INFO/CS 3300 Project 2 Report

Stephanie Zhou (sz244)
Soyeon Park (sp589)
Christina Wang (cw564)

A. A description of the data. Report where you got the data. Describe the variables. If you had to reformat the data or filter it in any way, provide enough details that someone could repeat your results. If you combined multiple datasets, specify how you integrated them. Mention any additional data that you used, such as shape files for maps. Editing is important! You are not required to use every part of the dataset. Selectively choosing a subset can improve usability. Describe any criteria you used for data selection. (10 pts)

We got our data from the Centers for Disease Control and Prevention. We looked at the National Vital Statistics Reports for the years 1998, 2000, 2006, 2010, 2013 to provide some range in the time we analyzed. We looked at the mortality data by race, gender, and state. Using svg elements, showed how the groups were affected by the top ten annual causes of death in the United States. For the map, we used a json file that delineates each state, and drew paths. We had many different causes and after parsing the data, selectively showed them in the visualization. The goal of the map was to show mortalities in the fifty states (continental United States, Alaska, and Hawaii), therefore territories like Puerto Rico, the Virgin Islands, Guam, American Samoa, and the Northern Mariana Islands were excluded in our visualization. There was data for statistics among different age groups, which we also excluded because we decided that we wanted to focus on the differences between gender and race in the pie chart form seen on our visualization.

2013 Report (http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_02.pdf)
2010 report (http://www.cdc.gov/nchs/data/nvsr/nvsr61/nvsr61_04.pdf)
2006 Report (http://www.cdc.gov/nchs/data/nvsr/nvsr57/nvsr57_14.pdf)
2000 Report (http://www.cdc.gov/nchs/data/nvsr/nvsr50/nvsr50_15.pdf)
1998 Report (http://www.cdc.gov/nchs/data/nvsr/nvsr48/nvs48_11.pdf)

B. A description of the mapping from data to visual elements. Describe the scales you used, such as position, color, or shape. Mention any transformations you performed, such as log scales. (10 pts)

We visualized mortalities in bubbles. The biggest bubble at the center without filter and view mode is the total number of mortalities recorded by the CDC in the United States in the selected year. Once clicked by the user, the main bubble shrinks to represent the mortalities claimed by causes of death not among the top ten (heart disease, cancer, …).But the label still displays the number of total death after shrinking because our test user found this to be more intuitive. The bubbles that appear around the center bubble represent the mortalities recorded for each of the top ten causes of death.

The size of these bubbles represents the proportion of mortalities caused by the name on the bubble. Once any of these bubbles are clicked, even smaller bubbles are created that branch out from their parent node, indicating more detailed categories of these causes of death when applicable. For instance, when the Cancer bubble is clicked, smaller bubbles representing leading types of cancer branch out from the parent bubble, indicating mortalities recorded for breast and lung cancer, leukemia, and all other cancers, in one bubble. Note that the proportion of smaller bubble in sub category with respect to larger bubble in the main category does not represent the ratio of mortality for display purpose,since some bubble would be too small for

Stephanie Zhou (sz244)
Soyeon Park (sp589)
Christina Wang (cw564)

interaction if the exact ratio is kept. but the ratio is kept for bubbles within the same sub-category.

When a view mode is selected, we use different colors to represent different groups. We had filters for gender and race as well to only display the statistics of the specified groups, as opposed to the overall mortalities recorded for both genders and all races. For gender, male mortalities are represented by a green ring, whereas female are represented by blue. For race, White Americans are colored green, Black Americans are colored blue, and all other minority Americans together are yellow. After the user click the first big bubble and show the categories, we use thicker border width and different color for the category bubble's border to indicate if the bubble is expandable to further show sub-categories or not.

When user enter gender/ethnicity view mode, pie chart is shown. For smaller ratio, we make the radius of the pie larger so that it would be easier for user to hover their mouse on that particular section of the pie chart. The filter will be disabled when user enter gender/ethnicity view mode, enabled when user exit to default mode.

We used different opacities on the map to represent an age-adjusted rate of each cause of death. We positioned the map such that when the user looks at the page full screen (width at maximum), the bubble visualization and map are visible at the same time. In this way, the user can see rate changes for each state as they select different causes of death.

C. The story. What does your visualization tell us? What was surprising about it? (5 pts)

Some surprising things we observed through our visualization are a prominence in death by heart disease in the southeastern states than other regions of the continental United States, Alaska, and Hawaii. Death by nephritis and other nephrotic kidney ailments were also relatively widespread in the Southeast. Suicide rate is troublingly high in Alaska compared to other states. However, cancer and stroke seem to be relatively widespread, not favoring specific regions over others.

From the bubble portion of the visualization, we can clearly see that heart disease and cancer are the top two leading causes of death every year. It was interesting to see that while normally the deaths caused by influenza alone (not including pneumonia) were only around 500, but that the sudden spike in influenza deaths in 2013 were thanks to the outbreak of swine flu, which many of us can remember clearly.