Daniel Connolly and Jillian MacGregor
Software Design Spring 2018

Mini Project 4 - Project Proposal

We intend to create an healthcare related interactive data visualization for our project. The data visual will allow users to mouse over a map of the United States and interact with each state. We hope to display the number and locations of all of the hospitals in each state, and will then proceed to either calculate and display the average distance to the nearest hospital state by state, focusing on rural areas, or perhaps display only data regarding medical costs in more urban areas, where hospitals are known to be nearer on average. Our minimum viable product can be considered the latter of these two options, as it would consist of an interactive state-by-state map of the US that displays information regarding medical costs and resulting information when the user scrolls over the state. At this stage in the process, calculating and displaying the average distance to the nearest hospital using census data and the Google Maps api can be considered a stretch goal, albeit a stretch goal we believe we can achieve given the right dataset.

Our learning goals are as follows:

Daniel Connolly → I mainly hope to learn about creating programs that interact with users and about data visualization as a whole. Additionally, I want to learn about programming with a partner on a project and want to continue improving my stylistic programming skills.

Jillian MacGregor → For this particular project, my learning goals revolve around the concepts of pair programming/development, as well as data capturing and visualization. I hope to gain a better understanding of what it is like to develop code working through GitHub (something touched upon in ModSim), and become familiar with using different libraries and methods of finding data to complete this project. I hope that I will also be able to build my skills by working with a partner to generate good quality code that I can refer back to later on.

We plan to primarily utilize pygame in order to create our interactive map of the country. Although we have found and considered other mapping libraries that would allow us to investigate the possibility of a county-by-county map of the US, we have decided to avoid these libraries until we have at least created our minimum viable product. Additionally, we are utilizing datasets found on the data.gov/health website to

amass hospital charge data and are in the process of locating the best dataset to use regarding the locations of all US hospitals. With regards to our stretch goal, we have considered looking through the US census bureau's website in order to find the locations of households throughout the US.

For the mid-project check-in, we hope to have created a minimum viable product-style interactive map of the United States using pygame. We also hope to have collected a non-trivial amount of the data we will need in order to complete our project. We may potentially have begun to integrate the data with the interactive map or have constructed a more interesting version of our map, although we realize this may not be possible given the time constraints. In the following week, we will continue to work on the interactive features of our map, collect any data we still require, and more heavily integrate that data with our map. The biggest goal for the mid-project check-in is to ensure that what we have decided to do lies within a proper scope of a 2-week project for our programming level.

The biggest risk we face with regards to this project is accurately deciding on the scope of our project. After some initial over-scoping, we hope we have found a balance in terms of deciding what constitutes our minimum viable product. Moreover, we are still working on finding the data necessary for our data visualization, which is proving more difficult than expected. Much of the risk surrounding this project lies in the ability to find and properly parse through meaningful data. Nevertheless, we realize that the amount of time we have to work on the project will quickly dwindle and believe this lack of time could be a major risk for our project, and are therefore willing to change our scope and proposal to fit the data that we can find.