXPLORATION OF HOW WALKING DURATION AFFECT MY HAPPINESS LEVEL

Final Project Report Yitian Liu DGM 6108

Content

1.	Introduction	2
2.	Overview of Data Collection	3
3.	Data Collection Chart	4
3	.1 From 10/11 to 10/25	4
3	.2 Data recorded for visualization	7
4.	Rationale and Process for Visualization	9
5.	Self-analysis and Conclusion1	2
Re	ference1	4

1. Introduction

The initial purpose of this project is to see if my daily steps would affect my happiness level as I noticed that my mood changes while I am walking outside from time to time, along with other significant events, to answer which factors have the most influence on my happiness level.

Through several weeks' data collection, from October 11 to December 12[,] 2022, I recorded data to observe and verify the pattern and correlation among these events and my happiness level. Below listed objective data and subjective data I tracked for three months:

Objective data:

- 1. Date
- 2. My daily walking steps (step)
- 3. Walking duration (minutes)
- 4. Weather conditions:
 - a. Humidity (%)
 - b. Maximum Temperature (F)
 - c. Minimum Temperature
- 5. The path where I have walked (category)
- 6. The duration of video call with my parents (minutes)
- 7. The number of gummies I eat (gummies)

Subjective data:

- 1. My happiness level of walking
- 2. My happiness level of video call

During the collection process and with the development of data visualization I learned, I have explored and observed a certain pattern that how to adjust the data scale and which data I should no longer track.

2. Overview of Data Collection

At beginning as I was still exploring the relationship among my data, I firstly built my data form from recording my steps and its duration every 30 minutes when I walk outside to see my happiness level transition, which was a heavy workload for me no matter on recording or assign them into JavaScript visualization. With the increasing of the walking data I collected, the pattern is not that obvious as I expected, as it would have more other factors may affect my happiness level while I was outside. To make it clearer to see how walking steps influence my happiness, I started to track the data when I was at home to compare it.

After two weeks of data recording, I firstly tried to put my data into d3 JavaScript visualization to observe, which I used color scale for path of walking. As the locations are so vary, which is too clutter and colorful (Seen the picture 1 below) to see the pattern. Since I also add the data of walking just at home, the path category is not that necessary for my observation. In addition, I could easily sense that my happiness level was always improve while I went to supermarket or shopping, which I already have the answer in mind, so I decided no longer to track them anymore.

For the data of No. gummies I eat, as I gradually aware of too much suage has negative influence on my body and skin, I intentionally controlled the gummies number I eat per day. I also bought health gummies like Hair, Skin, Nails gummies to help improving my health situation, which should not take account into my data list, as I think it has no pleasure function to affect my happiness level. In this case, I have no longer tracked this data since November.



Does my walking duration related to my happiness level?

As my walking duration increases, my happiness level increases

What's more, where I go might also positively affect my happiness level

1.Chart with location category

3. Data Collection Chart

3.1 From 10/11 to 10/25

Date	Sessio n Durati on (minut e)	io Session Walking ati Steps (step) nut	Weather Condition		Paths (Address)	Happiness Level for	Duration of video call
			Humidit y (%)	Temper ature (Celsius)	Location Category	(0-10)	parents (minute)
10/11/20 22	10	1292	36	17	Home; Train Station	1	
	6	972	36	17	Train Station; university	2	
	26	1316	76	12	university; Supermarket	3	
	30	2914	76	12	Supermarket; Market	10	
	30	995	76	11	Market; Restaurant	10	
	30	1450	76	11	Restaurant; Train Station	10	
	24	575	76	11	Street; Train Station	10	
	39	3717	76	11	Supermarket;	7	
							11
10/13/20 22	20	1814	77	18	Home; Train Station	0	
	17	1435	77	18	Train Station; university; university	2	
	27	1364	97	17	Convenience Store; Train Station	0	

	2011349717Train Station; Home		Train Station; Home	0			
							37
							9
10/14/20 22	24 1997 82 16 Home; Superma		Home; Supermarket	1			
	36	2742	82	16	Home; Supermarket	10	
							14
10/18/20 22	19	1733	66	18	Home; Train Station	0	
	20	1866	66	18	Convenience Store; university	0	
	32	1900	58	15	Supermarket	0	
	22	2132	58	15	Train Station	3	
	22	1826	58	15	Train Station; Home	2	
							22
10/19/20 22	32	2161	36	12	Home; Supermarket	2	
	26	854	36	12	Supermarket	3	
	38	2450	36	10	Home; Supermarket	2	
							23
10/20/20 22	10	1867	33	13	Home; Train Station	2	

	24	1129	33	13	university; Turf	2	
	21	1041	33	13	turf; university	3	
	44	2146	59	12	Train station	2	
	29	1475	59	12	Train station; store	3	
	18	1484	59	12	Home; Train Station	2	
							21
10/21/20 22	16	1716	52	16	Home; Train Station	1	
	26	1596	52	16	Train station; 3 restaurant		
	38	900	70	13	store;	2	
	40	1895	70	13	Store; Train station	3	
	11	1881	70	13	Home; Train Station	3	
10/23/20 22							10
10/24/20 22	18	1931	95	16	Home; restaurant	3	
	23	1972	98	13	Home; restaurant	4	
							8
10/25/20 22	17	2041	100	16	Train station	2	
	23	1899	100	16	university		

5	386	100	16	university	
61	3175	100	15	market	
47	1914	100	15	home	
					4

3.2 Data recorded for visualization

Date	No. of Gummi Candy I eat	Humidity (%) (average)	Duration of video call with parents (minute)	Happiness Level for walking (0-10)	Walking Steps (step)	Duration (minute)	min temp	max temp
10/11/2022	1	56	11	9	13231	195	30	49
10/13/2022	3	87	46	4	5747	84	56	67
10/14/2022	4	82	14	9	4739	60	59	68
10/18/2022	3	62	22	6	9745	115	52	67
10/19/2022	4	36	23	7	5465	96	45	57
10/20/2022	2	46	21	4	9142	146	42	58
10/21/2022	3	61	10	7	7988	131	46	61
10/24/2022	5	96.5	8	3	3909	41	52	60
10/25/2022	2	100	12	9	9415	153	57	61
10/26/2022	6	98	20	2	243	17	59	65
10/27/2022	0	38	24	6	5294	131	48	67
10/28/2022	0	60	0	8	11788	147	42	52
10/29/2022	2	68	0	6	10034	124	41	54
10/30/2022	1	58	8	8	14271	202	41	62
10/31/2022	4	94	18	2	113	15	45	68
11/1/2022	3	66	12	8	11032	158	57	64
11/2/2022	4	51	19	4	2947	34	53	69
11/3/2022	3	66	4	4	4449	42	47	60
11/4/2022	3	51	10	6	6454	28	51	71
11/5/2022	4	65	9	2	77	11	58	76
11/6/2022	3	65	35	4	5525	30	66	76

11/7/2022	1	27	7	8	12980	178	55	75
11/8/2022	3	26	26	6	4205	52	40	54
11/9/2022	0	62	18	9	14893	185	36	48
11/10/2022	2	45	10	5	4417	52	44	69
11/11/2022	4	72	24	1	90	32	55	70
11/12/2022	2	26	7	7	4760	68	56	76
11/13/2022	3	84	15	3	124	24	42	56
11/14/2022	5	26	10	2	124	15	35	44
11/15/2022	4	45	15	5	4276	40	43	52
11/16/2022	3	85	22	3	1217	32	37	57
11/17/2022		47	7	7	8651	68	36	46
11/18/2022		38	15	1	1514	12	33	45
11/19/2022		41	38	4	2330	36	32	43
11/20/2022		26	22	2	6733	74	29	38
11/21/2022		41	7	1	72	8	26	39
11/22/2022		39	10	8	10128	159	35	48
11/23/2022		58	24	3	329	10	36	50
11/24/2022		62	9	2	275	12	32	47
11/25/2022		82	10	2	2064	30	42	52
11/26/2022		53	0	8	10945	142	40	53
11/27/2022		50	21	4	5135	50	44	57
11/28/2022	5	72	0	0	136	12	44	48
11/29/2022		64	22	6	4698	52	39	43
11/30/2022	0	88	12	4	1342	32	41	57
12/01/2022		38	10	8	5658	73	37	43
12/02/2022		47	48	5	4742	52	34	45
12/03/2022	3	89	0	5	3026	35	39	59
12/04/2022	2	39	0	4	1575	21	34	43
12/05/2022		57	23	2	2507	22	30	48
12/06/2022	1	75	12	9	6730	68	41	50
12/07/2022	3	64	0	2	43	10	48	59
12/08/2022		48	18	7	6488	54	37	57
12/09/2022		64	15	5	155	17	34	46
12/10/2022		72	0	1	30	8	30	36
12/11/2022	0	74	20	3	1678	16	25	30

12/12/2022	1	40	7	2	2334	30	25	34
12/13/2022		39	45	9	6283	72	27	41
12/14/2022	4	48	20	0	44	10	23	36

4. Rationale and Process for Visualization

From all beginning when I tried to figure out how to express my data by coding, obviously I applied scatterplot as my first choice because this is the type I have learned from the class. It is true that scatterplot could show the relationship between data clearly and intuitively. However, it is easy to be limited by what you have learned from the class while I tried to be creative for my term project. Since changing r scale and color scale for scatterplot is a usual way, it is harder to break it and add more new scales as I have not yet researched at first.

Things changed while I was drawing my sketches. I have not expected that I would achieve them into code drawing in the sooner future, so I use my imagination and draw two visualizations different from scatterplot. Inspired by "Dear Data", which is an interesting and inspiring book, I have created an area chart (picture 2) and a chart with color leaves (picture 3) as my data expression.



2. Sketch 2





3. Sketch 3

From my second sketch, I started to do more research on how to draw area chart by using d3 as well began to consider which chart would express my data best.

The website "From Data to Viz" indicates that stacked area is suitable for observing the evolution of the whole and the related proportions of each group. In this way, comparing to single area chart, the stacked chart could better express and compare several data in visualization. Therefore, I tried different data to put them in a stacked area chart to see if it could be clearly compared and show the pattern as I expect.

For y-axis, it should range as number to calculate the total of a numeric variable, but also the expected range of each data set should be similar in case the areas are too narrow or too wide to observe.

For how to choose and build x-axis, I have done some research. According to *Fundamentals of Data Visualization*, data visualization also needs to consider aesthetics features, while all aesthetics fall into one of two categories: those that can represent continuous data, and those that cannot. To the point, the length of time is a continuous value (Wilke). In

addition, the date could show the overall changes of each day that how my duration changes, and how my happiness level changes. Therefore, I decided to apply date as my x-axis.

In addition, to better convey my visualization, I did research on how to assign color to each area. From website "Coolors", it showed trending color scheme and matching, I used them as my color reference.

When firstly building the draft of my visualization, the result is not as clear as I expected. I assigned walking duration, video call with parents' duration, humidity, happiness level as stacked area. Since the units for each data are different, it is hard to compare the relationship. Then I added a scatterplot to express my video call duration as it has narrow range in stacked area (picture 4).



As my walking duration increases, my happiness level increases.

In addition, humidity also positively affect my walking duration I also found that the longer time I spend on walking, the less time I spend on video call with my parents



I do believe that it is a creative work, but the pattern is still hard to tell unless to make a detailed observation of the trend of the line and area. Also, it is confusing to have circles on the line but need to see the trend of the line.

Finally, I realized that due to the non-uniformity of the y-axis units, it is difficult to compare these data on stacked area. So, I only left two durations as my stacked area, and built dual y-axis on right to express scatterplot of happiness level scale. Also adjusted the contrast of the color so that the narrow part can be shown.

To make my work more completed and easier to understand, I was thinking about adding

an interactive function. I successfully add tooltip for my scatterplot, but I was stuck on to add tooltip and zoom function for my area.

5. Self-analysis and Conclusion

At start point of this part, I would restate my hypothesis here: As my walking duration increases, my happiness level increases. During the process of learning d3 data visualization and collecting data, I have found several interesting discoveries by inverting different data into different charts.

My first and main discovery, which also is my conclusion, prove my hypothesis that the more time I spend on walking, the happier I am. Apart from my main hypothesis, I also observed from my data that the number of gummies I eat is inversely proportional to my walking duration (showed in my exploration plot 3), which I did not expect the correlation between them.

I also explored if weather condition (humidity and temperature) would affect my walking steps and duration. It is true that the most cases of long duration walking happened on weather condition with low humidity and high temperature. But for this point, I already knew that sunny day would make me happy and feel to walk more, and that is the reason I did not put weather conditions into my final visualization.

In addition, from my final project draft version, I surprisingly discovered that the walking duration on weekends is far lower than I expected, while I spend more time on walking on weekdays. It indicates that I prefer to relax at home on weekends rather than go out, which is a fun result I have not recognized before.

To summarize, other conclusions through project:

- ♦ The colder weather is, the less walking duration, from December, my walking duration obviously reduce compared to Oct and Nov.
- \diamond The more time I spend on walking, the less I spend on video call with my parents.
- \diamond As my walking duration increases, the number of gummies I eat decreases.
- \diamond I spend less time on walking on weekends compare to weekdays.

Before this class, I have tried to collect three good and happy things happened in one day for mindfulness. By writing down three good things to form a habit to think in a positive way and maintain emotional stability. This recorded event does not aim to prove or explore the relationship among several recorded data but do help me to form a recording habit. I would like to continue exploring other correlation in the future:

- ☆ Topic: The relationship between recording three good things and my mood. I would collect data of two months on recording three good things, the workout duration, the music duration, and my mood level, and another two months without recording the same data but without recording three good things every day to compare the possible difference on affecting my mood.
- ☆ Topic: Which factors mostly affect my sleeping quality. As I think sleeping quality and duration has a significant impact on my health, skin, and energy level, I want to use data-recording technique to explore which factors affect my sleep most. I would collect my sleeping duration and quality though sleeping tracking app, the cups of coffee I drink, the cups of milk I drink, my study duration, my workout duration.

If I have more time on my final project, I intend to do more research and figure out adding tooltip and zoom interactive function to my visualization. In addition, I also would like to add more data to the chart to observe more correlations. Thanks to this course, I already knew the basic JavaScript and D3 techniques to build my own data visualization, which also inspired my interest to learn more about data visualization knowledge and apply it for further future programs.

Reference

Healy, Yan Holtz and Conor. "Find the Graphic You Need." *From Data to Viz*, <u>https://www.data-to-viz.com/</u>.

"Trending Color Palettes - Coolors." Coolors.co, https://coolors.co/palettes/trending.

Dehghani, Zhamak. Data Mesh: Delivering Data-Driven Value at Scale. 1st ed., O'Reilly Media, 2022.

Wilke, Claus. Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures. 1st ed., O'Reilly Media, 2019.

Murray, Scott. Interactive Data Visualization for the Web: An Introduction to Designing With

D3. 2nd ed., O'Reilly Media, 2017.

Bostock, Mike. "D3 Gallery." Observable, 31 Aug. 2022, observablehq.com/@d3/gallery.

Stacks. using-d3js.com/05_06_stacks.html.