**Case Study: Azure → Snowflake with Snowpark, then Power BI**

**Scenario**

You’re the data engineer at **ItTechGenie Retail**. Sales teams drop monthly CSVs into an **Azure Storage** container. You must:

1. upload the CSV to Azure,
2. ingest it into **Snowflake** using **Snowpark**,
3. model it into proper **database/schema/table**, and
4. build a quick **Power BI** report for business users.

**Introduction**

The objective of this case study is to ingest sales data stored in Azure Blob Storage, load it into Snowflake using Databricks, and create an interactive Power BI dashboard for business users.

The dashboard provides key metrics, trends, and insights into sales, profit, and customer behavior.

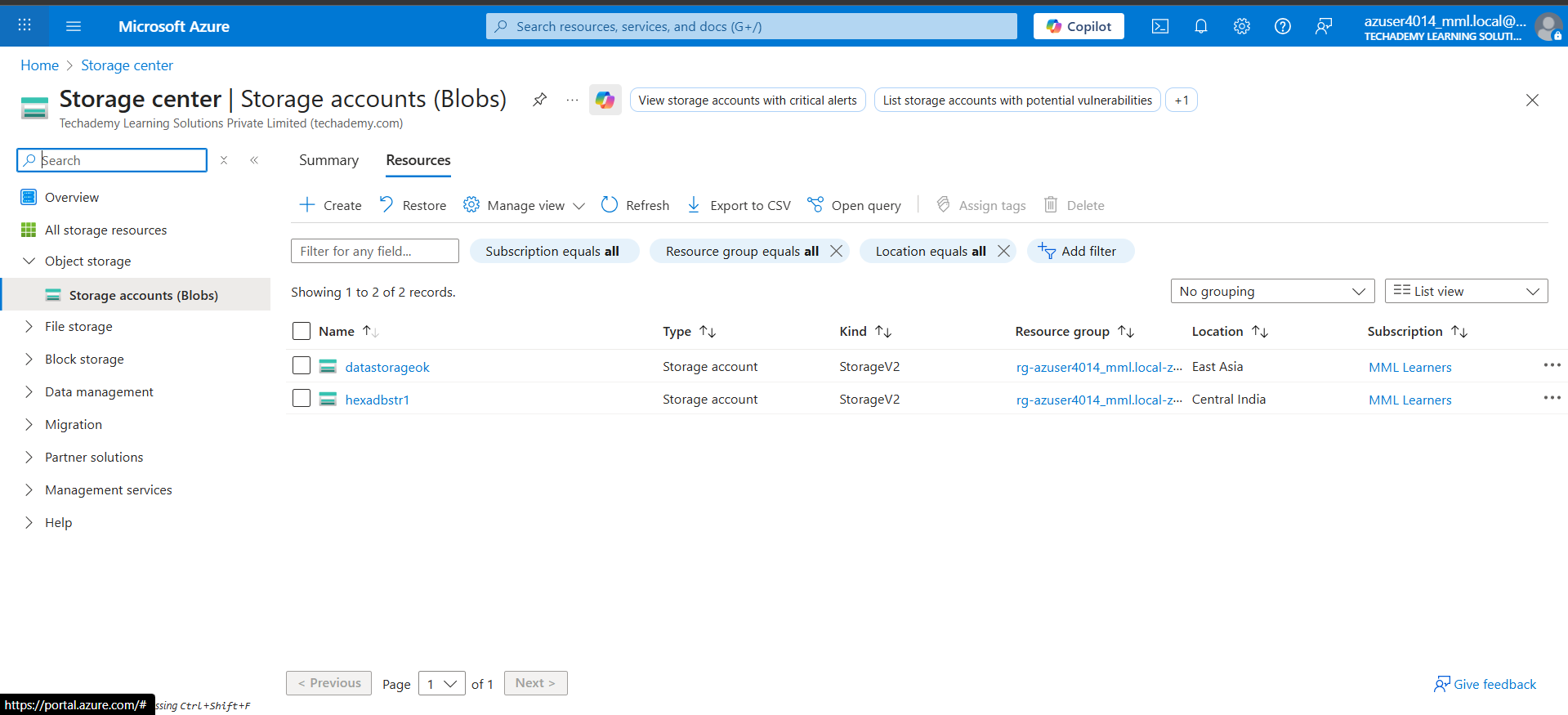
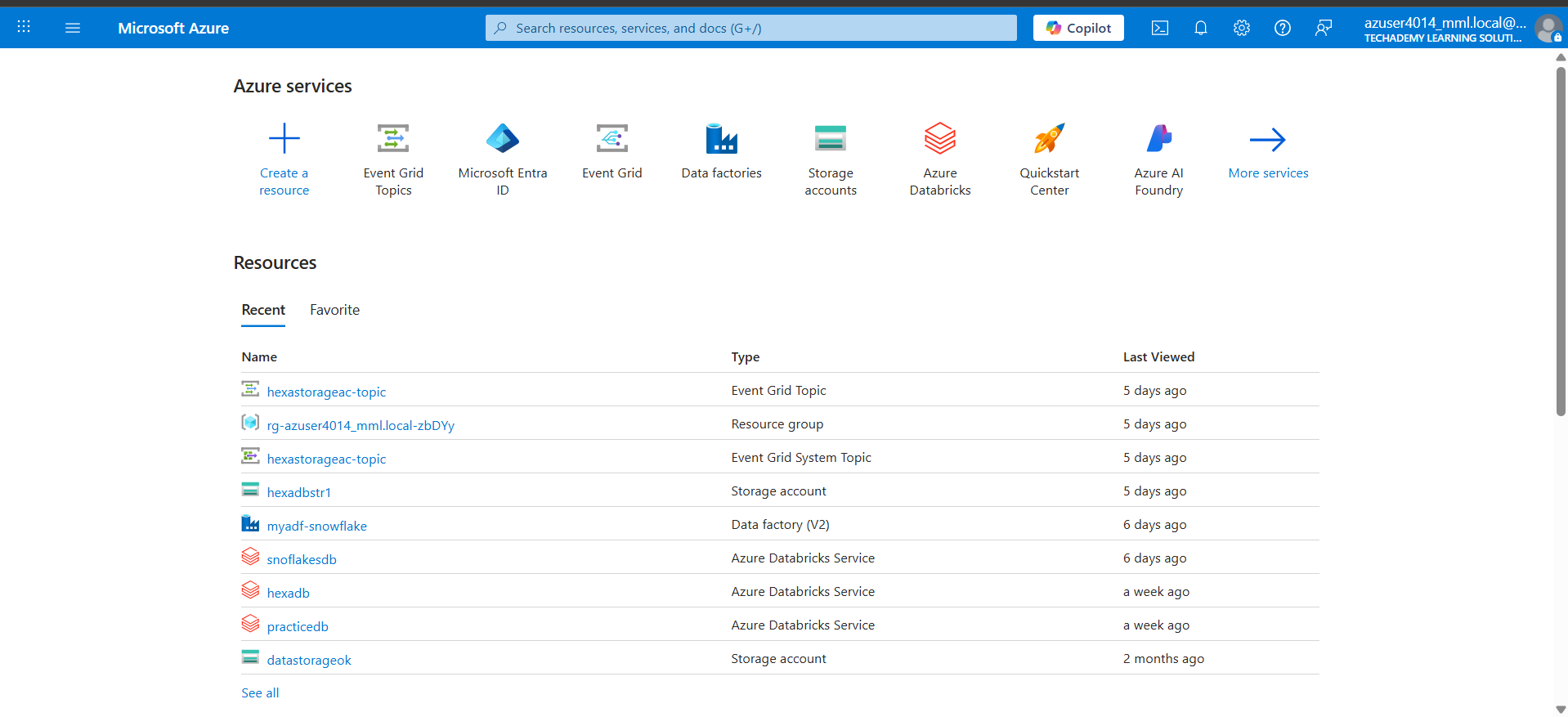
This workflow demonstrates the data pipeline:

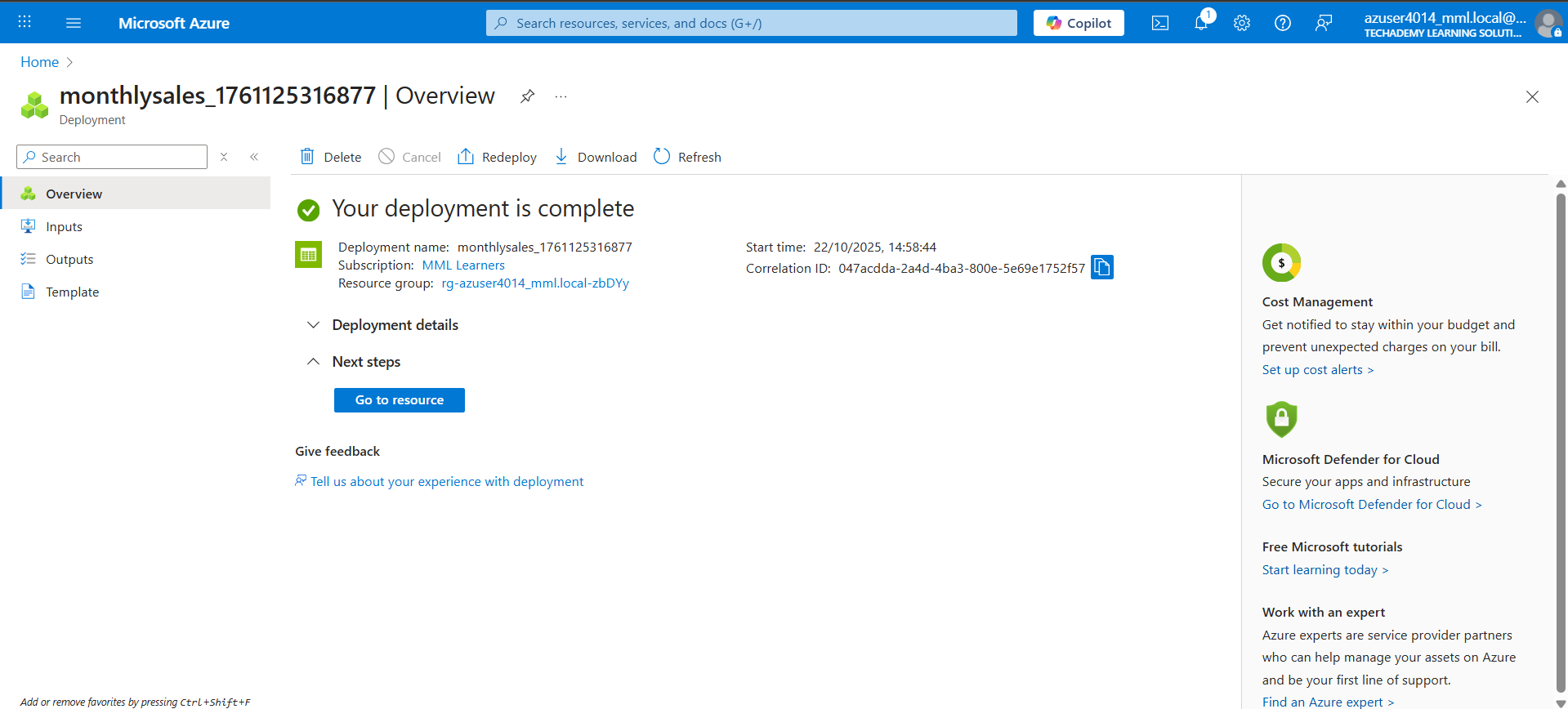
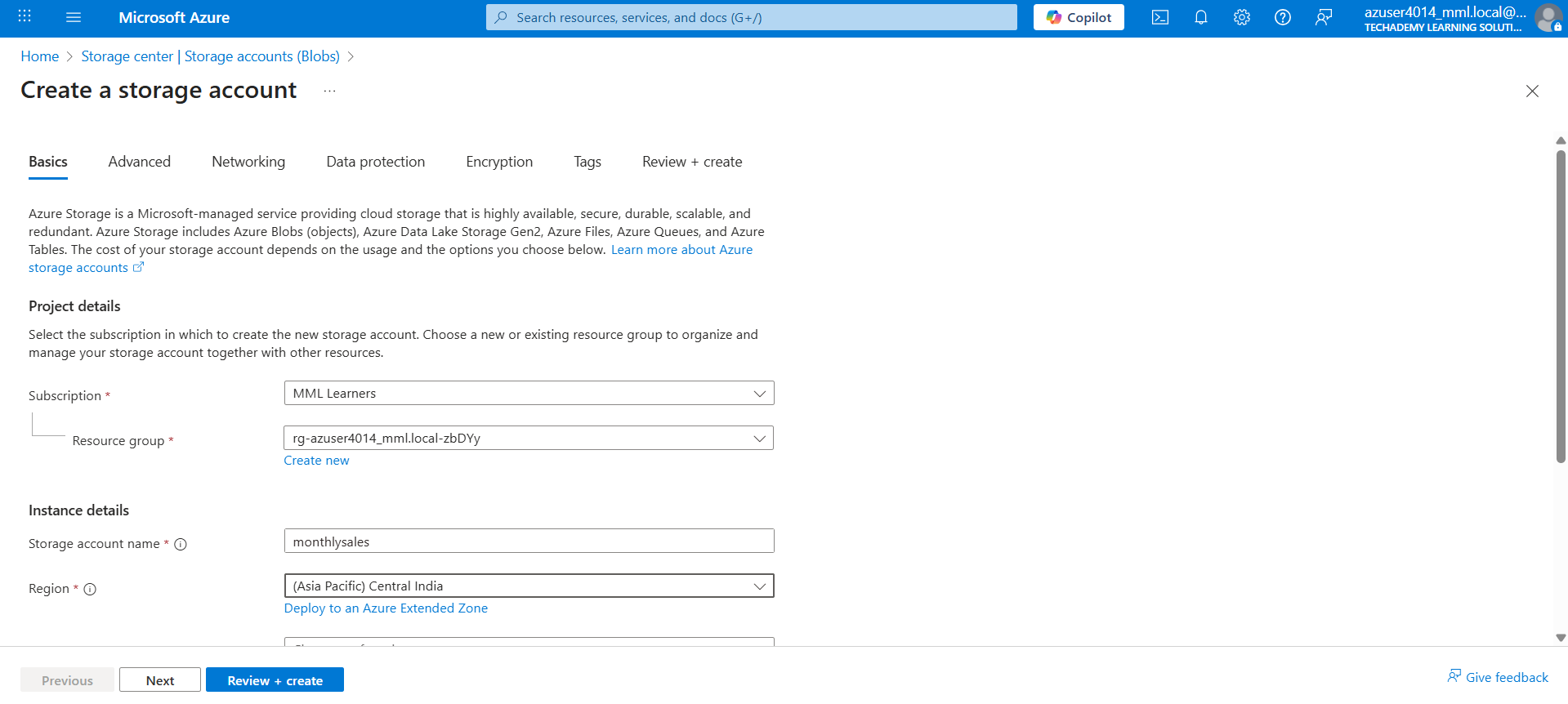
* Data Storage: Azure Blob Storage
* Data Ingestion: Azure Databricks using Snowflake connector
* Data Modeling: Snowflake (SALES table)
* Visualization: Power BI

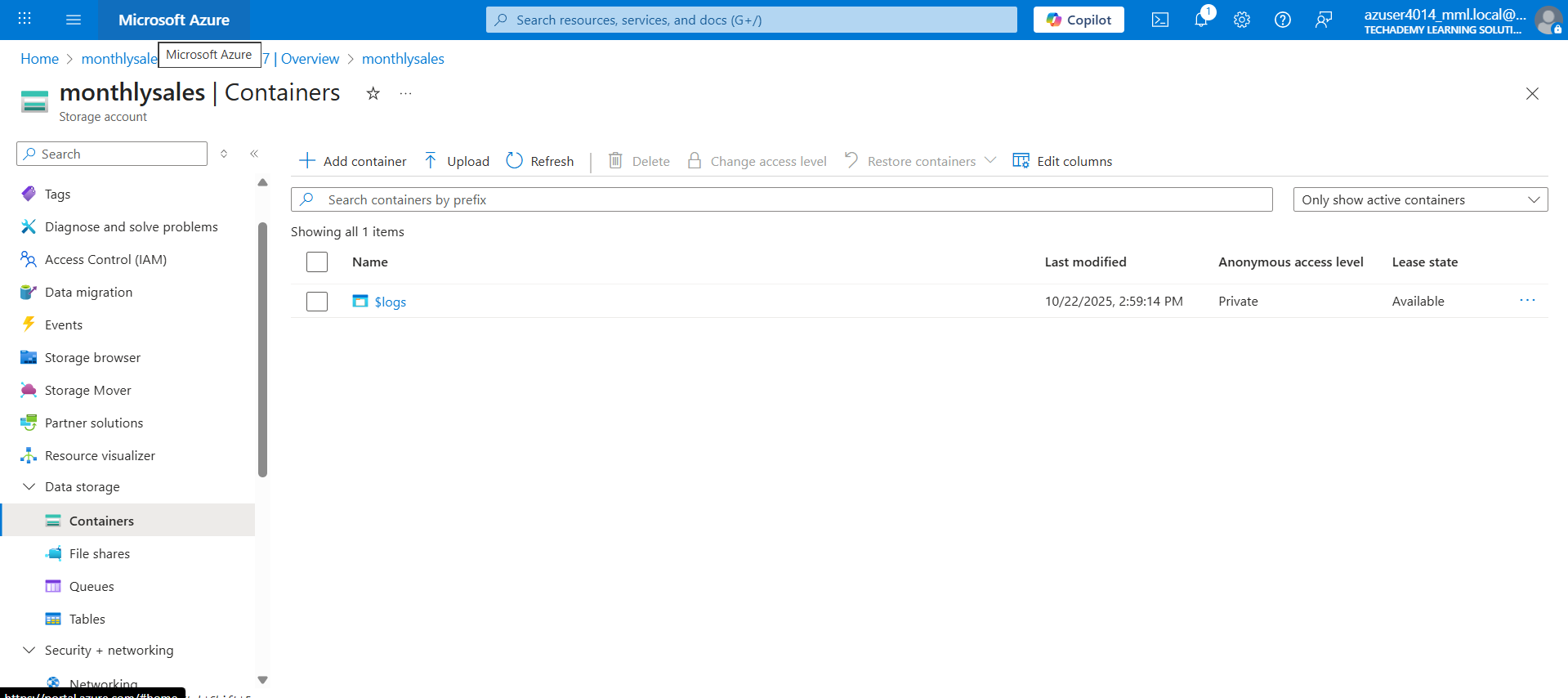
**Step 1: Azure Storage Setup**

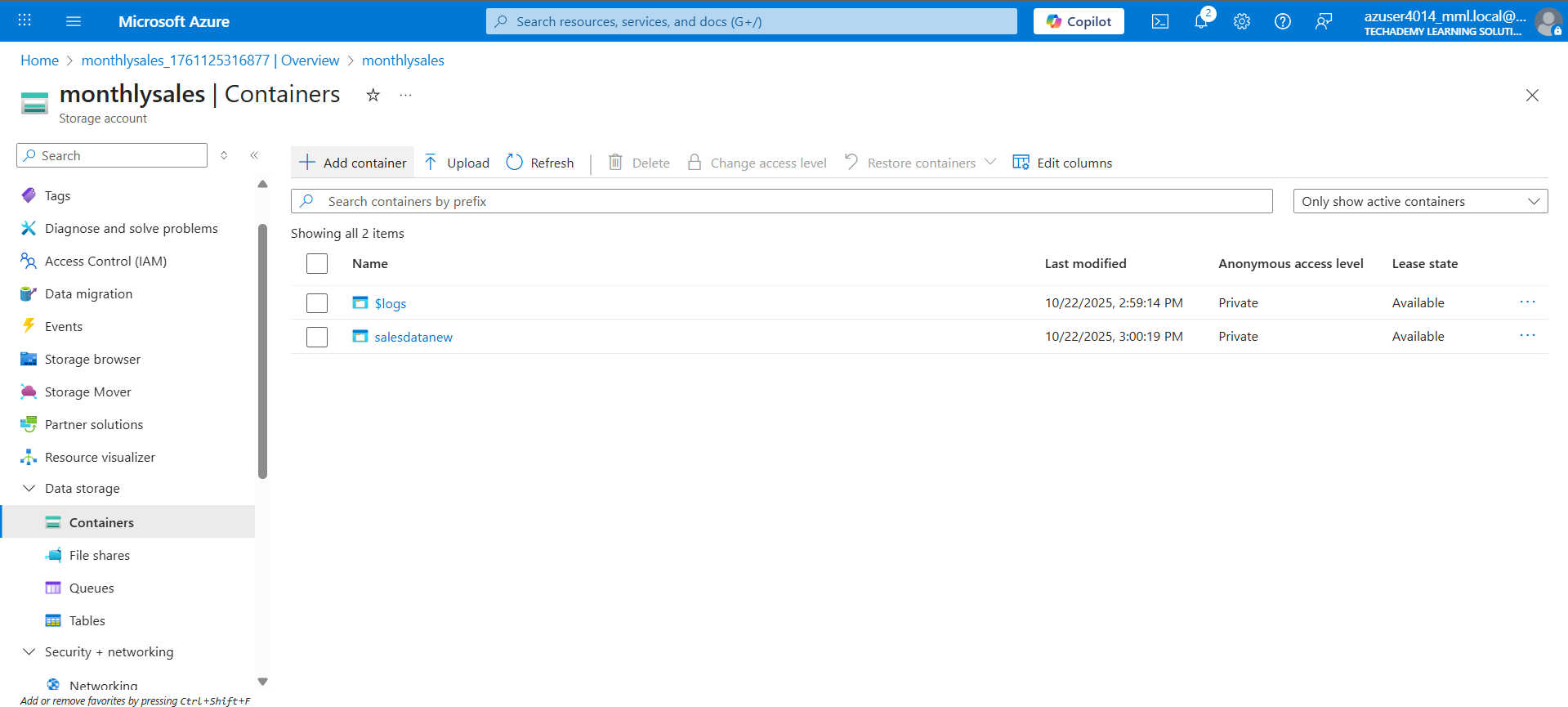
**Steps Taken:**

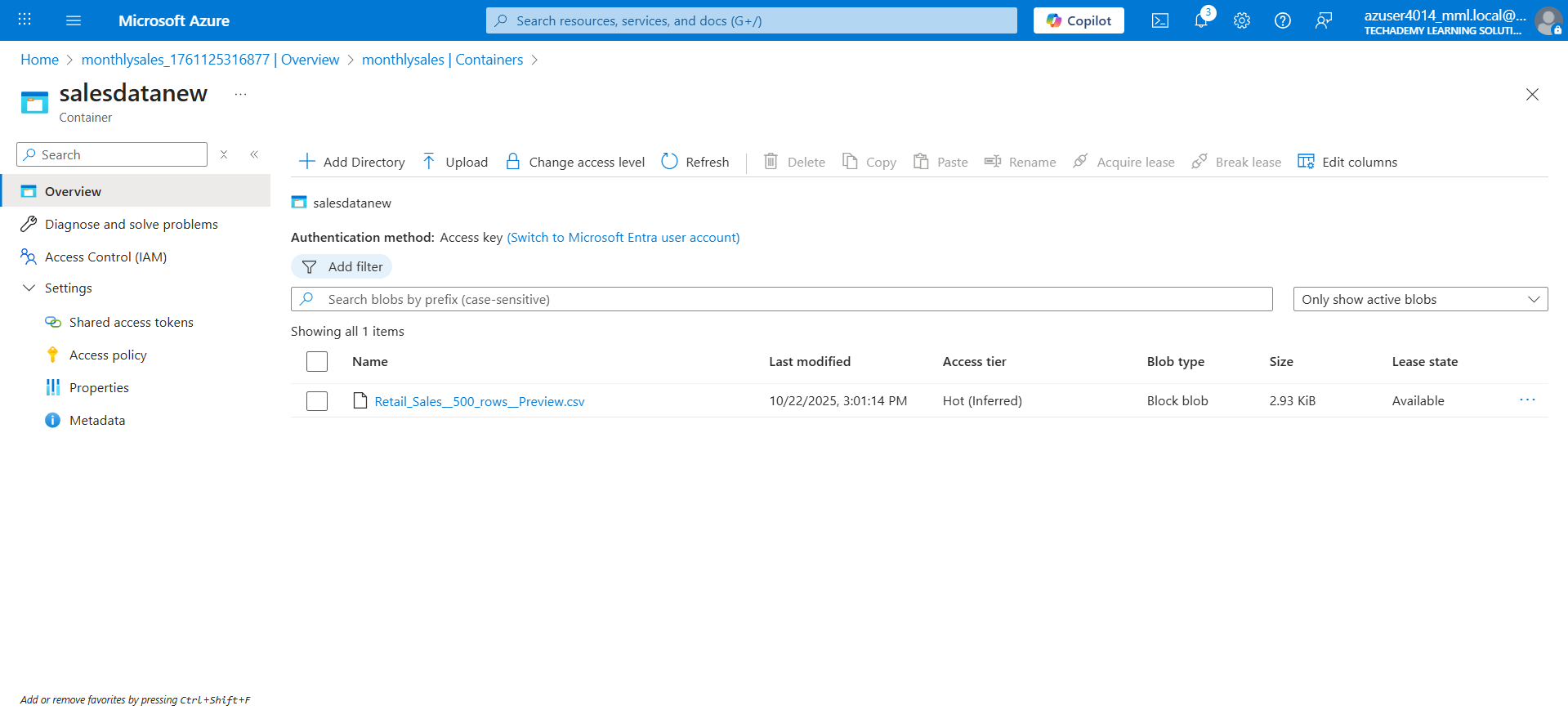
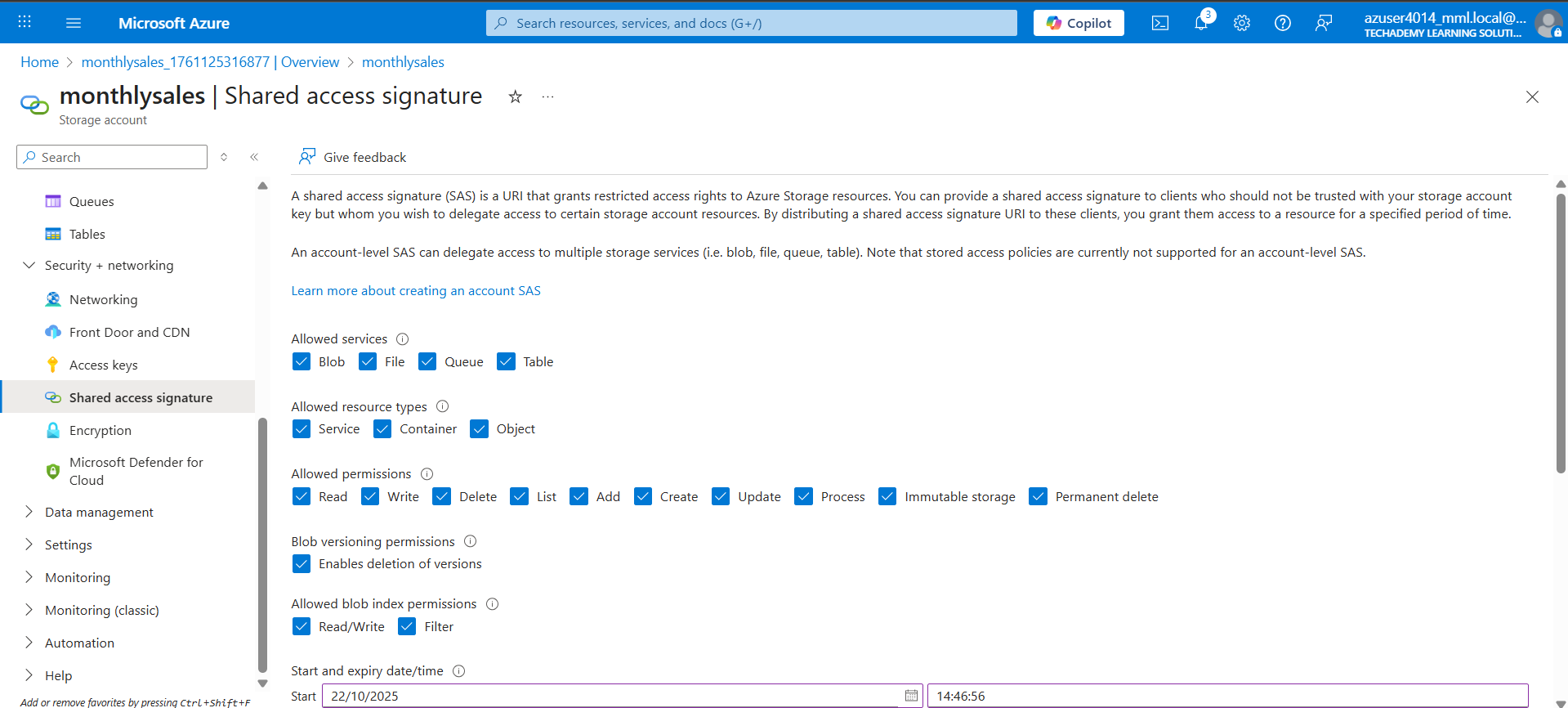
1. Logged into Azure Portal → Storage Accounts.
2. Clicked Create Storage Account, provided required details, and created the account.
3. Created a new container inside the storage account:
   * Clicked + New Container, entered a name, and created it.
4. Uploaded the CSV sales file to the container.
5. Generated a Shared Access Signature (SAS) token:
   * Navigated to Security + Networking → Shared Access Signature
   * Selected read/list permissions → copied SAS token and container URL.

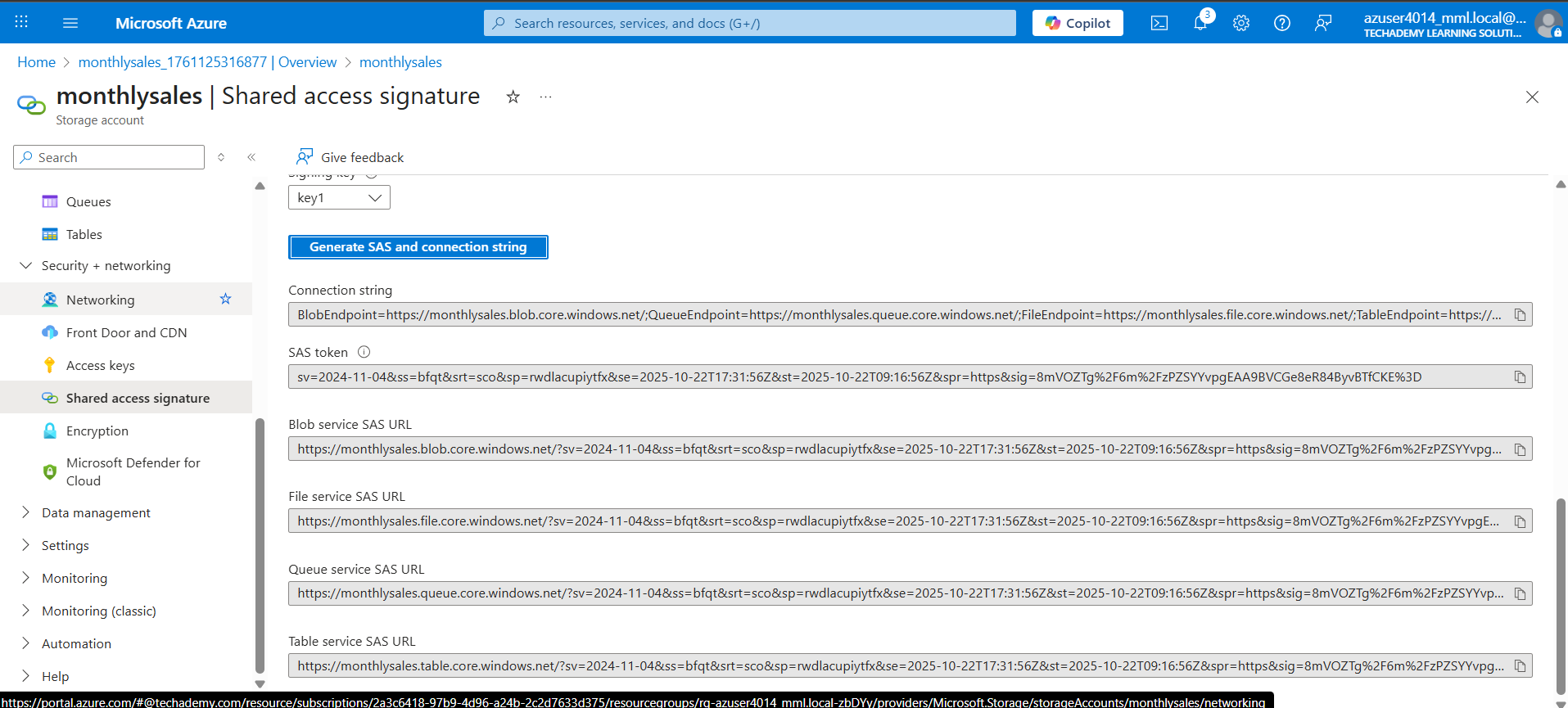








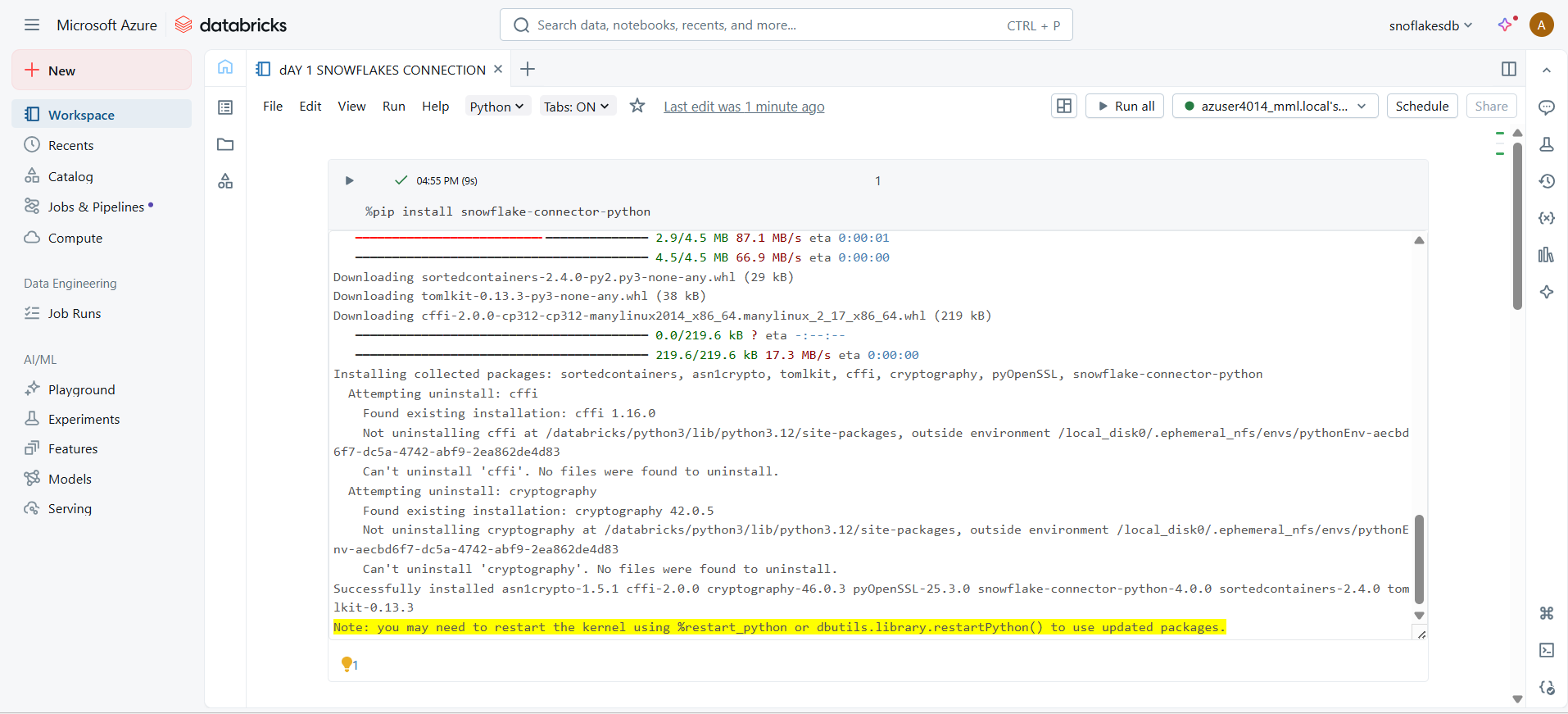




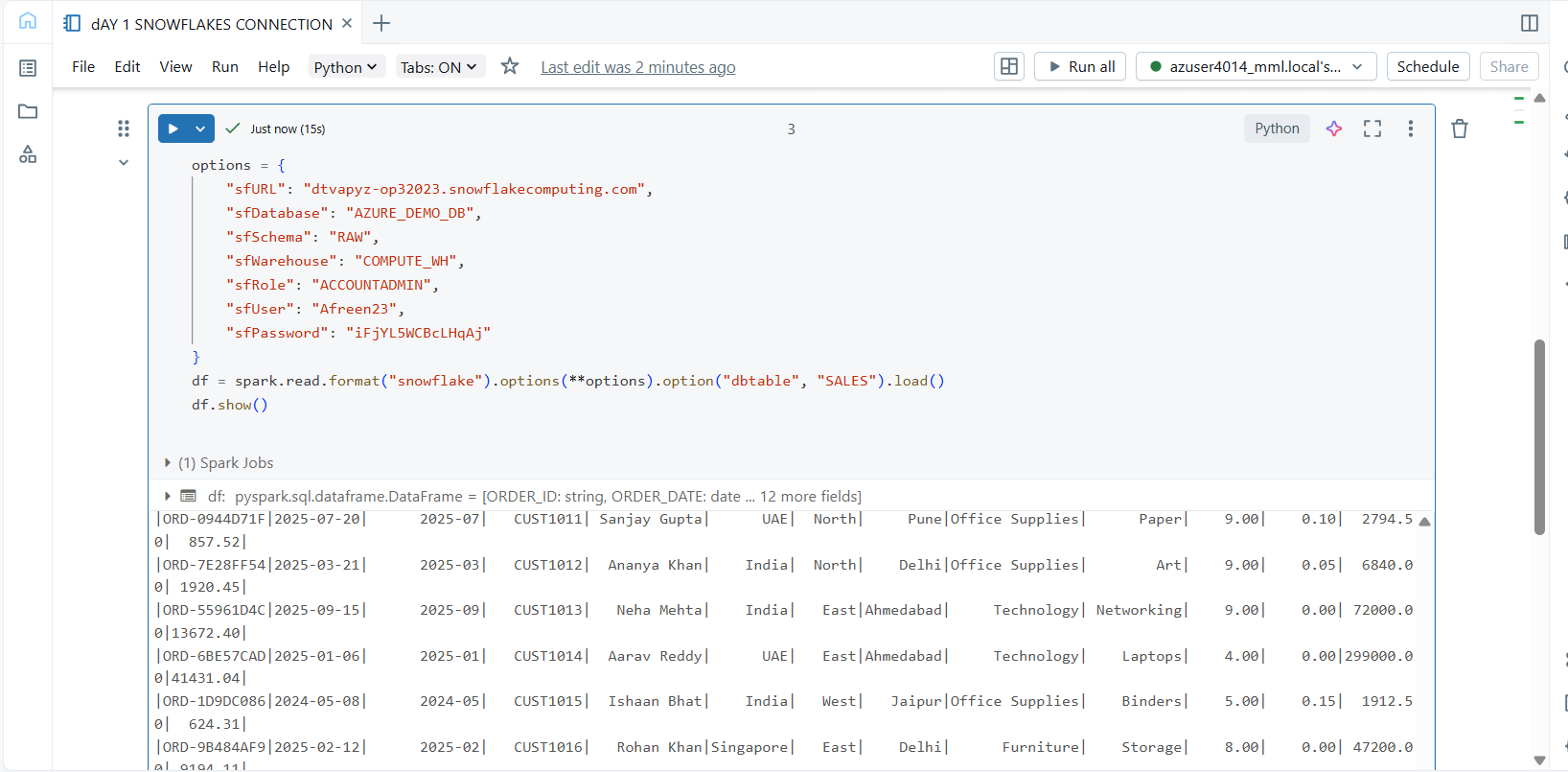
**Step 2: Data Ingestion into Snowflake using Databricks**

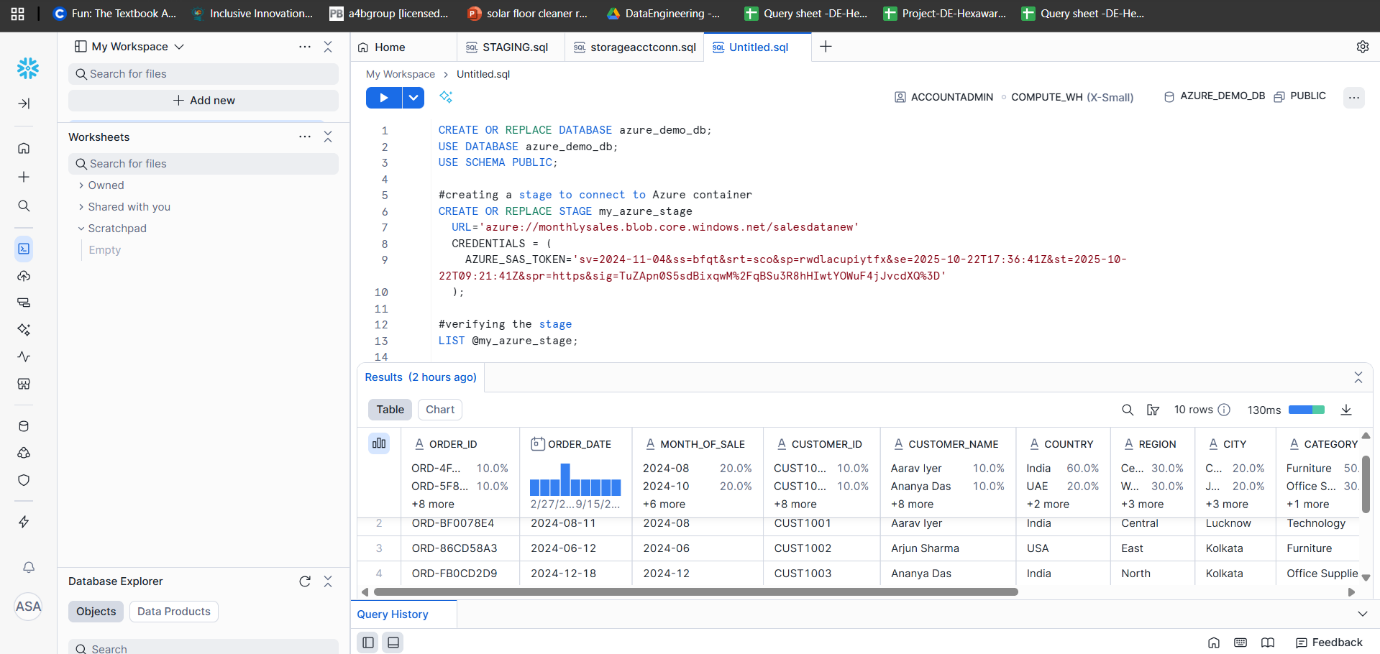
**Steps Taken:**

1. Opened Azure Databricks workspace.
2. Installed Snowflake connector library:



1. Configured connection options for Snowflake:
2. Created a **Spark DataFrame** reading from Snowflake SALES table:
3. Verified the loaded data:

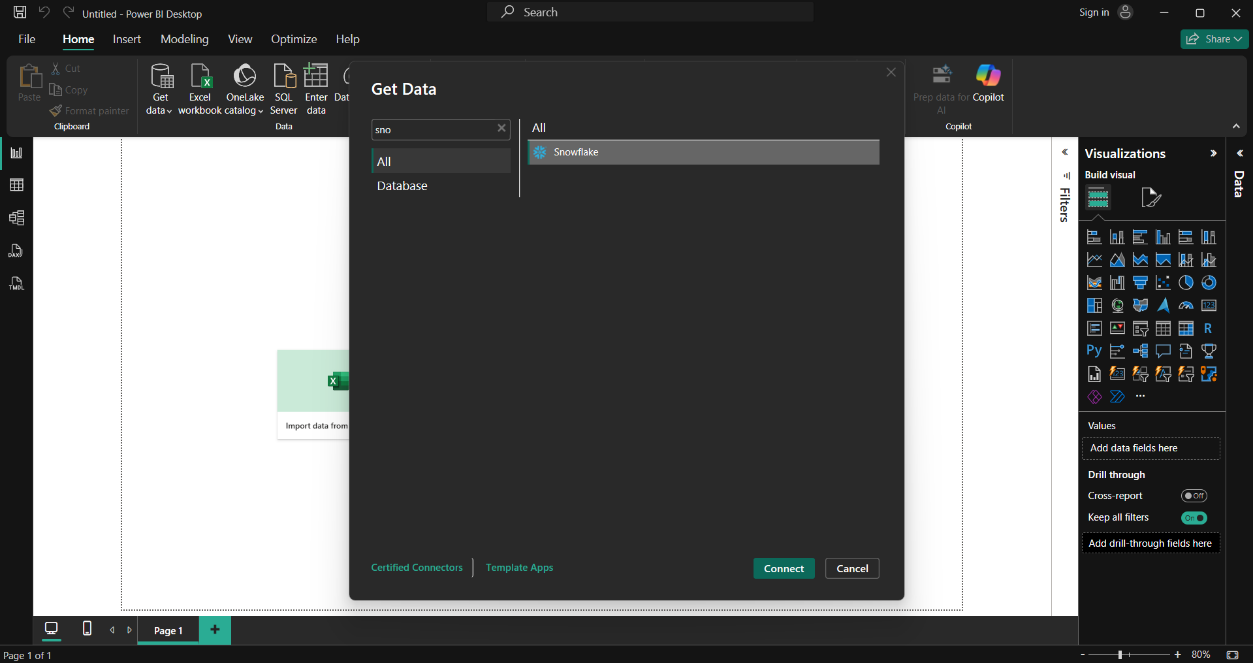


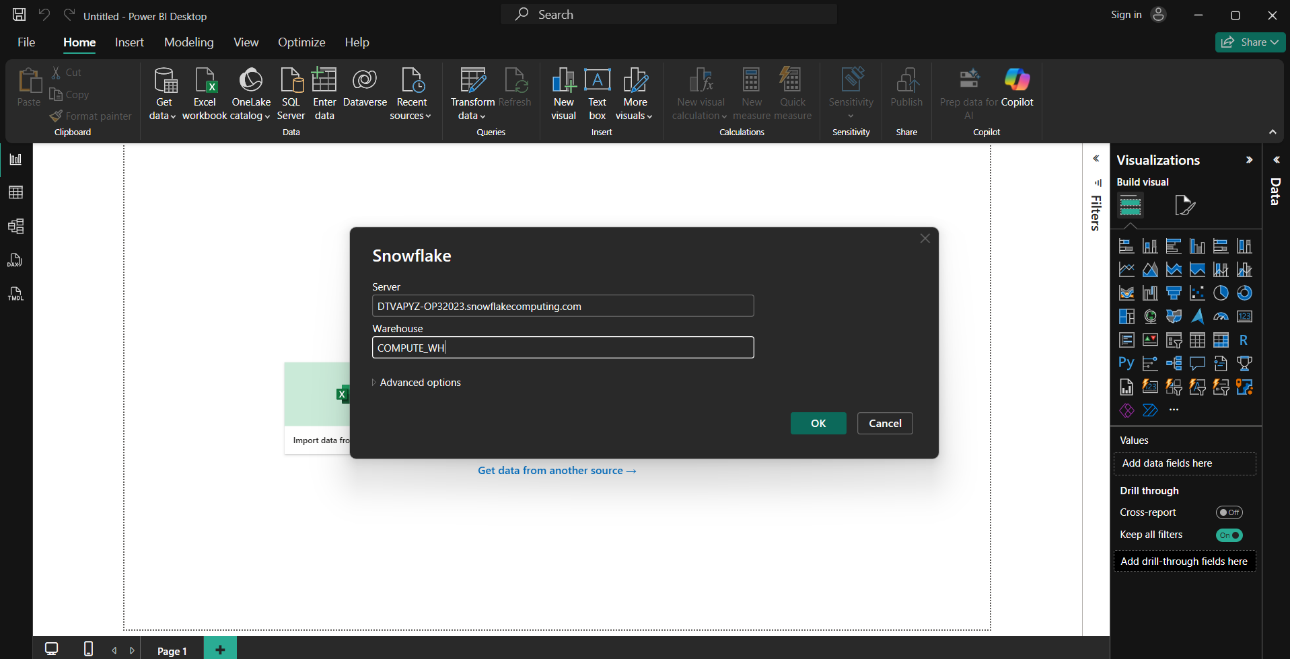


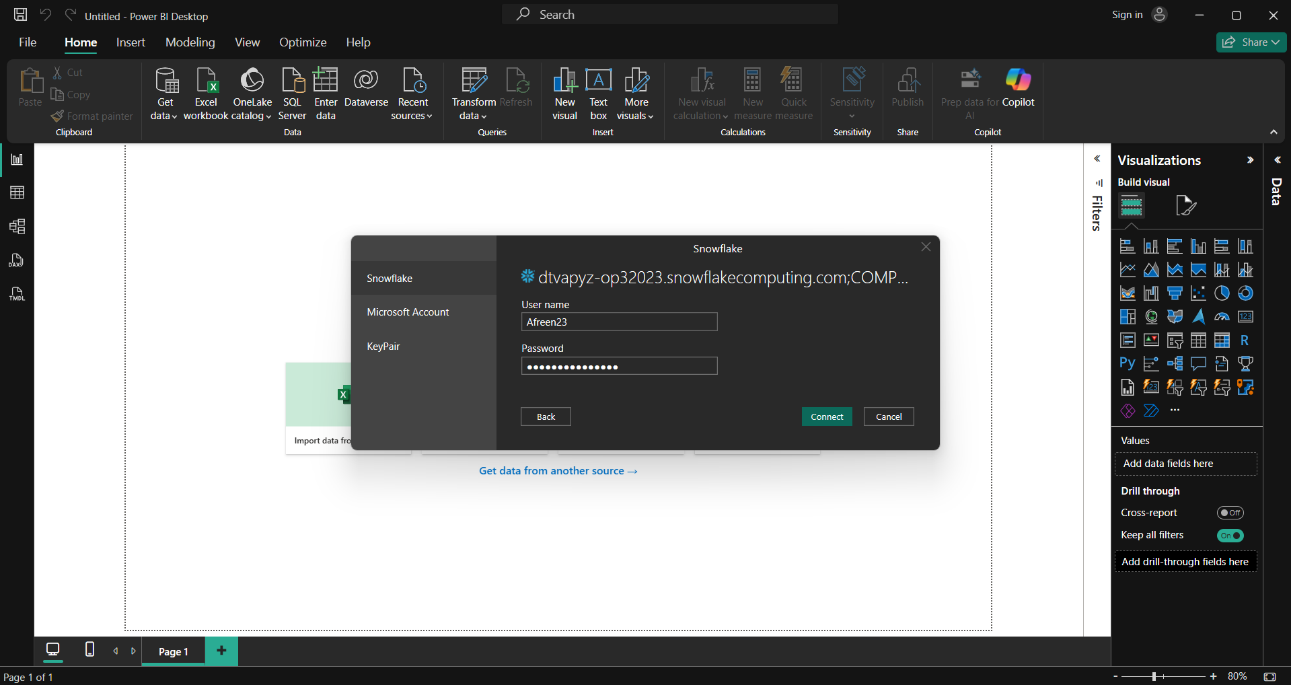
**Step 3: Power BI Report Creation**

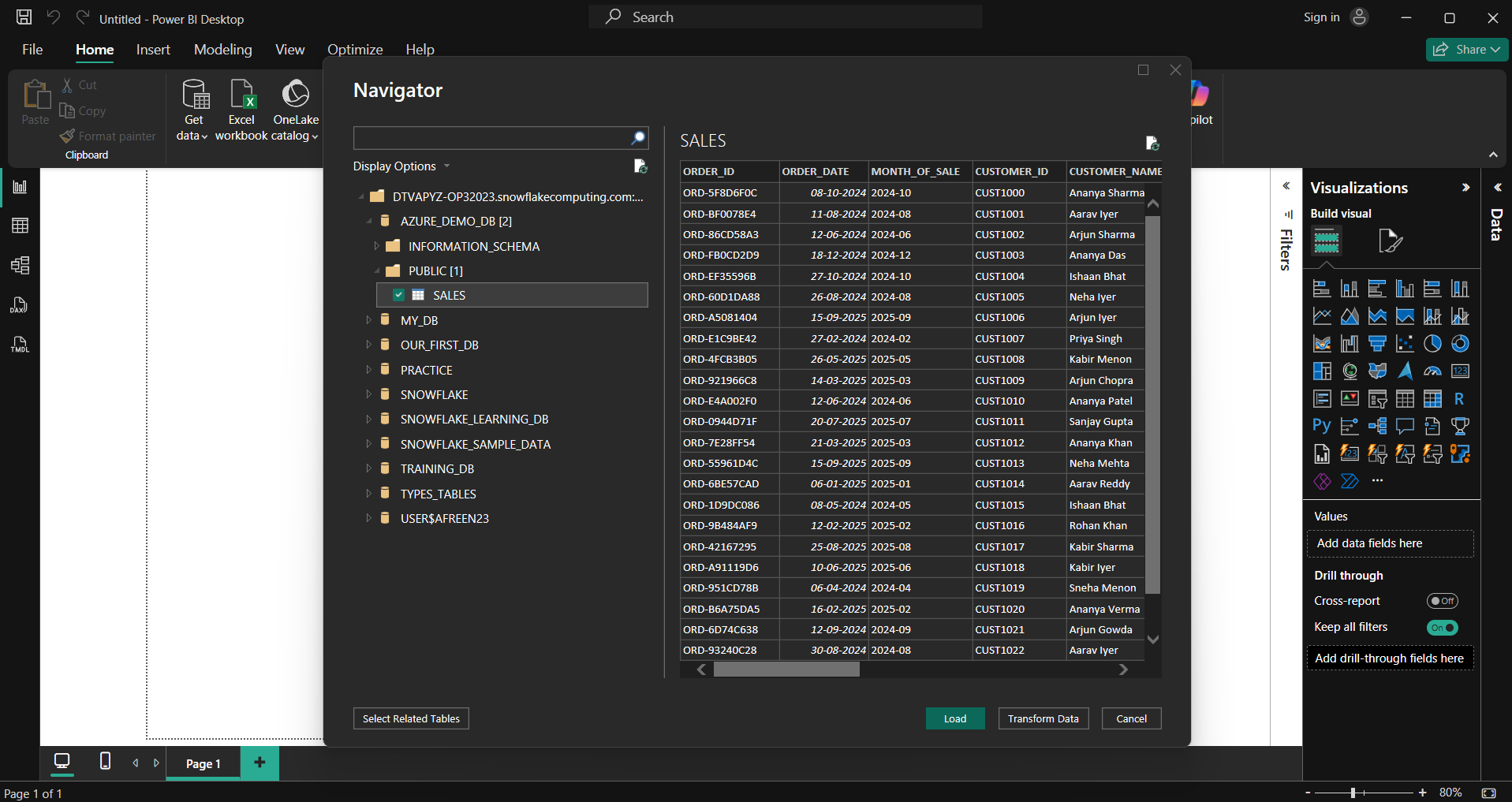
**Steps Taken:**

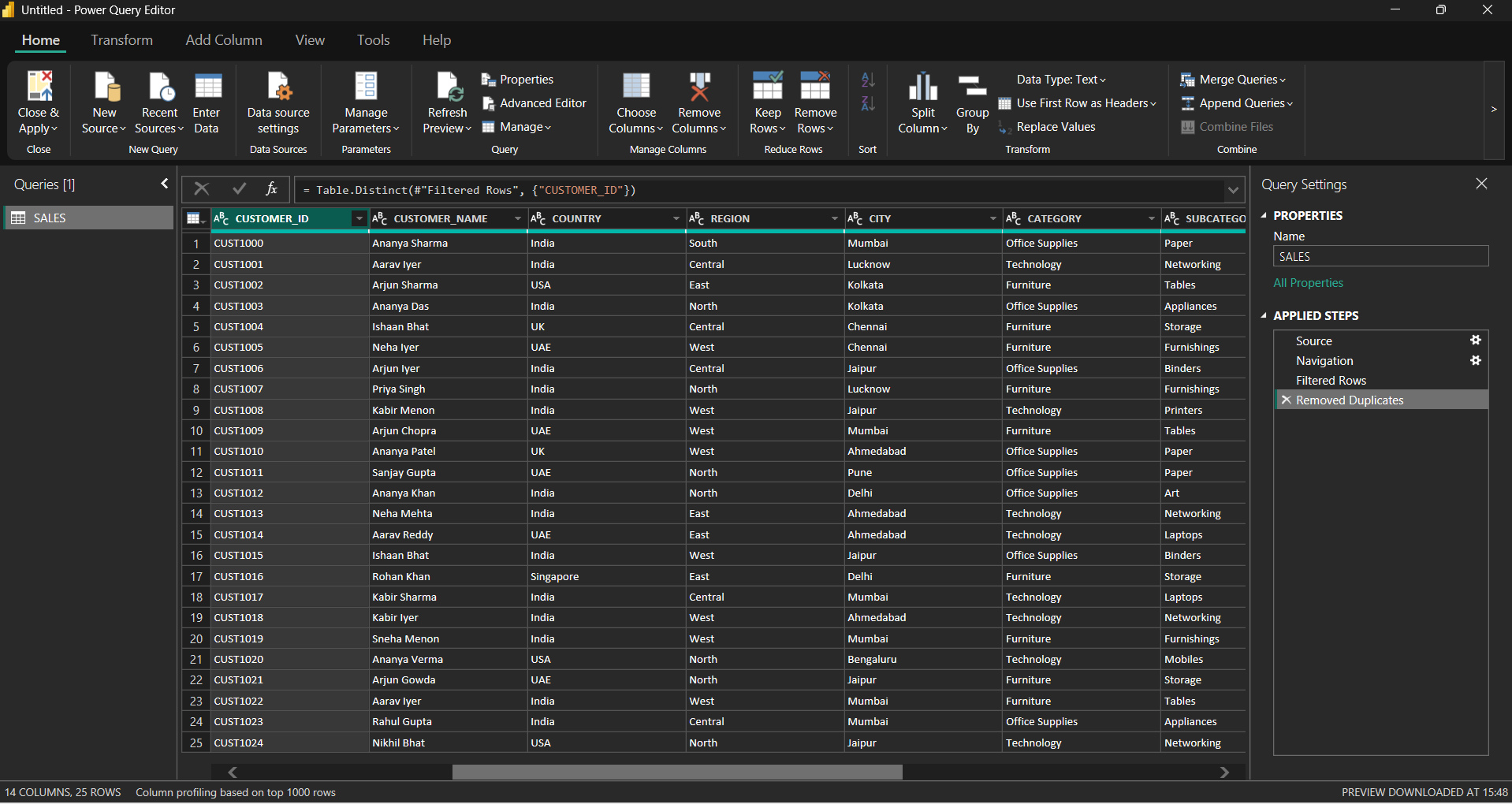
1. Opened Power BI Desktop.
2. Clicked Get Data → Snowflake, entered Snowflake server, warehouse, database, schema.
3. Selected Import mode to load SALES table.
4. Opened Power Query Editor:
   * Verified data types (Text, Date, Decimal).
   * Removed duplicate rows.
5. Loaded the cleaned data into Power BI.

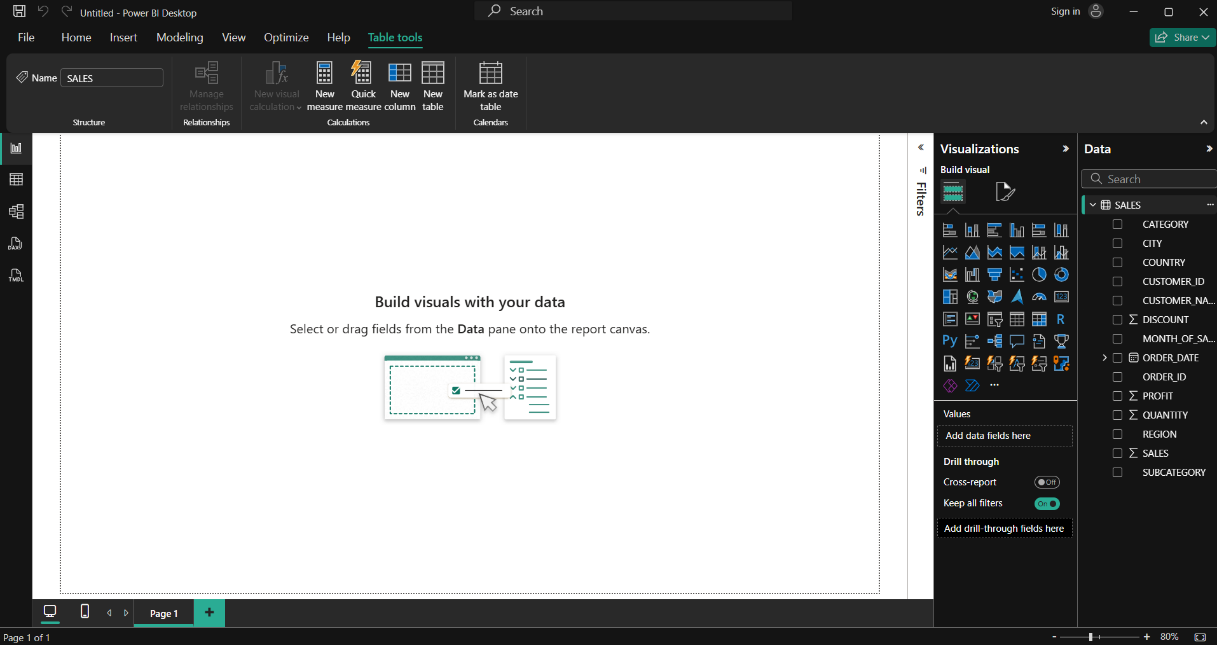








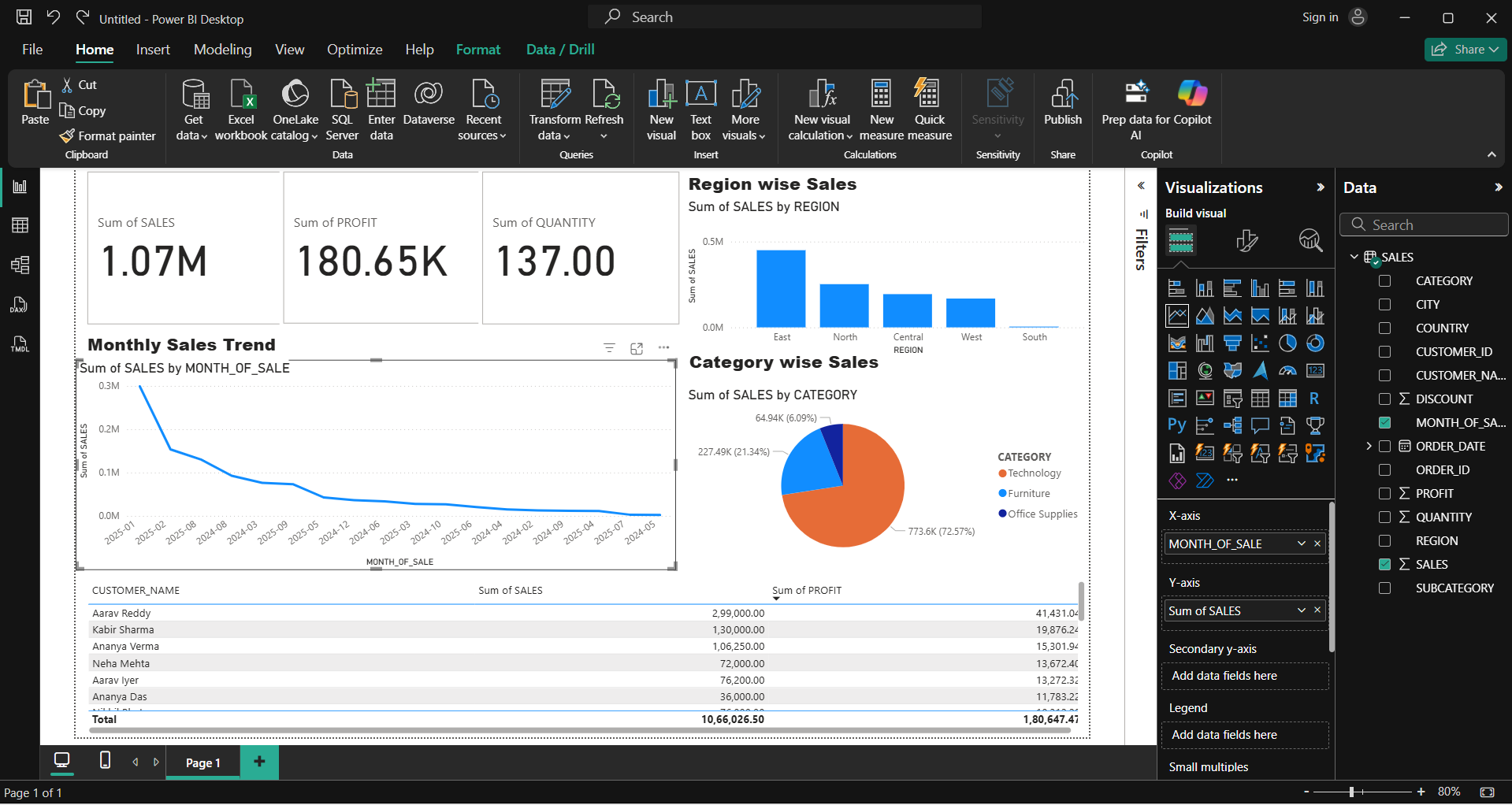




**6. Step 4: Dashboard Creation in Power BI**

**Visuals Created:**

1. **Top Metrics (Cards)**
   * Total Sales → Sum(SALES)
   * Total Profit → Sum(PROFIT)
   * Total Quantity Sold → Sum(QUANTITY)
2. **Monthly Sales Trend (Line Chart)**
   * X-axis: MONTH\_OF\_SALE
   * Y-axis: Sum(SALES)
3. **Region-wise Sales (Clustered Column Chart)**
   * X-axis: REGION
   * Values: Sum(SALES)
4. **Category-wise Sales (Pie Chart)**
   * Legend: CATEGORY
   * Values: Sum(SALES)
5. **Top Customers (Table)**
   * Rows: CUSTOMER\_NAME
   * Values: Sum(SALES), Sum(PROFIT)
   * Sorted descending by SALES



**Conclusion**

* Successfully loaded CSV from Azure Storage to Snowflake using Databricks.
* Verified the data in Snowflake using Databricks.
* Built a Power BI dashboard with key metrics and trends.
* Workflow demonstrates an end-to-end pipeline from Azure → Databricks → Snowflake → Power BI.