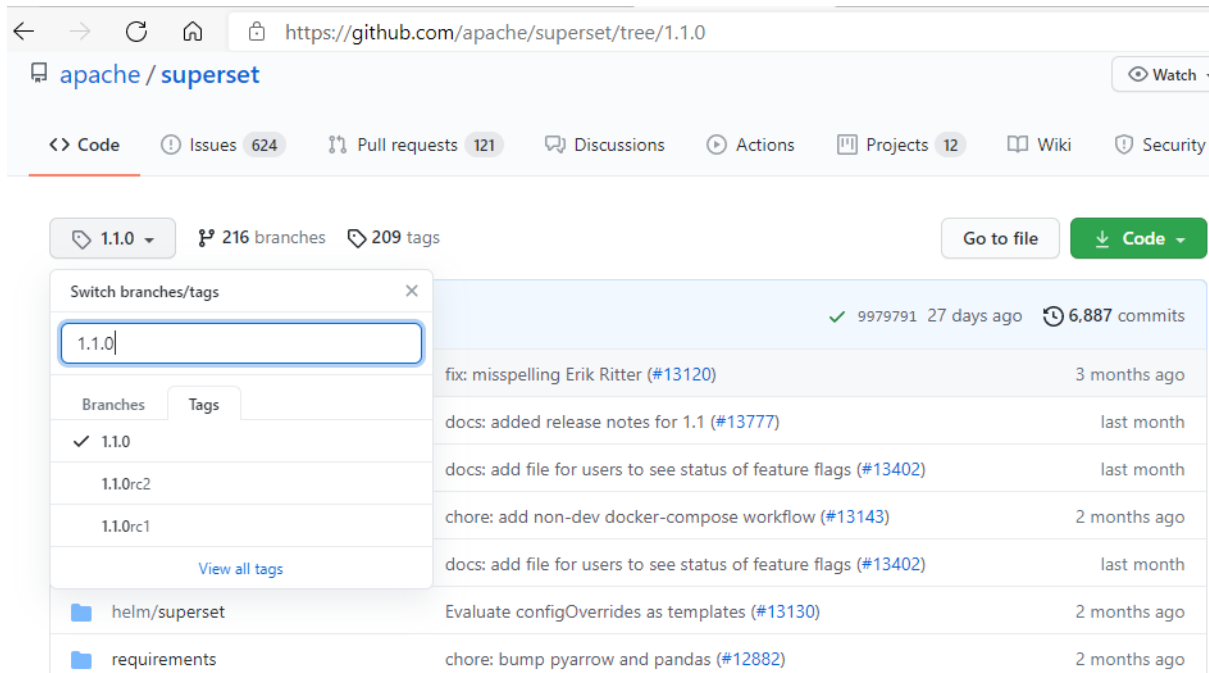


Cài đặt

Dưới đây là tóm tắt quá trình tôi cài đặt superset từ source code phiên bản 1.1.0 trên <https://github.com/apache/superset>, tag 1.1.0.



Tài liệu tham khảo chính thức ở đây:

<https://github.com/apache/superset/blob/master/CONTRIBUTING.md#fask-server>

Tuy nhiên trong quá trình cài đặt có phát sinh vài lỗi và phải giải quyết bằng cách bổ sung thêm vài thư viện (được **bôi vàng**).

Yêu cầu

Ubuntu 20: Nếu bạn định dùng máy ảo trên Windows để cài Ubuntu thì tham khảo [Bài 14 trong Ngày 4](#).

Trong trường hợp bạn dùng Ubuntu thì có khả năng Python được cài là 3.6.x. Bạn cần nâng cấp lên Python 3.8.x. Phiên bản Python tại thời điểm phần này được viết là 3.8.10. Hãy làm theo tài liệu sau:

<https://linuxize.com/post/how-to-install-python-3-8-on-ubuntu-18-04/>

Tiếp tục thực hiện các lệnh sau để backup file /usr/bin/python3 và ánh xạ python3.8 thành python3.

```
sudo cd /usr/bin
sudo mv ./python3 ./python3.bak
```

```
sudo ln -nsf ./python3.8 python3
```

Kiểm tra Python

Mặc định khi cài Ubuntu thì Python phiên bản 3 được cài. Kiểm tra lại bằng lệnh sau:

```
python3 -V
```

```
Python 3.8.5
```

Cài đặt back-end

Thực hiện các lệnh sau:

```
mkdir ~/projects && cd ~/projects

echo Checkout source code version 1.1.0
git clone https://github.com/apache/superset
cd superset
git checkout -b 1.1.0

export VENV_NAME=venv-superset
echo Create a new virtual environment with name: $VENV_NAME
sudo apt install python3-virtualenv
virtualenv $VENV_NAME
source $VENV_NAME/bin/activate

# Upgrade pip
sudo apt install python3-pip
python3 -m pip install --upgrade pip
sudo apt-get install build-essential libmysqlclient-dev
python3-dev python3-dev libsasl2-dev
pip3 install -r requirements/local.txt

pip3 install -e .
superset fab create-admin
superset db upgrade
superset init
superset load-examples
FLASK_ENV=development superset run -p 8088 --with-threads --
reload -debugger
```

Nếu bạn dùng Ubuntu 18 thì có thể thay dòng màu đỏ ở trên bằng 2 lệnh sau:

```
pip3 install virtualenv
sudo virtualenv $VENV_NAME
sudo apt-get install python-dev python3-dev libpython3.8-dev
```

Lệnh cuối cùng ở trên là để khởi động superset với port 8088. Bạn có thể truy cập <http://localhost:8088> ngay trong Ubuntu. Tuy nhiên lúc này superset chỉ có phần backend nên sẽ không có giao diện đầy đủ. Hãy nhấn Ctrl + C để tắt chương trình superset và chuyển sang phần [Cài đặt front-end](#) bên dưới.

Ghi chú

Nếu vì lý do nào đó mà quá trình làm việc của bạn bị gián đoạn, phải login lại thì thực hiện lệnh sau sau khi login để thực hiện tiếp các lệnh:

```
cd ~/projects/superset
export VENV_NAME=venv-superset
source $VENV_NAME/bin/activate
```

Cài đặt front-end

Chuẩn bị

- NodeJS
- npm

Một trong các cách chuẩn bị NodeJs và npm là dùng nvm (<https://github.com/nvm-sh/nvm>) với lệnh sau:

```
curl -o- https://raw.githubusercontent.com/nvm-sh/nvm/v0.37.0/install.sh | bash
```

Logout và login lại. Sau đó chạy lại các lệnh sau:

```
cd ~/projects/superset
export VENV_NAME=venv-superset
source $VENV_NAME/bin/activate
```

Chuẩn bị môi trường nodejs và npm:

```
cd superset-frontend
nvm install
nvm use
```

Cài đặt thư viện

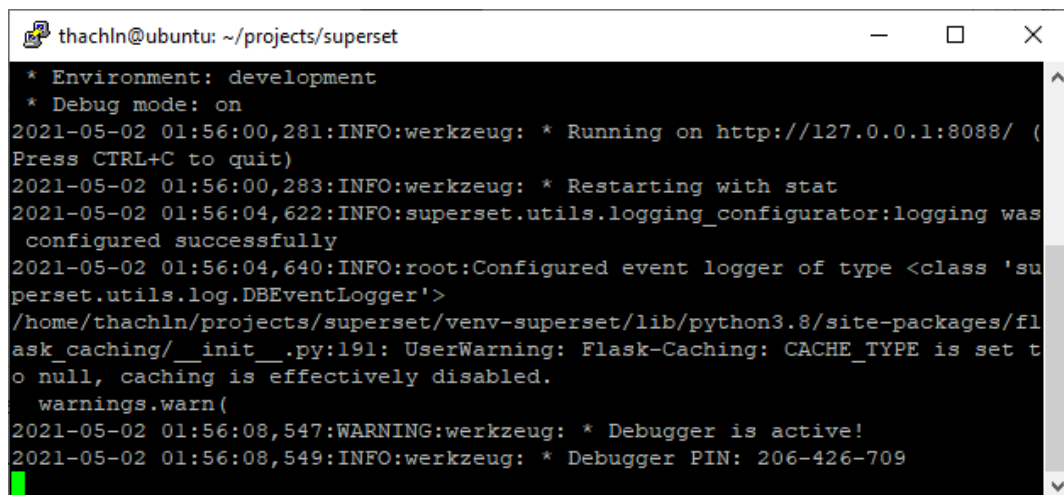
```
cd ~/projects/superset/superset-frontend
# Install dependencies from `package-lock.json`
npm ci
```

Build assets

```
npm run build
```

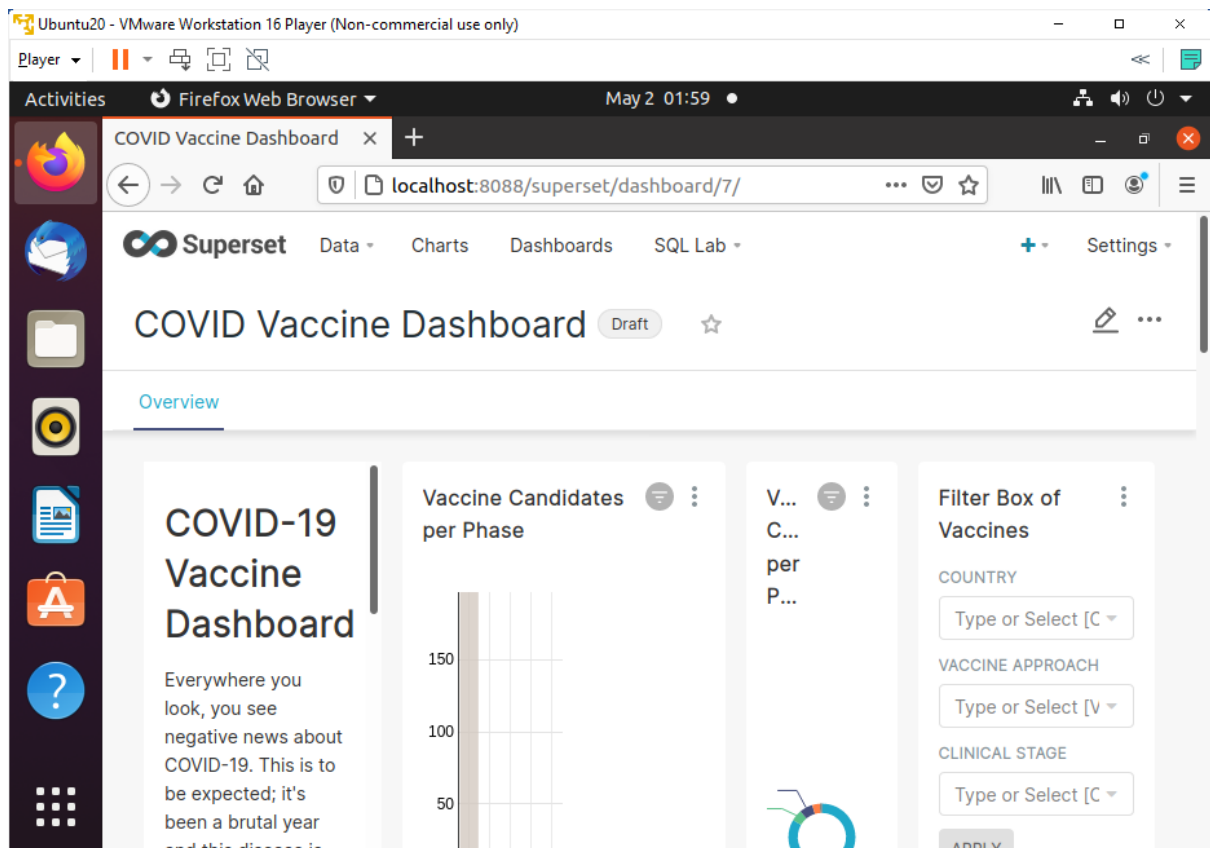
Sau khi biên dịch xong hãy khởi động lại superset bằng lệnh sau:

```
FLASK_ENV=development superset run -p 8088 --with-threads --
reload --debugger
```



```
thachln@ubuntu: ~/projects/superset
* Environment: development
* Debug mode: on
2021-05-02 01:56:00,281:INFO:werkzeug: * Running on http://127.0.0.1:8088/ (
Press CTRL+C to quit)
2021-05-02 01:56:00,283:INFO:werkzeug: * Restarting with stat
2021-05-02 01:56:04,622:INFO:superset.utils.logging_configurator:logging was
configured successfully
2021-05-02 01:56:04,640:INFO:root:Configured event logger of type <class 'su
perset.utils.log.DBEventLogger'>
/home/thachln/projects/superset/venv-superset/lib/python3.8/site-packages/fl
ask_caching/__init__.py:191: UserWarning: Flask-Caching: CACHE_TYPE is set t
o null, caching is effectively disabled.
  warnings.warn(
2021-05-02 01:56:08,547:WARNING:werkzeug: * Debugger is active!
2021-05-02 01:56:08,549:INFO:werkzeug: * Debugger PIN: 206-426-709
```

Nếu Ubuntu của bạn cài đặt ở chế độ có giao diện thì hãy đăng nhập và trải nghiệm superset với link <http://localhost:8088> ngay trên Ubuntu. Thông tin đăng nhập do bạn thiết lập trong quá trình cài đặt back-end.



Khởi động lại superset

Vì tôi chạy Ubuntu với VMware trên Windows khi cần sử dụng superset thì bật máy lên và thực thi các lệnh sau:

```
cd ~/projects/superset
export VENV_NAME=venv-superset
source $VENV_NAME/bin/activate
FLASK_ENV=development superset run -p 8088 --with-threads --
reload --debugger
```

Cài đặt và cấu hình Nginx

Phần trước bạn đã biết cách cài đặt và khai thác superset qua URL <http://localhost:8088> ngay trên máy Ubuntu đã cài superset. Tình huống đặt ra cho bạn là:

① Nếu bạn muốn truy ứng dụng superset thông qua web server như nginx với URL như <http://localhost> thì phải làm sao?

② Mở rộng một chút ý ở trên: Nếu bạn muốn truy cập từ một máy khác thông qua địa chỉ IP của máy đang chạy Ubuntu thì làm sao? Cụ thể trong trường hợp này là nếu dùng trình duyệt từ máy Windows để truy cập vào superset trên máy chạy Ubuntu thì làm sao?

Một trong cách giải quyết cho hai tình huống trên là sử dụng Nginx (như đã gợi ý trong tình huống ①).

Cài đặt Nginx

Chạy lệnh sau trong Ubuntu:

```
sudo apt install nginx
```

Cấu hình Nginx

Chỉnh sửa file cấu hình mặc định của nginx bằng lệnh sau:

```
sudo nano /etc/nginx/sites-available/default
```

```
#         location / {
#             # First attempt to serve request as file, then
#             # as directory, then fall back to displaying a 404.
#             try_files $uri $uri/ =404;
#         }
#         location / {
#             proxy_pass http://127.0.0.1:8088;
#             include /etc/nginx/proxy_params;
#         }
```

Phần **màu đỏ** là chú thích lại (thêm dấu # phía trước mỗi dòng).

Phần **màu xanh** là thêm mới.

Nhấn phím Ctrl+O để lưu file. Ctrl+X để thoát trình soạn thảo nano.

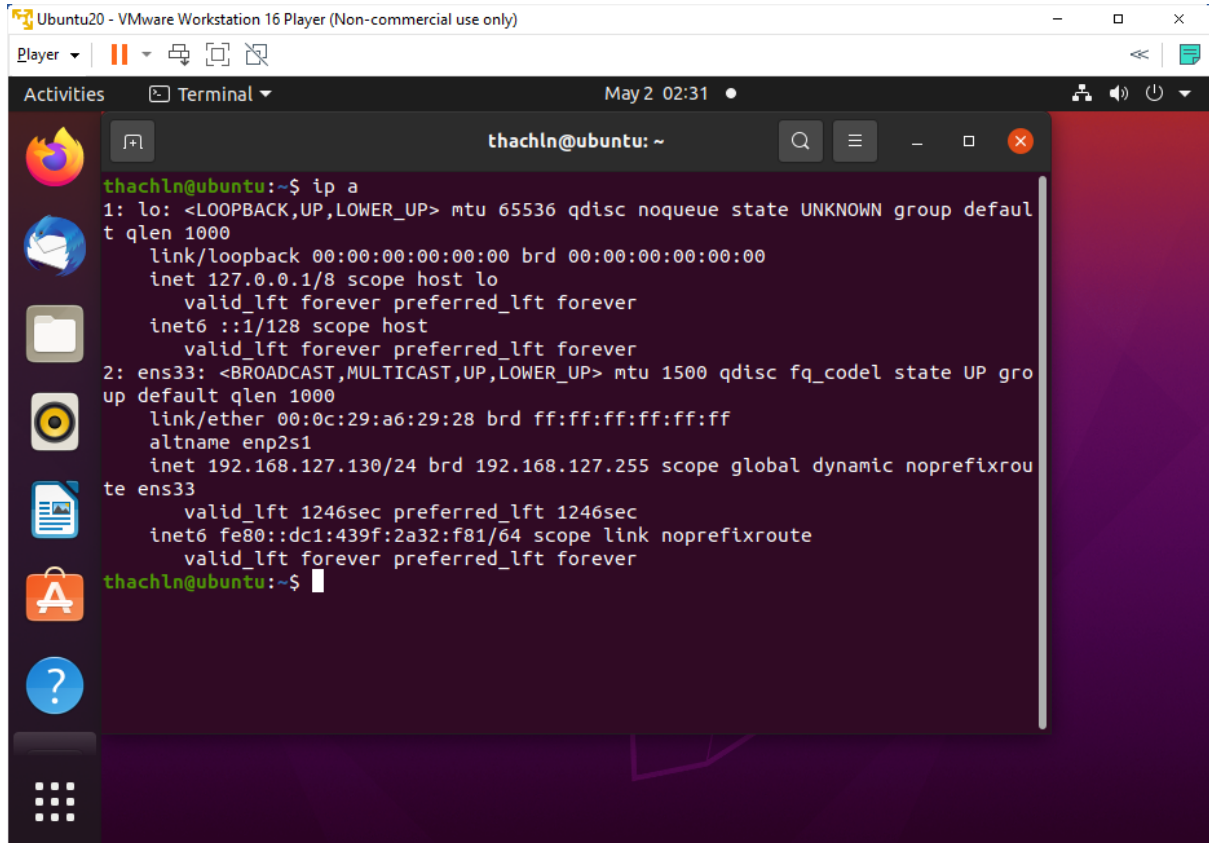
Khởi động lại nginx

```
sudo service nginx restart
```

Truy cập nginx trên Ubuntu từ máy Windows

Xem địa chỉ IP của máy Ubuntu bằng lệnh sau:

```
ip a
```

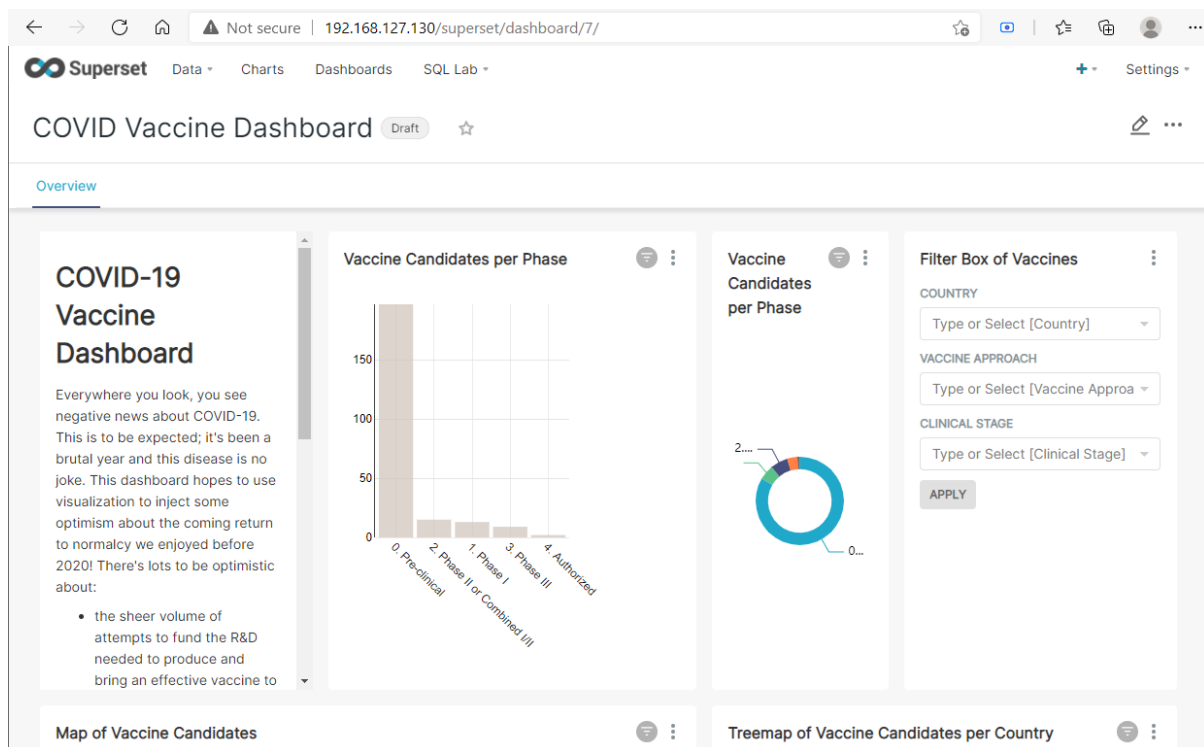


The screenshot shows a terminal window titled "thachln@ubuntu: ~" within a VMware Workstation 16 Player. The terminal displays the output of the command "ip a". The output shows details for the loopback interface "lo" (127.0.0.1) and the ethernet interface "ens33" (192.168.127.130). The IP address 192.168.127.130 is highlighted in the original image.

```
thachln@ubuntu:~$ ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: ens33: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 00:0c:29:a6:29:28 brd ff:ff:ff:ff:ff:ff
    altname enp2s1
    inet 192.168.127.130/24 brd 192.168.127.255 scope global dynamic noprefixroute ens33
        valid_lft 1246sec preferred_lft 1246sec
    inet6 fe80::dc1:439f:2a32:f81/64 scope link noprefixroute
        valid_lft forever preferred_lft forever
thachln@ubuntu:~$
```

Ubuntu trên máy tôi có địa chỉ IP là: 192.168.127.130.

Vì vậy đường dẫn để trải nghiệm superset là <http://192.168.127.130>



Khai thác superset

Cài đặt MySQL

[Tham khảo cài đặt MySQL trên Ubuntu tại đây.](#)

Tạo database

Thực hiện 3 lệnh sau trong cửa sổ lệnh MySQL:

```
CREATE DATABASE mysuperset DEFAULT CHARACTER SET 'UTF8MB4';
CREATE USER 'mysuperset_user'@'%' IDENTIFIED BY
'Mysuperset#123';
GRANT ALL PRIVILEGES ON mysuperset.* TO 'mysuperset_user'@'%'
WITH GRANT OPTION;
```

Thử kết nối database bằng mysqlclient

Thực hiện lệnh sau trong cửa sổ lệnh của hệ điều hành (Ubuntu):

```
mysql -u mysuperset_user -pMysuperset#123 mysuperset
```

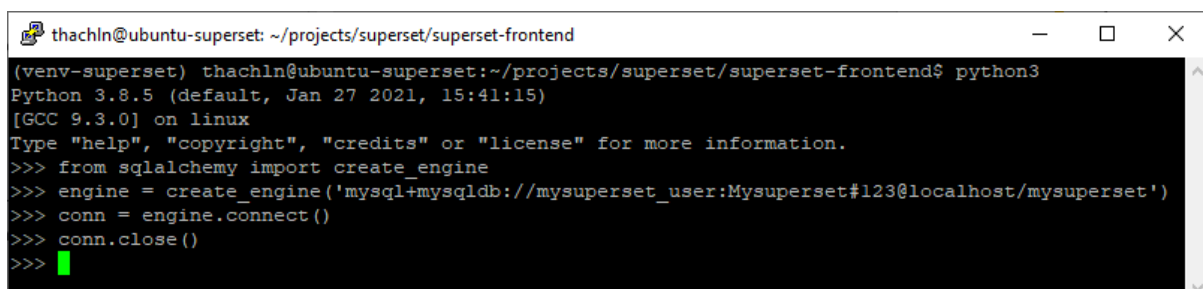
Thử test nối database bằng Python ngay trên máy Ubuntu

Mở cửa sổ lệnh Python bằng lệnh sau:

```
python3
```

Thực hiện các lệnh sau trong cửa sổ lệnh Python:

```
from sqlalchemy import create_engine
engine =
create_engine('mysql+mysqldb://mysuperset_user:Mysuperset#123@192.168.
127.131/mysuperset?charset=utf8mb4')
conn = engine.connect()
conn.close()
```



```
thachln@ubuntu-superset: ~/projects/superset/superset-frontend
(venv-superset) thachln@ubuntu-superset:~/projects/superset/superset-frontend$ python3
Python 3.8.5 (default, Jan 27 2021, 15:41:15)
[GCC 9.3.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> from sqlalchemy import create_engine
>>> engine = create_engine('mysql+mysqldb://mysuperset_user:Mysuperset#123@localhost/mysuperset')
>>> conn = engine.connect()
>>> conn.close()
>>>
```

Thử kết nối database bằng Python trên máy Windows

Cấu hình MySQL Server

Mặc định MySQL được cấu hình chỉ chạy trên localhost, tức địa chỉ IP là 127.0.0.1. Để MySQL được truy cập từ máy khác trong mạng thì sửa file cấu hình bằng lệnh sau:

```
sudo nano /etc/mysql/mysql.conf.d/mysqld.cnf
```

Tìm dòng **bind-address** sửa lại như sau:

```
bind-address            = 0.0.0.0
```

Sau đó khởi động lại mysql server bằng lệnh:

```
sudo service mysql restart
```

Cài thư viện cho môi trường python (trên client)

```
pip install mysqlclient
```

Mã nguồn Python

```
from sqlalchemy import create_engine
engine =
create_engine('mysql+mysqldb://mysuperset_user:Mysuperset#123@192.168.
127.131/mysuperset')
engine.connect()
engine.close()
```

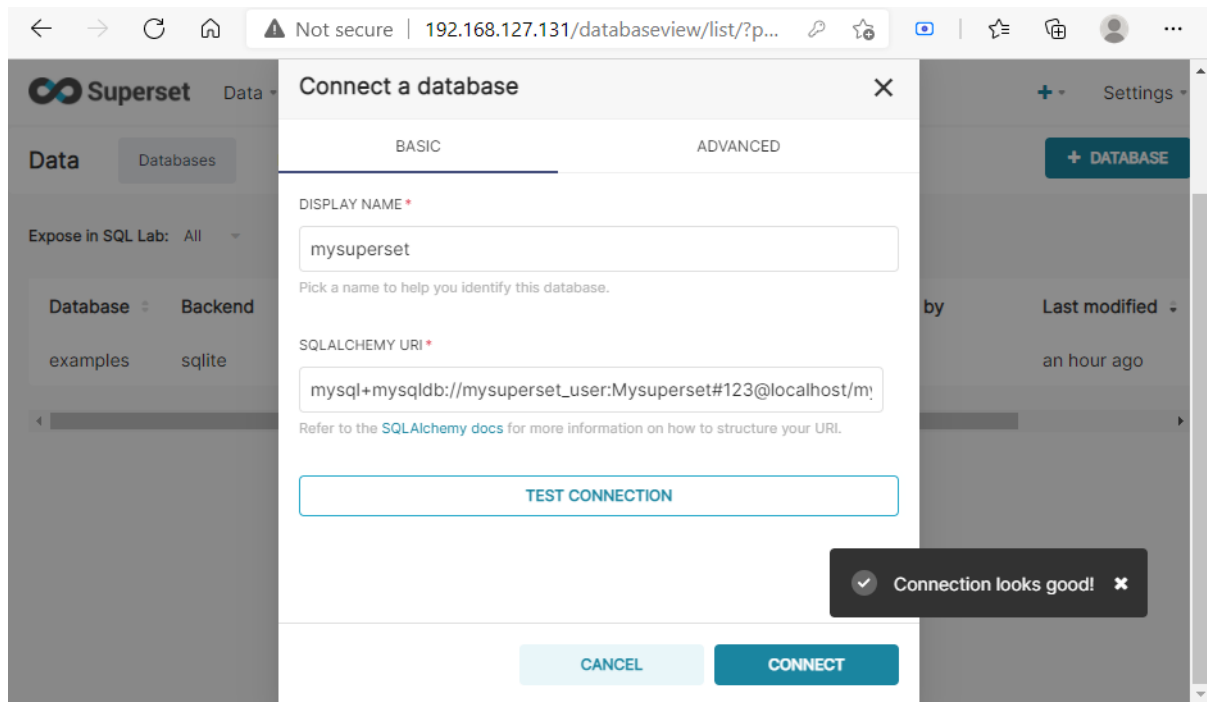
(Hãy thay địa chỉ IP bằng địa chỉ IP đúng trên máy Ubuntu của bạn)

Tạo Database trong superset

Tạo database với 2 thông số sau:

DISPLAY NAME	mysuperset (<i>bạn có thể đặt tên khác</i>)
SQL ALCHEM Y URI	mysql+mysqldb://mysuperset_user:Mysuperset#123@localhost/mysu perset

Bấm nút “TEST CONNECTION” để kiểm tra kết nối.



Thử bậc “Allow data upload”

Vào tab “ADVANCED”, đánh dấu vào mục “Allow data upload”.

Edit database

BASIC

ADVANCED

SQL Lab

Configure how this database will function in SQL Lab.

Performance

Adjust settings that will impact the performance of this database.

Security

Add connection information for other systems.

Other

Additional settings.

☐ Impersonate Logged In User (Presto & Hive)

☒ Allow data upload

EXTRA

```

1 {
2   "metadata_params": {},
3   "engine_params": {},
4   "metadata_cache_timeout": {},
5   "schemas_allowed_for_csv_upload": []
6 }
7

```

JSON string containing extra configuration elements.

1. The engine_params object gets unpacked into the sqlalchemy.create_engine call,

CANCEL

SAVE

Nếu bạn không gặp lỗi gì thì coi như đã rất may mắn 😊!