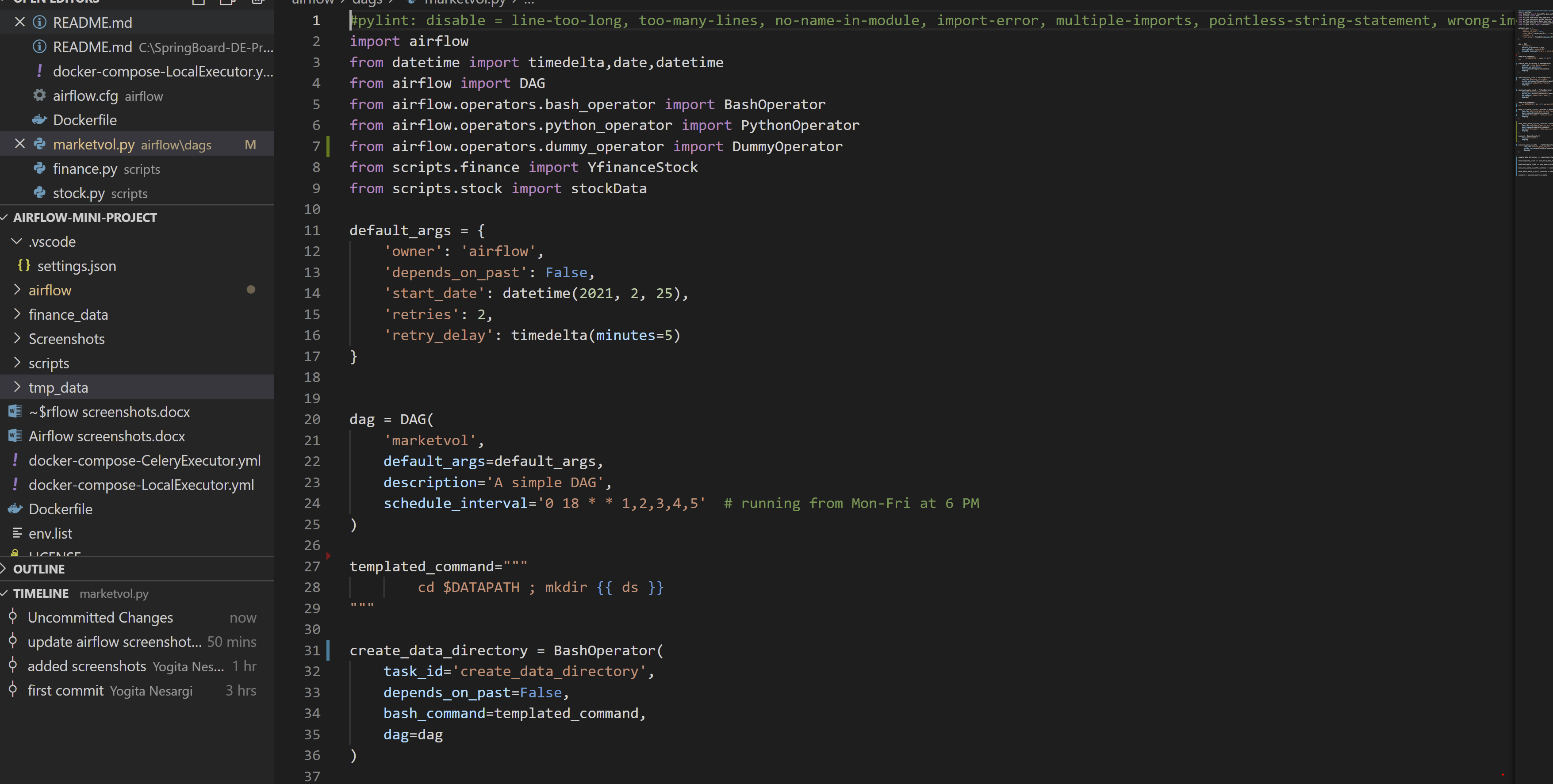
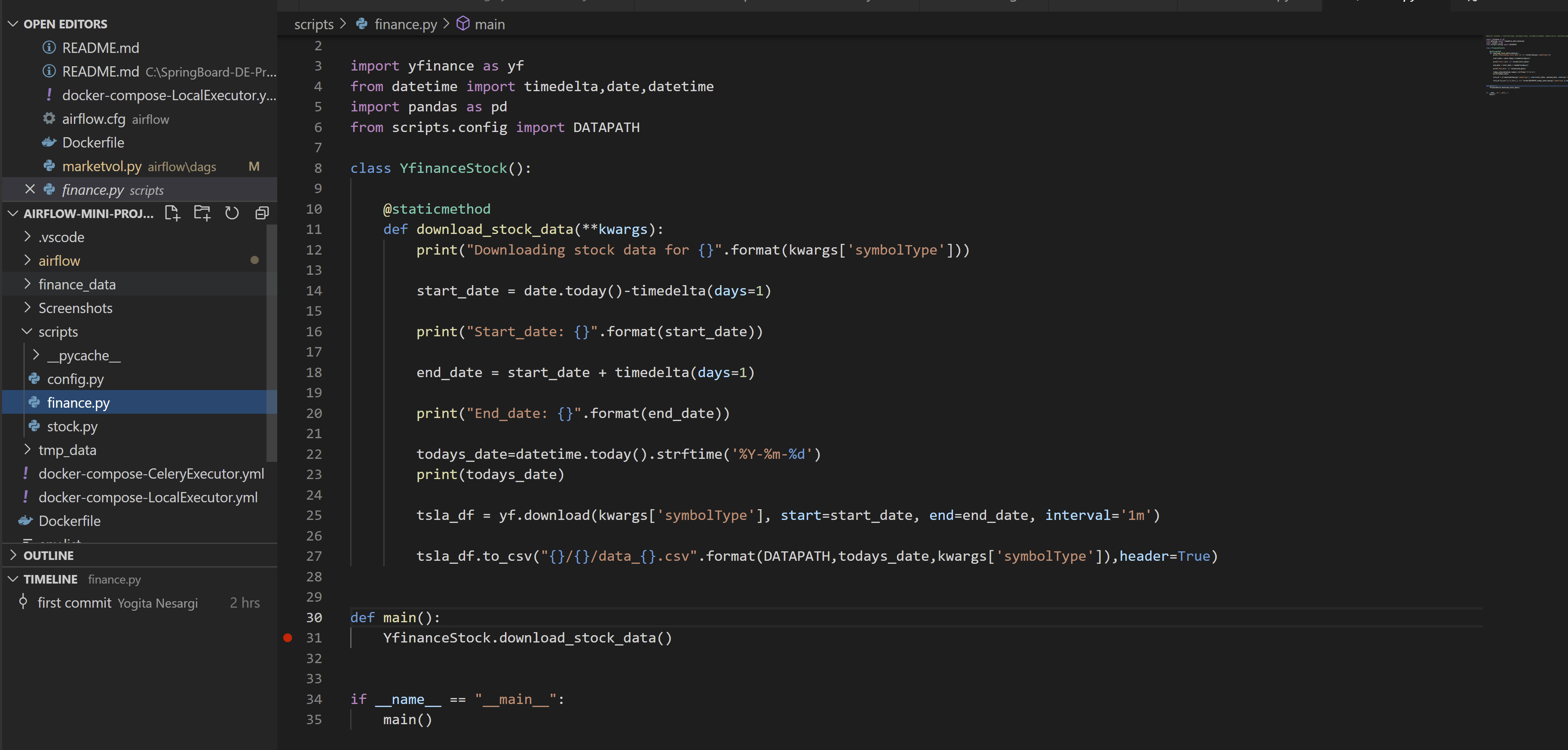
Project structure:

An Airflow dag ‘marketvol’ is defined under ‘./airflow/dags’ folder.

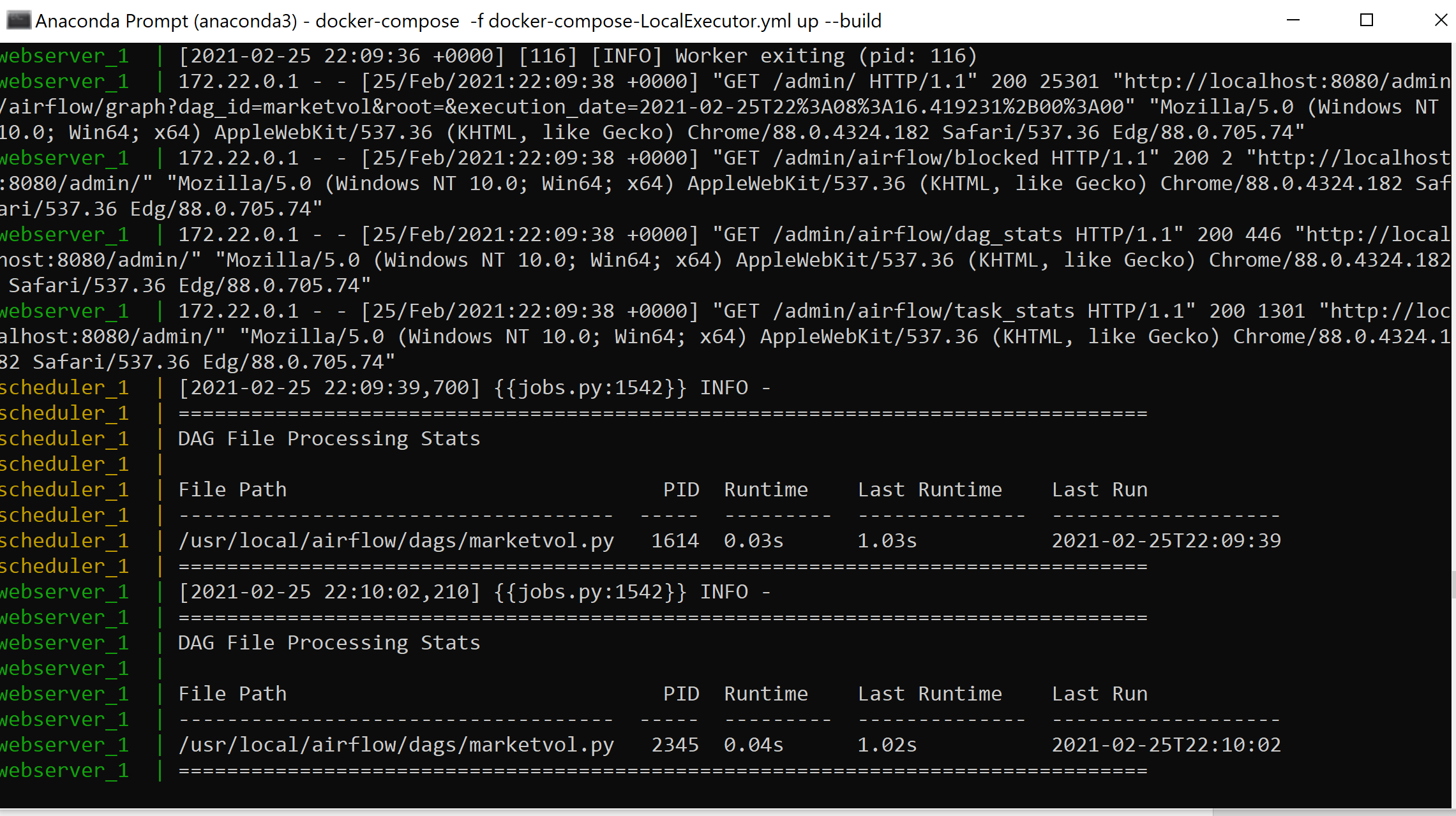


Scripts are under: ‘. /scripts’ folder

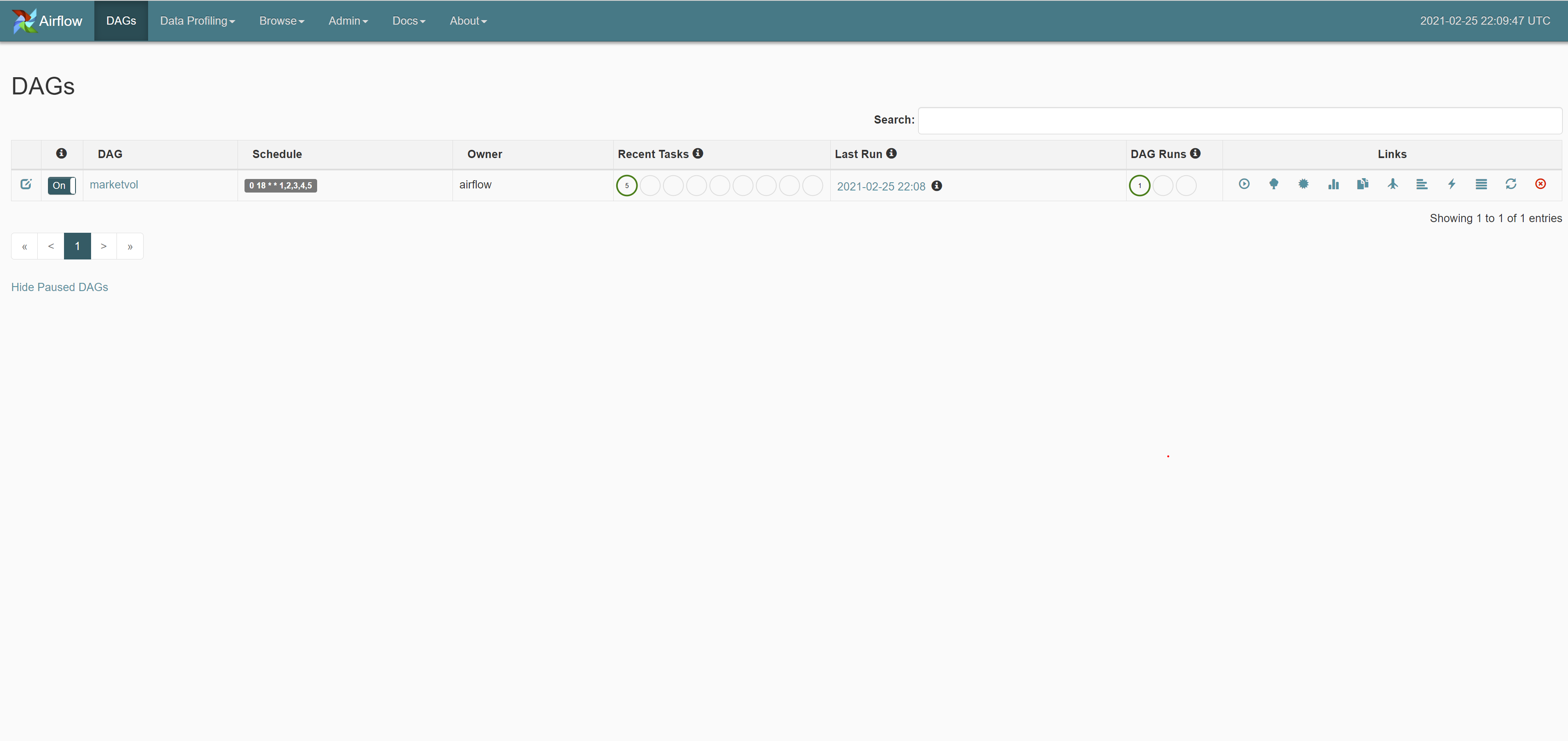


Starting docker container

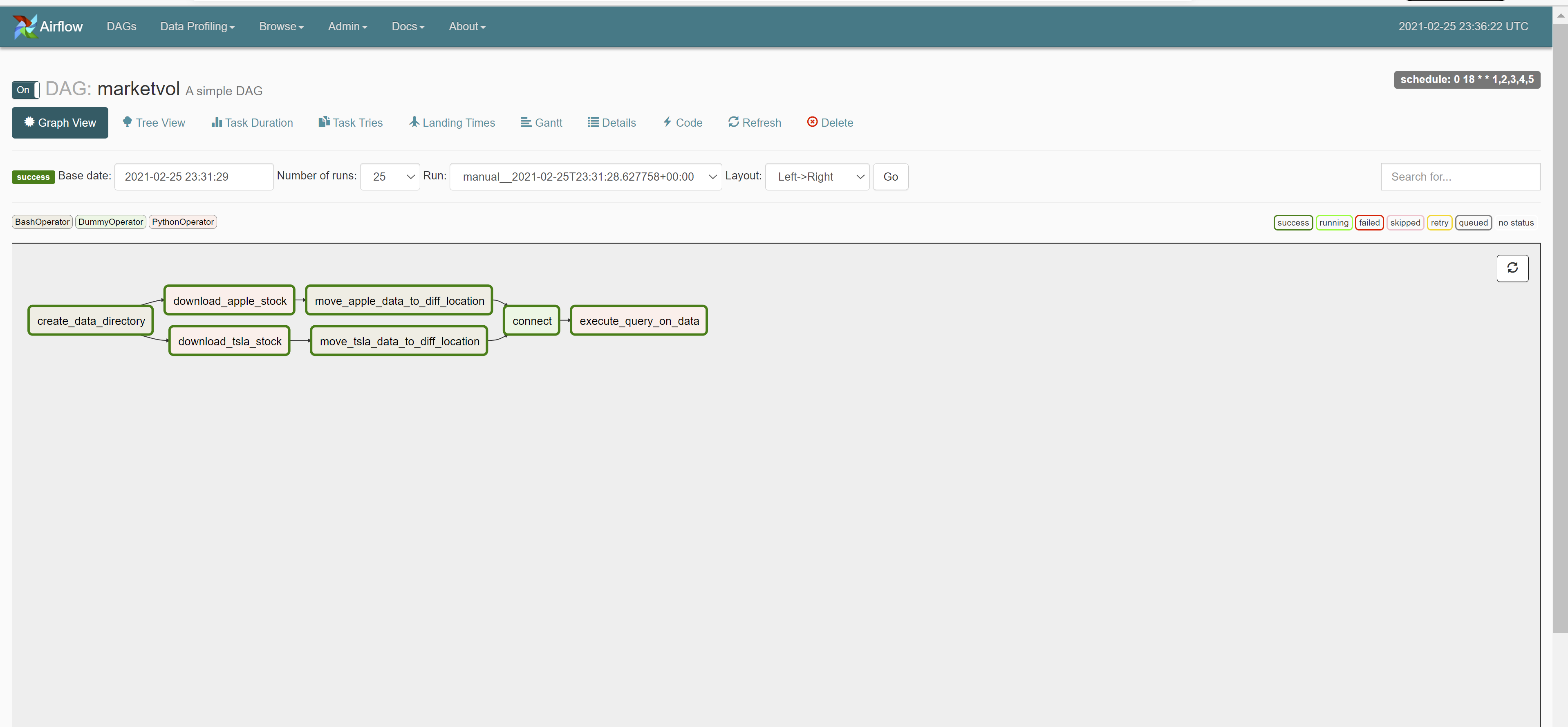
Command: docker-compose -f docker-compose-LocalExecutor.yml up --build



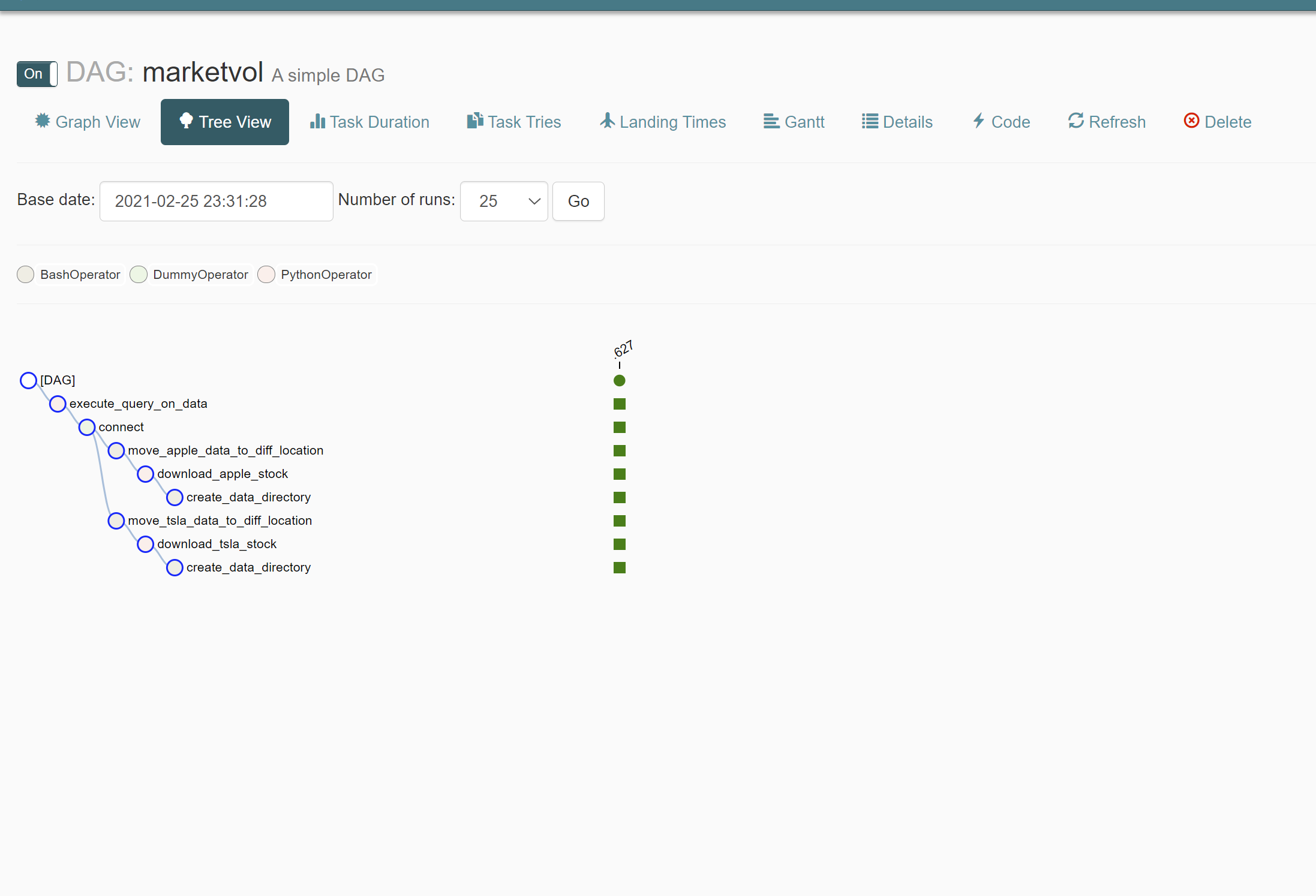
Access Airflow UI using ‘http://localhost:8080’



Graph View



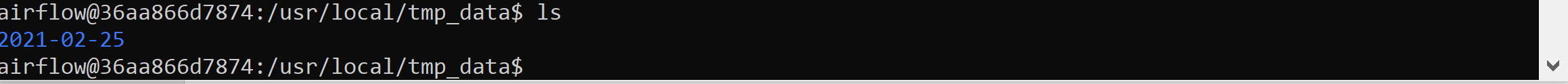
Tree View

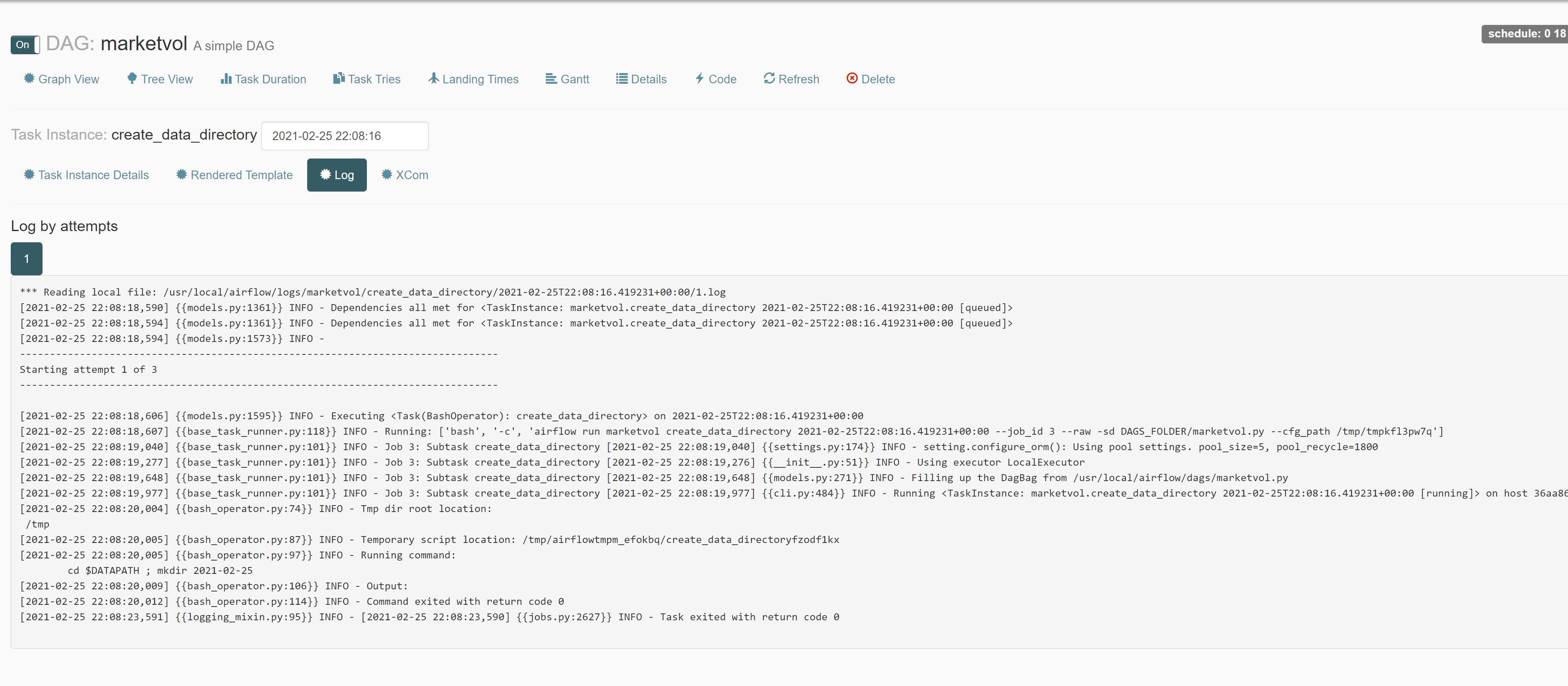


Logs of individual task

Task 1: create\_data\_directory

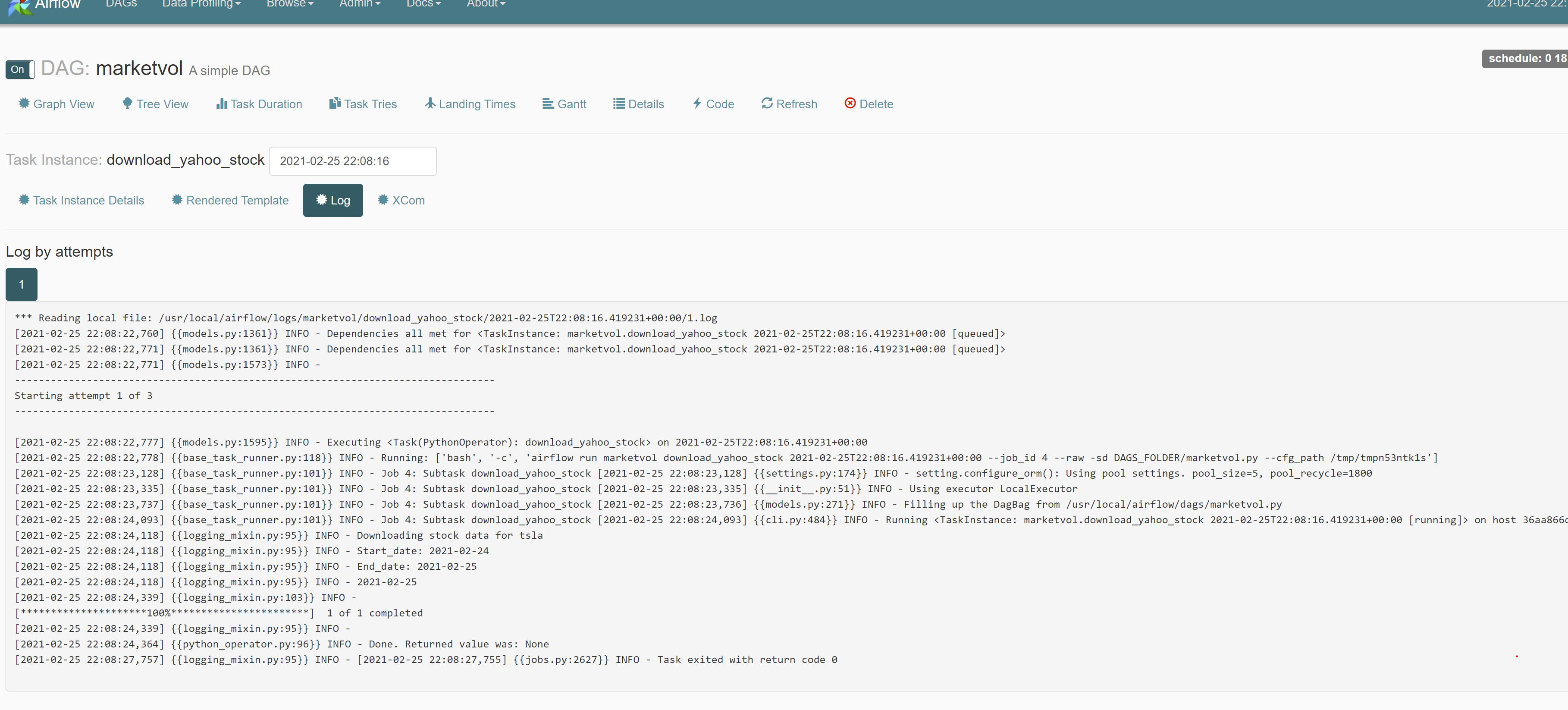
The Airflow BashOperator is used to create a directory on current\_date: 2021-02-25





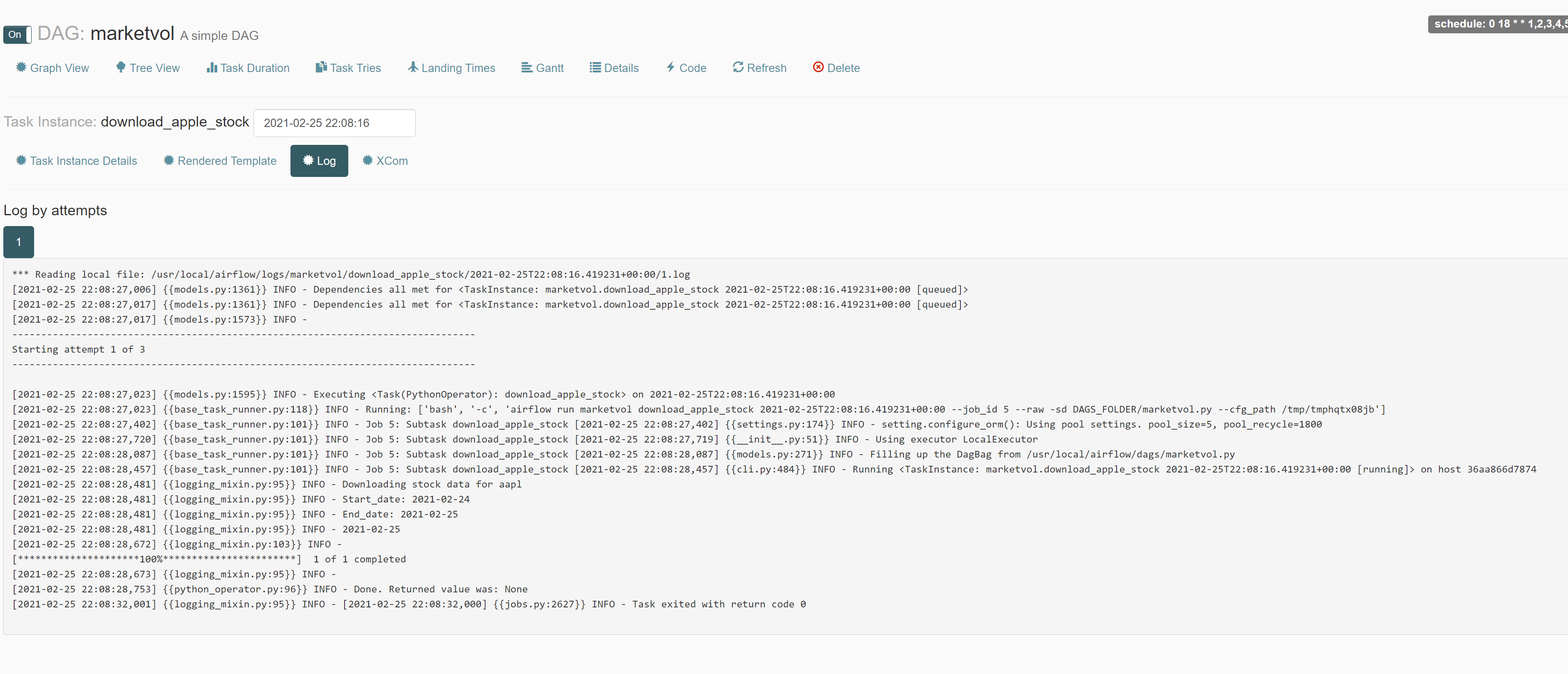
Task 2: download\_yahoo\_stock

The Airflow PythonOperator is used to download yahoo stock data to temporary path ‘/usr/local/tmp\_data/2021-02-25’



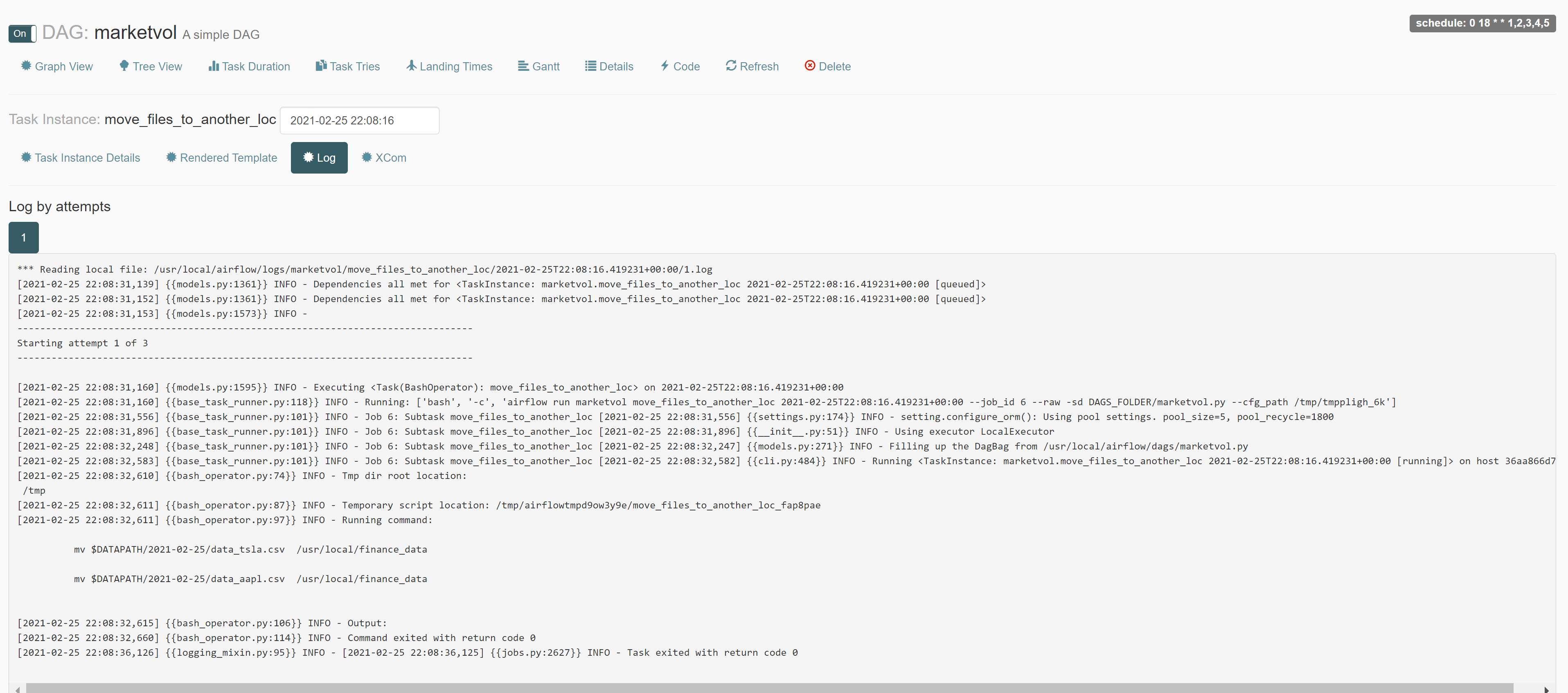
Task 3: download\_apple\_stock

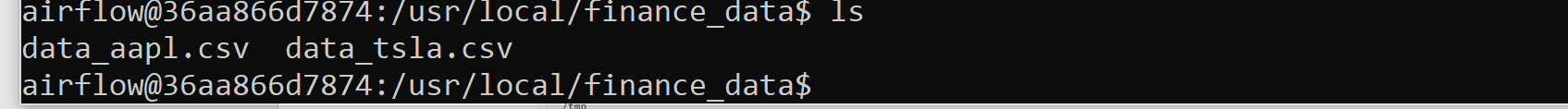
The Airflow PythonOperator is used to download apple stock to temporary path ‘/usr/local/tmp\_data/2021-02-25’



Task 4: move\_tsla\_data\_to\_diff\_location

The Airflow BashOperator is used to move the data from temporary storage to another location ‘/usr/local/finance\_data’





Task 5: move\_apple\_data\_to\_diff\_location

The Airflow BashOperator is used to move the data from temporary storage to another location ‘/usr/local/finance\_data’

Task 6: A dummy operator to connect the two previous task to last task

Task 7: execute\_query\_on\_data

Finally, Airflow ‘PythonOperator’ to run query on the above two dataframes.

