

## **QBUS6840: Tutorial 2 Tasks**

### **PM2.5 data of Beijing**

This hourly data set contains the PM2.5 data in Beijing. The time period is between Jan 1st, 2010 to Dec 31st, 2015. Missing data are denoted as NA.

Now you are required to write a python script to:

1. Read the dataset and store the data in a Dataframe. Use time information as the index of this Dataframe. Note that the time information could be found in the second column (year-month-day-hour) of the csv file.
2. Extract the PM\_US\_Post column and then store the data according to the year information (say, store each year's data into a single Series variable).
3. Extract the data in March of each year.
4. Calculate the mean value of each March's data, and then plot these mean value by using a line curve.