CS573 Data Visualization
Final Project Process Book

By Qiuzhe Ma & Qianyun Yang

A. Overview and Motivation

1. 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.

2. 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 Collection and Data Processing

1. Data Collection

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 12/13/2015. 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:

Seattle 911 Fire Alarm 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

Seattle Population Link:
http://www.arcgis.com/home/webmap/viewer.html?webmap=cdf24bce6363445e83670a6ea42f5dbe

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

2. Data Processing

1) Data cleaning

- 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 data is unreliable because there is always a data missing, and the quantity of the data 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.

2) Data calculating and aggregating

- To make it easier for us to design, we decide to select and calculate data by using EXCEL functions instead of calculating in d3 coding. Therefore, we have one dataset for each visualization.
- For Bar & Pie chart, we decide to make the data in json file to be called in our code, and the data includes seven types of 911 calls and relative numbers, and each year includes one json file.
- For area chart, we decide to use types and numbers, which is ordered by month and year.
- For maps, we want to design a map that shows the fire incidents density. And the selection button can change the type to show on the map. Therefore, we need to aggregate by types and save them in different type json files.
- For season map, we decide to design the dataset that separate by four seasons, and show the results in form and lines. The four seasons are Months 3.4.5 for spring, 6,7,8 for summer, 9,10,11 for autumn, 12,1,2 for winter.

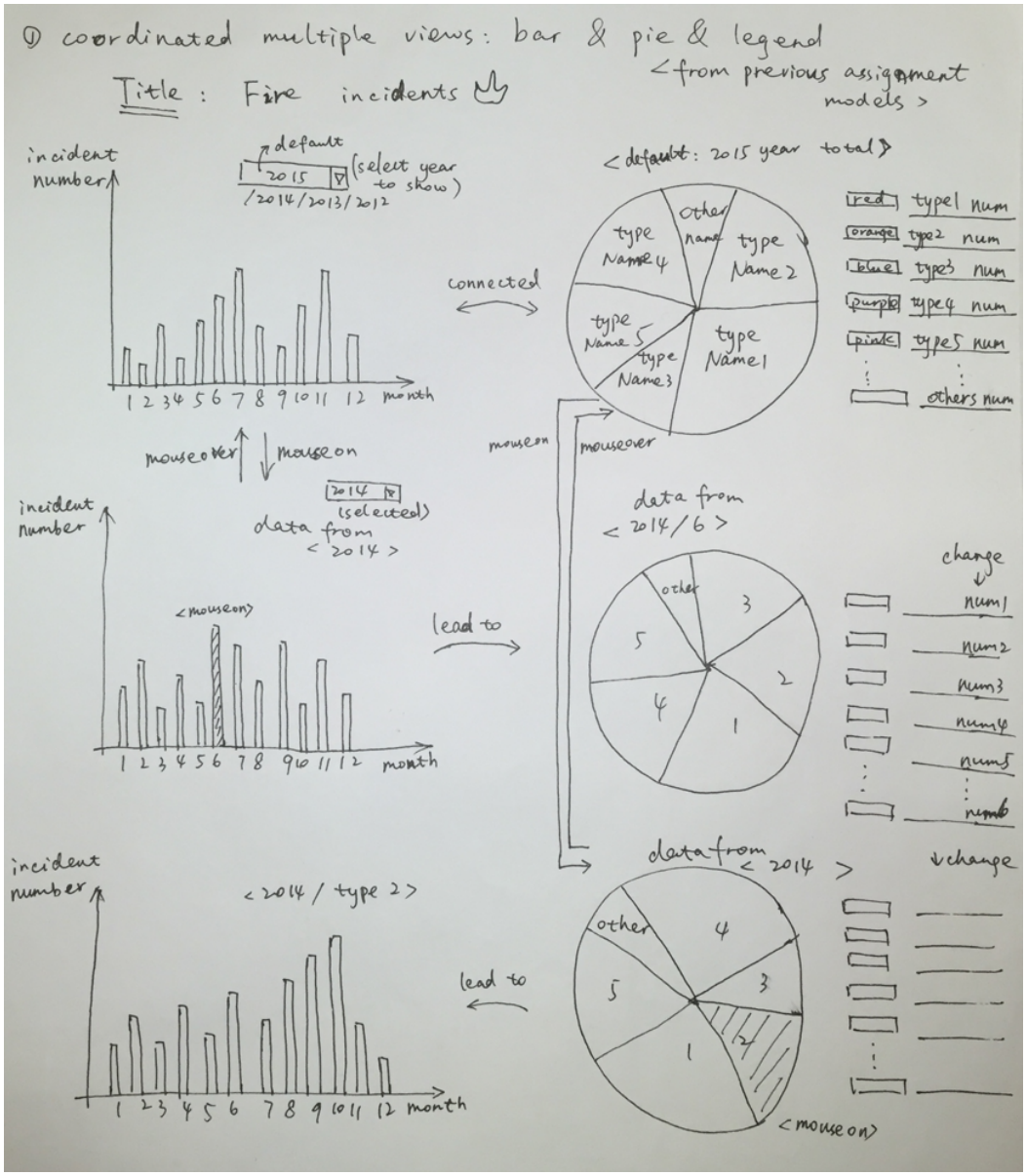
C. Prototype Design

◆ **Design 1:** Coordinated multiple views: bar, donut and related legend.

Design method:

This design contains three parts: one bar chart which contains 12 months of data, one donut chart which contains different types distribution, and legend connect to the donut chart to make the data clearer to show.
The select button on top of this figure needs to be able to choose year to visualize. And the default value is 2015.
The design figure below illustrates the basic functions designed. And this will explore data by mouseon and mouseover.

Visualization preview:



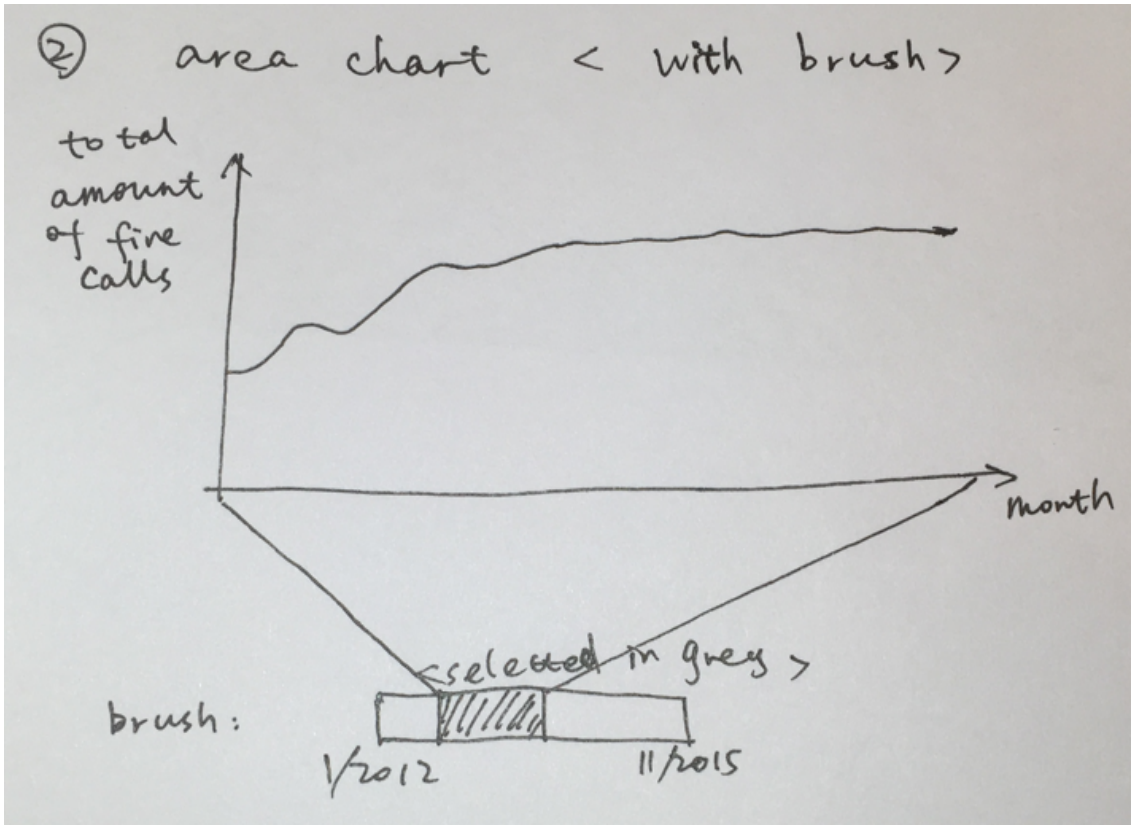
- Design 2: Area chart with time window shows comparison of different types.

Design method:

This design contains several compares to different types. The time window under the figure shows the detailed area chart in selected period of time. The time window is from 2012 to 2015, which contains every record in the dataset.

The figure below shows the imagination of prototype.

Visualization preview:



◆ Design 3: Maps

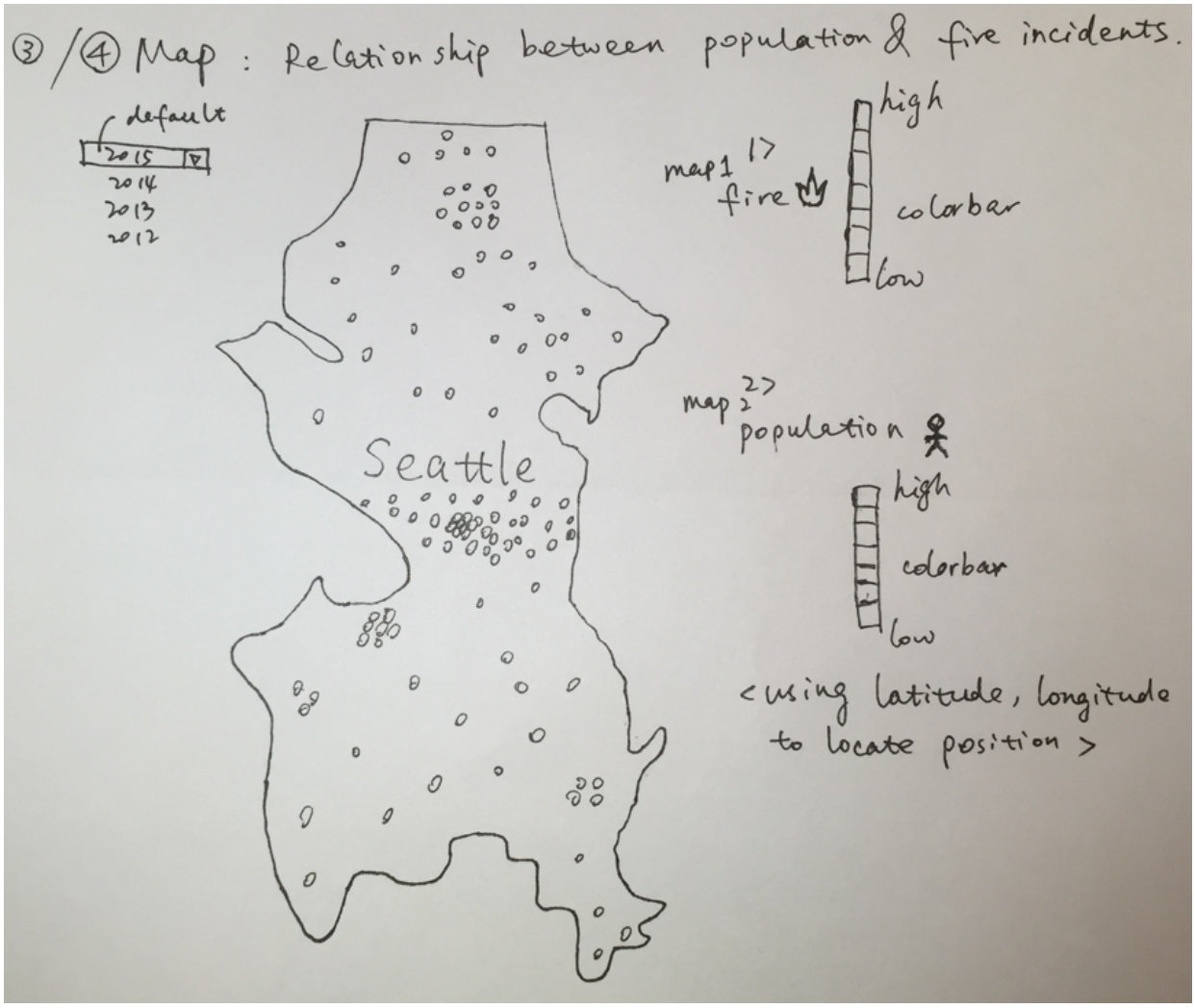
1) Population density 2) Fire incident density

Design method:

This design contains two maps, one is fire incident density map, another is population density map.
The incident density map is supposed to visualize the total numbers of 911 calls in one year. There is a selection button on top of the figure to choose the year. The default value is 2015. There is a color bar shows the frequency from high to low.
The population density map is related to fire incident density map. That is because it is common sense that the place has bigger population density is supposed to have bigger fire incidents density. The color bar shows the frequency from high to low.

The figure below shows the imagination of prototype.

Visualization preview:



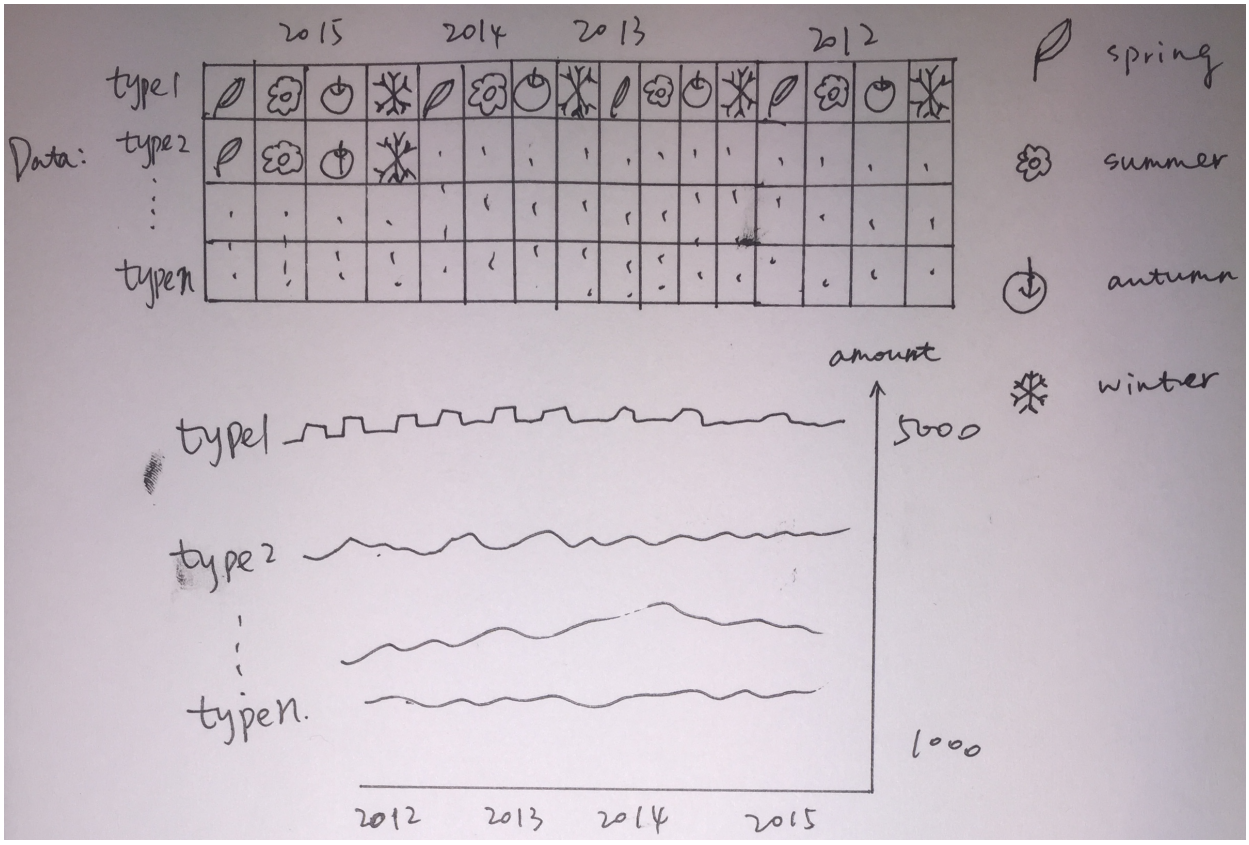
◆ Design 4: Season Distribution (***Added by suggestion of professor***)

Design method:

This design contains one chart that shows the exact data in form aggregated in four seasons. Because we guess that the the relation of different types of data may be reflected by seasons, the idea is also recommended by professor. The figure under the form shows the lines related to the form. By clicking one row of the form, the figure below will show the line related, it can show different lines in one figure so that it is easier to compare the relationship of different types. By clicking the selected row(marked in dark blue) again, the line on the figure will be removed form the figure. This will make it more clear fore the users to see and explore.

The figure below shows the imagination of prototype.

Visualization preview:

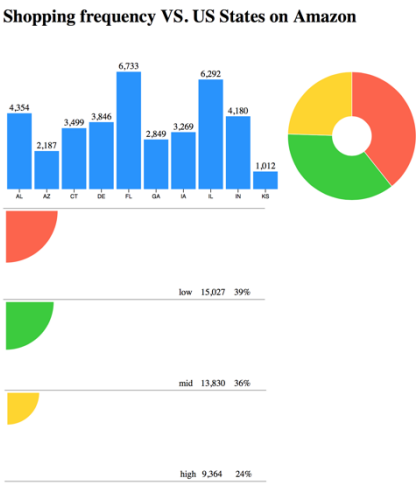


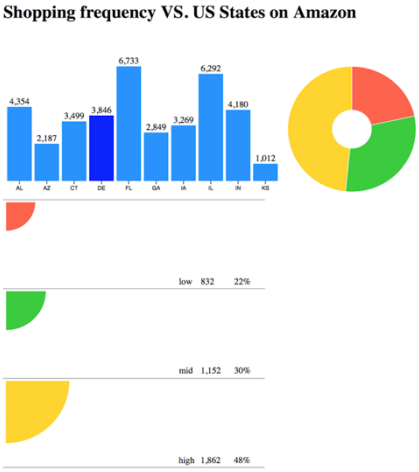
D. Implementing Four Parts of Visualizations

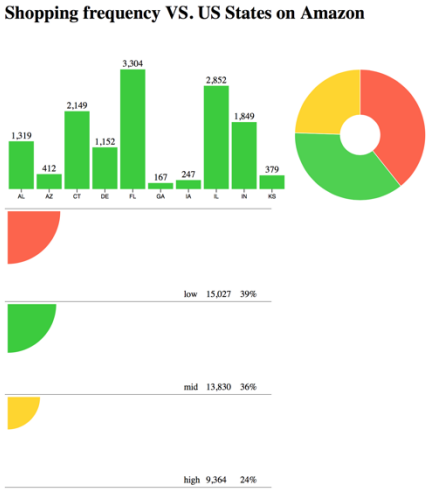
1. Coordinated multiple views: bar chart, pie chart and legend.

Related work:

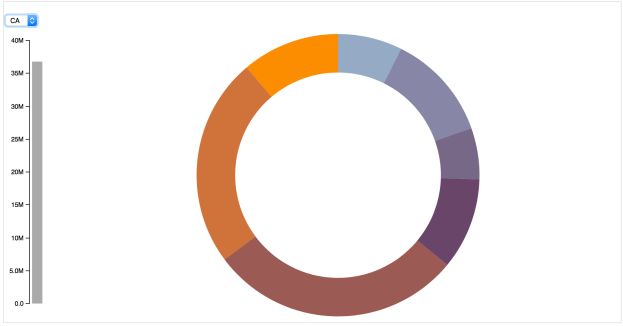
This visualization is inspired by codes from assignment 4, the three related visualizations. That visualization includes one bar chart, one pie chart and one semi-circle chart. The inspiration code is from Qiuzhe Ma (one of the final project author). The inspired design is as follows:







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

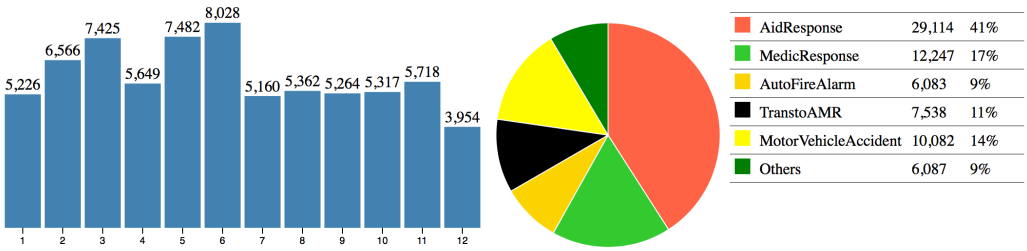


Detailed Functional Description：

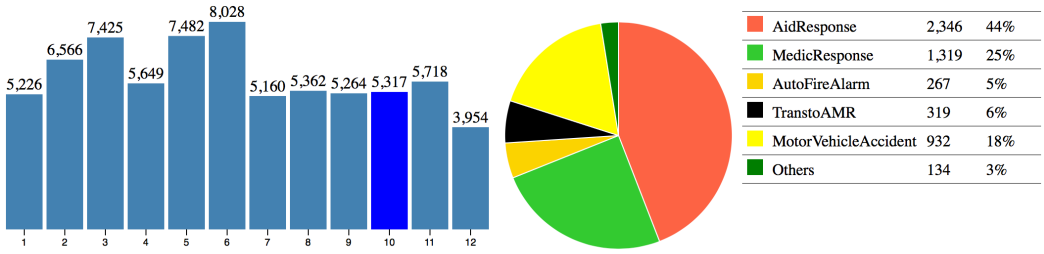
In this part, we want to show the coordinated multiple views about the fire alarm. This part is related to the assignment that I have done before, so that it is not a very hard part for us to deal with.
In the bar diagram, it is supposed to show the total 911 fire alarm calls for one selected year (2015/2014/2013/2012), 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.

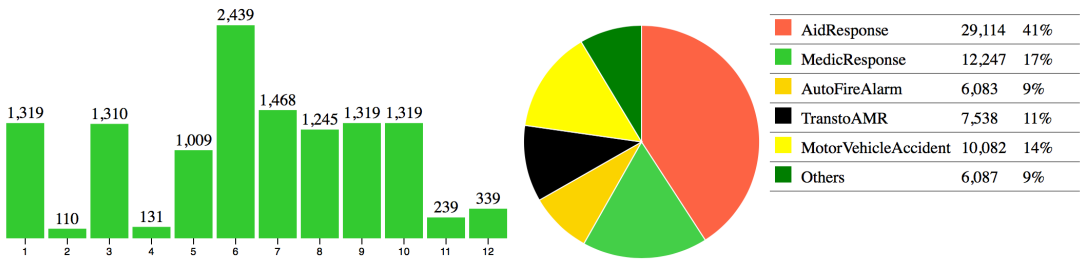
Original: The bar chart shows the total amount of 911 fire alarm calls in year 2015, and the pie chart shows the proportion of each type in year 2015.



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

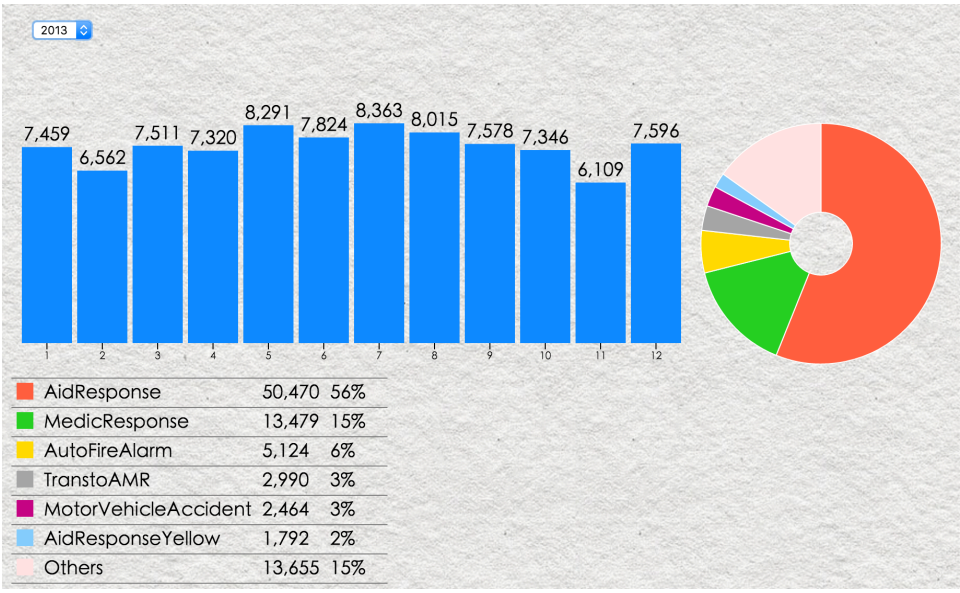


Data Exploratory: When hovering on the pie chart, the bar chart will change related. And the color of the selected pie will be colored darker as shown below.

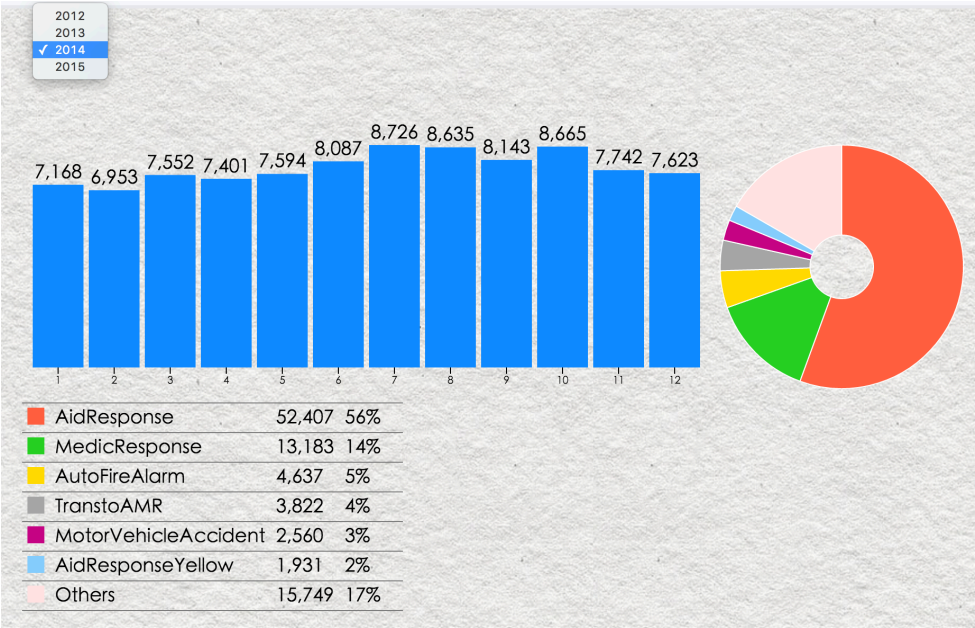


Design Evolution（final implementation）:

We added the selection button on top of the visualization as shown below. It is able to select a year's data to see. We added one type into the dataset of this visualization. We change the color of pie chart and make it more appealing. We also change the pie chart to donut chart to make it more appealing.



Change year selection:



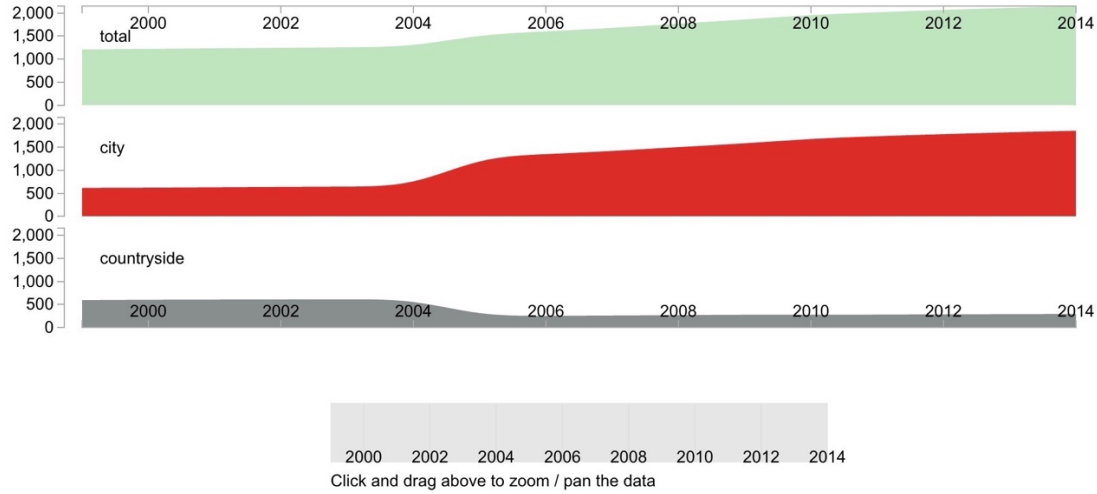
Visualization Analysis:

From the visualization, we can see that mostly, months 7, 8, 9 contains more 911 calls than other months in the four years. And as years grows, the average incidents number becomes bigger, we consider that this is because population is increasing continuously.

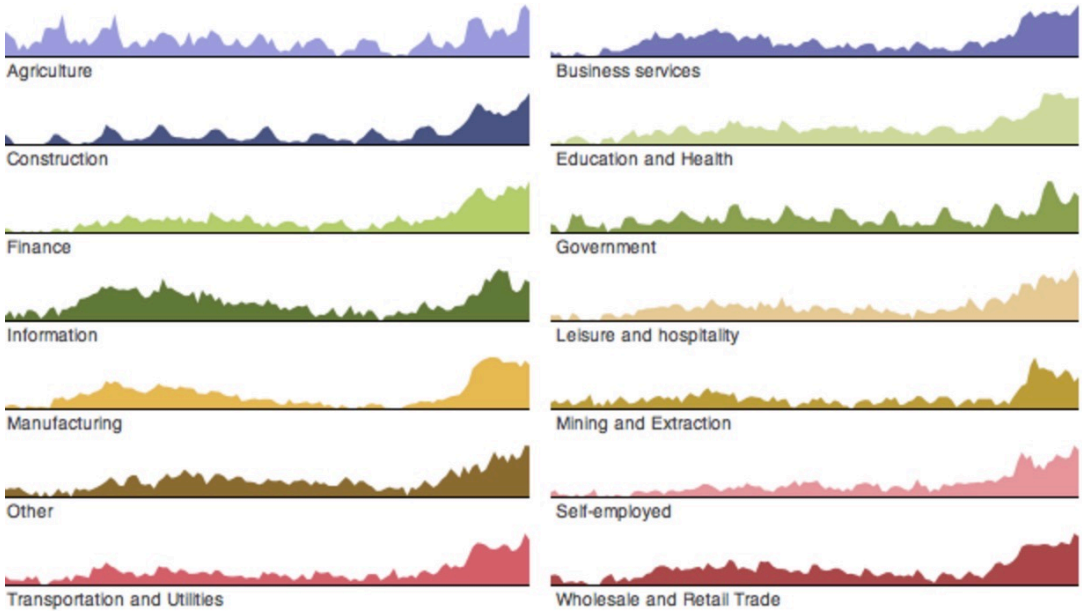
2. *Area chart with time window selection*

Related work:

This visualization is inspired by codes from assignment 4, the three related visualizations. The useful part of that visualization includes 3 area charts as shown below. The inspiration code is from Qianyun Yang (one of the final project author). The inspired design is as follows:

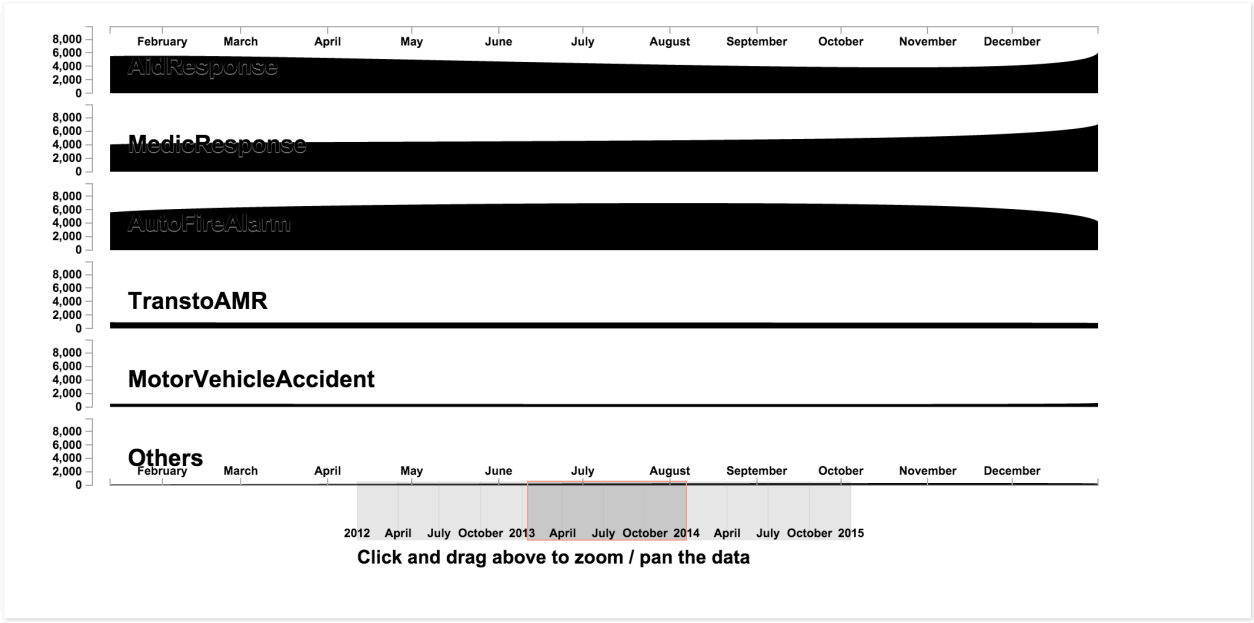


Another inspiration code is from d3 design on the website, which figure is shown as follows. This area chart design is more colorful and show more details:

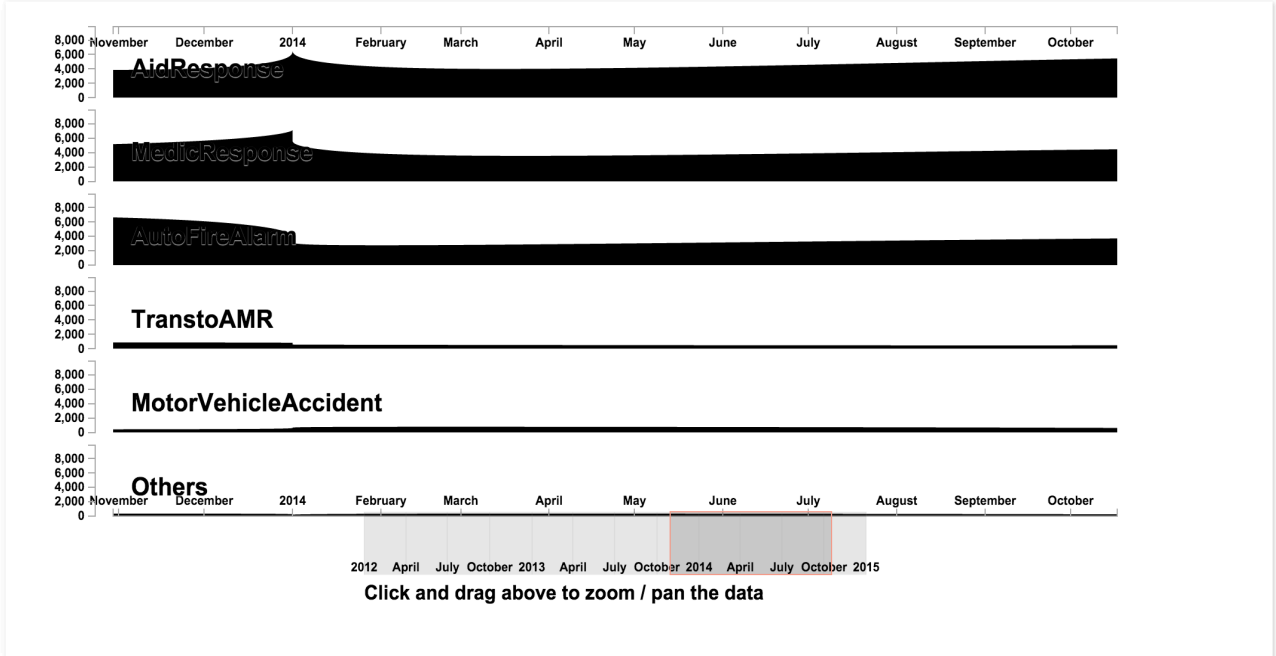


Detailed Functional Description：

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. This shows 6 types: Aid Response, Medic Response, Auto Fire Alarm, Trans to AMR, Motor Vehicle Accident, Others.

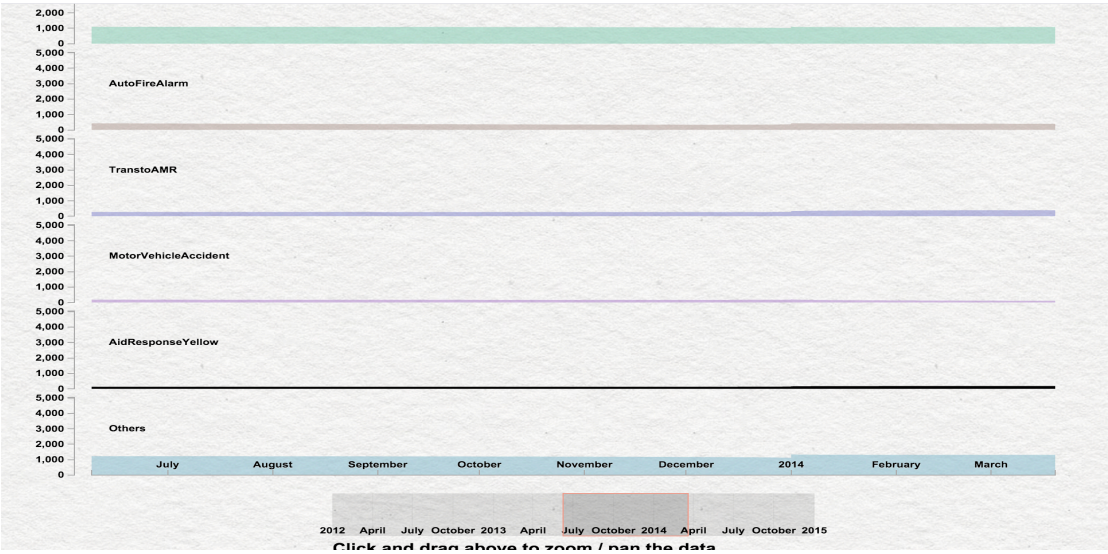


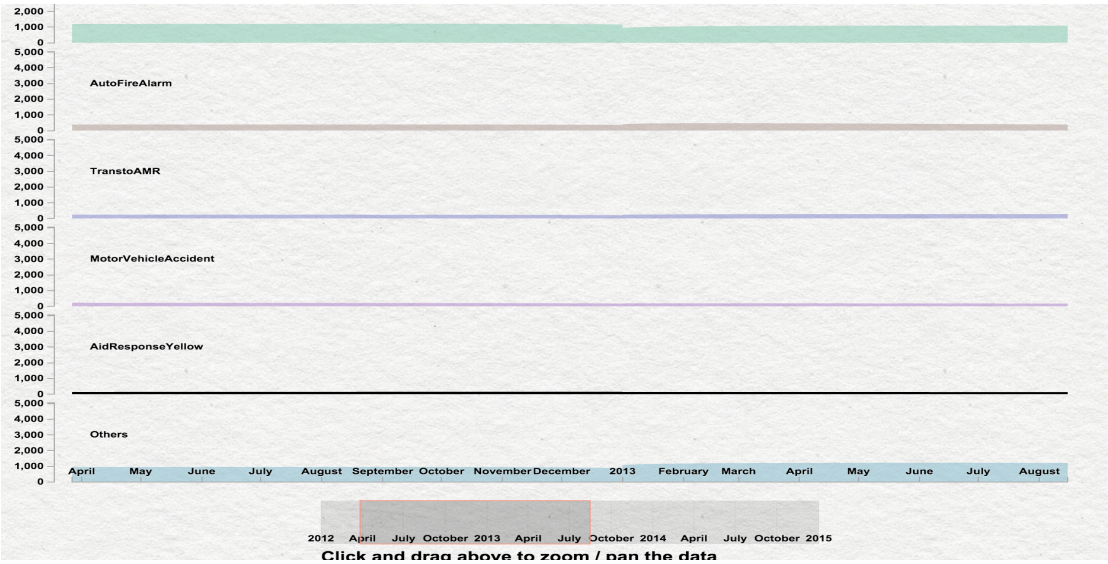
After choose a time period in time window, it shows as follows:



Design Evolution（final implementation）:

We changed the color of the figure and color of the legend to make it more colorful and readable. We also changed 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.
This shows 7 types (we added one type): Aid Response, Medic Response, Auto Fire Alarm, Trans to AMR, Motor Vehicle Accident, Others.





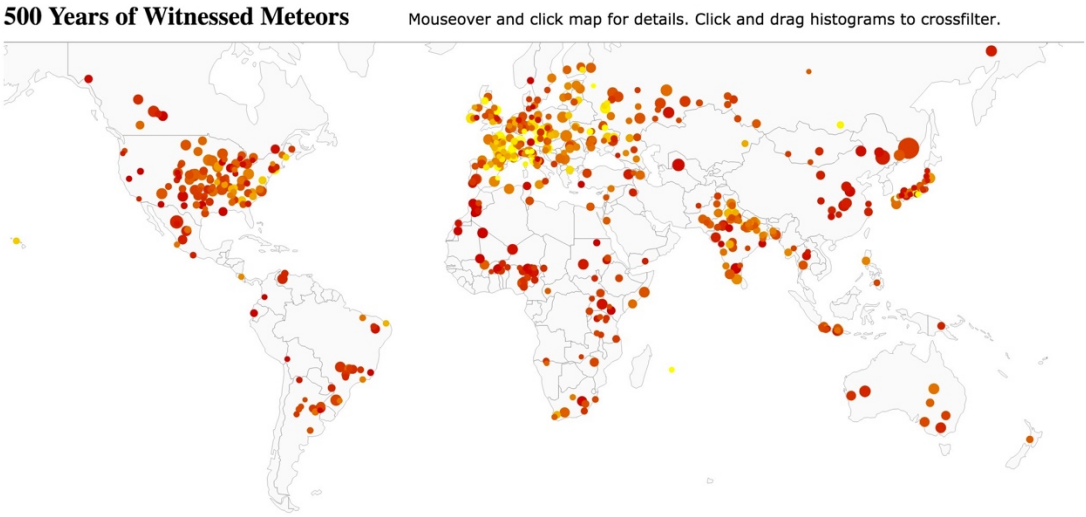
Visualization Analysis:

In this visualization, we find out that there are some sudden growths and falls. For example, in end of year 2012 and beginning of year 2013, there is a sudden fall in type Aid Response. While in end of year 2013 and beginning of 2014, there is a sudden growth in type Auto Fire Alarm and in type Trans to AMR as well.
We also find out that the data mostly changes smoothly and the amount of each type almost contains the same with a subtle changing over time.

3. Fire Incidents Density VS. Population Density Maps

Related work:

This design is inspired form the map design in d3 examples on the website. We choose this design to help design ourselves. From the figure below, we can see that there are different colors of points in the map and also different size of circles to show the amount in the dataset.



Detailed Functional Description:

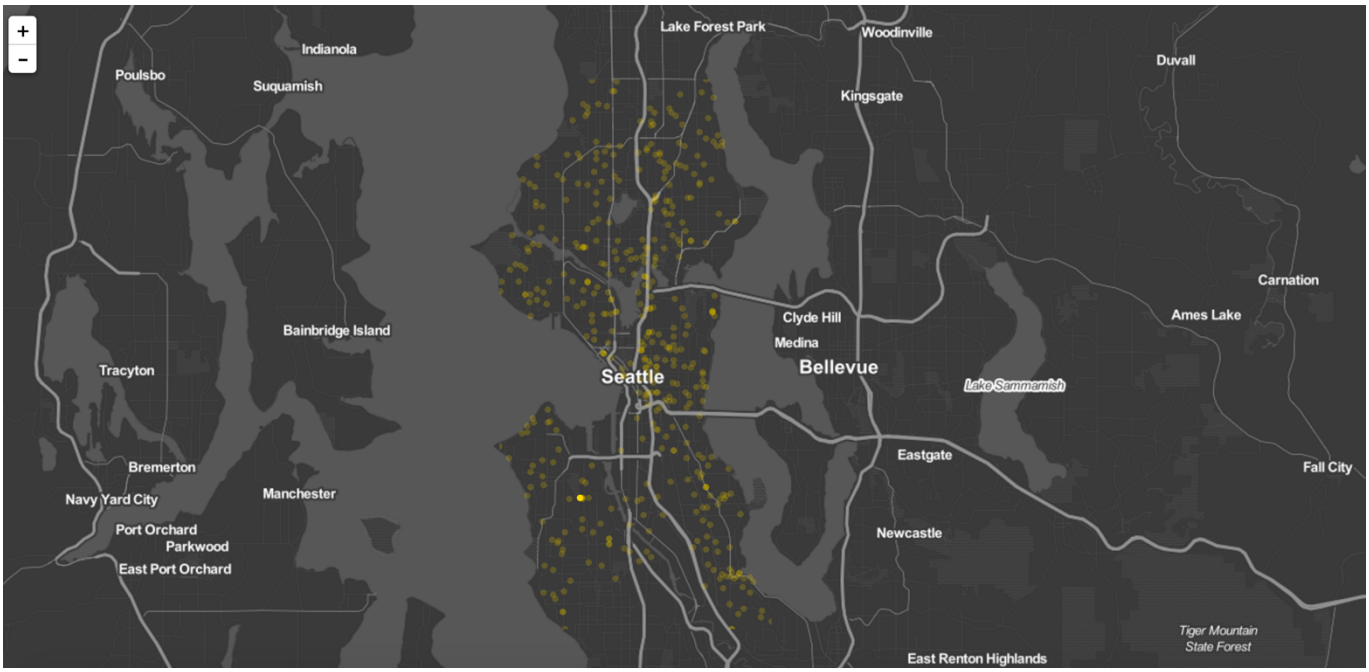
However, we find out that the map is a very hard one to design because it is not similar to common visualizations, it contains the json file that describe the area map, and it is hard to find Seattle's map that is completed already.

Therefore, in our design, we basically use three steps to design the map:

- ✓ We find a world map that contains latitude and longitude information inside. Because this resource is necessary so that we can draw points on the map.
- ✓ We center view the map to Seattle to make it a Seattle City view.
- ✓ We find the longitude and latitude information in the 911 calls dataset, and try to mark points on the map based on this information.

1) Fire incident density

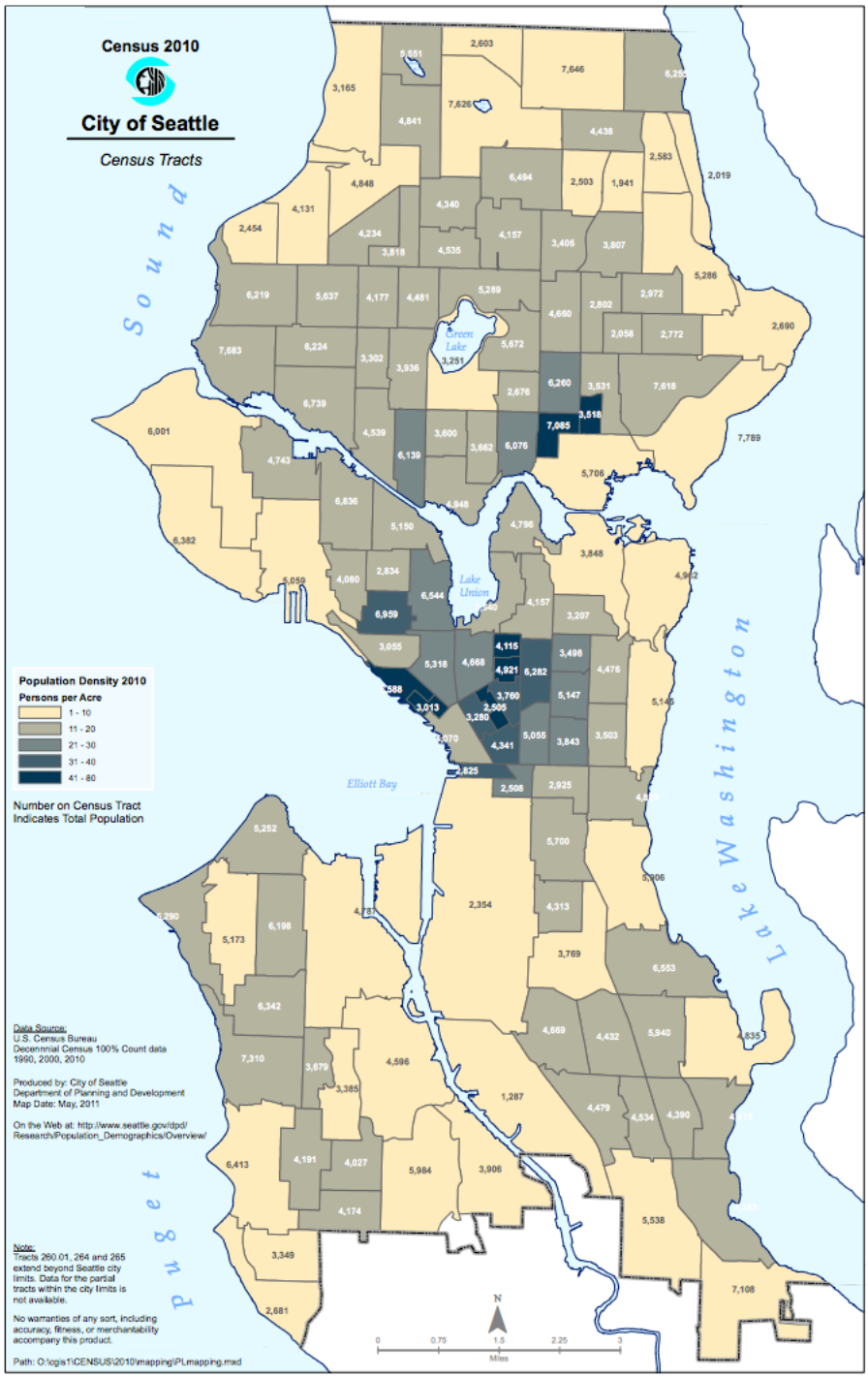
The figure below uses the dataset of one kind of Seattle 911 calls, named WireDown.csv. This figure is just to show the blueprint of the different kinds of types in points map.



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

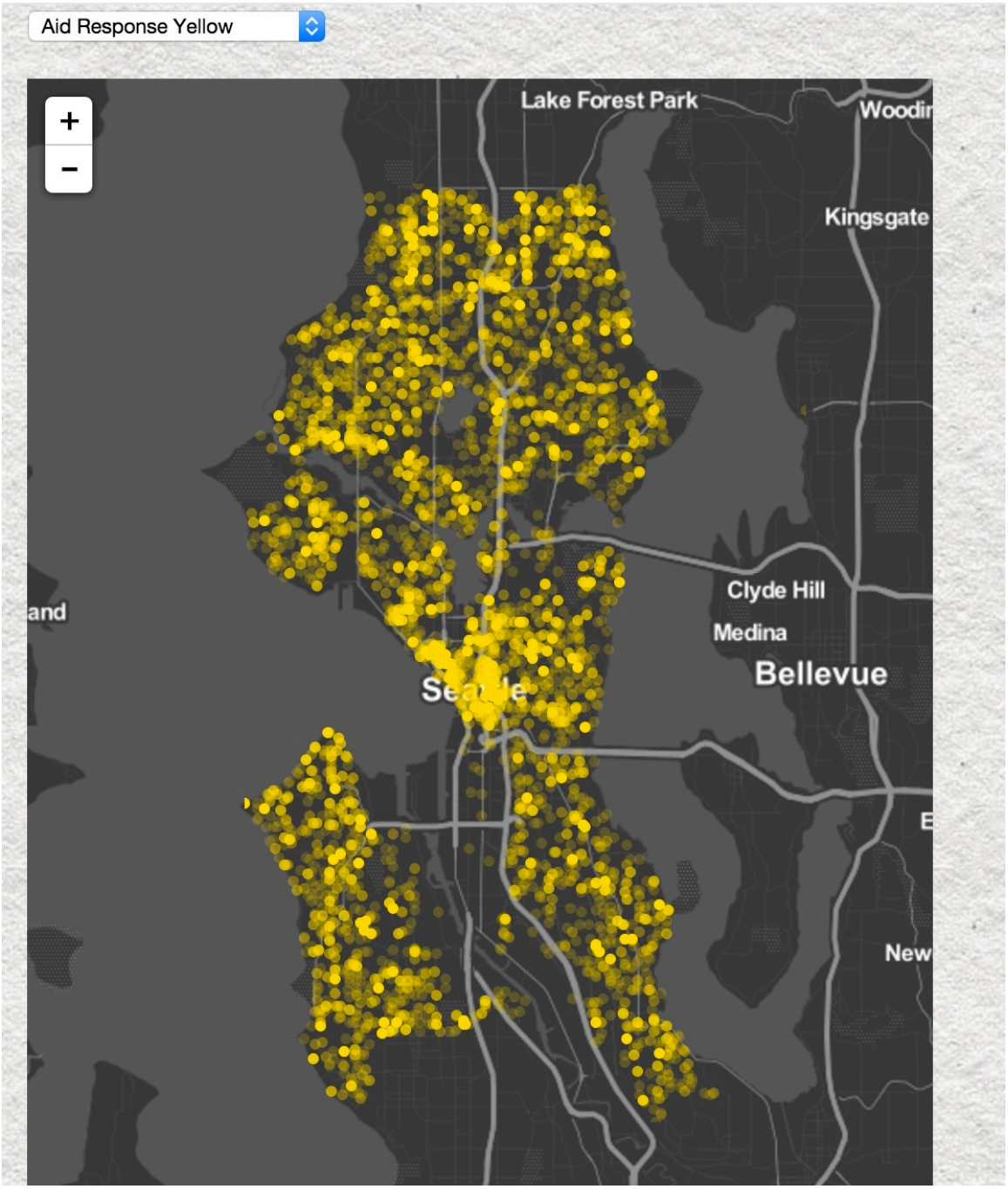
The figure above shows the population density of Seattle in year 2010. We find out that census will be calculated every 5 years. It is not every year. But it still can show and explain the relationship between fire alarm frequency and population density.



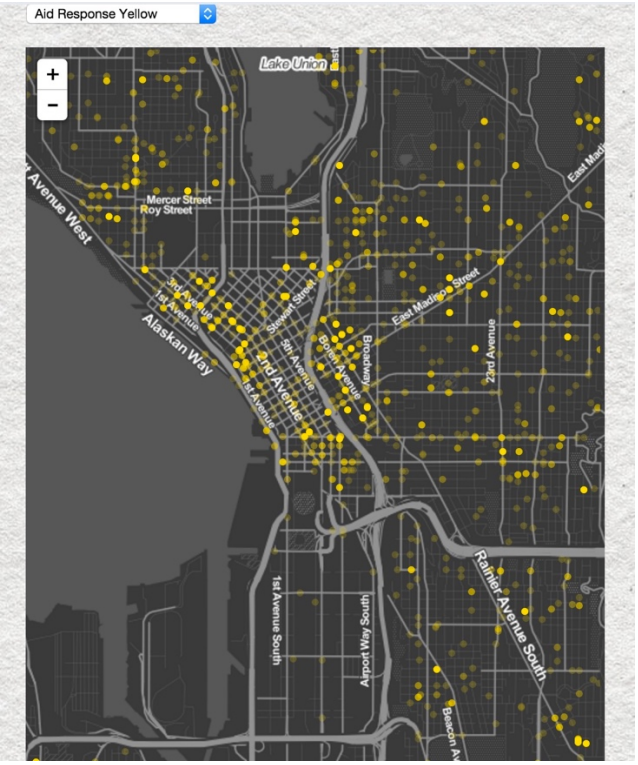
Design Evolution (final implementation):

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 improved is added several types selected from the dataset and show them by selection button on the map.

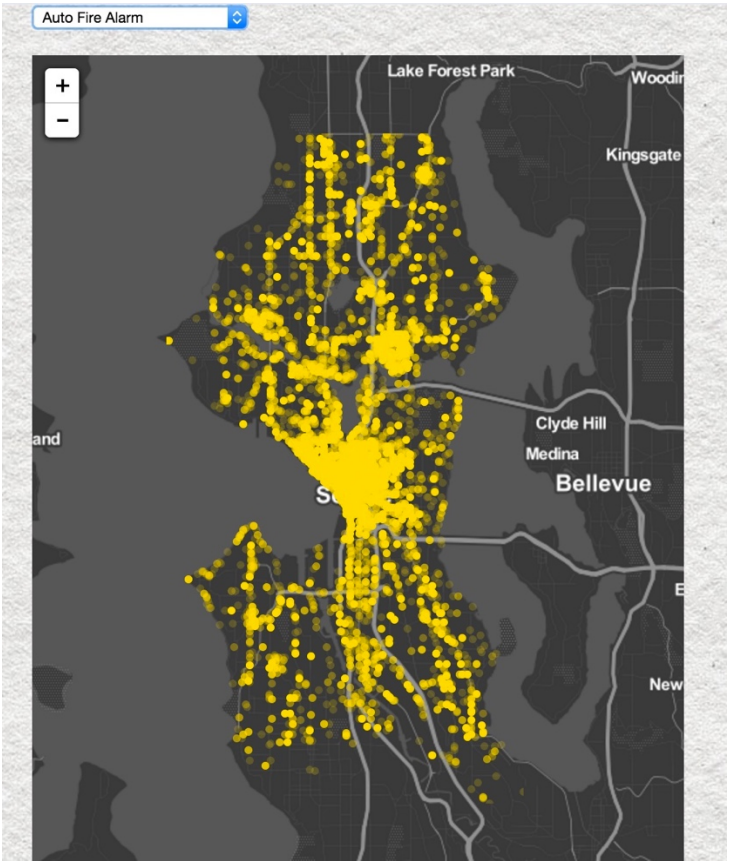
Default value for type selection is Aid Response Yellow. As shown below:



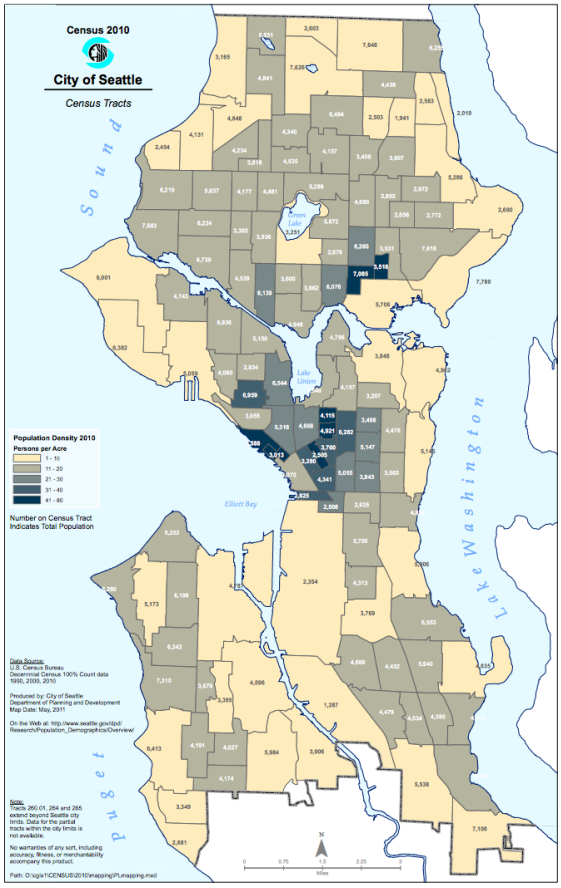
By zooming in and out we can see detailed marked points. I also select the Aid Response Yellow type to see.



By changing the selection button, we can see different types. Such as Auto Fire Alarm.



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. We find out that there is no directly related map and population density resources in the website for us to design. We can only find population information that is in year 2010 or before, current data needs to buy from the website at about 300 dollars. We also find a population density map in 2010. This may give some related information for our users to see that fire incident density is related to population density.



Visualization Analysis:

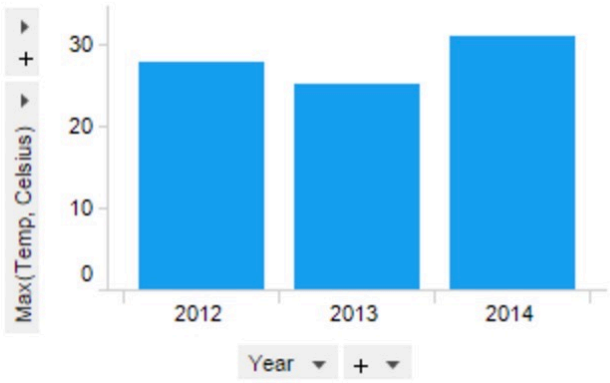
In this part, we can see that the different type has different density distribution. But it is all related to population density. This can be noticed by comparing two maps. It is also a very straightforward visualization to see the data in detail. It is very useful for government and fire work department to allocate resources rationally.

4. Season Distribution (Recommended by Professor)

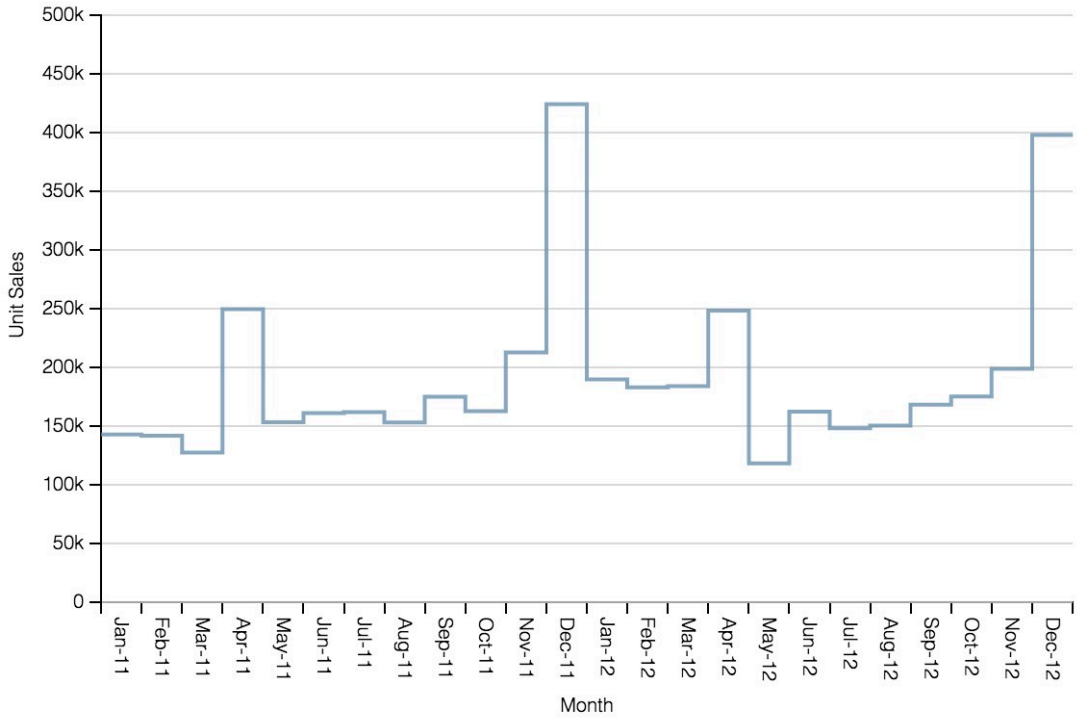
Related work:

In this visualization, we are inspired by these two visualizations. The first visualization contains a form which describes the four season and a visualization that show one year's season in bar chart. However, it can not show the variation tendency for each year.

Year	Season	Temp, Celsius
2012	Winter	3.90
2012	Spring	15.60
2012	Summer	27.80
2012	Fall	15.30
2013	Winter	-6.70
2013	Spring	22.00
2013	Summer	25.20
2013	Fall	16.70
2014	Winter	0.50
2014	Spring	21.00
2014	Summer	31.10
2014	Fall	19.50



Therefore, we find this visualization below to show all 4 years' season in one form and in one figure. By using this figure, we can make the visualization clearer.



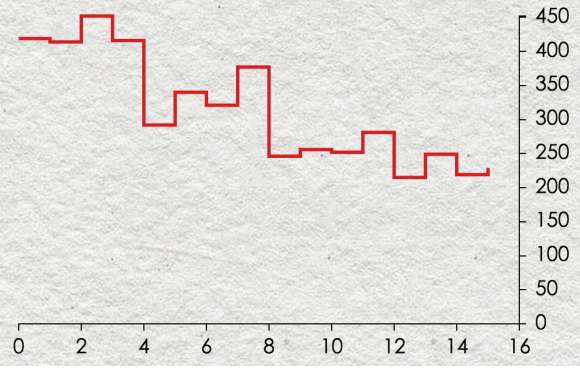
Detailed Functional Description:

This season part is recommended by professor in the proposal presentation. In this visualization, we combined form and figure together as one season distribution visualization.

	2012				2013				2014				2015			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
AidResponse	4,476	4,937	4,375	3,812	4,246	4,744	4,471	4,089	4,286	4,643	3,894	3,965	3,719	4,108	3,787	4,067
MedicReponse	1,232	1,247	1,214	1,096	1,074	1,127	1,154	1,111	1,228	1,101	1,004	1,062	1,252	1,254	1,212	1,182
AutoFireAlarm	369	424	409	366	366	386	429	382	454	396	422	416	372	432	420	377
TranstoAMR	419	414	452	416	292	340	321	377	246	256	252	281	215	249	219	229
MotorVehicleIncident	229	244	247	232	213	240	247	128	209	214	207	187	213	215	223	200
AidResponseYellow	160	190	184	156	169	159	157	173	149	156	137	143	161	151	161	151
Others	1,162	1,633	1,443	990	1,155	1,487	1,405	1,215	1,136	1,300	1,095	1,184	978	1,155	1,091	916

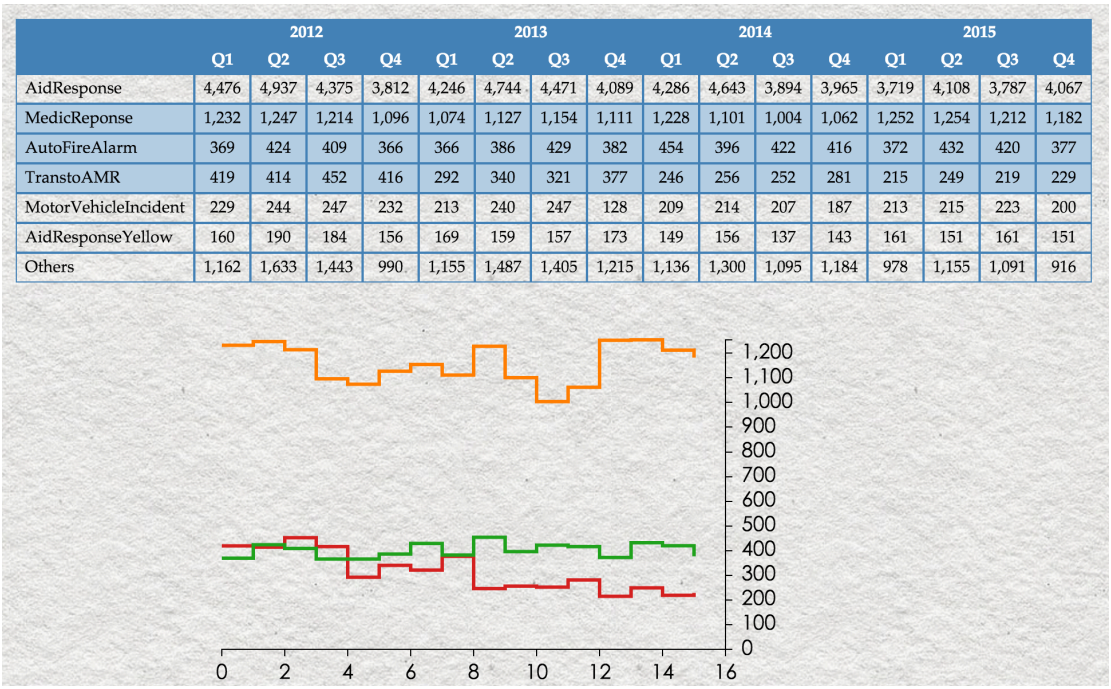
When mouseon one line on the form, and click it, it will draw the line underneath in the figure. The data in the figure shows not only the tendency of four seasons, but the tendency of each year.

	2012				2013				2014				2015			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
AidResponse	4,476	4,937	4,375	3,812	4,246	4,744	4,471	4,089	4,286	4,643	3,894	3,965	3,719	4,108	3,787	4,067
MedicReponse	1,232	1,247	1,214	1,096	1,074	1,127	1,154	1,111	1,228	1,101	1,004	1,062	1,252	1,254	1,212	1,182
AutoFireAlarm	369	424	409	366	366	386	429	382	454	396	422	416	372	432	420	377
TranstoAMR	419	414	452	416	292	340	321	377	246	256	252	281	215	249	219	229
MotorVehicleIncident	229	244	247	232	213	240	247	128	209	214	207	187	213	215	223	200
AidResponseYellow	160	190	184	156	169	159	157	173	149	156	137	143	161	151	161	151
Others	1,162	1,633	1,443	990	1,155	1,487	1,405	1,215	1,136	1,300	1,095	1,184	978	1,155	1,091	916



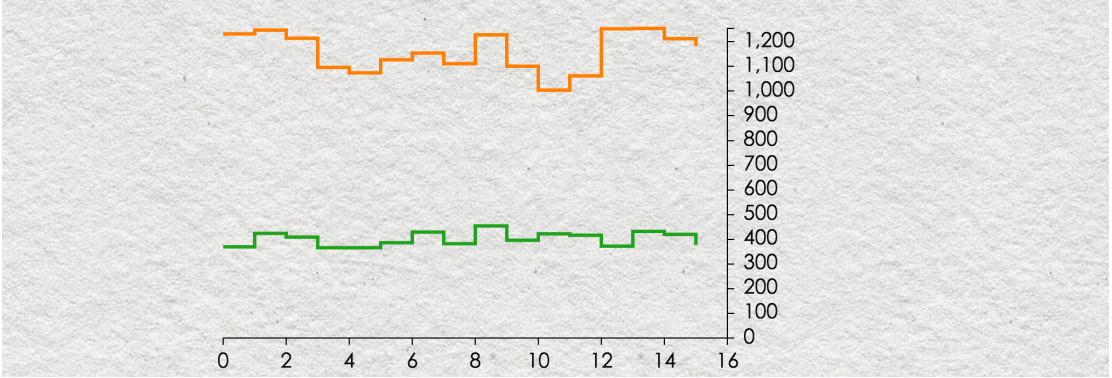
We can select several lines together to compare the variation tendency difference between different types.

	2012				2013				2014				2015			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
AidResponse	4,476	4,937	4,375	3,812	4,246	4,744	4,471	4,089	4,286	4,643	3,894	3,965	3,719	4,108	3,787	4,067
MedicReponse	1,232	1,247	1,214	1,096	1,074	1,127	1,154	1,111	1,228	1,101	1,004	1,062	1,252	1,254	1,212	1,182
AutoFireAlarm	369	424	409	366	366	386	429	382	454	396	422	416	372	432	420	377
TranstoAMR	419	414	452	416	292	340	321	377	246	256	252	281	215	249	219	229
MotorVehicleIncident	229	244	247	232	213	240	247	128	209	214	207	187	213	215	223	200
AidResponseYellow	160	190	184	156	169	159	157	173	149	156	137	143	161	151	161	151
Others	1,162	1,633	1,443	990	1,155	1,487	1,405	1,215	1,136	1,300	1,095	1,184	978	1,155	1,091	916



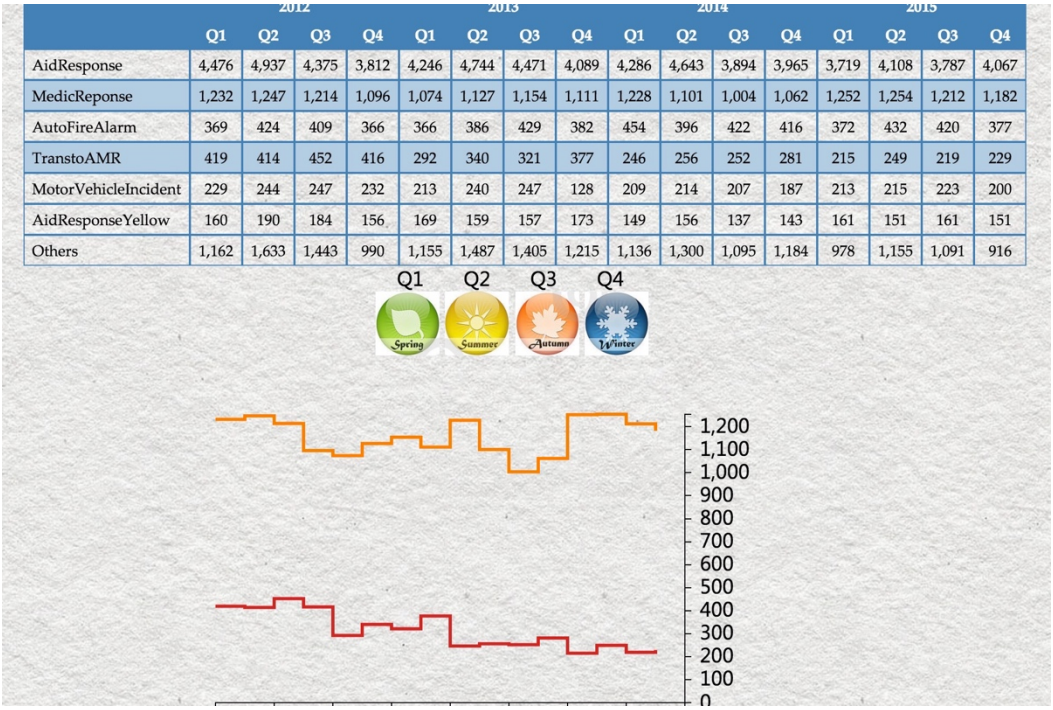
To remove the selected type, you just need to click the line in the data again. Different types are shown in different colors.

	2012				2013				2014				2015			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
AidResponse	4,476	4,937	4,375	3,812	4,246	4,744	4,471	4,089	4,286	4,643	3,894	3,965	3,719	4,108	3,787	4,067
MedicReponse	1,232	1,247	1,214	1,096	1,074	1,127	1,154	1,111	1,228	1,101	1,004	1,062	1,252	1,254	1,212	1,182
AutoFireAlarm	369	424	409	366	366	386	429	382	454	396	422	416	372	432	420	377
TranstoAMR	419	414	452	416	292	340	321	377	246	256	252	281	215	249	219	229
MotorVehicleIncident	229	244	247	232	213	240	247	128	209	214	207	187	213	215	223	200
AidResponseYellow	160	190	184	156	169	159	157	173	149	156	137	143	161	151	161	151
Others	1,162	1,633	1,443	990	1,155	1,487	1,405	1,215	1,136	1,300	1,095	1,184	978	1,155	1,091	916



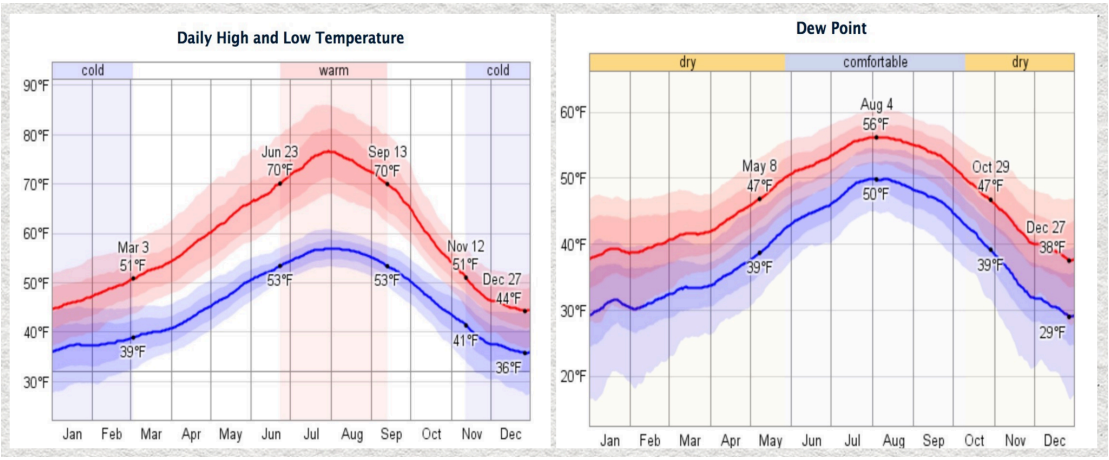
Design Evolution (final implementation):

We added some icons to show the meaning of Q1, Q2, Q3, Q4 in the form, which means spring, summer, autumn, winter orderly.



	2012				2013				2014				2015			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
AidResponse	4,476	4,937	4,375	3,812	4,246	4,744	4,471	4,089	4,286	4,643	3,894	3,965	3,719	4,108	3,787	4,067
MedicReponse	1,232	1,247	1,214	1,096	1,074	1,127	1,154	1,111	1,228	1,101	1,004	1,062	1,252	1,254	1,212	1,182
AutoFireAlarm	369	424	409	366	366	386	429	382	454	396	422	416	372	432	420	377
TranstoAMR	419	414	452	416	292	340	321	377	246	256	252	281	215	249	219	229
MotorVehicleIncident	229	244	247	232	213	240	247	128	209	214	207	187	213	215	223	200
AidResponseYellow	160	190	184	156	169	159	157	173	149	156	137	143	161	151	161	151
Others	1,162	1,633	1,443	990	1,155	1,487	1,405	1,215	1,136	1,300	1,095	1,184	978	1,155	1,091	916

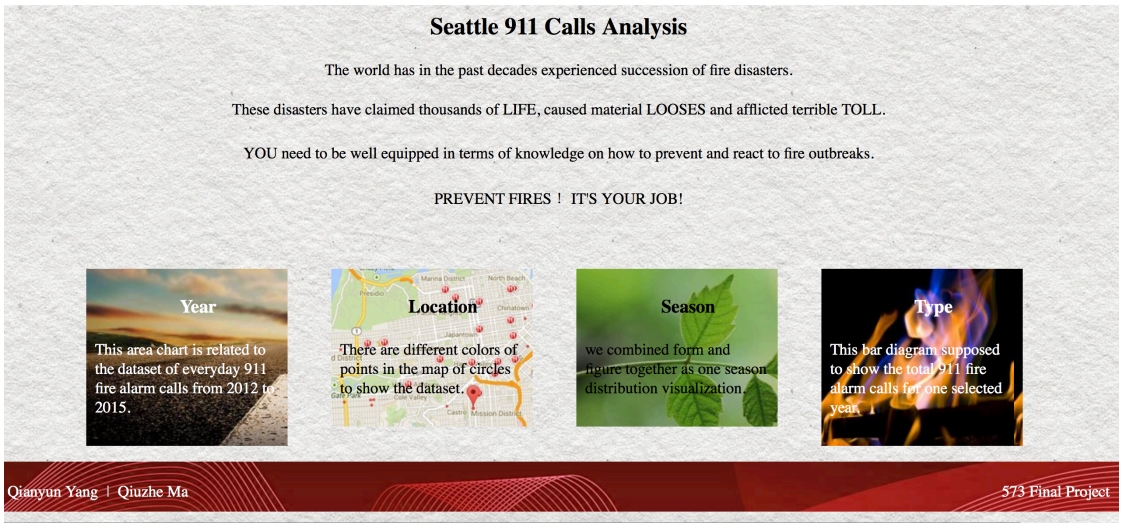
The figure below is downloaded from the website show the Seattle 2015 temperature and humidity. Because we consider that season analysis is related to temperature and humidity.



Visualization Analysis:

In this visualization, we can find that in summer, the temperature and humidity are both high. Compare to this, we find that most of the 911 fire calls happens in summer while winter contains the lowest frequency of 911 calls.

E. Home Page Design



F. Conclusion

In our final project, we try to design a website with several visualizations to help show the 911 fire alarm calls happened in Seattle, the main dataset is downloaded from Seattle official website. We design a very useful website that explore the data by year, season, type and location. We also combined this data with other related information online to help analyze the data in a reasonable manner.

Here are some useful conclusions we got from our four visualizations:

- ✓ From the ***type*** visualization, we can see that months 7, 8, 9 contains more 911 calls than other months in the four years mostly. And as years grows, the average incidents number becomes bigger, we consider that this is because population is increasing continuously.
- ✓ From the ***year*** visualization, we find out that there are some sudden growths and falls. For example, in end of year 2012 and beginning of year 2013, there is a sudden fall in type Aid Response. While in end of year 2013 and beginning of 2014, there is a sudden growth in type Auto Fire Alarm and in type Trans to AMR as well. We also find out that the data mostly changes smoothly and the amount of each type almost contains the same with a subtle changing over time.
- ✓ From the ***location*** visualization, we can see that the different type has different density distribution. But it is all related to population density. This can be noticed by comparing two maps. It is also a very straightforward visualization to see the data in detail. It is very useful for government and fire work department to allocate resources rationally.
- ✓ From the ***season*** visualization, we can find that in summer, the temperature and humidity are both high. Compare to this, we find that most of the 911 fire calls happens in summer while winter contains the lowest frequency of 911 calls.