CS573 Data Visualization
Final Project Process Book

By Qiuzhe Ma & Qianyun Yang

A. Project Motivation and Objectives

Background and Motivation

Our final project is about fire emergency happened in Seattle from 2012 to 2015. The existence of fire changes our life and behavior. We need fire to cook, to keep warm and fire is also necessary in manufacture. But if people are not cautious about fire, it may lead to disasters. Each year people die or are seriously injured as a result of fires at work. Besides loss of life, fire costs people millions of money from damage of property, loss of business, fire compensation and insurance premiums. Many of these fires can be avoided by taking fire precautions. If a fire breaks out a lot of the effects can be minimized by having effective controls and procedures in place. This project therefore intends to make an assessment of the factors influencing fire disaster preparedness. We consider that our analysis may provide valuable information to improve allocation of fire resources, and remind people about the fire prevention in effective ways.

Project Objectives

The world has in the past decades experienced succession of fire disasters. These disasters have claimed many thousands of life, caused material looses and afflicted terrible toll. Thus, people need to be well equipped in terms of knowledge on how to prevent and react to fire outbreaks. We try to analyze the fire emergency in Seattle and get some conclusions of which reason may cause fire emergency most, which time period may fire happens frequently, what locations may fire happens the most. We intend to use visualizations such as bar chart, pie chart, area chart and Seattle map together to design coordinated multiple views. Our design will make the fire data intuitionistic interface to users and is interesting to understand and learn.

B. Data Processing and Prototype Design (Week 1)

Data

The first data we collected is named Seattle Real Time Fire 911 calls. We find a website that describe Seattle real time fire 911 calls from 2010 to now. The data will update every minute and we choose the dataset from 1/1/2010 to 11/17/2015. The dataset includes information such as address, type, date time, latitude, longitude, report location and incident number. The second data we find is Seattle population because we also want to find the relationship between density of fire incidents and density of population of Seattle. The population density in Seattle includes population density and location information in 2015. The third data we find is the map of Seattle in json file. Here are the links to the website where we find our data:

Fire Calls Link: https://data.seattle.gov/Public-Safety/Seattle-Real-Time-Fire-911-Calls/kzjm-xkqj
Seattle Map Link: https://catalog.data.gov/dataset/seattle-json

SeattlePopulationLink:http://www.arcgis.com/home/webmap/viewer.html?webmap=cdf24bce6363445e83670a6ea42f5dbe

http://zipatlas.com/us/wa/seattle/zip-code-comparison/population-density.htm

Data Processing

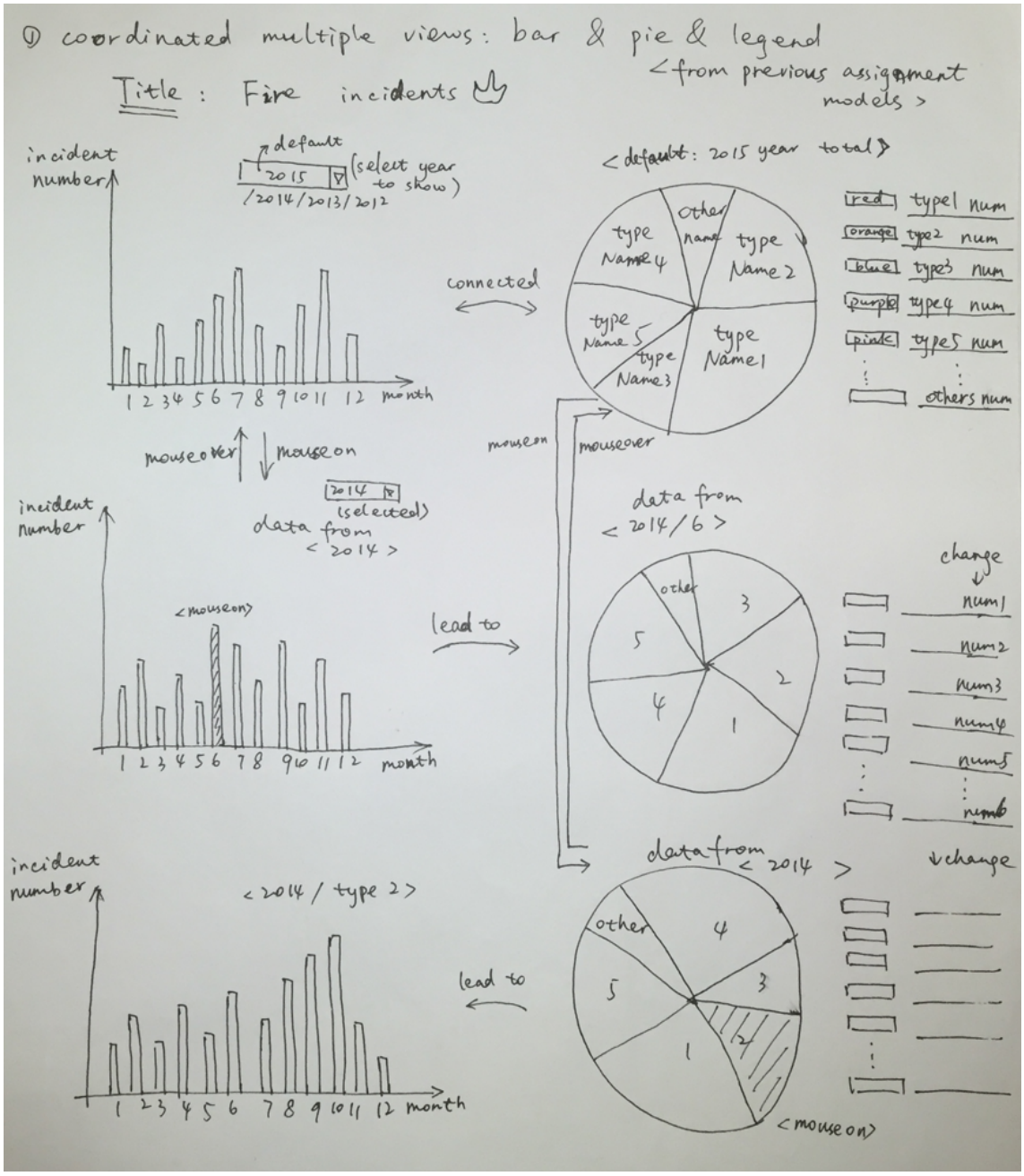
- We need time, type and location (longitude, latitude) information, but we do not need street location in our design. That is because longitude and latitude can be called directly in coding, and street location tells the same description in words.
- We also do not need the column named incident number. This is like the order number and we try to replace it by ordered integers.
- We remove the data from 1/1/2010 to 12/31/2011, which we figure out that the

quality of the data is low because there is always a data missing, and the quantity is not large compared to other years' data.

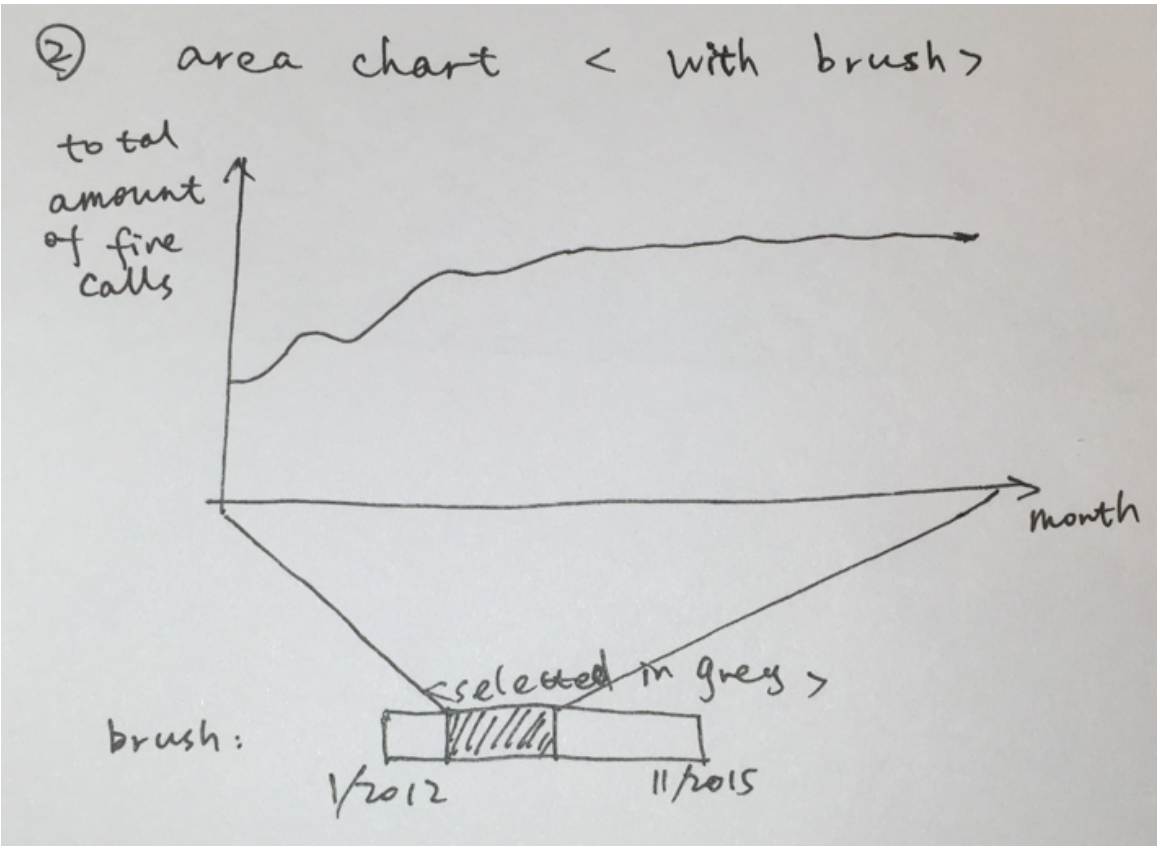
- After data clean up, we have more than 300 thousands of columns of data to process, which is an enormous dataset and is enough for our analysis.
- We want to design a map that shows the relationship between population density and fire incidents density.
- We decide to make the data in json file to be called in our code.

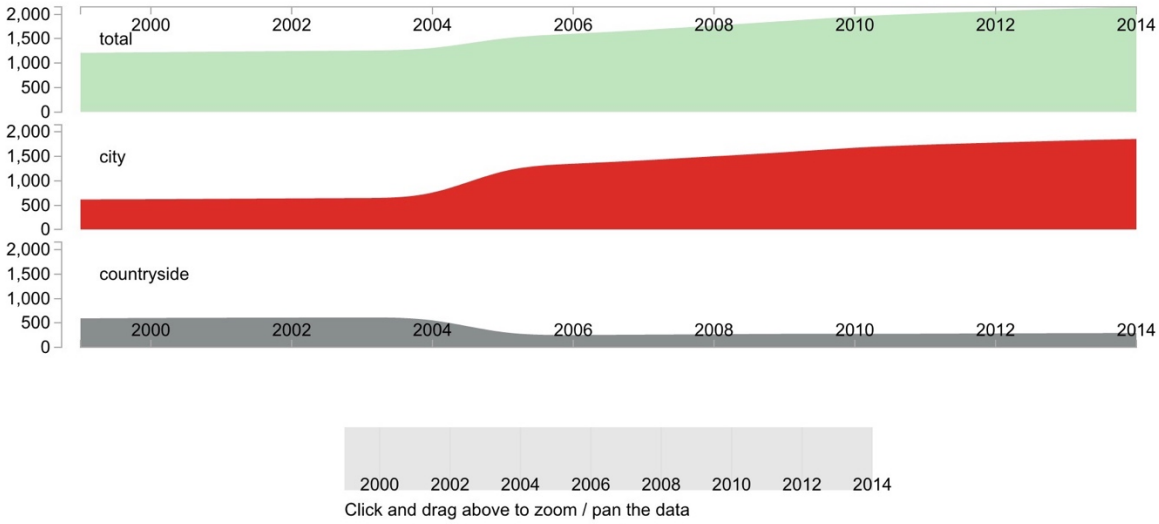
Visualization Design

- Design 1: Coordinated multiple views: bar chart, pie chart and legend.



◆ Design 2: Area chart with brush shows time varying specific type.

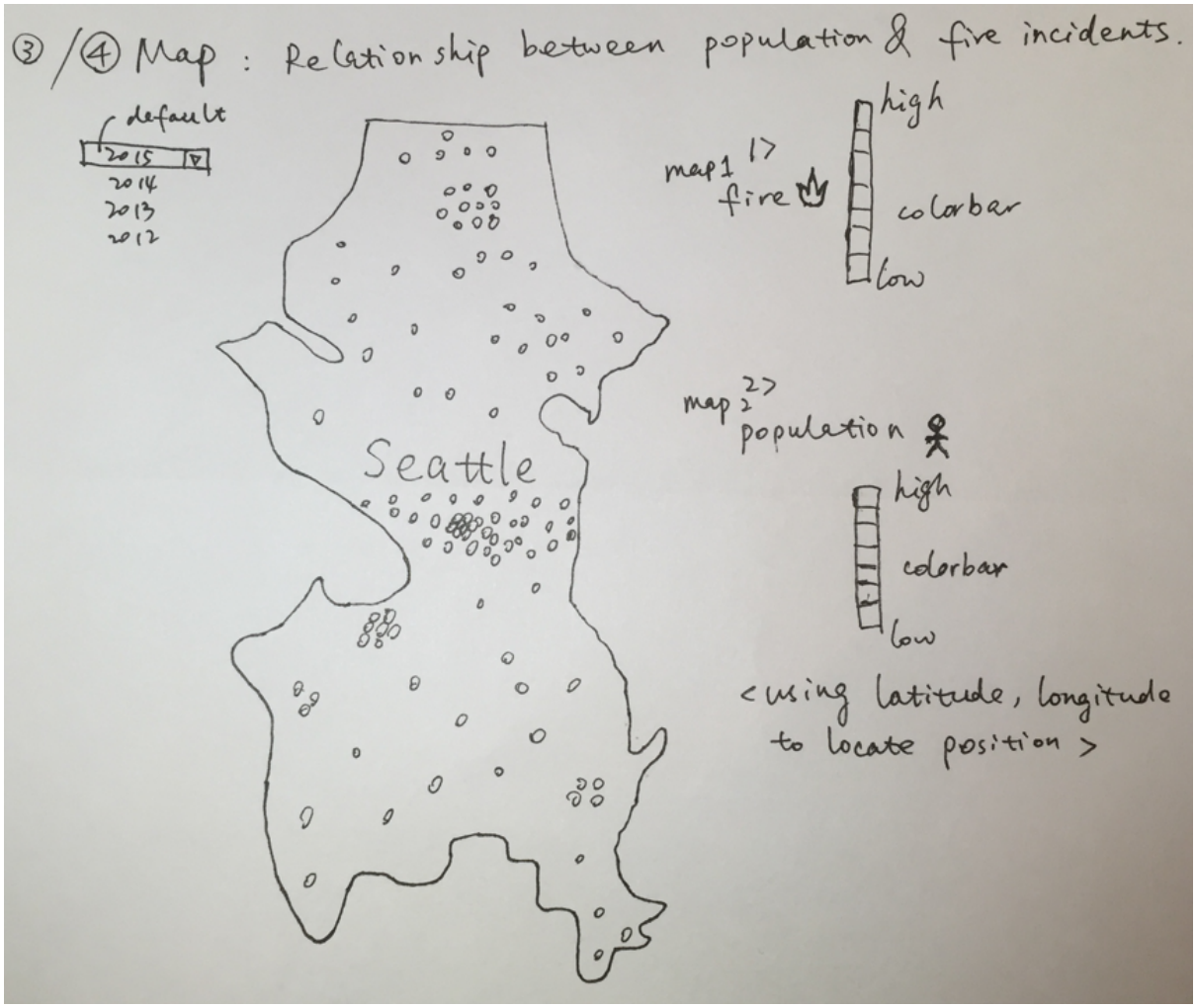




◆ Design 3: Heat Maps

1) Population density

2) Fire incident density



C. Implementing Three Parts of Visualizations (Week 2)

1. Coordinated multiple views: bar chart, pie chart and legend.

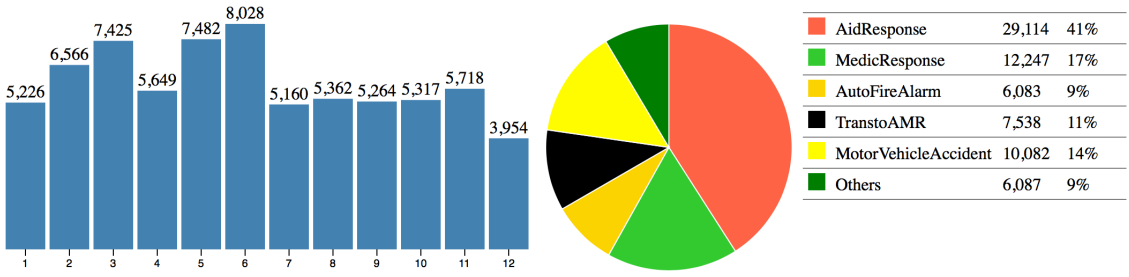
Related work:

In this part, we want to show the coordinated multiple views about the fire alarm. This part is related to the assignment that I have done before, so that it is not a very hard part for us to deal with.

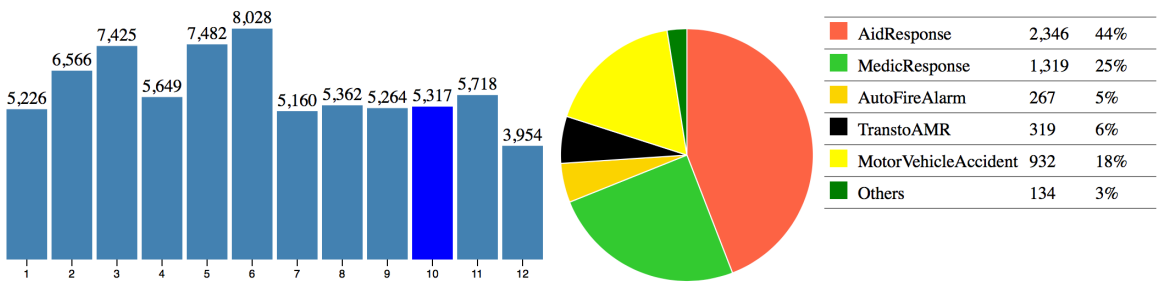
In the bar diagram, it is supposed to show the total 911 fire alarm calls for one selected year (2015/2014/2013/2012), and the default value of year is 2015. It shows the amount of 911 calls each month.

In the pie chart, it shows the proportion of each type and also the detailed number of each part. Because there are about 25 types of 911 calls, it is really hard to show them all because some of them contains very small amount. To compact the dataset, we deal with the dataset using python to count the numbers of each type. As a result, we find out that there are 5 types occupying most part of the dataset. Therefore, we decide to illustrate the 20 types that contains small values in one part named others. This can be seen in the pie chart below.

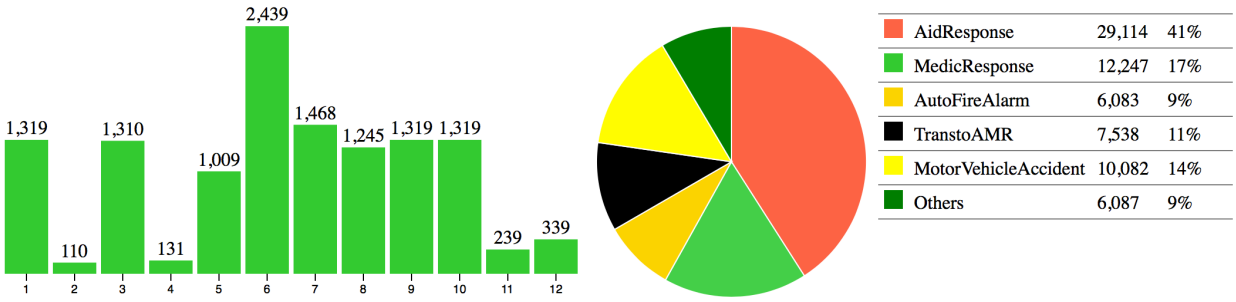
The bar chart shows the total amount of 911 alarm calls in year 2015, and the pie chart shows the proportion of each type in year 2015.



When hovering on the bar chart, the pie chart and the legend on the right side will change related. And the color of the selected bar will be colored blue as shown below.



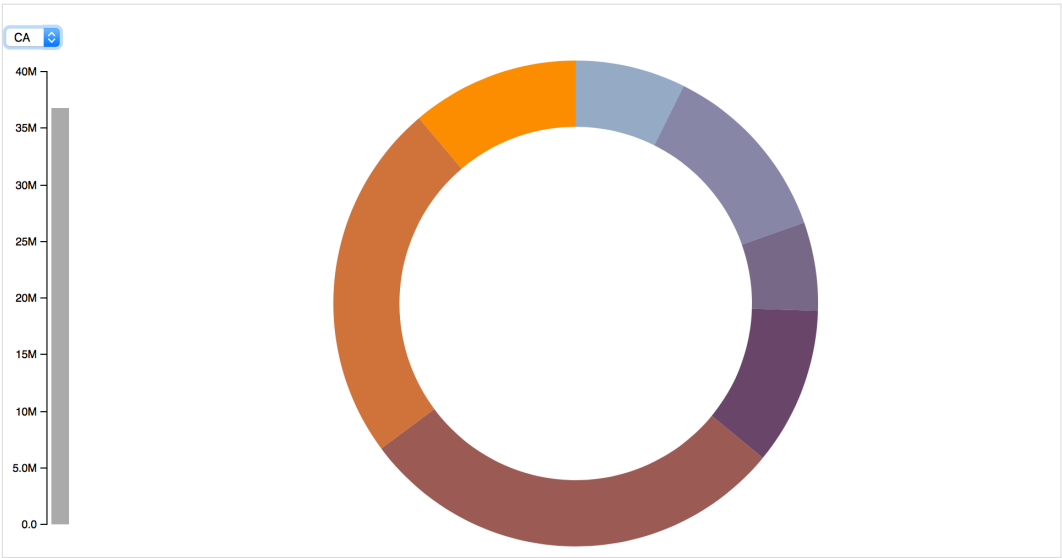
When hovering on the pie chart, the bar chart will change related. And the color of the selected pie will be colored darker as shown below.



Work needs to be done later:

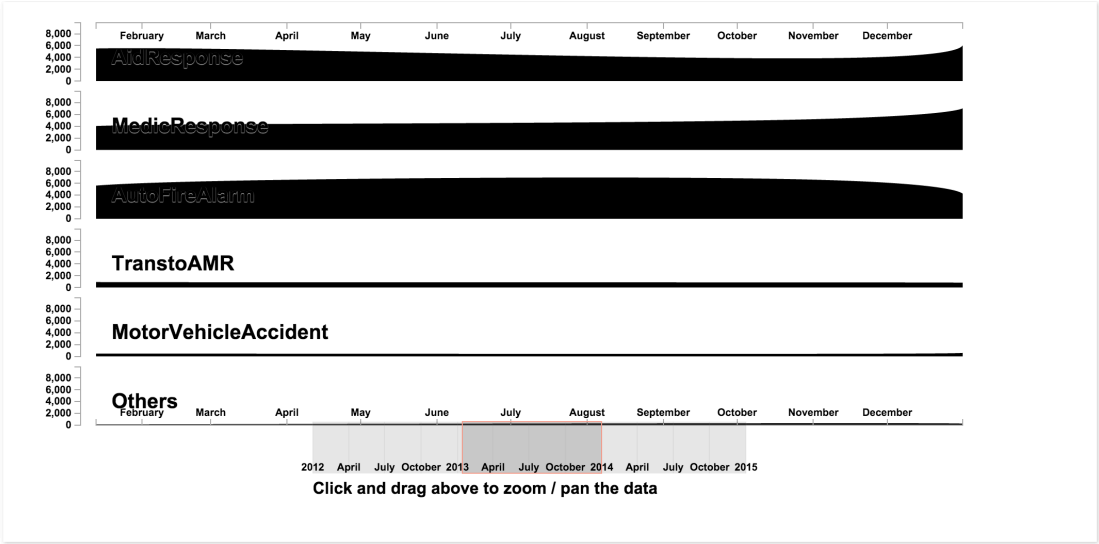
We only design the bar chart, the pie chart and the legend. We do not have the selection button offered to the users to choose the year they want to see. We have already get the dataset of year 2015, 2014, 2013, 2012 separately. We have already find an example that contains the selection button.

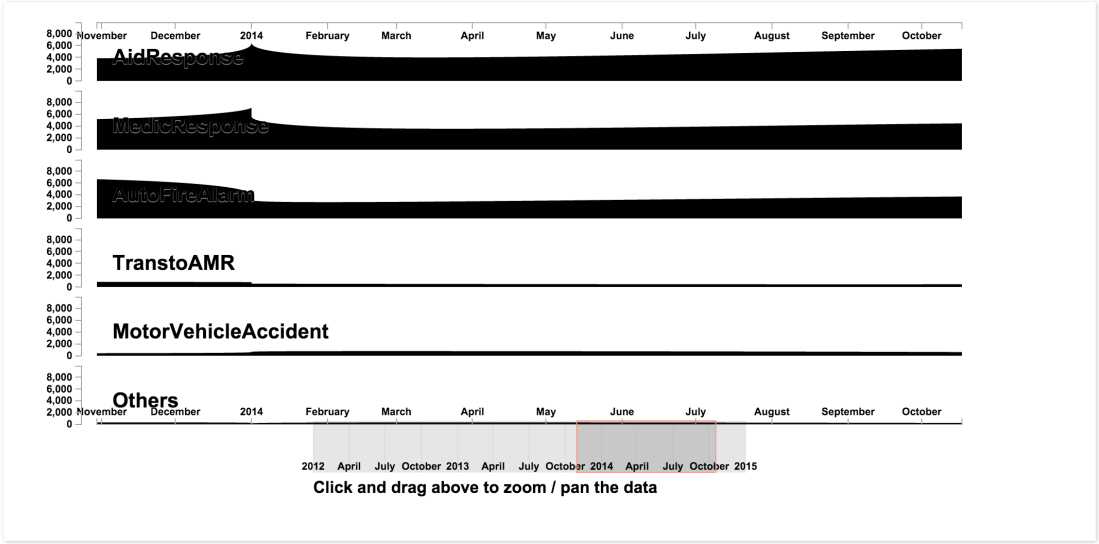
Reference: http://bl.ocks.org/mbostock/5872848



2. Area chart with brush shows time varying specific type

This area chart is related to the dataset of everyday 911 fire alarm calls from 2012 to 2015. There is a brush that shown below to give the users the scale to see.





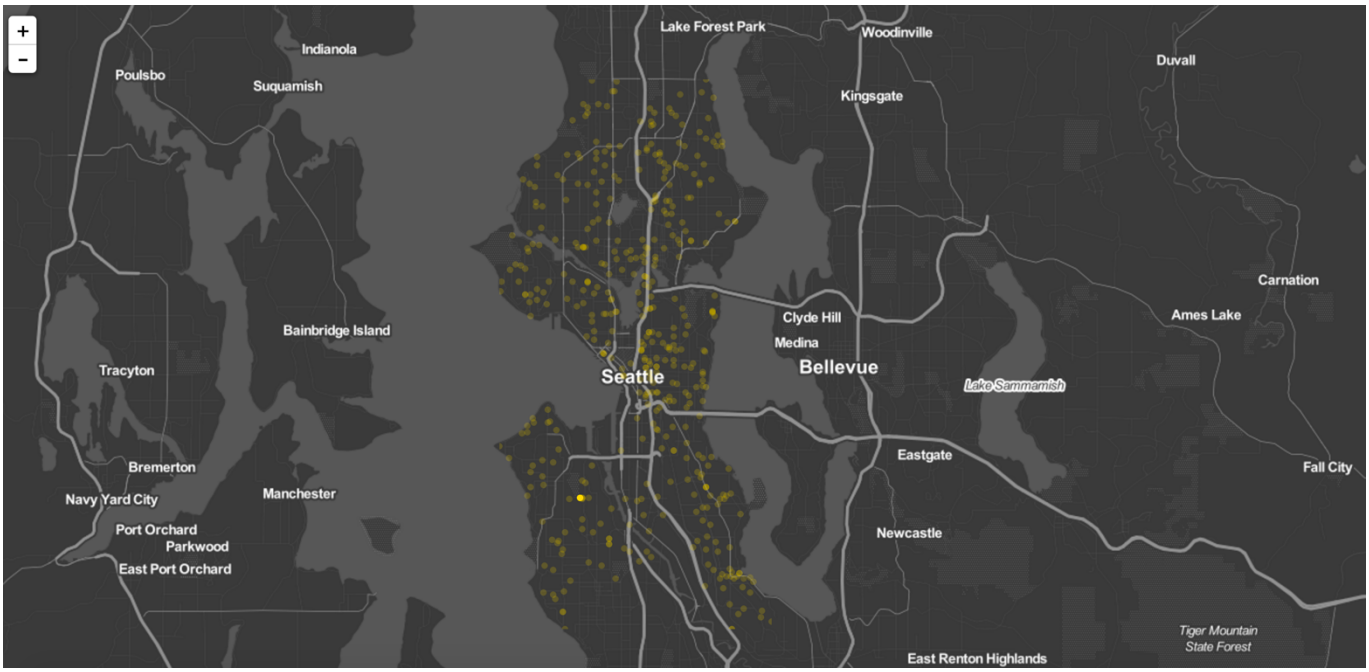
Work needs to be done later:

These figures above show the results of area figure, similar to the first design, we also need a selection button to show the year 2015/2014/2013/2012. Besides, we need to change the color of the figure and color of the legend to make it more colorful and readable. We also need to change the scale of the y-axis. Because some of them are too small to see, as the dataset contains a small number compared to others.

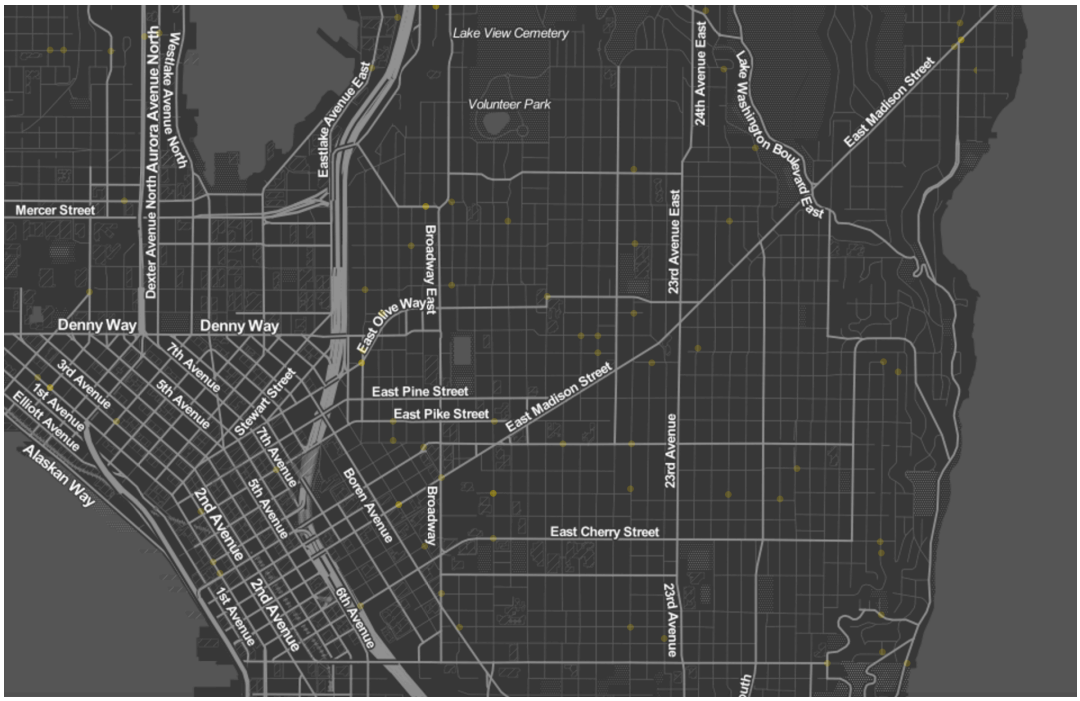
3. *Points Maps*

1) Fire incident density

The figure below uses the dataset of one kind of Seattle 911 calls, named WireDown.csv. This figure is just to show the blueprint of the different kinds of types in points map.

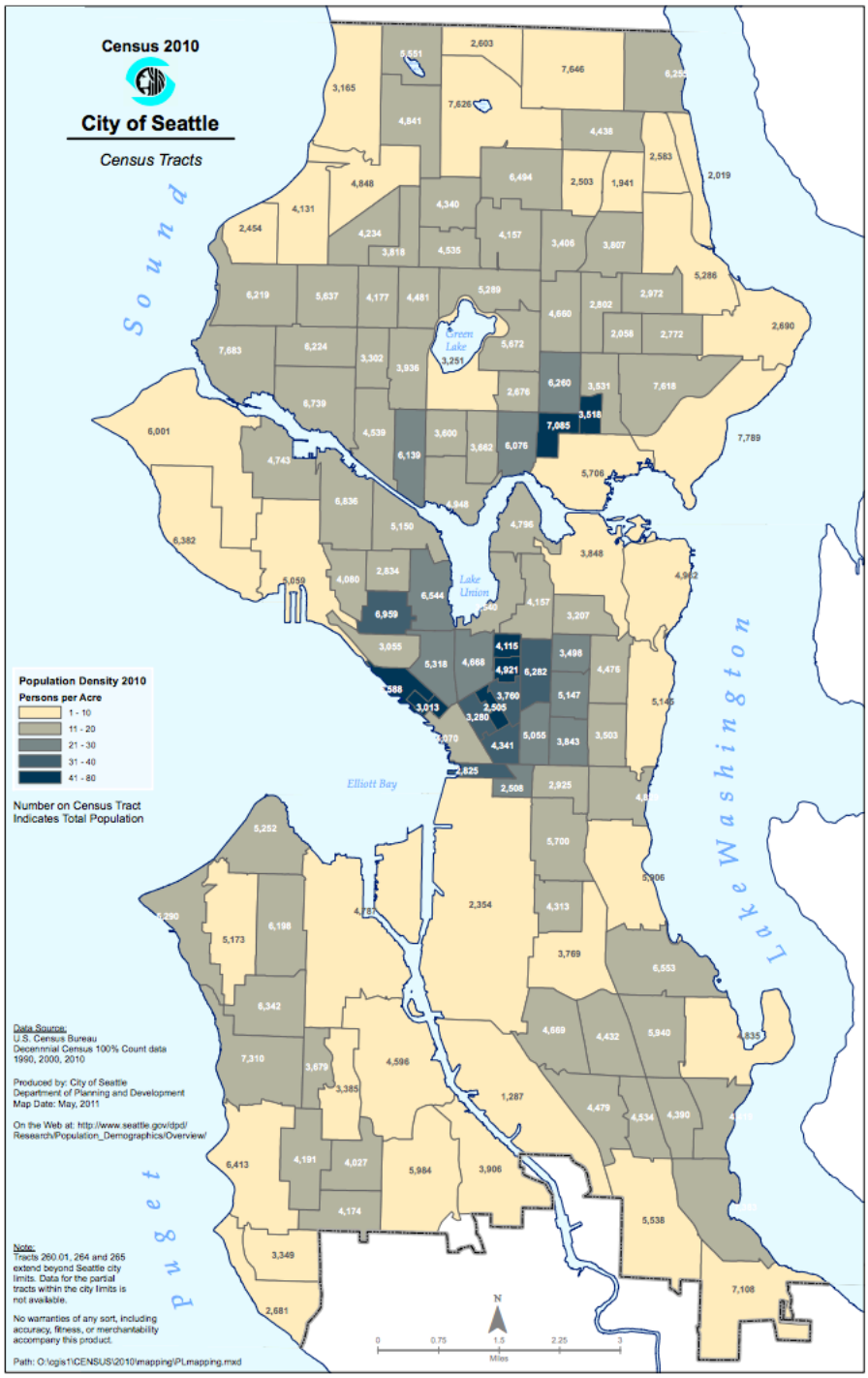


The figure below shows the zoom in outcomes of the figure above. We can see that it indicates the exact location point on the map using yellow points because we use the latitude and longitude data from the dataset.



2) Population density

The figure above shows the population density of Seattle in year 2010. We find out that census will be calculated every 5 years. It is not every year. But it still can show and explain the relationship between fire alarm frequency and population density.



Work needs to be done later:

For the first map, the fire incident density map, we need to mark every type in the

map. We already have the dataset, what we need to do next is to aggregate the dataset and show them together on the map. And it is a little inconvenient to zoom the map from US to Seattle, we might want to find a way to make it show the map of Seattle only.

For the second map, the population map, we already have the dataset of population density of Seattle in 2015. But the dataset contains only the information of zip code and exact location. What we need to do is to find the relationship between zip code and the district information. It will finally show the result map as the map of Seattle in 2010 above.