

ML in EDU

Homework 1

Getting Data From RESTful API & Visualization

Reference: <https://hackmd.io/@ysliu/H1vGPMTLr>

CONTENTS

01

**Data Collection
(30%)**

02

**Visualization
(40%)**

03

**Analysis
(30%)**

04

Hand in

<Hint> You maybe need these packages:

`requests, json, matplotlib, pandas(optional), scipy(optional)`

Data Collection (30%)

Request data from RESTful API

空氣品質指標(AQI)



政府資料開放平臺
DATA.GOV.TW



資料集評分:



平均 4 (38 人次投票)

資料集描述:

每小時提供各測站之空氣品質指標 (AQI)，原始資料版本公告於空氣品質監測網<https://taqm.epa.gov.tw>

主要欄位說明:

SiteName(測站名稱)、County(縣市)、AQI(空氣品質指標)、Pollutant(空氣污染指標物)、Status(狀態)、SO2(二氧化硫ppb)、CO(一氧化碳ppm)、CO_8hr(一氧化碳8小時移動平均ppm)、O3(臭氧ppb)、O3_8hr(臭氧8小時移動平均ppb)、PM10(懸浮微粒 $\mu\text{g} / \text{m}^3$)、PM2.5(細懸浮微粒 $\mu\text{g} / \text{m}^3$)、NO2(二氧化氮ppb)、NOx(氮氧化物ppb)、NO(一氧化氮ppb)、WindSpeed(風速 m / sec)、WindDirec(風向degrees)、PublishTime(資料發布時間)、PM2.5_AVG(細懸浮微粒移動平均值 $\mu\text{g} / \text{m}^3$)、PM10_AVG(懸浮微粒移動平均值 $\mu\text{g} / \text{m}^3$)、SO2_AVG(二氧化硫移動平均值ppb)、Longitude(經度)、Latitude(緯度)、SiteId(測站編號)

資料下載網址:

JSON



檢視資料

空氣品質指標(AQI)-JSON

XML



檢視資料

空氣品質指標(AQI)-XML

CSV



檢視資料

空氣品質指標(AQI)-CSV

01

Data Collection (30%)

Request data from RESTful API



政府資料開放平臺
DATA.GOV.TW

欄位說明參照：

<https://data.gov.tw/dataset/40448>

下載 json 檔：

[http://opendata.epa.gov.tw/ws/Data/AQI/?\\$format=json](http://opendata.epa.gov.tw/ws/Data/AQI/?$format=json)

(注意：網址請將 https 改成 http 以避免安全認證問題)

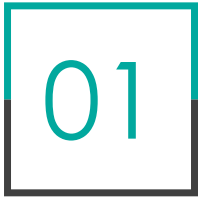
Attributes reference：

<https://data.gov.tw/dataset/40448>

Download .json file：

[http://opendata.epa.gov.tw/ws/Data/AQI/?\\$format=json](http://opendata.epa.gov.tw/ws/Data/AQI/?$format=json)

(<Notes> Please modify the url from https to http to avoid security issue)



Data Collection (30%)

Request data from RESTful API

請使用 HTTP request 直接將資料抓到程式變數內。

可以隨意使用套件 (e.g. requests, request, PycURL, urllib2 ...)

Please send HTTP request to fetch data into your variables

You may use whatever packages you like (e.g. requests, request, PycURL, urllib2)



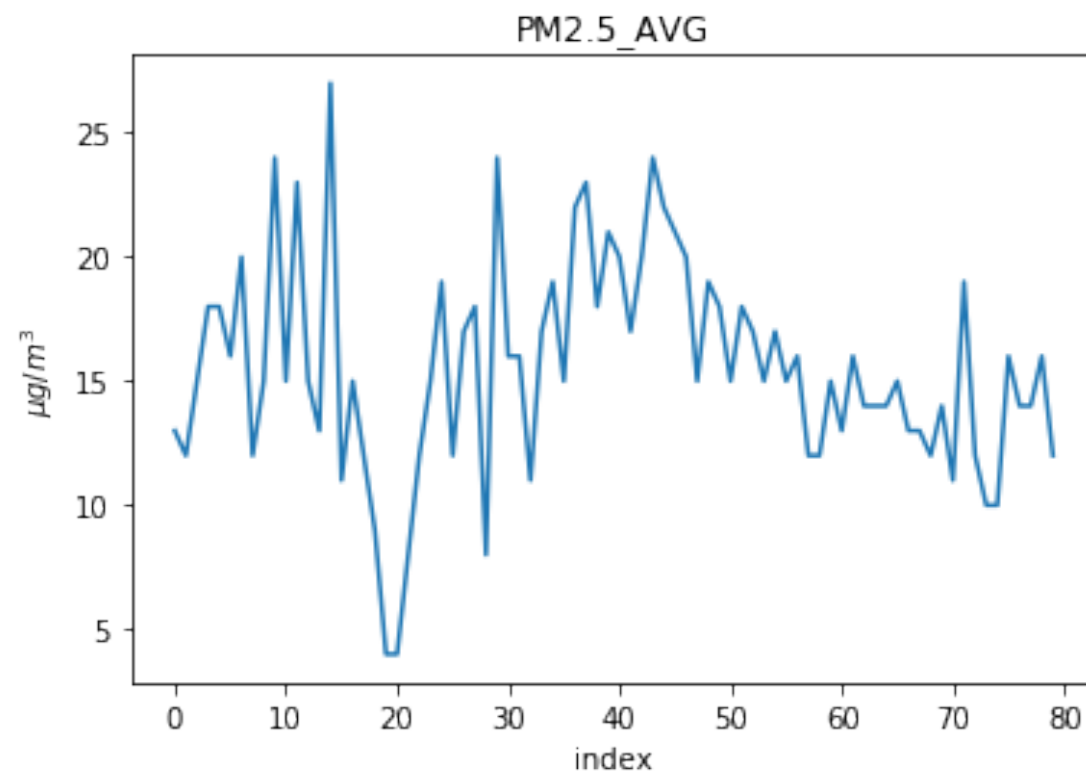
Visualization (40%)

Plot a line graph of an attribute

E.g. line graph of "PM2.5_AVG"

請對 AQI 欄位畫出折線圖。

Plot a line graph of the "AQI" column



<Hint> You may check the y-axis should be sorted

03

Analysis (30%)

Using 1st ~ 4th moments to analyze the distribution

請算出 AQI 的一階動差（平均數）、二階_{中心}動差（變異數）、
三階_{標準}動差（偏態）、四階_{標準}動差（峰態）並分析你觀察到的結果。

Please get mean, variance, skewness and kurtosis of the AQI,
and explain what you find.

$$\text{mean} = EX$$

$$\text{variance} = E(X - \mu_X)^2$$

$$\text{skewness} = \frac{E(X - \mu_X)^3}{\sigma^3}$$

$$\text{Kurtosis} = \frac{E(X - \mu_X)^4}{\sigma^4}$$

Ref :

[https://en.wikipedia.org/wiki/Moment_\(mathematics\)](https://en.wikipedia.org/wiki/Moment_(mathematics))

<https://zh.wikipedia.org/wiki/偏度>

<https://zh.wikipedia.org/zh-tw/峰度>

04

Hand in your homework to e3

Hand in your report & code to e3(<https://e3new.nctu.edu.tw>)

You should only hand in 2 files:

`hw1_<student_id>.pdf`

`hw1_<student_id>.zip`

(e.g. hw1_0123456.pdf, hw1_0123456.zip)

Due: 10/8 (Tue.) 11:00 a.m.

TA:

蔡旻均 dollars9256741@gmail.com

劉昱劭 ysl@cs.nctu.edu.tw