Data Visualization Final Project

By Jian Sun, YiLing Chen

1.Choose Dataset

Three datasets are chosen by us to do visualization.

The first is **"GEDtesternum.csv"**. It's about the Female and Male GED testers in year 2010. We plan to show the number of testers for both gender in different age groups.
The chosen plot is **Stacked Column HighChart**.
GED is **General Educational Development.** Here is hyperlink for the test, https://ged.com/

The Second is **"PM25valueChina.csv"**. It's about the concentration of PM2.5 among five main Chinese cities from May, 2013 to September, 2016. We plan to show the changes for the PM 2.5 value of 5 cities in the past 5 years
The chosen plot is **Basic Line HighChart**.

The third is **"horsepowerandpower.csv"**. It's about the relationship between vehicle's horsepower and price.
The chosen plot is **Area HighChart**.

2. Load the Datasets

Step1. Open MySQL and create a new Schema called charts.

Step2. Right Click the charts, click Table Data Import Wizard, then browse our computer and load all three datasets.

Step3. Fresh the charts Schema.

3. Code three charts

Stacked Column HighChart, Basic Line High Chart and Area High Chart are three target charts.
According to our datasets, we coded three JavaScript files to plot and three HTML files to control the size of the plots. the "expressRoutesMySql.js" is used to connect the dataset loaded in MySQL, HighChart and Node Server.

4. Build Node Server

Step1. Open the file "Visualization Final Project" in the Brackets.

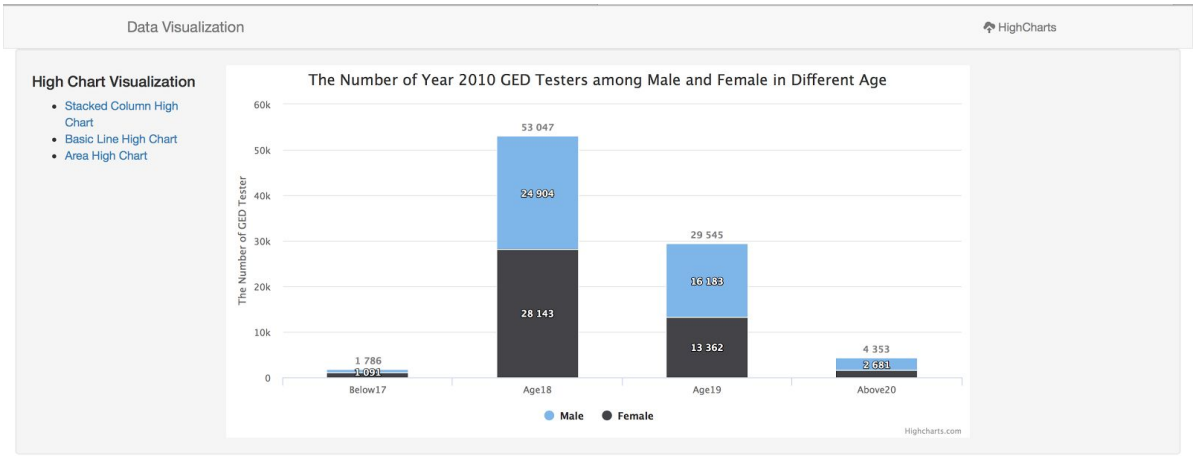
Step2. Open terminal in MacBook, go to file "Visualization Final Project" → file "final", which includes file "config", "filestoload", "package.json", "public", "server.js", "servermodules" and "node_modules".

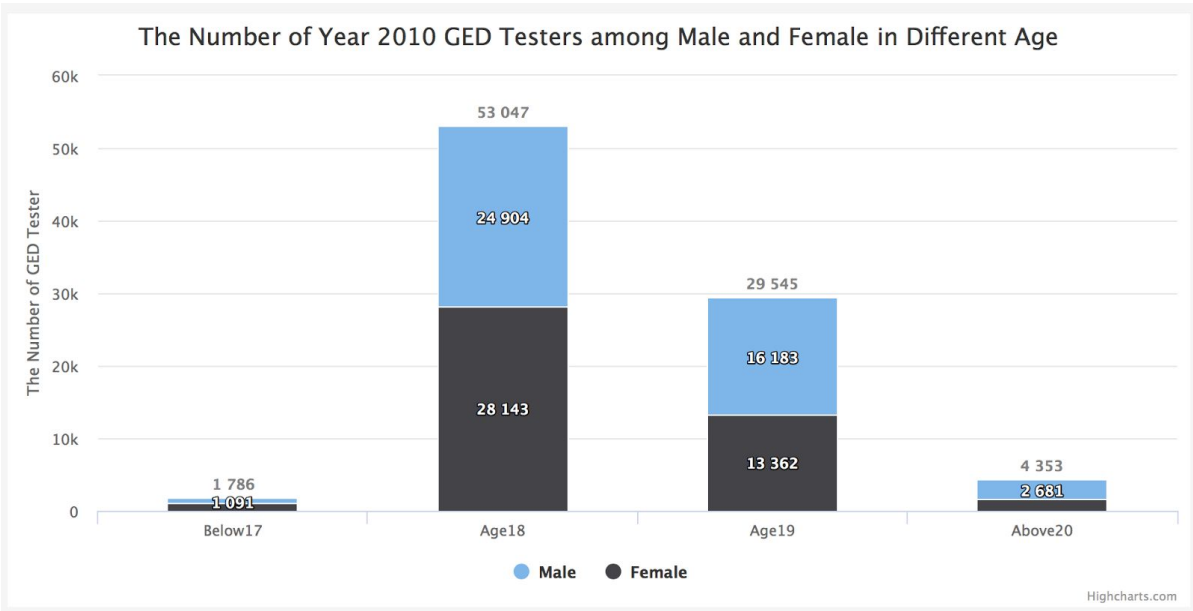
Step3. Build Node Server under this file with the code like "npm init ; npm install express mongoose morgan body-parser method-override mysql ; node server.js". Finally, enter localhost:8088 to see results in Safari or Chrome.

5. Plot Explanation

5.1 Stacked Column HighChart

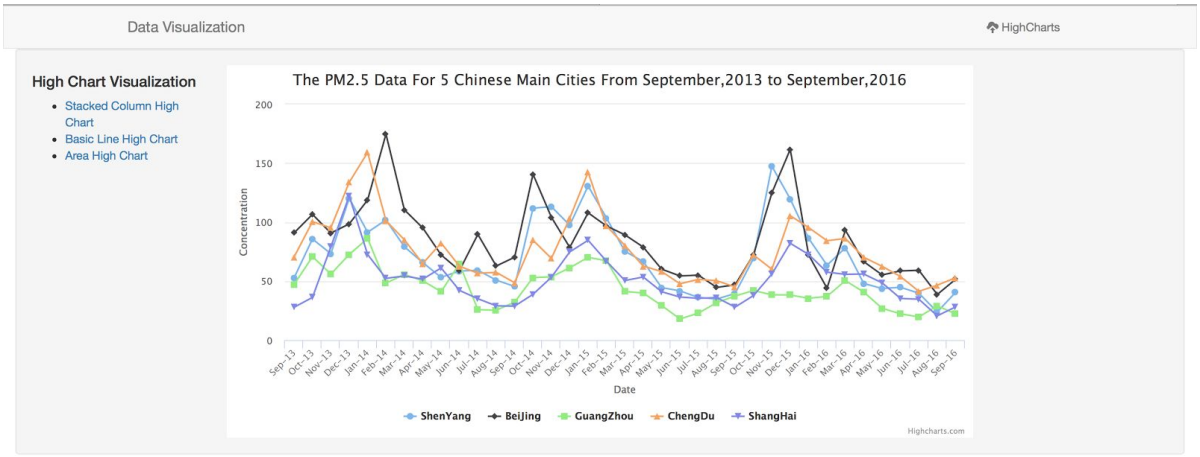
The following plot is to show the number of GED Testers for both genders in different age groups. The total number is 88731. The largest group is age 18, which occupies 53047 people. Within the group, the Female testers are 28143, which is bigger than Male testers, 24904. The second largest group is age 19, which has 29545 people. In general, the Female testers are more than the Male testers. That's mean, in 2010, Female is more keen on GED test than Male.

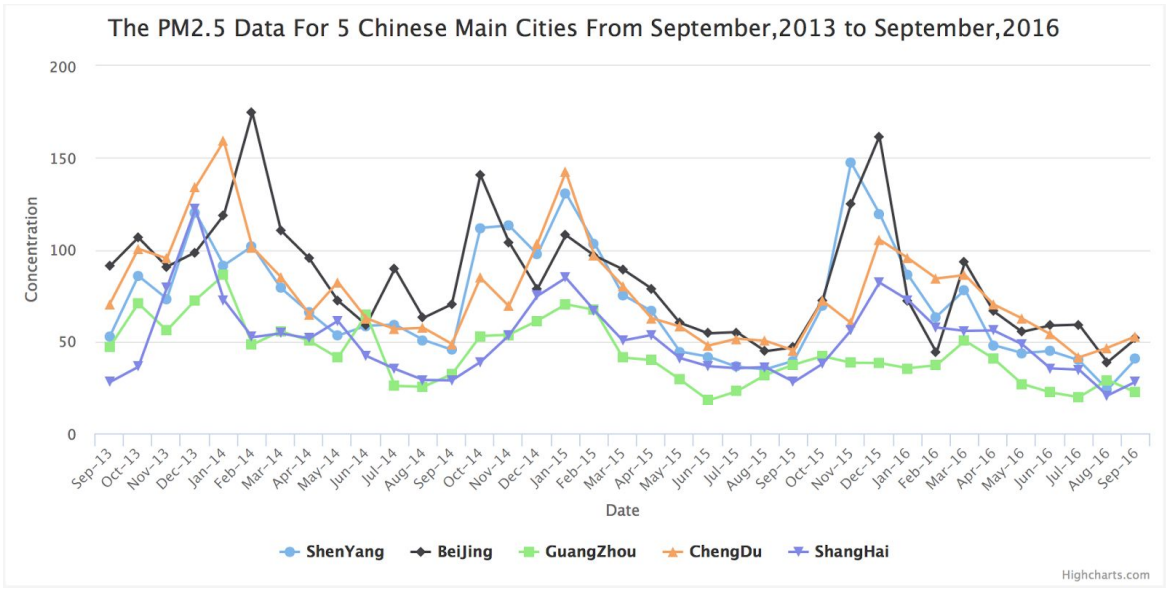




5.2 Basic Line HighChart

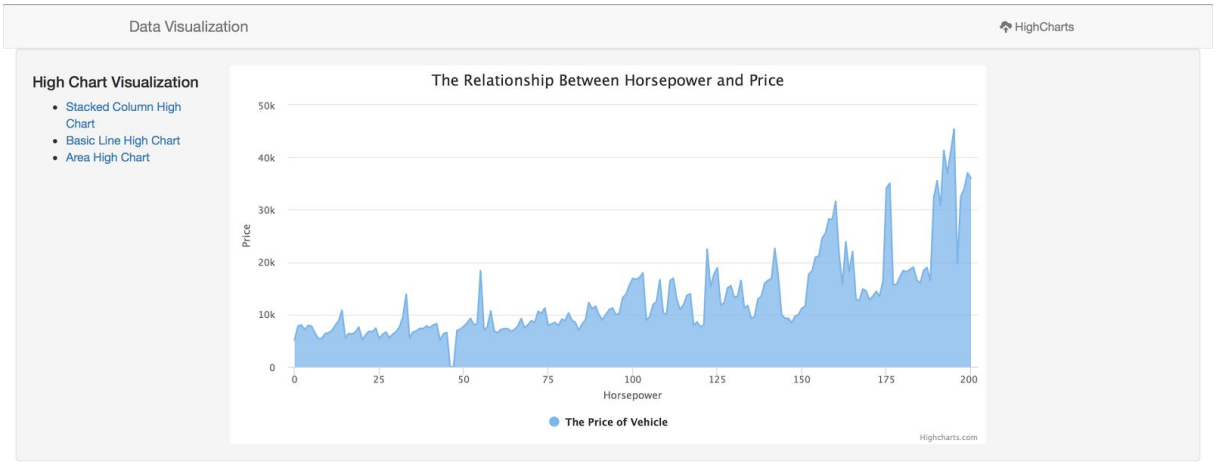
The following plot is to show the changes of PM 2.5 for 5 Chinese main cities, ShenYang, BeiJing, GuangZhou, ChengDu and ShangHai, from 9, 2013 to 9, 2016. It's obvious to see that the BeiJing, black line, has the relative highest PM 2.5 value, which fluctuates around 75 ug/m3, for the past 4 years. Conversely, GuangZhou, green line, is in relative lowest value, which fluctuates around 30 ug/m3, for the past 4 years. All five cities have peak value in winter, that's mean the winter has the largest PM 2.5 concentration and the worst air quality among the year. It's good for people to wear breathing mask while walking down the street.

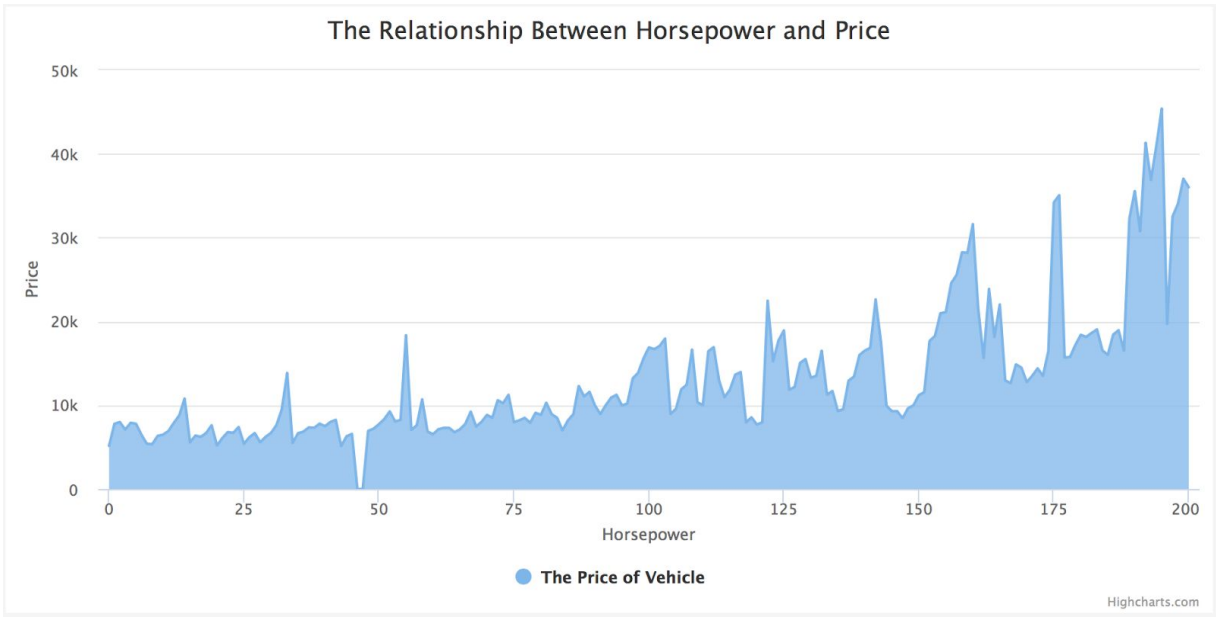




5.3 Area HighChart

The following plot is to show the price of vehicles for different horsepower. The total number of vehicles is 201. The largest price is 45400 dollars, which have 184 HP. The smallest price is 5151 dollars, which have 48 HP. In general, the more horsepower vehicle have, the more expensive the vehicle will be. What's more, when the horsepower of vehicle is between 50HP to 150HP, the price of the vehicle improve slowly. Hence by carefully comparison, the client could purchase a huge-power vehicle with low price, if he chooses a vehicle between the range 50HP and 150HP.


Data Visualization
HighCharts
High Chart Visualization
Stacked Column High Chart
Basic Line High Chart
Area High Chart
The Relationship Between Horsepower and Price
50k
40k
30k
20k
10k
0
Price
0
25
50
75
100
125
150
175
200
Horsepower
The Price of Vehicle
Highcharts.com


The Relationship Between Horsepower and Price
50k
40k
30k
20k
10k
0
Price
0
25
50
75
100
125
150
175
200
Horsepower
The Price of Vehicle
Highcharts.com