

TRỰC QUAN HOÁ DỮ LIỆU

PHÂN TÍCH THÔNG KÊ DỮ LIỆU NHIỀU BIẾN



Hướng dẫn: Nguyễn Mạnh Hùng

MỤC LỤC

I.	THÔNG TIN	2
II.	MÔI TRƯỜNG	2
III.	HƯỚNG DẪN CÀI ĐẶT	2
IV.	THỰC NGHIỆM	4
V.	TỔNG KẾT	5
VI.	THAM KHẢO	5

I. THÔNG TIN

Cá Nhân

MSSV	HỌ TÊN	EMAIL	SĐT
19127517	Hồ Thiên Phước	htphuoc19@clc.fitus.edu.vn	-

II. MÔI TRƯỜNG

CÔNG CỤ	NGÔN NGỮ
Visual Studio Code	Python

III. HƯỚNG DẪN CÀI ĐẶT

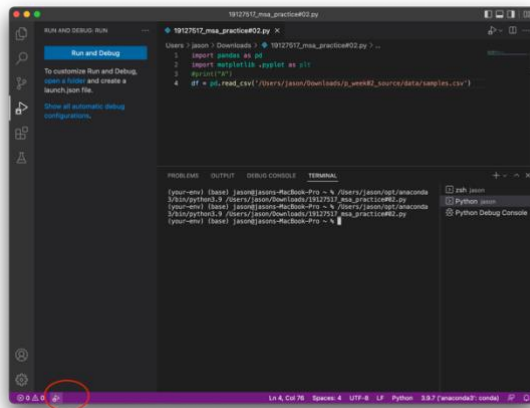
1. Cài đặt pip

```
python get-pip.py  
hoặc  
python -m pip install --upgrade pip
```

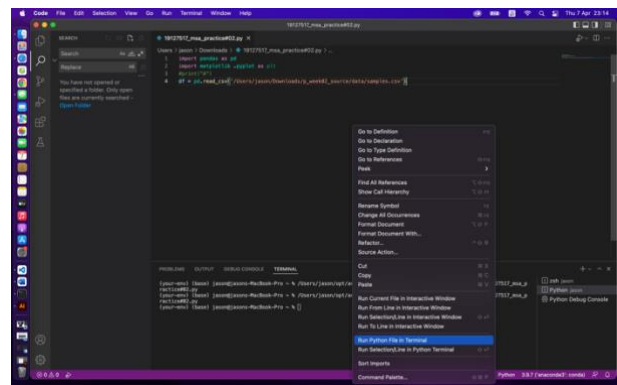
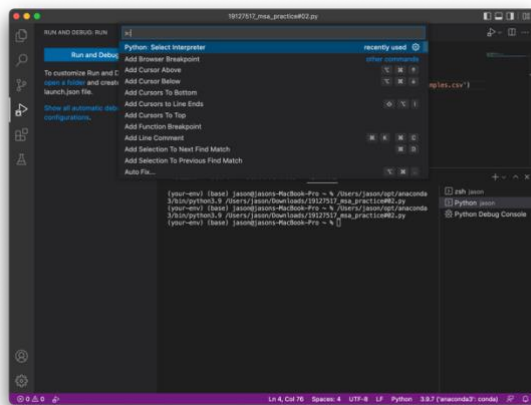
2. Cài đặt venv

```
pip install virtualenv  
virtualenv your-env  
source your-env/bin/activate  
pip install numpy  
pip install pandas  
pip install matplotlib
```

MSA-LAB02: TRỰC QUAN HOÁ DỮ LIỆU



Sau khi tạo env, chọn interpreter. (Lưu ý không cần cài extension Python)



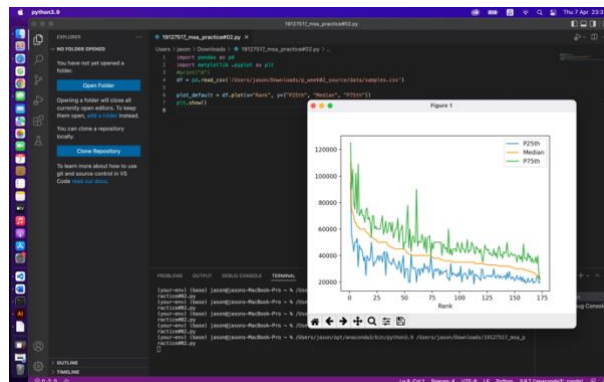
Chọn path your-env/bin/python3.9

Chạy code có import thư viện và không báo lỗi

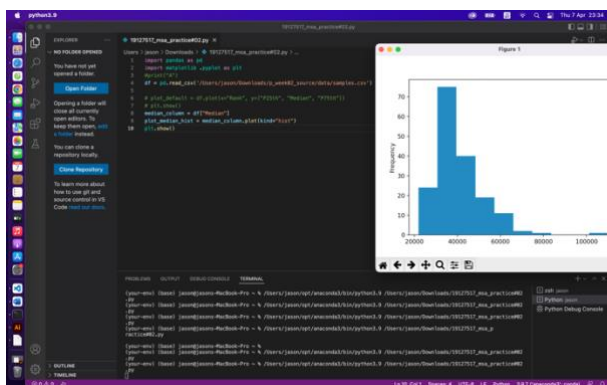
[illegible]

Tải sample.csv <https://courses.ctda.hcmus.edu.vn/mod/resource/view.php?id=51437> và sửa lại path

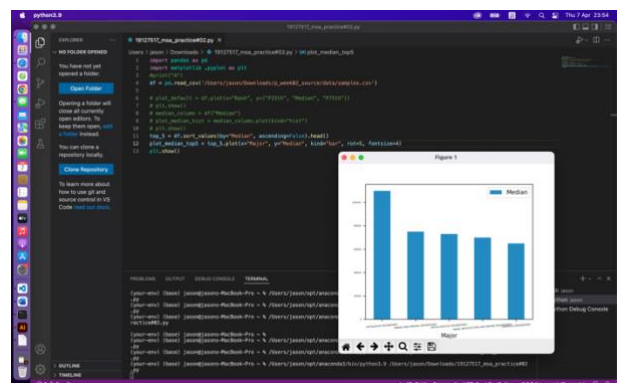
IV. THỰC NGHIỆM



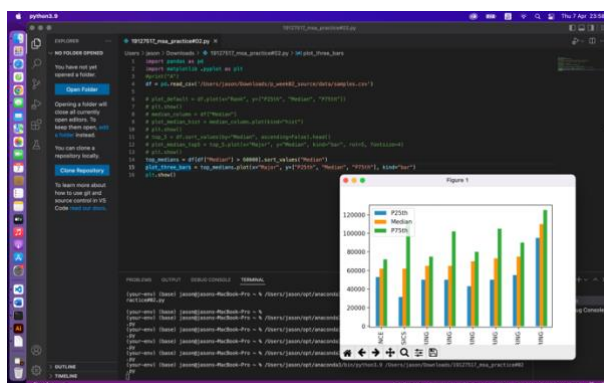
⌘+S trước khi chạy code ở Terminal. Biểu đồ mặc định với info chọn sẵn



Biểu đồ histogram

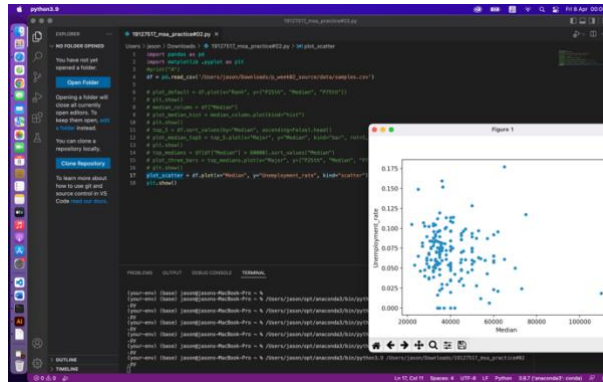


Sort và lấy top 5 giá trị

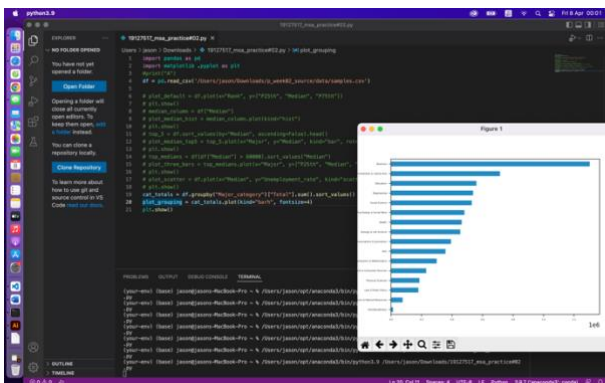


Biểu đồ cột tam

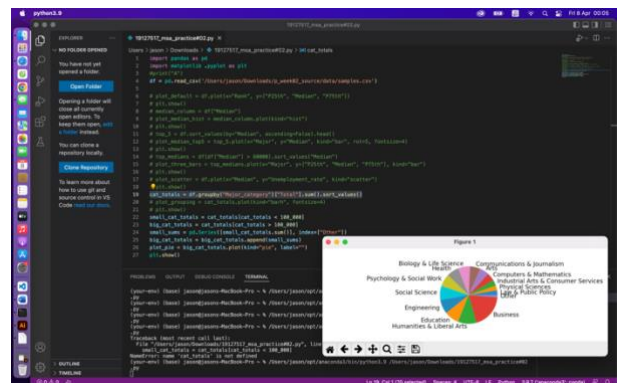
MSA-LAB02: TRỰC QUAN HOÁ DỮ LIỆU



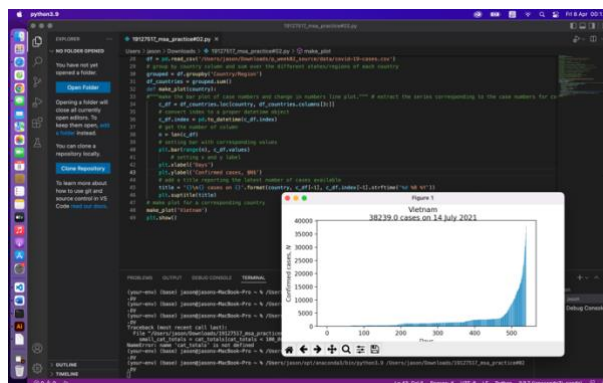
Scatter với điểm trung vị



Gom nhóm



Mục bổ sung Other là tổng các mục nhỏ



covid-19-cases.csv

V. TỔNG KẾT

Qua bài thực hành này, em được tiếp cận thêm một công cụ Visual Studio Code và file .csv mà đồ giờ em không biết xài.

VI. THAM KHẢO

[1] <https://courses.ctda.hcmus.edu.vn/mod/resource/view.php?id=51436>