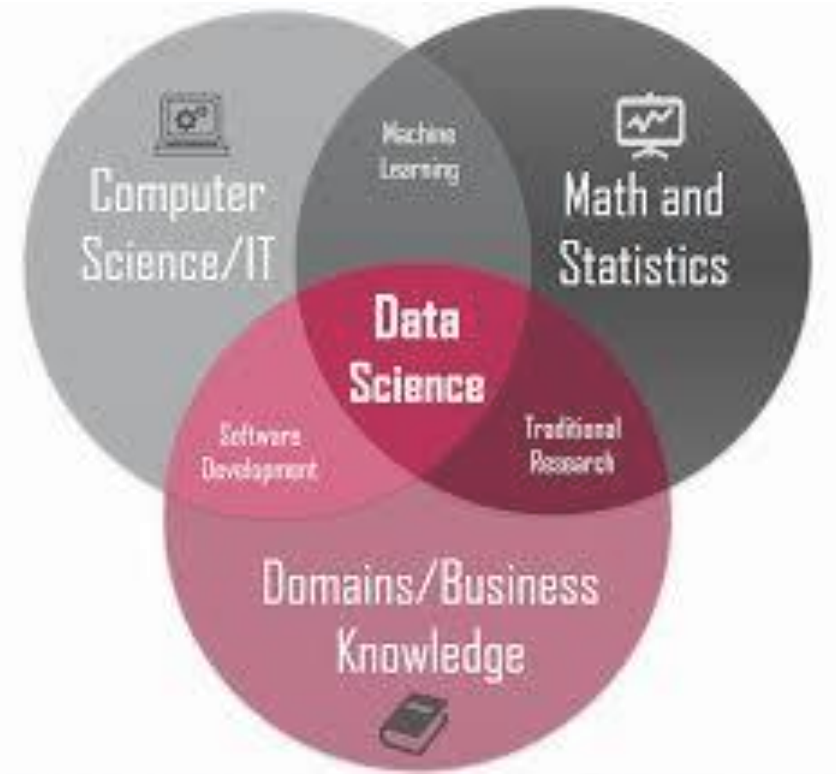


AI & DS



U01-環境安裝與設定

2023.10_V1.2

Data
Science

Artificial
Intelligence

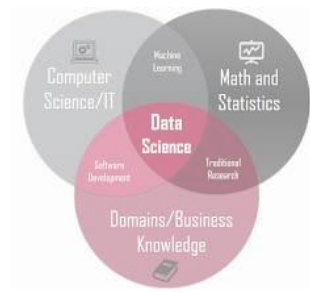
Machine
Learning

Deep
Learning

Statistics

單元大綱

- Python 機器學習IDE 安裝
 1. Anaconda 篇
 2. Python 機器學習套件
 3. Google Colaboratory篇
- Practice Labs





ANACONDA



Google Colaboratory



Visual Studio Code

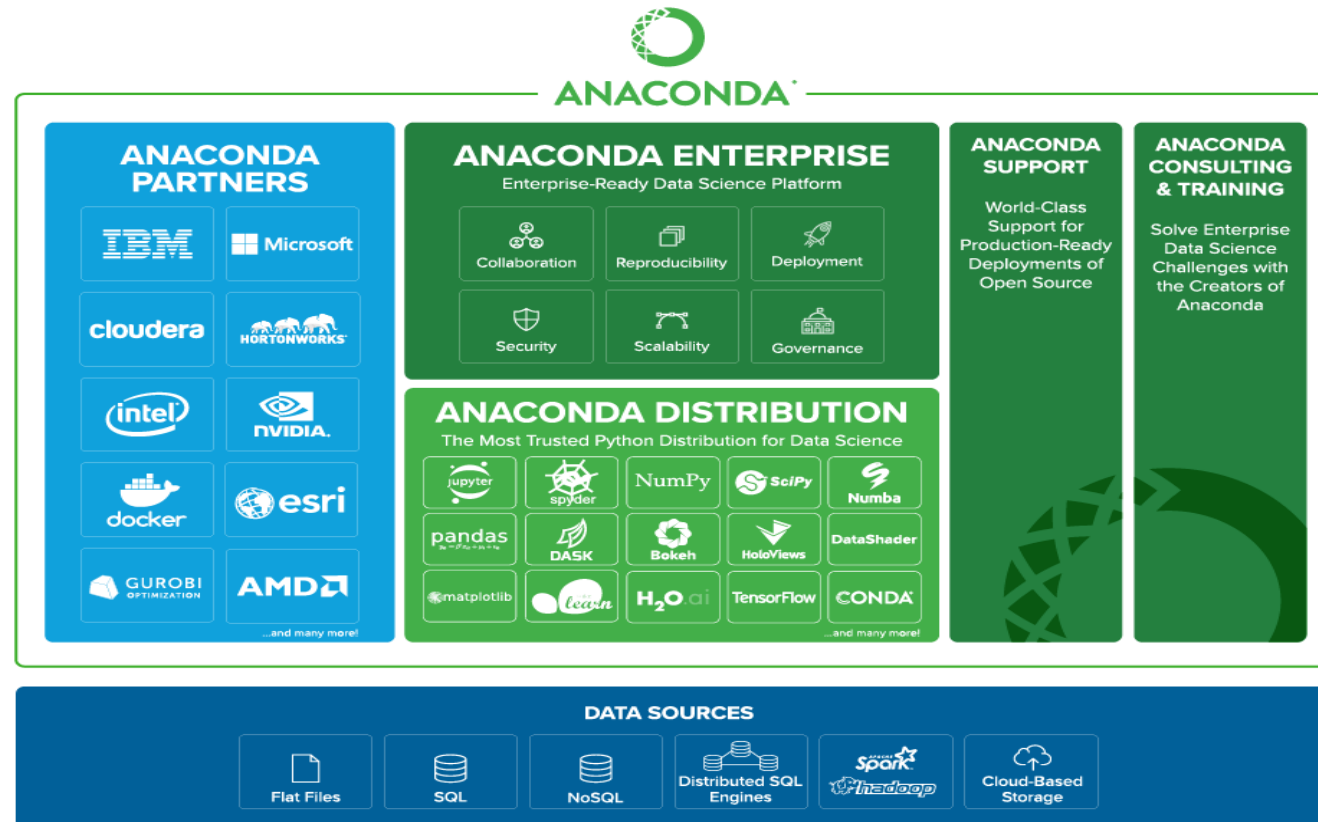
Python 機器學習IDE 安裝



ANACONDA

Anaconda 篇

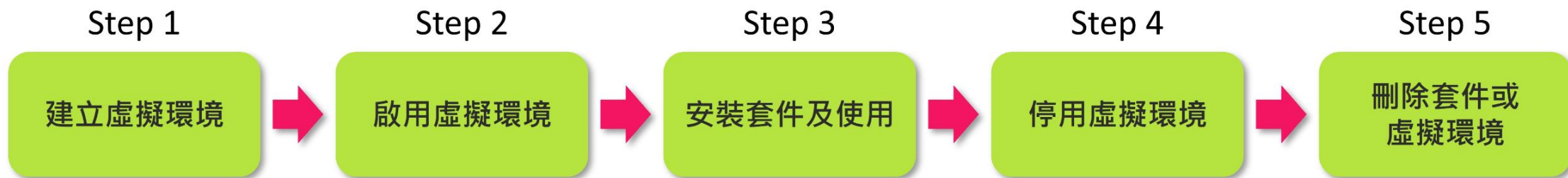
Anaconda (2023版)安裝



- 安裝參考 → <https://pse.is/4tsv2c>
- Anaconda 官網 → <https://www.anaconda.com/>
- Anaconda指令應用總整理 → <https://pse.is/5bxwkp>

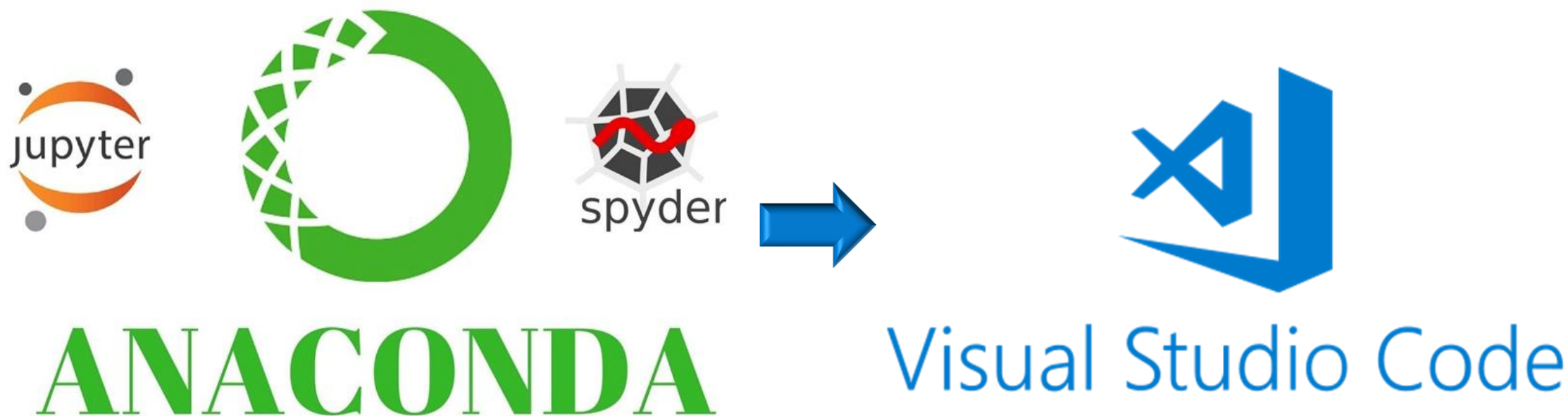
Anaconda 虛擬環境設定

- Anaconda Navigator 是 Anaconda 發行版 (Anaconda distribution) 中包含桌面圖形用戶界面 (Graphical User Interface, GUI)，它允許用戶可以在不使用輸入命令(在命令行環境中)的情況下啟動應用程序並輕鬆管理 conda 套件 (Package)、環境和通道。
- Navigator 可以在 Anaconda.org 或本地 Anaconda 儲存庫 (Repository) 中搜尋套件 (Packages)。它適用於 Windows、macOS 和 Linux。



虛擬環境建置參考 → <https://pse.is/4tvpjp>

Anaconda搭建 vscode 開發環境



搭建參考 ➔ <https://www.woodowlab.com/python-tutorial-0-anaconda/>

Miniconda -Anaconda 的輕量版

- Anaconda (適合電腦空間大，預先安裝 720 多個 Python 套件) 所占容量約 3.0 Gb
- Miniconda (適合電腦空間有限，可自由控制需要安裝的 Python 套件) 所占容量約 600 Mb
- 安裝步驟：

01 | 下載 Miniconda

02 | 安裝 Miniconda

03 | conda 自動加入 PATH (重要步驟)

04 | 確認 conda 版本以及是否安裝成功

05 | 更新 conda

06 | 安裝 Jupyter Botebook

07 | 執行 Jupyter Notebook

08 | 修改 Jupyter Notebook 預設路徑

Latest Miniconda Installer Links

Latest - Conda 4.10.3 Python 3.9.5 released July 21, 2021

Platform	Name	SHA256 hash
Windows	Miniconda3 Windows 64-bit	b33797064593ab2229a0135dc69001bea05cb56a20c2f243b1231213642e260a
	Miniconda3 Windows 32-bit	24f438e57ff2ef1ce1e93050d4e9d13f5050955f759f448d84a4018d3cd12d6b
MacOSX	Miniconda3 MacOSX 64-bit bash	786de9721f43e2c7d2803144c635f5f6e4823483536dc141ccd82dbb927cd508
	Miniconda3 MacOSX 64-bit pkg	8fa371ae97218c3c005cd5f04b1f40156d1506a9bd1d5c078f89d563fd416816
Linux	Miniconda3 Linux 64-bit	1ea2f885b4dbc3098662845560bc64271eb17085387a70c2ba3f29fff6f8d52f
	Miniconda3 Linux-aarch64 64-bit	4879820a10718743f945d88ef142c3a4b30dfc8e448d1ca08e019586374b773f
	Miniconda3 Linux-ppc64le 64-bit	fa92ee4773611f58ed9333f977d32bbb64769292f605d518732183be1f3321fa
	Miniconda3 Linux-s390x 64-bit	1faed9abecf4a4ddd4e0d8891fc2cdaa3394c51e877af14ad6b9d4aadba4e90d8

Windows installers

Windows

Python version	Name	Size	SHA256 hash
Python 3.9	Miniconda3 Windows 64-bit	58.1 MiB	b33797064593ab2229a0135dc69001bea05cb56a20c2f243b1231213642e260a
Python 3.8	Miniconda3 Windows 64-bit	57.3 MiB	8940cdd621557bc55743d6bb4518c6d343a4587127e76de808fb07e51df03fea
Python 3.7	Miniconda3 Windows 64-bit	55.8 MiB	9c031506bfc0428a0ac46c9152f9bdd48d5bdaa83046691bf8e0a4480663c05
Python 3.9	Miniconda3 Windows 32-bit	55.3 MiB	24f438e57ff2ef1ce1e93050d4e9d13f5050955f759f448d84a4018d3cd12d6b
Python 3.8	Miniconda3 Windows 32-bit	54.5 MiB	f81c165384c18d1986e2ba2f86cef384bc62266c46b34cd3d274e751ff5d91ed
Python 3.7	Miniconda3 Windows 32-bit	55.3 MiB	a1bb8338be12ee09dbd4cab9dccc2fbdcc99f65d99281dd2c07d24ad0f23dd1f7c

<https://www.1989wolfe.com/2019/07/miniCONDAwithPython.html>

Anaconda 人工智慧與機器學習套件

- 查詢所有套件: `pip list`
- 更新某一個套件: `pip install -U 套件名稱`

Ex: `pip install -U NumPy`

- Anaconda指令集參考網址: <https://reurl.cc/V5ejz5>

- 常見的AI機器學習套件:
- [Numpy](#): 向量與矩陣運算
- [Pandas](#): 資料處理分析與資料視覺化
- [Scipy](#):
- SciPy 是以 Numpy 為基礎建構的開放原始碼科學運算套件(Numpy 進階版)
- [Statsmodels](#):經典統計模型
- [Matplotlib](#): 資料視覺化
- [Seaborn](#): 以Matplotlib 為基礎建構的高階繪圖套件(Matplotlib進階版)
- [Scikit-learn](#):主流機器學習演算法套件, 包含分類, 回歸, 分群, 降維, 隨機森林, 支持向量機, KNN, K-means 等, 並提供了資料預處理, 特徵提取, 超參數最佳化, 以及模型評估等所需的函數模組

```
(base) C:\Users\C7>pip install -U Scikit-learn
Collecting Scikit-learn
  Downloading scikit_learn-1.0.1-cp38-cp38-win_amd64.whl (7.2 MB)
    7.2 MB 3.2 MB/s
Requirement already satisfied, skipping upgrade: scipy>=1.1.0 in c:\programdata\anaconda3\lib\site-packages (from Scikit-learn) (1.5.2)
Requirement already satisfied, skipping upgrade: joblib>=0.11 in c:\programdata\anaconda3\lib\site-packages (from Scikit-learn) (0.17.0)
Requirement already satisfied, skipping upgrade: threadpoolctl>=2.0.0 in c:\programdata\anaconda3\lib\site-packages (from Scikit-learn) (2.1.0)
Requirement already satisfied, skipping upgrade: numpy>=1.14.6 in c:\programdata\anaconda3\lib\site-packages (from Scikit-learn) (1.19.2)
Installing collected packages: Scikit-learn
  Attempting uninstall: Scikit-learn
    Found existing installation: scikit-learn 0.23.2
    Uninstalling scikit-learn-0.23.2:
      ERROR: Could not install packages due to an EnvironmentError: [WinError 5] 存取被拒: 'c:\programdata\anaconda3\lib\site-packages\scikit_learn-0.23.2.dist-info\COPYING'
      Consider using the '--user' option or check the permissions.

(base) C:\Users\C7>
```

使用 Jupyter Notebook 開發Python程式

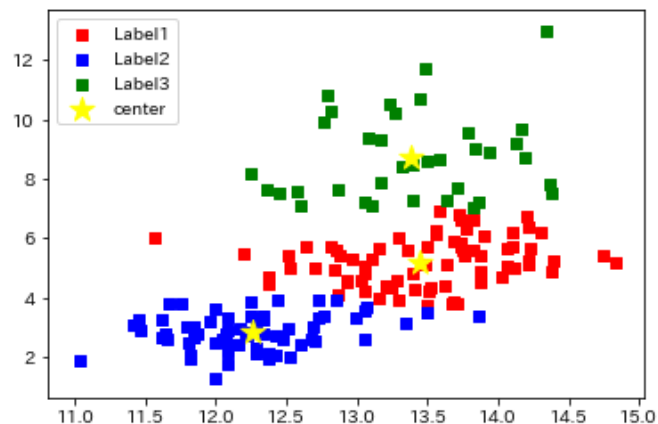
実装方法

```
In [15]: from sklearn.cluster import KMeans  
n_clusters = 3  
model = KMeans(n_clusters=n_clusters)
```

```
In [16]: pred = model.fit_predict(X)
```

結果の表示

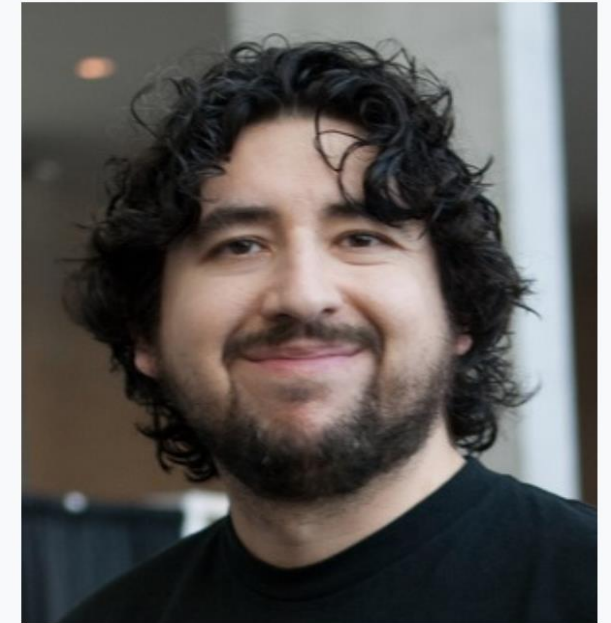
```
In [13]: %matplotlib inline  
import matplotlib.pyplot as plt  
fig, ax = plt.subplots()  
ax.scatter(X[pred==0, 0], X[pred==0, 1], color='red', marker='s', label='Label1')  
ax.scatter(X[pred==1, 0], X[pred==1, 1], color='blue', marker='s', label='Label2')  
ax.scatter(X[pred==2, 0], X[pred==2, 1], color='green', marker='s', label='Label3')  
ax.scatter(model.cluster_centers_[0], model.cluster_centers_[1], s=200, color='yellow', marker="*", label="center")  
ax.legend()  
plt.show()
```



Markdown 語法介紹

- Markdown是一種輕量級標記式語言，創始人為約翰·格魯伯。它允許人們使用易讀易寫的純文字格式編寫文件，然後轉換成有效的XHTML（或者HTML）文件。這種語言吸收了很多在電子郵件中已有的純文字標記的特性。
- 由於Markdown的輕量化、易讀易寫特性，並且對於圖片、圖表、數學式都有支援，目前許多網站都廣泛使用Markdown來撰寫說明文件或是用於論壇上發表訊息。如GitHub、Reddit、Discord、Diaspora、Stack Exchange、OpenStreetMap、SourceForge、簡書等，甚至還能被用來撰寫電子書。
- Markdown 語法介紹:
 1. Markdown語法大全 → <https://pse.is/4ua42p>
 2. 十分鐘快速掌握 Markdown → <https://pse.is/4ujtk4>

約翰·格魯伯
John Gruber



出生	1973年 (49 - 50歲)
職業	專欄作家
公民權	 美國
主題	設計、科技、蘋果公司
代表作	Markdown、《Daring Fireball》、The Talk Show、Vesper
配偶	Amy Jane Gruber
	官方網站
	daringfireball.net  (英文)



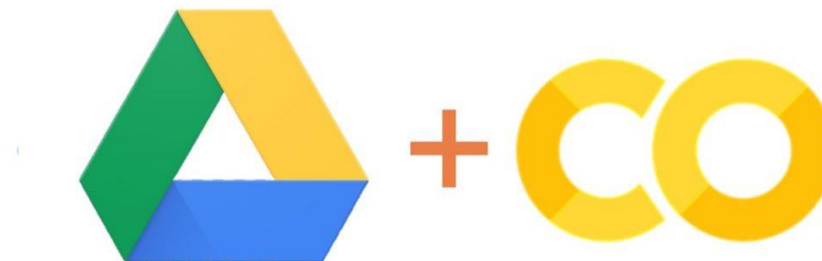
Google Colaboratory

Google Colaboratory篇

Google Colab 安裝

Colab 簡易操作介紹

- Google Colab 是 Google 所開發類似 Jupyter Notebook 的 Python 線上執行環境，它整合於 Google Drive 中，以外掛的形式存在。



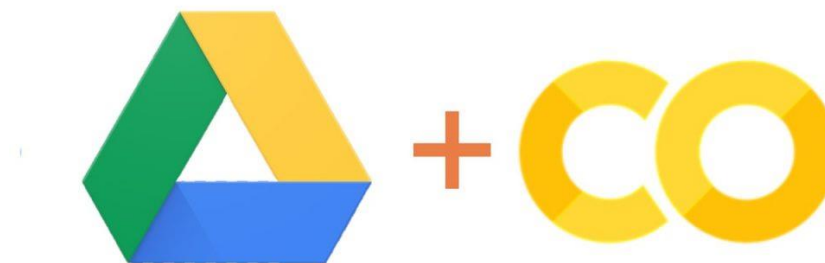
- Google Colab 的優點：
 - 跨平台、作業系統，不需要建立環境，打開瀏覽器登入 Google 就可以使用。
 - 預先安裝了常用的套件和模組，只要 import 就能使用。
 - 免費的 GPU、TPU 和記憶體可以使用，不用擔心自己的電腦等級不足以負荷。
 - 依附 Google Drive 中，方便備份、存取，也方便和他人一起協作。

Google Colab 安裝

Colab 簡易操作介紹

- Google Colab 的限制：

- 最多 12 小時的連續運算。
- 在閒置一段時間後，虛擬機會被停止並回收運算資源(包括上傳的檔案)，此時只需再重新連接，但檔案要重新上傳。



Google Colab 安裝

Colab 簡易操作介紹

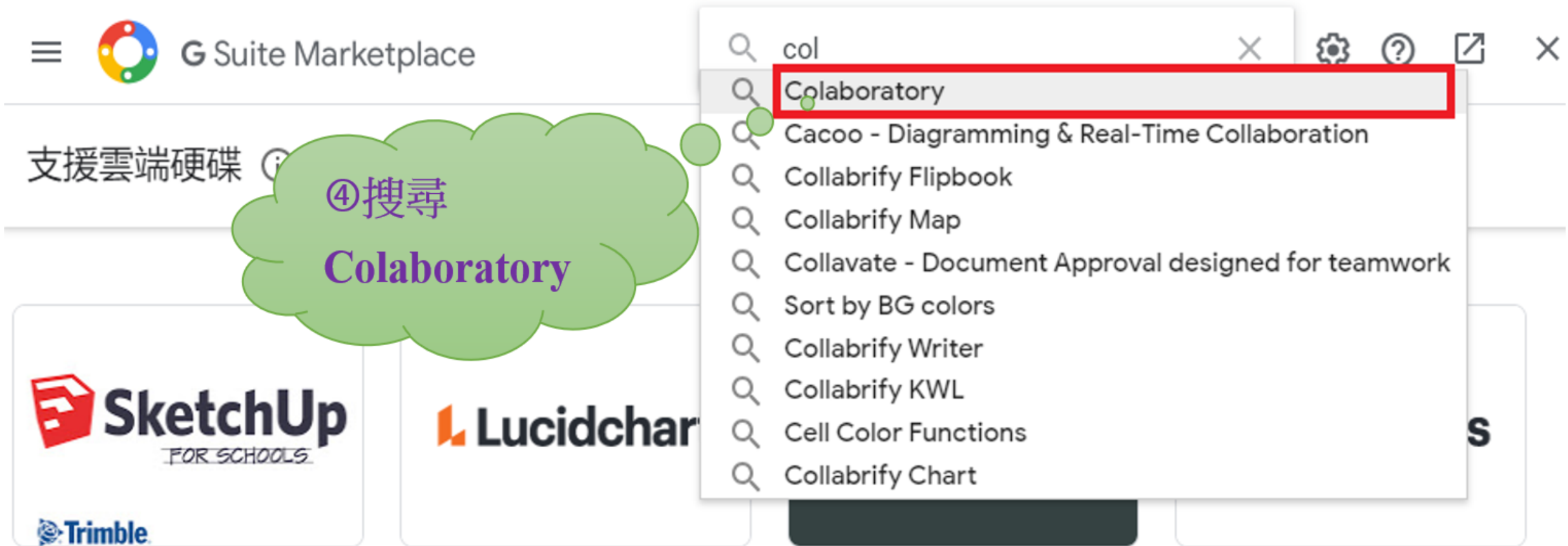
- 如何連結 Colab 應用程式：



Google Colab 安裝

Colab 簡易操作介紹

- 如何連結 Colab 應用程式：



Google Colab 安裝

Colab 簡易操作介紹

- 如何連結 Colab 應用程式：

以下查詢的搜尋結果： Colaboratory



Colaboratory

4.7 ★★★★★ (2460)

👤 1,882,448



Colaboratory ★★★★★ (2460) · 👤 1,882,448

colab-team

☁️ 雲端硬碟外掛程式

安裝

Google Colab 安裝

Colab 簡易操作介紹

- 如何連結 Colab 應用程式：

CO 可以開始安裝了

⑦繼續

[CO] 您已授權安裝。

點選 [繼續]，即表示您瞭解這個應用程式會根據相關的[服務條款](#)和[隱私權政策](#)使用您的資訊。

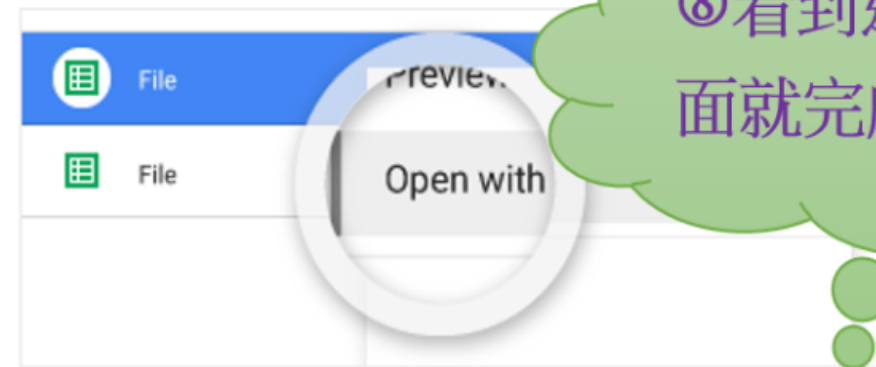
取消

繼續

CO 您已成功安裝「Colaboratory」！

「Colaboratory」啟動位置

在雲端硬碟中選擇 [選擇開啟工具] 即可找到「Colaboratory」：



⑧看到這個畫面就完成了

完成

Colab 簡易操作介紹

- Colab 掛載雲端硬碟資料夾：

- 在 Colab 中運行 Python 程式如果需要讀取資料集檔案或字型，透過 Google 提供的套件，我們可以讓 Colab 上的程式直接讀取自己的雲端硬碟。
 - ◆ 優點：檔案存在於自己的雲端硬碟，就隨時都可以直接存取，不用每次上傳。
 - ◆ 缺點：程式運行時要輸入連結 Google Drive 所需要的授權碼。

準備實作環境

Colab 簡易操作介紹

● Colab 掛載雲端硬碟資料夾：

■ 步驟：



Colab 簡易操作介紹

- 當開啟或新增一個 Colab 檔案，其附檔名為 ipynb，雲端伺服器將自動啟動一個獨立執行程序 RunTime。
- Menu 中的 RunTime 功能包含管理這個執行程序的命令。
- 當執行程序重新啟動 Restart 時，所有與導入套件、變量賦值等相關的內存都會丟失。
- 若發生無預期錯誤，可以 Restart RunTime，重新再執行程式。

執行階段	工具	說明
全部執行		Ctrl+F9
執行上方的儲存格		Ctrl+F8
執行聚焦的儲存格		Ctrl+Enter
執行選取範圍		Ctrl+Shift+Enter
執行下方的儲存格		Ctrl+F10
中斷執行		Ctrl+M I
重新啟動執行階段		Ctrl+M .
重新啟動並執行所有儲存格		
恢復原廠設定的執行階段		

Google Colab 安裝

Colab 簡易操作介紹

- ipynb 筆記本由文本 Text 和代碼 Code 兩種“單元格”組成：
 - 文本 Text：可以編輯程式重要的說明。
 - 代碼 Code：程式的主體，點擊筆記本左邊的運行按鈕或鍵盤快捷鍵[CTRL+ENTER]，將執行該單元格中的程式碼，執行完畢會顯示執行編號數值。

```
import numpy as np
from IPython import display
import matplotlib.pyplot as plt
import matplotlib.font_manager as plt_font
twfont1 = plt_font.FontProperties(fname="drive/My Drive/colab/kaiu.ttf")
```

TT B I <> 🔗 🖼️ 📄 📋 📌 ⋮

(2) 將原始資料分為訓練資料集和驗證資料集

(2) 將原始資料分為訓練資料集和驗證資料集

PRACTICE



Practice Labs

Lab01: 建立Anaconda虛擬環境

1. 在C:\使用者\ 建立一資料夾:
[MyPyProject](#)
2. 在Anaconda中設定一新的虛擬環境:
(1)名稱: [Test_Matplot](#)
(2) 新增: **Matplotlib** 套件
(3) 設定為現行使用環境
3. 使用Jupyter Notebook([Test_Matplot](#))
撰寫Python程式碼(如左圖),並
分別存檔為:
[MyFirstML.ipynb](#)
與 [MyFirstML.py](#)
4. 搭設Anaconda虛擬環境與
VSCode連結, 並開啟
[MyFirstML.py](#) 執行

+ 程式碼 + 文字



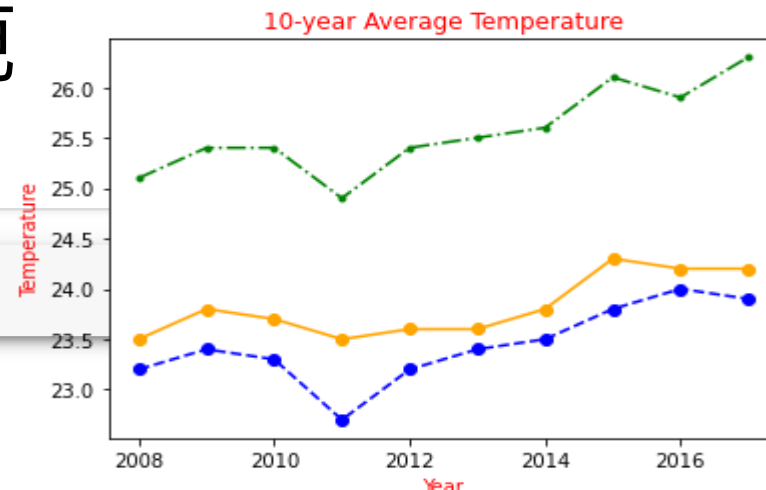
```
1 import matplotlib.pyplot as plt
```

輸入 matplotlib 套件, 並更名為 plt


```
[ ] 1 Taipei_temp = [23.2, 23.4, 23.3, 22.7, 23.2, 23.4, 23.5, 23.8, 24, 23.9]
    2 Taichung_temp = [23.5, 23.8, 23.7, 23.5, 23.6, 23.6, 23.8, 24.3, 24.2, 24.2]
    3 Kaohsiung_temp = [25.1, 25.4, 25.4, 24.9, 25.4, 25.5, 25.6, 26.1, 25.9, 26.3]
    4 year = range(2008, 2018)
    5
```

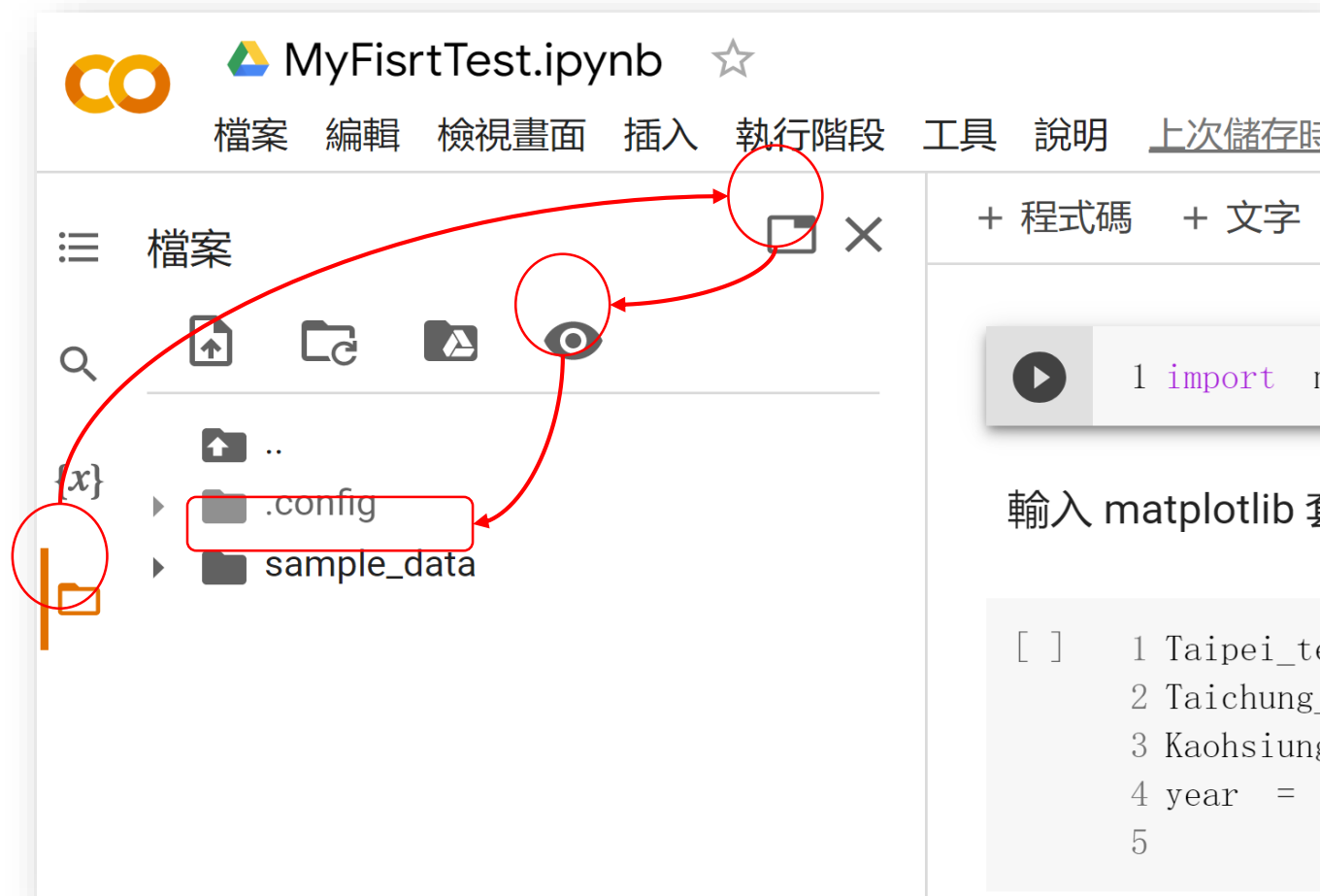
設定 Taipei, Taichung, Kaohsiung 三個城市, 10 個月份的溫度。並設定年度(year)

```
[ ] 1 plt.plot(year, Taipei_temp, color = 'blue', marker='o', linestyle = '--')
    2 plt.plot(year, Taichung_temp, color = 'orange', marker='o', linestyle = '-')
    3 plt.plot(year, Kaohsiung_temp, color = 'green', marker='.', linestyle = '-.')
    4 plt.xlabel('Year', color = 'red')
    5 plt.ylabel('Temperature', color = 'red')
    6 plt.title('10-year Average Temperature', color = 'red')
    7 plt.show()
```



Lab02(更新): 建立Google Colaboratory環境

1. 開啟Google 雲端硬碟
2. 建立一新資料夾: MyPyProj
3. 開啟 Google Colaboratory
4. 上傳 MyFirstML.ipynb執行
5. [檔案]→[移動]，將本檔從原目錄: Colab Notebooks 移動至新目錄: MyPyProj
6. 開啟Colab側邊欄的  (檔案), 查看未掛接雲端硬碟前的檔案結構



Lab02(更新): 建立Google Colaboratory環境

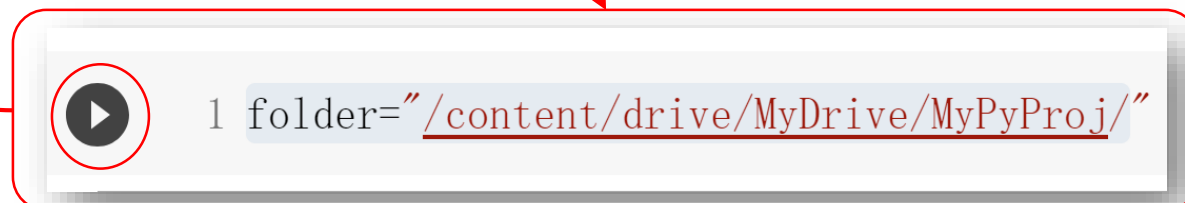
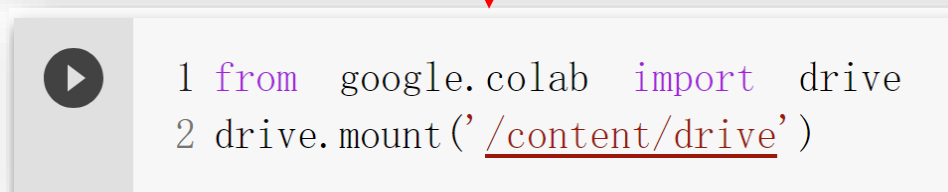
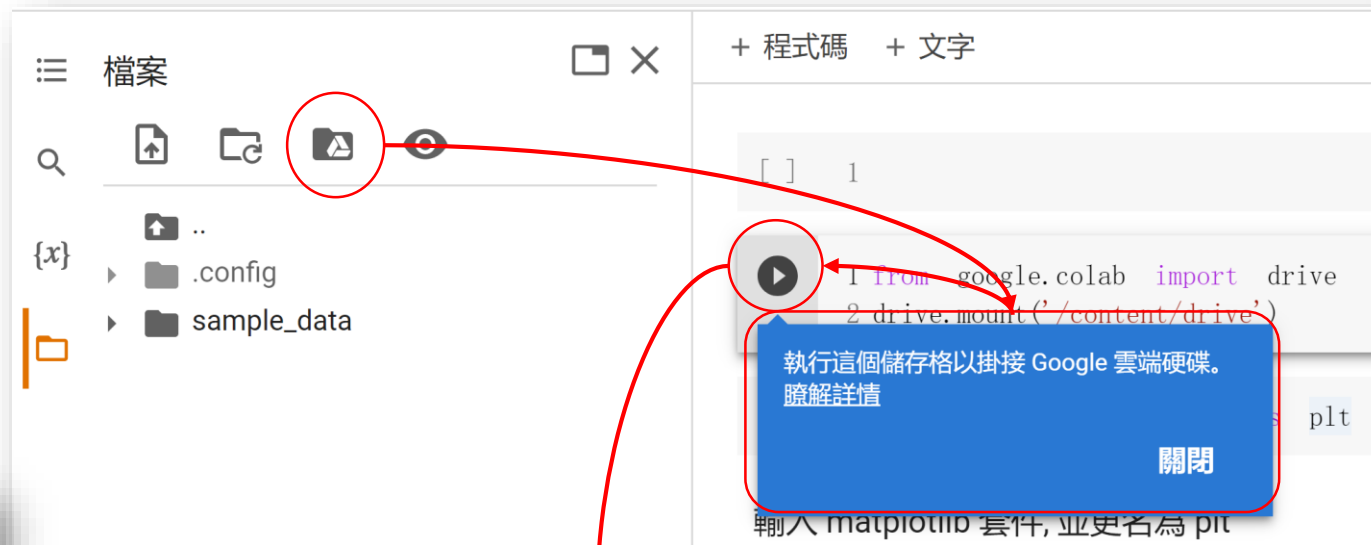
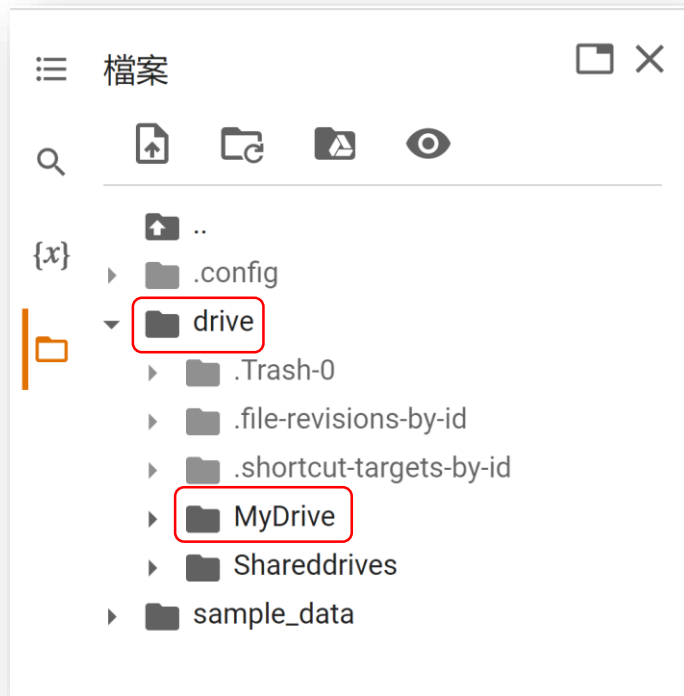
7. 按  進行雲端硬碟掛接

8. 再新增一行[程式碼]

9. 輸入:

`folder="/content/drive/MyDrive/MyPyProj/ "`

10. 再查看目錄結構



Lab03: 建立本書專用Colab開發執行環境

1. 開啟Google 雲端硬碟
2. 建立一新資料夾: [資料科學自學聖經](#)
3. 進入[資料科學自學聖經](#)資料夾後，將所下載之範例，從Ch02-Ch14都上傳至此。

[注意] Ch07內容所佔據空間有656MB, 上傳時間較久!!

4. 上傳完畢後，請進入Ch02, 以Colab開啟檔案:
「[資料科學神器：Numpy與Pandas.ipynb](#)」
5. 依照本教材存放於雲端硬碟的檔案路徑,進行雲端硬碟掛載與資料夾路徑設定。