Research on China Air Quality

Haoxun Zhao

1. Motivation

This map provides information on particulate matter air pollution less than 2.5 microns in diameter (PM2.5). Under typical conditions, PM2.5 is the most damaging form of air pollution likely to be present, contributing to heart disease, stroke, lung cancer, respiratory infections, and other diseases.

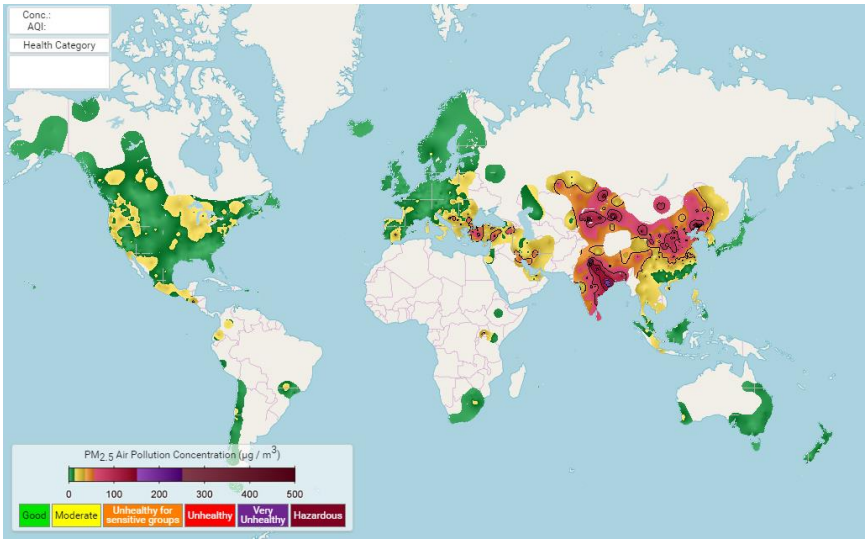

Figure 1: Global PM2.5 Concentration
http://berkeleyearth.org/air-quality-real-time-map/

As we can see from the map, China is one of the most seriously polluted country in the world. From 2013 to 2017, I spent 4 years to get my bachelor's degree in Beijing which is the darkest central area on the map in China. During these four years, I found that air was mostly polluted ever. People can only see the blue sky for less than 10 days in a whole year. They are surrounded by all kinds of pollutant which is like a dome blocking everyone. In China, 500,000 people die every year at a very young age directly due to the air pollution. As above, it is urgent for people to investigate and find ways to control and clean our air. The goal of this project is finding what challenge we are facing.

2. Data Source

There are three main data sources in this project

1. http://aqi.cga.harvard.edu/china/cumulative/.This data includes air pollutant concentration of most of the cities in China through Feb. 2014 to Feb. 2016. There are 2542745 records in it. Important variables include 'stationname-str', 'latitude-float', 'longitude-float', 'pm25-str', 'pm10-str', 'o3-str', 'no2-str', 'so2-str', 'co-str'. The last 6 variables represent the concentration of corresponding pollutant at a specific time.
2. http://www.stateair.net/web/historical/1/1.html. This data includes hourly PM2.5

concentration in Beijing through Apr. 2008 to July 2017. There are 79559 records in it. Important variables include 'Data-str', 'Value-float'. Value means the concentration of PM2.5 in a specific time in Beijing.

3. http://www.stats.gov.cn/tjgz/wzlj/dftjwz/ This data is provided by Chinese government and include GDP of each capital city in China from 2013 to 2017. There are 36 records represent 36 capital cities in total.

3. Analysis and Results

3.1 Is there any relationship between each kind of pollutant and GDP?

Figure 2 shows that PM2.5 and PM10 are highly linearly related. They also have linear relationship with NO2 and SO2. It is also clear that there is no obvious linear relationship between air pollution level and GDP.

PM2.5 and PM10 are Particle Air Pollution with diameter less than 2.5um and 10um respectively. They have the same sources including incomplete combustion, automobile emissions, dust, cooking, so it is reasonable that they have high linear relationship. On the other hand, GDP can be affected by many different factors. Some cities like Taiyuan and Zhengzhou are mainly depend on mining industry, some cities like Beijing and Shijiazhuang depend on iron industry. Air quality in these cities are the lowest. However, some costal cities like Shanghai and Guangzhou depend on trade, so their air quality are relatively higher.

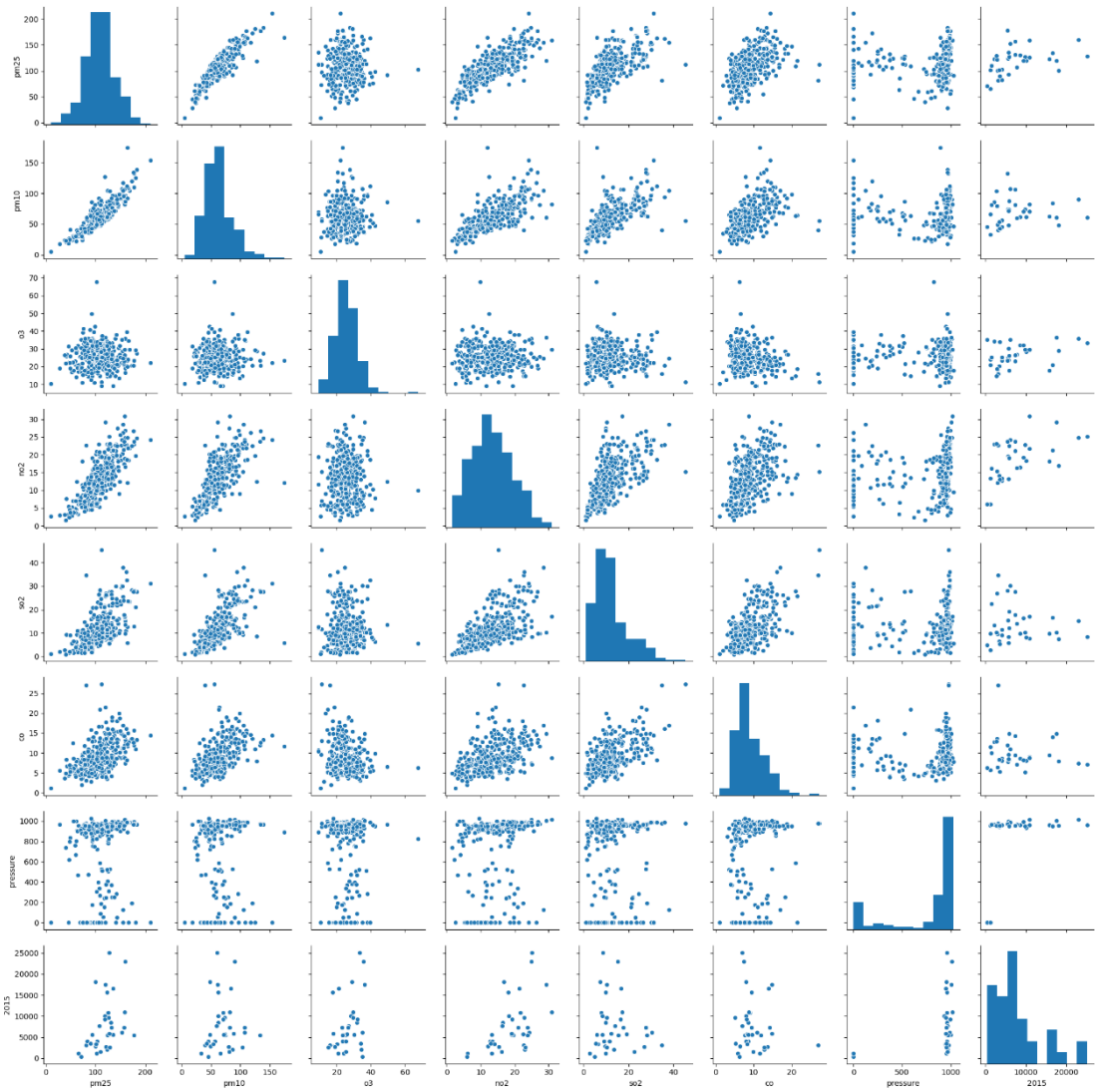

Figure 2. Pairplot of all kinds of pollutant and GDP

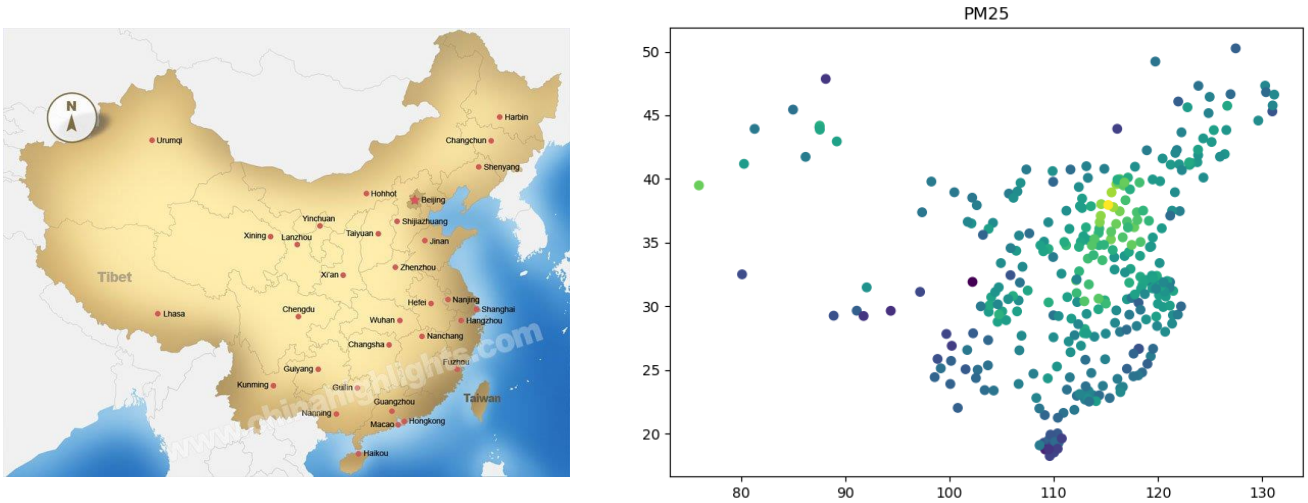

Figure 3. Geographic distribution of PM2.5. Yellow: High polluted, Blue: Low polluted

4.2 How is the pollutant level changed geographically?

Figure 3 is a Chinese map and a PM2.5 geographic distribution. It is clear that the most polluted point is Beijing. PM2.5 mainly come from incomplete combustion of fossil fuels. It is reported that there are 21 million people living in Beijing and 6 million cars. These cars produce lots of particles and exhaust into air. There are more than 2000 iron factories around Beijing, every year 380 million tons of coals are consumed here. Lack of advanced clean and control technique, lots of particles are exhausted into air. At the same time, Beijing is surrounded by high mountains at north, east and west. These mountains keep any wind away from north and east, so the polluted air cannot flow away. All these factors make the Beijing's air so poor. In contrast, Haikou is the city at the bottom of China. It's economy mainly depends on tourist industry. There is fewer people and fewer factories, so Haikou's air quality is much higher than Beijing.

4.3 Are there obvious differences between leading cities and other cities in terms of pollutant level?

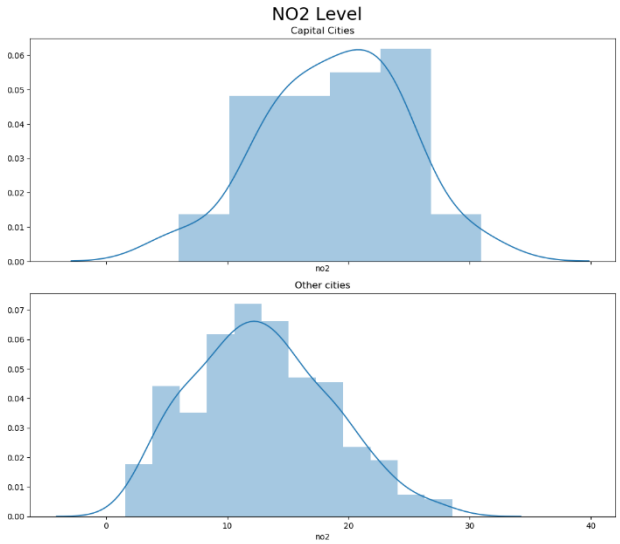

Figure 4 NO2 level in Capital Cities and Other Cities

From t-test and distribution plot, discrepancy of mean pollutant level between capital cities and other cities is ranked in descent order as: NO2 > PM2.5 > PM10 > SO2 > CO > O3. NO2 is mainly come from exhaust emission of vehicles. In 2017, there are totally 217 million cars in China. Nearly 90 million cars are running in these capital cities. So many cars squeezing in cities results in traffic jam. When cars running with low speed, they will produce NO2 as 2 times as normal.

4.4 How does the pollutant change during different time scale in Beijing?

There are many outputs in Notebook. Figure 5 shows that the pollutant level reach top from 20 to 3 o'clock. There are so many people living together in a city. They need all kinds of supplies such as foods, materials, fuels. They also produce lots of trash. All of these are carried in and out by heavy trucks at night because most cities have policies that trucks cannot enter during daytime. At the same time, China's diesel standard is much lower than developed country. The diesel is not clean enough.

These factors make a truck produce pollutant more than 10000 times as much as a car. This is the reason why the air gets most polluted when most people are sleeping.

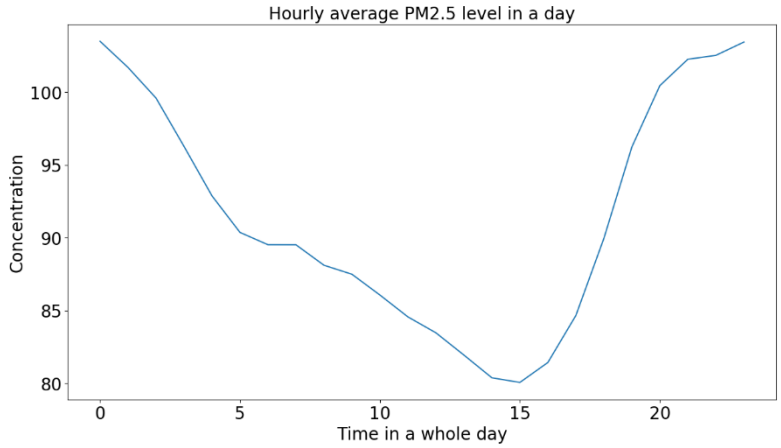

Figure 5 Hourly average PM2.5 level in a day

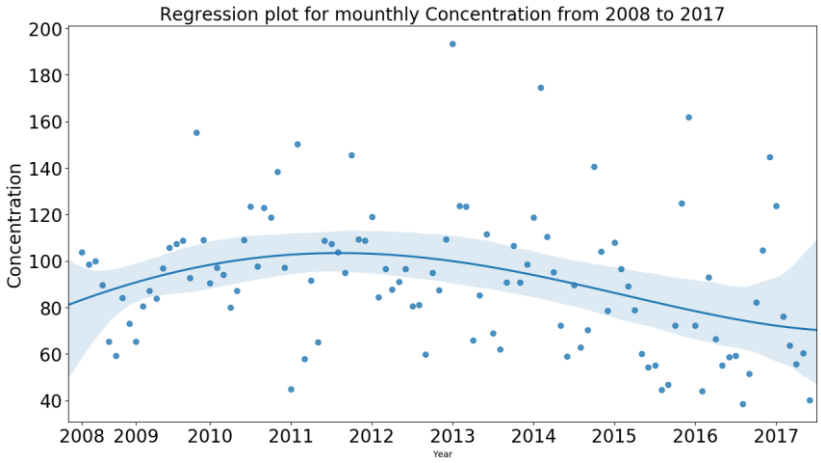

Figure 6 Regression plot monthly concentration from 2008 to 2017

From Figure 6 we can see that air pollution goes up from 2008 to 2012 and goes down from 2012 to 2017. Chinese government has made a series of laws to control the pollutant. More and more people are living in a healthy style. Government highly recommends people take public transportations instead of private cars. Thousands of iron factories are closed in these years because of violating the environment law and excess capacity.

At last, what I want to say is that, environment protection is not responsibilities of a certain amount of people, it is about the whole human beings. There are still so many people focusing on economy development at the cost of environment. I hope we can work together to make change happen.