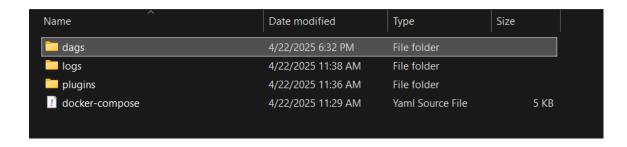# TP8 : Apache Airflow

1. do `docker-compose up`

2. you will get a structure like this:



3. then inside the dags folder , create a python file `main.py` with this code:

```python
import json
import os
import requests
from datetime import datetime
from airflow import DAG
from airflow.operators.bash import BashOperator
from airflow.operators.python import PythonOperator

# Paths inside the Docker container
ROOT_DIR = "/opt/airflow/dags/tmp"
LAUNCHES_FILE = os.path.join(ROOT_DIR, "launches.json")
IMAGES_DIR = os.path.join(ROOT_DIR, "images")

default_args = {
    'owner': 'tp8',
    'start_date': datetime(2024, 5, 10),
    'retries': 1
}

dag = DAG(
    dag_id='tp8',
    default_args=default_args,
    schedule_interval='@daily'
```

```python
)

# Task 1: Download launches.json
download_launches = BashOperator(
    task_id="download_launches",
    bash_command=f'mkdir -p {ROOT_DIR} && curl -Lk "<https://ll.thesp
acedevs.com/2.0.0/launch/upcoming>" > {LAUNCHES_FILE}',
    dag=dag
)

# Task 2: Download images
def get_pictures():
    # Create images directory if it doesn't exist
    os.makedirs(IMAGES_DIR, exist_ok=True)

    # Read launches.json
    with open(LAUNCHES_FILE, "r") as f:
        launches = json.load(f)

    # Download images
    for launch in launches.get("results", []):
        image_url = launch.get("image")
        if image_url:
            image_name = os.path.basename(image_url)
            image_path = os.path.join(IMAGES_DIR, image_name)
            response = requests.get(image_url)
            with open(image_path, "wb") as img_file:
                img_file.write(response.content)

download_images = PythonOperator(
    task_id="download_images",
    python_callable=get_pictures,
    dag=dag
)

# Task 3: Notify with image count
notify = BashOperator(
    task_id="notify",
```

```
    bash_command=f'echo "Number of images: $(ls {IMAGES_DIR} | wc
-l)"',
    dag=dag
)

# Set task dependencies
download_launches >> download_images >> notify
```

4. in the same directory (dags) create a folder called `tmp`

5. give the permission to tmp folder to write: `chmod -R 777 dags/tmp`

6. go to `http://localhost:8080/`

7. Then just follow the pages in the lab.