

Change the World Project Proposal

Emily Wang & Brenna Manning
Data Science Spring 2016

Summary of the Project

Main Idea - Exploring depression and bipolar disorder, and investigating different axes of diversity seen in structural neuroimaging studies.

Axes of Diversity:

- Era
- Age
- Socioeconomic Status
- Gender
- Ethnicity
- Geographic Location - Cultural Impact
- Ability

We will be exploring the data and creating visualizations about what we find, with a focus on increasing awareness as well as delving deeper into potential opportunities for impact.

MVP

- Exploration and deeper understanding of dataset with different axes of diversity
- Visualizations of data across our axes of diversity in a polished jupyter notebook

Stretch Goals

Create visualizations for potential opportunities for impact.

- Outreach programs
- Group therapy
- Families
- Medications
- Spaces
- And more!

In what way will it change the world?

- Making the abstract stuff more tangible/visible to people, and extracting it into an artifact. Raising awareness of depression and bipolar disorder, where they exist and how they are often present.
- Exploring opportunities for impact. What has and has not been implemented, and what has had short term and/or long term positive impact?

What do we want to learn?

- Data cleaning/engineering

- Dataset explorations
- Representing our insights
- Communication skills
 - Infographics/Portfolio
 - Presentation (Perhaps video)
- Design skills
 - Graphic design
 - Creating clear concise visualizations

Where will we get the data for our project?

- Primary sources:
 - <http://archpsyc.jamanetwork.com/article.aspx?articleid=1107416>
 - <https://sites.google.com/site/depressiondatabase/>
- We may use more if we find more relevant sources.

Additional resources

- *Visual Display of Quantitative Information* - Edward Tufte

Final Deliverable:

We would like to create a bridge to bring this information into a larger context.

- jupyter notebooks: the prototyping ground for our explorations and data-driven graphics
- Posters or Portfolio of Visualizations: compelling graphics that tell a story to the population at large and future policymakers.
- Video Presentation: potentially TED talk style, 15-20 minutes of us narrating the story and opportunities in this topic.

Workflow and Schedule

Suggested Schedule (will reflect and edit accordingly throughout the project)

- **Fri Feb 26 (Class)**
 - Project Proposal due before 9am
- **Saturday Feb-27 Meeting**
 - Read documents
 - Working with the data
 - (Independently) Background Info from other sources
- **Monday Feb-29 Meeting**
 - jupyter Notebook prototypes of data visualization
- **Tues March 1 (Class)**
 - Mystery
 - Sanity check on progress: Edit the plan and rescope if needed
- **Wed March-2 Meeting**
 - Dependent on what we learn the first few meetings
 - Should be making good progress towards mid-project check in (see below)

- **Thursday-Saturday March 3-5 Independent Work**
 - Should be making good progress towards mid-project check in (see below)
- **Fri March 4 (Class) Mid Project Check-In**
 - `mid_project_checkin.md` in GitHub repo
 - Scope out the following week (portfolio, video presentation, both/either, etc.)
- **Tues March 8 (Class)**
 - TBD
- **Fri March 11 (Class) Final Output and Reflection Due**

Support for this schedule and collaboration style:

- Includes both independent work time (for reading up on external sources) and team meetings for exploration, discussion, and prototyping deliverables (first in jupyter notebooks, then in a more formal portfolio and/or presentation)
- Significant progress made before every Data Science class session to prevent all of it happening at the last minute.

Mid-project Check-in (Fri March 4)

- Some visualizations inspired by our axes of diversity
- Some qualitative resources/papers
- Read paper associated with dataset
- Jupyter Notebooks as prototyping ground

Assessment

Note: Currently these seem more like feedback/discussion points rather than explicit properties to be assessed. We hope to have more concrete assessment points soon; perhaps we can discuss this list with Paul more in person. :-)

Jupyter Notebooks

- Feedback on software engineering practices (helper functions, readability, etc.)
- Narration of code, process, learning about the techniques, and learning about the data

Visualization Portfolio

- Should be polished
- Narrative feedback on whether the visualizations are misleading, easy to understand, etc. -- design decisions for communicating via graphics

Video Presentation

- Comments on presentation style
- Was the audience confused? Did the audience understand our message?