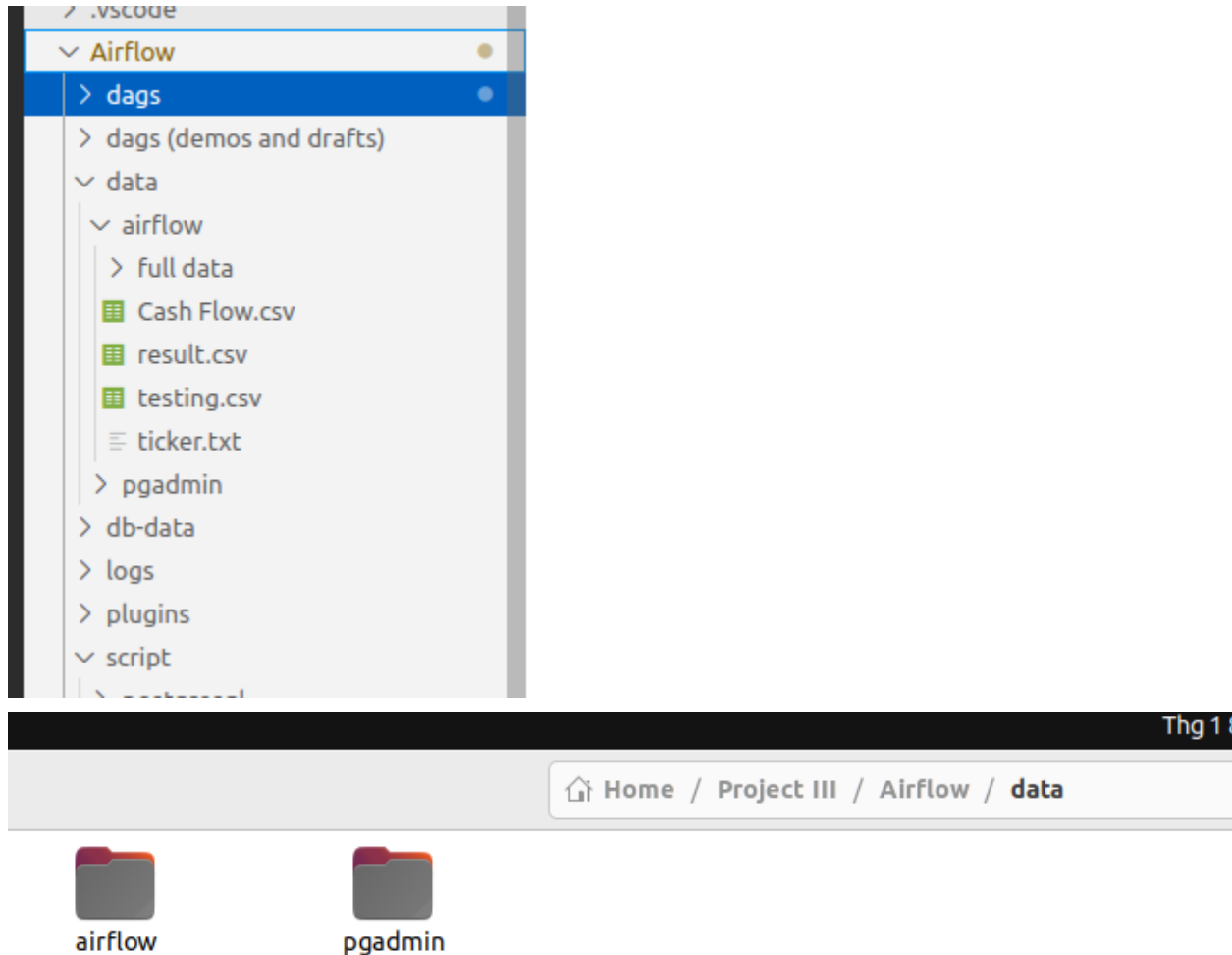


Setup để chạy các DAG lấy dữ liệu và nhập vào bảng

1. Tạo các folder mới
 - Tạo các folder như này để sau làm nơi lưu data trên Airflow (Mọi người có thể xem trên repo của mình)



2. Docker-compose.yaml
 - Thêm các lib python sẽ sử dụng trong code DAG

```
AIRFLOW__CORE__DAGS_ARE_FETCHED_AT_CREATION: 'true'
AIRFLOW__CORE__LOAD_EXAMPLES: 'false'
AIRFLOW__API__AUTH_BACKENDS: 'airflow.api.auth.backend.basic_auth'
_PIP_ADDITIONAL_REQUIREMENTS: 'vnstock wheel psycpg2 sqlalchemy'
volumes:
  - ./dags:/opt/airflow/dags
  - ./logs:/opt/airflow/logs
```

- Mount thêm 1 folder mới dùng để lưu các file csv sẽ tạo

```
_PIP_ADDITIONAL_REQUIREMENTS: 'vnstock wheel psycpg2 sqlalchemy'
volumes:
  - ./dags:/opt/airflow/dags
  - ./logs:/opt/airflow/logs
  - ./plugins:/opt/airflow/plugins
  - ./data/airflow:/opt/airflow/data
```

- SAU KHI SỬA FILE NÀY CẦN CHẠY docker-compose up lại để nó nhận các thay đổi
3. Tạo file csv chứa danh sách ticker

- Chạy DAG “ticker_list” trên Airflow hoặc là copy luôn file ticker.txt trong /Airflow/data/airflow
 - Copy các file csv trong /Airflow/data/airflow/full data, mấy file này để sau khi mọi người viết và chạy thử DAG thì có thể chạy DAG “import_full_database” để nhập lại toàn bộ dữ liệu ban đầu
4. Tạo DAG
- Trong folder đã có file cash_flow.py, file này dùng để nhập dữ liệu bảng cash_flow, mọi người sẽ lấy nó làm template để tạo thêm dag cho các bảng còn lại
 - 1 số trường cần thay đổi
 - Dag id và description, để sau này lên web ui của airflow nhìn không bị lẫn các dag với nhau

```
with DAG(
    dag_id='import_full_database',
    description='Import all data to database',
    start_date=pendulum.yesterday()
```

- schedule_interval, dùng để lên lịch chạy các dag, tùy vào bảng sẽ có thời gian chạy khác nhau (Bảng theo quý thì chạy 10h sáng ngày đầu tháng 2,5,8,11, bảng theo ngày thì chạy hàng ngày), mọi người xem cách để điền cái trường này tại đây <https://crontab.guru/>

```
description='import to cash flow table',
start_date=pendulum.yesterday(),
schedule_interval='0 10 1 1-12/3 *',
catchup=True
```

5. Chia việc

- Nguyễn Minh Mạnh
 - income_statement
 - balance_sheet
 - financial_ratio
 - stock_history
- Nguyễn Văn Hùng
 - general_rating
 - business_model_rating
 - business_operation_rating
 - financial_health_rating
- Phạm Đình Gia Dũng
 - cash_flow
 - stock_intraday_transaction
 - valuation_rating
 - industry_financial_health