Deirdre Offenheiser­

DSGN 352 Design with Data

Final Project Proposal

(Due) 11/3/2015

* **What is the purpose of the visualization? Is there a question driving the creation of the visualization? What do you intend for users to learn from your visualization?**

While sifting through sites of collected datasets I discovered a csv file containing a list of all the passengers onboard the Titanic at the time it collided with an iceberg in the North Atlantic Ocean. Though I’m sure there’s little new information to glean from the data collected in the csv file, it demonstrates how a diverse group of people was impacted by a singular event. What’s more the measure of impact from this event is clear-cut rather than descriptive. A passenger either survived or she didn’t. The absence of ambiguity in this respect allows one to consider other passenger characteristics as well as how they intersected with survival rates. For instance I would like to show a series of visualizations that demonstrate that women survived in greater numbers then men, and that the those in the lower classes were much less likely to survive. These assertions are born out by a series of rough visualizations I created in excel to compare these characteristics. I would also like to show the composition of the various passenger boats. I believe this visualization demonstrates how likely different passenger groups were to mix, as well as how passengers were prioritized. Other stories I’m interested in exploring include the age composition of the passenger list, potentially broken up by class, as well as how age intersected with survival. Additionally, I would like to map the homes and final destinations of the passengers to a map to demonstrate their diverse makeup. My goal in creating this visualization is to interrogate the data to give a richer picture of the characteristics of those aboard the Titanic as well as to demonstrate how certain passenger characteristics intersected with survival rates.

* **From where will you obtain a dataset? How much work will it take to acquire, sift, and format your dataset?**

Sources:

Contains data sets for statistics projects including the Titanic passenger list data set:

<http://biostat.mc.vanderbilt.edu/wiki/Main/DataSets?CGISESSID=10713f6d891653ddcbb7ddbdd9cffb79>

Basic description of dataset and methodology:

<http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic.html>

Includes more information about voyage and passengers:

<http://www.encyclopedia-titanica.org/>

* **What is the scope of functionality or features that you wish to implement?**

I would like to create a site that is composed of an introductory or home page with links to three separate subpages that explore the data in distinct ways. I intend to use a set of icons on my homepage to allow visitors to navigate to these distinct icons. I’m also interested in the idea of having my content and navigational tools fairly well integrated into the page. Another idea I’m interested in exploring besides those I described above, is to select a set of passengers (maybe two or three from each class) to be featured in a detail view if a visitor were to hover over a data point (perhaps on the map of passenger homes).

* **Are there similar extant visualizations that inspire your proposed visualized or might serve as a starting point?**

I have found some extant visualizations of this topic:

<https://www.jasondavies.com/parallel-sets/>

<https://www.behance.net/gallery/26632637/Surviving-Passengers-of-Titanic-data-visualization>

<http://radar.oreilly.com/2012/04/titanic-visualization.html>

<https://www.kaggle.com/c/titanic>

* **What challenges are you anticipating? Can you think of some ways to address these challenges at the outset?**

The data set consists of about 1,300 data points (each representing an individual passenger). While this data set is certainly large there are fairly simple ways to categorize the data to make it more readable and address a particular question. I will also need to separate the ‘home.destination’ column into separate characteristics and edit for errors. The idea of mapping passengers to particular locations will also require me to decide whether to map passengers to their home cities or home countries as well as figure what information I need to create a map visualization.