Dataset: orders\_data\_header.csv

order\_id,customer\_id,purchase\_date,amountORDER1,CUST1,1/1/2021,100 ORDER2,CUST2,2/2/2022,200 ORDER3,CUST3,3/3/2023,300 ➡️ Env vars

Environment variables:

AIRFLOW\_\_CORE\_\_ENABLE\_XCOM\_PICKLING=True

AIRFLOW\_\_ASTRO\_SDK\_\_SQL\_SCHEMA=ASTRO\_SDK\_SCHEMA

SQL requests:

CREATE DATABASE ASTRO\_SDK\_DB;

CREATE WAREHOUSE ASTRO\_SDK\_DW;

CREATE SCHEMA ASTRO\_SDK\_SCHEMA;

CREATE OR REPLACE TABLE customers\_table (customer\_id CHAR(10), customer\_name VARCHAR(100), type VARCHAR(10) );

INSERT INTO customers\_table (CUSTOMER\_ID, CUSTOMER\_NAME,TYPE) VALUES ('CUST1','NAME1','TYPE1'),('CUST2','NAME2','TYPE1'),('CUST3','NAME3','TYPE2');

CREATE OR REPLACE TABLE reporting\_table (CUSTOMER\_ID CHAR(30), CUSTOMER\_NAME VARCHAR(100), ORDER\_ID CHAR(10), PURCHASE\_DATE DATE, AMOUNT FLOAT, TYPE CHAR(10));

INSERT INTO reporting\_table (CUSTOMER\_ID, CUSTOMER\_NAME, ORDER\_ID, PURCHASE\_DATE, AMOUNT, TYPE) VALUES ('INCORRECT\_CUSTOMER\_ID','INCORRECT\_CUSTOMER\_NAME','ORDER2','2/2/2022',200,'TYPE1'), ('CUST3','NAME3','ORDER3','3/3/2023',300,'TYPE2'), ('CUST4','NAME4','ORDER4','4/4/2022',400,'TYPE2');