

## Cài đặt môi trường sử dụng thuật toán phân cụm

\*Note\*: Để cài đặt được Python trên Visual Studio Code, người dùng có thể tham khảo theo những hướng dẫn sau:

- + Truy cập đường link sau và chọn phiên bản phù hợp với máy tính của người dùng sau đó tải về: <https://code.visualstudio.com>
- + Hướng dẫn cài đặt Python bản mới nhất trên Visual Studio Code người dùng có thể tham khảo theo đường link sau: <https://tek4.vn/khoa-hoc/lap-trinh-python-can-ban/cai-dat-moi-truong-phat-trien-python-tren-visual-studio-code>

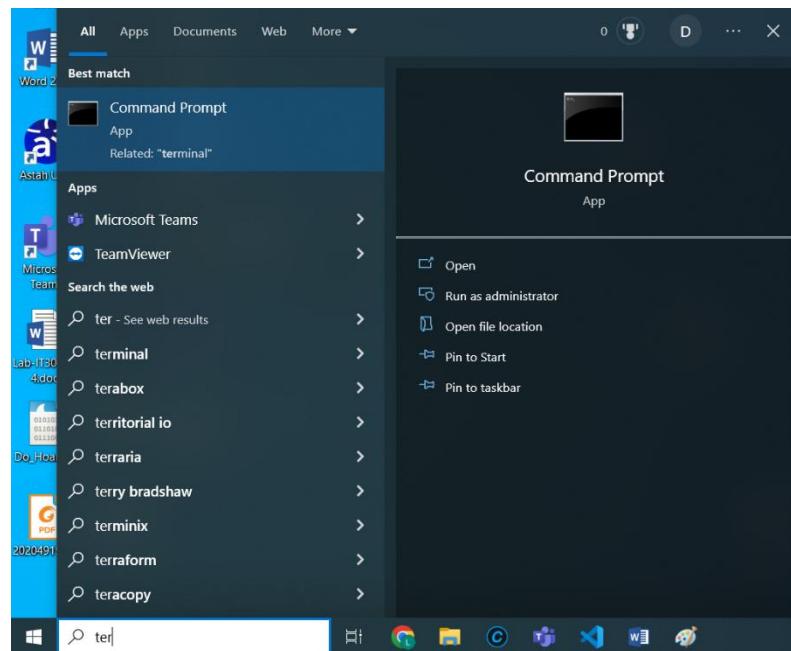
Sau khi đã hoàn thành cài đặt Visual Code Studio và Python, người dùng có thể thực hiện những bước sau đây để chạy thuật toán

### 1. Chuẩn bị thư viện

Trong thuật toán này, chúng em xây dựng dựa trên những thư viện sau:

- pandas
- math
- matplotlib
- tqdm

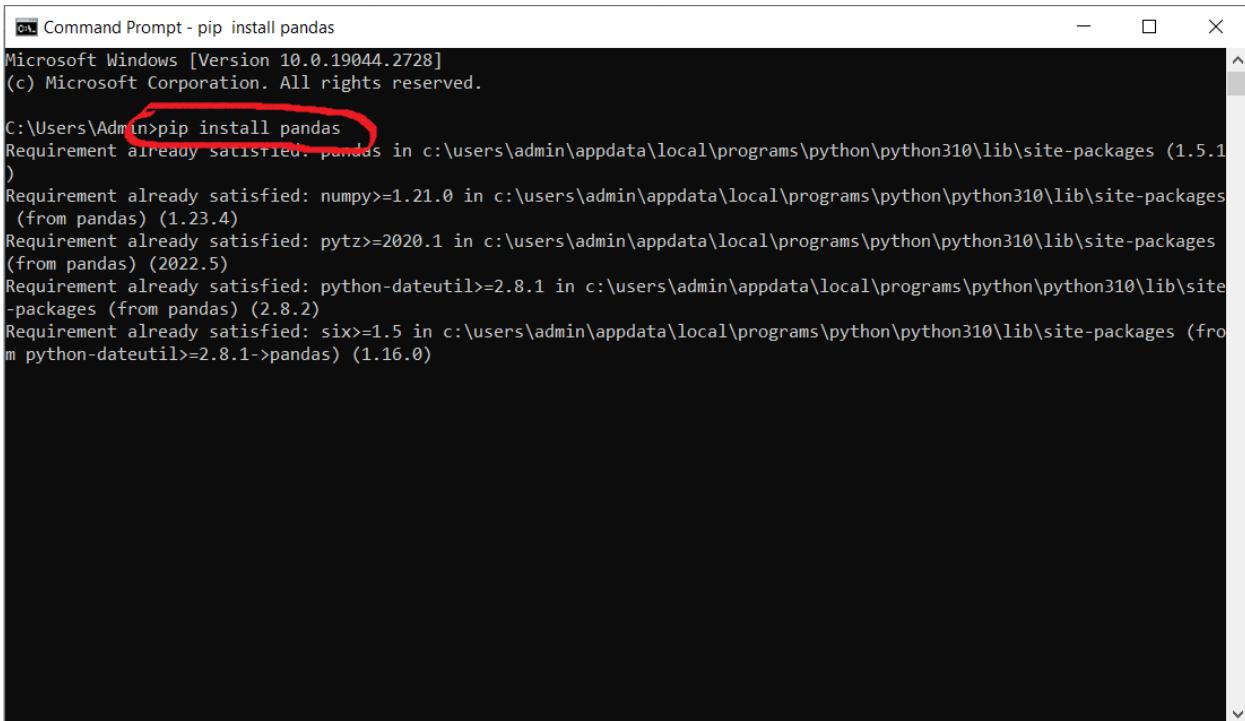
Để cài đặt thư viện, mọi người có thể vào ô tìm kiếm của Windows, gõ tìm Command Prompt hoặc Terminal rồi sau đó chọn Command Prompt:



Sau khi đã chọn vào Command Prompt, mọi người có thể cài đặt thư viện bằng cách gõ câu lệnh sau đây:

pip install “tên thư viện ở bên trên”

Ví dụ nếu muốn cài đặt thư viện pandas:



```
Command Prompt - pip install pandas
Microsoft Windows [Version 10.0.19044.2728]
(c) Microsoft Corporation. All rights reserved.

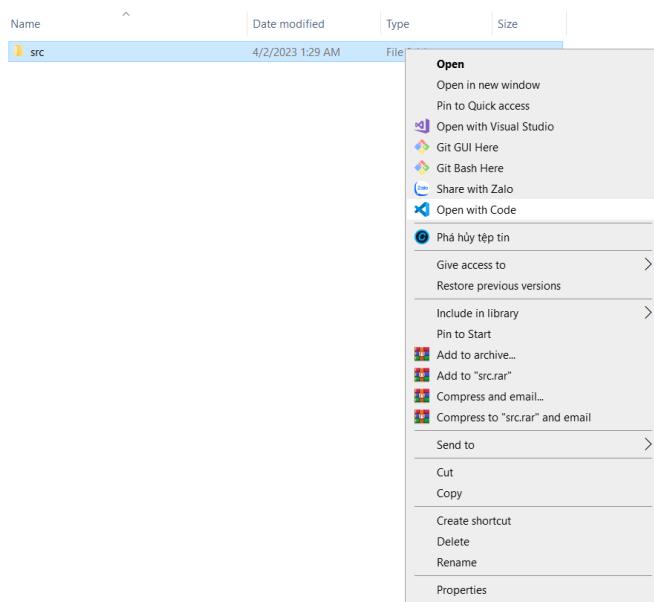
C:\Users\Admin>pip install pandas
Requirement already satisfied: pandas in c:\users\admin\appdata\local\programs\python\python310\lib\site-packages (1.5.1)
Requirement already satisfied: numpy>=1.21.0 in c:\users\admin\appdata\local\programs\python\python310\lib\site-packages (from pandas) (1.23.4)
Requirement already satisfied: pytz>=2020.1 in c:\users\admin\appdata\local\programs\python\python310\lib\site-packages (from pandas) (2022.5)
Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\admin\appdata\local\programs\python\python310\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\admin\appdata\local\programs\python\python310\lib\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
```

Sau khi lần lượt cài đặt xong hết tất cả các thư viện ở trên, chúng ta sẽ đi vào phần chạy thuật toán.

## 2. Chạy thuật toán

Bước 1: Tải toàn bộ folder Source Code trên Google Drive về trên máy tính

Bước 2: Nhấn chuột phải vào Folder “Source Code” vừa tải về và chọn “Open with Code” với biểu tượng của Visual Code Studio bên cạnh.



Bước 3: Sau khi thực hiện các bước trên, bây giờ chúng ta chỉ cần nhập vào lamda và filename để có thể chạy chương trình.

Sau khi đã nhập đầy đủ input, nhấn vào đây để chạy

Nhập filename là tên tập dataset

Nhập lambda để chạy thuật toán

Các file trong source code

Sau khi đã nhấn nút “Run Python File” thì thuật toán sẽ được thực hiện và in ra số cụm như sau:

```
PS C:\Users\Admin\Desktop\Project_2\src> & C:/Users/Admin/AppData/Local/Programs/Python/Python310/python.exe "c:/Users/Admin/Desktop/Project_2/src/Clustering.py"
Computing matrix S: 100%
Computing matrix M: 100%
Epoch 1: 100%
-----Checking-----
-----Done-----
Epoch 2: 100%
-----Checking-----
-----Done-----
Checking completely!!!
100%
Time: 68.18341016769409
5
PS C:\Users\Admin\Desktop\Project_2\src>
```

Trên đây là quá trình hướng dẫn thực thi thuật toán, nếu có gì bổ sung thì chúng em sẽ cập nhật sớm nhất.