USE DATABASE DEV;

USE SCHEMA DATASCIENCE;

CREATE OR REPLACE FILE FORMAT my\_parquet\_format

TYPE = 'PARQUET';

CREATE OR REPLACE STAGE bronze\_xml

URL = 's3://datascience-output-bucket/bronze/xml/'

STORAGE\_INTEGRATION = aws\_snowflake\_connect

FILE\_FORMAT = my\_parquet\_format;

LIST @DEV.DATASCIENCE.BRONZE\_XML

SELECT \* FROM @DEV.DATASCIENCE.BRONZE\_XML

CREATE OR REPLACE TABLE customers (

customer\_id INT,

customer\_name STRING,

start\_date TIMESTAMP\_NTZ

);

CREATE OR REPLACE TABLE products (

product\_id STRING,

product\_name STRING,

start\_date TIMESTAMP\_NTZ

);

CREATE OR REPLACE TABLE orders (

order\_id STRING,

customer\_id INT,

product\_id STRING,

start\_date TIMESTAMP\_NTZ

);

COPY INTO customers

FROM @DEV.DATASCIENCE.BRONZE\_XML/2025/Jul/05/

FILE\_FORMAT = (FORMAT\_NAME = my\_parquet\_format)

MATCH\_BY\_COLUMN\_NAME = CASE\_INSENSITIVE

PATTERN = '.\*Customers\_.\*\\.parquet'

COPY INTO products

FROM @DEV.DATASCIENCE.BRONZE\_XML/2025/Jul/05/

FILE\_FORMAT = (FORMAT\_NAME = my\_parquet\_format)

MATCH\_BY\_COLUMN\_NAME = CASE\_INSENSITIVE

PATTERN = '.\*Products\_.\*\\.parquet';

COPY INTO orders

FROM @DEV.DATASCIENCE.BRONZE\_XML/2025/Jul/05/

FILE\_FORMAT = (FORMAT\_NAME = my\_parquet\_format)

MATCH\_BY\_COLUMN\_NAME = CASE\_INSENSITIVE

PATTERN = '.\*Orders\_.\*\\.parquet';

SELECT \* FROM customers LIMIT 10;

SELECT \* FROM products LIMIT 10;

SELECT \* FROM orders LIMIT 10;