



巨集程式與資料分析應用 Macro programming and data analytics

Section 3

網路爬蟲與資料分析

陳彥佑



課前說明

- 延續家耀老師與明居老師的內容
 - 1項作業、1項Term project(2024/01/04)
 - 2024/01/04 Term Project報告方式與第一階段相同，同學必須將報告內容錄製成影片繳交，當天會選擇幾組上台報告
 - ✓ 原則上以前兩階段未報告的組別為優先
 - ✓ 前兩階段已經報告過組別，自願上台報告者加分



課程說明

- Section 1
 - Excel VBA
 - Excel儀表板與資料視覺化
 - ...
- Section 2
 - Python常用語法與觀念
 - JSON與CSV檔案下載分析
 - 資料繪圖(Python)
 - ...
- Section 3 (實例應用)
 - Python常用語法與觀念
 - 巨量資料下載(XML檔)
 - 巨量資料分析實例應用
 - ...



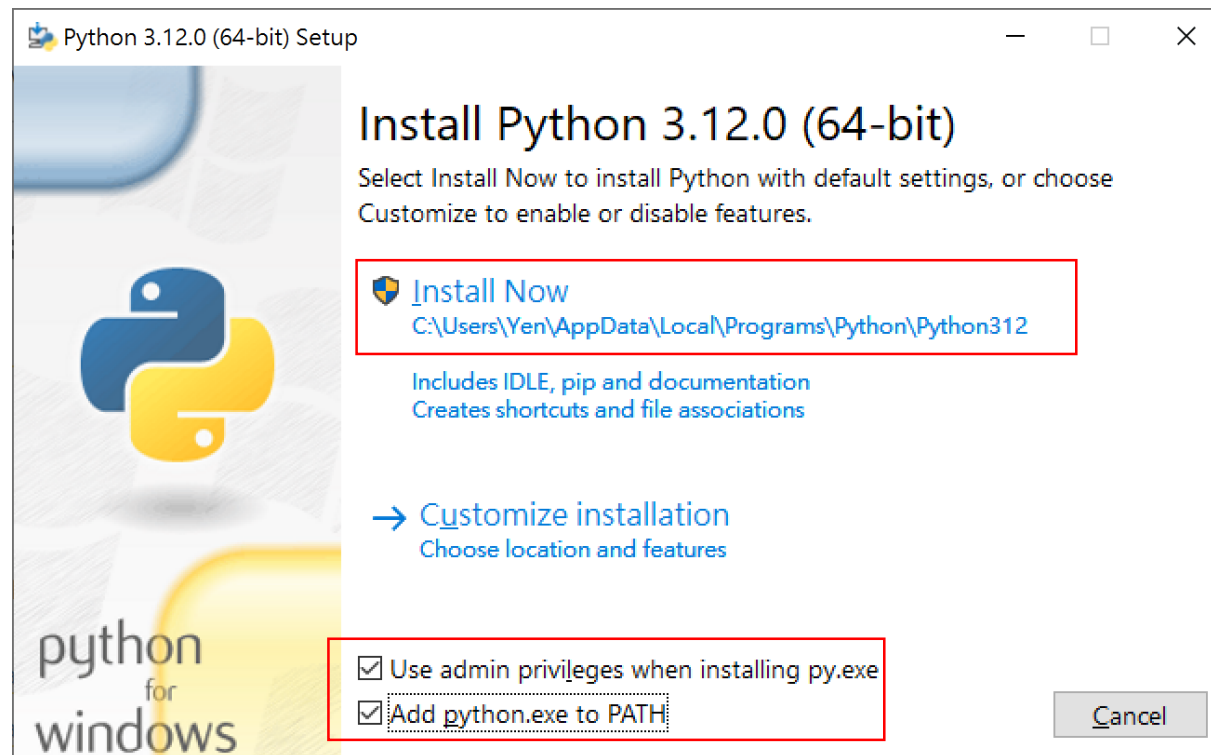
課程進度 – Section 3

日期	課程主題
12/07	高公局資料格式說明與下載
12/14	偵測器資料讀取、展示與分析
12/21	電子收費資料讀取、展示與分析
12/28	偵測器資料及電子收費資料整合應用
01/04 (2024)	Term Project 3 – 期末成果發表



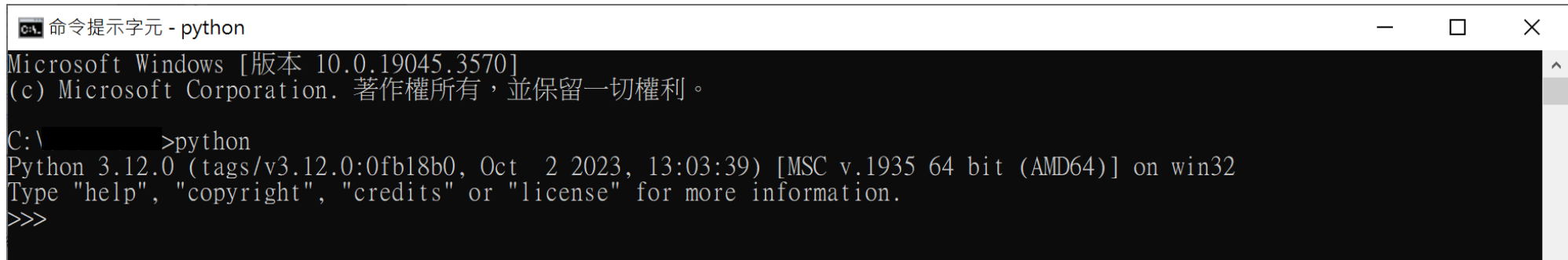
環境建置 (安裝python)

- 下載並安裝python
 - <https://www.python.org/>





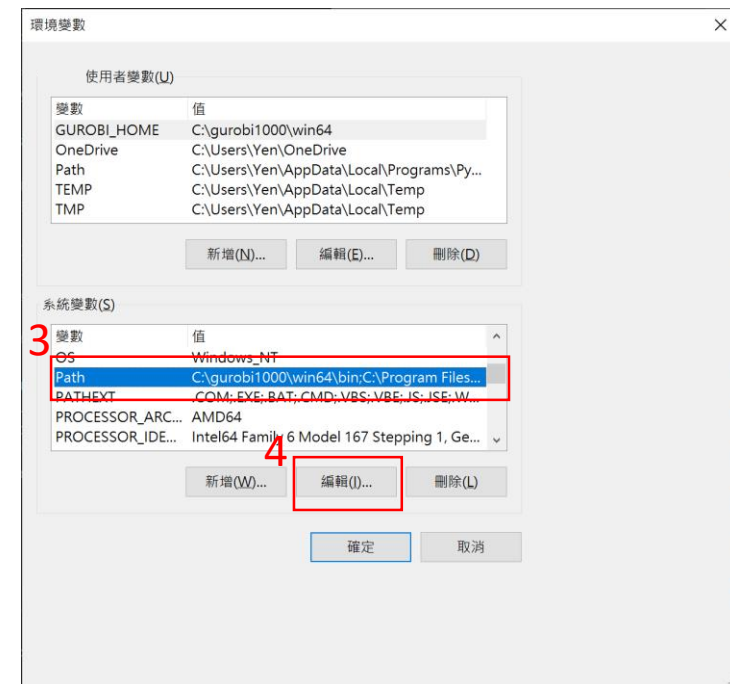
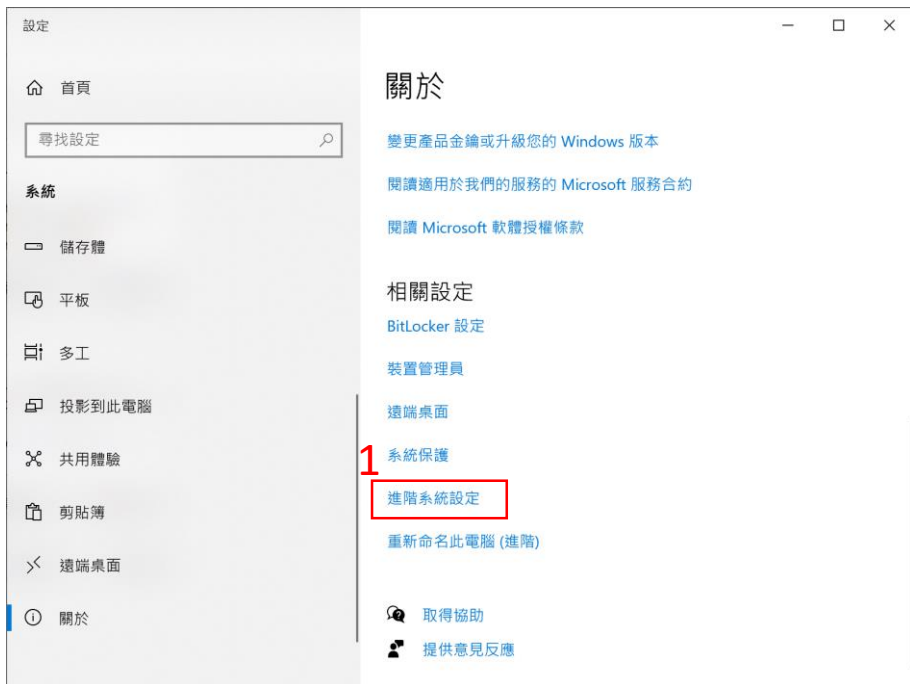
- 測試是否有安裝成功
 - 開啟命令提示字元，輸入python，按Enter



```
命令提示字元 - python
Microsoft Windows [版本 10.0.19045.3570]
(c) Microsoft Corporation. 著作權所有，並保留一切權利。
C:\>python
Python 3.12.0 (tags/v3.12.0:0fb18b0, Oct 2 2023, 13:03:39) [MSC v.1935 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

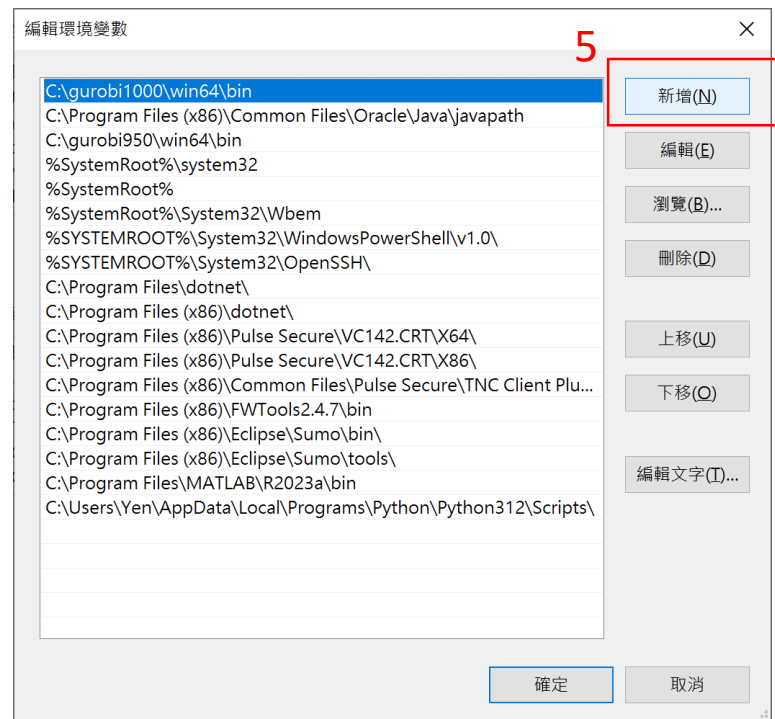


- 在「我的電腦」上按右鍵->內容->進階系統設定->環境變數->選取系統變數Path->編輯->加入指令路徑->確認設定
 - 確認以下兩個路徑有被加入(路徑中之**user**為本身電腦的使用者名稱，需依照個別電腦修改)
C:\Users**user**\AppData\Local\Programs\Python\Python312\Scripts
C:\Users**user**\AppData\Local\Programs\Python\Python312





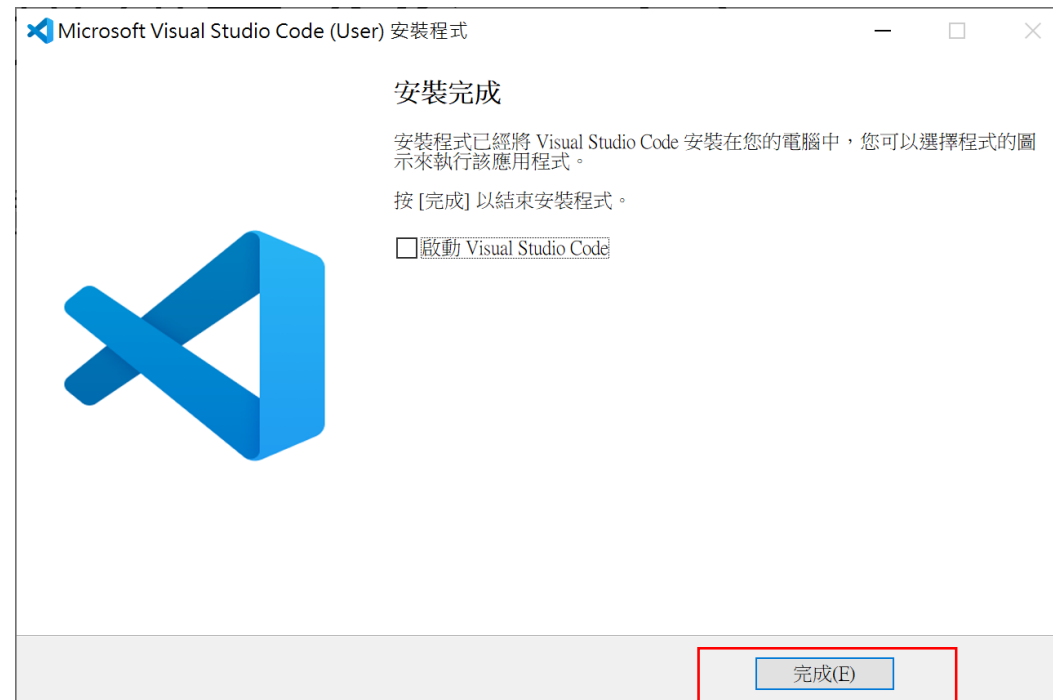
- 若無，則選新增，加入以下兩路徑(**user**為本身電腦的使用者名稱)
 - C:\Users**user**\AppData\Local\Programs\Python\Python312\Scripts
 - C:\Users**user**\AppData\Local\Programs\Python\Python312





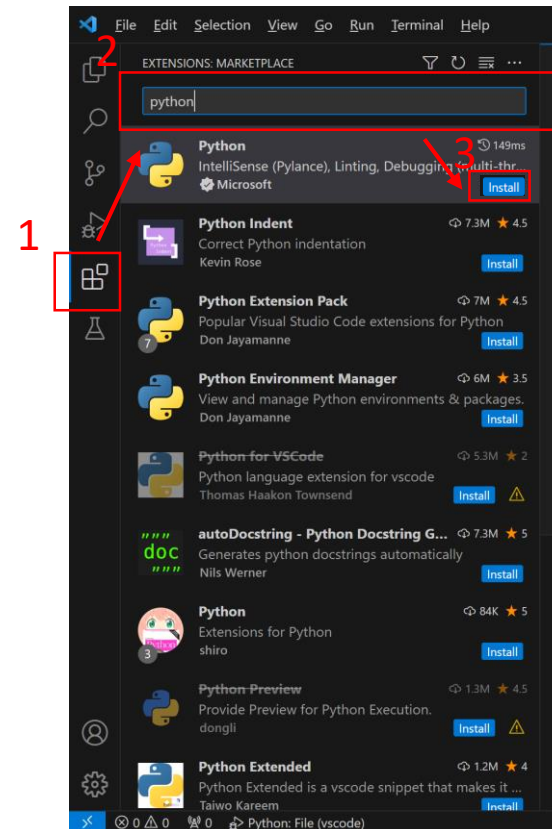
環境建置 (VS code)

- 下載VS code並安裝
 - <https://code.visualstudio.com/download>



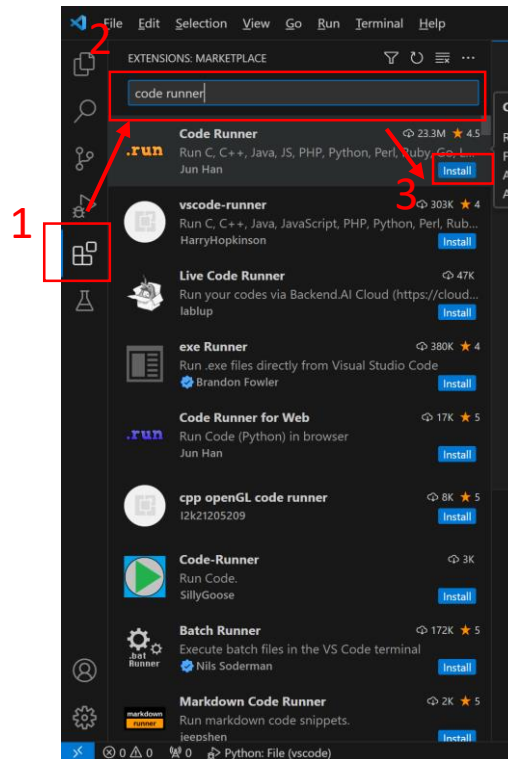
環境建置 (VS code)

- 開啟VSCode，點擊Extensions圖示。在搜尋的地方輸入【Python】，並點選install

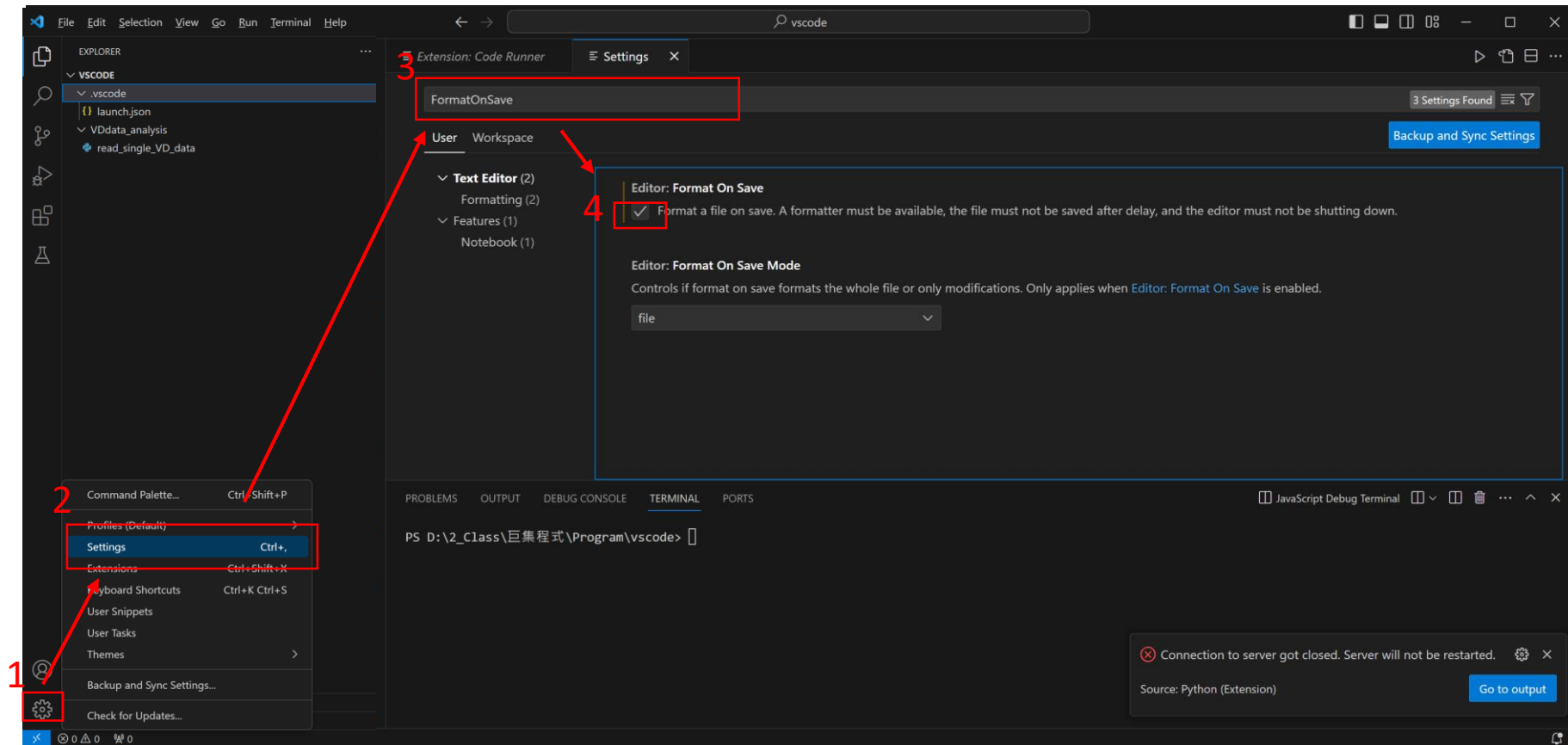


環境建置 (VS code)

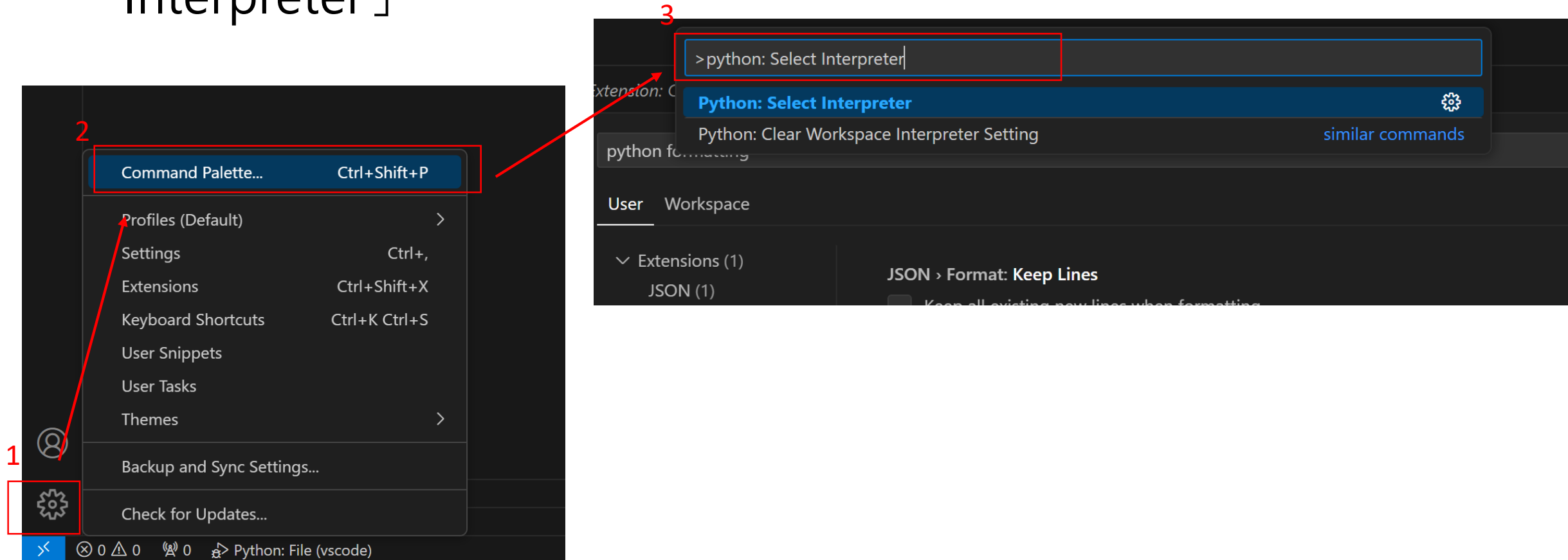
- 安裝Code Runner套件(用來方便我們執行程式碼)。
- 搜尋輸入【Code Runner】，點擊【Install】進行安裝



- 選擇左下角齒輪圖案，並選擇setting，在搜尋的地方輸入【FormatOnSave】，勾選【Format a file on save】選項，讓VSCode在我們儲存檔案的時候，自動依照指定程式碼風格原則幫我們排版程式碼。

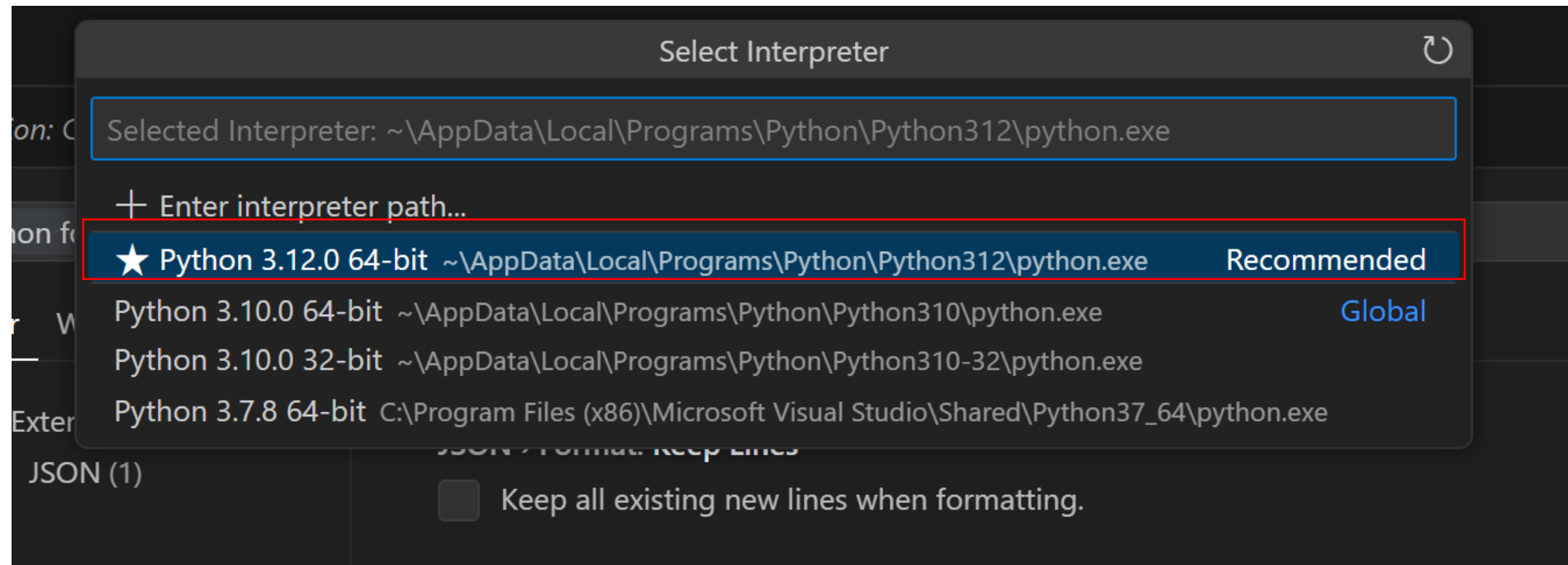


- 最後，於VSCode選擇Python的直譯器，也就是位於剛剛所安裝的Python路徑下
- 輸入「python: Select Interpreter」並選擇「python: Select Interpreter」





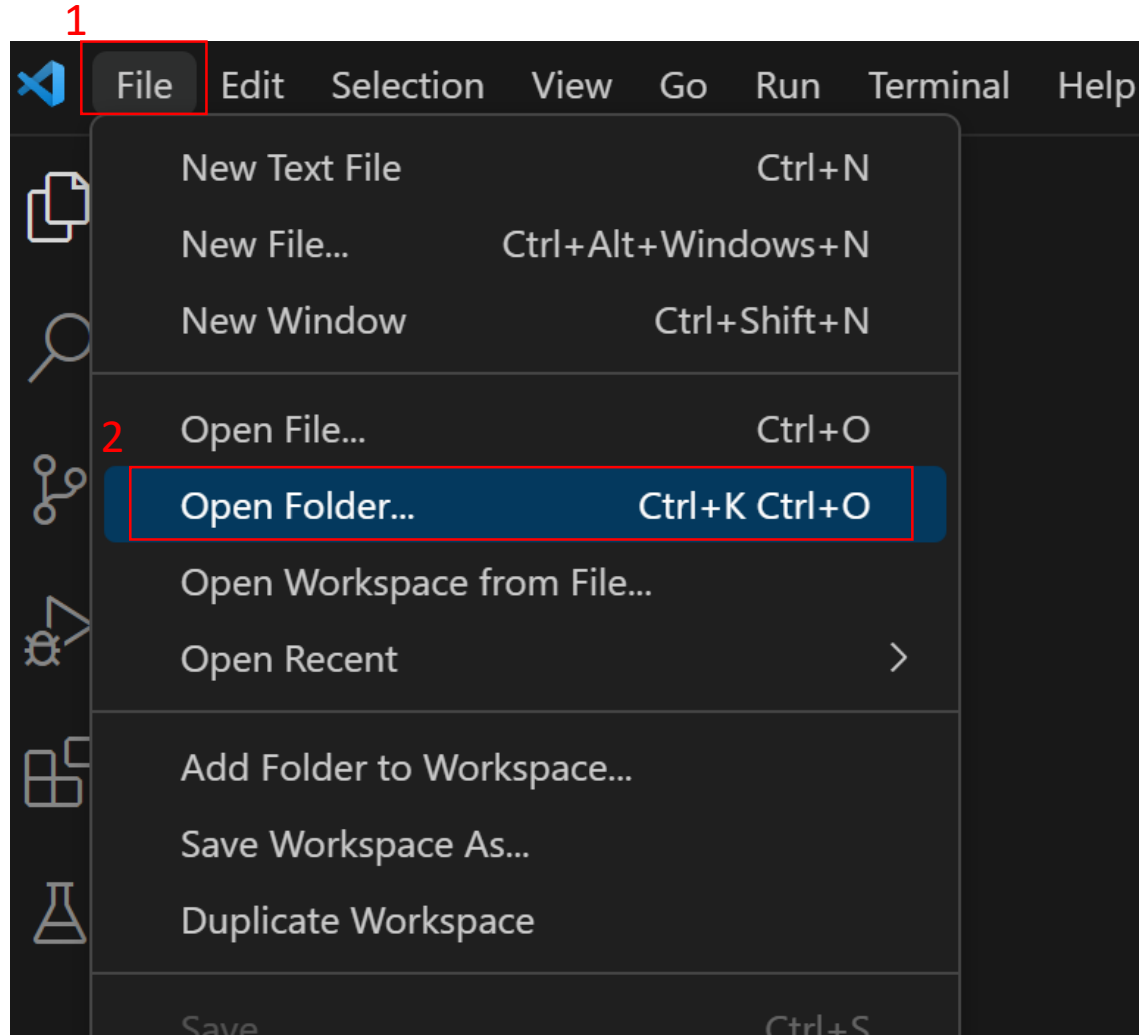
- 選擇剛剛安裝的python 3.12.0





開始使用VS code

- File → Open Folder



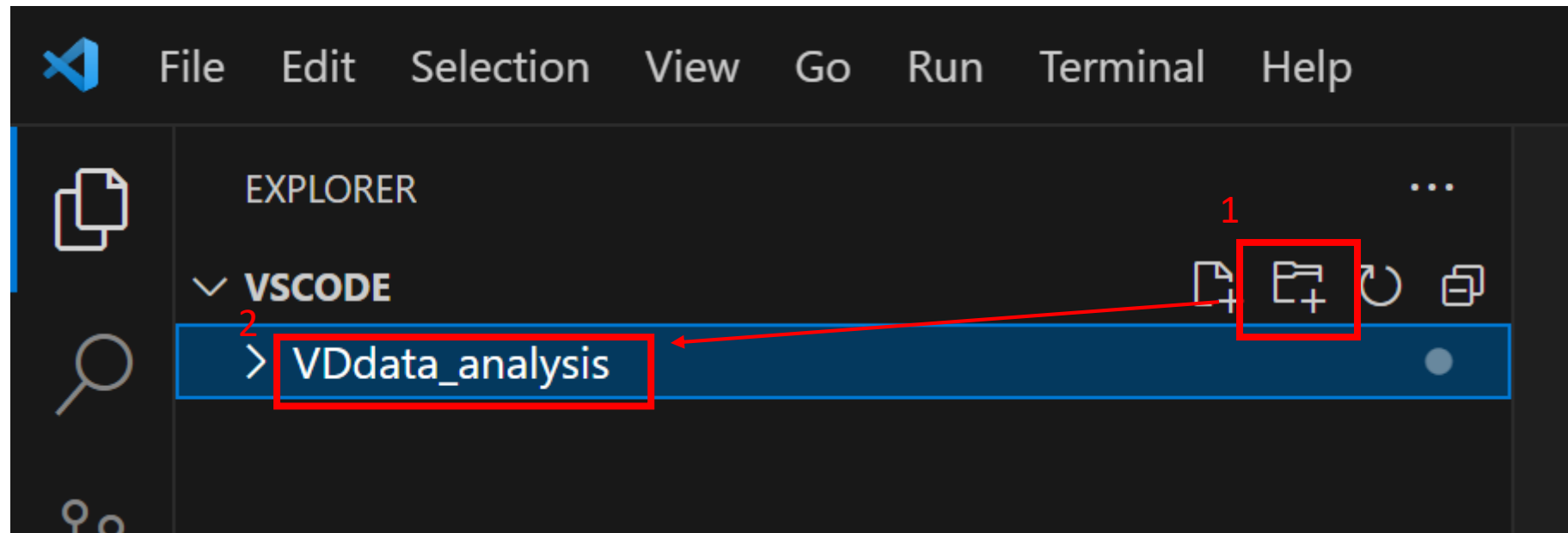
- 點選右鍵新建一個資料夾





開始使用VS code

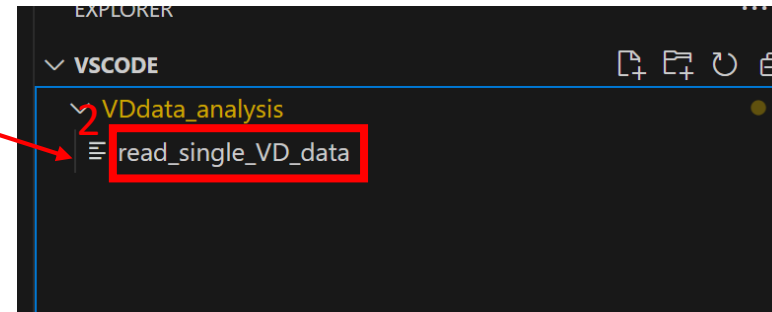
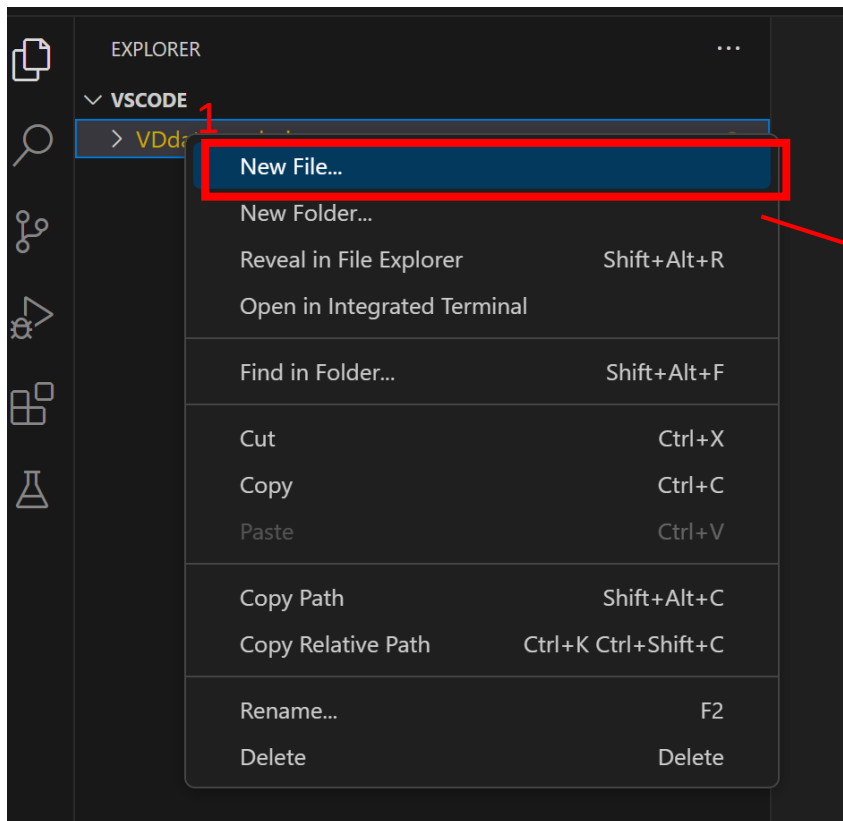
- 選擇New Folder，並輸入資料夾名稱





開始使用VS code

- 在新建的資料夾上按右鍵，選擇New File，並輸入檔案名稱





高公局即時路況 -- 1968

- <https://1968.freeway.gov.tw/>

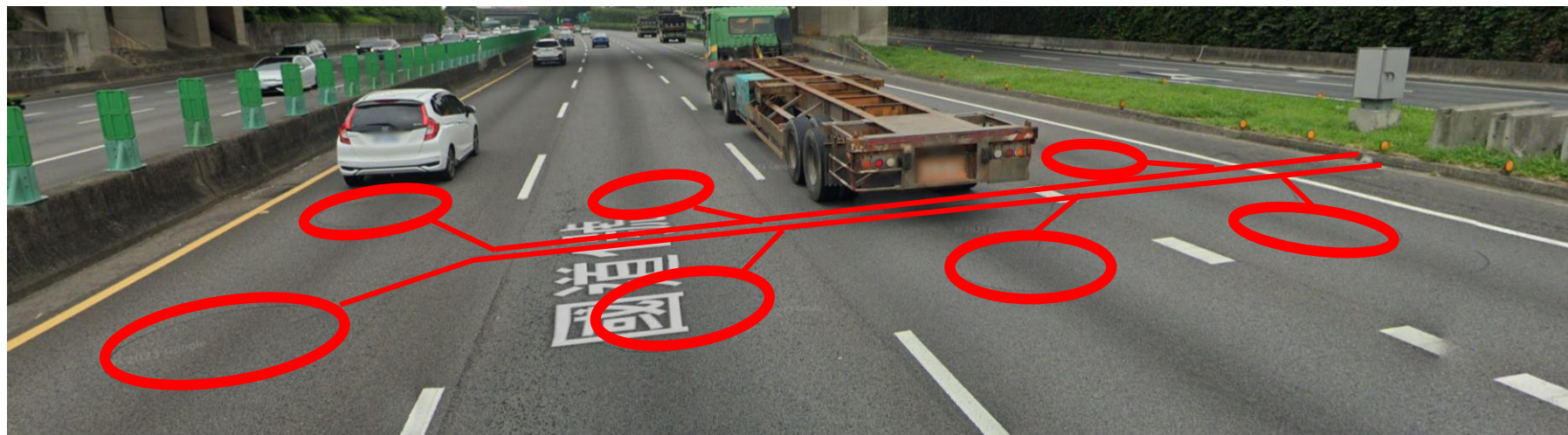


高公局資料來源

- Vehicle detectors (VD)
- Electronic toll collection (ETC)
- 中央氣象局
- 即時影像
- 事件資料庫
-



[Source: Google Map]



[Source: Google Map]



高公局資料來源

- 即時資料 & 歷史資料
 - 即時資料
 - 提供資訊(旅客、控制中心、數位雙生(digital twins))
 - 即時檢視控制策略績效
 -
 - 歷史資料
 - 數位雙生應用(digital twins)
 - 塞車原因分析
 - 控制策略擬定(定時、AI)
 - 偵測器功能是否異常之檢驗與確認
 - ...

交通部高速公路局VD資料庫介紹

<https://tisvcloud.freeway.gov.tw/>

交通部高速公路局
FREEWAY BUREAU, MOTC

交通資料庫

訂定目的

為利促進及活化政府資料應用並提升政府資料品質及價值，交通部高速公路局(以下簡稱本局)彙整國道即時及歷史交通資訊，開放「交通資料庫」(以下簡稱本資料庫)供產、官、學、研各界免費下載加值應用，以利用路人獲取更多元之服務，特訂定本注意事項。

使用下載資料應注意事項

- 1 本資料庫資料來源為「高快速公路整體路網交通管理系統」與「高速公路電子收費系統」，資料免費提供下載使用。
- 2 本系統資料格式係依據交通部發布之「即時路況資料標準2.0」及本局「高速公路電子收費交通資料蒐集支援系統(TDCS, Traffic Data Collection System)使用手冊」。
- 3 下載資料內容為系統自動產出，系統無法預估斷線或設備故障等情況，故資料僅供參考，下載者皆不得以本局 資料庫所提供之資訊主張任何權利。
- 4 使用由本資料庫下載之資料所加值之服務或研究成果，於對外提供時均應聲明介接本局「交通資料庫」，若未能執行本局有權限制連線。(顯名聲明：交通部高速公路局「交通資料庫」)
- 5 使用者重複擷取資料週期間應大於40秒，以免影響他人下載權利。若發現有違反情事，本局得逕行中斷其連線，使用者不得異議。
- 6 本資料庫資料不得進行任何違反中華民國法律之使用及應用，並須遵循國發會政府資料開放相關規範。
- 7 使用者開始進行資料下載即視同同意本注意事項之相關規定。
- 8 為配合交通部政策，本平台路況資訊1.1版已於111年10月31日終止服務，請使用者儘速改接2.0版
- 9 交通資料庫訂於8/1 19:00-8/2 01:30將進行系統維護，期間資料不定時暫停更新，造成不便，敬請見諒。

交通資料庫開放項目及說明

歷史資料庫

項目	更新週期(檔案/內容)	格式	說明	欄位資訊	下載
VD 動態資訊(v2.0)	1day/1min	xml	1分鐘定時更新及統整單一VD的動態偵測值		
VD 靜態資訊(v2.0)	1day/1day	xml	提供VD之空間位置描述及其他相關欄位		
最新消息(v2.0)	1day/1min	xml	提供即時路況事件資料		
路段即時路況動態資訊(v2.0)	1day/1min	xml	提供機關發布路段或基礎路段的即時路況動態資料		
eTag 配對路徑動態資訊(v2.0)	1day/5min	xml	提供eTag之配對路徑動態資訊		
eTag 配對路徑靜態資訊(v2.0)	1day/1day	xml	提供eTag之配對路徑靜態資料		
eTag 靜態資訊(v2.0)	1day/1day	xml	提供eTag之空間位置描述資訊		
CMS 一分鐘動態資訊(v2.0)	1day/1min	xml	提供該CMS的即時訊息		
CMS 靜態資訊(v2.0)	1day/1day	xml	提供CMS之空間位置描述及其他相關欄位		
CCTV 靜態資訊(v2.0)	1day/1day	xml	提供CCTV之空間位置描述及其他相關欄位		

交通資料庫開放項目及說明 歷史資料庫					
項目	更新週期(檔案/內容)	格式	說明	欄位資訊	下載
VD 動態資訊(v2.0)	1day/1min	xml	1分鐘定時更新及統整單一VD的動態偵測值		
VD 靜態資訊(v2.0)	1day/1day	xml	提供VD之空間位置描述及其他相關欄位		
最新消息(v2.0)	1day/1min	xml	提供即時路況事件資料		
路段即時路況動態資訊(v2.0)	1day/1min	xml	提供機關發布路段或基礎路段的即時路況動態資料		
eTag 配對路徑動態資訊(v2.0)	1day/5min	xml	提供eTag之配對路徑動態資訊		
eTag 配對路徑靜態資訊(v2.0)	1day/1day	xml	提供eTag之配對路徑靜態資料		
eTag 靜態資訊(v2.0)	1day/1day	xml	提供eTag之空間位置描述資訊		
CMS 一分鐘動態資訊(v2.0)	1day/1min	xml	提供該CMS的即時訊息		
CMS 靜態資訊(v2.0)	1day/1day	xml	提供CMS之空間位置描述及其他相關欄位		
CCTV 靜態資訊(v2.0)	1day/1day	xml	提供CCTV之空間位置描述及其他相關欄位		



交通部高速公路局VD資料庫介紹

歷史資料庫檔案下載 [上一頁](#) [歷史資料庫](#)

Name	Last modified	Size
20231204/	2023-12-04 19:52	-
20231203/	2023-12-04 00:00	-
20231202/	2023-12-03 00:00	-
20231201/	2023-12-02 00:00	-
20231130/	2023-12-01 00:00	-
20231129/	2023-11-30 00:00	-
20231128/	2023-11-29 00:00	-
20231127/	2023-11-28 00:00	-
20231126/	2023-11-27 00:00	-
20231125/	2023-11-26 00:00	-
20231124/	2023-11-25 00:00	-

歷史資料庫檔案下載 [上一頁](#) [歷史資料庫](#)

Name	Last modified	Size
VDLive_2359.xml.gz	2023-11-29 23:59	168K
VDLive_2358.xml.gz	2023-11-29 23:58	168K
VDLive_2357.xml.gz	2023-11-29 23:57	169K
VDLive_2356.xml.gz	2023-11-29 23:56	168K
VDLive_2355.xml.gz	2023-11-29 23:55	168K
VDLive_2354.xml.gz	2023-11-29 23:54	168K
VDLive_2353.xml.gz	2023-11-29 23:53	169K
VDLive_2352.xml.gz	2023-11-29 23:52	168K
VDLive_2351.xml.gz	2023-11-29 23:51	169K
VDLive_2350.xml.gz	2023-11-29 23:50	169K



交通部高速公路局VD資料庫介紹

- XML (Extensible Markup Language , 可擴展標記語言) 為一種用於結構化數據的標記語言。其純文字檔之格式，有利其用於在各種應用程式和系統之間，交換和存儲數據。
 - 結構化
 - XML允許將數據結構化為元素和節點的層次結構。
 - 可擴展性
 - XML 是可擴展的，允許自定義標記和結構。
 - 平台獨立性
 - XML 是純文字格式，不依賴特定的平台。它可以在不同的系統和應用程序之間進行數據交換。



交通部高速公路局VD資料庫介紹

- 易於讀取和編寫
 - XML 通常使用標記語法，其結構和內容易於理解和編寫。
- 與網頁技術結合
 - XML 可以與 HTML 以及其他技術（例如 XPath、XSLT 和 XML Schema）結合使用，用於數據交換、網頁內容和數據驗證等。



交通部高速公路局VD資料庫介紹

XML declaration

Root
element

```
<?xml version="1.0" encoding="UTF-8"?>
<VDLiveList xsi:schemaLocation="http://traffic.transportd
  <UpdateTime>2023-12-04T19:48:00+08:00</UpdateTime>
  <UpdateInterval>60</UpdateInterval>
  <AuthorityCode>NFB</AuthorityCode>
  <LinkVersion>23.05.1</LinkVersion>
</VDLiveList>
```

Open tag

Root
element
closure

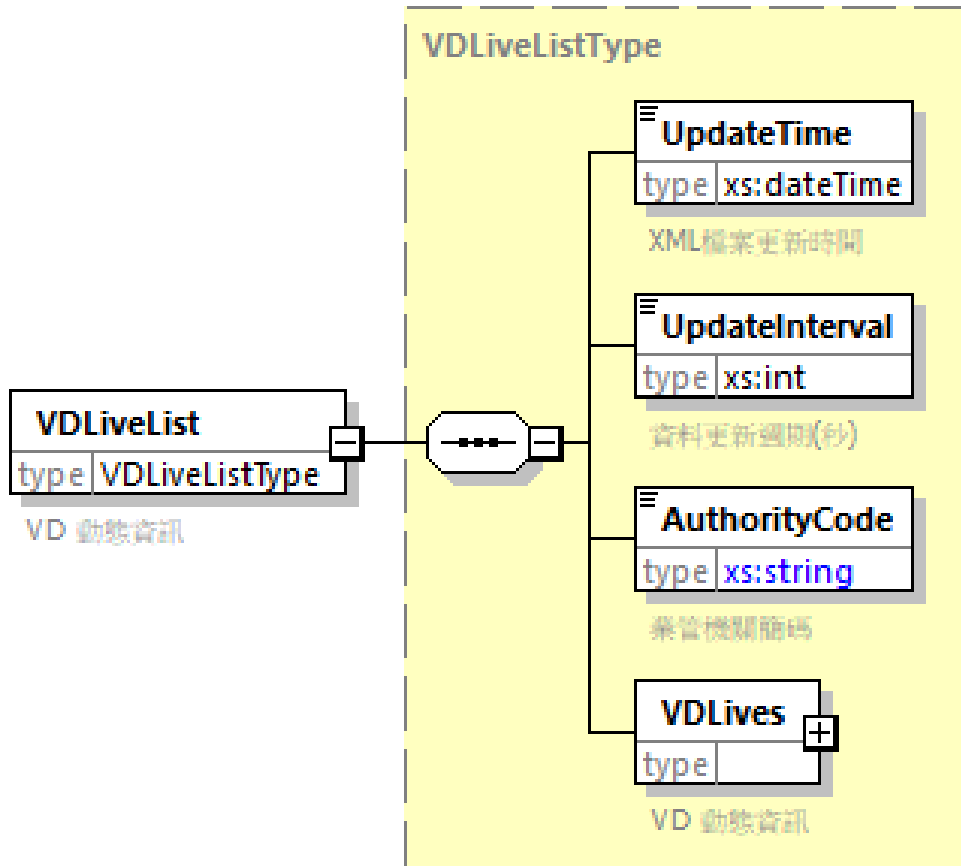
Content

Node

Close tag



交通部高速公路局 VD資料庫介紹



```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <VDLiveList xsi:schemaLocation="http://traffic.transportdata.tw/standar
3 <UpdateTime>2023-12-04T19:48:00+08:00</UpdateTime>
4 <UpdateInterval>60</UpdateInterval>
5 <AuthorityCode>NFB</AuthorityCode>
6 <LinkVersion>23.05.1</LinkVersion>
7 <VDLives>
8   <VDLive>
9     <VDID>VD-N1-N-198-O-SE-21-北上出口-彰化</VDID>
10    <LinkFlows>
11      <LinkFlow>
12        <LinkID>0000101145010N</LinkID>
13        <Lanes>
14          <Lane>
15            <LaneID>0</LaneID>
16            <LaneType>2</LaneType>
17            <Speed>61</Speed>
18            <Occupancy>2</Occupancy>
19            <Vehicles>
20              <Vehicle>
21                <VehicleType>S</VehicleType>
22                <Volume>6</Volume>
23                <Speed>60</Speed>
24              </Vehicle>
25              <Vehicle>
26                <VehicleType>L</VehicleType>
27                <Volume>0</Volume>
28                <Speed>0</Speed>
29              </Vehicle>
30              <Vehicle>
31                <VehicleType>T</VehicleType>
32                <Volume>0</Volume>
33                <Speed>0</Speed>
34              </Vehicle>
35            </Vehicles>
36          </Lane>
37        </Lanes>
38      </LinkFlow>
39    </LinkFlows>
40    <Status>0</Status>
41    <DataCollectTime>2023-12-04T19:46:00+08:00</DataCollectTime>
42  </VDLive>
```

檔案

常用

插入

版面配置

公式

資料

校閱

檢視

開發人員

Acrobat

小組

設計

告訴我您想要執行的動作...

剪下

複製

貼上

複製格式

剪貼簿

新細明體

12

A

A

B

I

U

字型

對齊方式

自動換列

通用格式

數值

樣式

設定格式化的條件

格式化為表格

一般

中等

好

壞

計算方式

連結的儲...

備註

說明文字

插入

刪除

格式

儲存格

自動加總

填滿

清除

排序與篩選

尋找與選取

建立 PDF 並共用連結

Adobe Acrobat

A1

fx

ns1:UpdateTime

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	ns1:UpdateTime	ns1:UpdateInterval	ns1:AuthorityCode	ns1:VDID	ns1:LinkID	ns1:LaneID	ns1:LaneType	ns1:Speed	ns1:Occupancy	ns1:VehicleType	ns1:Volume	ns1:Speed	ns1>Status	ns1:DataCollectTime			
2	2022-01-01T00:03:00+08:00	60	NFB	VD-N1-N-198-O-SE-21-北上出口-彰化	0000101145010N	0	2	50	2	S	3	50	0	2022-01-01T00:00:00+08:00			
3	2022-01-01T00:03:00+08:00	60	NFB	VD-N1-N-198-O-SE-21-北上出口-彰化	0000101145010N	0	2	50	2	L	1	46	0	2022-01-01T00:00:00+08:00			
4	2022-01-01T00:03:00+08:00	60	NFB	VD-N1-N-198-O-SE-21-北上出口-彰化	0000101145010N	0	2	50	2	T	0	0	0	2022-01-01T00:00:00+08:00			
5	2022-01-01T00:03:00+08:00	60	NFB	VD-T82-W-10-O-EN-1-Loop	1008200100940Q	0	2	59	1	S	2	58	0	2022-01-01T00:00:00+08:00			
6	2022-01-01T00:03:00+08:00	60	NFB	VD-T82-W-10-O-EN-1-Loop	1008200100940Q	0	2	59	1	L	0	0	0	2022-01-01T00:00:00+08:00			
7	2022-01-01T00:03:00+08:00	60	NFB	VD-T82-W-10-O-EN-1-Loop	1008200100940Q	0	2	59	1	T	0	0	0	2022-01-01T00:00:00+08:00			
8	2022-01-01T00:03:00+08:00	60	NFB	VD-T82-W-14-O-EN-1-Loop	1008201102010Q	0	2	74	1	S	1	73	0	2022-01-01T00:00:00+08:00			
9	2022-01-01T00:03:00+08:00	60	NFB	VD-T82-W-14-O-EN-1-Loop	1008201102010Q	0	2	74	1	L	0	0	0	2022-01-01T00:00:00+08:00			
10	2022-01-01T00:03:00+08:00	60	NFB	VD-T82-W-14-O-EN-1-Loop	1008201102010Q	0	2	74	1	T	0	0	0	2022-01-01T00:00:00+08:00			
11	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	0	2	-99	-99	S	-99	-99	0	2022-01-01T00:00:00+08:00			
12	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	0	2	-99	-99	L	-99	-99	0	2022-01-01T00:00:00+08:00			
13	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	0	2	-99	-99	T	-99	-99	0	2022-01-01T00:00:00+08:00			
14	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	1	2	0	0	S	0	0	0	2022-01-01T00:00:00+08:00			
15	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	1	2	0	0	L	0	0	0	2022-01-01T00:00:00+08:00			
16	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	1	2	0	0	T	0	0	0	2022-01-01T00:00:00+08:00			
17	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	2	2	-99	-99	S	-99	-99	0	2022-01-01T00:00:00+08:00			
18	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	2	2	-99	-99	L	-99	-99	0	2022-01-01T00:00:00+08:00			
19	2022-01-01T00:03:00+08:00	60	NFB	VD-N2-W-11.344-N-LOOP	0000200101100H	2	2	-99	-99	T	-99	-99	0	2022-01-01T00:00:00+08:00			
20	2022-01-01T00:03:00+08:00	60	NFB	VD-N1-N-33-I-EN-1-五股	0000101115080F												



交通部高速公路局VD資料庫介紹

- 所有國道上的VD個數：3763個
- VD資料欄位說明：<https://tisvcloud.freeway.gov.tw/history-list.php>

UpdateTime :	本平台資料更新時間(ISO8601格式:yyyy-MM-ddTHH:mm:sszzzz)
UpdateInterval :	本平台資料更新週期(秒)
AuthorityCode :	業管機關簡碼
VDID :	VD設備代碼
LinkID :	基礎路段代碼, 請參閱[基礎路段代碼表] https://traffic-api-documentation.gitbook.io/traffic/xiang-dai-zhao-biao
LaneID :	車道代碼
LaneType :	車道種類 = ['1: 一般車道', '2: 快車道', '3: 慢車道', '4: 機車道', '5: 高承載車道', '6: 公車專用道', '7: 轉向車道', '8: 路肩', '9: 輔助車道', '10: 調撥車道', '11: 其他']
Speed :	平均速率偵測值, ['-99=資料異常', '-1=道路封閉']
Occupancy :	佔有率偵測值
VehicleType :	車種代碼, 請參閱[車種代碼] https://traffic-api-documentation.gitbook.io/traffic/xiang-dai-zhao-biao
Volume :	流量偵測值, -99=資料異常
speed(2) :	平均速率偵測值, -99=資料異常
Status :	設備狀態 = ['0: 正常', '1: 通訊異常', '2: 停用或施工中', '3: 設備故障']
DataCollectTime :	資料蒐集時間(ISO8601格式:yyyy-MM-ddTHH:mm:sszzzz)



交通部高速公路局VD資料庫介紹

- VDID編碼方式
 - 主線
 - VD-國道-方向-里程-主線-偵測器種類
 - VD-N3-N-393.234-M-Loop
 - 匝道
 - VD-國道-方向-里程-上或下匝道-方向-中文說明
 - VD-N1-N-110-O-SW-1-北上出口



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

```
2 #指定下載的網頁連結
3 url = "https://tisvcloud.freeway.gov.tw/history/motc20/VD/20231024/VDLive_2357.xml.gz"
```

https://tisvcloud.freeway.gov.tw/history/motc20/VD/20231024/VDLive_2357.xml.gz



VD資料下載- 單一檔案

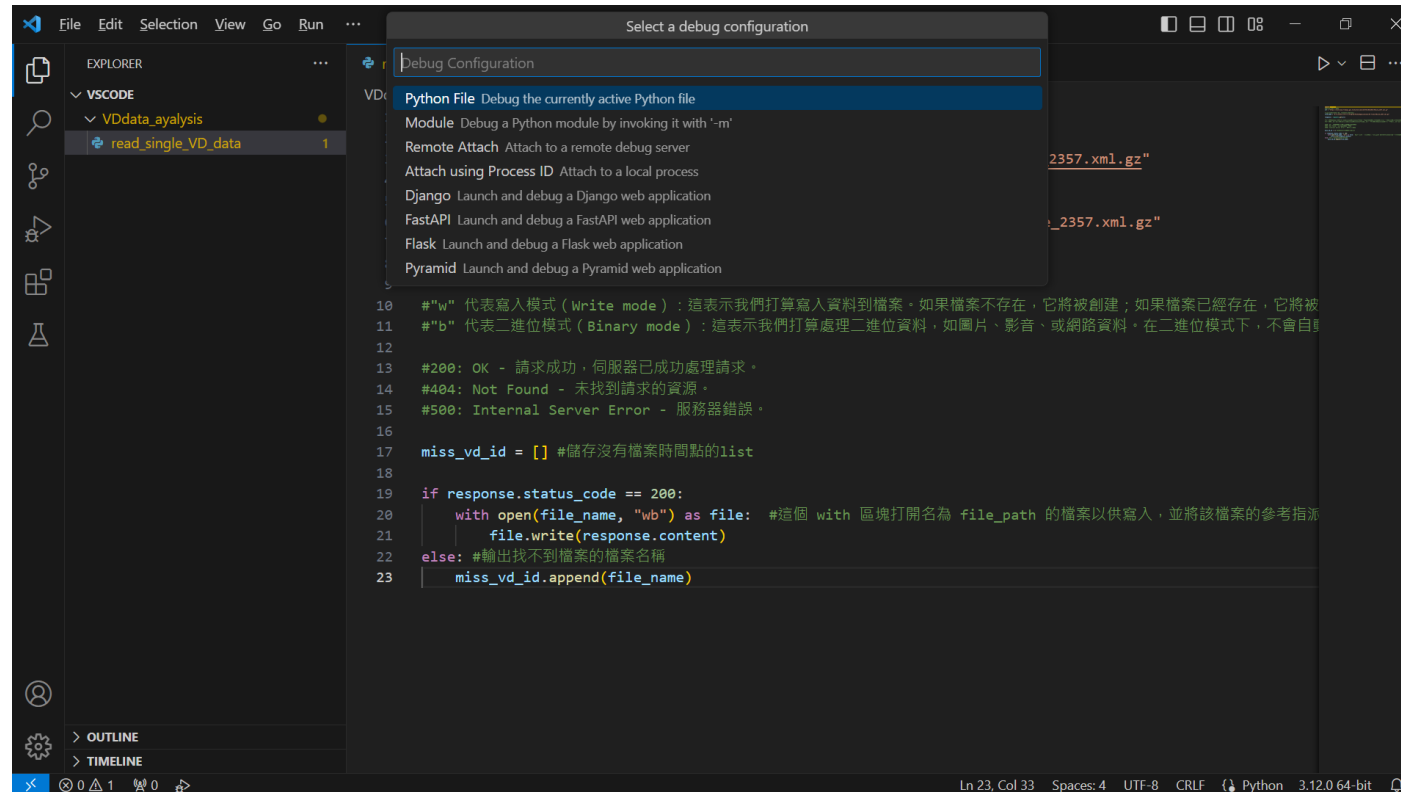
- Requests函式庫
 - 抓取網頁的資料

Requests方法	說明
requests.get(url)	向指定URL提交取得資料之請求
requests.post(url)	向指定URL提交Post之請求
requests.put(url)	向指定URL提供最新內容
requests.delete(url)	請求刪除指定的URL
requests.head(url)	請求提供url的回應標頭（不含內容）。
requests.options(url)	請求伺服器提供url可用的功能選項。



VD資料下載-單一檔案

- Terminal的設置



```
10 # "w" 代表寫入模式 (Write mode) : 這表示我們打算寫入資料到檔案。如果檔案不存在，它將被創建；如果檔案已經存在，它將被
11 # "b" 代表二進位模式 (Binary mode) : 這表示我們打算處理二進位資料，如圖片、影音、或網路資料。在二進位模式下，不會自
12
13 #200: OK - 請求成功，伺服器已成功處理請求。
14 #404: Not Found - 未找到請求的資源。
15 #500: Internal Server Error - 伺服器錯誤。
16
17 miss_vd_id = [] #儲存沒有檔案時間點的list
18
19 if response.status_code == 200:
20     with open(file_name, "wb") as file: #這個 with 區塊打開名為 file_path 的檔案以供寫入，並將該檔案的參考指派
21         file.write(response.content)
22 else: #輸出找不到檔案的檔案名稱
23     miss_vd_id.append(file_name)
```



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

PS D:\2_Class\巨集程式\Program\vscode> & 'C:\Users\Yen\AppData\Local\Microsoft\Windows\UniversalDebugger\adapter/../../debugpy\launcher' '60123' '--' 'D:\2_Class\巨集程式\Program\vscode'
讀取失敗
PS D:\2_Class\巨集程式\Program\vscode>
* History restored

Windows PowerShell
Copyright (C) Microsoft Corporation. 著作權所有，並保留一切權利。
請嘗試新的跨平台 PowerShell https://aka.ms/pscore6

PS D:\2_Class\巨集程式\Program\vscode> pip install requests
```

```
2
3 import requests
```



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

```
2 #指定下載的網頁連結  
3 url = "https://tisvcloud.freeway.gov.tw/history/motc20/VD/20231024/VDLive_2357.xml.gz"
```

```
5 response = requests.get(url)
```



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

- `open()` 函數
 - `open(file, mode, buffering, ...)`
 - `file`: 要打開的文件名稱或路徑
 - `mode`: 指定打開文件的模式。它可以是 'r' (預設，只讀模式)、'w' (寫入模式，會覆蓋現有文件，如果文件不存在則創建)、'a' (附加模式，在文件末尾寫入，如果不存在則創建)、'r+' (讀寫模式)、'b' (二進制模式)
 - `buffering`: 控制文件的緩衝設置。通常可以使用預設值 -1 或其他正整數來設置緩衝大小



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

- with 語句
 - 在 Python 中用於確保執行完特定塊代碼後釋放資源。通常搭配 as 關鍵字使用，特別是在需要管理文件、網路連接或資源時。這個語法可以自動管理資源的分配和釋放。

```
with open('example.txt', 'r') as file:
```

```
    content = file.read()
```

```
    # 在這裡使用文件內容，不需要擔心文件的關閉問題
```

```
    # 當語句執行完畢後，文件會自動關閉，即使在上面的代碼中  
    出現異常
```



VD資料下載- 單一檔案

- 伺服器收到 requests 方法所發出之請求後，會傳回一個 Response 物件
 - response 物件裡包含回應的訊息資訊，可以透過下列的屬性與方法，查詢相關內容。

Response 物件屬性	說明
url	資源的 URL 位址
content	回應訊息的內容
text	回應訊息的內容字串
raw	原始回應訊息串流
status_code	回應的狀態
encoding	回應訊息的編碼



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

```
2 #指定下載的網頁連結
3 url = "https://tisvcloud.freeway.gov.tw/history/motc20/VD/20231024/VDLive_2357.xml.gz"
```

```
5 response = requests.get(url)
```

```
7 # 指定儲存檔案的路徑和檔名 (包括附檔名)
8 file_name = "D:/2_Class/巨集程式/Program/PythonApplication1/VD files/VDLive_2357.xml.gz"
```

```
20 with open(file_name, "wb") as file: #這個 with 區塊打開名為 file_path 的檔案以供寫入
21     file.write(response.content)
```

"w" 代表寫入模式 (Write mode)：這表示我們打算寫入資料到檔案。如果檔案不存在，它將被創建；如果檔案已經存在，它將被清空並重新寫入。

"b" 代表二進位模式 (Binary mode)：這表示我們打算處理二進位資料，如圖片、影音、或網路資料。在二進位模式下，不會自動將換行符號轉換為特定的行尾符號（如 `\n` 或 `\r\n`），這對處理二進位資料非常重要。



- 練習：
 - 下載下列檔案
 - url =
"https://tisvcloud.freeway.gov.tw/history/motc20/VD/20231024/VDLive_12345.xml.gz"
 - 若是請求成功，則儲存檔案；若是不成功，則輸出 “讀取失敗”
- Hint:
 - if response.XXX == 200;
 - else:

status_code:

200: OK - 請求成功，伺服器已成功處理請求

404: Not Found - 未找到請求的資源

500: Internal Server Error - 伺服器錯誤



- 練習解答：

```
if response.status_code == 200:  
    with open(file_name, "wb") as file: #這個 with 區塊打開名為 file_path 的檔案以供寫入  
        file.write(response.content)  
else: #輸出找不到檔案的檔案名稱  
    print('讀取失敗')
```



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

- 使用到的module

- shutil

- Python 標準庫中的一個module，提供了對文件和目錄操作的各種功能，包括拷貝、移動、刪除等
 - 文件移動：shutil.move(path/file, path/file)
 - 文件拷貝：shutil.copyfileobj(path/file, path/file)
 - ...

- OS

- Python 的一個標準庫module，提供了許多與操作系統相關的功能。這些功能包括文件操作、目錄操作、環境變量設置等
 - 獲取當前工作目錄：os.getcwd()
 - 創建目錄：os.mkdir()
 - 組合多個路徑：os.path.join(path1, path2, path3,...)



VD資料下載－單一檔案

打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

- 使用到的module
 - gzip
 - Python 標準庫中的一個module，用於壓縮和解壓縮文件或數據流，通常用於處理 .gz 格式的文件
 - 壓縮文件：gzip.compress(data)
 - 打開文件：gzip.open(file,mode)
 - ...



打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

```
import shutil
import os
import gzip

# 設定.gz文件的路徑
gz_file_path = "D:/2_Class/巨集程式/Program/PythonApplication1/VD files/" # 换成你的.rar文件的存放路徑

# 設定解壓縮文件的路徑
extracted_folder_path = "D:/2_Class/巨集程式/Program/PythonApplication1/VD files unzip1/" # 替换成你想要的

# 檢查文件夾是否存在，若無，則建立之
if not os.path.exists(extracted_folder_path):
    os.makedirs(extracted_folder_path)

#指定gz文件檔名
gz_file_name = "VDLive_202310242357.xml.gz"
uncompressed_file_name = "VDLive_202310242357.xml"

# 開啟.gz文件並解壓縮至目標文件夾，且更改名稱
with gzip.open(os.path.join(gz_file_path, gz_file_name), 'rb') as gz_file, open(os.path.join(extracted_folder_path, uncompressed_file_name), 'wb') as uncompressed_file:
    shutil.copyfileobj(gz_file, uncompressed_file)

print("解壓縮並稱改名稱完成")
```



- 解壓縮既有檔案
 - `os.path.join(gz_file_path, gz_file_name)`
 - `gzip.open(os.path.join(gz_file_path, gz_file_name), 'rb')` as `gz_file`
- 指定複製目的地的檔案
 - `os.path.join(extracted_folder_path, uncompressed_file_name)`
 - `open(os.path.join(extracted_folder_path, uncompressed_file_name), 'wb')` as `uncompressed_file`
- 移動檔案
 - `shutil.copyfileobj(gz_file, uncompressed_file)`



打開網頁，找到
要下載的檔案

點選檔案

指定儲存位置並
存檔

解壓縮檔案

```
import shutil
import os
import gzip

# 設定.gz文件的路徑
gz_file_path = "D:/2_Class/巨集程式/Program/PythonApplication1/VD files/" # 换成你的.rar文件的存放路徑

# 設定解壓縮文件的路徑
extracted_folder_path = "D:/2_Class/巨集程式/Program/PythonApplication1/VD files unzip1/" # 替换成你想要的

# 檢查文件夾是否存在，若無，則建立之
if not os.path.exists(extracted_folder_path):
    os.makedirs(extracted_folder_path)

#指定gz文件檔名
gz_file_name = "VDLive_202310242357.xml.gz"
uncompressed_file_name = "VDLive_202310242357.xml"

# 開啟.gz文件並解壓縮至目標文件夾，且更改名稱
with gzip.open(os.path.join(gz_file_path, gz_file_name), 'rb') as gz_file, open(os.path.join(extracted_folder_path, uncompressed_file_name), 'wb') as uncompressed_file:
    shutil.copyfileobj(gz_file, uncompressed_file)

print("解壓縮並稱改名稱完成")
```



VD資料下載－多個檔案

- 利用for迴圈變動下載網頁連結與儲存位置

```
2 #指定下載的網頁連結  
3 url = "https://tisvcloud.freeway.gov.tw/history/motc20/VD/20231024/VDLive_2357.xml.gz"
```

需變動之處

```
#指定gz文件檔名  
gz_file_name = "VDLive_202310242357.xml.gz"  
uncompressed_file_name = "VDLive_202310242357.xml"
```



VD資料下載－多個檔案

- 字串相加：+
 str1 = "Hello, "
 str2 = "world!"
 result = str1 + str2
 print(result) # 將會print出：Hello, world!
- Integer to string
 - `str(int)`
- For loop
 - `for index in range(x1,x2)`



VD資料下載- 多個檔案

- Print出2023/01/01至2023/01/02的字串(可複製的程式碼在下一頁)

```
year_start = 2022
year_end = 2023
start_month = 1
end_month = 1
start_day = 1
end_day = 3

for year_index in range(year_start
    for month_index in range(start
        for day_index in range(sta
            print(str(year_index)+
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
y\adapter/../../debugpy\launcher' '61986' '--' 'D:\2_Class\巨集程式\Program
PS D:\2_Class\巨集程式\Program\vscode> d:; cd 'd:\2_Class\巨集程式\Program
ython-2023.20.0\pythonFiles\lib\python\debugpy\adapter/../../debugpy\launch
PS D:\2_Class\巨集程式\Program\vscode> d:; cd 'd:\2_Class\巨集程式\Program
ython-2023.20.0\pythonFiles\lib\python\debugpy\adapter/../../debugpy\launch
```

```
202211
202212
202213
202311
202312
202313
```

```
PS D:\2_Class\巨集程式\Program\vscode> █
```



VD資料下載－多個檔案

- Print出2023/01/01至2023/01/02的字串(程式碼)

```
year_start = 2022
```

```
year_end = 2023
```

```
start_month = 1
```

```
end_month = 1
```

```
start_day = 1
```

```
end_day = 3
```

```
for year_index in range(year_start, year_end + 1):
```

```
    for month_index in range(start_month, end_month + 1):
```

```
        for day_index in range(start_day, end_day + 1):
```

```
            print(str(year_index)+str(month_index)+str(day_index))
```



```
17 for year_index in range(year_start, year_end + 1):
18     for month_index in range(start_month, end_month + 1):
19
20         if month_index <= 9:
21             month_str = "0" + str(month_index)
22         else:
23             month_str = str(month_index)
24
25         for day_index in range(start_day, end_day + 1):
26
27             if day_index <= 9:
28                 day_str = "0" + str(day_index)
29             else:
30                 day_str = str(day_index)
31
32         print(str(year_index) + month_str + day_str)
```



```
if month_index <= 9:  
    month_str = "0" + str(month_index)  
else:  
    month_str = str(month_index)
```