

etl\_movies\_series  
\*\*\* Reading local file: /opt/airflow/logs/dag\_id=etl\_streaming\_smtp/run\_id=scheduled\_\_2023-07-12T00:00:00+00:00/task\_id=etl\_movies\_series/attempt=9.log

[2023-07-13, 15:47:01 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.etl\_movies\_series scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:01 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.etl\_movies\_series scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:01 UTC] {taskinstance.py:1376} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:01 UTC] {taskinstance.py:1377} INFO - Starting attempt 9 of 14

[2023-07-13, 15:47:01 UTC] {taskinstance.py:1378} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:01 UTC] {taskinstance.py:1397} INFO - Executing <Task(PythonOperator): etl\_movies\_series> on 2023-07-12 00:00:00+00:00

[2023-07-13, 15:47:01 UTC] {standard\_task\_runner.py:52} INFO - Started process 11708 to run task

[2023-07-13, 15:47:01 UTC] {standard\_task\_runner.py:79} INFO - Running: ['\*\*\*', 'tasks', 'run', 'etl\_streaming\_smtp', 'etl\_movies\_series', 'scheduled\_\_2023-07-12T00:00:00+00:00', '--job-id', '9962', '--raw', '--subdir', 'DAGS\_FOLDER/etl\_streaming\_SMTP.py', '--cfg-path', '/tmp/tmppbc\_6nc6', '--error-file', '/tmp/tmp0l4euf\_v']

[2023-07-13, 15:47:01 UTC] {standard\_task\_runner.py:80} INFO - Job 9962: Subtask etl\_movies\_series

[2023-07-13, 15:47:01 UTC] {task\_command.py:371} INFO - Running <TaskInstance: etl\_streaming\_smtp.etl\_movies\_series scheduled\_\_2023-07-12T00:00:00+00:00 [running]> on host 37e3a29a6183

[2023-07-13, 15:47:01 UTC] {taskinstance.py:1591} INFO - Exporting the following env vars:

AIRFLOW\_CTX\_DAG\_OWNER=AdriPrietoPulido

AIRFLOW\_CTX\_DAG\_ID=etl\_streaming\_smtp

AIRFLOW\_CTX\_TASK\_ID=etl\_movies\_series

AIRFLOW\_CTX\_EXECUTION\_DATE=2023-07-12T00:00:00+00:00

AIRFLOW\_CTX\_TRY\_NUMBER=9

AIRFLOW\_CTX\_DAG\_RUN\_ID=scheduled\_\_2023-07-12T00:00:00+00:00

[2023-07-13, 15:47:06 UTC] {logging\_mixin.py:115} INFO - Se ha exportado el Dataframe

[2023-07-13, 15:47:06 UTC] {python.py:173} INFO - Done. Returned value was: None

[2023-07-13, 15:47:06 UTC] {taskinstance.py:1420} INFO - Marking task as SUCCESS. dag\_id=etl\_streaming\_smtp, task\_id=etl\_movies\_series, execution\_date=20230712T000000, start\_date=20230713T154701, end\_date=20230713T154706

[2023-07-13, 15:47:06 UTC] {local\_task\_job.py:156} INFO - Task exited with return code 0

[2023-07-13, 15:47:06 UTC] {local\_task\_job.py:273} INFO - 1 downstream tasks scheduled from follow-on schedule check

envio\_info  
\*\*\* Reading local file: /opt/airflow/logs/dag\_id=etl\_streaming\_smtp/run\_id=scheduled\_\_2023-07-12T00:00:00+00:00/task\_id=envio\_info/attempt=18.log

[2023-07-13, 15:47:08 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.envio\_info scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:08 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.envio\_info scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:08 UTC] {taskinstance.py:1376} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:08 UTC] {taskinstance.py:1377} INFO - Starting attempt 18 of 23

[2023-07-13, 15:47:08 UTC] {taskinstance.py:1378} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:08 UTC] {taskinstance.py:1397} INFO - Executing <Task(PythonOperator): envio\_info> on 2023-07-12 00:00:00+00:00

[2023-07-13, 15:47:08 UTC] {standard\_task\_runner.py:52} INFO - Started process 11721 to run task

[2023-07-13, 15:47:08 UTC] {standard\_task\_runner.py:79} INFO - Running: ['\*\*\*', 'tasks', 'run', 'etl\_streaming\_smtp', 'envio\_info', 'scheduled\_\_2023-07-12T00:00:00+00:00', '--job-id', '9963', '--raw', '--subdir', 'DAGS\_FOLDER/etl\_streaming\_SMTP.py', '--cfg-path', '/tmp/tmpo4vwdvr6', '--error-file', '/tmp/tmpk4vm1zsr']

[2023-07-13, 15:47:08 UTC] {standard\_task\_runner.py:80} INFO - Job 9963: Subtask envio\_info

[2023-07-13, 15:47:08 UTC] {task\_command.py:371} INFO - Running <TaskInstance: etl\_streaming\_smtp.envio\_info scheduled\_\_2023-07-12T00:00:00+00:00 [running]> on host 37e3a29a6183

[2023-07-13, 15:47:08 UTC] {taskinstance.py:1591} INFO - Exporting the following env vars:

AIRFLOW\_CTX\_DAG\_OWNER=AdriPrietoPulido

AIRFLOW\_CTX\_DAG\_ID=etl\_streaming\_smtp

AIRFLOW\_CTX\_TASK\_ID=envio\_info

AIRFLOW\_CTX\_EXECUTION\_DATE=2023-07-12T00:00:00+00:00

AIRFLOW\_CTX\_TRY\_NUMBER=18

AIRFLOW\_CTX\_DAG\_RUN\_ID=scheduled\_\_2023-07-12T00:00:00+00:00

[2023-07-13, 15:47:08 UTC] {logging\_mixin.py:115} INFO - Index(['genero'], dtype='object')

[2023-07-13, 15:47:08 UTC] {logging\_mixin.py:115} INFO - Hola

La informacion de streaming se cargo con valores nulos

recomendamos revisar la columna: genero

[2023-07-13, 15:47:11 UTC] {logging\_mixin.py:115} INFO - Exito

[2023-07-13, 15:47:11 UTC] {python.py:173} INFO - Done. Returned value was: None

[2023-07-13, 15:47:11 UTC] {taskinstance.py:1420} INFO - Marking task as SUCCESS. dag\_id=etl\_streaming\_smtp, task\_id=envio\_info, execution\_date=20230712T000000, start\_date=20230713T154708, end\_date=20230713T154711

[2023-07-13, 15:47:11 UTC] {local\_task\_job.py:156} INFO - Task exited with return code 0

[2023-07-13, 15:47:11 UTC] {local\_task\_job.py:273} INFO - 1 downstream tasks scheduled from follow-on schedule check

conect\_RedShift

* [1](http://localhost:8080/log?dag_id=etl_streaming_smtp&task_id=conect_RedShift&execution_date=2023-07-12T00%3A00%3A00%2B00%3A00#1)

Jump To End Toggle Wrap [Download](http://localhost:8080/get_logs_with_metadata?dag_id=etl_streaming_smtp&task_id=conect_RedShift&execution_date=2023-07-12T00%3A00%3A00%2B00%3A00&try_number=1&metadata=null&format=file)

\*\*\* Reading local file: /opt/airflow/logs/dag\_id=etl\_streaming\_smtp/run\_id=scheduled\_\_2023-07-12T00:00:00+00:00/task\_id=conect\_RedShift/attempt=1.log

[2023-07-13, 15:47:12 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.conect\_RedShift scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:12 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.conect\_RedShift scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:12 UTC] {taskinstance.py:1376} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:12 UTC] {taskinstance.py:1377} INFO - Starting attempt 1 of 6

[2023-07-13, 15:47:12 UTC] {taskinstance.py:1378} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:12 UTC] {taskinstance.py:1397} INFO - Executing <Task(PythonOperator): conect\_RedShift> on 2023-07-12 00:00:00+00:00

[2023-07-13, 15:47:12 UTC] {standard\_task\_runner.py:52} INFO - Started process 11735 to run task

[2023-07-13, 15:47:12 UTC] {standard\_task\_runner.py:79} INFO - Running: ['\*\*\*', 'tasks', 'run', 'etl\_streaming\_smtp', 'conect\_RedShift', 'scheduled\_\_2023-07-12T00:00:00+00:00', '--job-id', '9964', '--raw', '--subdir', 'DAGS\_FOLDER/etl\_streaming\_SMTP.py', '--cfg-path', '/tmp/tmp7dr0lefd', '--error-file', '/tmp/tmpxm3mokua']

[2023-07-13, 15:47:12 UTC] {standard\_task\_runner.py:80} INFO - Job 9964: Subtask conect\_RedShift

[2023-07-13, 15:47:12 UTC] {task\_command.py:371} INFO - Running <TaskInstance: etl\_streaming\_smtp.conect\_RedShift scheduled\_\_2023-07-12T00:00:00+00:00 [running]> on host 37e3a29a6183

[2023-07-13, 15:47:12 UTC] {taskinstance.py:1591} INFO - Exporting the following env vars:

AIRFLOW\_CTX\_DAG\_OWNER=AdriPrietoPulido

AIRFLOW\_CTX\_DAG\_ID=etl\_streaming\_smtp

AIRFLOW\_CTX\_TASK\_ID=conect\_RedShift

AIRFLOW\_CTX\_EXECUTION\_DATE=2023-07-12T00:00:00+00:00

AIRFLOW\_CTX\_TRY\_NUMBER=1

AIRFLOW\_CTX\_DAG\_RUN\_ID=scheduled\_\_2023-07-12T00:00:00+00:00

[2023-07-13, 15:47:14 UTC] {logging\_mixin.py:115} INFO - Connected to Redshift successfully!

[2023-07-13, 15:47:14 UTC] {python.py:173} INFO - Done. Returned value was: None

[2023-07-13, 15:47:14 UTC] {taskinstance.py:1420} INFO - Marking task as SUCCESS. dag\_id=etl\_streaming\_smtp, task\_id=conect\_RedShift, execution\_date=20230712T000000, start\_date=20230713T154712, end\_date=20230713T154714

[2023-07-13, 15:47:14 UTC] {local\_task\_job.py:156} INFO - Task exited with return code 0

[2023-07-13, 15:47:14 UTC] {local\_task\_job.py:273} INFO - 1 downstream tasks scheduled from follow-on schedule check

cargar\_tabla

\*\*\* Reading local file: /opt/airflow/logs/dag\_id=etl\_streaming\_smtp/run\_id=scheduled\_\_2023-07-12T00:00:00+00:00/task\_id=cargar\_tabla/attempt=1.log

[2023-07-13, 15:47:15 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.cargar\_tabla scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:15 UTC] {taskinstance.py:1179} INFO - Dependencies all met for <TaskInstance: etl\_streaming\_smtp.cargar\_tabla scheduled\_\_2023-07-12T00:00:00+00:00 [queued]>

[2023-07-13, 15:47:15 UTC] {taskinstance.py:1376} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:15 UTC] {taskinstance.py:1377} INFO - Starting attempt 1 of 6

[2023-07-13, 15:47:15 UTC] {taskinstance.py:1378} INFO -

--------------------------------------------------------------------------------

[2023-07-13, 15:47:15 UTC] {taskinstance.py:1397} INFO - Executing <Task(PythonOperator): cargar\_tabla> on 2023-07-12 00:00:00+00:00

[2023-07-13, 15:47:15 UTC] {standard\_task\_runner.py:52} INFO - Started process 11756 to run task

[2023-07-13, 15:47:15 UTC] {standard\_task\_runner.py:79} INFO - Running: ['\*\*\*', 'tasks', 'run', 'etl\_streaming\_smtp', 'cargar\_tabla', 'scheduled\_\_2023-07-12T00:00:00+00:00', '--job-id', '9965', '--raw', '--subdir', 'DAGS\_FOLDER/etl\_streaming\_SMTP.py', '--cfg-path', '/tmp/tmpbz0x5xjw', '--error-file', '/tmp/tmpeu1\_xb18']

[2023-07-13, 15:47:15 UTC] {standard\_task\_runner.py:80} INFO - Job 9965: Subtask cargar\_tabla

[2023-07-13, 15:47:15 UTC] {task\_command.py:371} INFO - Running <TaskInstance: etl\_streaming\_smtp.cargar\_tabla scheduled\_\_2023-07-12T00:00:00+00:00 [running]> on host 37e3a29a6183

[2023-07-13, 15:47:15 UTC] {taskinstance.py:1591} INFO - Exporting the following env vars:

AIRFLOW\_CTX\_DAG\_OWNER=AdriPrietoPulido

AIRFLOW\_CTX\_DAG\_ID=etl\_streaming\_smtp

AIRFLOW\_CTX\_TASK\_ID=cargar\_tabla

AIRFLOW\_CTX\_EXECUTION\_DATE=2023-07-12T00:00:00+00:00

AIRFLOW\_CTX\_TRY\_NUMBER=1

AIRFLOW\_CTX\_DAG\_RUN\_ID=scheduled\_\_2023-07-12T00:00:00+00:00

[2023-07-13, 15:47:18 UTC] {logging\_mixin.py:115} INFO - Se ha cargado la Tabla

[2023-07-13, 15:47:18 UTC] {python.py:173} INFO - Done. Returned value was: None

[2023-07-13, 15:47:18 UTC] {taskinstance.py:1420} INFO - Marking task as SUCCESS. dag\_id=etl\_streaming\_smtp, task\_id=cargar\_tabla, execution\_date=20230712T000000, start\_date=20230713T154715, end\_date=20230713T154718

[2023-07-13, 15:47:18 UTC] {local\_task\_job.py:156} INFO - Task exited with return code 0

[2023-07-13, 15:47:18 UTC] {local\_task\_job.py:273} INFO - 0 downstream tasks scheduled from follow-on schedule check