LINK:  
  
<https://www.youtube.com/watch?v=j7fNG-V4aGE>

Purpose/Aim:  
  
Simple Sanction screening: sanction screening service is a tool that helps companies comply with regulations related to AML and KYC by detecting financial crimes (Sanction Scanner).

In essence, we need to:

Ensure that our business does not transact with anyone listed.

To achieve this, we need to implement a screening system to flag possible persons or entities with sanctions.

We will create a data pipeline which:

Extracts the Sanction list and used as our reference of the people that we should not be transacting with.

SOURCE: Office of Foreign Assets Control (OFAC) <https://ofac.treasury.gov/>

<https://ofac.treasury.gov/specially-designated-nationals-list-data-formats-data-schemas>

<https://ofac.treasury.gov/media/29976/download?inline>

FTP – File Transfer Protocol

Way to download, upload, and transfer files from one location to another on the internet and between computer systems.

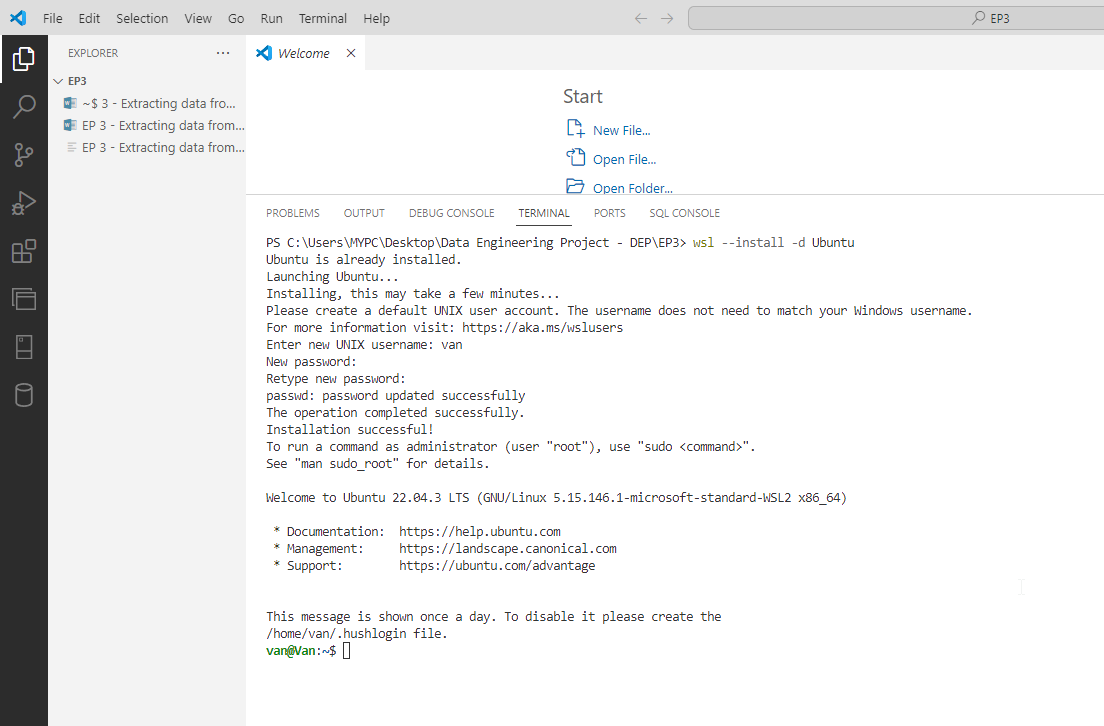
We will setup an FTP server on WSL using vsftpd.

Setting up WSL

PS C:\Users\MYPC\Desktop\Data Engineering Project - DEP\EP3> wsl –list

PS C:\Users\MYPC\Desktop\Data Engineering Project - DEP\EP3> wsl –install –d Ubuntu

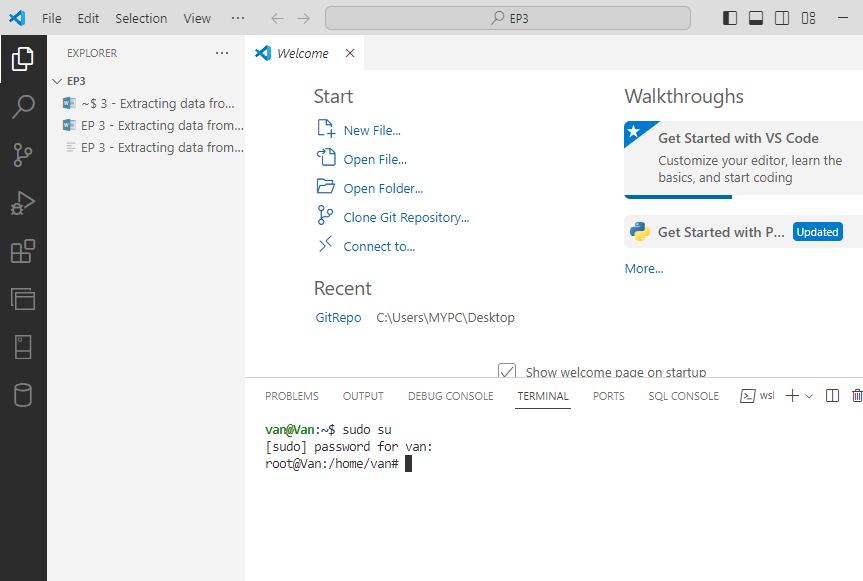
User: van  
Password: password



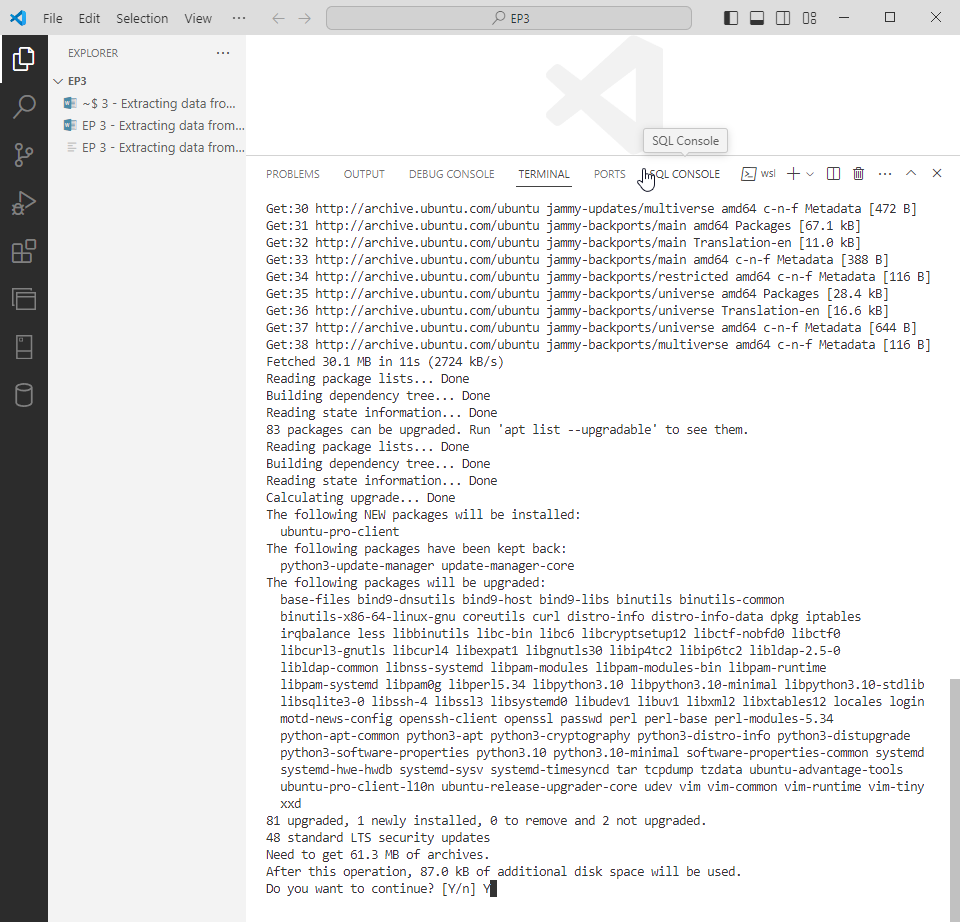
Linux commands will be used: wsl

Run   
van@Van:~$ sudo su

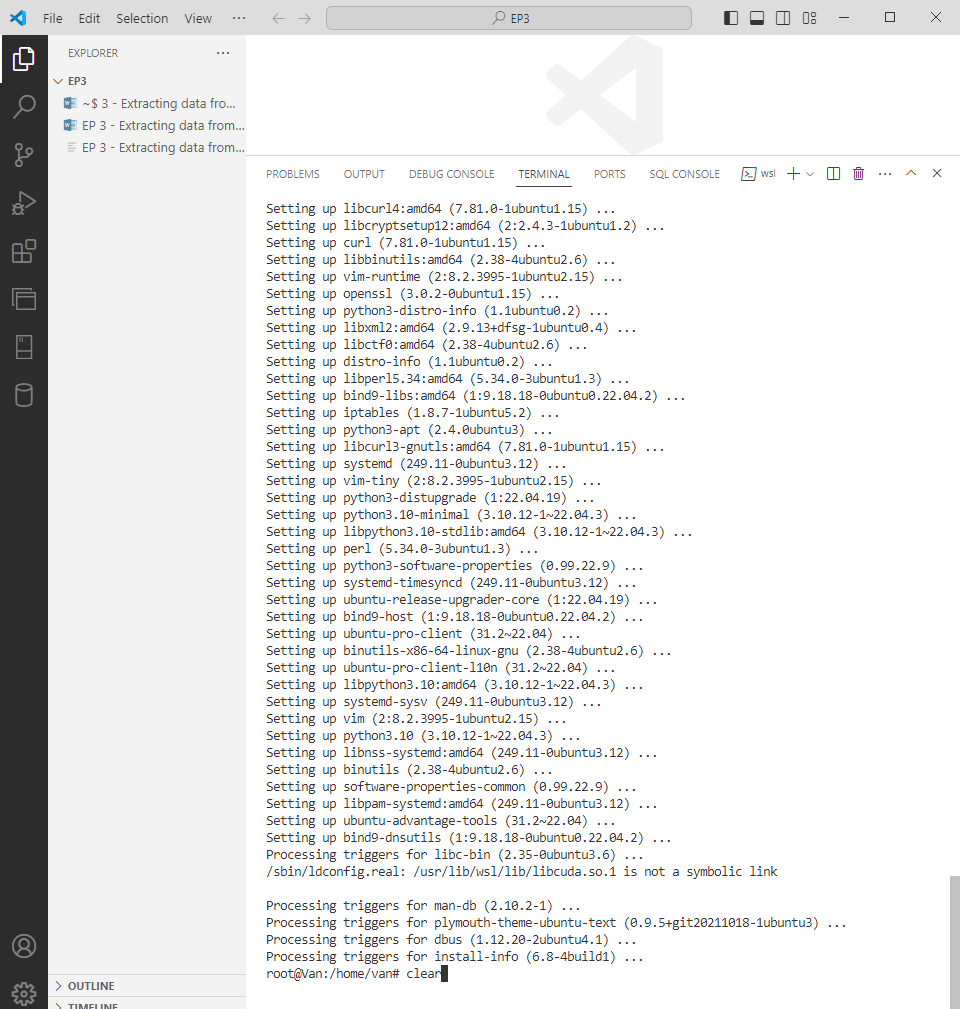
Will be switch to root user: admin user super user



Type password

Then type this command to see the latest update and upgrade the latest packages you have   
  
root@Van:/home/van# apt update && apt upgrade  
  


Type Y for Yes

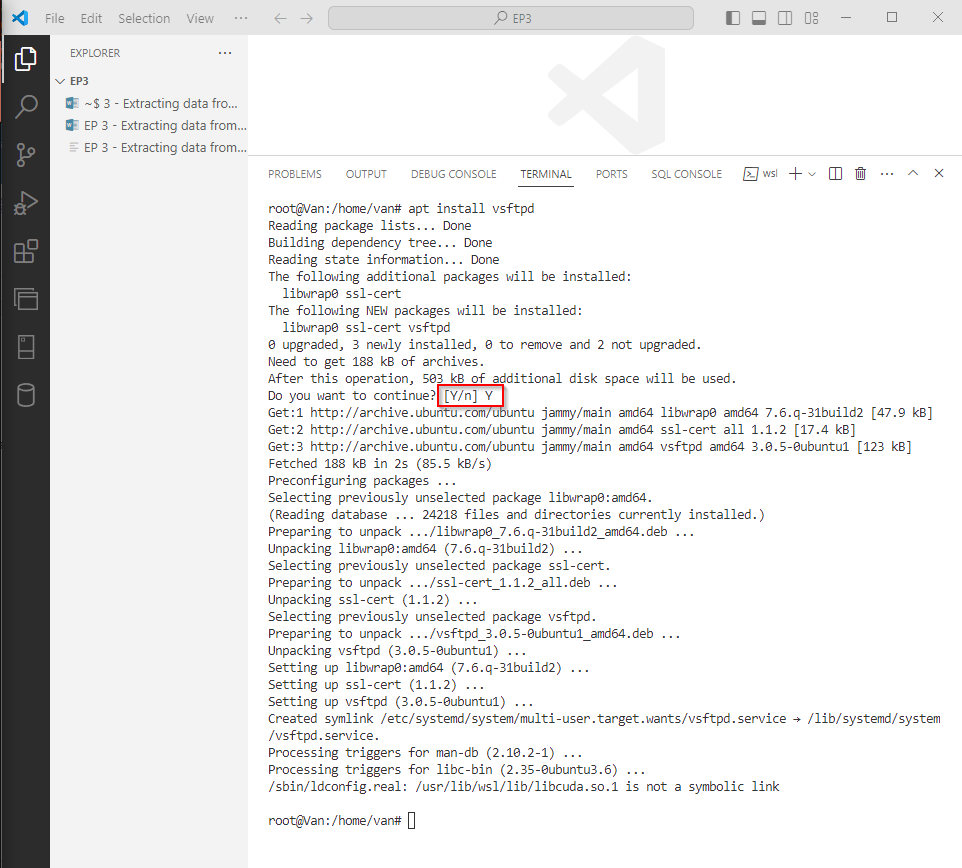


Clear

2. we will install vsftpd

root@Van:/home/van# apt install vsftpd

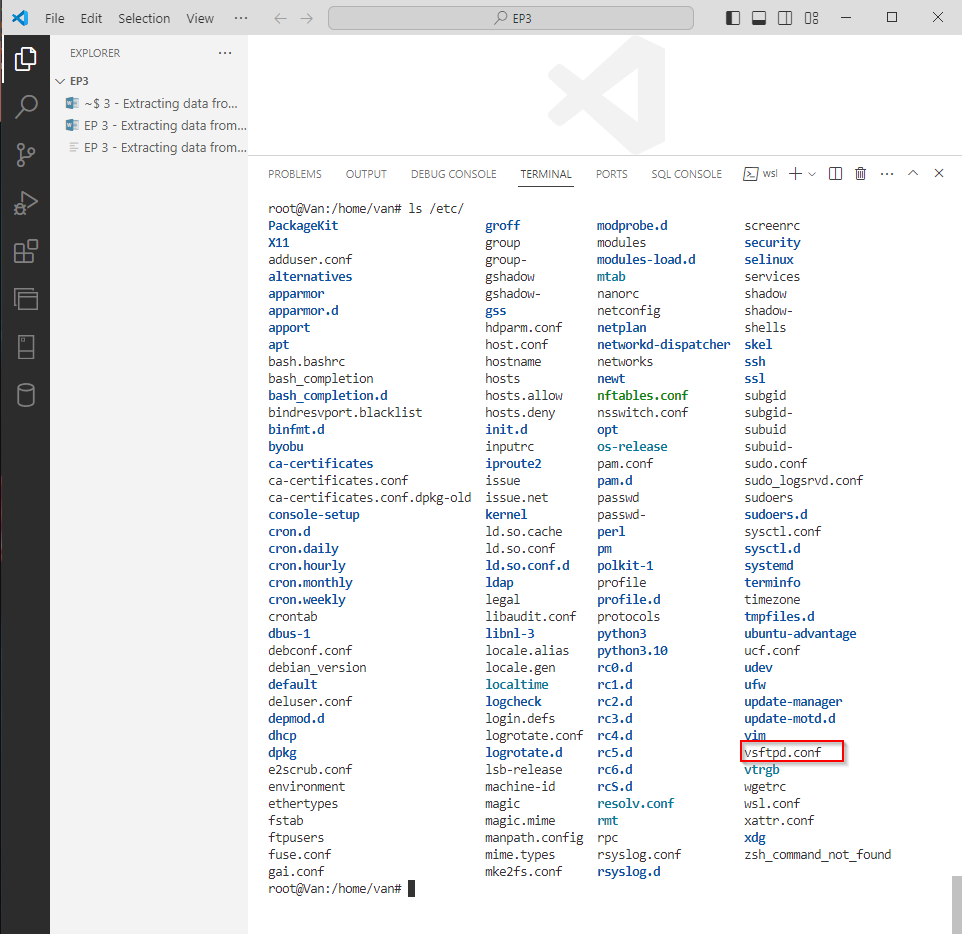
then type Y



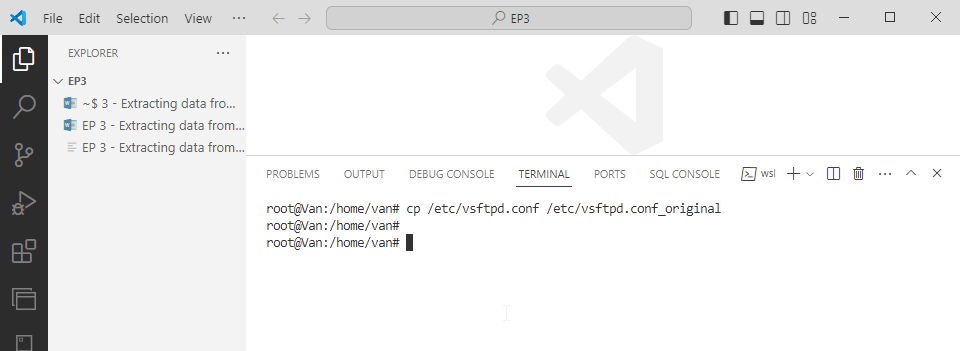
Then, we will configure the vsftpd, we can create a specific user that we can use to log in to sftp server

root@Van:/home/van# ls /etc/

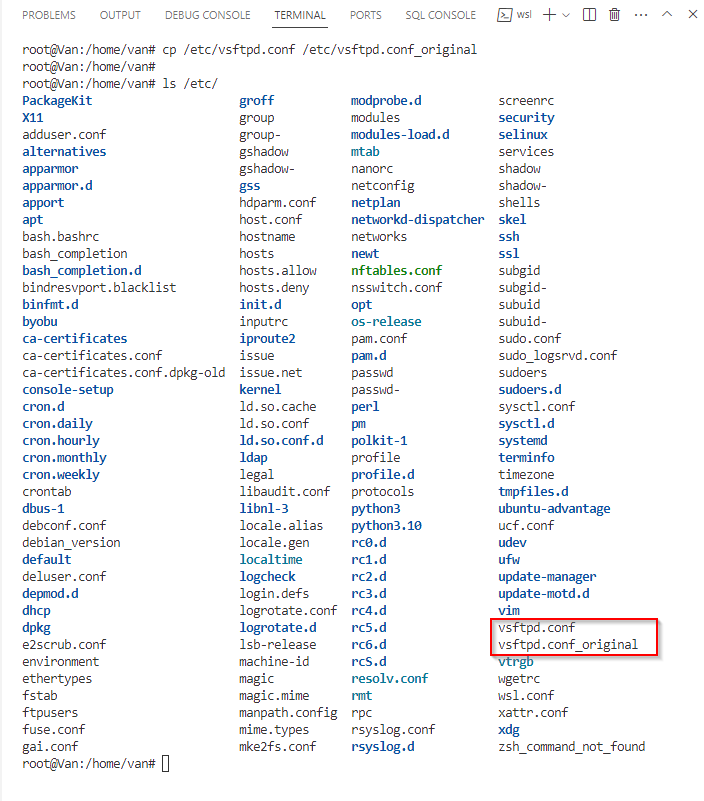
forward slash “/” used to navigate to the parent root folder of the Ubunta OS



We will edit or configure this vsftpd.conf file – so we can access outside Ubuntu going to Linux. So we will need a copy of this as backup before we edit it.

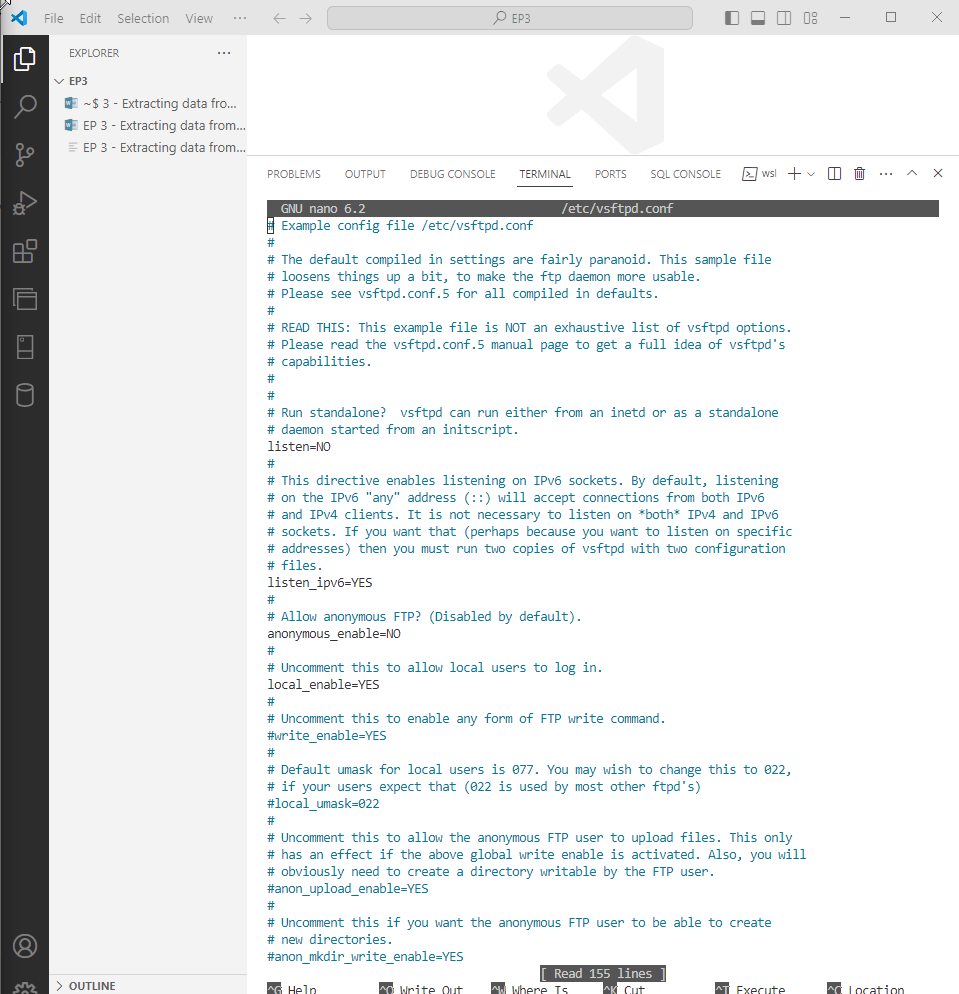


Then ls to view if its been copied



root@Van:/home/van# nano /etc/vsftpd.conf

nano – embedded text editor in Ubuntu



We will edit or add these parameters  
  
local\_enable=YES

write\_enable=YES

chroot\_local\_user=YES # there are lines of this, we will uncomment the 2nd only

chroot\_list\_enable=YES

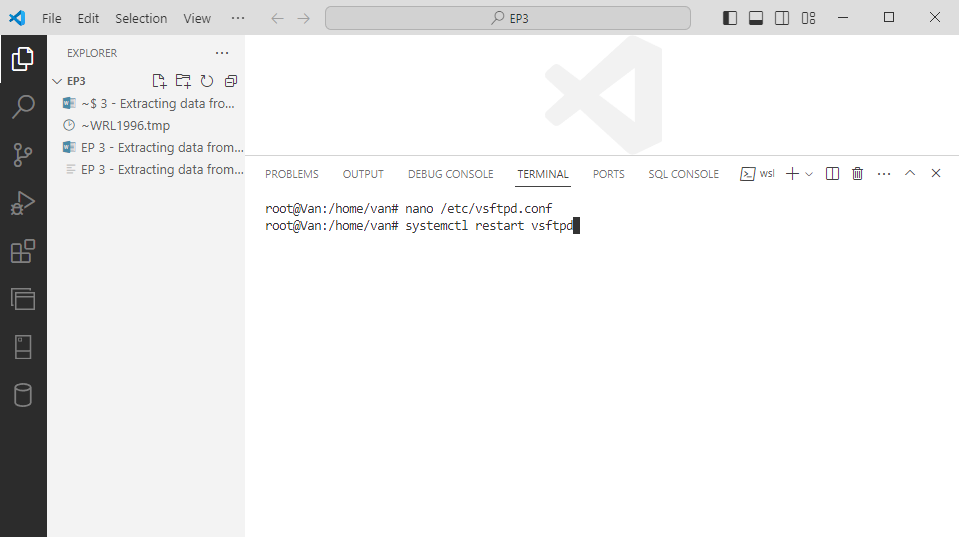
chroot\_list\_file=/etc/vsftpd.chroot\_list # or just uncomment

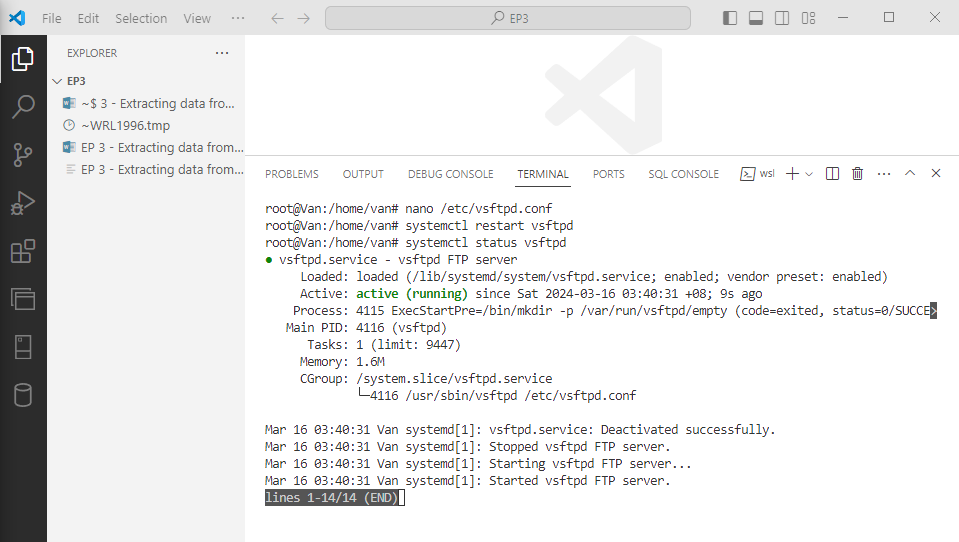
ssl\_enable=YES

require\_ssl\_reuse=NO #add to the bottom of the file

Commands used:  
arrow keys to navigate in the file

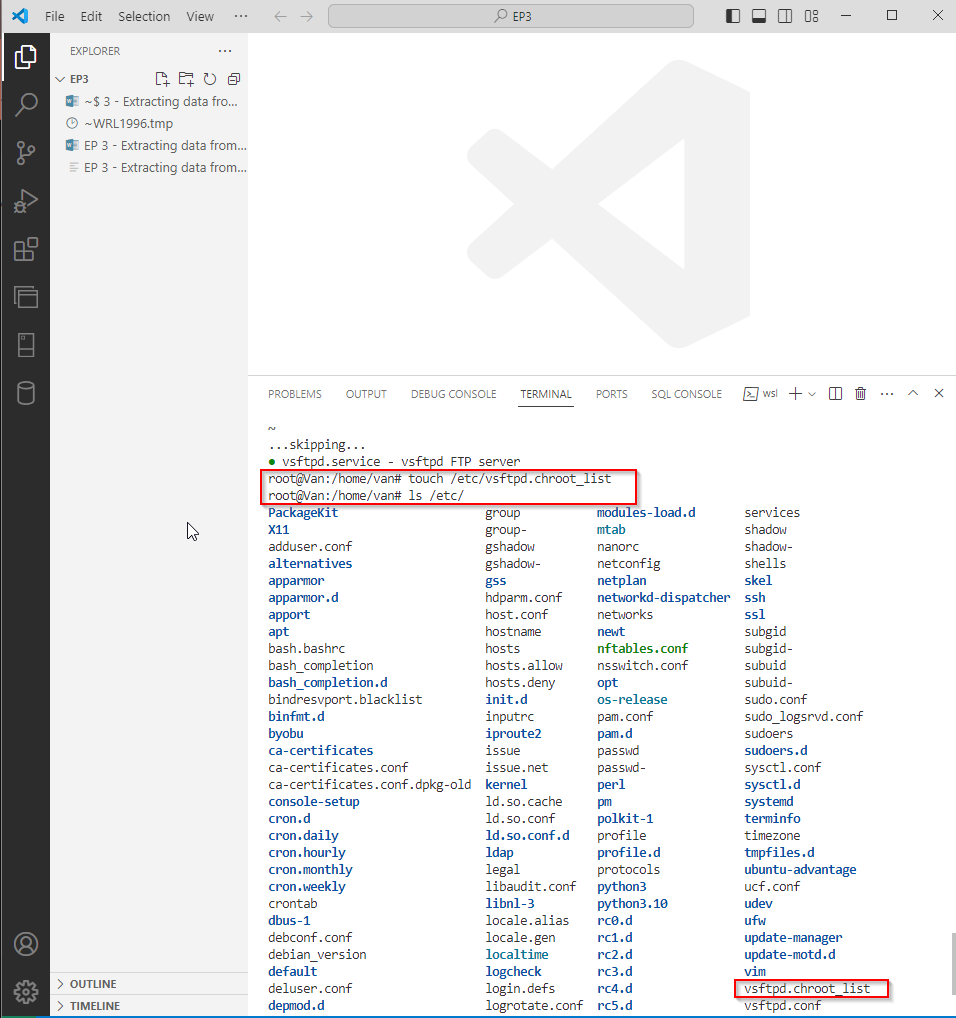
Ctrl + w – to search in the filere

after editing it, then save (ctrl+s) and (ctrl+x) to close the file.  
  
then restart, to refresh the configuration file we edited.  
  


then we run this command to see the status  
root@Van:/home/van# systemctl status vsftpd  
  


press q – to quit the systemctl if you stuck lol

then we will use the touch command to create a blank file for this vsftpd.chroot\_list

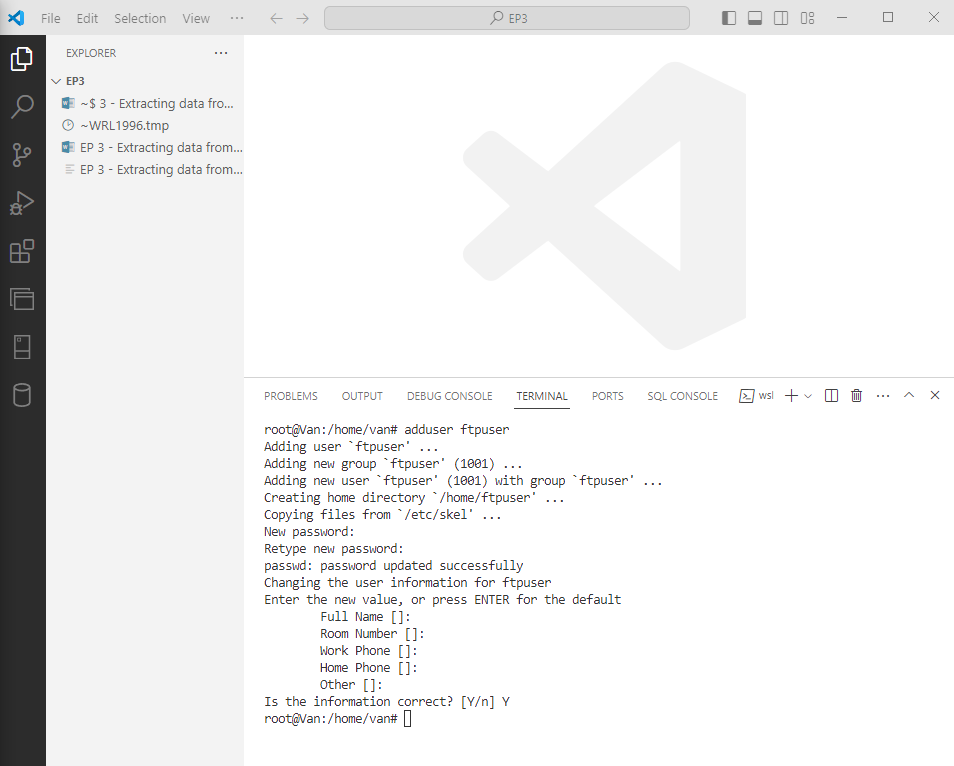


This file will be the file to list all the users that has chroot, locking the users for particular folder that they has. (security purposes)

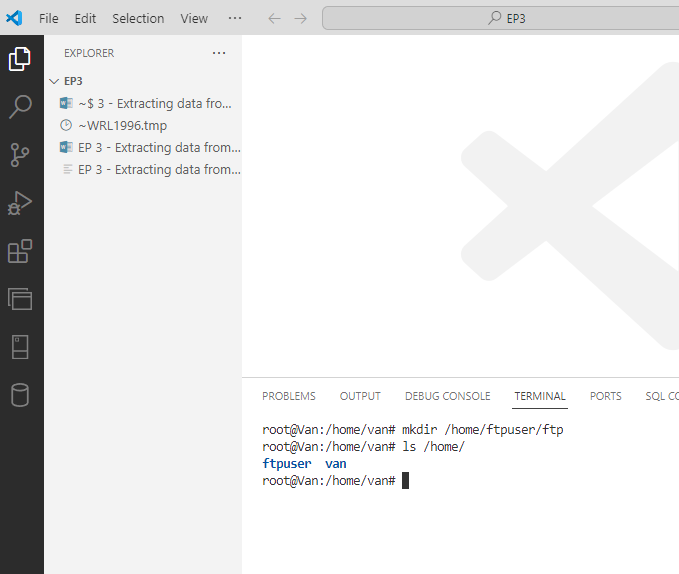
Next, we will create FTP user

I create a password:  
Password: password

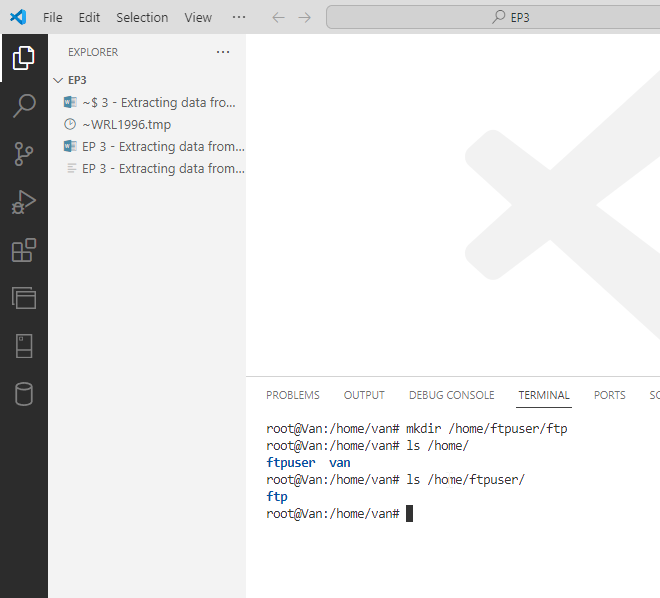
and gonna leave blank on the info user:

then type Y  
  
  


Next, we will create a folder/directory.

root@Van:/home/van# mkdir /home/ftpuser/ftp  
  


then you can see the ftp folder inside, this is where we gonna dump the files we downloaded or uploaded from web



Then, we gonna change the access level, we are changing the ownership of the folder to nobody

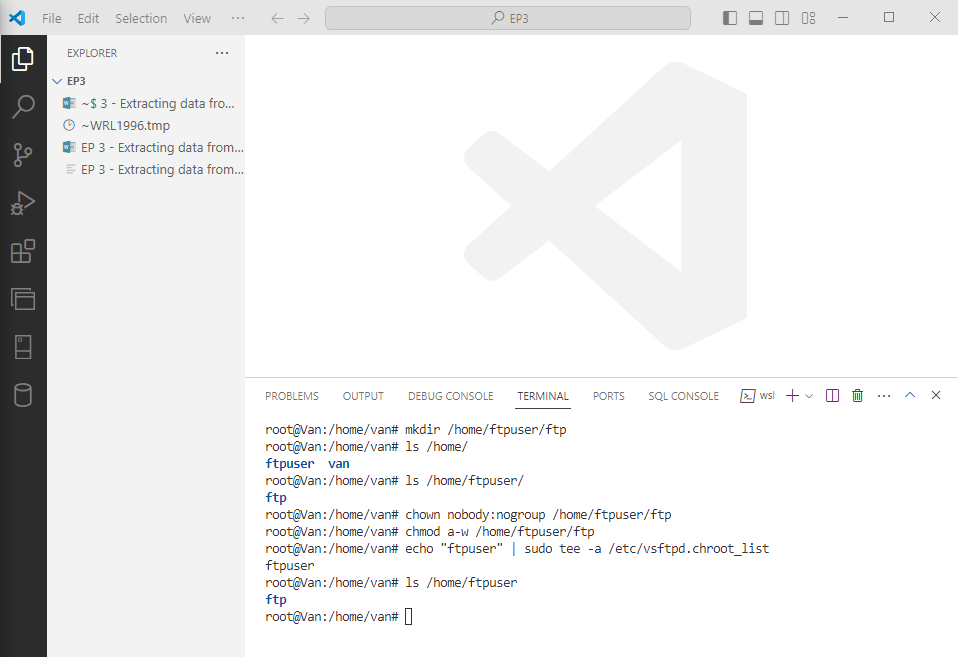
root@Van:/home/van# chown nobody:nogroup /home/ftpuser/ftp

and, we gonna change the access level or revoking the access of other users to our FTP folder

root@Van:/home/van# chmod a-w /home/ftpuser/ftp

and last is, we gonna add the ftpuser inside the file that we touch or created earlier (blank file). This is to make sure that the ftpuser can only access the home directory.

root@Van:/home/van# echo "ftpuser" | sudo tee -a /etc/vsftpd.chroot\_list



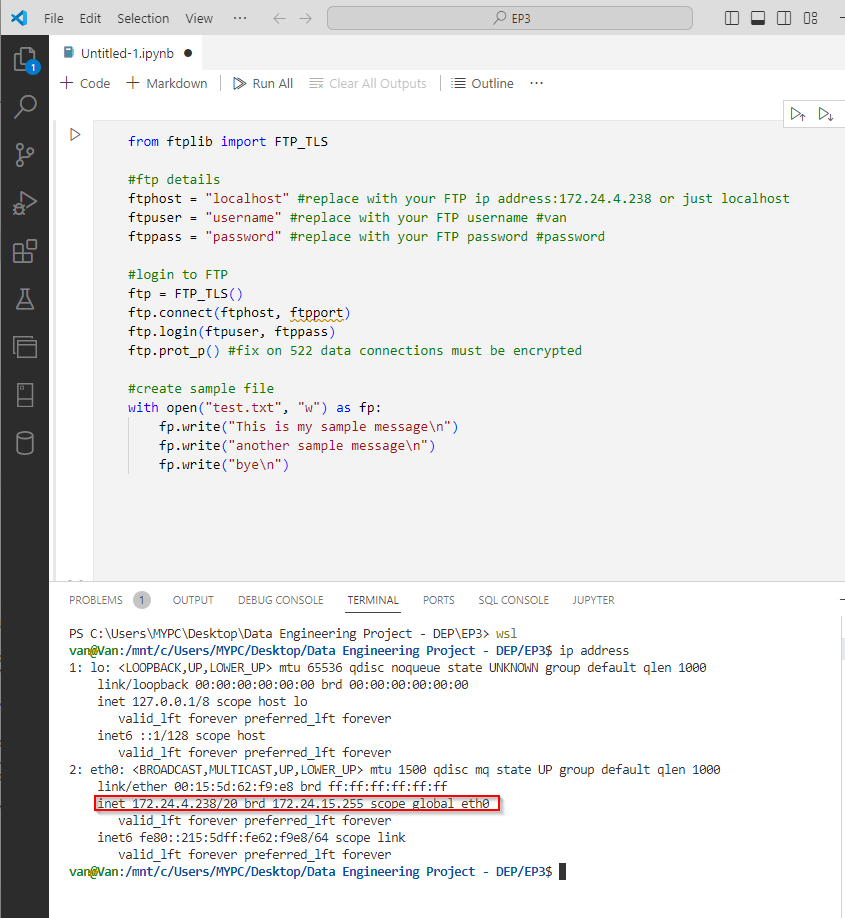
3. We will connect our FTP server to Python. Open a jupyter notebook file. Install jupyter if you have not yet installed.

To get the ip address: type wsl

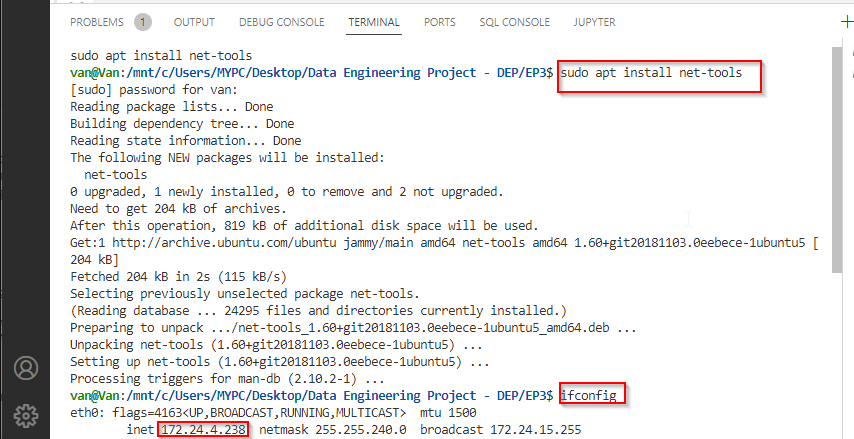
And ip address

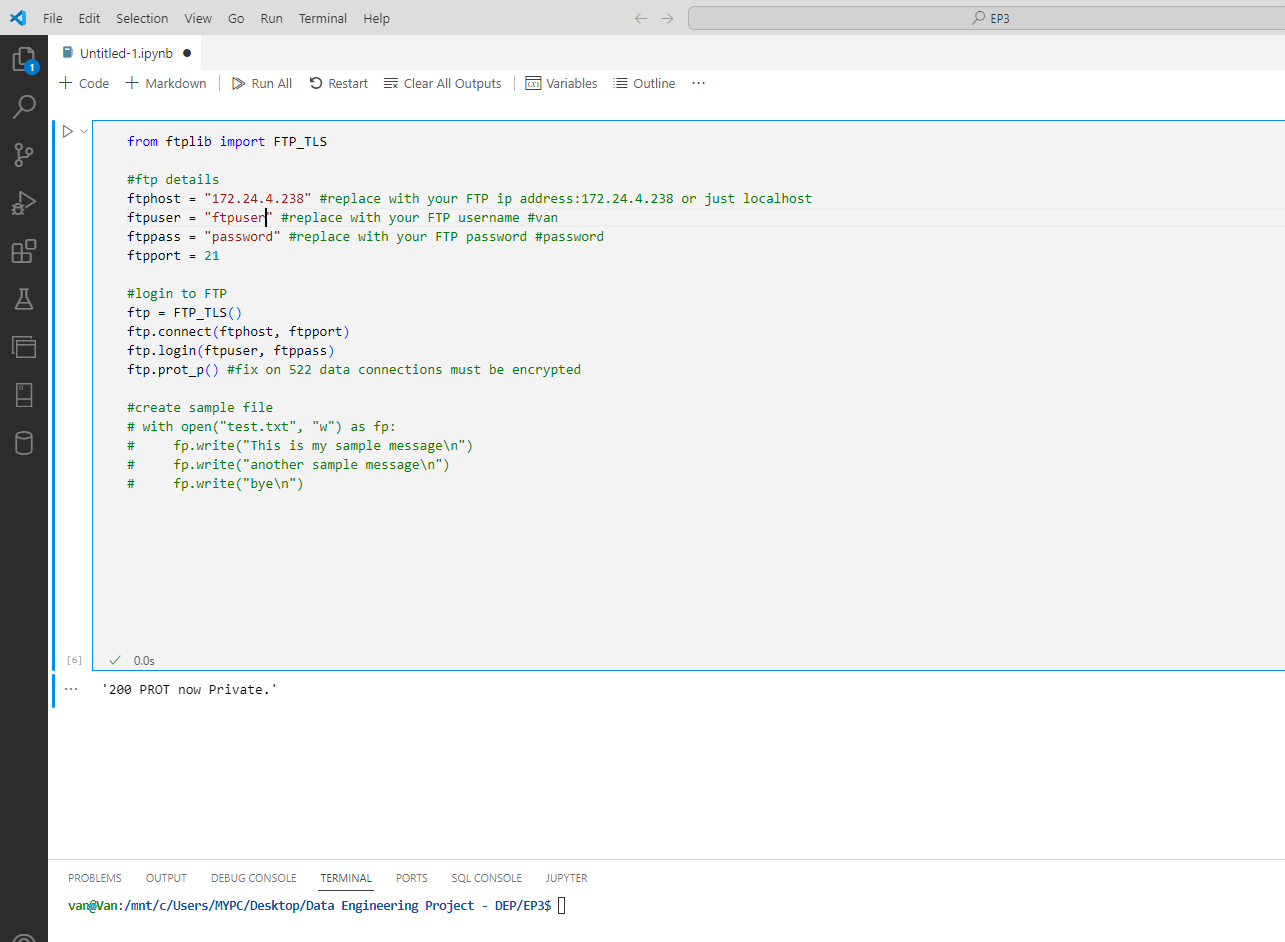
If in case you are using “ifconfig”, that’s also a way to look for ip address of wsl, and if the ifconfig is not yet installed. Just type: sudo apt install net-tools

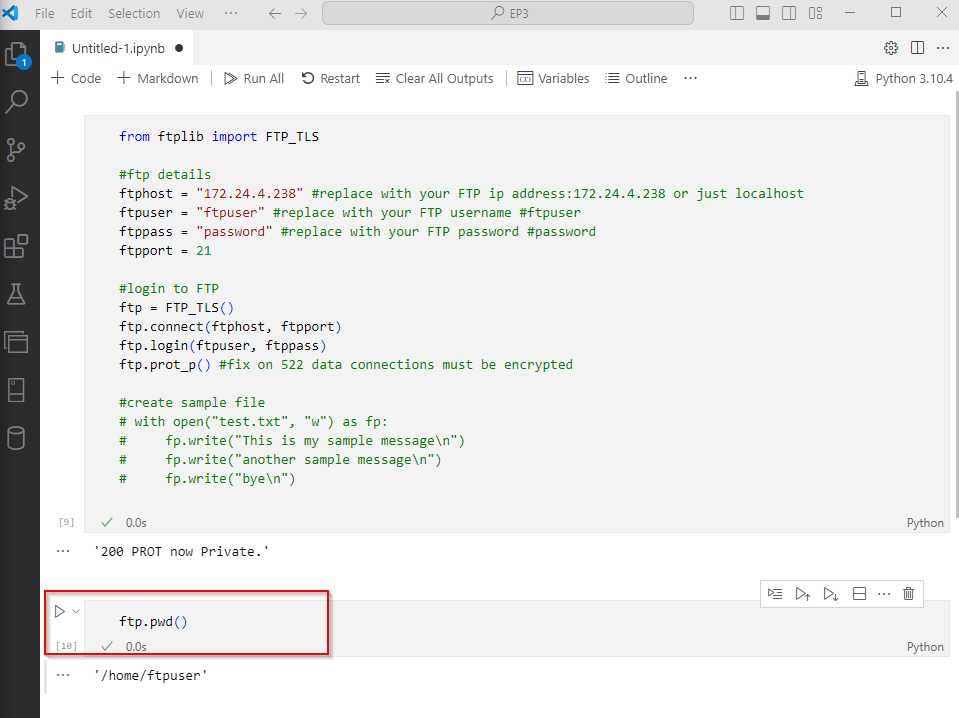
Then run again: ifconfig



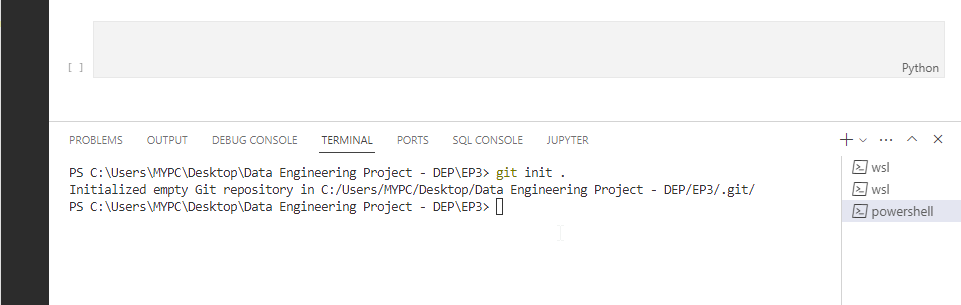
IF ifconfig used

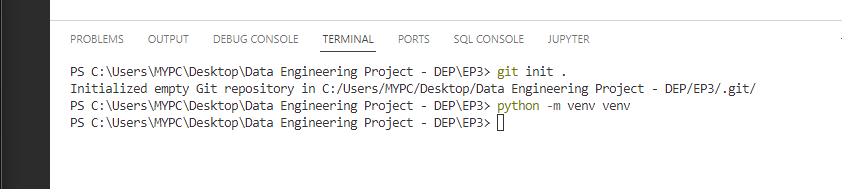


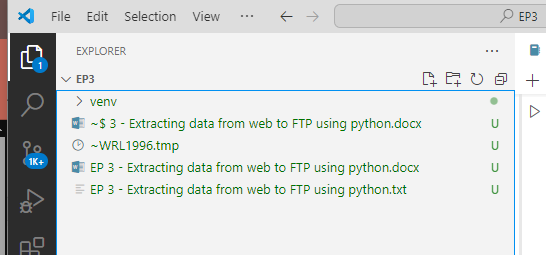
To test if we can connect to the FTP  
  


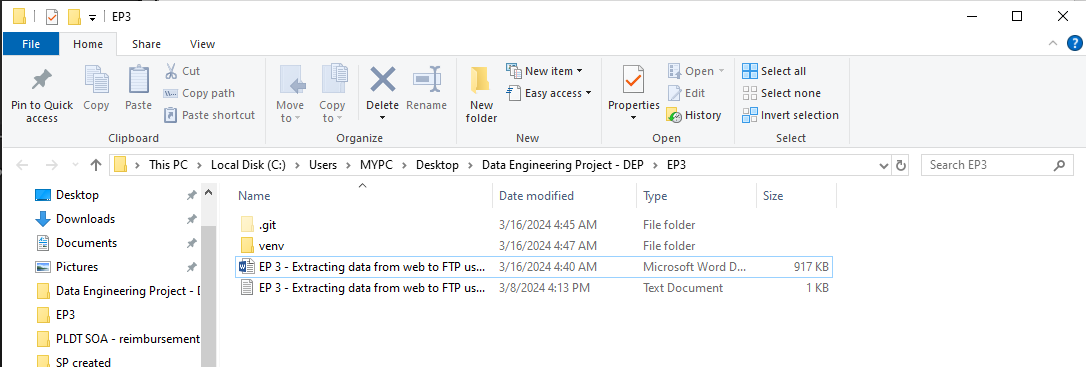
Also check if we are inside the ftp working directory  
  


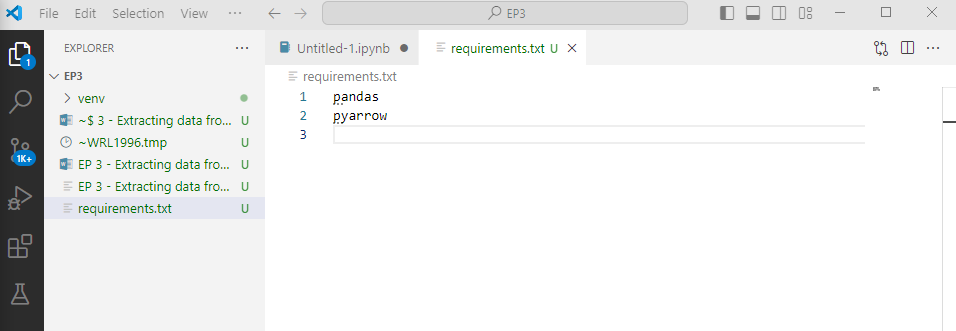
We need to create a repository, so for this episode

Use command: PS C:\Users\MYPC\Desktop\Data Engineering Project - DEP\EP3> git init .  
  


Then we create a virtual environment  
  


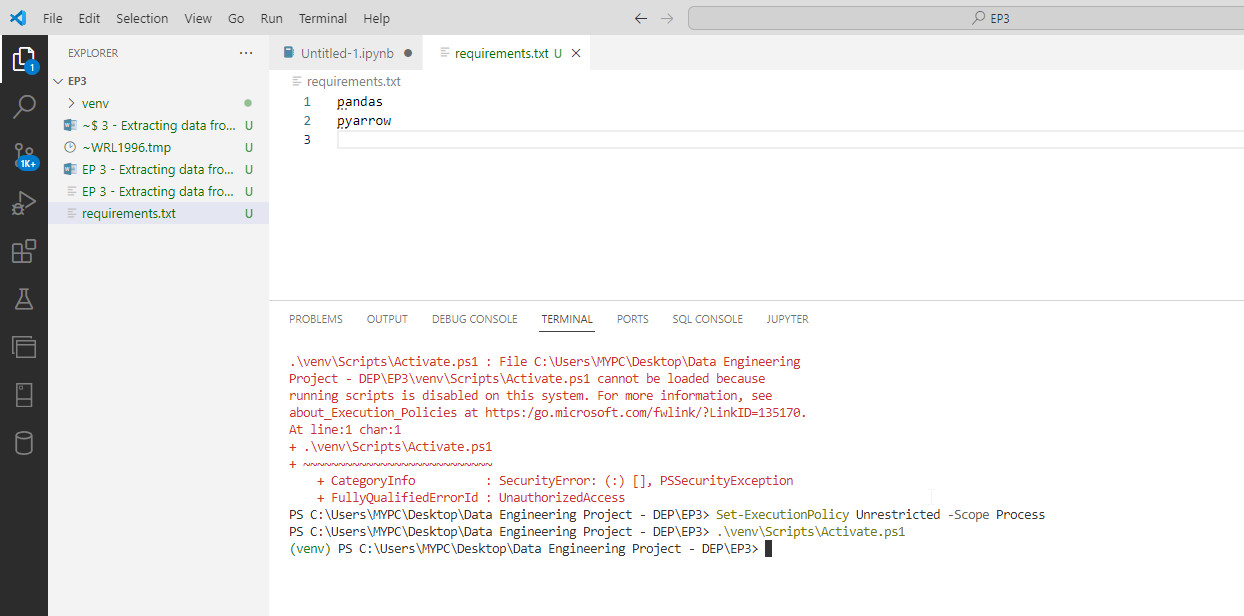




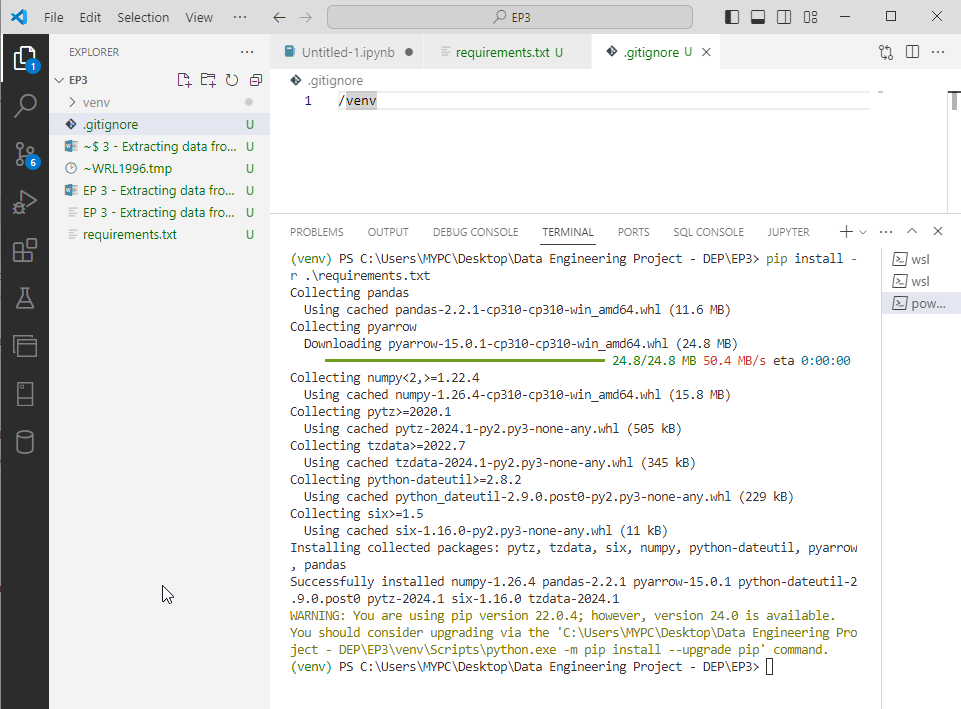
Then we create a requirements.txt which includes the packages we wanna install and used  
  


Then, activate first the activate the virtual environment

Use this command if unauthorized: Set-ExecutionPolicy Unrestricted -Scope Process



And then don’t forget to create a .gitignore file and add the environment you created. Example: /venv

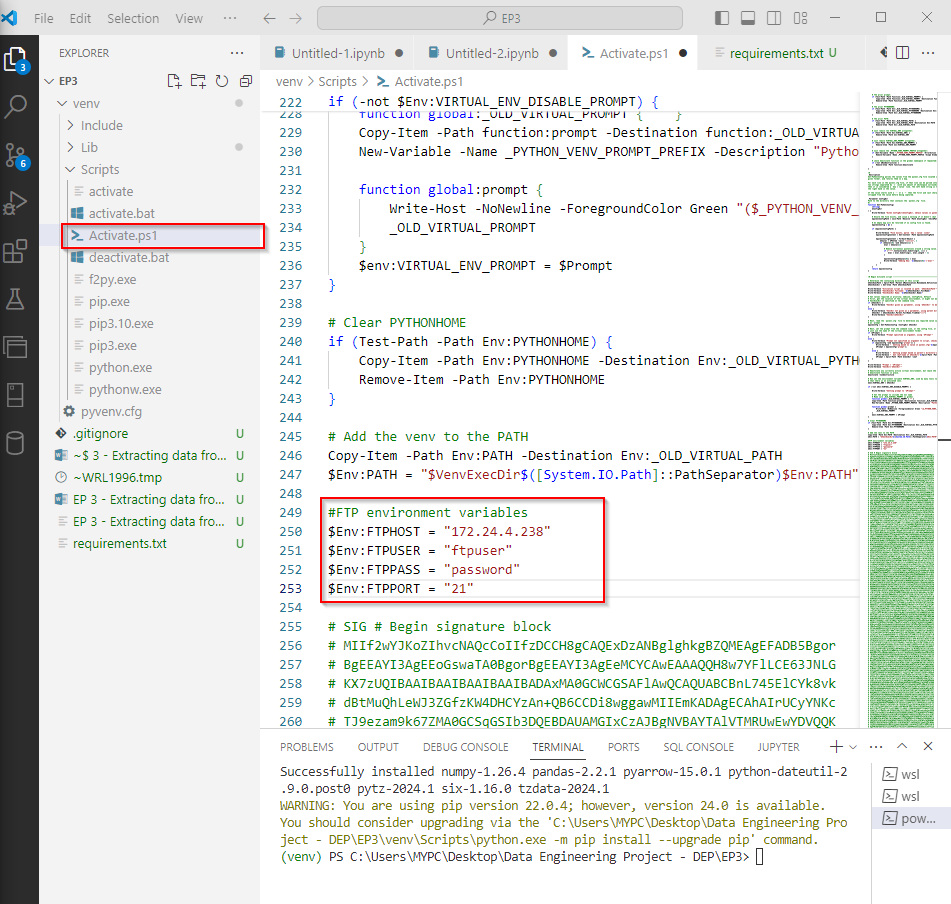


Rule of thumb: we can’t show the credentials upfront, so we need to protect it.

Go to the /venv/

Activate.ps1 since we used powershell as our terminal to activate the virtual environment

Open the Activate.ps1 file, then add the variables (credentials)



So we will just to test it out, if we can call the environment variables we input in the Activate.ps1. Open a file (jupyter notebook), activate the environment (don’t forget to use the command to gain authorize access)

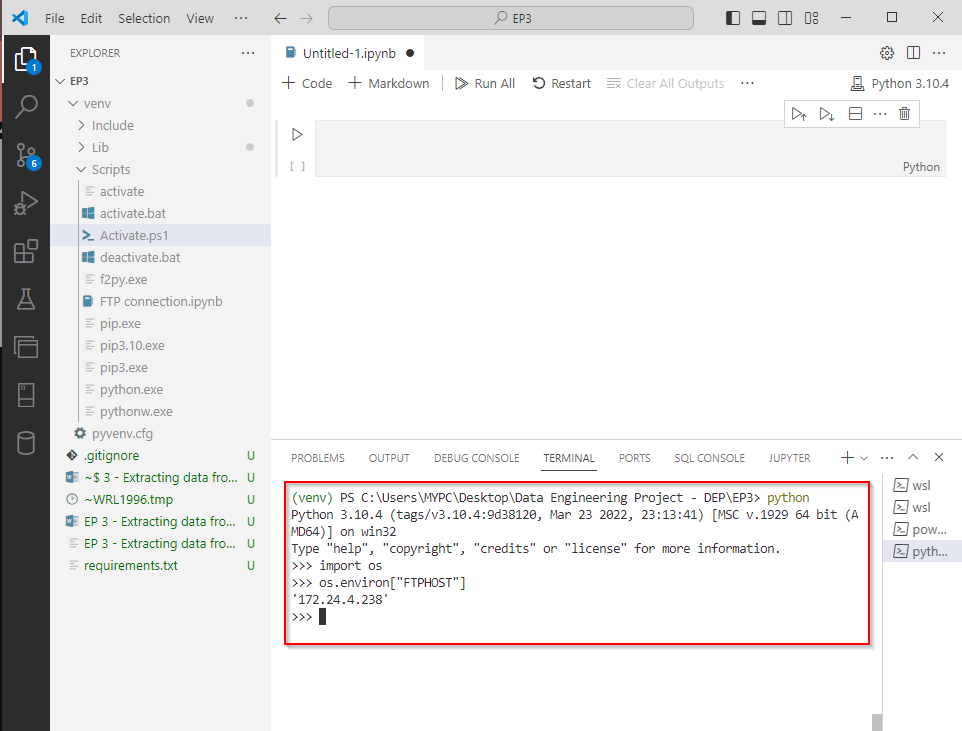
Set-ExecutionPolicy Unrestricted -Scope Process

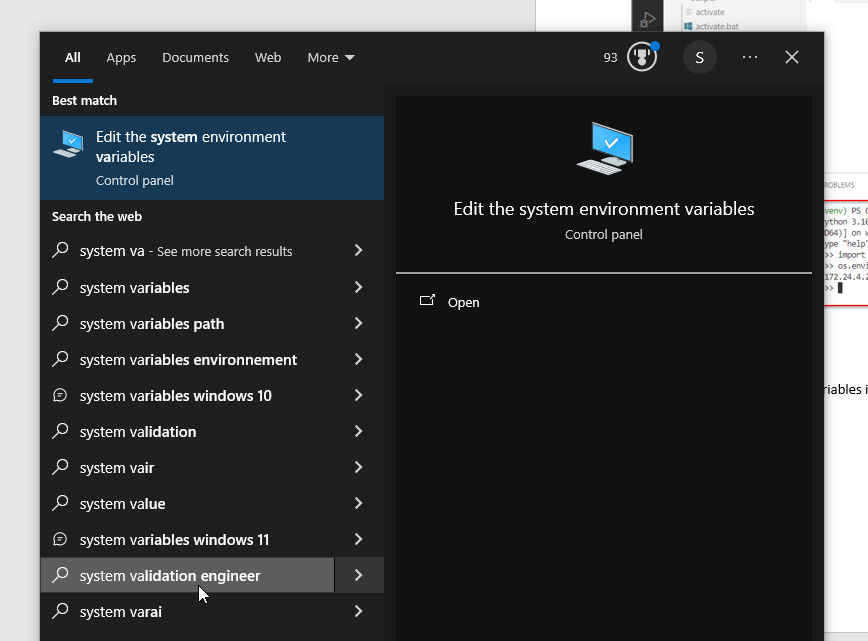
Then, type python and then type

import os

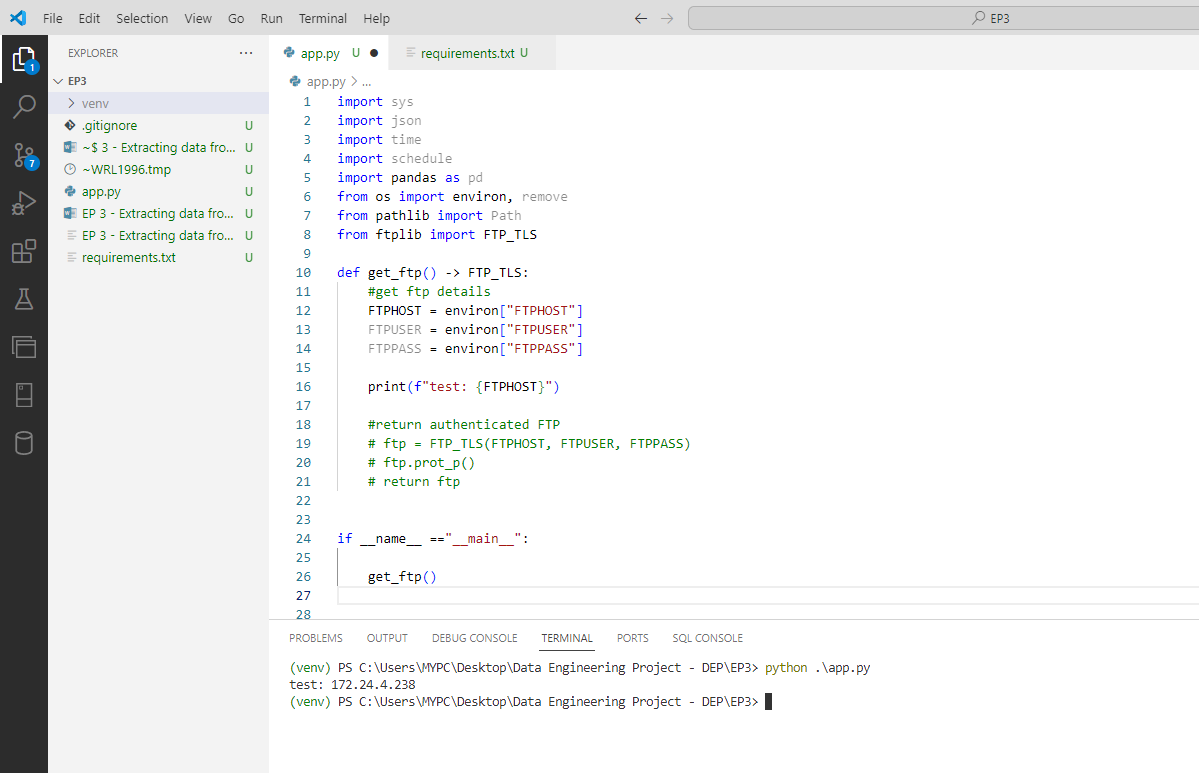
os.environ[“FTPHOST”]

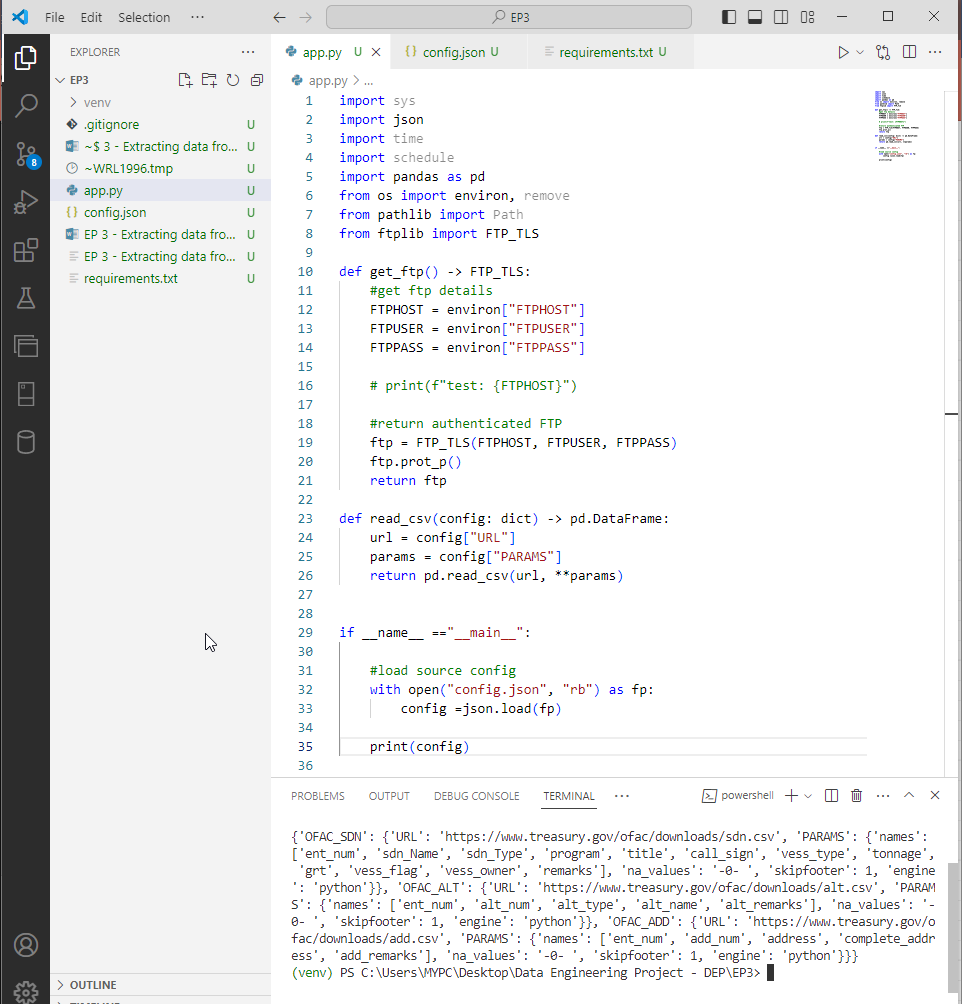
if you see the value of the variable you created/added in the Activate.ps1 it means it ran successfully

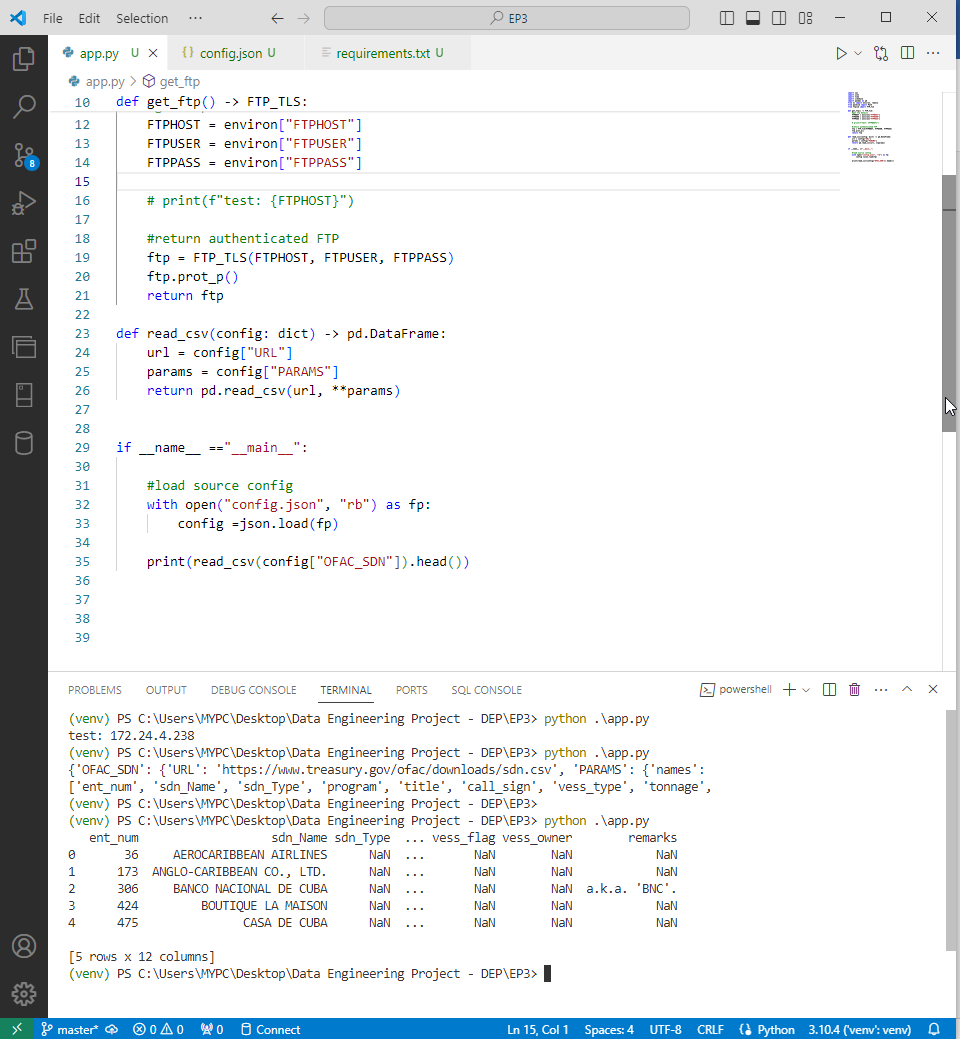


OR you can use or create you variables in your system variables  
  


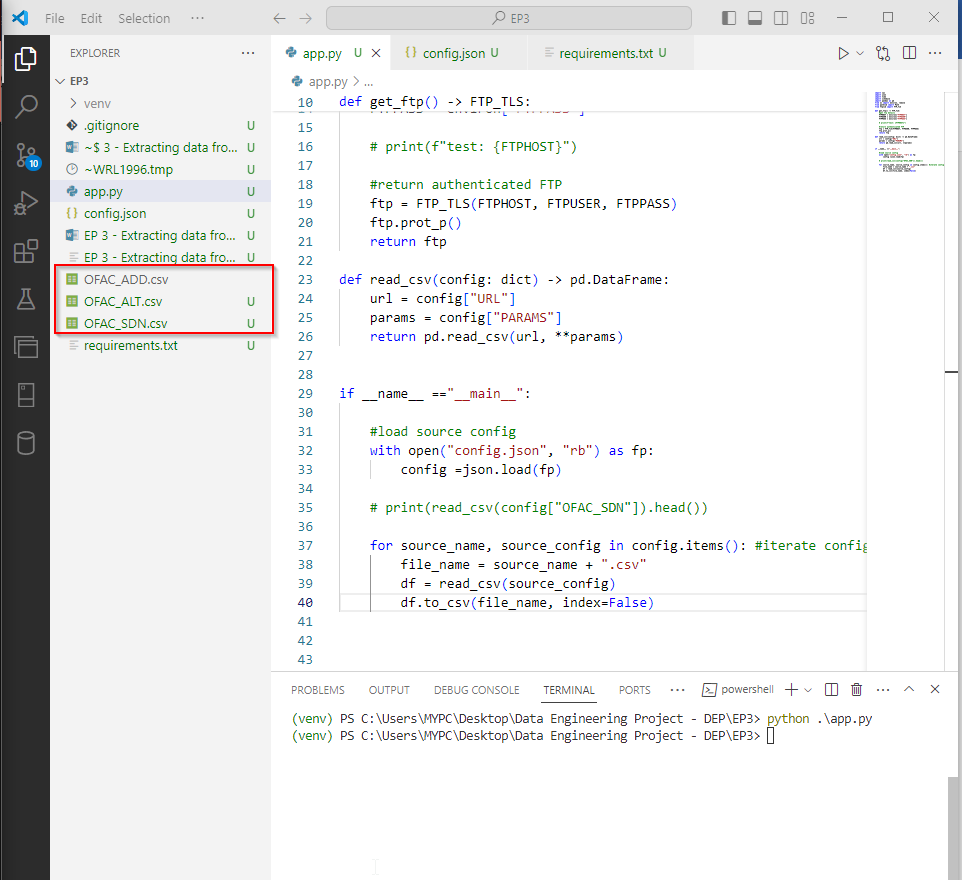
If you have encountering issue on modules that are missing, you can just restart the vscode

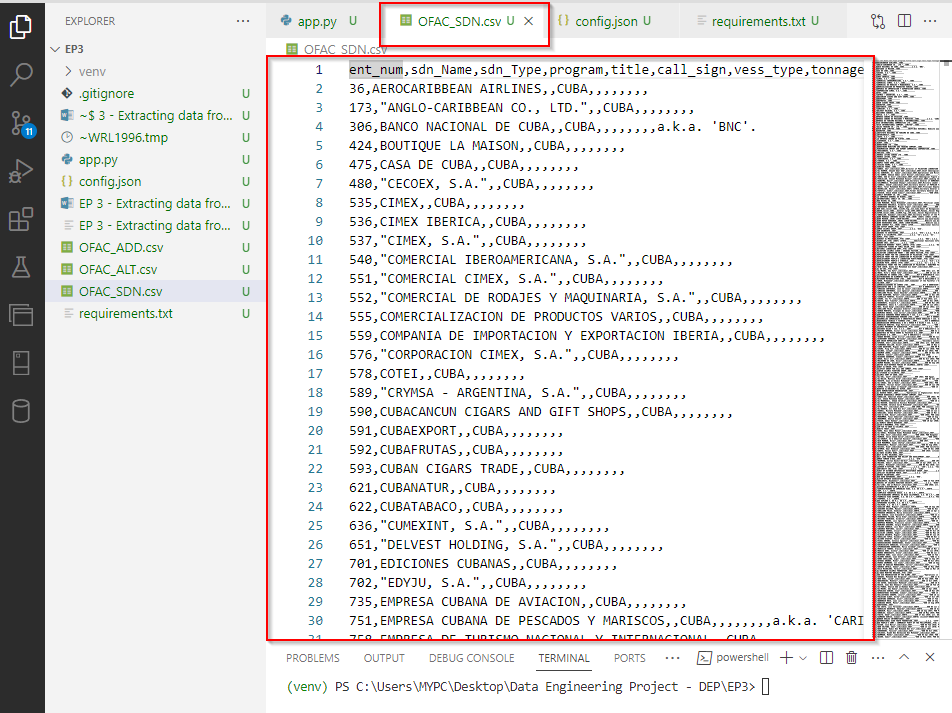




Then we try to test it out by reading and printing the file, the OFAC\_SDN  
  


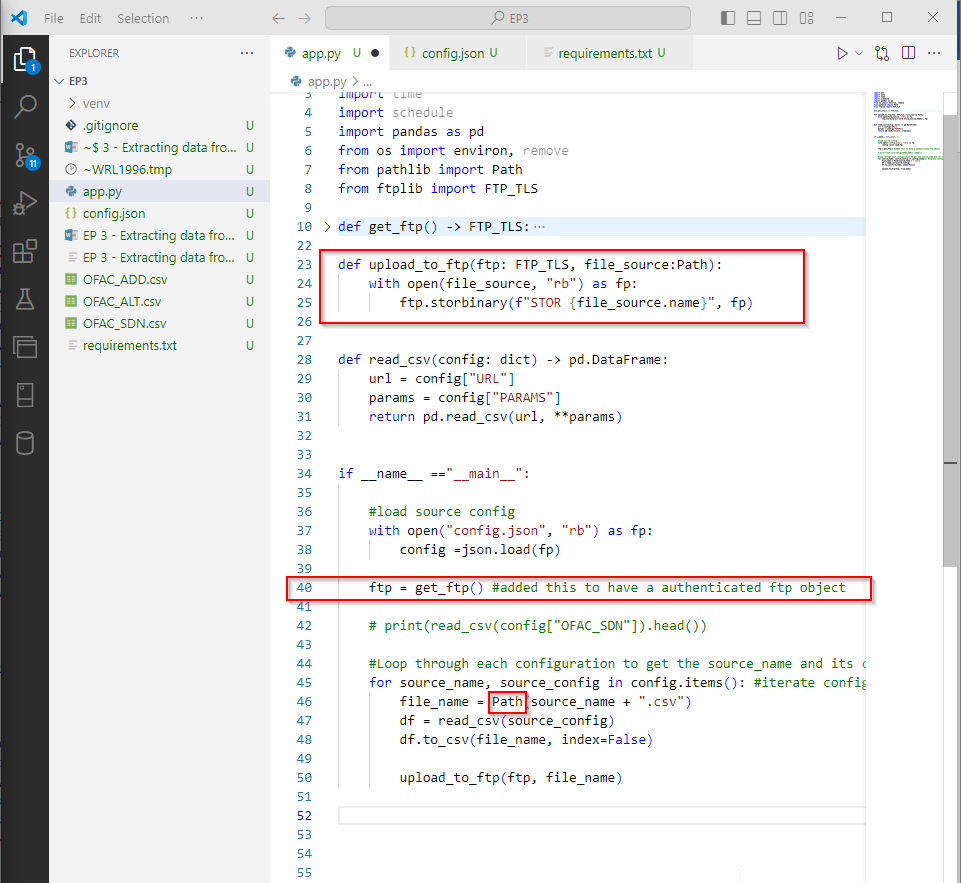
#load source config, “rb” – read binary

Then we try to test it out, we iterate the config.json dict using for loop and read and save it as csv  
  


And then, will try to open the csv. As you can see it’s clean and has column headers, NULL values which is 0 last time. Index = False, to not use the index in the dataframe.  
  


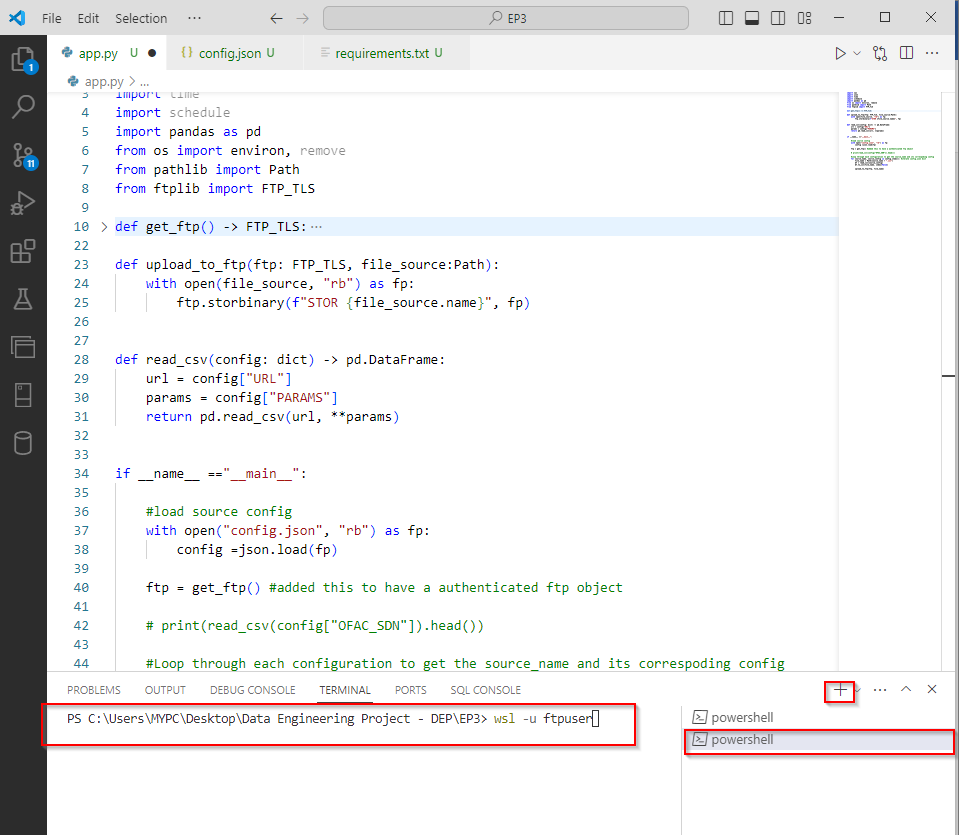
Next, we will upload these csv files to our FTP server so we will create a function to upload these

We added ftp = get\_ftp to have a authenticated ftp object, and also we converted our file\_name into pathlib object.   
  
so the upload\_to\_ftp (ftp, file\_name) -------------🡪 (ftp authenticated object, pathlib object)

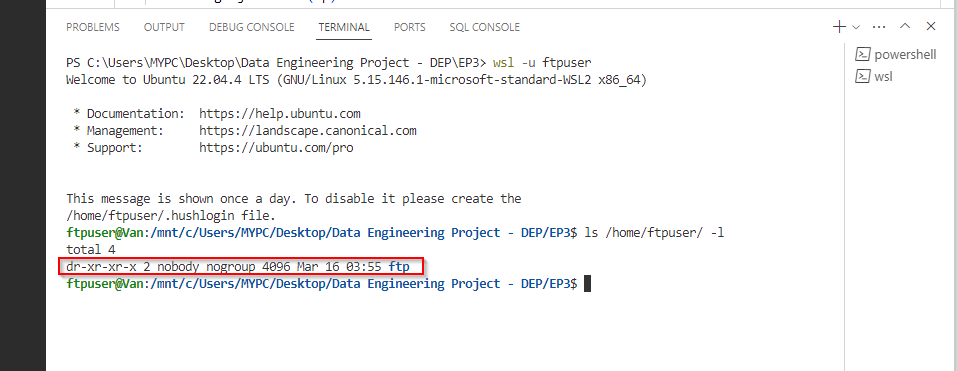


How can we determine if the files were uploaded to the ftp? To find out:  
  
Open a another terminal tab, and then type

PS C:\Users\MYPC\Desktop\Data Engineering Project - DEP\EP3> wsl -u ftpuser



Then, we will run below command to see if the files were uploaded after we run the app.py again



As you can see, we already uploaded the csv files to our FTP server

