

CS171 Process Book

Lennon Bensaou and Andrew Zhou

1 November 7, 2016

We have completed our detailed project plan, along with 3 sketches of visualizations (Figures 1 and 2), an interaction storyboard (Figure 3), and a preliminary website layout (Figure 4). We have additionally begun searching for data and have found the National Comorbidity Survey (NCS), National Survey on Drug Use and Health (NSDUH), National Mental Health Services Survey (NMHSS), and Google Ngram Viewer to be the most promising sources of data.

We have identified seven tasks and three goals for our project, which will serve as a framework as we proceed. We will enumerate these tasks and goals here:

Our tasks are addressing: 1) Introduction and Scientific Background; 2) Demographic Information; 3) Symptoms, Causes, and Other Associations; 4) Treatments; 5) Comorbidity; 6) Social Understandings and Perceptions of Mental Illness; and 7) Mental Health Resources.

Our goals are to: 1) Raise awareness of mental health issues afflicting Americans; 2) Change the general public's perception of mental illness and decrease societal stigma; and 3) To provide scientific and sociological information about the history of both mental illness and our understanding of it.

For a more detailed explanation of these tasks and goals see our project plan.

We have included our sketches in this entry of the process book to give an idea of the starting point of our visualizations. We provide sketches of our demographic/association-based choropleth and bar chart in Figure 1, our comorbidity pie chart in Figure 2, our pie chart interactions in Figure 2, and our preliminary website layout in Figure 3. (Figure 1 is 2 sketches.) Note that some of our images are displaced due to the typesetting environment.

We have added Figures 1, 2, and 3 in this update.

2 November 14, 2016

We have completed our primary data scraping and cleaning, and partially implemented one of our visualizations. Additionally, we have blank frameworks in place for two more. We have created our first design of an innovative view, implemented our rough webpage design and structure, and designed our interactions.

We have obtained and processed our primary datasets, the Collaborative Psychiatric Epidemiology Surveys (CPES) and the National Survey on Drug Use and Health 2014. Both are very large files with many survey questions, so we have filtered them down to the most relevant ones: currently 22 for the CPES and 9 for the NSDUH, all having to do with particular mental health diagnoses and demographic data. These data are located at cpescondensed.tsv and nsduh_14_condensed.tsv. We will likely end up using more questions from the survey in the end, but we have written a script (processing/processData.py) to automatically filter and clean our data, so it will be trivial to get information for more fields. We plan to use the CPES for our comorbidity pie chart and demographic bar chart and the NSDUH for our choropleth, since the NSDUH has specific state data and the CPES does not. Accordingly, we are tentatively unlinking our choropleth from the selector and confining it to simple hover interactions as the NSDUH dataset is much less comprehensive in terms of demographics. We will note the usage of different data sets explicitly on our website. We are also using Google Ngram Viewer for our final visualization; we have not finalized the specific terms we are using so we will include that data later.

We have partially implemented our pie chart, as can be seen in Figure 4. This chart will be interactive (See Figure 2, right side) when we finish it; currently it displays the number of people who suffer from certain numbers of mental illnesses. The data for this chart are from the CPES. We have also put in place placeholders for our choropleth and our demographic bar chart, which we illustrate in Figure 1. Additionally, we have included a spot for our innovative visualization for Mental Illness Terminology Over Time, an evolving interactive word cloud, which we illustrate in Figure 5. We illustrate interactions for our pie chart in Figure 2; for our bar chart in Figure 6; and for our word cloud in Figure 5.

We have begun the implementation of our layout. It largely follows our proposed structure, though our tentative unlinking of our choropleth and bar chart has shifted the configuration slightly. We add important quotations about and information concerning mental health between each section. See Figure 7 for a screenshot of this layout. Our storytelling approach will follow this layout, giving a brief scientific and historical overview of mental health at the top and progressing through a sequence of visualizations as the user scrolls down. We first present the choropleth to give the viewer a tangible glimpse at mental health in America; then to the bar chart to allow them to see specific data concerning the incidence of mental illness; then to the comorbidity pie chart to indicate how many Americans suffer from multiple illnesses, and how this comorbidity affects them; and finally to the word cloud to display the historical evolution of mental health perceptions. We believe that by beginning with a choropleth and ending with a timeline-based word cloud, the viewer will appreciate the extent of the problem both in modern society and historically.

We have added Figures 4, 5, 6, and 7 in this update.

3 November 21, 2016

We have reassessed our datasets and changed some of our objectives of the project. The text-based datasets have been too much effort for us to attempt to unravel, and so we are relying principally on the CPES for our visualizations. Additionally, we reviewed the NSDUH and realized we were mistaken about it having sufficient geographic information for us to attempt a choropleth. We

have thus modified the focus of our project.

We are going to implement a bar chart indicating incidence of illness by demographic. Additionally, we are going to extend our pie chart with an interactive function: if we select a “central illness,” the pie chart will indicate the comorbidity rates of individuals with that illness; thus, we may notice which illnesses are more strongly correlated with comorbidity. Finally, we are implementing a different creative visualization. This visualization is an adaptation of a sunburst chart; however, rather than strictly portraying hierarchical data, it has been repurposed to indicate subsets. We do impose some measure of hierarchy on the data by deciding which illness will correspond to which ring, but we do not need inner rings to be filled in for a particular angle for the outer rings to be filled in. Thus, we may visualize the incidence of any particular subset of n illnesses: each of the 2^n subsets will have its own sector. We will also allow the user to select a “central illness” (tied to the pie chart) and thus filter the data to only individuals with that illness. See a mockup of our design in Figure 8.

4 November 28, 2016

We have finished implementing our principal charts: the demographic bar chart, the pie chart, and the sunburst chart. Though this is fewer charts than we would have liked for this checkpoint, we consider ourselves to have implemented the “must-have” components of our project. We would ideally like to have another chart for associations/symptoms, which would likely be a modified version of the first bar chart; we will aim to implement this component prior to the final due date and presentation.

We have implemented interactivity for all of our charts, and linked our pie chart and our sunburst chart. Our bar chart allows the user to select demographic and illness to display, while the pie chart and sunburst chart allow the user to select the “central illness” by which to filter the dataset. Additionally, with a series of checkboxes the user may select a particular combination of illnesses; our site then displays the incidence of individuals with that combination of illnesses (either within the general population or within the population with the “central illness”) and highlights the relevant sector on the sunburst chart. We would perhaps like to add the ability to reorder the hierarchy of illnesses (currently in order of prevalence) before the final due date. See Figures 9 and 10 for screenshots of our progress.

We have additionally finished a preliminary writeup of the text for the website. Currently we have focused on explanations of the interactions and charts on the site, but we will be working on adding more scientific explanations in the future.

For our final checkpoint, we would like to hopefully finish the association visualization, add different demographics by which to filter (which simply requires a bit more data processing), and allow the implicit “hierarchy” in our sunburst chart to be ordered by the viewer.

5 December 5, 2016

We have finished implementing our association visualization, which can be seen in Figure 11. In addition, we have continued to refine the text of our website and have added information on the DSM and links to outside resources for viewers. Furthermore, we have refined the sunburst chart to allow the user to reorder the rings of the sunburst chart so as to look at more specific subsets (given that our sunburst chart imposes a hierarchy on the ordering of the diseases).

Additionally, we have received our peer evaluations, which may be seen in Figures 13 and 14. We took tips from our peer evaluators on how to refine the storytelling of the website in particular to highlight information that we felt was prominent.

6 December 12, 2016

We edited our text further, commented our code more thoroughly, created our screencast, and uploaded the website. We turned in the assignment!

Some reflections: we changed the scope and focus of our project drastically over the past month or two due to difficulties with datasets and unearthing additional insights from our own. Probably the most significant change was pivoting from a text-based “creative visualization” to our sunset sunburst chart, which took significant amounts of thinking and math to design and implement. It was a fruitful process overall, and we learned a lot about the iterative design process and about incorporating feedback and additional brainstorming into a finished product, as well as about an important topic: mental health. We hope the website will be informative to all viewers!

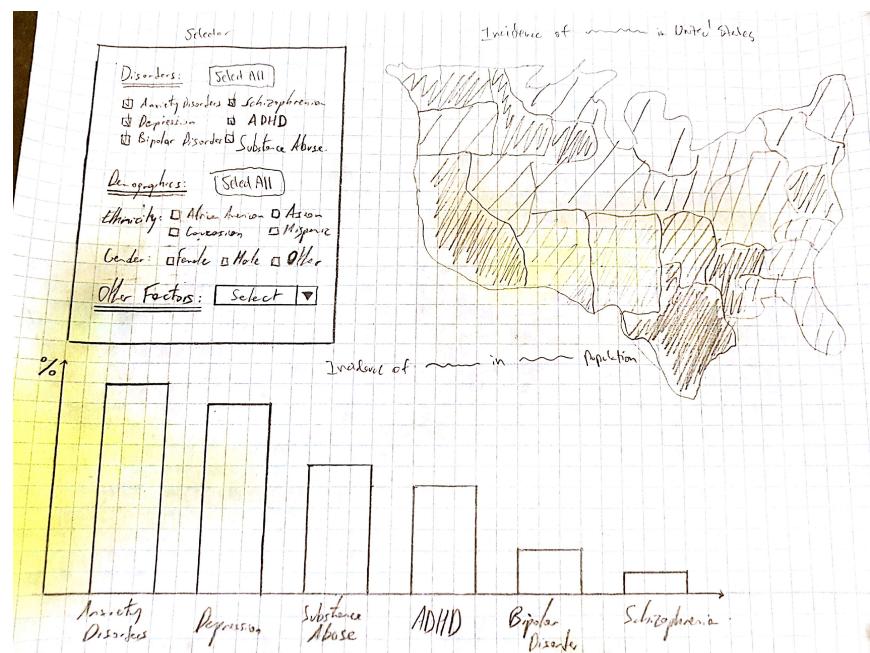


Figure 1: Bar Chart and Choropleth of Mental Illness by Demographic and Association

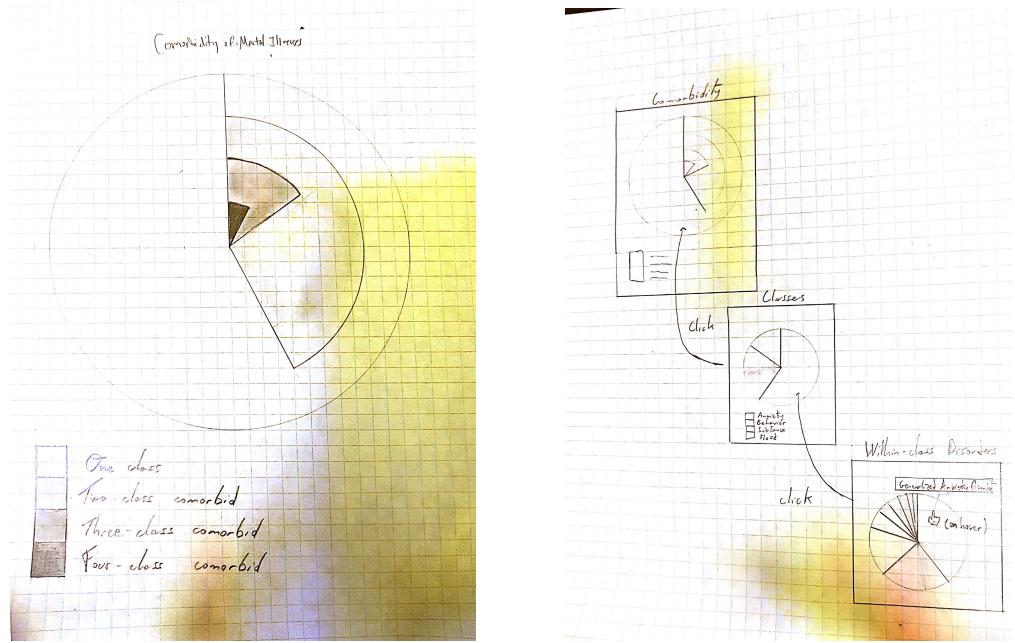


Figure 2: Comorbidity Pie Chart (left) and Pie Chart Interaction (right)



Figure 3: Website Layout and Storytelling

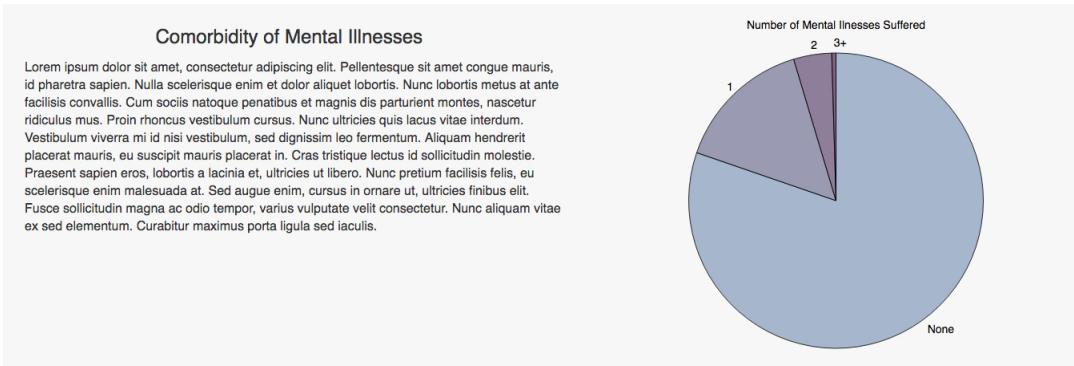


Figure 4: Pie Chart Preliminary Implementation

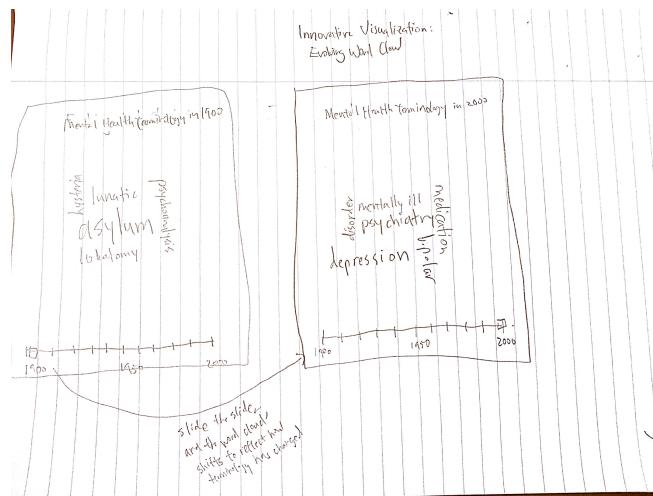


Figure 5: "Innovative Design" - Evolving Interactive Word Cloud

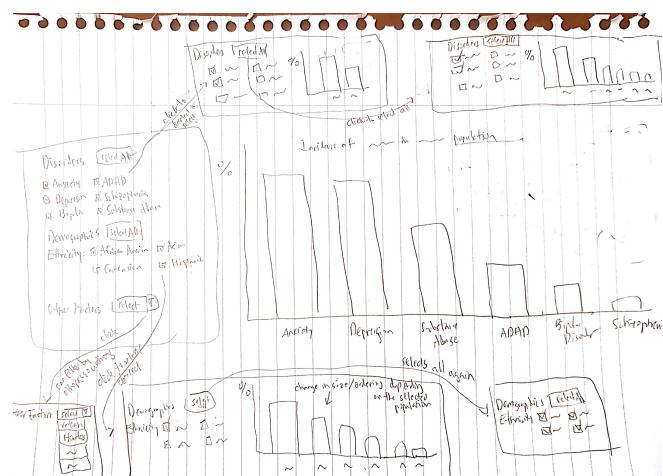


Figure 6: Bar Chart Interaction Storyboard

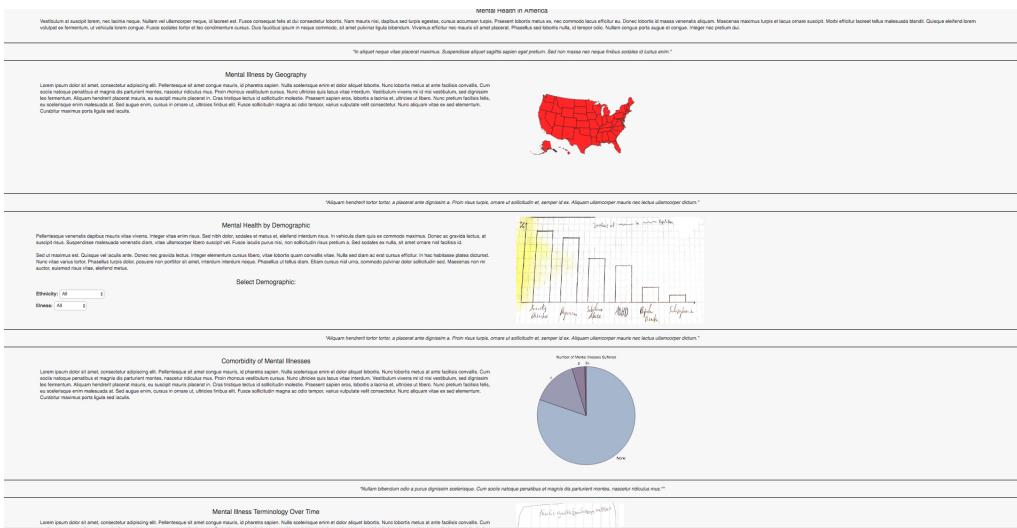


Figure 7: Storytelling and Website Layout Screenshot

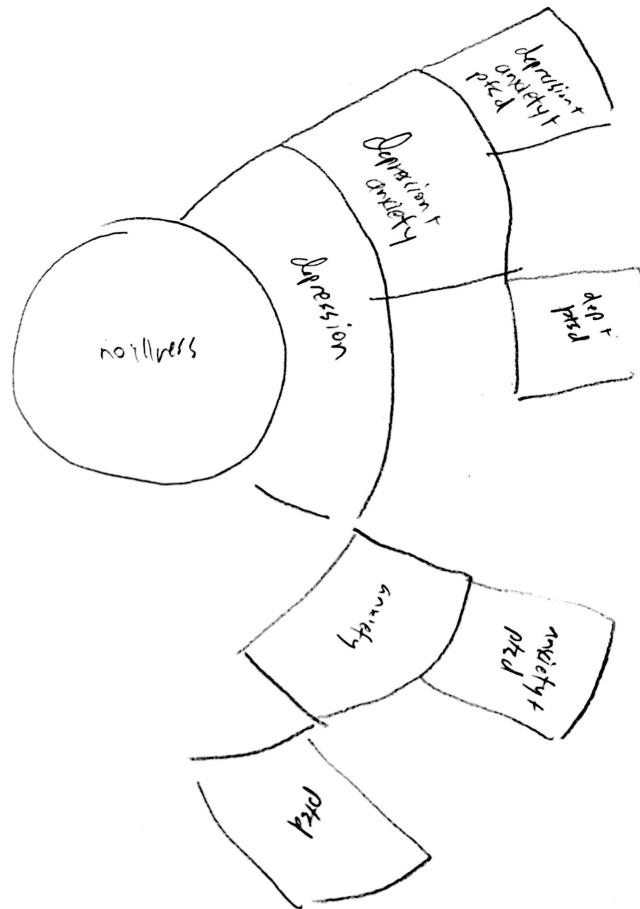


Figure 8: Sunburst Subset Chart

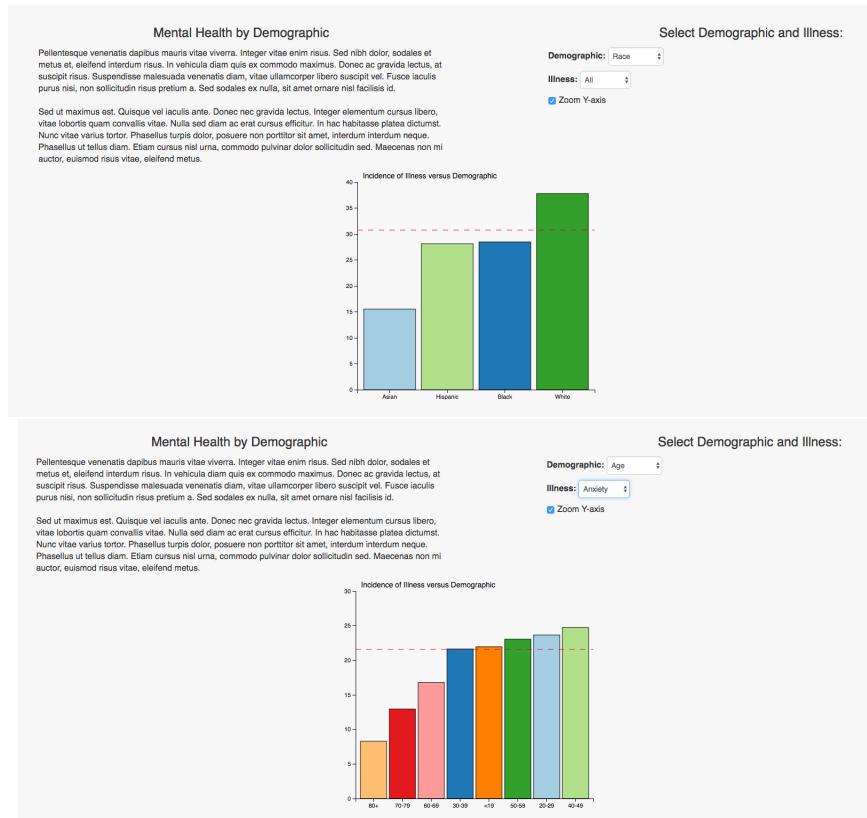


Figure 9: Demographic Bar Chart and Interaction

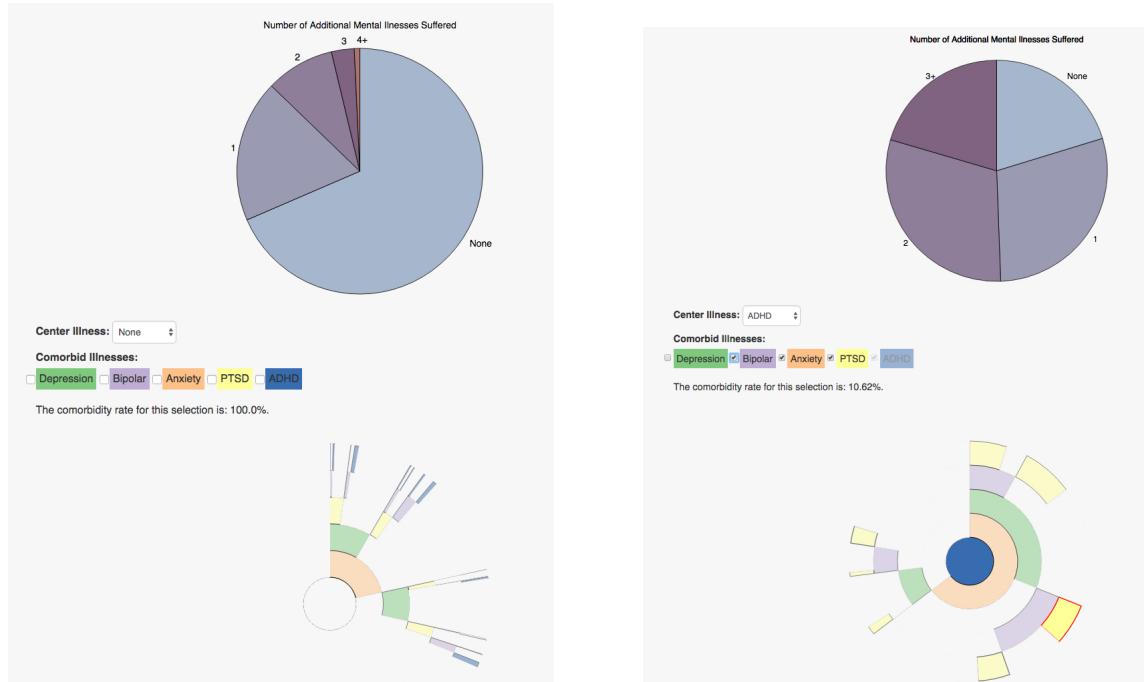


Figure 10: Pie Chart, Sunburst Subset Chart, Interaction, and Linking

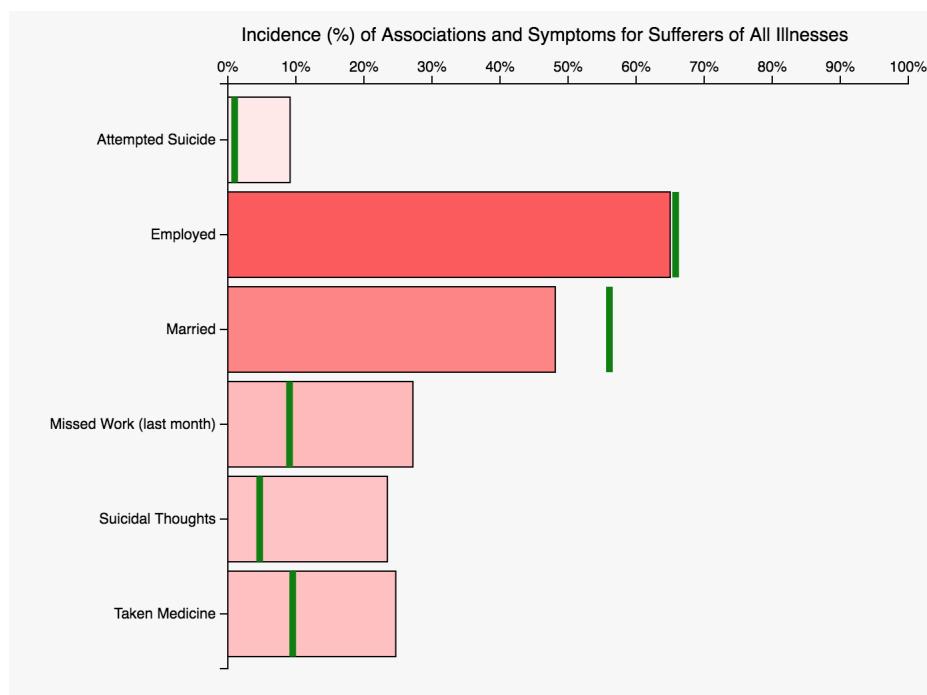


Figure 11: Association Visualization

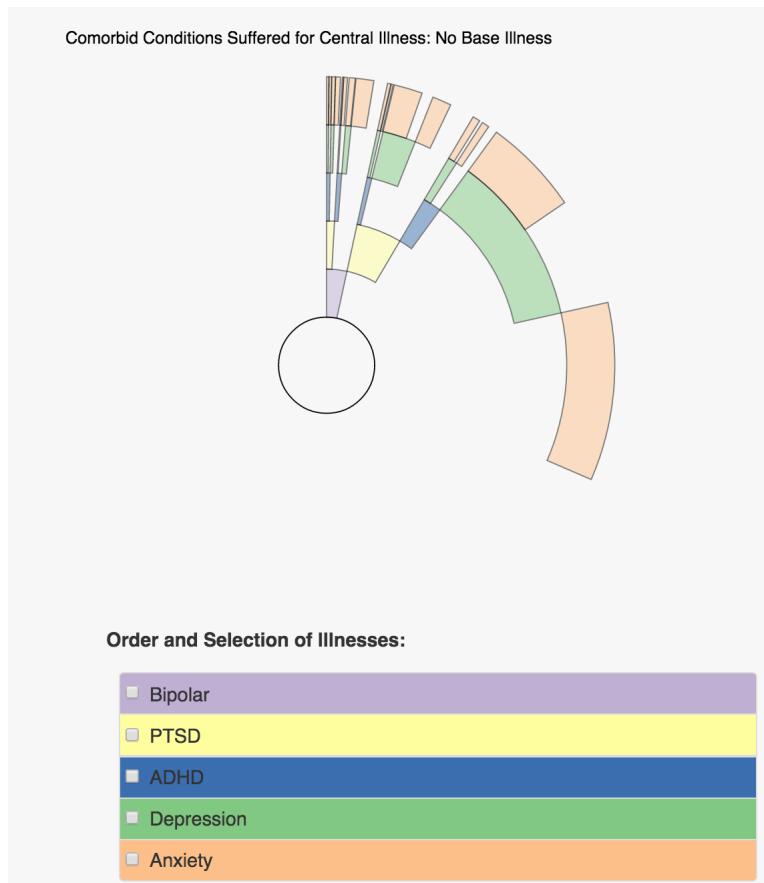


Figure 12: Sunburst Visualization, Revised

| | |
|---|--------|
| 3. Rubric: Please give scores are from 0-5 (0 means bad, 5 means good) | |
| Criteria | Points |
| Technical requirement 1 fulfilled (multiple coordinated views) | 5 |
| Technical requirement 2 fulfilled (one innovative view, either an extension of an existing visualization type or a novel visualization) | 4 |
| Effective visualizations | 3 |
| Innovative visualizations | 4 |
| Level of technical difficulty | 5 |
| Clear storytelling | 3 |
| Visual design (including website) | 3 |
| Addresses the goals | 4 |
| Sensible and effective interaction | 4 |

4. Please write down any other comments you might have:

Excellent. I really like the interactivity and simplicity.

| | |
|---|--------|
| 3. Rubric: Please give scores are from 0-5 (0 means bad, 5 means good) | |
| Criteria | Points |
| Technical requirement 1 fulfilled (multiple coordinated views) | 3 |
| Technical requirement 2 fulfilled (one innovative view, either an extension of an existing visualization type or a novel visualization) | 4 |
| Effective visualizations | 4 |
| Innovative visualizations | 4 |
| Level of technical difficulty | 4 |
| Clear storytelling | 3 |
| Visual design (including website) | 4 |
| Addresses the goals | 3 |
| Sensible and effective interaction | 4 |

4. Please write down any other comments you might have:

- Life effects of mental illness
 - How many people might experience this more?
 - Could there be an integration & explanation of genetics

| | |
|---|--------|
| 3. Rubric: Please give scores are from 0-5 (0 means bad, 5 means good) | |
| Criteria | Points |
| Technical requirement 1 fulfilled (multiple coordinated views) | 3 |
| Technical requirement 2 fulfilled (one innovative view, either an extension of an existing visualization type or a novel visualization) | 4 |
| Effective visualizations | 4 |
| Innovative visualizations | 4 |
| Level of technical difficulty | 4 |
| Clear storytelling | 3 |
| Visual design (including website) | 4 |
| Addresses the goals | 3 |
| Sensible and effective interaction | 4 |

4. Please write down any other comments you might have:

Focus on storytelling. Looks good!

Figure 13: Peer Evaluations 1, 2, and 3

3. Rubric:
Please give scores are from 0-5 (0 means bad, 5 means good)

| Criteria | Points |
|---|--------|
| Technical requirement 1 fulfilled (multiple coordinated views) | 4 |
| Technical requirement 2 fulfilled (one innovative view, either an extension of an existing visualization type or a novel visualization) | 4 |
| Effective visualizations | 5 |
| Innovative visualizations | 5 |
| Level of technical difficulty | 4 |
| Clear storytelling | 4 |
| Visual design (including website) | 3 |
| Addresses the goals | 5 |
| Sensible and effective interaction | 5 |

4. Please write down any other comments you might have:

- Face incoherent point numbering
- big introduction, 40% + have a consistent color scheme
- add more prints

2

3. Rubric:
Please give scores are from 0-5 (0 means bad, 5 means good)

| Criteria | Points |
|---|--------|
| Technical requirement 1 fulfilled (multiple coordinated views) | 5 |
| Technical requirement 2 fulfilled (one innovative view, either an extension of an existing visualization type or a novel visualization) | |
| Effective visualizations | |
| Innovative visualizations | |
| Level of technical difficulty | 4 |
| Clear storytelling | 5 |
| Visual design (including website) | |
| Addresses the goals | |
| Sensible and effective interaction | 5 |

4. Please write down any other comments you might have:

color palette could be a lot more diverse / content interaction brings randomness and makes the audience want to explore more

2

Figure 14: Peer Evaluations 4 and 5