Matthew Gabor, Hide Shidara, Jason Blevins

ECS 163, Final Project

Traffic Fatalities in the United States

For our final project, we will be using a combination of closely related datasets that are centered around traffic safety. Specifically, fatal crashes over time and the circumstances surrounding their occurrence broken down to the state level. The sources of the datasets are the government NHTSA (National Highway Traffic Safety Administration)and IIHS (Insurance Institute for Highway Safety).

We plan to create a single page with three connected visualizations that the user can switch between. The user first sees a map of the United States that shows an overview, and can then select two states or regions for comparison. Once two states / regions are selected, they can choose one of the three visualizations to view. The three visualizations will show: (1) fatalities over time with significant events such as laws or weather events, (2) an animation of the fatalities and corresponding factors, and (3) a customizable scatterplot with regression line that plots fatalities with a multitude of user selectable factors. The third visualization will support filtering of data along with axis selection.

By examining these visualizations, we can ask questions such as what factors correlate to an increase or decrease in traffic fatalities. This is facilitated by the 50 different states having differing laws, timelines, and events. By selecting states that are either similar (either both low or both high fatalities), or different levels of fatalities, the user can explore differences and identify trends that are not obvious when viewing the raw data.

Three specific tasks that our visualization can perform are listed as follows: The user can select two states or regions and based on that selection can do one of the three following things: (1) Compare the number of fatalities between the states while showing the correlating factors such as alcohol, distraction, or speeding. (2) View a timeline of traffic fatalities in each state with significant events pointed out and a compare button that overlays the two timelines for easy comparison. (3) Select different variables to display against total fatalities and view a regression line. For example: the user can select “Number of DUIs” and plot that against traffic fatalities while comparing two states.

We will have 3 main vis and an abstracting, “cover page” vis. The first 2 visualizations will directly answer the questions by allowing the user to see correlations that we find within the data. For example, if the question is about correlations between 2 states and the driving fatalities in California, the user will be able to see the amount of fatalities relative to another state in the first vis and the causation of different fatality trends in the second. The third vis will be exploratory, allowing the user freedom to explore new trends of the data and discovering correlations that even we can’t have come up with.

Hue will be used to show the number of fatalities for each state, the darker, the more deaths. Vis 3 will utilize position of each dot to show the relationship between the attributes selected. Movement in Vis 1 will show what the fatalities were caused by. Further we will color the cars to show the categorical attributes that a state’s fatality is caused by. Interaction will play into our visual design by the user being able to interact with the US Map and choosing the states that they want to compare. The visualization parameters that will be under user control are: selection of state's, selection of causes of fatalities, and attributes for Vis 3.

The US map allows users to easily navigate through the dataset by giving them a way to easily select and compare data across states. The scatterplot regression also provides a way to navigate the data through filtering options. For example, if a user wanted to see how New York compared to California in overall fatalities, they would just select both states on the map. If they wanted to drill down further and look at how certain factors, such as laws affected these rates they could do so in the scatterplot/regression area.

The interaction between the US Map and all three visualizations is important because it allows the user to filter through the data by state pairs rather than trying to view it all at once. The interaction between the factors animation and the time chart are important because the date selected in the time chart dictates which year we show data for.

Responsibilities and Tasks:

Matt:

* Layout to navigate through each viz
* Animation of fatalities with factors

Jason:

* Selectable US Map/Intro
* Fatalities over time chart

Hide:

* Regression/Heatmap/Scatterplot chart
* Data aggregation and filtering