U.S. Pollution Data

Peng Yan
Zhixing Chang

Xiaoyang Dou

Tianlun Zhao

Description

- Name of our dataset:

    - "US pollution"
- Dealing with four main kinds of pollutions in the US
- Attributes containing following attribute for each pollutants

    - mean

    - 1st max value

    - 1st max hours

    - AQI (Air Quality Index, lower is better)
- From 2000 - 2016

Questions sought to answer

1. What are the main pollutants in the U.S.?
2. Does the overall level of environmental pollution in the United States show an upward trend or a downward trend?
3. Colorado compares to other developed states(New York, Florida, and California), the pollution is more serious or lighter?
4. What causes the increase or decrease of pollutants?
5. How to solve air pollution?

Data Preparation work
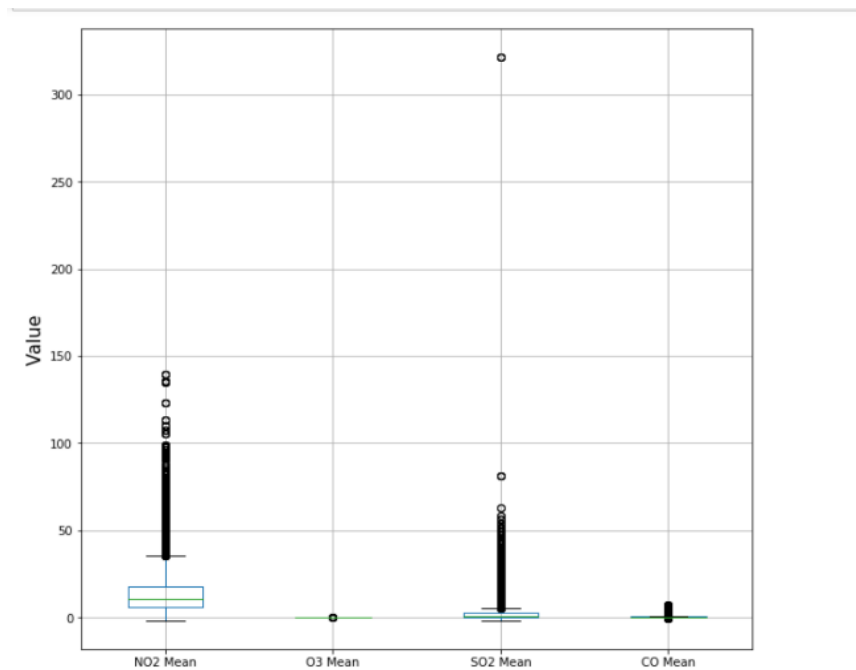
U.S. Pollution Data, Pollution in the U.S. since 2000

https://www.kaggle.com/sogun3/uspollution

Tools used

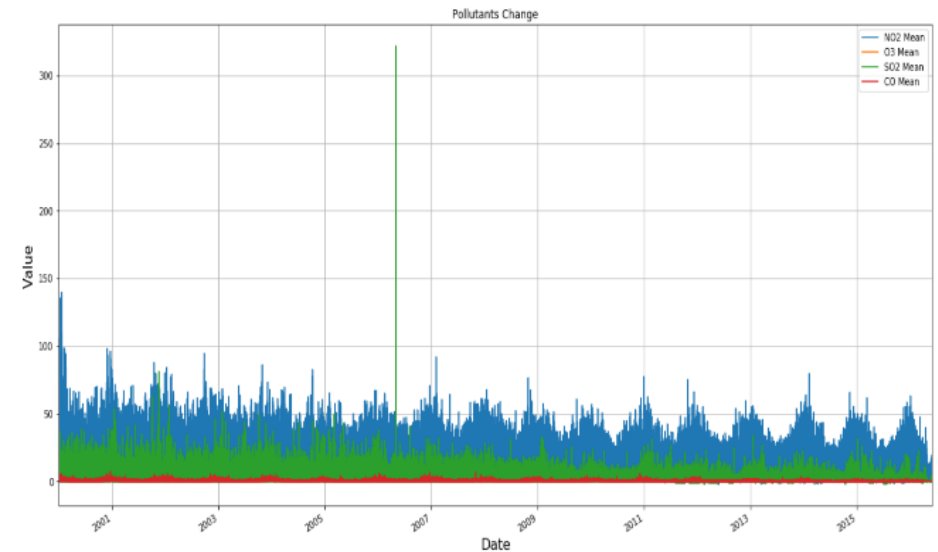- Pyhon3
  - Pandas
  - Numpy
  - Matplotlib

- D3
- Excel

MAIN TECHNIQUES APPLIED

Data Cleaning and Preprocessing

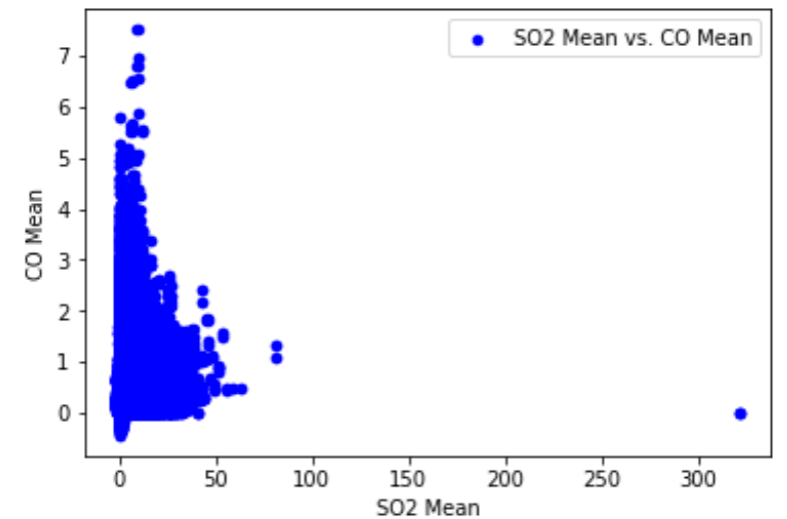- Delete unused columns
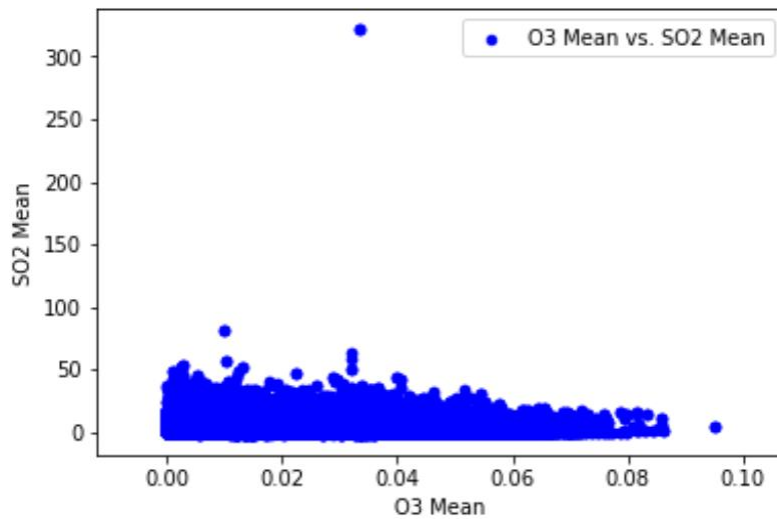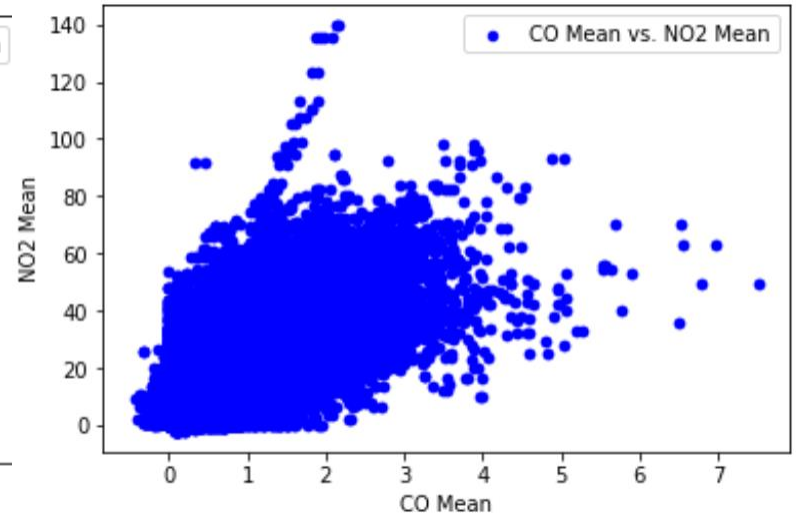
- Delete zeros and NaNs

Correlation
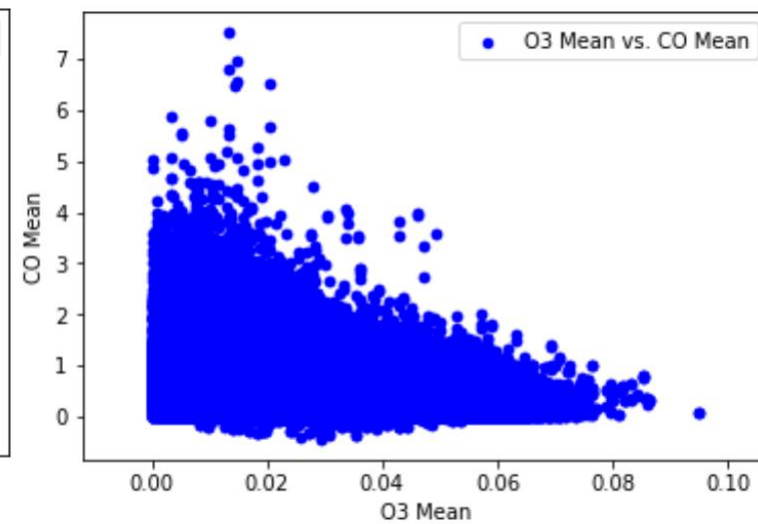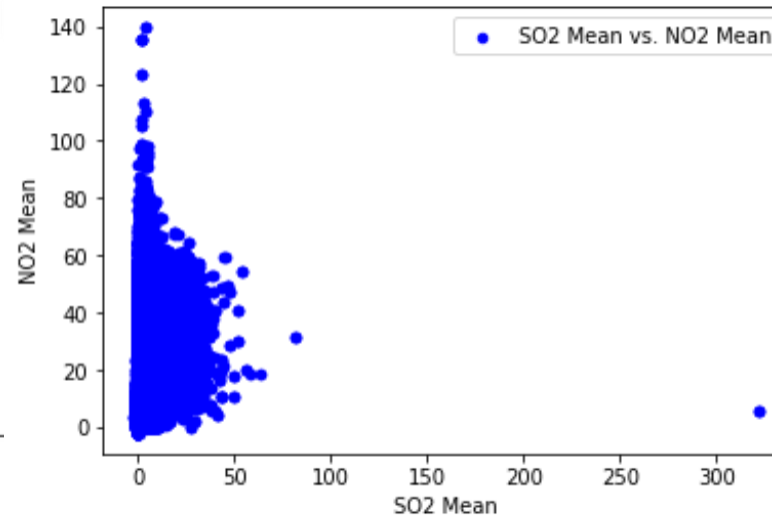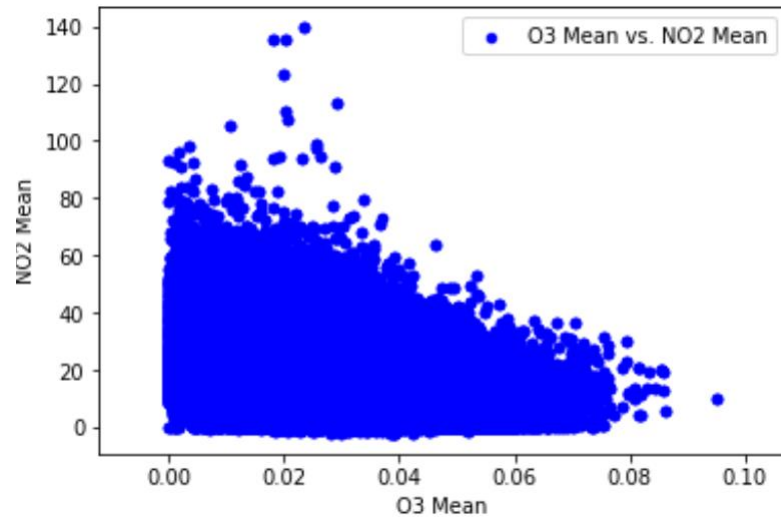
Data Transformation

- Min-max Normalization

Max value:

```
State                      Arizona
Date Local     2000-01-19 00:00:00
NO2 Mean                    139.542
Name: 1468, dtype: object
State                      Arizona
Date Local         2000-01-10 00:00:00
NO2 1st Max Value              267
Name: 1432, dtype: object
State                      Arizona
Date Local     2000-01-10 00:00:00
NO2 AQI                        132
Name: 1432, dtype: object
```

```
State                  Pennsylvania
Date Local     2000-06-10 00:00:00
O3 Mean                    0.095083
Name: 77152, dtype: object
State              Country Of Mexico
Date Local       2007-07-04 00:00:00
O3 1st Max Value            0.141
Name: 744381, dtype: object
State                   California
Date Local     2013-06-29 00:00:00
O3 AQI                        218
Name: 1364788, dtype: object
```

```
State                   California
Date Local     2000-12-20 00:00:00
CO Mean                    7.50833
Name: 10505, dtype: object
State                   California
Date Local     2000-12-20 00:00:00
CO 1st Max Value              19.9
Name: 10504, dtype: object
State                   California
Date Local     2000-12-20 00:00:00
CO AQI                        201
Name: 10505, dtype: object
```

```
State                     Oklahoma
Date Local     2006-05-04 00:00:00
SO2 Mean                   321.625
Name: 596159, dtype: object
State                     Oklahoma
Date Local     2006-05-04 00:00:00
SO2 1st Max Value             351
Name: 596159, dtype: object
State                     Illinois
Date Local     2002-02-14 00:00:00
SO2 AQI                       200
Name: 229463, dtype: object
```

Min value:

```
State   Pennsylvania
Date Local   2012-12-11 00:00:00
NO2 Mean   9.1375
Name: 1299522, dtype: object
State   Arizona
Date Local   2000-05-16 00:00:00
NO2 AQI   1
Name: 528, dtype: object
```

```
State   California
Date Local   2007-02-05 00:00:00
O3 Mean   0.013792
Name: 636115, dtype: object
State   California
Date Local   2000-12-07 00:00:00
O3 AQI   8
Name: 12030, dtype: object
```

```
State   Kansas
Date Local   2015-11-02 00:00:00
SO2 Mean   0.591667
Name: 1663965, dtype: object
State   Arizona
Date Local   2000-03-17 00:00:00
SO2 AQI   NaN
Name: 306, dtype: object
```
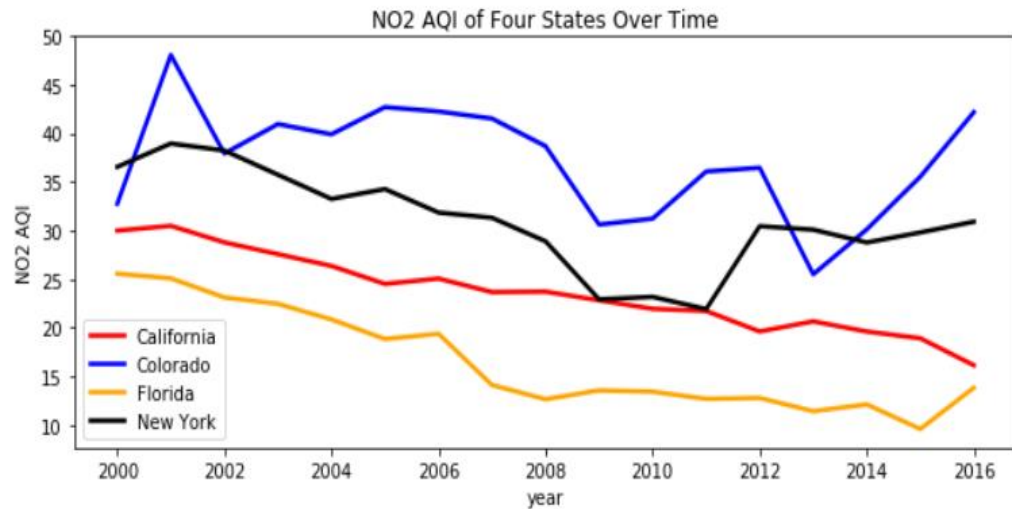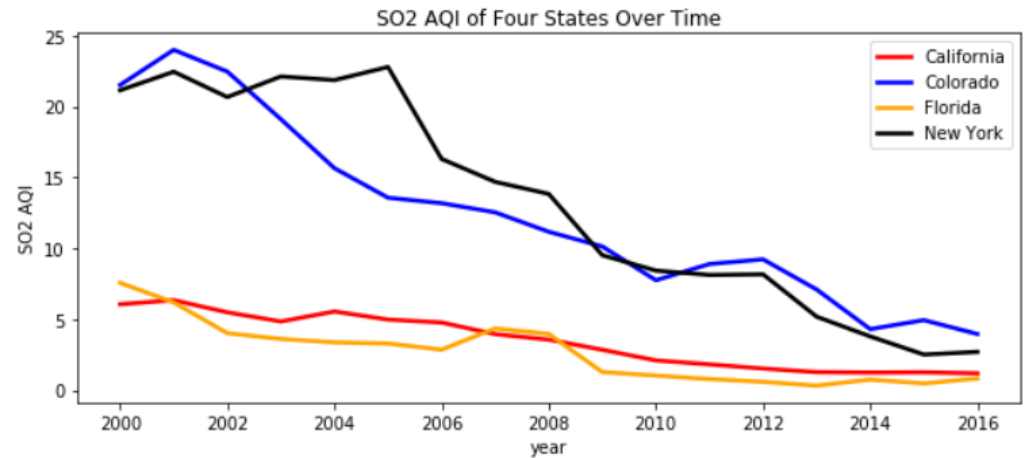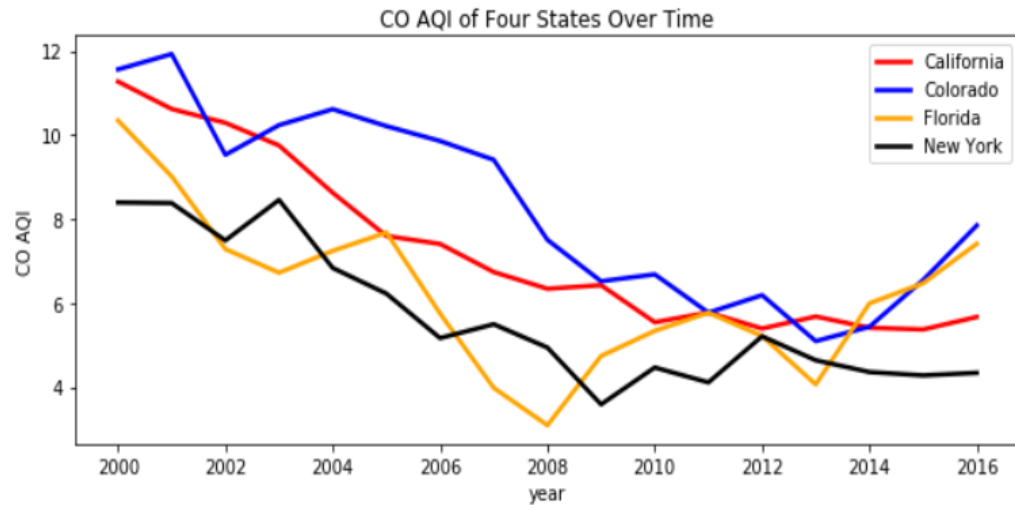
```
State   Colorado
Date Local   2014-03-06 00:00:00
CO Mean   0.145833
Name: 1513982, dtype: object
State   Arizona
Date Local   2000-10-09 00:00:00
CO AQI   1
Name: 1081, dtype: object
```
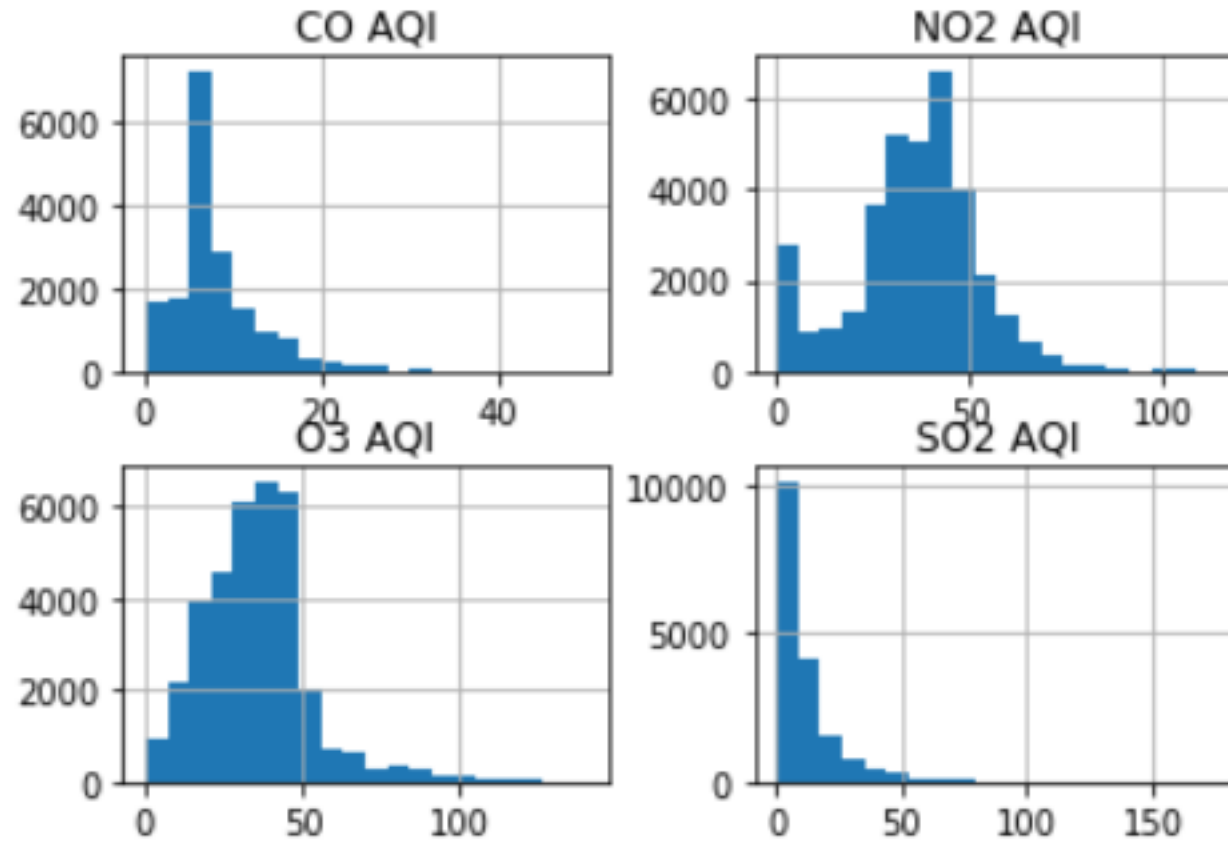
# Key result (Comparing AQI of 4 States)
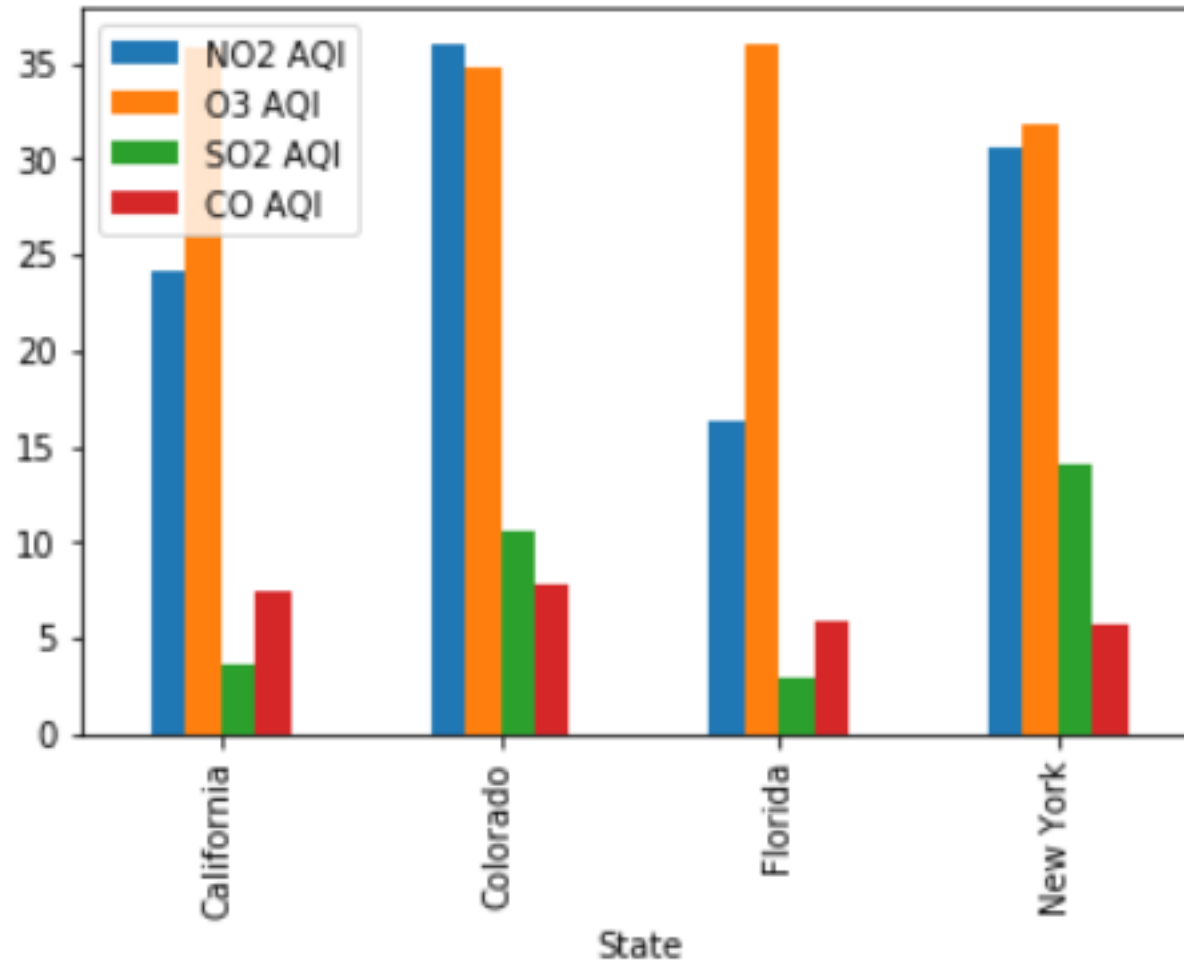
# Key result (4 pollutants AQI of Four States Over Time)
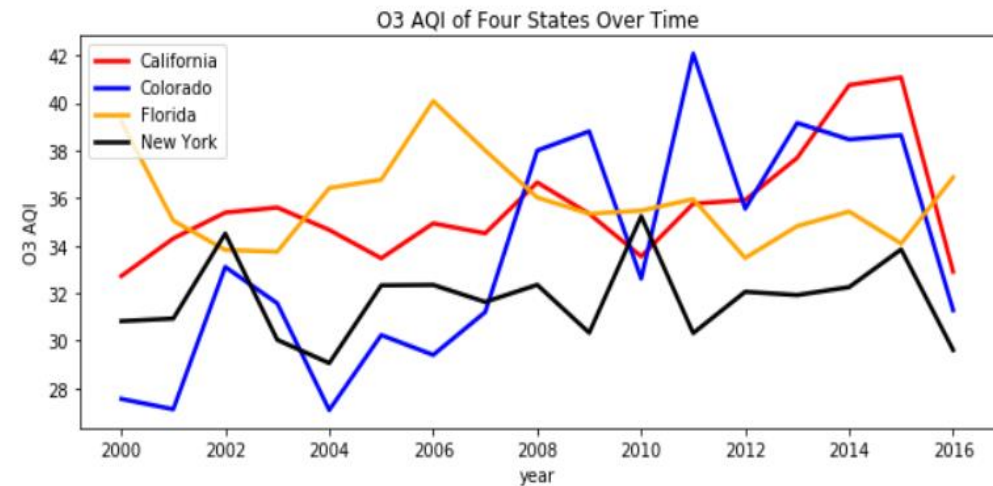
# Key result (Frequency of AQI all over the US)

# Knowledge gained

1. U.S. air quality is better than before

2. NO2 and O3 are the main pollutants for us

   a. According to the example four states:

      i. California

      ii. New York

      iii. Colorado

      iv. Florida

# How that knowledge can be applied

- Awareness of environmental pollution problem and its impact on people's lives.
- Focus on controlling the main pollutants(CO, SO2, O3, NO2)

  - Government can make policy to reduce pollution

  - People can have a daily life plan to reduce pollution

- Results of AQI for different pollutant in different states

  - Help future pollution detection

  - Prevent from producing more pollution

- Reuse of O3

# Reference

https://airnow.gov/index.cfm?action=aqibasics.aqi

https://www.opendatasoft.com/2017/03/14/open-data-air-pollution-a-powerful-tool-in-the-struggle-for-cleaner-air/

http://www.ipc.cas.cn/kxcb/kpwz/201503/t20150304_4317094.html