

PYTHON OPEN DATA

抓取即時口罩數量
實作練習

吳智鴻

實作練習

用 **PYTHON**，抓取政府開放資料，
實作出即時口罩資料

目標：找出國立臺中教育大學附近的口罩資訊

參考來源：

OPEN ROBOT [如何用Python 連線政府開放資料平台之健保特約機構口罩剩餘數量明細清單 API](#)

GOOGLE COLAB

- Google提供的免費的線上AI開發工具。
- 好處：
 - 不需安裝軟體，雲端瀏覽器即可以開發
 - 套件不需要事先安裝
 - 可以運用GPU/TPU的運算資源

政府開放資料平台

- [Data.gov.tw/dataset/116285](https://data.gov.tw/dataset/116285)

政府資料開放平臺 DATA.GOV.TW Q | 智能客服 | 線上客服 | 登入/註冊

首頁 > 資料集 > 健保特約機構口罩剩餘數量明細清單

健保特約機構口罩剩餘數量明細清單

[相關資料集](#)

資料集評分: ★★★★★
平均 4.7 (16 人次投票)

資料集描述: 健保特約機構口罩剩餘數量明細清單

主要欄位說明: 醫事機構代碼、醫事機構名稱、醫事機構地址、醫事機構電話、成人口罩剩餘數、兒童口罩剩餘數、來源資料時間

資料下載網址: [CSV](#) [檢視資料](#) 健保特約機構口罩剩餘數量明細清單

提供機關: 衛生福利部中央健康保險署

提供機關聯絡人姓名: 李宛臻 (02-27065866#3608)

更新頻率: 每3分鐘

授權方式: 政府資料開放授權條款-第1版

計費方式: 免費

上架日期: 2020/02/05

資料集類型: 系統介接程式

詮釋資料更新時間: 2020/03/13 15:24

關鍵字:

主題分類: 其他

服務分類: 公共資訊

資料集分類: 開放資料

注冊應用:

檢視資料->資料下載網址

資料資源欄位:

醫事機構代碼、醫事機構名稱、醫事機構地址、醫事機構電話、成人口罩剩餘數、兒童口罩剩餘數、來源資料時間

檔案格式: CSV

編碼格式: UTF-8

資料量:

資料下載網址:

[http://data.nhi.gov.tw/Datasets/Download.ashx?rid=A210300001-D50001-001&l=<https://data.nhi.gov.tw/resource/mask/maskdata.csv>](http://data.nhi.gov.tw/Datasets/Download.ashx?rid=A210300001-D50001-001&l=https://data.nhi.gov.tw/resource/mask/maskdata.csv)

資料資源描述: 健保特約機構口罩剩餘數量明細清單

資料資源更新時間: 2020/03/13 15:25

經本平臺品質檢測後符合表格化資料之資料資源將會轉為CSV、XLSX、ODS、XML、JSON檔案格式，若所轉出之資料內容與機關所提供之原始資料有所出入，請以機關所提供之原始資料為主。

CSV 下載:

https://quality.data.gov.tw/dq_download_csv.php?nid=116285&md5_url=2150b333756e64325bdb4a5fd45fad1

XLSX 下載:

https://quality.data.gov.tw/dq_download_xlsx.php?nid=116285&md5_url=2150b333756e64325bdb4a5fd45fad1

ODS 下載:

https://quality.data.gov.tw/dq_download_ods.php?nid=116285&md5_url=2150b333756e64325bdb4a5fd45fad1

XML 下載:

https://quality.data.gov.tw/dq_download_xml.php?nid=116285&md5_url=2150b333756e64325bdb4a5fd45fad1

JSON 下載:

https://quality.data.gov.tw/dq_download_json.php?nid=116285&md5_url=2150b333756e64325bdb4a5fd45fad1

用GOOGLE COLAB來雲端開發

HTTPS://COLAB.RESEARCH.GOOGLE.COM/

- Google 搜尋Colab



The image shows a Google search interface with the search term 'colab' entered. The search results are displayed below the search bar, with a red box highlighting the top result. The top result is for 'Google Colab' from 'colab.research.google.com'. The description states: 'Colab notebooks allow you to combine executable code and rich text in a single document, along with images, HTML, LaTeX and more. When you create your ...'. Below the description, there are two columns of text: 'Google Colaboratory' and '歡迎使用Colaboratory!'. The 'Google Colaboratory' section says 'Colaboratory is a free Jupyter notebook environment that ...'. The '歡迎使用Colaboratory!' section says '借助Colaboratory, 您可以编写和执行代码、保存和共享分析结果, 以 ...'. Below the highlighted result, there is a link to 'google.com 的其他相關資訊 »'. At the bottom, there is another search result from 'medium.com' titled '透過Google Colaboratory 學習使用Python 做機器學習等科學 ...' with a description: '使用Google Colab 進行以Python 為基礎的專案研究與開發. 前言: 現在資料科學或是人工智慧 (包含機器學習等) 的話題正熱, 而想要投入這些領域研究或開發應用的 ...'.

Google

colab

全部 圖片 影片 新聞 購物 更多 設定 工具

約有 7,120,000 項結果 (搜尋時間: 0.34 秒)

colab.research.google.com ▾ 翻譯這個網頁

Google Colab

Colab notebooks allow you to combine executable code and rich text in a single document, along with images, HTML, LaTeX and more. When you create your ...

Google Colaboratory
Colaboratory is a free Jupyter notebook environment that ...

歡迎使用Colaboratory!
借助Colaboratory, 您可以编写和执行代码、保存和共享分析结果, 以 ...

[google.com 的其他相關資訊 »](#)

medium.com ▾ 透過-google-colaboratory-學習使用-python-做機器學...

透過Google Colaboratory 學習使用Python 做機器學習等科學 ...

使用Google Colab 進行以Python 為基礎的專案研究與開發. 前言: 現在資料科學或是人工智慧 (包含機器學習等) 的話題正熱, 而想要投入這些領域研究或開發應用的 ...

進入後按 **NEW NOTEBOOK**

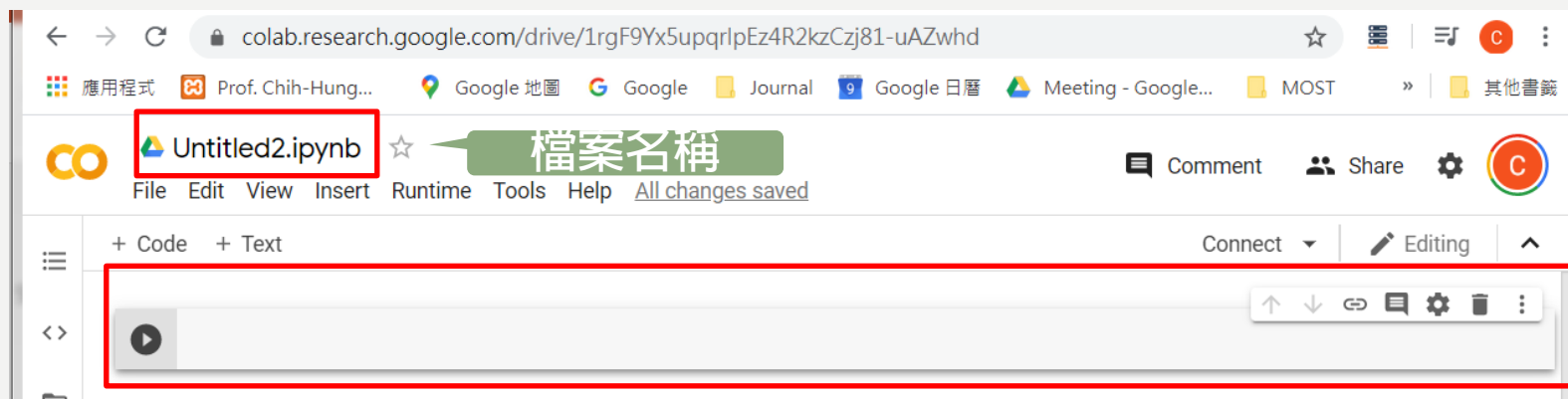
The screenshot shows the Google Colaboratory interface. At the top, it says "Welcome To Colaboratory" with a menu bar (File, Edit, View, Insert, Runtime, Tools, Help). On the left is a "Table of contents" sidebar. The main content area displays "What is Colaboratory?" with introductory text and a list of features: "Zero configuration required", "Free access to GPUs", and "Easy sharing".

An overlay window is open, showing a list of notebooks. The overlay has tabs for "Examples", "Recent", "Google Drive", "GitHub", and "Upload". The "Recent" tab is active. Below the tabs is a search bar "Filter notebooks" and a table of notebooks:

Title	First opened	Last opened	
Welcome To Colaboratory	4 minutes ago	0 minutes ago	
Taiwanmask_OK.ipynb	4 minutes ago	4 minutes ago	
object_detection_tutorial.ipynb	Jan 3, 2020	Jan 3, 2020	
CV2_examples.ipynb	Dec 22, 2019	Dec 23, 2019	
Welcome To Colaboratory	Mar 28, 2019	Dec 23, 2019	

At the bottom right of the overlay, there are two buttons: "NEW NOTEBOOK" (highlighted with a red box) and "CANCEL".

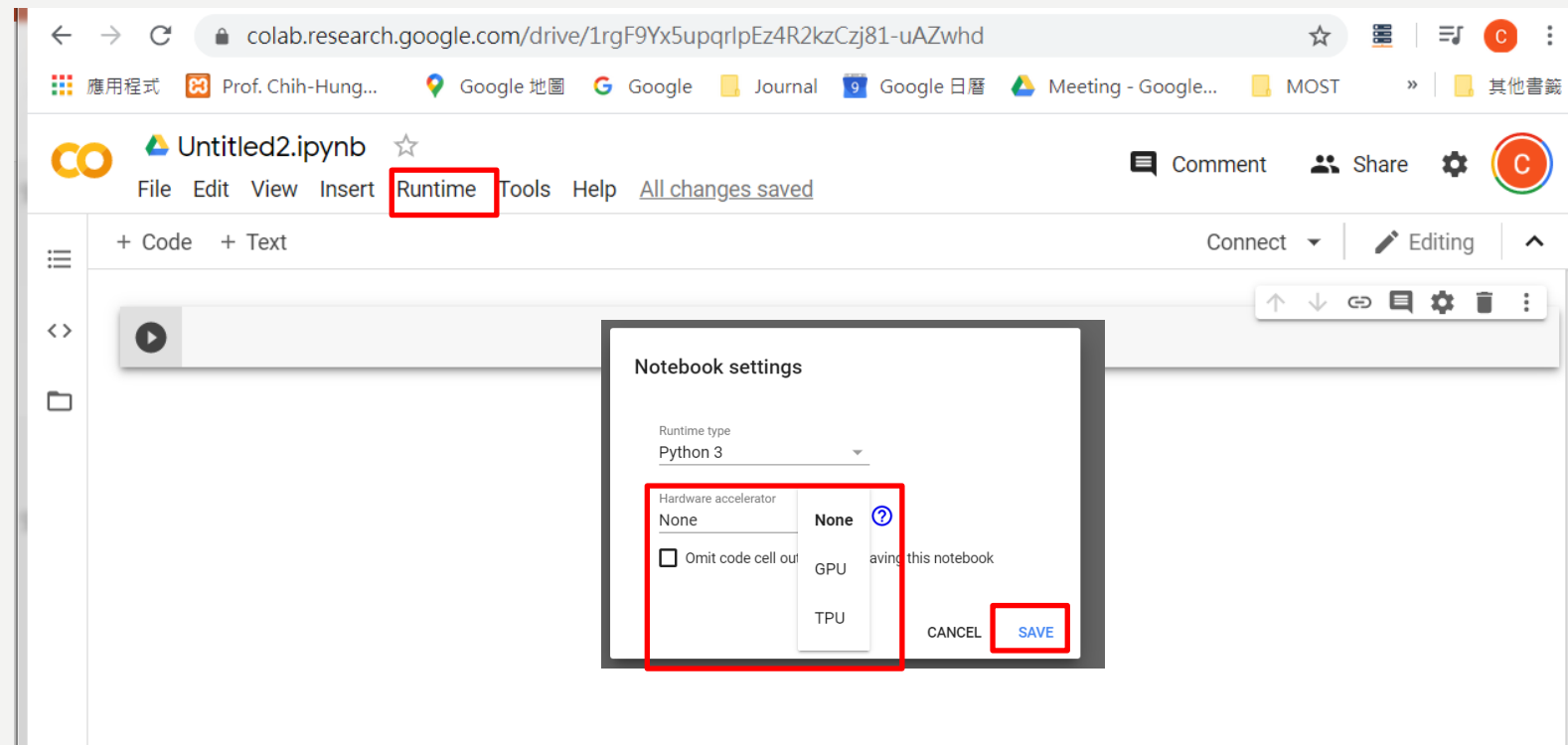
開始在雲端寫程式



把程式寫在這邊

RUNTIME -> CHANGE RUNTIME TYPE

可以選**GPU/TPU**來加速運算





程式#1 簡易程式

抓取口罩開放資料，並
顯示。

程式請參考 [TAIWANMASK-1](#)

TAIWANMASK-1

(讀入口罩即時資料OPENDATA，
並顯示)

```
import csv
import requests
import pandas as pd

#prefix = '臺北市大安區'
#prefix = '臺中市'
prefix='新竹市'

# 需設定抓取的開放資料url位置
url = "XXXXXX"

download = requests.get(url) #取得url的下載資料
decoded_content = download.content.decode('utf-8') #把下載資料編碼成utf8
cr = csv.reader(decoded_content.splitlines(), delimiter=',') #用csv.reader把資料分行並存入cr變數中

original_list = list(cr)
print(original_list)

[9] mask_df = pd.DataFrame(original_list) #把資料轉換成pandas的表格格式
mask_df.head() #列印前五筆資料
```

修改這個地方。指定分析的地區

修改這個地方。指定抓取的opendata網址

完成後畫面

- 印出前五筆資料。

口罩剩餘數

0	1	2	3	4	5	6
醫事機構代碼	醫事機構名稱	醫事機構地址	醫事機構電話	成人口罩剩餘數	兒童口罩剩餘數	來源資料時間
0145080011	衛生福利部花蓮醫院豐濱原住民分院	花蓮縣豐濱鄉豐濱村光豐路41號	(03)8358141	1508	2	2020/04/13 22:53:34
2101010013	松山健康服務中心	臺北市松山區八德路4段692號6樓	(02)27671757	5	0	2020/04/13 22:53:34
2101100227	中山健康服務中心	臺北市中山區松江路367號7樓	(02)25014616	3	2	2020/04/13 22:53:34
2101110027	內湖健康服務中心	臺北市內湖區民權東路6段99號2樓	(02)27911162	3	4	2020/04/13 22:53:34



程式#2 進階程式

可以依照地區、並依照
口罩剩餘數量由大到
小排列

程式請參閱

TAIWANMASK-2

PYTHON程式#1 把雲端程式碼貼上

```
import csv
import requests
import pandas as pd
```

匯入所需套件 (csv, request, pandas)

```
#prefix = '臺北市大安區'
#prefix = '臺中市'
prefix = '新竹市'
```

定義prefix變數，設定篩選區域

```
url = "https://data.nhi.gov.tw/resource/mask/maskdata.csv"
prefix_list = []
row_count = 0
```

設定抓取的開放資料CSV網址

```
with requests.Session() as s:
    download = s.get(url) #取得url的下載資料
    decoded_content = download.content.decode('utf-8') #把下載資料編碼成utf8
    cr = csv.reader(decoded_content.splitlines(), delimiter=',') #用csv.reader把資料分行並存入cr變數中
    original_list = list(cr) #把csv讀入的資料轉換成方便處理的list格式

    print('original_list len = %d' % (len(original_list)-1)) #列印有幾筆資料

    # 計算符合prefix 區域條件的有幾筆
    # -----
    for row in original_list:
        row_count = row_count + 1
        if row_count != 1:
            if prefix == '':
                prefix_list.append(row)
            elif row[2].startswith(prefix): #如果資料符合prefix，就把資料加入prefix_list中
                prefix_list.append(row)
    print('prefix_list len= %d' % len(prefix_list)) #列印符合條件的有幾筆
```

```
original_list len = 6158
prefix_list len= 82
```

執行結果

PYTHON程式#2

```
def myFunc(row):  
    return int(row[4]) #row[4]代表成人口罩剩餘數  
  
prefix_list.sort(key=myFunc, reverse=True) #按照口罩剩餘數由大到小排序
```

```
prefix_list.insert(0, original_list[0]) #在prefix_list資料中，加入表頭欄位資訊，方便閱讀
```

```
mask_df = pd.DataFrame(prefix_list) #把prefix_list資料轉換成pdDataFrame格式，並存入mask_df  
with pd.option_context('display.max_rows', None, 'display.max_columns', None):  
    display(mask_df) #列印mask_df資料
```

	0	1	2	3	4	5	6
0	醫事機構代碼	醫事機構名稱	醫事機構地址	醫事機構電話	成人口罩剩餘數	兒童口罩剩餘數	來源資料時間
1	5912051412	元培藥局	新竹市香山區元培街306號光曜樓1、2樓	(03)5937289	600	2235	2020/03/16 14:52:47
2	5912012120	仁美藥局	新竹市東區振興里振興路52號	(03)5625398	576	60	2020/03/16 14:52:47
3	5912040473	台大藥局	新竹市北區西大路618號一樓	(03)5282932	573	203	2020/03/16 14:52:47
4	5912041827	正大藥局	新竹市北區中正路三九〇號	(03)5356056	489	1078	2020/03/16 14:52:47
5	5912012353	康田大藥局	新竹市東區金山街102號1樓	(03)5638217	432	155	2020/03/16 14:52:47
6	5912013421	上允藥局	新竹市東區光復路1段360之17號1樓	(03)6687664	345	136	2020/03/16 14:52:47

```
# original_list
```

```
#Note: "https://data.nhi.gov.tw/resource/mask/maskdata.csv"
```

讓資料依照口罩剩餘數量由大到排序

列印資料

執行結果

程式修改地方

- 設定抓取的URL
- 設定prefix 想要篩選的地區

完成結果如下： 新竹地區即時口罩資訊

- 結果如下

```
original_list len = 6158  
prefix_list len= 82
```

0	1	2	3	4	5	6
醫事機構代碼	醫事機構名稱	醫事機構地址	醫事機構電話	成人口罩剩餘數	兒童口罩剩餘數	來源資料時間
1 5912051412	元培藥局	新竹市香山區元培街306號光曜樓1、2樓	(03)5937289	600	2235	2020/03/16 14:52:47
2 5912012120	仁美藥局	新竹市東區振興里振興路52號	(03)5625398	576	60	2020/03/16 14:52:47
3 5912040473	台大藥局	新竹市北區西大路618號一樓	(03)5282932	573	203	2020/03/16 14:52:47
4 5912041827	正大藥局	新竹市北區中正路三九〇號	(03)5356056	489	1078	2020/03/16 14:52:47
5 5912012353	康田大藥局	新竹市東區金山街102號1樓	(03)5638217	432	155	2020/03/16 14:52:47
6 5912013421	上允藥局	新竹市東區光復路1段360之17號1樓	(03)6687664	345	136	2020/03/16 14:52:47

HOMework#1 (個人作業)

恭喜你已經完成即時口罩查詢程式

- 請將此截圖張貼至FB社團Homework#I 留言下方。範例如下:

姓名：XXX 學號：XXXX 班級:XXX

- 地區：台中市

```
original_list len = 2689  
prefix_list len= 10
```

0	1	2	3	4	5	6
醫事機構代碼	醫事機構名稱	醫事機構地址	醫事機構電話	成人口罩剩餘數	兒童口罩剩餘數	來源資料時間
1 5917032551	好思康藥局	臺中市西區五權八街一一一號一樓	(04)23730826	543	66	2020/04/13 23:20:37
2 5903240394	好好誠美藥局	臺中市西區向上北路1 8 6號1樓	(04)23022838	6	0	2020/04/13 23:20:37
3 5903240358	早齋藥局	臺中市西區美村路一段3 4 2號1樓	(04)23029168	5	0	2020/04/13 23:20:37
4 5917030128	太信藥局	臺中市西區三民路一段9 5 - 1號1樓	(04)23718318	5	0	2020/04/13 23:20:37
5 5917032220	匯康藥局	臺中市西區向上路一段二七七號一樓	(04)23014038	5	2	2020/04/13 23:20:37
6 2303240013	臺中市中西區衛生所	臺中市西區民權路1 0 5號	(04)22223811	0	6	2020/04/13 23:20:37
7 5917030262	明裕藥局	臺中市西區向上路一段5 0 7號1樓	(04)24723480	0	4	2020/04/13 23:20:37
8 5917030842	向上中西藥局	臺中市西區向上路一段3 5 9號1樓	(04)24756688	0	3	2020/04/13 23:20:37
9 5917032293	祥鶴大藥局	臺中市西區美村路一段六七三號一樓	(04)23710969	0	3	2020/04/13 23:20:37
10 5917032739	捷康藥妝藥局	臺中市西區民生路1 5 9號1樓	(04)22298777	0	2	2020/04/13 23:20:37