

## Implementación de flujo automatizado del proceso de carga de datos a la base de datos SQL Server.

**Docker-Windows:** No me permitió ejecutar la configuración correctamente para levantar el servidor de Apache-Airflow

### Implementación de Apache-Airflow en Linux Ubuntu (WSL-2)

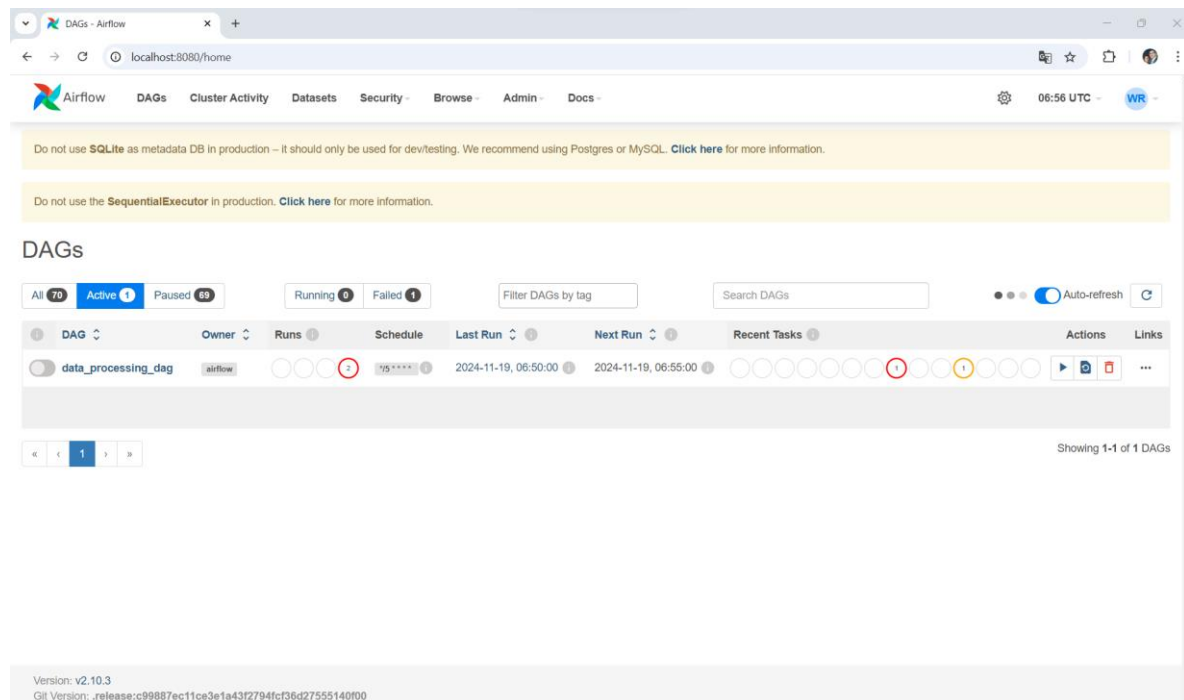
Ruta DAG: [\\wsl.localhost\Ubuntu\home\williamxlr\airflow\entorno\\_airflow\dags](\\wsl.localhost\Ubuntu\home\williamxlr\airflow\entorno_airflow\dags)

Ruta Scripts: [\\wsl.localhost\Ubuntu\home\williamxlr\airflow\entorno\\_airflow\scripts](\\wsl.localhost\Ubuntu\home\williamxlr\airflow\entorno_airflow\scripts)

Se ejecuta con errores:

Posible causa: Sistema Operativo Linux

Todo el proceso se debe mover y probar teniendo en cuenta las rutas de los scripts y el ambiente para la base de datos y la ejecución de los subprocessos.



List Dag Run - Airflow

localhost:8080/dagrun/list/?flt\_3\_dag\_id=data\_processing\_dag&flt\_3\_state=failed

Airflow

DAGsCluster ActivityDatasetsSecurityBrowseAdminDocs

06:58 UTCWR

List Dag Run

Search

Actions

Record Count: 2

	State	Dag Id	Logical Date	Run Id	Run Type	Queued At	Start Date	End Date	Note	External Trigger	Conf	Duration
<input type="checkbox"/>	failed	data_processing_dag	2024-11-19, 06:50:00	scheduled__2024-11-19T06:50:00+00:00	scheduled	2024-11-19, 06:55:01	2024-11-19, 06:55:01	2024-11-19, 06:55:03	False			1s
<input type="checkbox"/>	failed	data_processing_dag	2024-11-19, 06:45:00	scheduled__2024-11-19T06:45:00+00:00	scheduled	2024-11-19, 06:54:14	2024-11-19, 06:54:14	2024-11-19, 06:54:16	False			2s

Version: v2.10.3  
Git Version: .release:c99887ec11ce3e1a43f2794fcf36d27555140f00

data\_processing\_dag - Grid - Airflow

localhost:8080/dags/data\_processing\_dag/grid?execution\_date=2024-11-19+06%3A50%3A00%2B00%3A00&tab=graph

Airflow

DAGsCluster ActivityDatasetsSecurityBrowseAdminDocs

06:58 UTCWR

DAG: data\_processing\_dag DAG para limpiar y cargar datos

Schedule: \*/5 \* \* \* \*Next Run ID: 2024-11-19, 06:55:00 UTC

19/11/202406:50 a. m.

All Run TypesAll Run StatesClear Filters

Auto-refresh25

deferredfailedqueuedremovedrestartingrunningscheduledshutdownskippedsuccessup\_for\_rescheduleup\_for\_retryupstream\_failedno\_status

data\_processing\_dag

Duration

00:00:00

00:00:01

00:00:02

clear\_data\_task

load\_data\_task

DetailsGraphGanttCodeEvent LogRun DurationTask DurationCalendar

clear\_data\_taskPythonOperator

load\_data\_taskPythonOperator

LayoutLeft -> Right

React Flow

data\_processing\_dag - Grid - Airflow

localhost:8080/dags/data\_processing\_dag/grid?execution\_date=2024-11-19+06%3A50%3A00%2B00%3A00&tab=logs&dag\_run\_id=scheduled\_2024-11-19T06%3A50%3A00%2B00%3A00&ta...

Airflow DAGs Cluster Activity Datasets Security Browse Admin Docs

06:59 UTC

DAG: data\_processing\_dag DAG para limpiar y cargar datos

Schedule: \*/5 \* \* \* \* Next Run ID: 2024-11-19, 06:55:00 UTC

data\_processing\_dag Run 2024-11-19, 06:50:00 UTC Task clear\_data\_task

Clear task Mark state as... Filter DAG by task

Details Graph Gantt Code Event Log Logs XCom Task Duration

All Levels All File Sources Wrap Download See More

```
LAPTOP-0070/287.
*** Found local files:
*** + /home/williamlr/airflow/logs/dag_id=data_processing_dag/run_id=scheduled_2024-11-19T06:50:00/task_id=clear_data_task/attempt=1.log
[2024-11-19, 06:55:02 UTC] [local_task_job_runner.py:123] ▶ Pre task execution logs
[2024-11-19, 06:55:02 UTC] [taskinstance.py:3311] ERROR - Task failed with exception
Traceback (most recent call last):
  File "/home/williamlr/.local/lib/python3.10/site-packages/airflow/models/taskinstance.py", line 767, in _execute_task
    result = _execute_callable(context=context, **execute_callable_kwargs)
  File "/home/williamlr/.local/lib/python3.10/site-packages/airflow/models/taskinstance.py", line 733, in _execute_callable
    return ExecutionCallableRunner(
  File "/home/williamlr/.local/lib/python3.10/site-packages/airflow/utils/operator_helpers.py", line 252, in run
    return self.func(*args, **kwargs)
  File "/home/williamlr/.local/lib/python3.10/site-packages/airflow/models/basemanager.py", line 417, in wrapper
    return func(self, *args, **kwargs)
  File "/home/williamlr/.local/lib/python3.10/site-packages/airflow/operators/python.py", line 238, in execute
    return_value = self.execute_callable()
  File "/home/williamlr/.local/lib/python3.10/site-packages/airflow/operators/python.py", line 236, in execute_callable
    return runner.run(self.op_args, **self.op_kwargs)
  File "/home/williamlr/.local/lib/python3.10/site-packages/airflow/utils/operator_helpers.py", line 252, in run
    return self.func(*args, **kwargs)
  File "/home/williamlr/airflow/entorno/airflow/my_dag.py", line 10, in run_cleardata
    subprocess.run(['python', '-c', 'home/williamlr/airflow/entorno/airflow/scripts/cleardata.py'], check=True)
  File "/usr/lib/python3.10/subprocess.py", line 526, in run
    raise CalledProcessError(retcode, process.args,
subprocess.CalledProcessError: Command '['python', '-c', 'home/williamlr/airflow/entorno/airflow/scripts/cleardata.py']' returned non-zero exit status 2.
[2024-11-19, 06:55:02 UTC] [logging_mixin.py:190] INFO - Task instance in failure state
[2024-11-19, 06:55:02 UTC] [logging_mixin.py:190] INFO - Task start:2024-11-19 06:55:02.472135+00:00 end:2024-11-19 06:55:02.583643+00:00 duration:0.111508
```