VYPER 2.0

DOCUMENTATION

Vyper to Databricks

Vyper to Local Machine

XLXS from Databricks to Local Machine

Table of Contents

1) Vyper to Databricks:

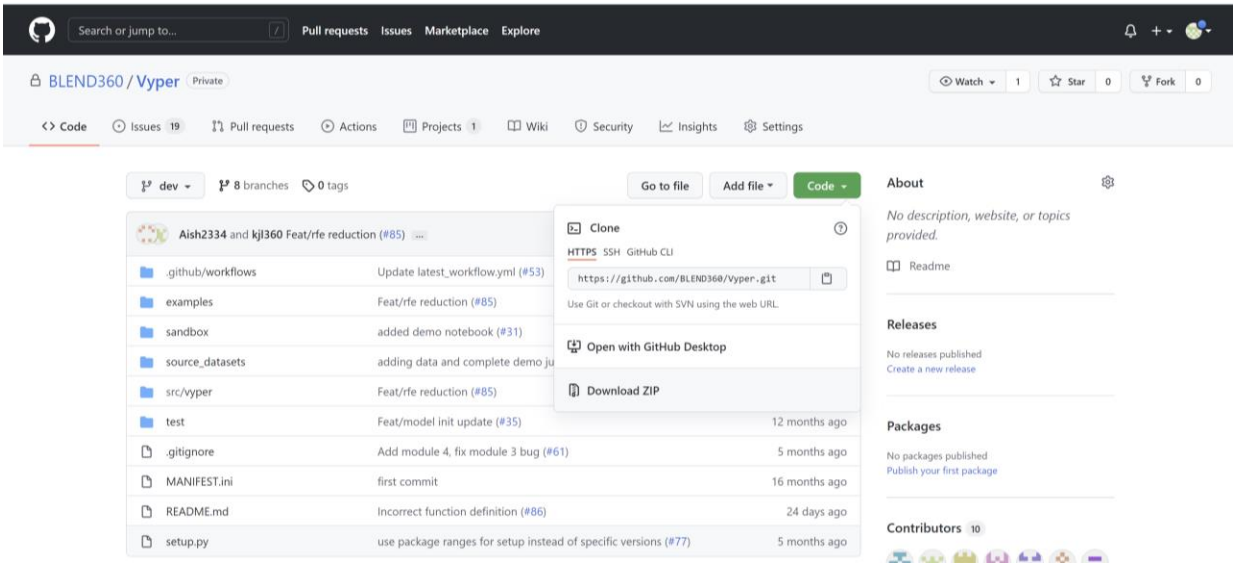
This part of the documentation will help you about how to get the in-house tool Vyper setup into Databricks.

a. Download Vyper repo from GitHub

i. Contact Kevin Lee - kevin.lee@Blend360.com for GitHub access

ii. Contact Ziyue Qi - ziyue.qi@Blend360.com for Databricks access

iii. Download the Vyper repository in your machine



b. Open Anaconda Powershell Prompt or any other Command Line Interface

i. Open the repository using cd command

Anaconda Powershell Prompt (anaconda3)

```
(base) PS C:\Users\RishabhUpadhye> cd Vyper-dev
(base) PS C:\Users\RishabhUpadhye\Vyper-dev>
```

ii. Install the wheel library – 'pip install wheel'

```
(base) PS C:\Users\RishabhUpadhye\Vyper-dev> pip install wheel
Requirement already satisfied: wheel in c:\users\rishabhupadhye\anaconda3\lib\site-packages (0.37.0)
(base) PS C:\Users\RishabhUpadhye\Vyper-dev>
```

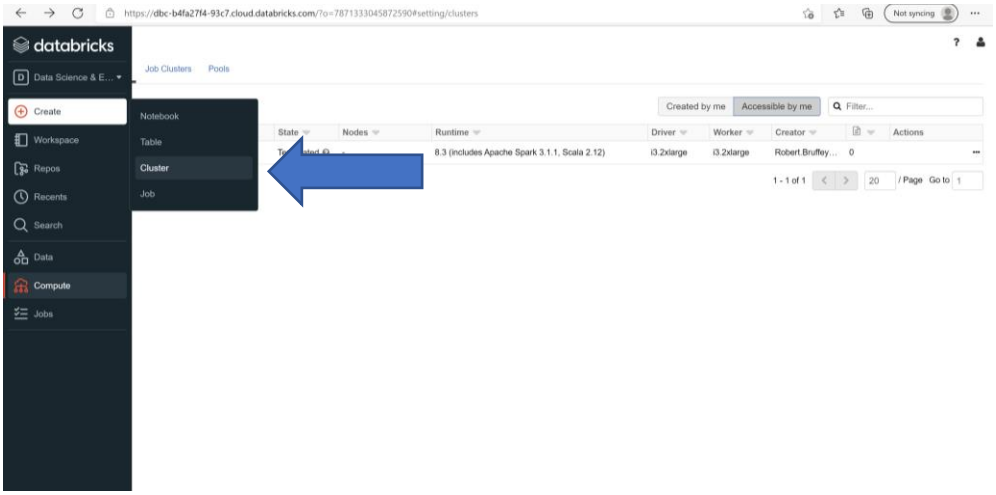
iii. Run - ‘python setup.py bdist_wheel’ command

```
(base) PS C:\Users\RishabhUpadhye\Vyper-dev> python setup.py bdist_wheel
running bdist_wheel
running build
running build_py
running egg_info
writing src\vyper.egg-info\PKG-INFO
writing dependency_links to src\vyper.egg-info\dependency_links.txt
writing requirements to src\vyper.egg-info\requires.txt
writing top-level names to src\vyper.egg-info\top_level.txt
reading manifest file 'src\vyper.egg-info\SOURCES.txt'
writing manifest file 'src\vyper.egg-info\SOURCES.txt'
installing to build\bdist.win-amd64\wheel
running install
running install_lib
creating build\bdist.win-amd64\wheel
creating build\bdist.win-amd64\wheel\vyper
copying build\lib\vyper\base.py -> build\bdist.win-amd64\wheel\.\vyper
creating build\bdist.win-amd64\wheel\vyper\batch
copying build\lib\vyper\batch\__init__.py -> build\bdist.win-amd64\wheel\.\vyper\batch
creating build\bdist.win-amd64\wheel\vyper\metadata
copying build\lib\vyper\metadata\__init__.py -> build\bdist.win-amd64\wheel\.\vyper\metadata
creating build\bdist.win-amd64\wheel\vyper\user
copying build\lib\vyper\user\explorer.py -> build\bdist.win-amd64\wheel\.\vyper\user
copying build\lib\vyper\user\model.py -> build\bdist.win-amd64\wheel\.\vyper\user
copying build\lib\vyper\user\__init__.py -> build\bdist.win-amd64\wheel\.\vyper\user
creating build\bdist.win-amd64\wheel\vyper\utils
copying build\lib\vyper\utils\connectors.py -> build\bdist.win-amd64\wheel\.\vyper\utils
copying build\lib\vyper\utils\tools.py -> build\bdist.win-amd64\wheel\.\vyper\utils
copying build\lib\vyper\utils\__init__.py -> build\bdist.win-amd64\wheel\.\vyper\utils
copying build\lib\vyper\__init__.py -> build\bdist.win-amd64\wheel\.\vyper
running install_egg_info
Copying src\vyper.egg-info to build\bdist.win-amd64\wheel\.\vyper-2.0.0-py3.8.egg-info
running install_scripts
creating build\bdist.win-amd64\wheel\vyper-2.0.0.dist-info\WHEEL
creating 'dist\vyper-2.0.0-py3-none-any.whl' and adding 'build\bdist.win-amd64\wheel' to it
adding 'vyper/__init__.py'
adding 'vyper/base.py'
adding 'vyper/batch/__init__.py'
adding 'vyper/metadata/__init__.py'
adding 'vyper/user/__init__.py'
adding 'vyper/user/explorer.py'
adding 'vyper/user/model.py'
adding 'vyper/utils/__init__.py'
adding 'vyper/utils/connectors.py'
adding 'vyper/utils/tools.py'
adding 'vyper-2.0.0.dist-info/METADATA'
adding 'vyper-2.0.0.dist-info/WHEEL'
adding 'vyper-2.0.0.dist-info/top_level.txt'
adding 'vyper-2.0.0.dist-info/RECORD'
removing build\bdist.win-amd64\wheel
(base) PS C:\Users\RishabhUpadhye\Vyper-dev>
```

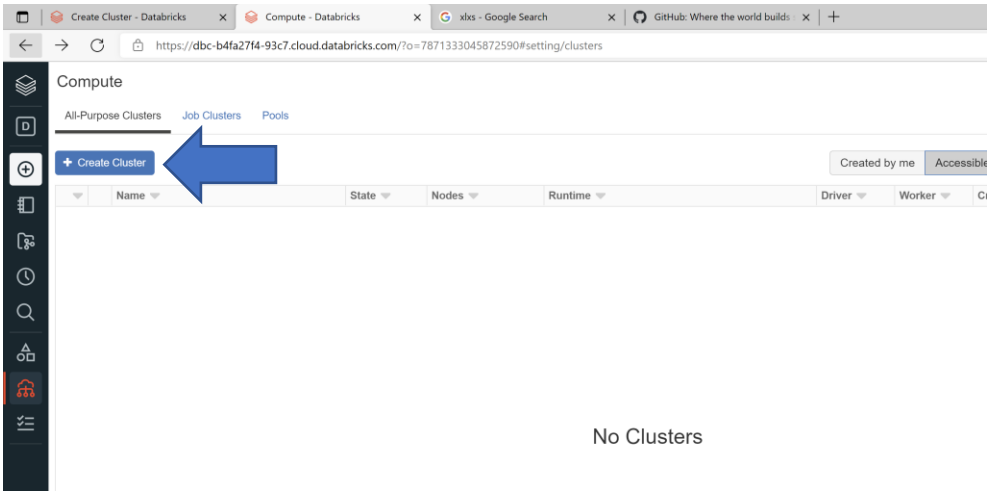
iv. We have the wheel file in **dist** folder of our repository

c. Open Databricks - https://dbc-b4fa27f4-93c7.cloud.databricks.com/

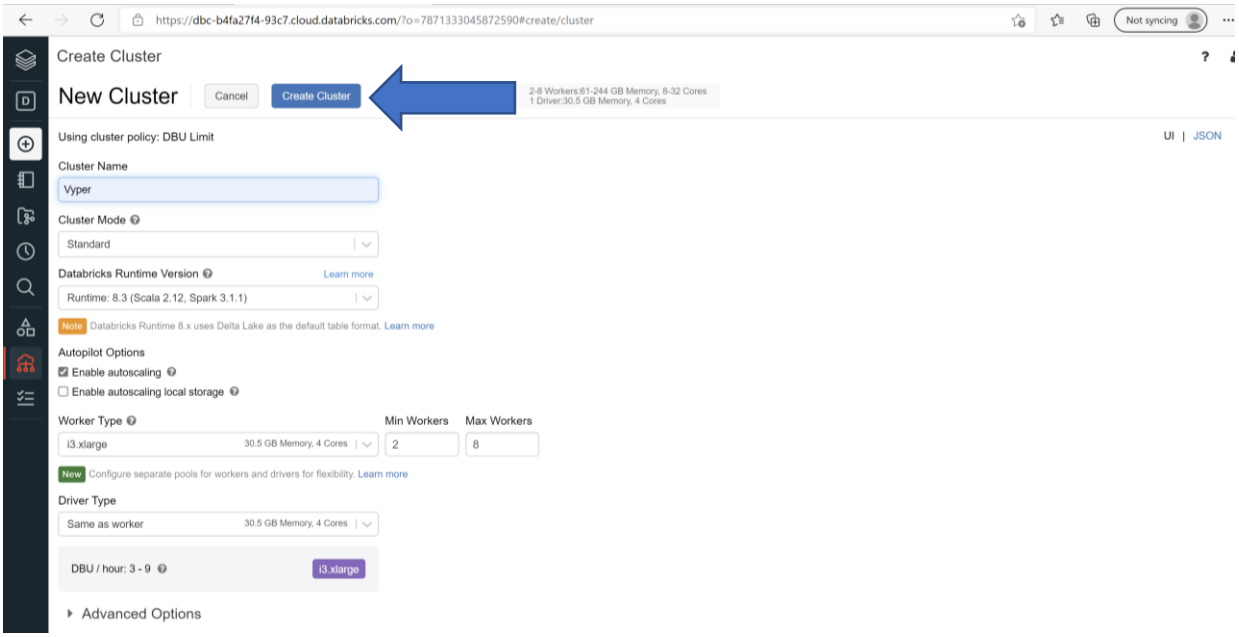
i. Click on Create -> Cluster



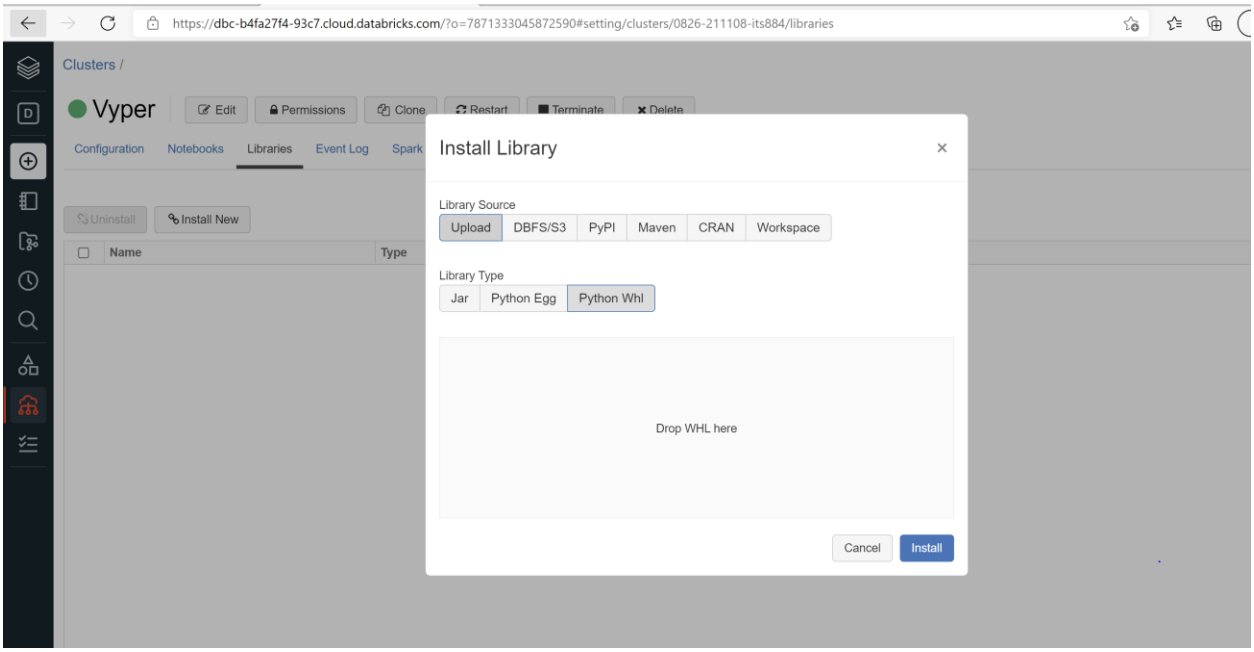
ii. Click on Create Cluster



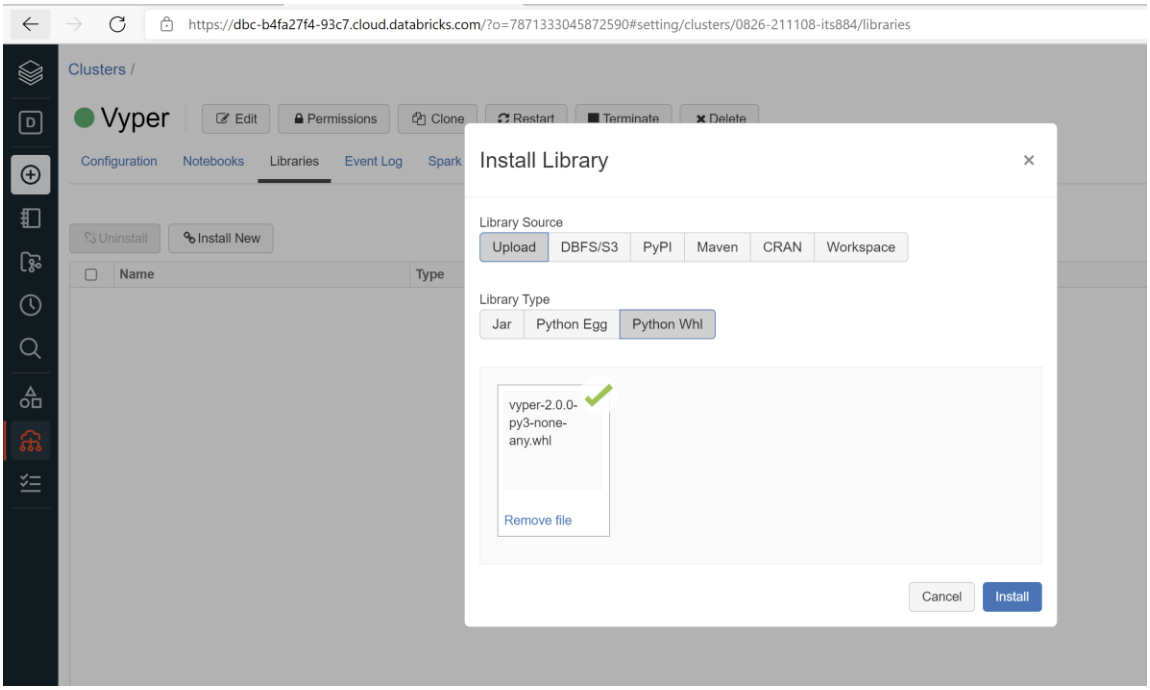
iii. Name the Cluster and select the other specifications as per your requirements. Click on Create Cluster – It take a few minutes for a cluster to be created.



iv. Click on **Libraries** tab and then on **Install New** button. An Install Library window will popup. Select **Library Type** as **Python Whl**.



v. Upload the wheel file from the **dist** folder of the Vyper Repo. Click on Install.



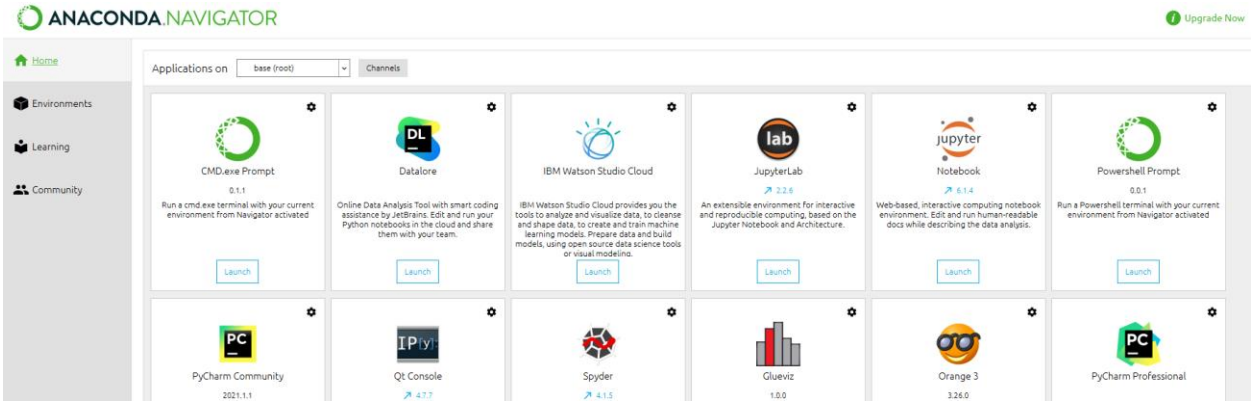
vi. Now you can successfully run Vyper in Databricks using this cluster.

2) Vyper 2.0 on Local Machine

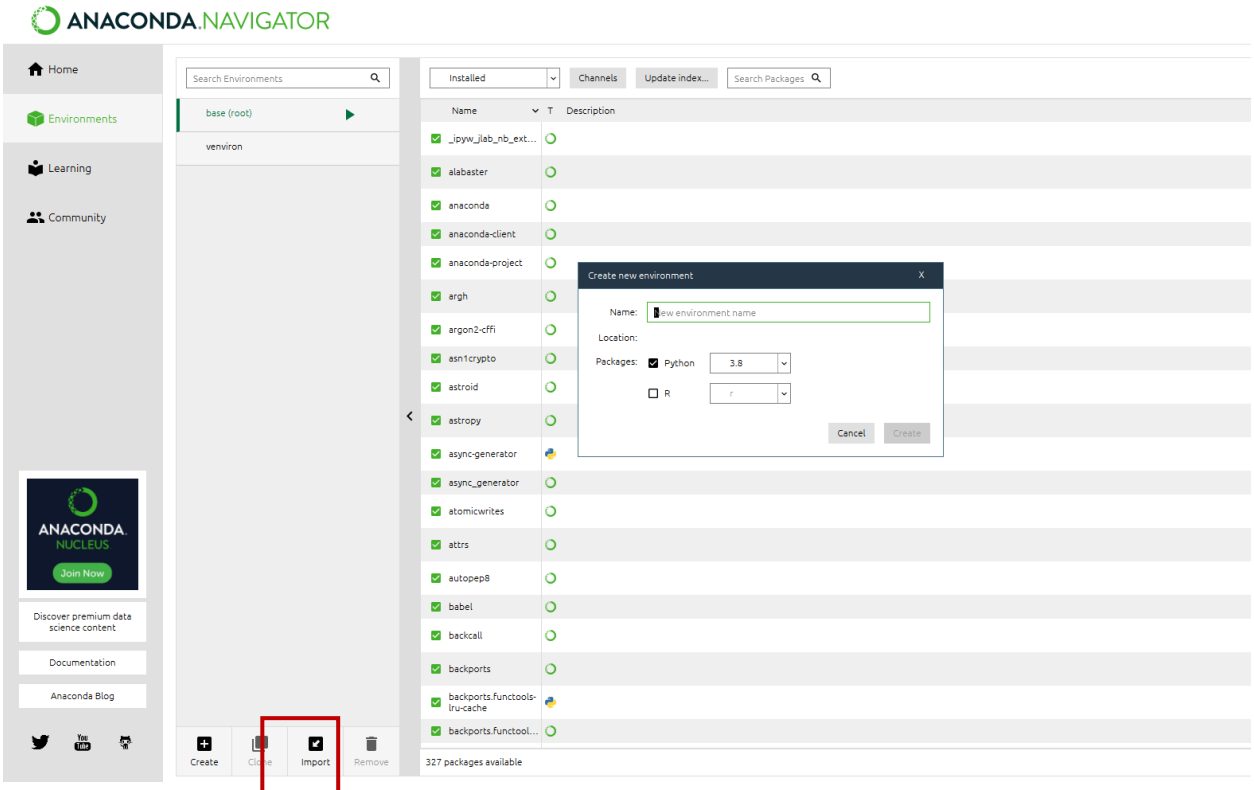
a. Anaconda Navigator

i. Install Anaconda Navigator (Anaconda | Individual Edition)

ii. Once installed, please launch Anaconda Navigator

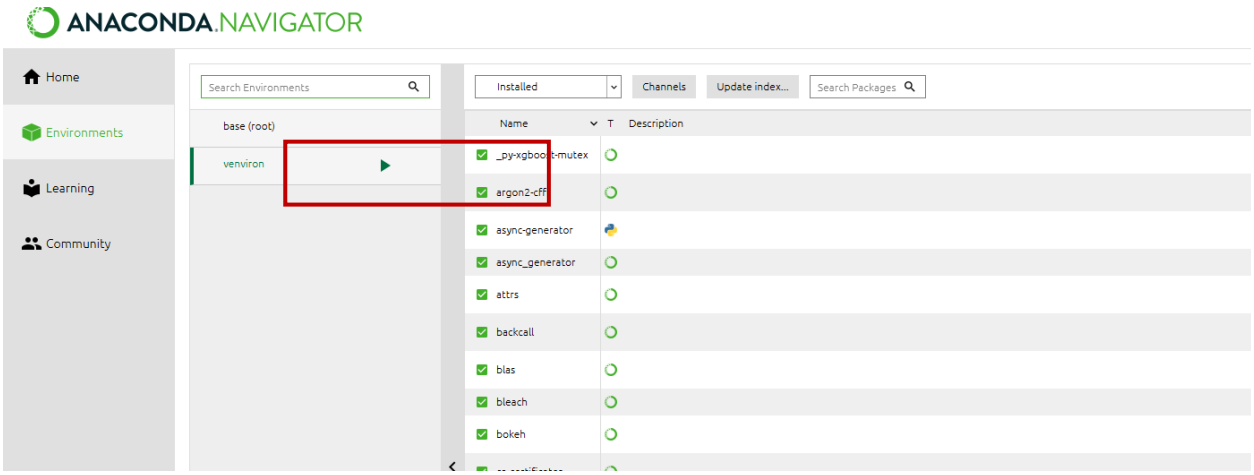


iii. Navigate to the "Environments" Tab on the left and click on the "Create" option at the bottom of the page to create a virtual environment to work with Vyper in

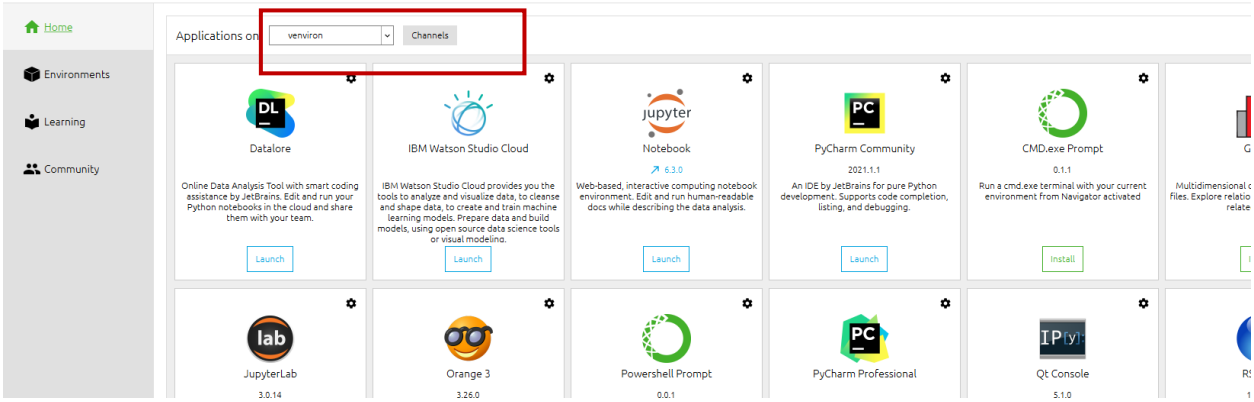


iv. Give a name to your environment, select Python 3.8 and click on Create

v. Once the environment is ready, click on that environment to activate it

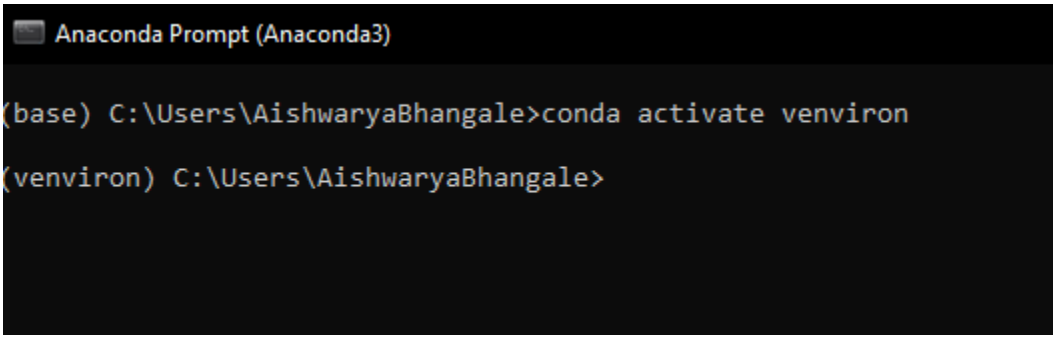


vi. Once activated, click on "Home" tab on the left. You would notice your virtual environment activated as follows :



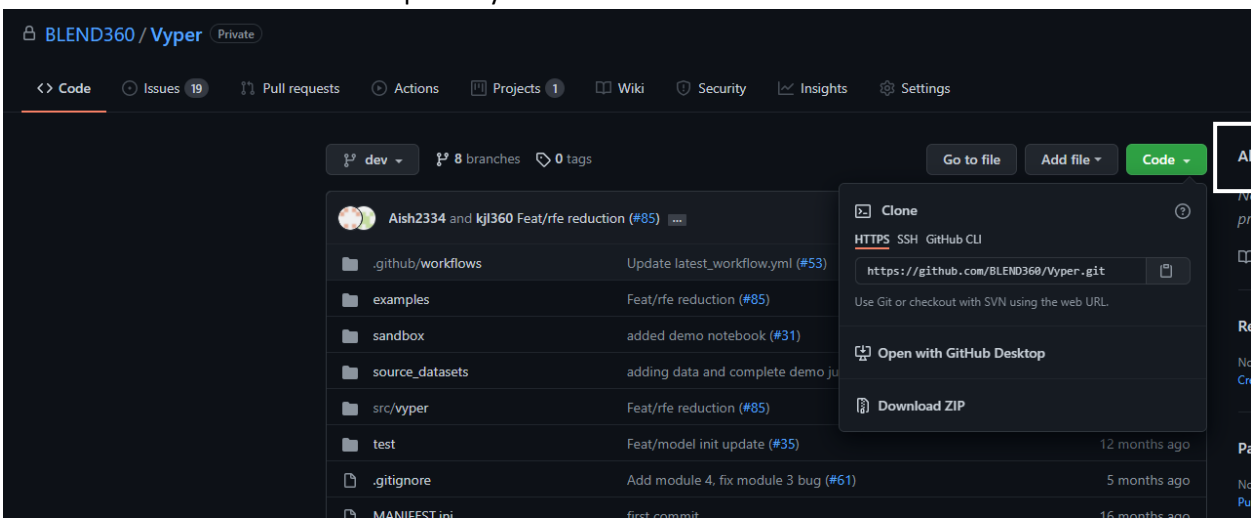
b. Anaconda Powershell Prompt

i. The same environment can be activated from Anaconda prompt as well. Open Anaconda prompt and run the following command:
conda activate your_venv_name



ii. Now, navigate to the directory where you have downloaded and saved your cloned github repository of Vyper 2.0 (BLEND360/Vyper (github.com)). Click on Code button in Github and select "Download Zip". Then unzip the folder in your designated directory(which is the same directory to navigate to from

Anaconda prompt) Following is a screenshot of how you can clone/download a Github repository:




```
(base) C:\Users\AishwaryaBhangale>conda activate venviron

(venviron) C:\Users\AishwaryaBhangale>cd Documents

(venviron) C:\Users\AishwaryaBhangale\Documents>cd Github

(venviron) C:\Users\AishwaryaBhangale\Documents\GitHub>
```

iii. Now, run the following command in Anaconda Prompt to install the Vyper 2.0 Python dependencies:

```
(venviron) C:\Users\AishwaryaBhangale\Documents\GitHub>python -m pip install -e Vyper
Obtaining file:///C:/Users/AishwaryaBhangale/Documents/GitHub/Vyper
Requirement already satisfied: pandas<=1.2.3,>=1.0.3 in c:\users\aishwaryabhangale\.conda\envs\venviron\lib\site-packag
s (from vyper==2.0.0) (1.0.5)
Requirement already satisfied: statsmodels<=0.12.2,>=0.11.1 in c:\users\aishwaryabhangale\.conda\envs\venviron\lib\site
packages (from vyper==2.0.0) (0.12.2)
Collecting scikit-learn<=0.24.1,>=0.23.0
  Using cached scikit_learn-0.24.1-cp38-cp38-win_amd64.whl (6.9 MB)
Requirement already satisfied: scipy<=1.6.2,>=1.5.0 in c:\users\aishwaryabhangale\.conda\envs\venviron\lib\site-package
 (from vyper==2.0.0) (1.6.2)
Requirement already satisfied: openpyxl<=3.0.7,>=2.6.2 in c:\users\aishwaryabhangale\.conda\envs\venviron\lib\site-pack
ges (from vyper==2.0.0) (3.0.7)
Requirement already satisfied: et-xmlfile in c:\users\aishwaryabhangale\.conda\envs\venviron\lib\site-packages (from op
npyxl<=3.0.7,>=2.6.2->vyper==2.0.0) (1.0.1)
Requirement already satisfied: numpy>=1.13.3 in c:\users\aishwaryabhangale\.conda\envs\venviron\lib\site-packages (from
```

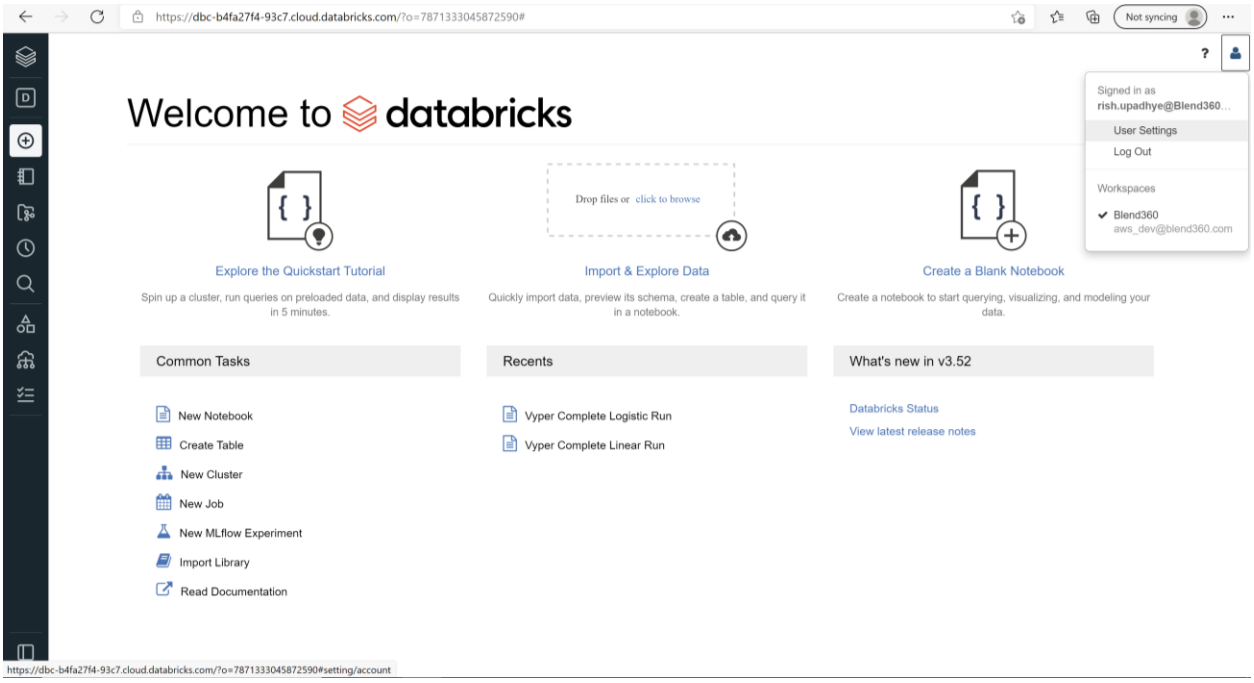
iv. Once you have installed Vyper successfully in your local machine, you can run the following command to launch jupyter notebook from within Anaconda Prompt itself (or you can launch it from Anaconda Navigator too)

```
(venviron) C:\Users\AishwaryaBhangale\Documents\GitHub>jupyter notebook
```

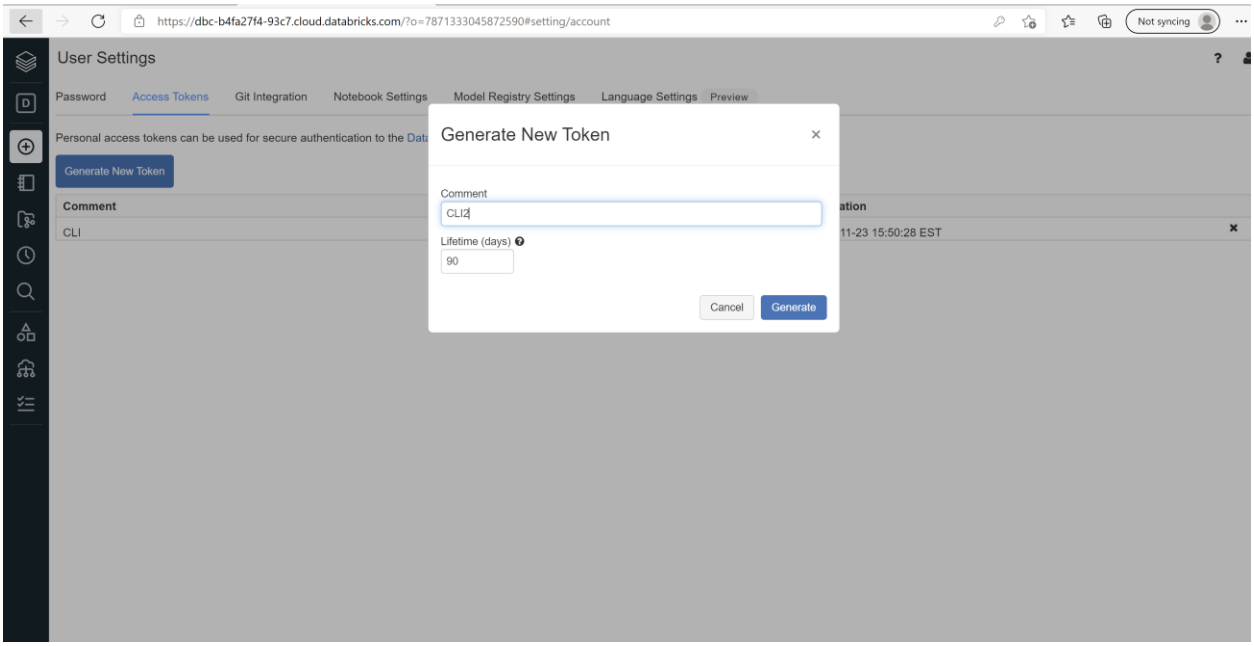
3) XLXS from Databricks to Local Machine

a. Token Creation

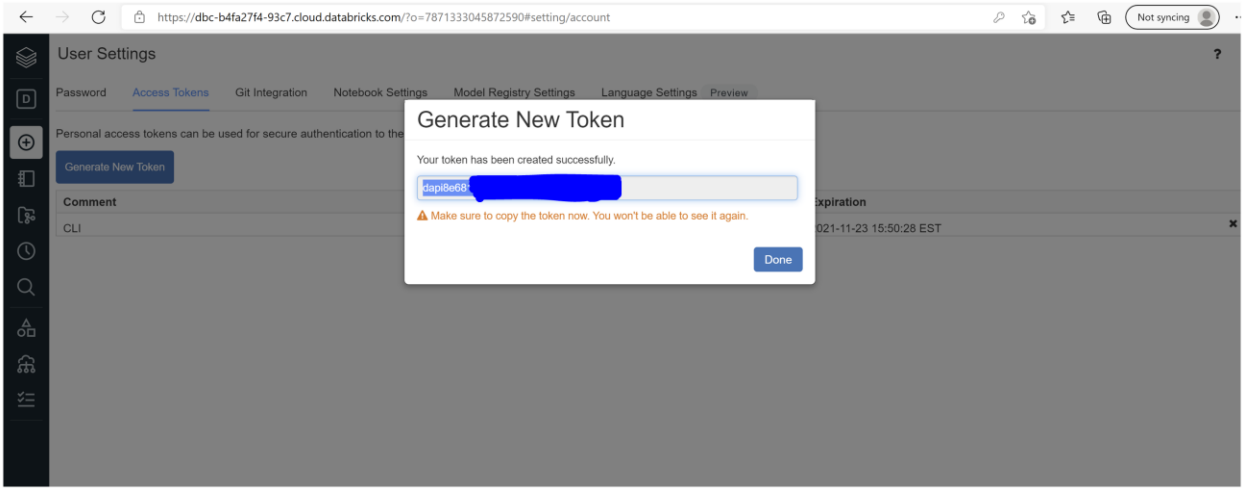
i. Click on **Account** and then **User Settings**



ii. Click on **Access Tokens** Tab and then click on **Generate New Token** button. Name the token and set the Lifetime as per your requirements.

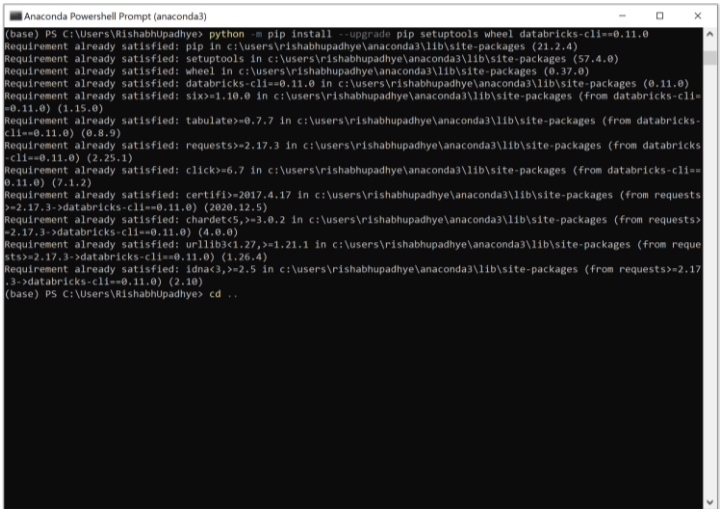


iii. **COPY THE TOKEN. YOU WON'T BE ABLE TO SEE IT AGAIN.**



b. Open Anaconda Powershell Prompt or any other Command Line Interface

i. Install Databricks CLI using this command ‘python -m pip install --upgrade pip setuptools wheel databricks-cli==0.11.0’.



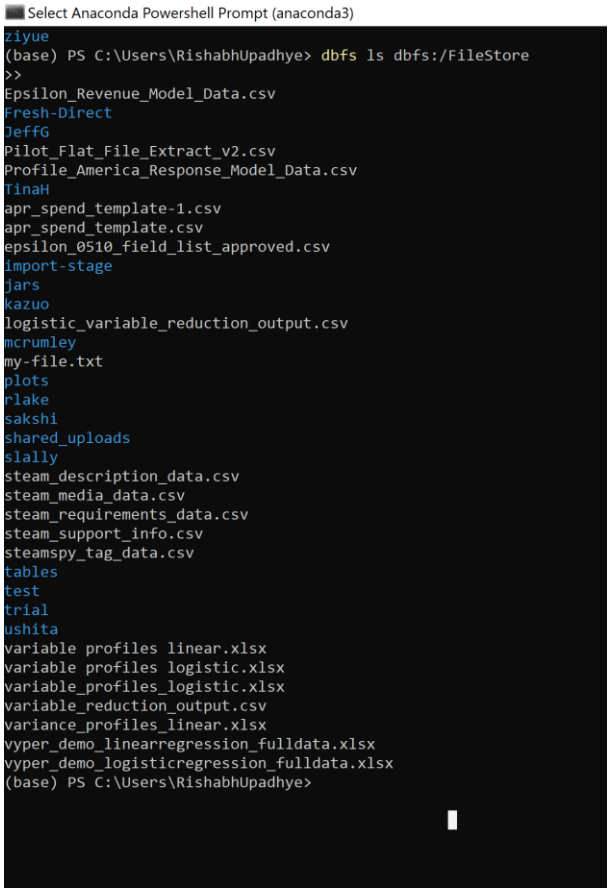
ii. Configure Databricks using token in CLI using this command ‘databricks configure - -token’

```
(base) PS C:\Users\RishabhUpadhye\Documents> databricks configure --token
>>
Databricks Host (should begin with https://):
```

iii. Enter the host as https://dbc-b4fa27f4-93c7.cloud.databricks.com/ and then enter the token.

```
(base) PS C:\Users\RishabhUpadhye\Documents> databricks configure --token
>>
Databricks Host (should begin with https://): https://dbc-b4fa27f4-93c7.cloud.databricks.com/
Token: dapi5[redacted]
```

iv. List all the files in FileStore using ‘dbfs ls dbfs:/FileStore’



v. We will transfer the ‘vyper_demo_linearregression_fulldata.xls’ in our machine. Same process could be followed for the other files.

vi. Run command
‘dbfs cp dbfs:/FileStore/vyper_demo_linearregression_fulldata.xlsx vyper_demo_linearregression_fulldata.xls’

```
(base) PS C:\Users\RishabhUpadhye\Documents> dbfs cp dbfs:/FileStore/vyper_demo_linearregression_fulldata.xlsx vyper_demo_linearregression_fulldata.xls'
```

vii. The file will successfully be copied to your local machine in the Documents folder.

4) References

- Vyper to Databricks meeting with Charles Du - *https://wgcp-my.sharepoint.com/:v:/g/personal/aishwarya_bhangale_blend360_com/EcJCzeN0oURGoiuejCET-ncBPdWuksuJfpysLVUn8Us7dA*
- DBFS Documentation - *https://docs.gcp.databricks.com/data/databricks-file-system.html*
- Databricks Forum - *https://forums.databricks.com/questions/13438/download-a-dbfsfilestore-file-to-my-local-machine.html*
- Azarudeen Shahul Youtube - *https://www.youtube.com/watch?v=PdLpXhK4u8w*