**SNOWFLAKE ASSIGNMENT 1**

**Scenario: Migrate an on-premises SQL Server database to Snowflake via Azure Databricks.**

**Task: Extract data from SQL Server, transform it in Databricks using Snowpark, and load it into Snowflake.**

**Step 1: Using Azure Portal Query Editor**

1. **Open Azure Portal**
   * Go to https://portal.azure.com
   * Navigate to your **SQL Database** (e.g., sqlsf) → Click **Query editor (preview)**
2. **Login**
   * Use **SQL Authentication**
   * Username: sa-user
   * Password: Poojashree@307
3. **Create Table**

CREATE TABLE CUSTOMER (

id INT PRIMARY KEY,

customer VARCHAR(50),

region VARCHAR(50),

amount DECIMAL(10,2),

);

1. **Insert Data Directly**

INSERT INTO CUSTOMER (id, customer, region, amount) VALUES

(1, 'Amit Kumar', 'North', 12000),

(2, 'Priya Sharma', 'South', 18000)

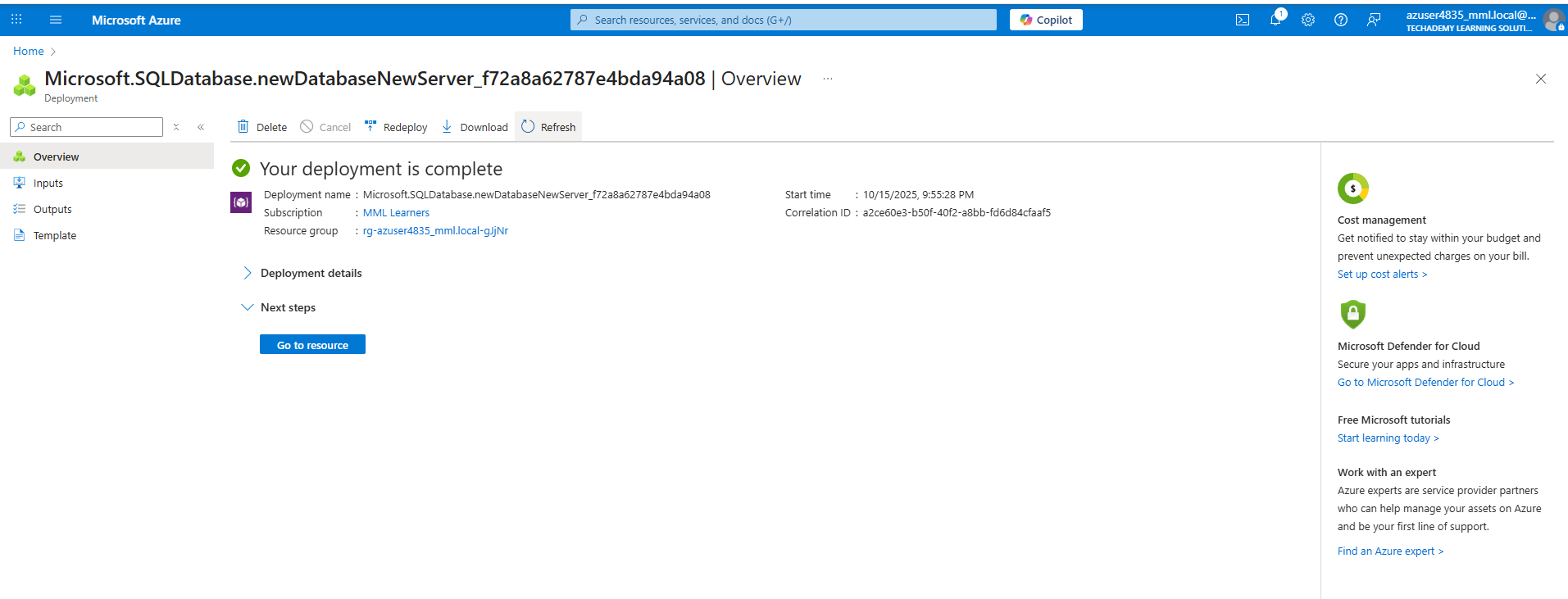
(3, 'Rahul Mehta', 'East', 15000),

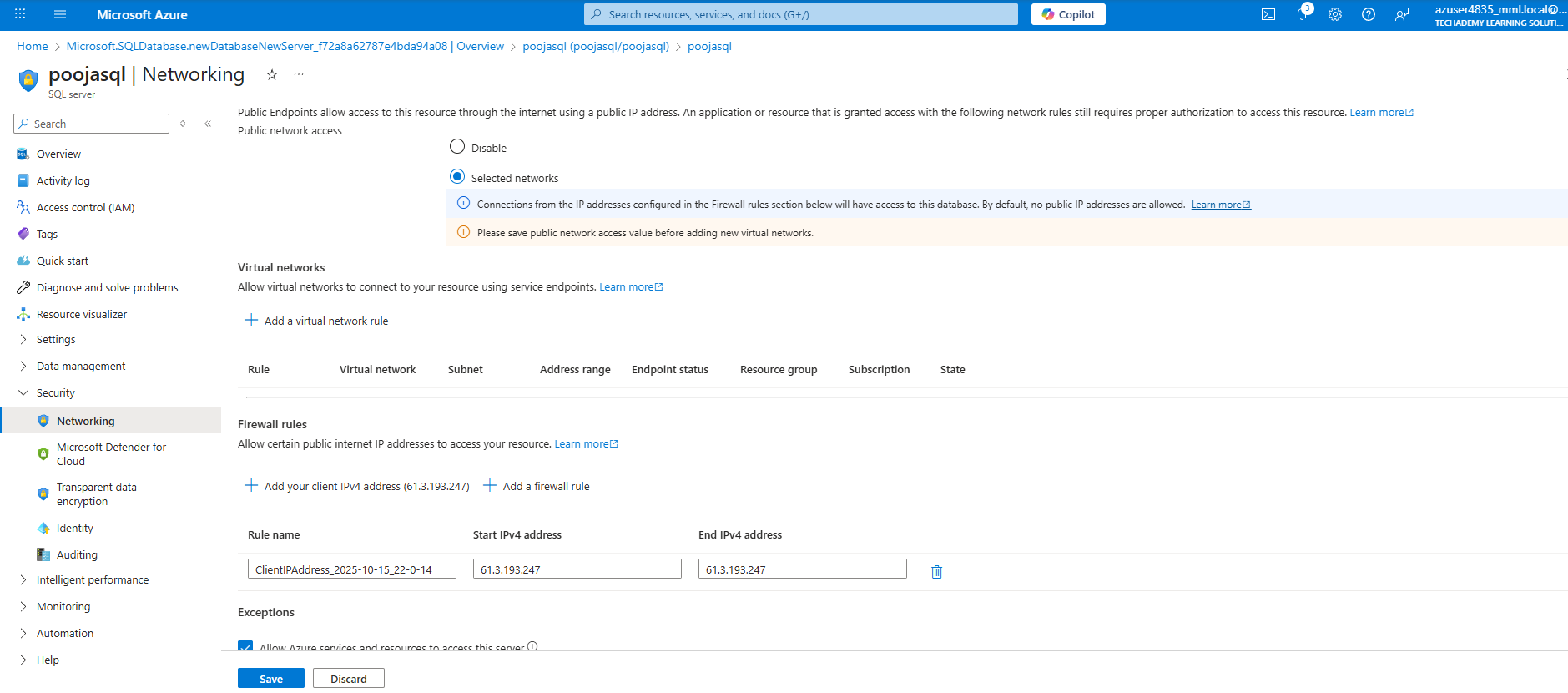
(4, 'Sneha Rao', 'West', 22000),

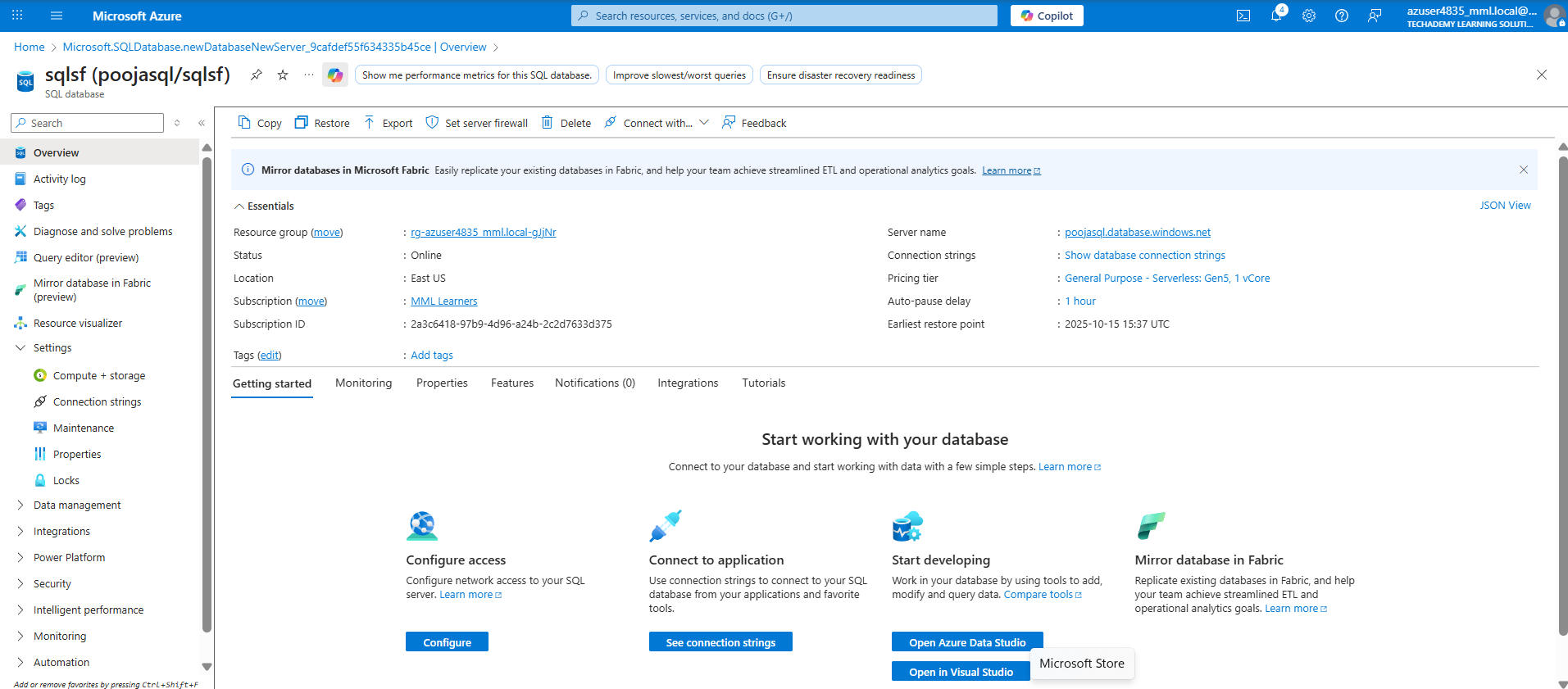
(5, 'Ravi Patel', 'Central', 17000);

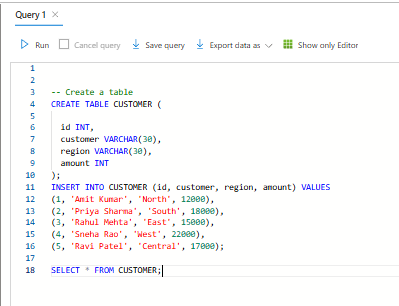
1. **Verify Data**

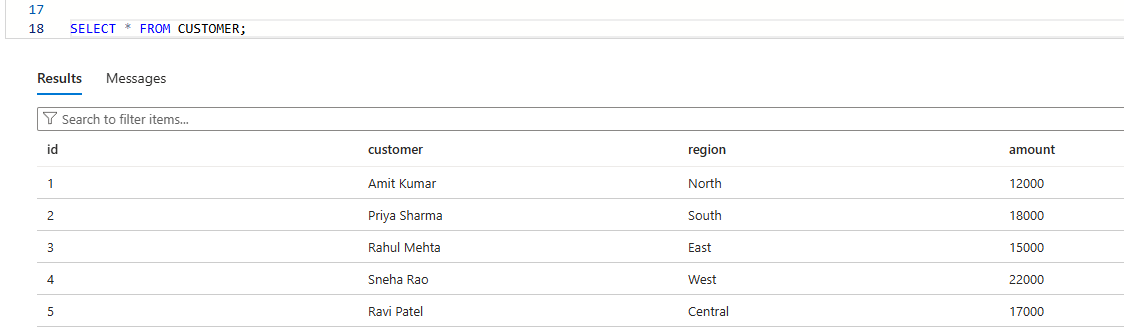
SELECT \* FROM CUSTOMER;



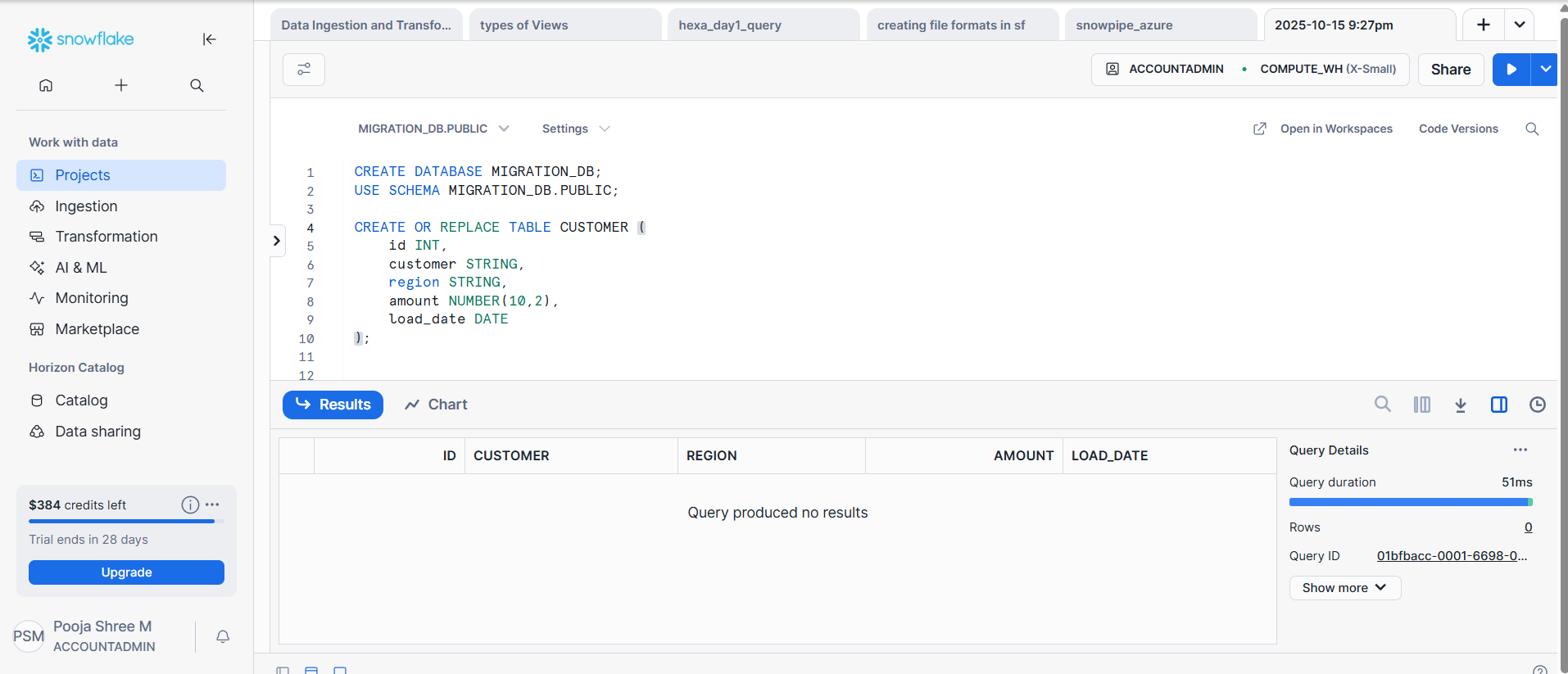








Before transforming and connecting sql op in snowflake



**Step 2: Prepare Azure Databricks workspace**

**Open:**  
 Go to https://portal.azure.com

**Actions:**

1. In the search bar, type **"Databricks"**
2. Click **"Azure Databricks" → Create**
3. Choose:
   * Subscription: your Azure subscription
   * Resource group: create a new one (e.g., data-migration-rg)
   * Workspace name: databricks-sql-migration
   * Region: same as your storage or Snowflake region (e.g., East US)
4. Click **Review + Create → Create**
5. Once deployed, click **“Go to resource” → “Launch Workspace”**
6. It opens **Databricks Web UI** (in a new tab).

**Step 3: Create a Databricks cluster**

**In Databricks Workspace:**

1. Click **Compute** → **Create Cluster**
2. Name: migration-cluster
3. Runtime: choose **“11.x LTS (includes Apache Spark 3.5 + Scala 2.12)”**
4. Click **Create Cluster** (wait a few minutes until cluster = Running )

**Step 4: Connect Databricks to SQL Server**

We’ll use **JDBC** connection.

**In Databricks:**

1. Click **Workspace → Create → Notebook**
   * Name: SQL\_to\_Snowflake\_Migration
   * Language: Python
2. Attach to your cluster.

**Step 5: Transform data in Databricks using Snowpark / PySpark**

**Step 6: Connect to Snowflake**

Now, we’ll write data into Snowflake.

**Pre-setup (in Snowflake Web UI):**

1. Login to Snowflake Snowsight
2. Create target database and schema:

CREATE DATABASE MIGRATION\_DB;

USE SCHEMA MIGRATION\_DB.PUBLIC;

Create target table:

CREATE OR REPLACE TABLE EMPLOYEES\_SNOW (

EMP\_ID INT,

EMP\_NAME STRING,

DEPT STRING,

LOAD\_DATE DATE

);

**Step 7: Load data into Snowflake from Databricks**

In your Databricks notebook, add:

sfOptions = {

  "sfURL": "sqishot-fa68768.snowflakecomputing.com",

  "sfDatabase": "MIGRATION\_DB",

  "sfSchema": "PUBLIC",

  "sfWarehouse": "COMPUTE\_WH",

  "sfRole": "ACCOUNTADMIN",

  "sfUser": "poojashree",

  "sfPassword": "Poojashree@307"

}

# Write transformed data into Snowflake

transformed\_df.write \

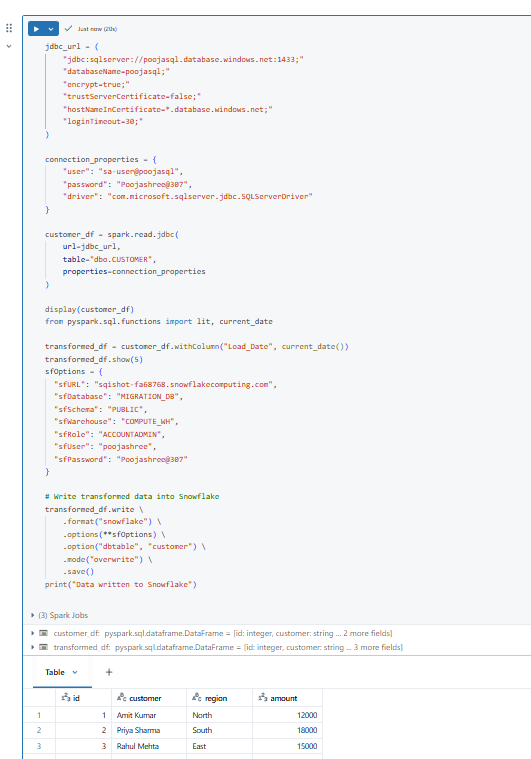
.format("snowflake") \

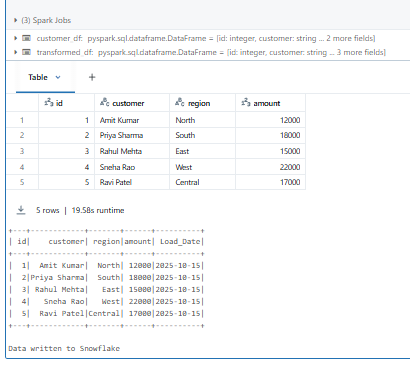
.options(\*\*sfOptions) \

.option("dbtable", "EMPLOYEES\_SNOW") \

.mode("overwrite") \

.save()





**Step 8: Verify in Snowflake**

**Go back to Snowsight**, and run:

USE DATABASE MIGRATION\_DB;

SELECT \* FROM CUSTOMER;

You’ll see your migrated data now available in Snowflake.

