

```
eugen@DESKTOP-L6093R3 MINGW64 /c/Windows/system32
$ docker exec -it edvai_hadoop bash
root@94908946c352:/# su hadoop
hadoop@94908946c352:/# hdfs dfs -mkdir -p /data/tripdata/warehouse/airport_trips
hadoop@94908946c352:/# hdfs dfs -chmod -R 775 /data/tripdata
hadoop@94908946c352:/# hive -e "CREATE DATABASE IF NOT EXISTS tripdata LOCATION '/data/tripdata/warehouse';"
hive> CREATE EXTERNAL TABLE IF NOT EXISTS tripdata.airport_trips (
    >     tpep_pickup_datetime TIMESTAMP,
    >     airport_fee           DECIMAL(10,2),
    >     payment_type          INT,
    >     tolls_amount          DECIMAL(10,2),
    >     total_amount          DECIMAL(10,2)
    > )
    > STORED AS PARQUET
    > LOCATION 'hdfs://data/tripdata/warehouse/airport_trips';
OK
```

```
hive> DESCRIBE tripdata.airport_trips;
OK
tpep_pickup_datetime      timestamp
airport_fee                decimal(10,2)
payment_type                int
tolls_amount                decimal(10,2)
total_amount                decimal(10,2)
Time taken: 0.102 seconds, Fetched: 5 row(s)
```

```
hadoop@94908946c352:/# /home/hadoop/ingest_yellow_2021_01_02.sh
[INFO] Descargando archivos Parquet...
--2025-11-02 16:17:16-- https://d37ci6vzurychx.cloudfront.net/trip-data/yellow_tripdata_2021-01.parquet
Resolving d37ci6vzurychx.cloudfront.net (d37ci6vzurychx.cloudfront.net)... 65.8.245.171, 65.8.245.178, 65.8.245.51, ...
Connecting to d37ci6vzurychx.cloudfront.net (d37ci6vzurychx.cloudfront.net)|65.8.245.171|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 21686067 (21M) [application/x-www-form-urlencoded]
Saving to: '/home/hadoop/landing/tripdata/yellow_tripdata_2021-01.parquet'

/home/hadoop/landing/tripdata 100%[=====] 20.68M 12.1MB/s in 1.7s
2025-11-02 16:17:18 (12.1 MB/s) - '/home/hadoop/landing/tripdata/yellow_tripdata_2021-01.parquet' saved [21686067/21686067]

--2025-11-02 16:17:18-- https://d37ci6vzurychx.cloudfront.net/trip-data/yellow_tripdata_2021-02.parquet
Resolving d37ci6vzurychx.cloudfront.net (d37ci6vzurychx.cloudfront.net)... 65.8.245.171, 65.8.245.178, 65.8.245.51, ...
Connecting to d37ci6vzurychx.cloudfront.net (d37ci6vzurychx.cloudfront.net)|65.8.245.171|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 21777258 (21M) [application/x-www-form-urlencoded]
Saving to: '/home/hadoop/landing/tripdata/yellow_tripdata_2021-02.parquet'

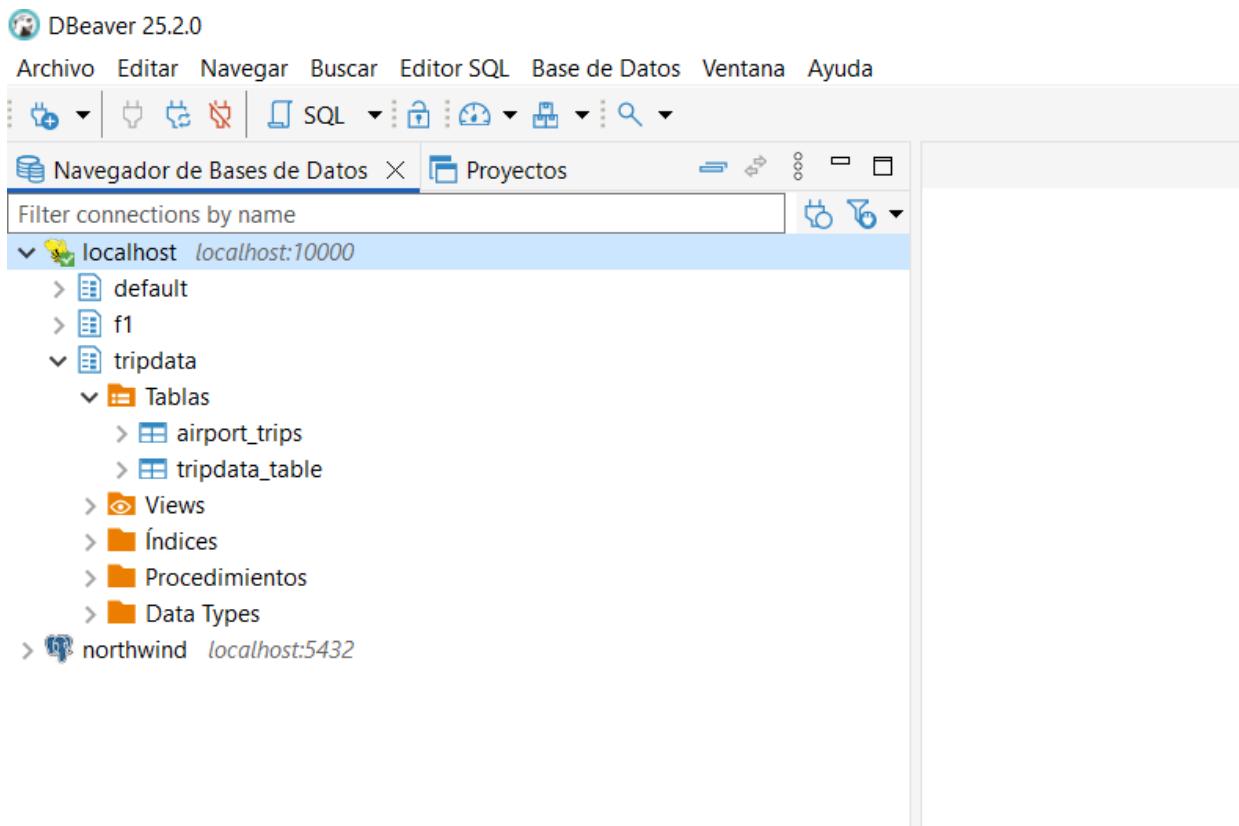
/home/hadoop/landing/tripdata 100%[=====] 20.77M 12.4MB/s in 1.7s
2025-11-02 16:17:21 (12.4 MB/s) - '/home/hadoop/landing/tripdata/yellow_tripdata_2021-02.parquet' saved [21777258/21777258]

[INFO] Creando directorios en HDFS...
[INFO] Copiando archivos a HDFS...
✓ Archivos copiados exitosamente a HDFS:
drwxr-xr-x - hadoop supergroup 0 2025-11-02 16:17 /data/tripdata/raw/yellow/2021/01
-rw-r--r-- 1 hadoop supergroup 21686067 2025-11-02 16:17 /data/tripdata/raw/yellow/2021/01/yellow_tripdata_2021-01.parquet
drwxr-xr-x - hadoop supergroup 0 2025-11-02 16:17 /data/tripdata/raw/yellow/2021/02
-rw-r--r-- 1 hadoop supergroup 21777258 2025-11-02 16:17 /data/tripdata/raw/yellow/2021/02/yellow_tripdata_2021-02.parquet
hadoop@94908946c352:/# |
```

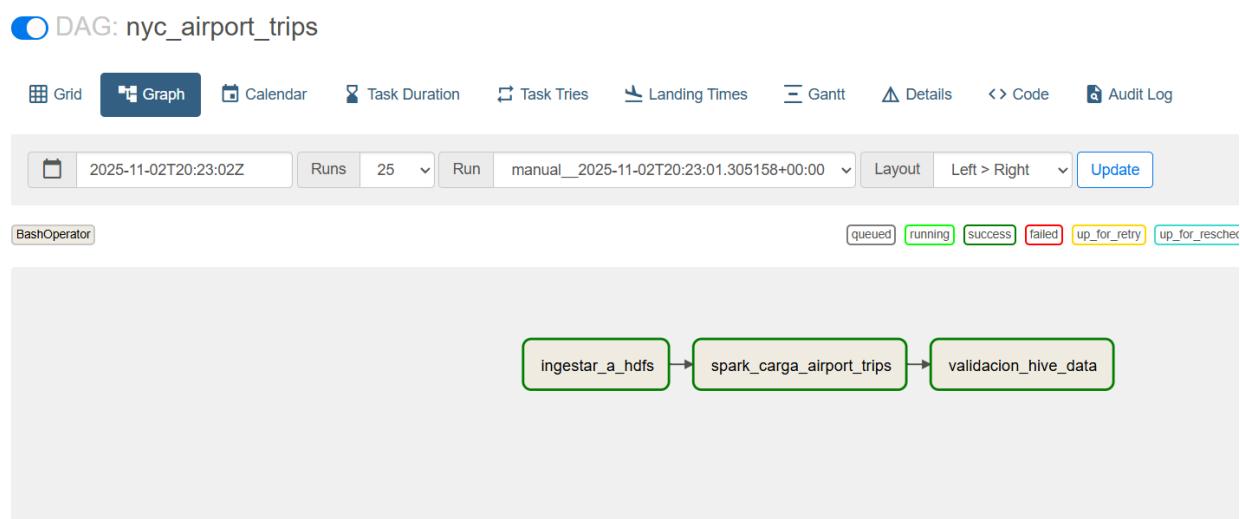
```

2025-11-02 16:23:59,101 INFO cluster.FairScheduler.
2025-11-02 16:23:59,101 INFO scheduler.DAGScheduler:
    ✓ Inserción completada. Filas cargadas: 5757
2025-11-02 16:23:59,112 INFO server.AbstractConnector
2025-11-02 16:23:59,112 INFO ui.SparkUI: Stopped at 2025-11-02 16:23:59

```



## Clase 8



DBeaver 25.2.0

Archivo Editar Navegar Buscar Editor SQL Base de Datos Ventana Ayuda

Navegador de Bases de Datos Proyectos

Filter connections by name

localhost localhost:10000

- > default
- > f1
  - > Tablas
    - > constructor\_results
      - AZ constructorref (STRING)
      - AZ cons\_name (STRING)
      - AZ cons\_nationality (STRING)
      - AZ url (STRING)
      - 123 points (DOUBLE)
    - > Claves
    - > Columnas de clave externa
    - > Índices
    - > Referencias
  - > driver\_results
    - > Columnas
      - AZ driver\_forename (STRING)
      - AZ driver\_surname (STRING)
      - AZ driver\_nationality (STRING)
      - 123 points (DOUBLE)
    - > Claves