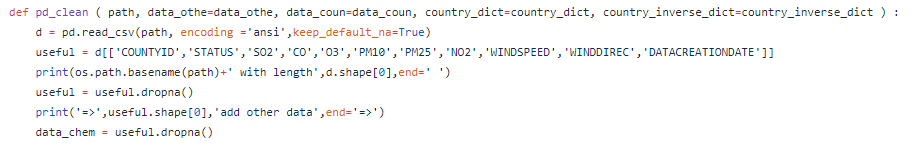
**Data**

In order to do the air quality index (AQI) prediction for cities in Taiwan during the year from 2015 to 2017, we consider the following data:

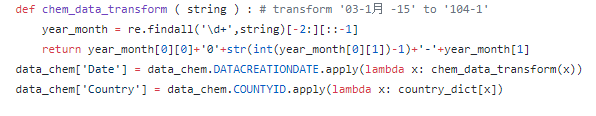
* Average income per city (Dept. of Household Registration, M.O.I. 內政部戶政司)
* Registration number of vehicle per city (Directorate General of Highways, MOTC 中華民國交通部公路總局)
* Population per city (Dept. of Household Registration, M.O.I. 內政部戶政司)
* AQI data (Civil IoT Taiwan)

**Data Preprocessing**

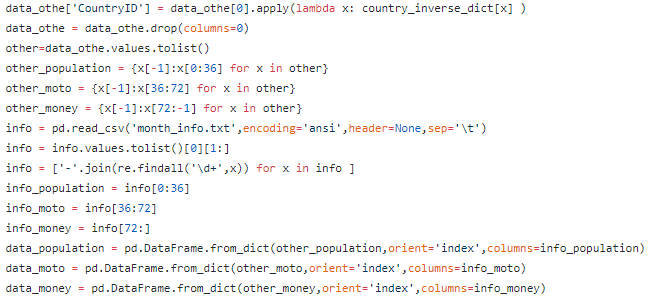
1. Keep useful features from AQI data



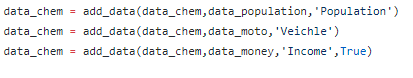
1. Transform '03-1月 -15' to '104-1'



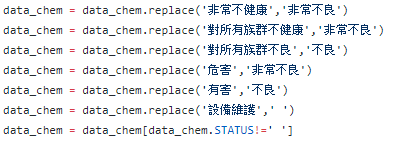
1. Since the country in ‘交通36\_人口36\_所得3.xlsx’ are chinese and the country in ‘Country\_ID.xlsx' and the data of PSI are number, we make a map of chinese country with its index. We also transform the datadate to the same form : year/month, so that we can merge several data together. Finally we save data\_population, data\_moto, and data\_money to record the data.



1. Combine population, vehicle and income data and AQI data together



1. Label processing

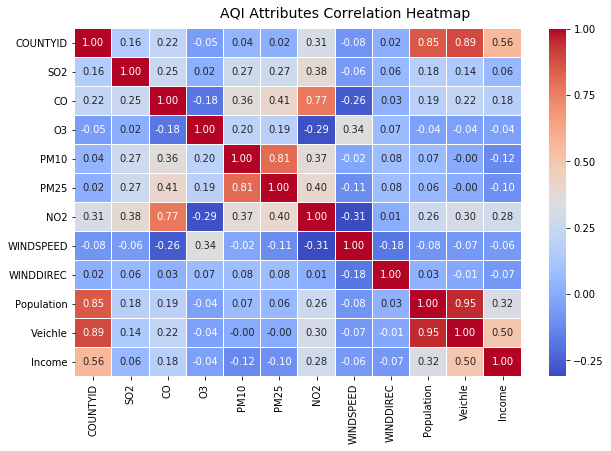


1. Normalization data values

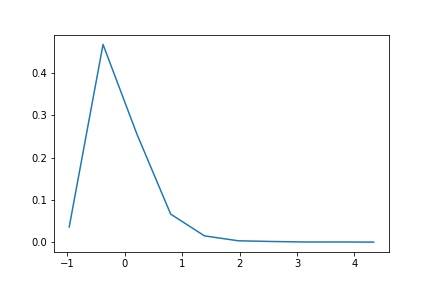


**Result**

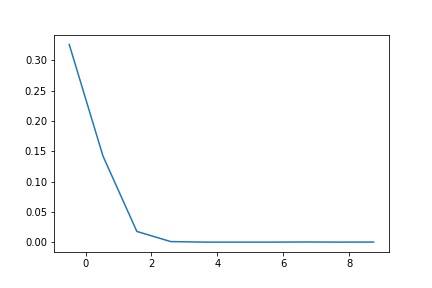
1. AQI Attributes correlation heatmap



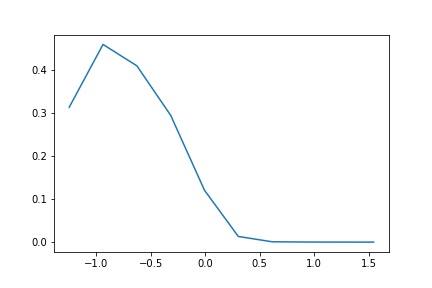
1. Probability density function of CO for good air quality

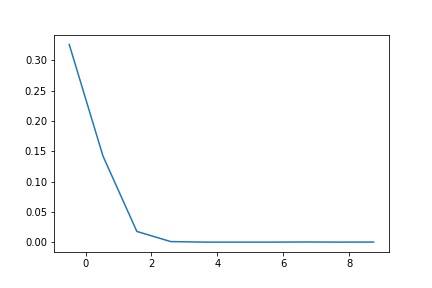


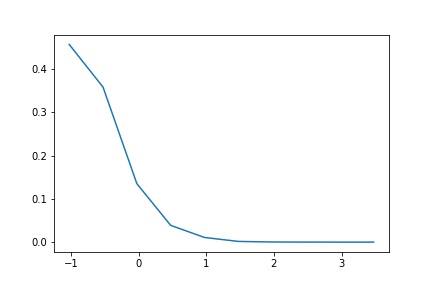
Probability density function of CO for bad air quality



1. Probability density function of O3 for good air quality



Probability density function of O3 for bad air quality

1. Probability density function of NO2 for good air quality

Probability density function of NO2 for bad air quality

