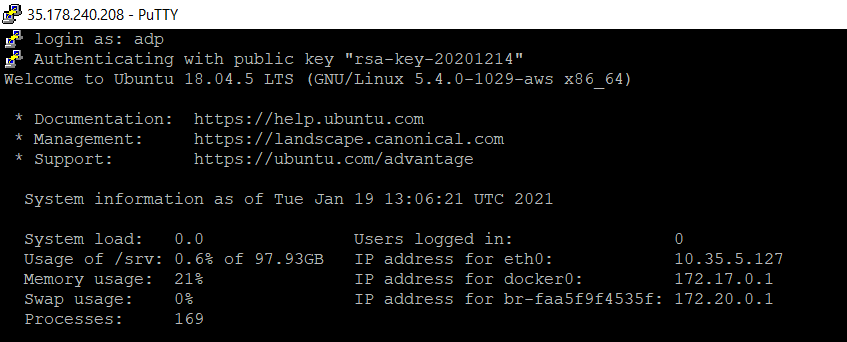
**Apache superset and airflow installation Dev Environment**

**Superset installation steps**

1. Login with dev environment using putty with username and password provided by Alex from medic.



1. Docker and Docker compose is installed on Dev environment by Medic.
2. Change the directory to /srv and create a new directory as apache-superset and airflow

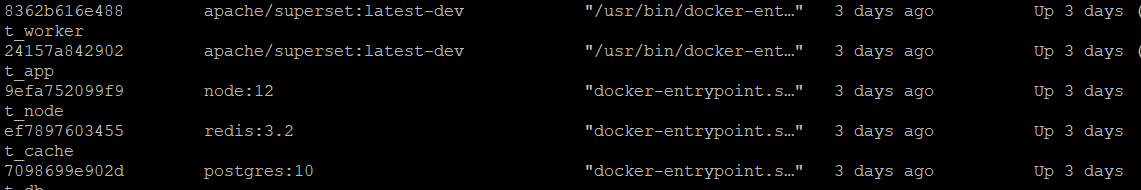
* cd /srv/
* sudo mkdir apache-superset
* sudo mkdir airflow

1. Execute the below commands for superset:

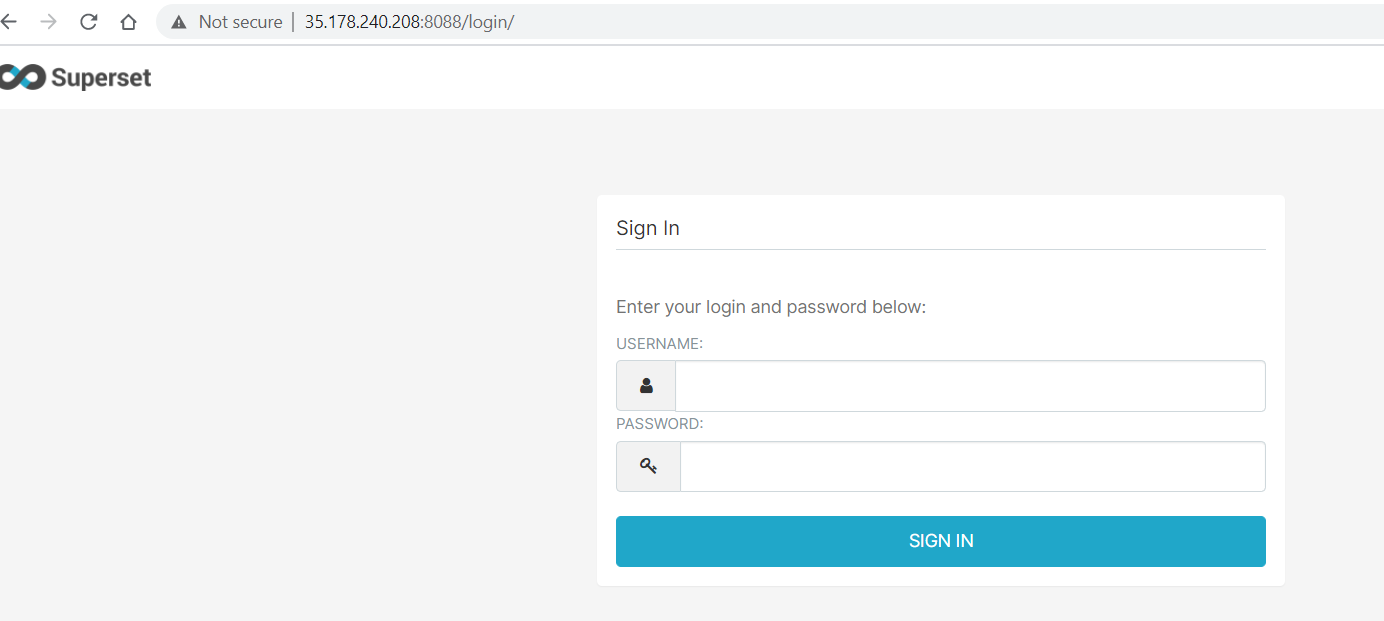
* cd apache-superset
* Clone the apache superset git repo
* git clone <https://github.com/apache/superset.git>
* cd superset
* sudo docker-compose up (already the port mapping to the host is 8088)

All dependant containers for superset are also created.

* sudo docker ps (check all the containers are up and running)



* Curl -v <http://localhost:8088>
* Finally access the superset outside using <http://35.178.240.208:8088> on the browser.



Username: admin

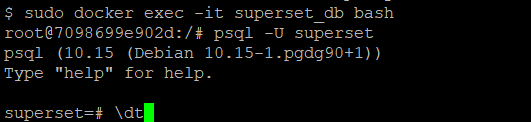
Password: admin

* To access the postgres db container to execute the db scripts below are the commands
* sudo docker exec -it superset\_db bash (and you will be in the db container)
* psql -U superset

username: superset (Postgres)

password: superset (postgres)

database: superset (postgres)



Issues faced during superset installation.

1. Tried creating superset and postgres containers individually with respective images was unable to set up connection between them since they both were having different networks.

* So, went ahead with docker-compose installation as suggested by the apache-superset official document. Since the docker-compose refers to folders locally so cloned the entire git repo(apache-superset).

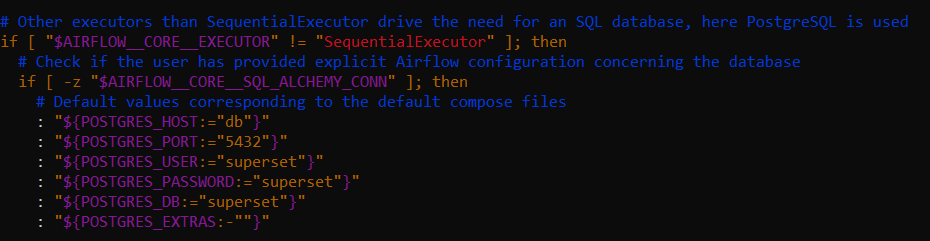
**Airflow Installation steps**

1. Once logged into dev environment, go to the folder under /srv/airflow

* cd /srv/airflow

1. Cloned the airflow repo to do the changes to accommodate db connectivity in airflow.cfg and in docker entrypoint.sh files.

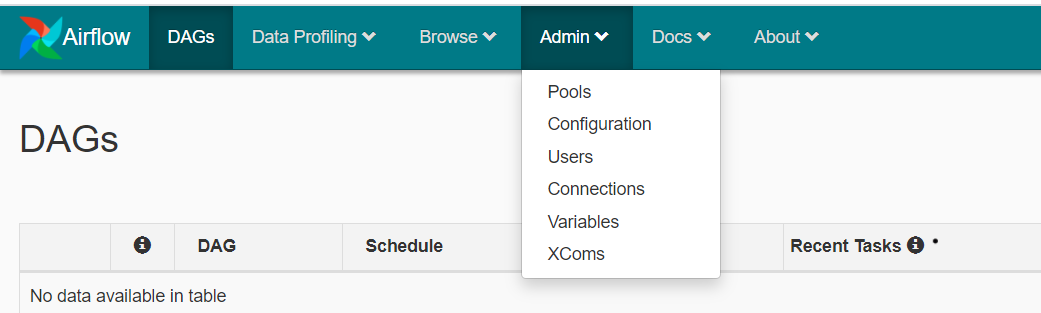
* git clone <https://github.com/puckel/docker-airflow.git>
* cd config
* sudo vi airflow.cfg and modify the postgres connection details
* 
* Similarly, in entrypoint.sh under “script” folder.



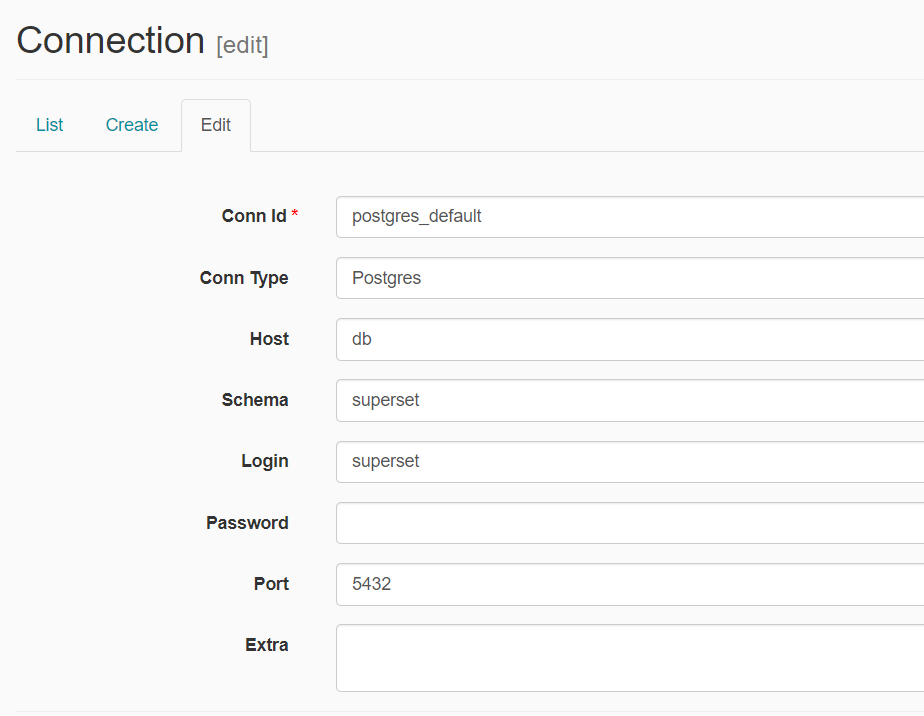
* Build the docker image now after the changes
* sudo docker image build -t airflow-test .
* Create the container from the image created on previous step.
* sudo docker run -d -p 9000:8081 --network=superset\_default --link superset\_db:db --name airflow airflow-test
* Check if the container is up and running
* sudo docker ps



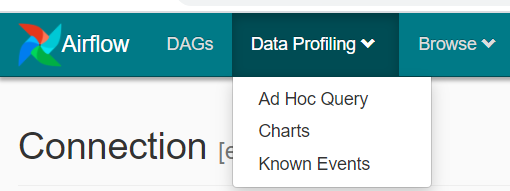
* Finally browse the url , <http://35.178.240.208:9000>
* And set postgres connection , under the connections tab



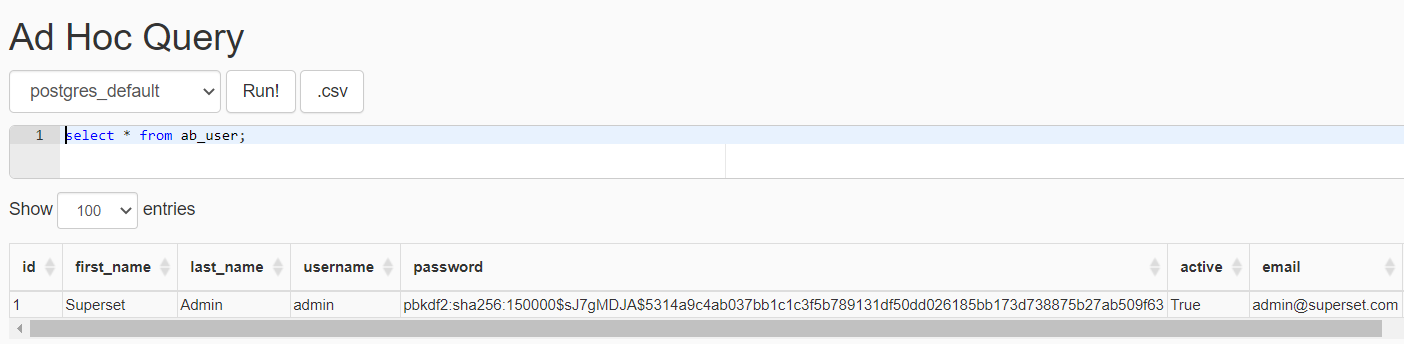
* Then choose “postgres\_default” and edit to connect to the superset db.



* Then browse to the tab “Ad hoc query”.



* Finally test the connection selecting the “postgres\_default” from the dropdown.



Issues encountered

1. Airflow container was not able to connect the postgres container.

* Had to execute the airflow container to join the postgres container network.

sudo docker run -d -p 9000:8081 --network=superset\_default --link superset\_db:db --name airflow airflow-test

1. Container was throwing sqlalchemy error .

* Modified the Dockerfile. Added the below lines to Dockerfile.

