

2016 COMP7507 PROJECT

Social Trends of China

By Group 13

Jiang Xinhou 3035347990

Yang Yonggui 3035348633

Yu Qingtian 3035348918

Wang Kaili 3035348114

1. Project objectives:

- 1) Display various aspects (agriculture, industry, tourism and transportation etc.) of China over the past decade to look back the changes of the whole society vividly.
- 2) Design proper visualization diagrams and interaction ways to show the data and its inter-connections effectively.
- 3) Analyze the relations of these aspects and the overall trend of them to draw valuable conclusions.

2. Data source and visualization tools:

The data, specifically the agriculture/industry/environment/tourism/transportation data in the past ten years, is selected from <http://data.stats.gov.cn/> and it focus on the annual data module. This project selects some topics such as production of vegetables and fruits, GDP, visitors flow, industrial distribution and amount of harbors from 2005 to 2014 for visualization and analysis. D3.js, JQuery, Tableau and Google charts were used for visualizations.

3. Labor division:

1). Tourism Data from 2005 to 2014 in China by Jiang Xinhou.

Task: Visualize the tourism data of China (How many foreign tourists come to China over the past decade? How about them and where are they from?)with D3.js and JQuery.

About

The group selected several views of past China including agriculture, industry, tourism and transportation, and here is about the tourism part. In this topic, foreign tourist data visualization was chosen to be done and three typical aspects of that were further picked out, namely the amount of tourists of different age groups, from different countries, and from different areas.

The reason of choosing to do visualization with d3.js is that programming's flexibility makes it possible to achieve creative diagram design and interaction design.

A visualization tool for the data of this part was self-designed and developed. "Tag Cloud" and combined bar and line chart are adopted as the main visualization elements, and the dynamic data interactions are achieved with D3.js transition animation and JQuery event.

Design details

Tag cloud: The index page shows tag clouds of the three selected aspects, the text sizes are corresponding to the ten years average data, the text color contains no information because they are randomly generated.

When the mouse hover the orange topic title, the corresponding tag cloud will be generated and come out with a animation, every time when the tag cloud is regenerated its layout will be updated. Click the index title and the index page will change to another three detailed diagrams.

Coordinate system: In the combined chart, two symmetric spaces above or below the 'x' axis were drawn to make full use of the screen space to display more information. The two 'y' axis is reversed and axisymmetric.

Line chart: The line charts are displayed in the 'above space' (above the x axis), showing the data trends over the past ten years. The line shows all the properties' (age or area) trends by dynamically change from the current property's line chart to another one. Besides there is also a wider and static line in the above space which shows the average data and this line will stay there all the time after first drawing. To keep the monotony on 'y' axis, the 'monotone' function is selected on the line chart and I also draw dots on every key point of the line.

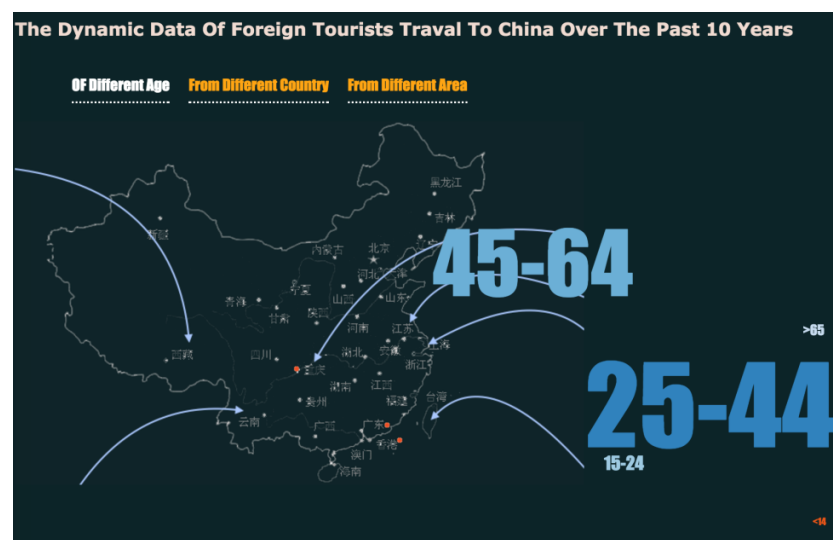
Bar chart: Initially, all the bar charts are draw in the below space, each year's bars are overlapped but will not hide each other totally. The data have been arranged by descending order so the shorter bar will draw on other longer ones. The bars are deployed in that way to make up the missing proportion information. Pie charts of every year are also expected to be shown to make the proportion information clear but failed to be done properly.

Color: Two contrasting colors of the color wheel are selected as the start color and end color to amplify the overall color contrast. The medium colors are interpolated between them. The color also contains amount information of the data.

Interaction: When the mouse hover the legends on top of the chart, the whole diagram's layout will change: the line chart will reshape to show the right data and the corresponding bar chart will move up to the above space, there is also an icon on top-right to show the current data. When the mouse moves out, the bar chart will slide back to the original space. Besides, when the mouse hover the line dots or the bars the corresponding value tip will show up.

Data analysis

Index page:

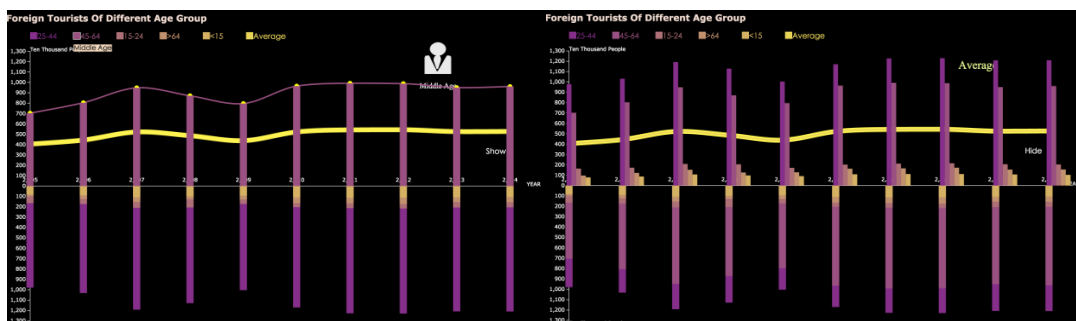
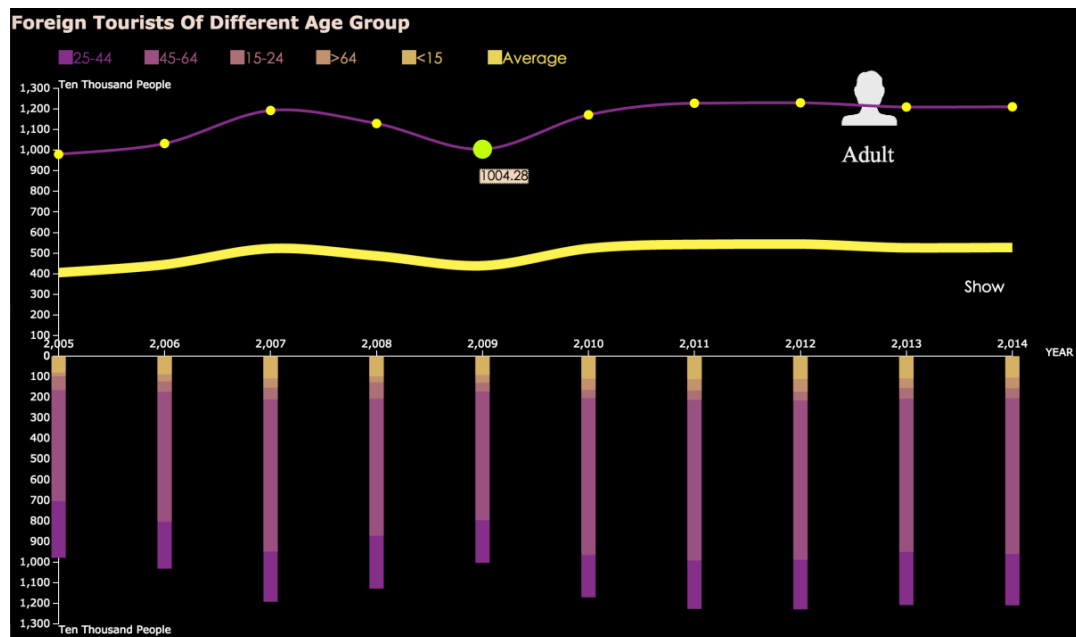


i) The age group tag cloud reflects the average amount of foreign tourists over the past decade, it can be easily found out that people who traveled to China are mainly adults (age from 25 to 44) and middle ages (age from 45 to 64). That is to say, children and old people share a little proportion of the total travelers.



ii) When it comes to travelers' countries and areas, it's clear that people who traveled to china are mainly from Asia, typical countries are Japan, Korea and Russia. A conclusion is that distance is a very important factor that effects people's travel destination selection. That means people tend to travel to a relatively near country because time and money consuming is an important factor that will always be taken into consideration.

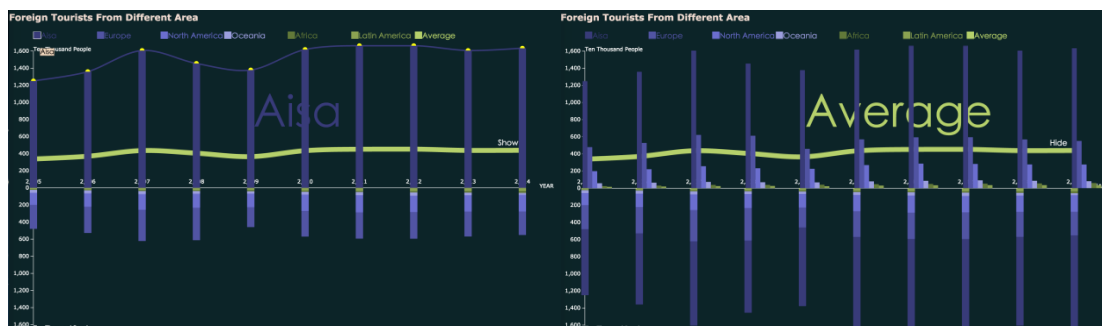
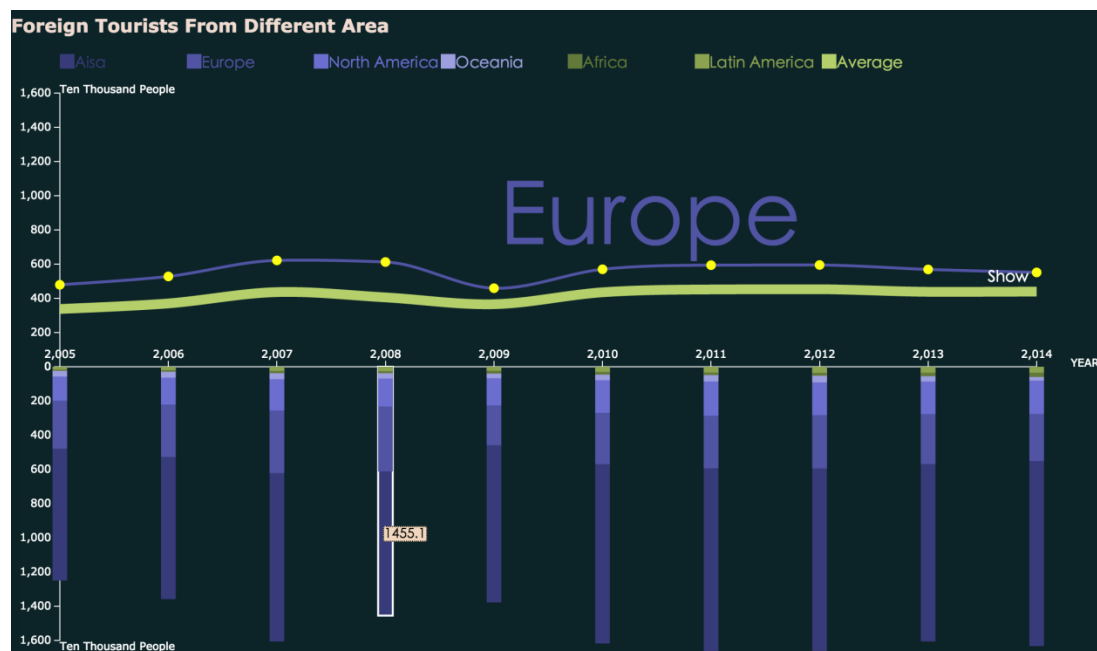
Tourists of different age group: Children(<15), Young(15~24), Adult(25~44), Middle Age(45~64), Elder(>64)



From the diagram of age group data, there is no sharp up and downs form 2005 to 2014. According to the average line, the amount of people traveled to china increased slightly in 2007 and decreased back in 2009 before another slow rise in 2010. Overall the amount is stable with a very slow increasing

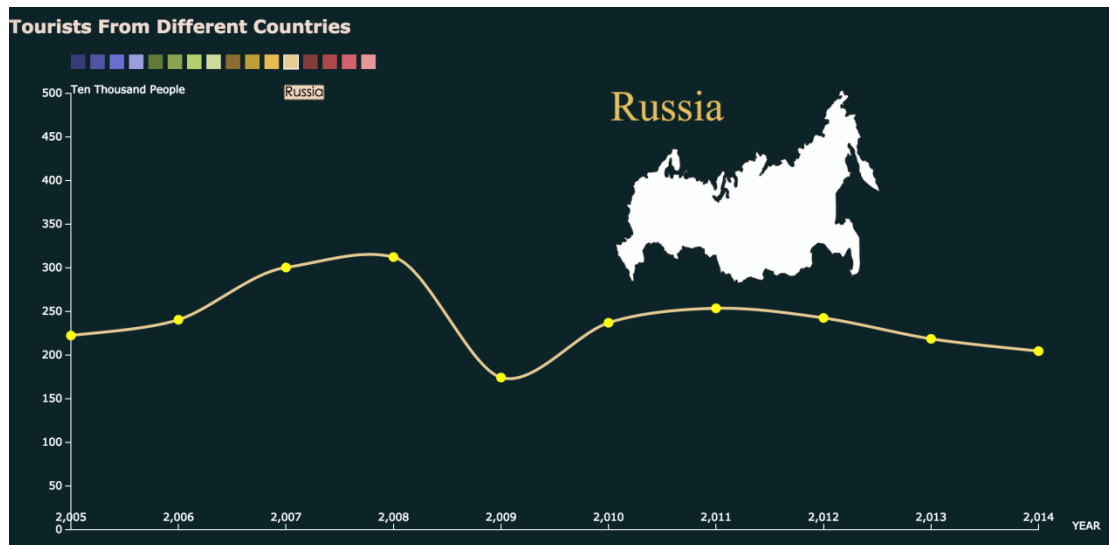
trend. The bar charts can also give us the same conclusion that most foreign travelers are adults and middle ages.

Tourists from different area: Europe, North America, Oceania, Africa, Latin America



This part is similar to part2, using the same diagram. From this chart people can clearly see how many foreign travelers have come to china from these six areas. Most tourists are from Asia mainly because of the distance convenience, and there are also quite a lot of tourists come from Europe, North America these kind of developed areas. The overall amount of tourists is quite stable in the past decade.

Tourist from different countries: Japan, America, Korea, Australia, etc.



In the countries data visualization part, the bar charts are removed. The whole space is used to show the trending data of each country. For most countries, the corresponding maps are displayed on the top-right to help people know which country's data is currently showing now, otherwise a single name of the country will be displayed instead because these countries' map are uneasy to recognize, which is of little help for information giving.

Insight:

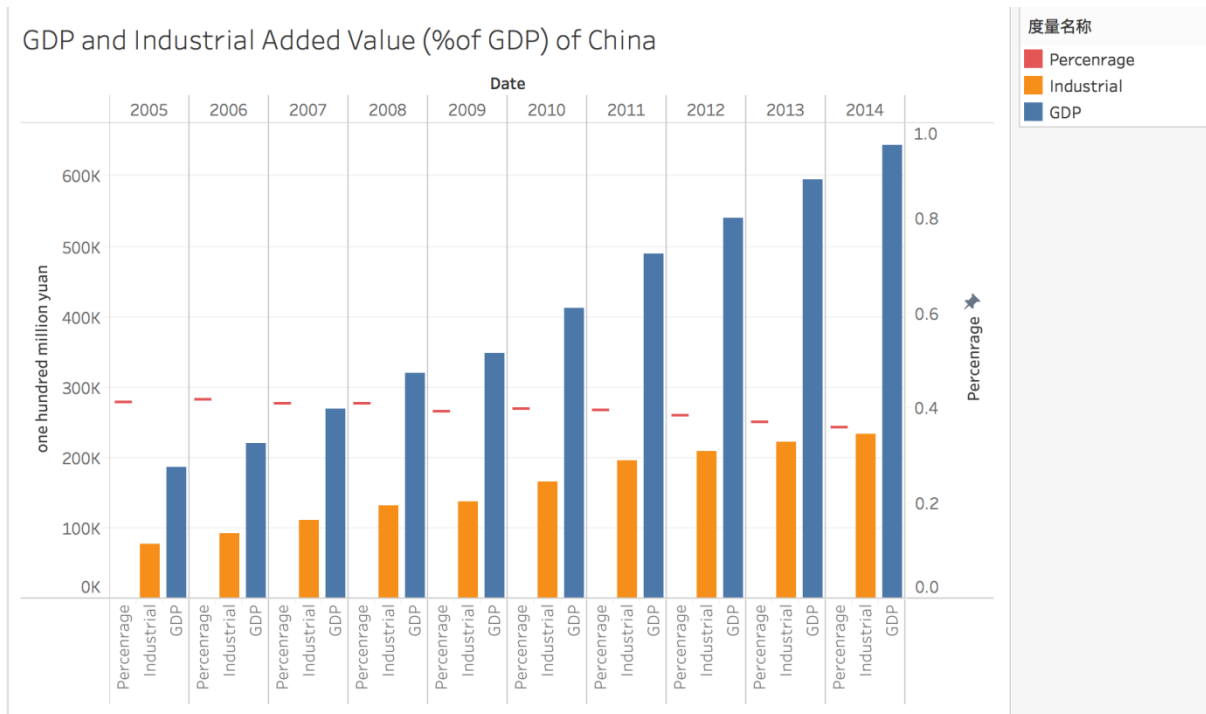
Based on the previous line charts and the final part - each country's tourism data, a common fact is that there was rapidly increasing foreign tourists travel to china in 2007 and 2008, when was just before the 2008 Olympic Game in Beijing.

Difficulties and Limitations:

Visualizing data by programming with d3.js is adopted to be more flexible to achieve the design, but it takes me too much time to learn d3.js that time is very limited to select and analyze data. Some other relevant data sources are also taken into consideration but their data formats are differ greatly and hard to integrate and display them together with d3.js so this part is focus on how to show the current data efficiently and user friendly to interact. The data used is typical but quite limited and not completed.

2).Development of industry of China by Yu Qingtian.

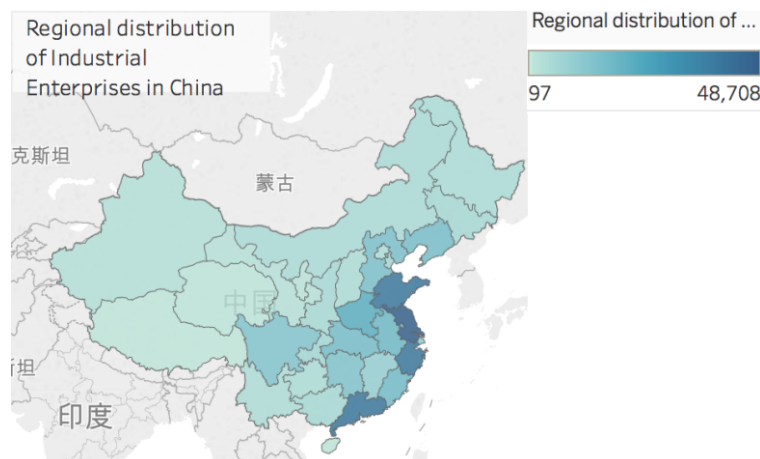
The purpose of my visualizations was to display the development of industry of China in the past. First of all, visualization consisted of bar chart and Gantt bar chart was created to present variation trend of GDP and industrial added value (including the percentage of Industrial added value in GDP) in China as shown in picture 1:



Picture 1: GDP and Industrial Added Value in China.

As can be seen in the picture, GDP of China kept raising from 2005 to 2014. Industrial added value of China increased at the same time but the growth decreased as a whole. This also can be concluded from decline of the percentage industrial added value in GDP. The phenomenon was called De-industrialization due to increase of labor cost, resource depletion and development of technology.

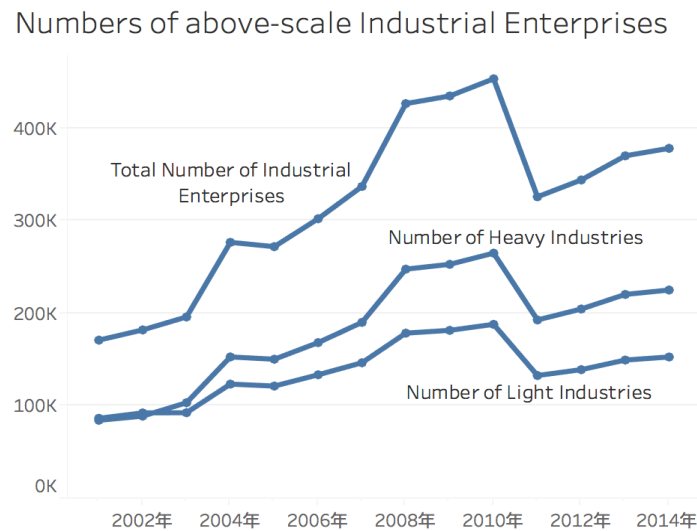
Secondly, I display regional distribution of industrial enterprises in China in 2014 as shown in picture 2 below:



Picture 2: Regional distribution of industrial enterprises in China

It can be concluded that industrial enterprises were mainly distributed in south-east coastal areas such as Jiangsu, Zhejiang and Guangdong. The convenient of transportation contributed to development of industry.

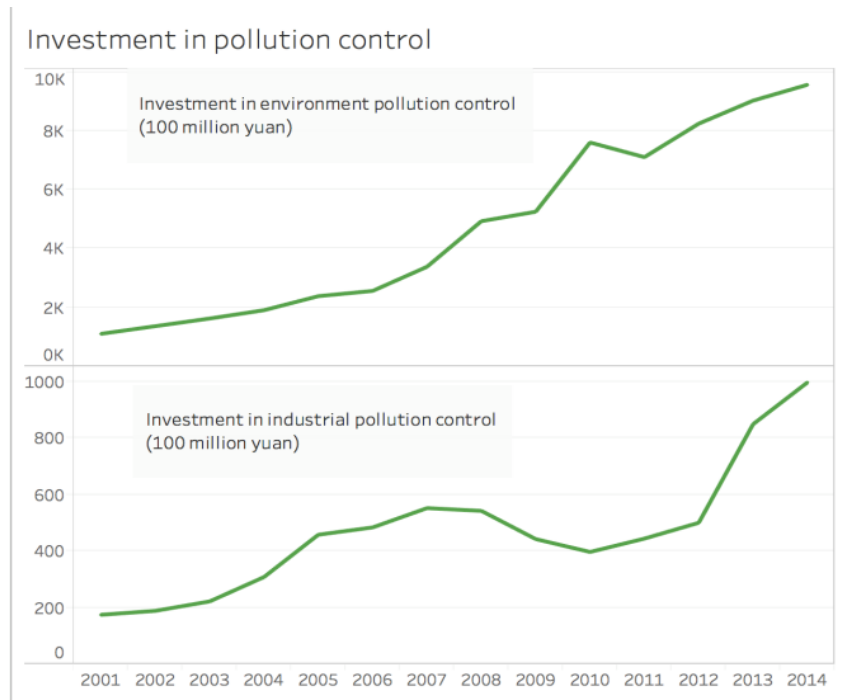
Thirdly, the numbers of above-scale industrial enterprises is displayed in picture 3:



Picture 3: Numbers of above-scale industrial enterprises

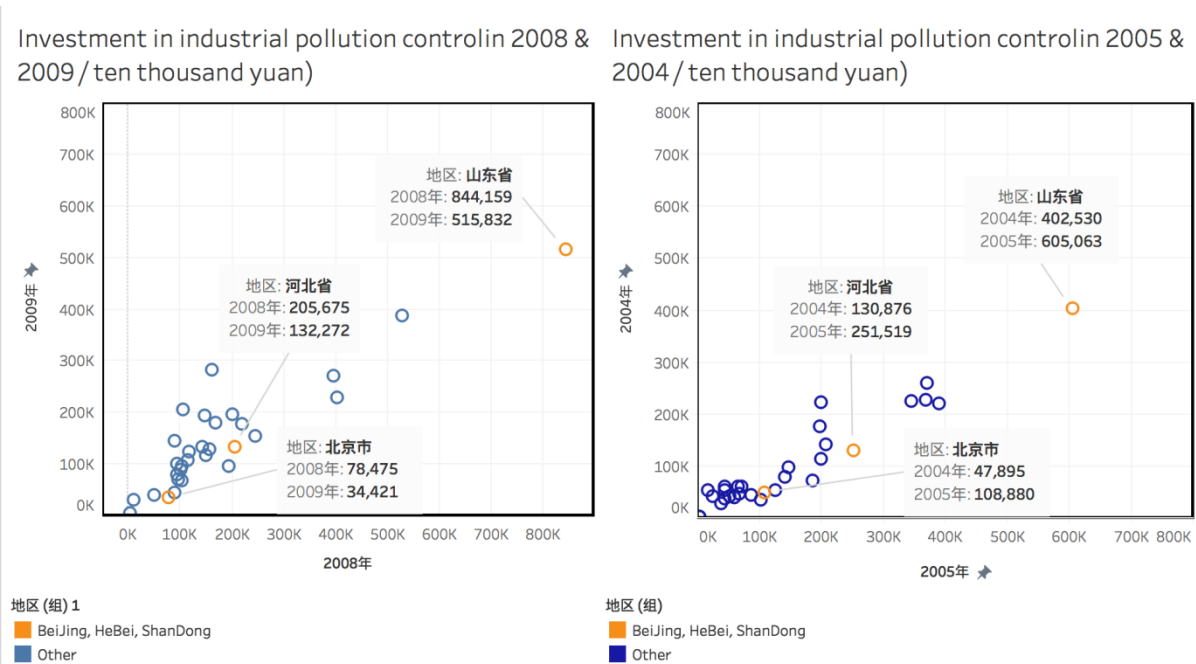
The definition of above-scale was revenue from main operations above 5 million Yuan before 2011, but it was changed to revenue from main operation above 20 million Yuan after 2011. As a result, sharp decline existed between 2010 and 2011. The growth of industrial enterprises slowed down from 2005 to 2008 and even reverted from 2004 to 2005, but it raised sharply from 2008 to 2009 especially the growth of heavy industries. This was mainly because the government limited the establishment of industrial enterprises with heavy pollution and shut down pollution exceeded enterprises after 2005 for preparation of 2008 Beijing Olympic Games. Then the policy was loosened after 2008.

The impacts of Beijing Olympic Games can also be seen from the line chart of investment in environment pollution and industrial pollution control below:



Picture 4: Investment in pollution control

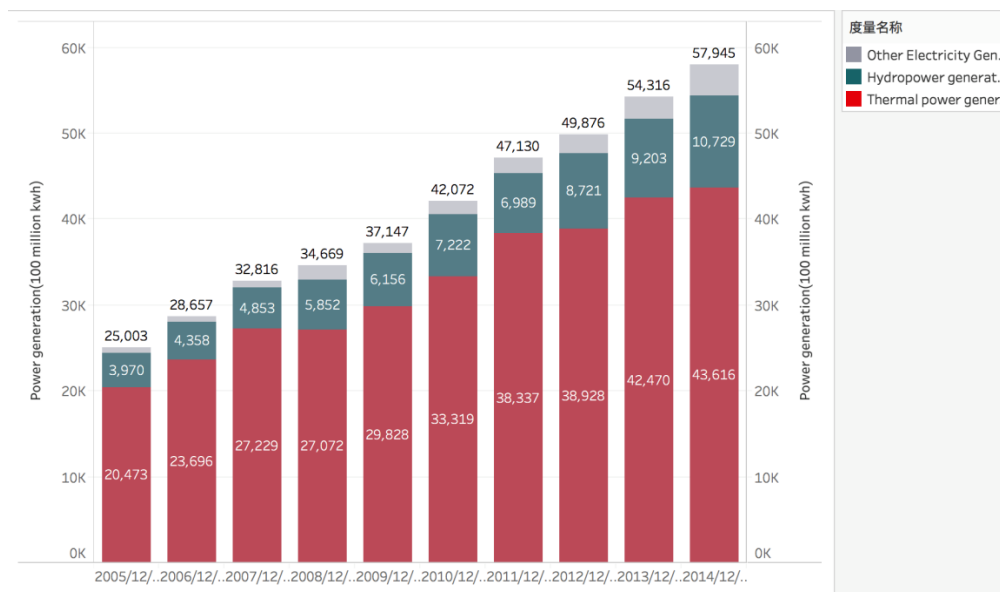
The total investment in environment control was increased as a whole, but the investment in industrial pollution control rose rapidly from 2001 to 2005 and kept a high level from 2005 to 2008. Then it fall down after 2008 due to the end of Beijing Olympic Games in August. For deeper insight, scatter plots of comparison between investment in industrial pollution control in 2005 and 2004 and comparison between investment in industrial pollution control in 2008 and 2009 were shown below:



Picture 5: Scatter plots of comparison between investments in industrial pollution control

It can be concluded that investment in industrial pollution control in the most provinces increased sharply from 2004 to 2005 due to the preparation of 2008 Beijing Olympic Games. Then it decreased from 2008 to 2009 due to the end and the continued ramifications of high level investment before. The above phenomenon was shown obviously in Beijing (the host of 2008 Olympic Games), Hebei and Shandong (close to Beijing).

The development of industry relies on electricity generation, thus a bar chart of power generation in China from 2005 to 2014 was created:



Picture 6: Power generation in China

As can be seen in picture 6, electricity generation kept going up from 2005 to 2014. The main method was thermal power generation but the proportion of it in electricity generation decreased in total. Moreover, the ratio of Hydropower generation and other method such as nuclear power and wind power generation increased, which were more environmental friendly.

New Insights: The impacts of 2008 Beijing Olympic Games on development of Industry of China were deeply analyzed. It contributed to standardized development and environmental protection.

Difficulties: There were some difficulties during the visualization. For instance, combining data in the same format and showing it with the same dimensions. However, these were solved by creating different worksheets and import them into one work package.

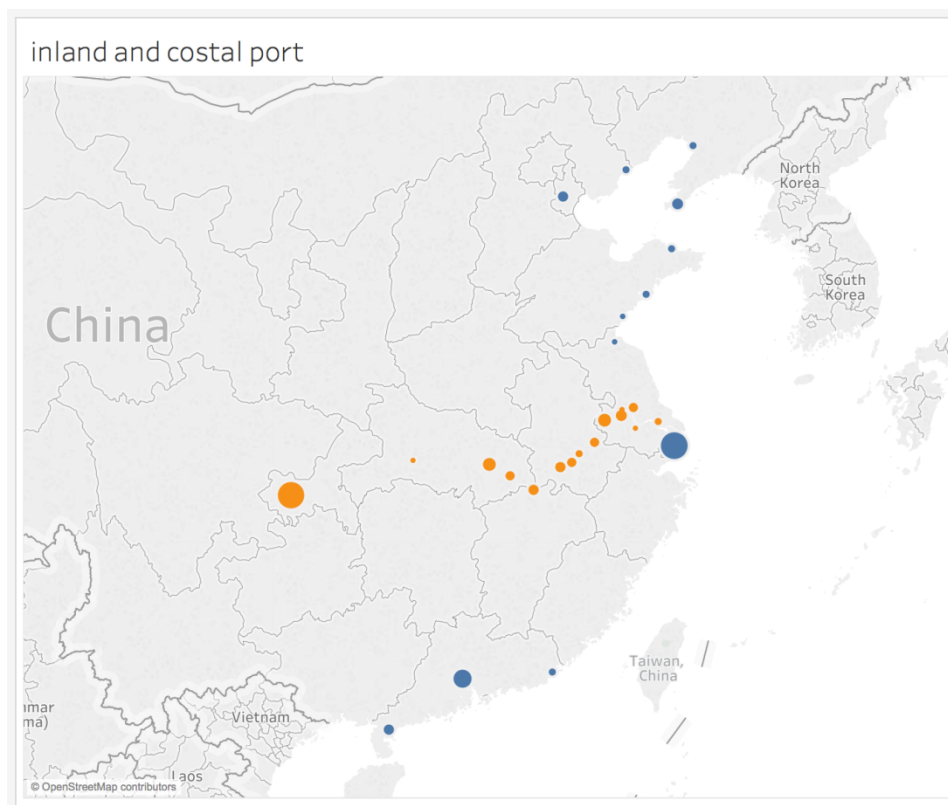
Methods and Justification: Various methods were tried such as bar chart, line chart, Gantt chart, Scatter chart and map. Proper methods were selected to display different aspects of data. For example, Gantt chart in picture 1 was used to show percentage distinguished from valves using bar chart; Map in picture 2 was used to display regional distribution; line charts in picture 3 and 4 were used to show variation trend; scatter charts in picture 5 was used to

display comparison of large amount of data with two metrics; stack diagram in picture 6 was used to present number and proportion.

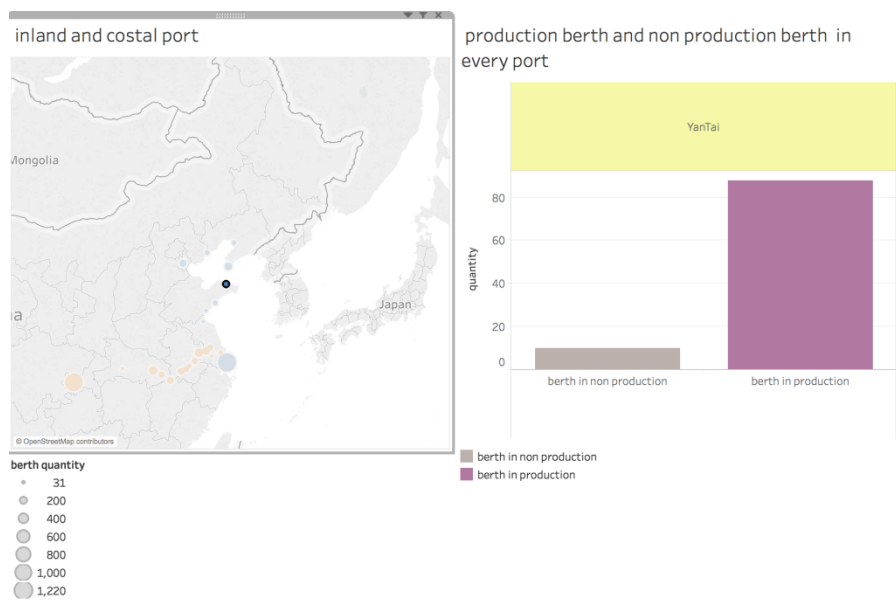
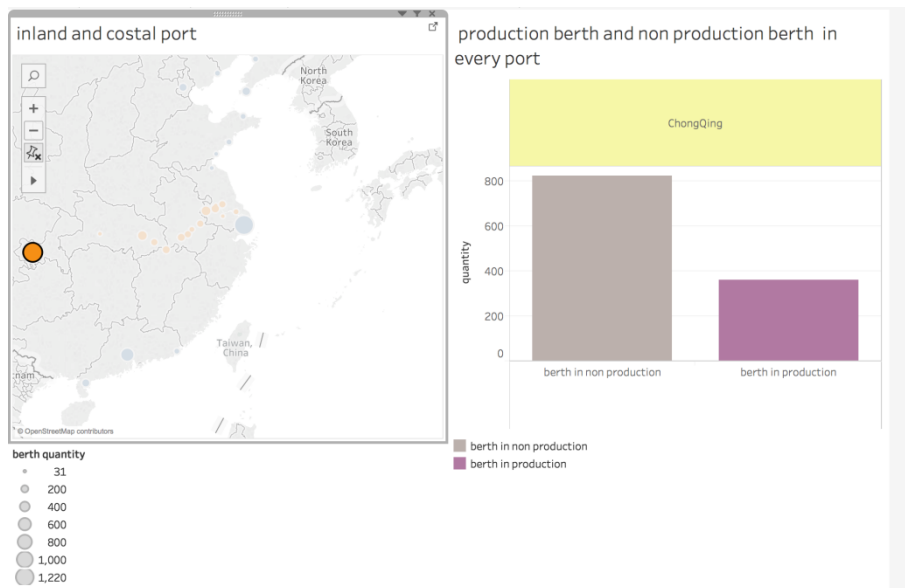
Undone Plans and Limitations: The changes of regional distributions were planned to be displayed, but the data was too much and proper method could not be found. The Tableau was used for visualization, however, it was inconvenient to show united data and difficult to combine various charts in one diagram.

3). Transportation (specifically harbors) Analysis by Wang Kaili.

The quantity of inland port berths and coastal port berths in 2005-2014 and the average number of inland and coastal cities in 2009-2014 were used to do the visualization.



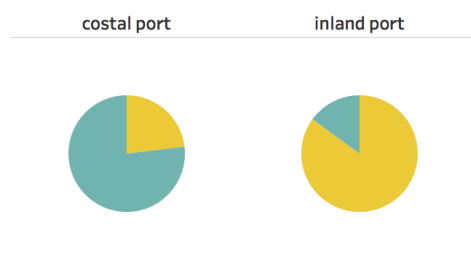
First of all, I use the average data in 2009-2014 to do a horizontal analysis. I set blue color for costal port and orange for inland port. And the circle size represent the quantity of berth (including the production and non- production berth) in that port. When the button clicked on the location of that port, another bar chart will show the quantity of production port and non-production port. And I find when clicking the inland port, the quantity of non-production berth is larger than the quantity of production port. And the quantity of production berth is larger than that of non- production berth in coastal port.



And the pie chart shows the same law.

Non-production ports account for a high proportion in inland port, production port accounted for a high proportion in coastal city.

different berth quantity in costal and inland port

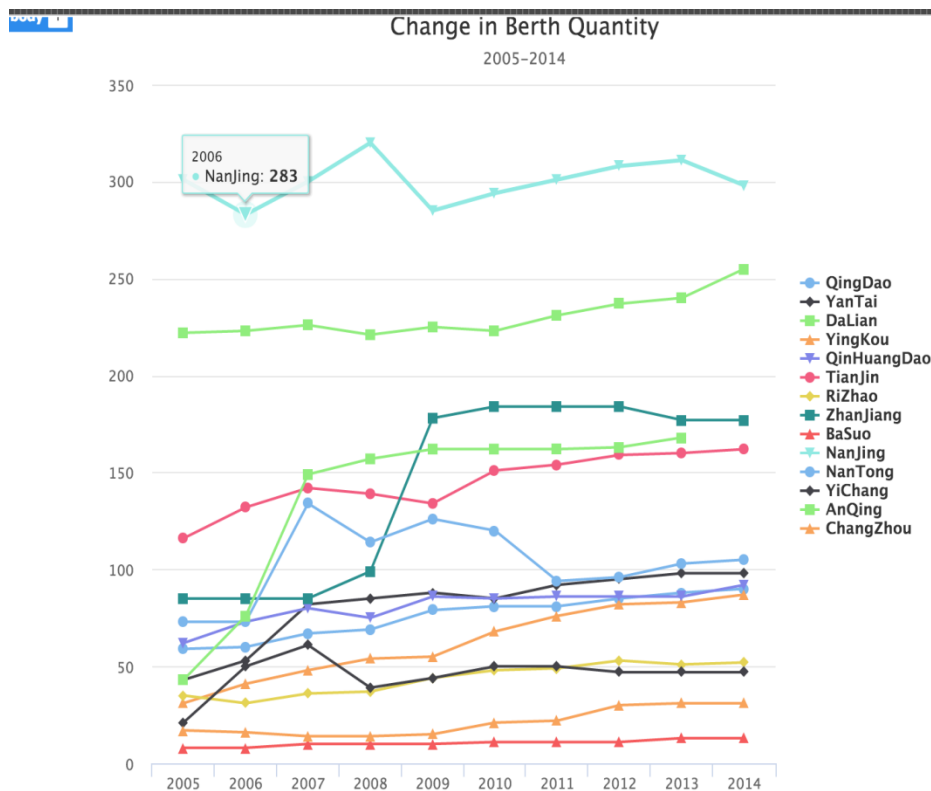


Category

berth in non producti..

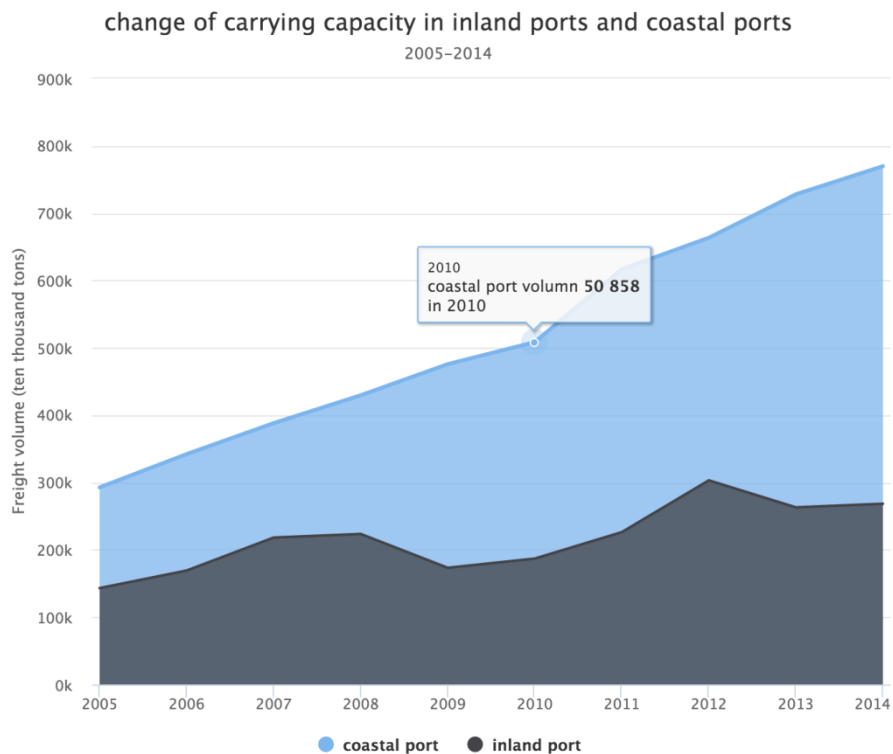
berth in production

And I use high-chart char set to show the change of quantity of berth in these ports, I find that almost every port shows a rising trend in these 10 years.

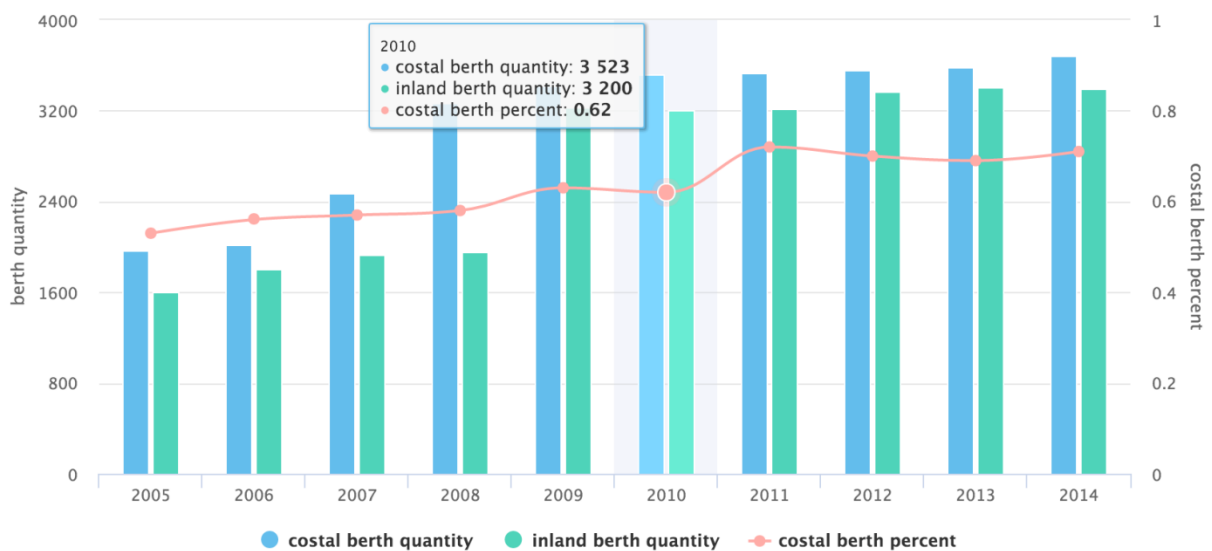


I also do the visualization of the change of volume of carrying capacity in inland and coastal ports.

I find that both in inland and costal port, the volume of carrying capacity is rising. And the volume of carrying capacity in costal port is almost twice of that in inland port. The volume of carrying capacity in inland port is higher than it following years in 2007 and 2008. The tourist transportation during 008 Olympic may cause the increase.



And I contrast the costal berth quantity and inland berth quantity, the line shows the costal berth percent. It shows that the percent of costal berth percent is around 50 percent and it shows a rising trend. Both the inland berth quantity and the costal berth quantity increase with time.



So the conclusion is inland ports are mainly for non-production use, like tourist transportation. And the costal ports are mainly for production use, such as transporting the production goods.

And the quantity of berth is show a rising trend, with the increase of GDP and the development of productive force. The quantity of berth used for production increases faster.

I think the amount of berth will increase and the percent of production berth will also be high in the coming years.

I use tableau and the high-chart char set to accomplish the graph.

What I think is poor is that my dataset is single. And the graph cannot integrate well.

I wanted to show the amount of production berth and non-production berth in every port in every year. But I find it is too cumbersome and it does not help the visualization and conclusion.

4). Agriculture Analysis by Yang Yonggui.

The visualization that I have done include 5 parts:

The plastic usage and poison usage in Beijing, Shanghai, Gansu, Zhejiang and Heilongjiang from 1995 to 2014;

The distribution of per capita disposable personal income in China;

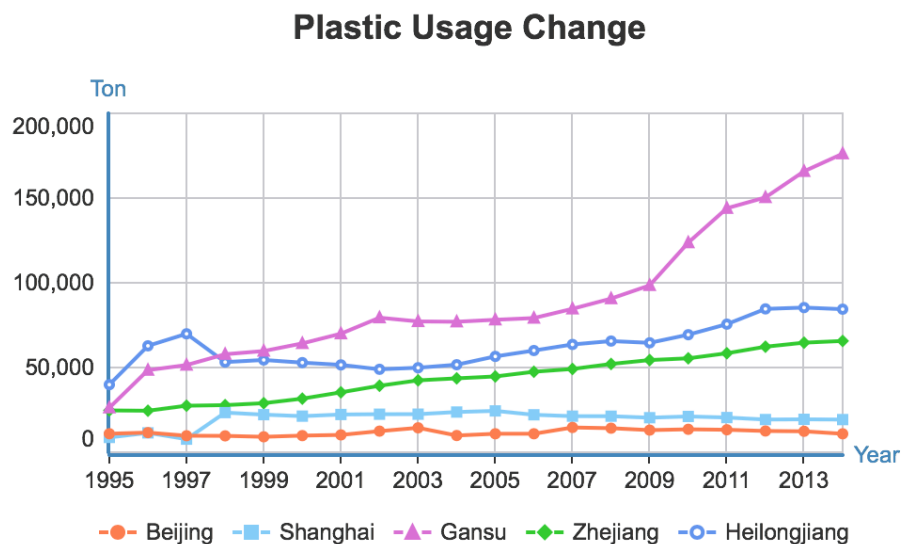
The agriculture gross output in Beijing and Shanghai from 1995 to 2014;

The agriculture gross output in Gansu and Heilongjiang from 1995 to 2014;

The primary industry structure in Beijing.

I used line chart, bar chart, stacked bar chart and graph chart to show the usage trend of certain objective.

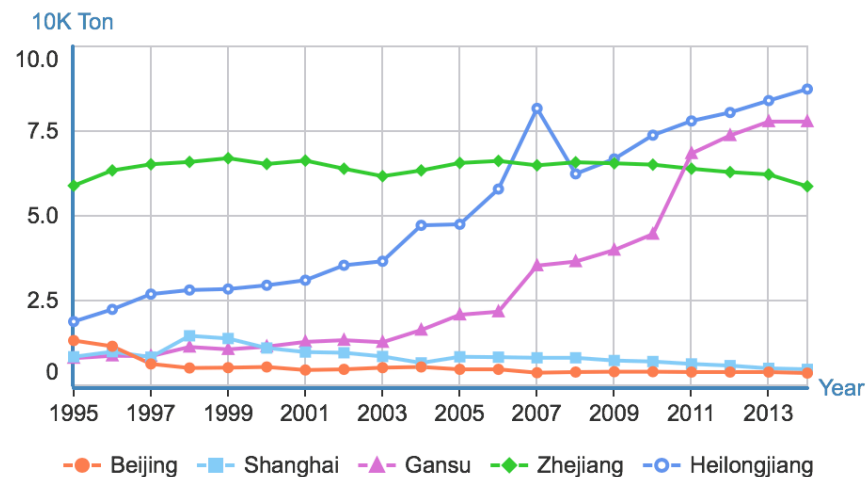
1. Identify the relationship between economy and agriculture structure.



The first picture above shows the plastic usage change from 1995 to 2014. The purpose is to show how the amount of environment-unfriendly material changes together with the development of agriculture economy. Each different color represents a city or province and the line chart shows the plastic usage change from 1995 to 2014. I chose 5 different provinces

and noticed that the more developed the city is, the less plastic it relies on. So I want to choose another different environment-unfriendly material and show how it changed during the same time.

Poison Usage Change



This picture above shows the poison usage change of the same 5 provinces in the first picture. Different color represents a different city or province. We can easily observe from the chart that although Zhejiang province used a large amount of poison as always, the figure stayed stable. Beijing and Shanghai also remained unchanged even drop down a little. By contrast, the figure of Gansu and Heilongjiang increased dramatically since 2003. Then I wonder how the economy of these provinces distributes.



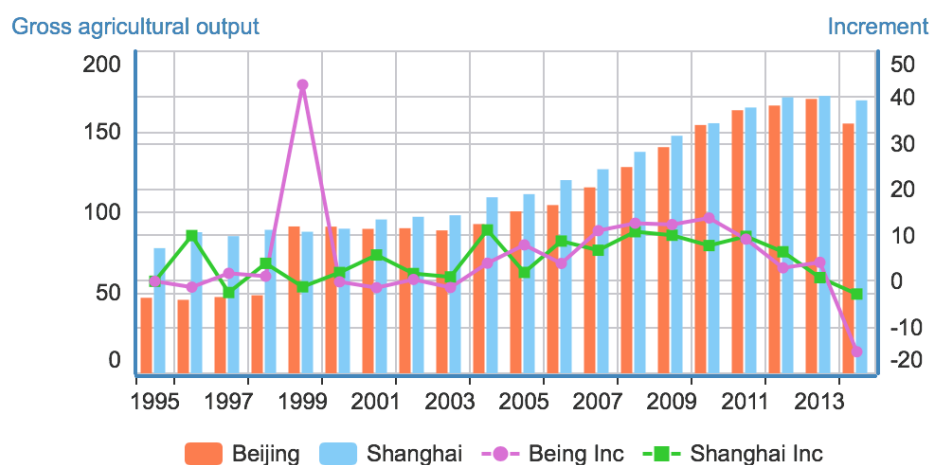
The picture above illustrates the per capita disposable personal income based on a Chinese map. The color shows the income difference between different provinces and cities. Income becomes higher when the color becomes deeper. And I get the information I want. Zhejiang,

Shanghai and Beijing are all very developed city/province, but Heilongjiang and Gansu both have a comparably low income level. So by the visualization above it seems that as the development of the economy, people will use less and less environment-unfriendly material in agriculture.

2. Identify the role of agriculture in different economy structure.

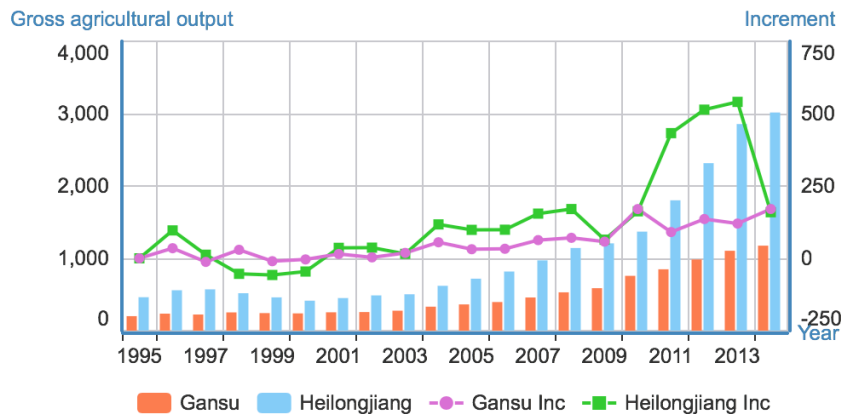
In my opinion, development determines the economy structure. So I decide to dig much deeper into the agriculture development of the most developed cities in China which are Beijing and Shanghai.

Shanghai VS Beijing Agriculture Development



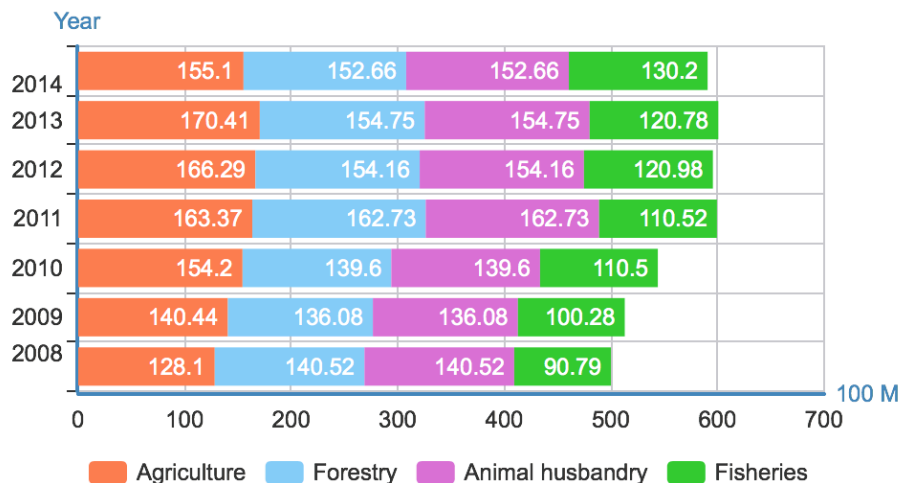
As you can see from the visualization showed above. I visualized the gross agriculture output of Beijing and Shanghai. The bar chart represents the gross agriculture output and the line chart shows the increment every year compared to the former year. We can easily observed that the gross agricultural output of Beijing and Shanghai both gradually increase but at a slower and slower pace. In the end, the increment become negative, which means the agriculture becomes a less and less important part of the cities' economy.

Gansu & Heilongjiang Agriculture Development



By contrast, I visualize the data of Gansu and Heilongjiang, the visualization result is showed above. Each bar represents the gross output of Gansu and Heilongjiang while each line shows the increment of the output compared to the former year. I find that both of these two provinces' gross agriculture output are increasing steadily. This means that the agriculture development of these two cities are still in an upward trend. So in some provinces that are not as developed as Beijing and Shanghai, agriculture seems to play a much more important role in local economy.

Beijing Agriculture Proportion Change



This picture above shows the different categories of primary industry in Beijing. This is a stacked bar chart and each bar represents the gross output of a certain year. Different color represents a different category in primary industry. As we can easily observe that three out of four categories remains stable in gross output and it can reflect my conclusion above. What's interesting is that Fishery seems to develop well in Beijing.

Conclusions:

i) The environment-unfriendly material usage in different provinces or cities has a different trend. For those developed city, the usage of environment-unfriendly material becomes less and less. While in some province whose economy performance is not that good. There is still a upward trend in usage of these materials;

ii)Agriculture plays a less and less important role in the city which has a fantastic economy status. However, for those cities whose economy is not that good, agriculture still plays a more and more important role in the local economy;

iii)Beijing's agriculture development remains stable these years as it is one of the most developed cities in China. However, what's interesting is that fishery is still developing fast these years.