pre-processing PM2.5(tp)

October 20, 2019

[39]:
```
import pandas as pd
import math
import csv

k = 4
PM_raw = [0]*k
path = './data_tp/'
#print(PM_raw)

for i in range(k):
    PM_raw[i] = pd.read_csv('./data_tp/'+ str(104+i) + '  _utf8.csv')
    #print(i)
#PM18_raw = pd.read_csv('./data/107  _20190315_utf16.csv', encoding = 'utf-16')
```

[40]:
```
PMd = {}

for kase in range(k):
    l_PM = len(PM_raw[kase])
    for i in range(l_PM) :
        if (PM_raw[kase][" "][i] == "PM2.5") & str(PM_raw[kase]["12"][i]).
 ↪isdigit() :
            PMd[PM_raw[kase][" "][i]] = PM_raw[kase]["12"][i]
```

[41]:
```
len(PMd)
```

[41]:
```
1401
```

[42]:
```
import csv
csv_file = path+"PM2.5_all.csv"
#csv_columns = ['date', '12AM PM2.5']
#csv_data = [{'date': key , '12AM PM2.5':PM25[key] } for key in PM25_18]

try:
    with open(csv_file, 'w') as csvfile:
        w = csv.writer(csvfile)
        w.writerow(['date','PM2.5'])
        for key, val in PMd.items():
```

```
            w.writerow([key, val])

except IOError:
    print("I/O error")
```

[ ]:

[ ]: