



Social Interaction Tracker

Link to the Working Prototype:
<http://socialtracker.someshrahul.com/>

SI 649

INFORMATION VISUALIZATION

By:

SOMESH RAHUL

Introduction

Social Interaction Tracker is a visualization tool that leverages the data about a person's social interactions in figuring out the following

1. How much social interaction an individual is having each day and with whom?
2. How much social interaction is optimum to get a good productive day?
3. Does social interaction with a specific group/sphere of people impact productivity?

Notes about the current state of the tool:

Currently, the tool is highly personalized to me, hence, I have made the following assumptions

1. The social interactions are categorized in three social spheres:
 - a. Friends
 - b. Family
 - c. Colleagues
2. The productivity of each day is determined by how much work I get done on that day and how I felt that day. I have categorized it in
 - a. High
 - b. Medium
 - c. Low
3. The primary consumer of this data is I.

The above assumptions in no way undermine the functionality of the tool. The visualizations that the tool shows are well equipped to answer the primary questions that the tool was designed for.

Data Collection Methodology

I collected my body data in terms of how much interaction I was doing with each individual. Towards the end of each day, I tagged these individuals with one of the three social sphere category that I was using for the purpose of this research. There were instances when an individual did not satisfy the basic requirements for any of the three social spheres, in those cases I chose not to include that data. Also, some of the individuals overlapped in two social spheres. In those cases I chose the social sphere that was more dominant (personal choice). For the purpose of this research, I chose to enter only the aggregated data in the csv file that I was using. I didn't enter the individual data.

Hence, for each day I had the following set of data.

1. Day Number
2. Number of hours I interacted that day with my social spheres
3. Number of hours I interacted with the social sphere of Friends
4. Number of hours I interacted with the social sphere of Family

5. Number of hours I interacted with the social sphere of Colleagues
6. What was my overall productivity for that day

I collected this data for a period of fifteen days. Here is a partial snapshot of the csv file:

day	hours	friends	family	colleagues	productivity
1	8.5	2.5	2	4	med
2	9	4	1	4	high
3	10	3.5	2.5	4	high
4	4	2	1	1	low
5	3	0.5	0	2.5	low
6	2.5	1	0	1.5	med
7	7.5	2.5	3	2	high
8	8	2	3	3	high
9	5	1	1	4	low

I chose to collapse the individual collected data into three social spheres, as that would help me in answering the question about impact of different social spheres on my daily productivity. Also, collapsing would help me understand my behavior as a whole, similar to what was revealed in the paper *Calendar View*, van Wijk et al., 1999.

Since I collapsed the data into three categories, hence, the social interaction data becomes categorical / grouped. Days are represented as ordinal (ordered data) and productivity is nominal (labels of high, med, low).

The key relationships within the data are:

1. Each day is associated with the number of hours of social interaction and the productivity on that day.
2. Number of hours of social interaction are subdivided into interactions with – Friends, Family and Colleagues

Design Methodology

The data that I have collected is temporal in nature. There are days and each day corresponds to definitive set of data. I was primarily interested in data and not the event; hence I have collected the data for all the days and not for particular events. Moreover, no such event happened while I was collecting the data and if there had been an event then I would have chosen to ignore the data for that day, as it would have skewed by research.

Therefore, to represent the data that is categorical, temporal, linear and static I shortlisted three choices [Muller + Schumann, 2003]

1. Sector Graphs

2. Stacked Bar Charts
3. Circle Graphs

Next, the research questions that I wanted to answer were about pattern detection and decision support. Hence, the visualization that I would choose should facilitate the above.

The problems with the circle graphs were that for categorical dataset there was too many cross edges and that was making it hard for the user to understand the data. Second, in case of same values for the variables on the same day, there was overlapping, which made the parts of data illegible.

Sector Graphs and Stacked Bar Charts both suited my purpose, but I preferred the Stacked bar charts because I wanted to have a quick comparison between the days in terms of overall social interactions. Further, there was this issue of scalability with Sector Graphs. If I wanted to represent more than 10 days on the graph then each sector was becoming really thin. Another problem with sector graphs was that the fill inside each sector was giving the wrong sense of area. Moreover, with stacked bar graphs the comparison between the interactions that I had on each day was easier since length is perceived as it is. [Length estimated correctly – <http://makingmaps.net/2007/08/28/perceptual-scaling-of-map-symbols/>]

Each bar in my stacked bar chart represents my overall social interaction for that day. It has three stacks (categories), corresponding to the social interaction with friends, with family and with colleagues respectively. Based on the Bertin's "Level of Organization", I used visual encoding of color to differentiate between the stacks of friends, family and colleagues. Further, I used the color hue to make the differentiation easy, I didn't use color intensity as that might be leading for some users. Finally, I chose the days as my x-domain and hours for my y-domain.

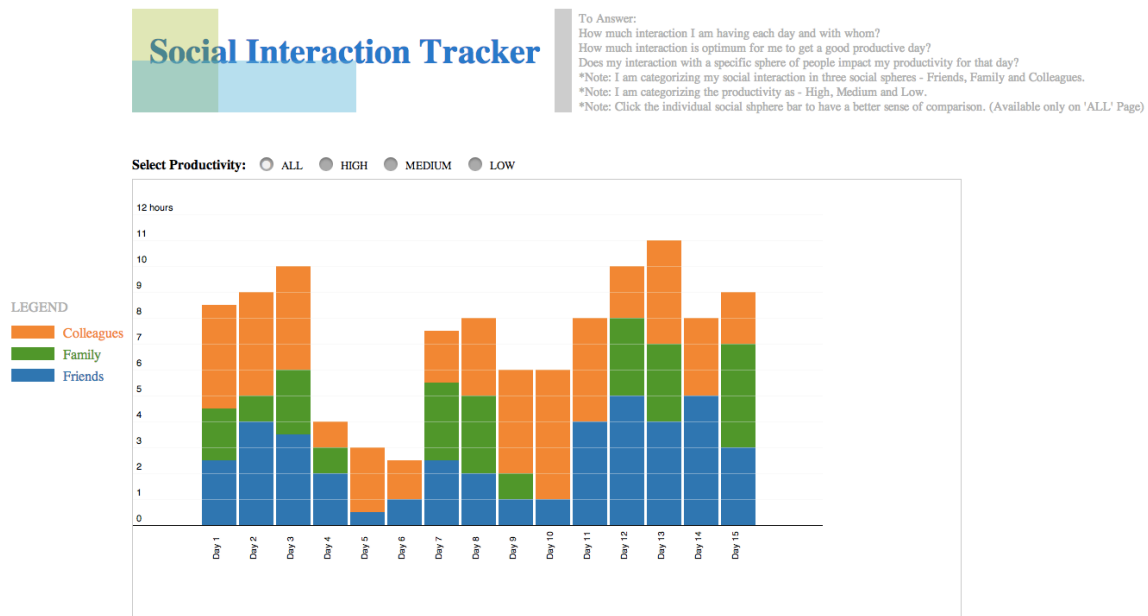
Another problem that I wanted to solve was the easy comparison between the stacks of each interaction type. That is, if I want to compare the amount of interaction that I had with my colleagues on each day, then I should be able to do that easily. With stacked bar charts this comparison is hard to make since the baseline is different for each stack (category), except for the first stacks (category). To solve this problem I made use of the solution provided by Jane Greystoke [<http://www.meandeviation.com/dancing-histograms/hist.html>].

Design Description

The home page of my tool shows the visualization, legend for the visualization, radio buttons and my research questions.

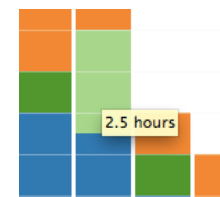
The radio buttons govern the data that will be presented by the visualization. The governing rule is – Productivity. That is, if 'High' is selected in the radio button then the data of social interaction for the high productive days will be shown. By default,

the 'ALL' radio button is checked. This means that the entire data of social interaction for all the days will be presented in the visualization.

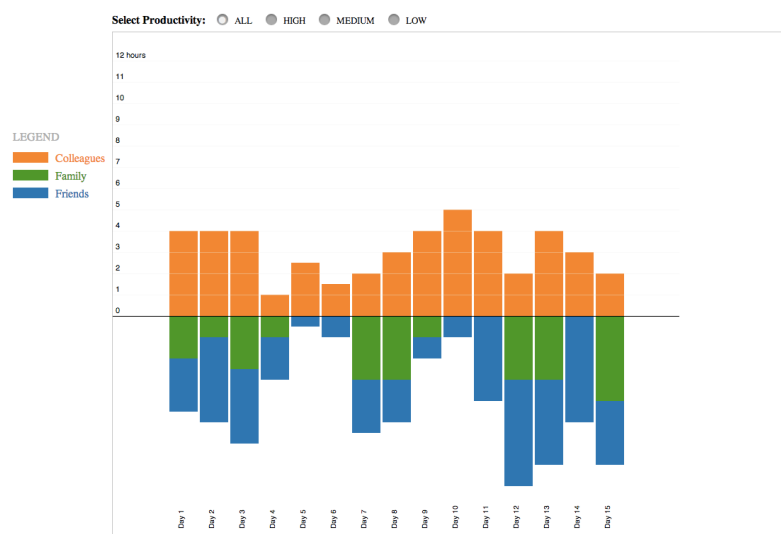


Each bar corresponds to the number of hours I had interaction on that day and each stack (category) in the bar corresponds to the number of hours I had interaction with that social sphere, for that day.

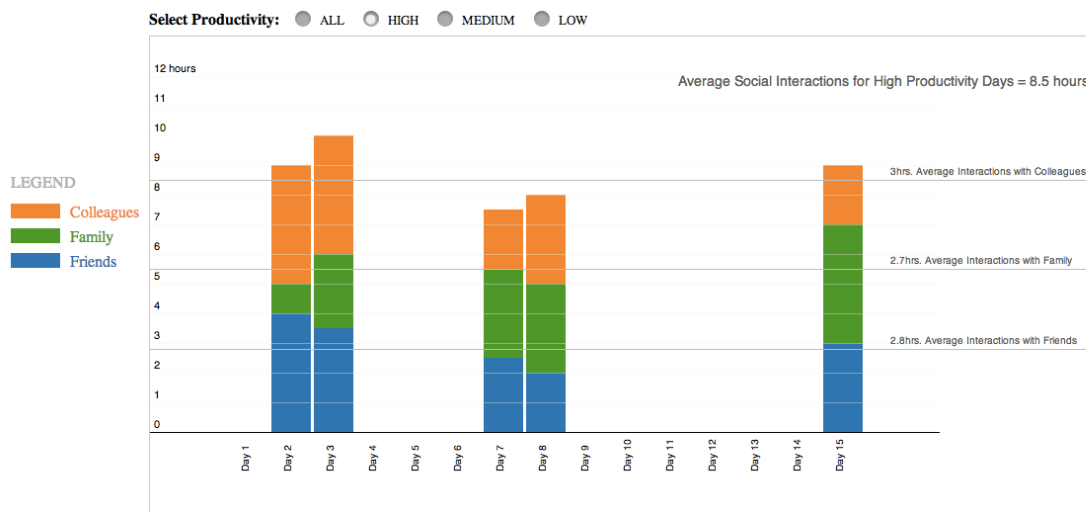
If you hover over any stack, then the hours of interaction will be popped up for that stack. On hover, the color of the stack also changes, so that the users can relate the block with the hovered value.



Next, to enable the comparison of a particular type social sphere, I implemented the click functionality. So if you hover, over a particular stack element and you click on it, then social sphere corresponding to that stack will come to the baseline, for easy comparison. Example is shown with social sphere - Colleagues.



To spot the patterns for the high productive days, you need to click on the radio button that says 'HIGH'.



Here, the user will be able to see the additional information of

1. Average Interactions with Friends: On high productive days how much average interaction did the user has with his friends. In this case, the user had around 2.8 hours of social interaction per day with his friends.
2. Average Interactions with Family: On high productive days user had 2.7 hours of average interaction with his family.
3. Average Interactions with Colleagues: On high productive days user had 3 hours of average interaction with his colleagues.
4. Average Social Interactions for High Productivity Days: This shows the average value of social interactions that the user has on high productive days.

Similarly, the data can be viewed for the medium and low productive days by clicking the 'MEDIUM' and 'LOW' radio buttons.

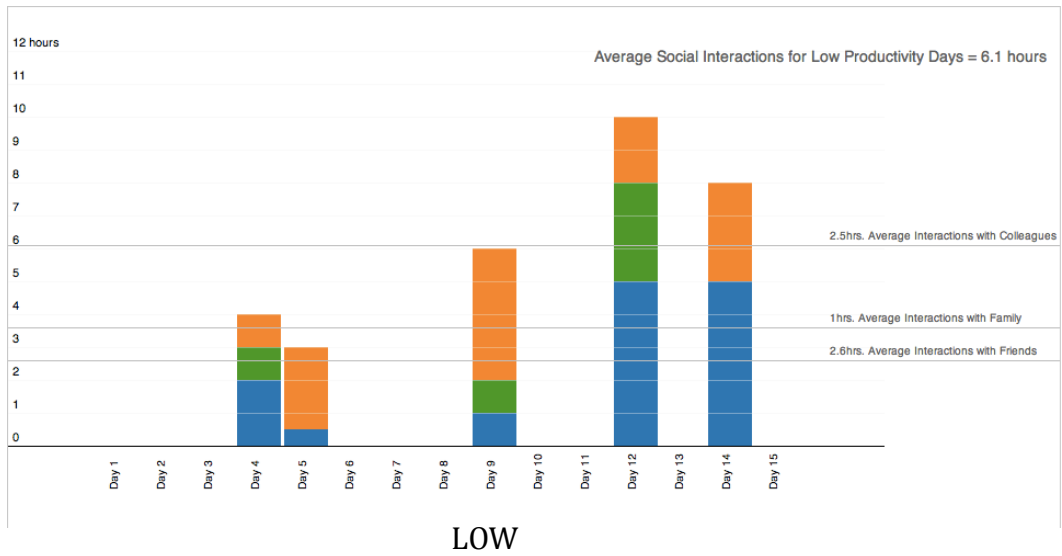
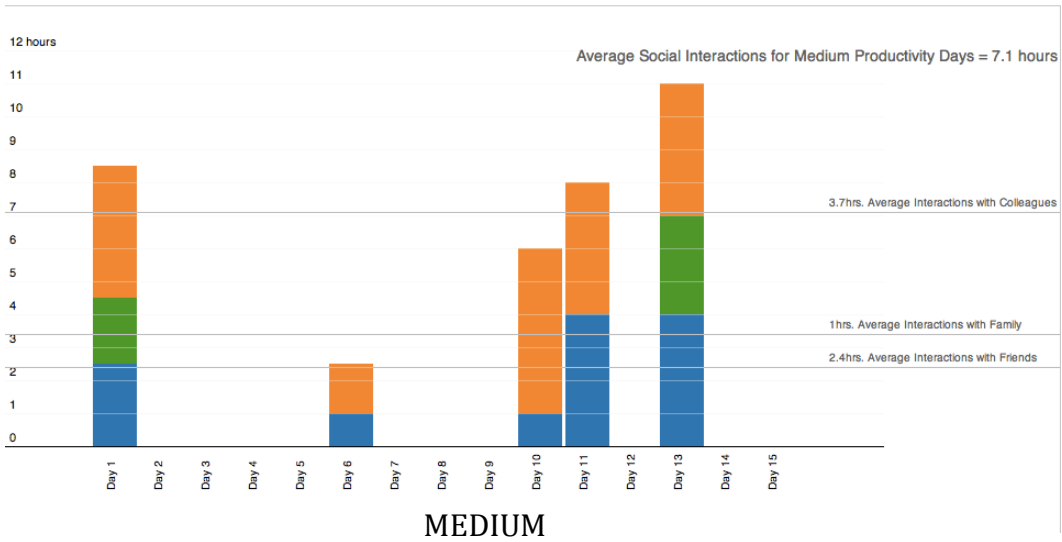
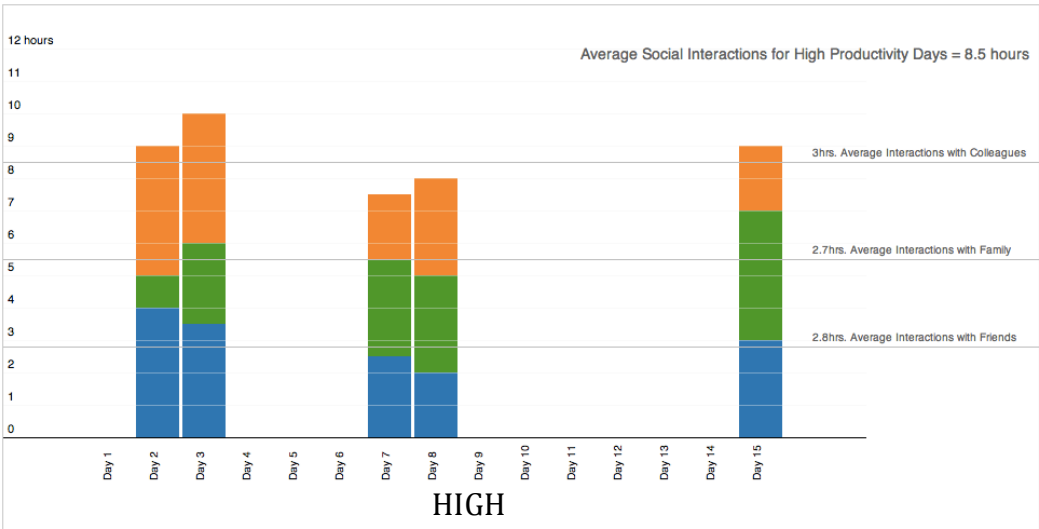
This complete visualization helps the user to understand how much interaction he has on a daily basis and how those interactions are affecting his productivity. Based on the data, it seems that I have high productive days when I have higher amount of social interactions. Further, I can do a microanalysis from this data to find out that within the social spheres, interaction with which of the spheres is affecting my productivity the most. Looking at the data I found out that on high productive days I have large amount of interactions with my family.

Below is the comparison of my interactions on three types of productivity days:

Select Productivity: ● ALL ● HIGH ● MEDIUM ● LOW

LEGEND

- Colleagues
- Family
- Friends



Future Work

At present the Graphical Integrity as described by Edward Tufte, breaks at a couple of places in my visualization.

1. In the high, medium and low productivity pages the lines that represent the averages are at disconnect with the values on the Y-axis. I chose to ignore this disconnect because putting all the average lines matching the axis values will result in a greater disconnect with respect to the social spheres for whom the averages are. But, I intend to maneuver this disconnect with more research in future.
2. Clicking the stacks on the 'ALL' page results in base lining all the stacks that correspond to the clicked social sphere. This doesn't happen on the other productivity pages. I chose not to resolve this currently as it required greater amount of time and efforts.

Further, I plan to include animations when the changes in graph happen based on the radio buttons or clicking.

Conclusion

The entire process of tracking my social interactions and then creating the visualizations based on that was extremely valuable. I learnt a lot about myself and also was able to answer the major research questions that I had before starting the research.

Now I know that I am having around six to eight hours of social interaction daily. Further, on high productivity days I have more social interactions. Also, on high productive days I have more interactions with my family.

In a nutshell, I am an extrovert, and I need more family interactions to have high productivity days.