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

Aarthi. P - 22/10/25

Scenario:

You're the data engineer at ItTechGenie Retail. Sales teams drop monthly CSVs into an Azure Storage container.

You must:

- upload the CSV to Azure
- ingest it into Snowflake using Snowpark
- model it into proper database/schema/table
- build a quick Power BI report for business users.

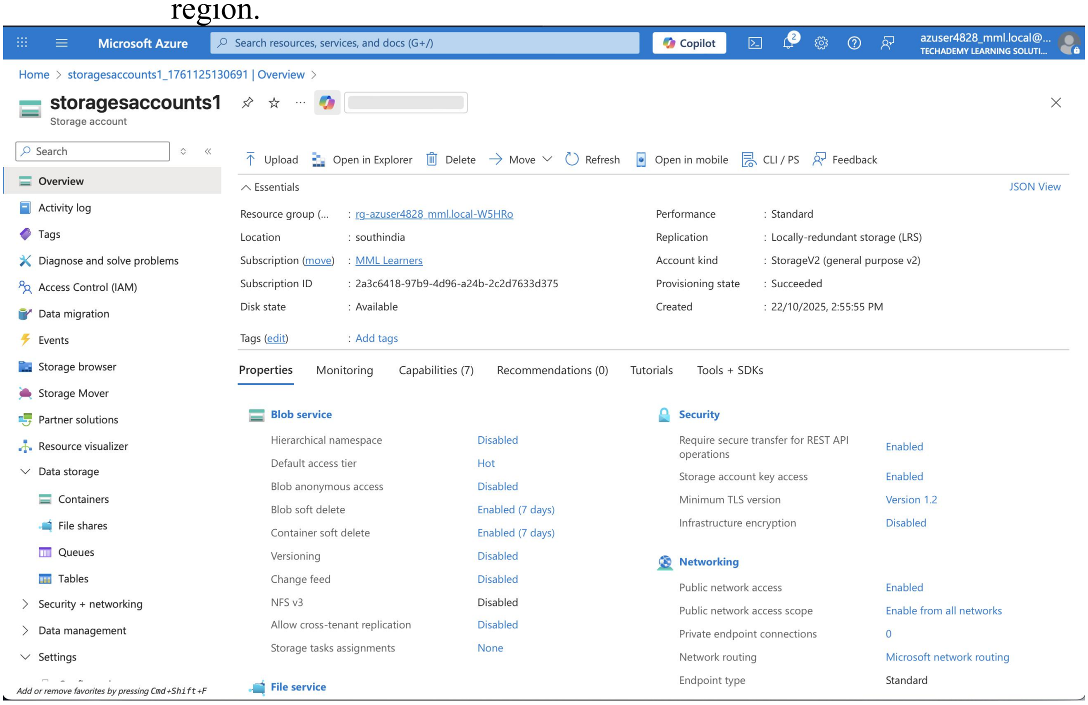
Goal:

Ensure that business users have access to clean, structured, and visualized sales data for reporting and analysis.

Step 1: Azure Storage Setup

Create an Azure Storage Account:

- Go to the Azure Portal → Storage Accounts → Create.
- Choose the subscription, resource group, storage account name, and region.



storagesaccounts1
Storage account

Essentials

Resource group (...	: rg-azuser4828_mml.local-W5HRo	Performance	: Standard
Location	: southindia	Replication	: Locally-redundant storage (LRS)
Subscription (move)	: MML Learners	Account kind	: StorageV2 (general purpose v2)
Subscription ID	: 2a3c6418-97b9-4d96-a24b-2c2d7633d375	Provisioning state	: Succeeded
Disk state	: Available	Created	: 22/10/2025, 2:55:55 PM
Tags (edit)	: Add tags		

Blob service

Hierarchical namespace	Disabled
Default access tier	Hot
Blob anonymous access	Disabled
Blob soft delete	Enabled (7 days)
Container soft delete	Enabled (7 days)
Versioning	Disabled
Change feed	Disabled
NFS v3	Disabled
Allow cross-tenant replication	Disabled
Storage tasks assignments	None

File service

Security

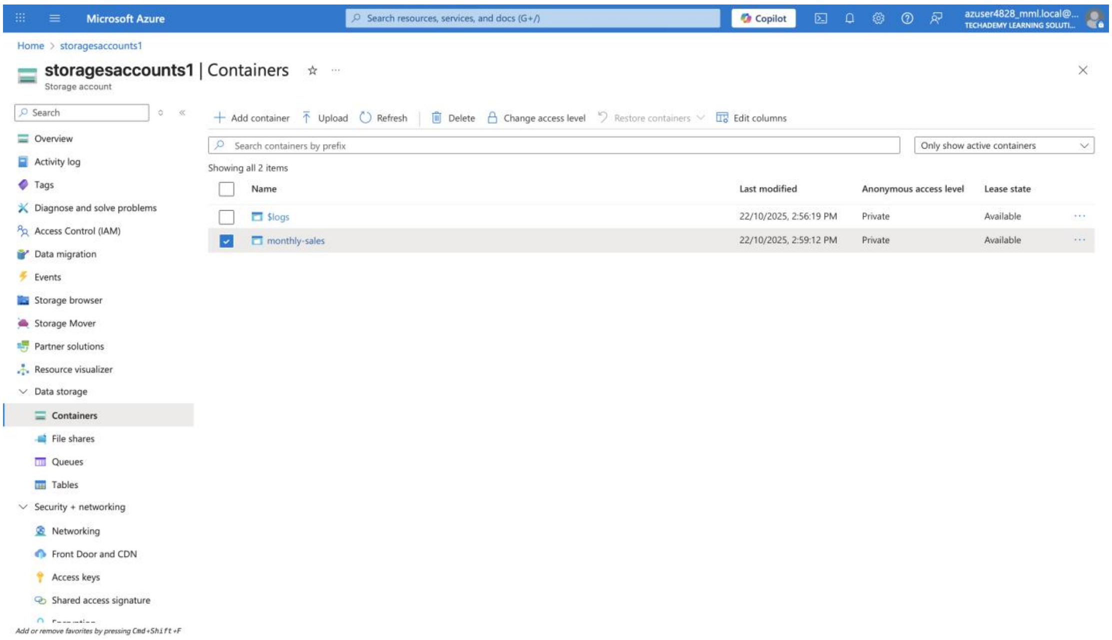
Require secure transfer for REST API operations	Enabled
Storage account key access	Enabled
Minimum TLS version	Version 1.2
Infrastructure encryption	Disabled

Networking

Public network access	Enabled
Public network access scope	Enable from all networks
Private endpoint connections	0
Network routing	Microsoft network routing
Endpoint type	Standard

Create a Container:

- Inside the storage account, create a container (e.g., sales-monthly).
- Set access level to private.

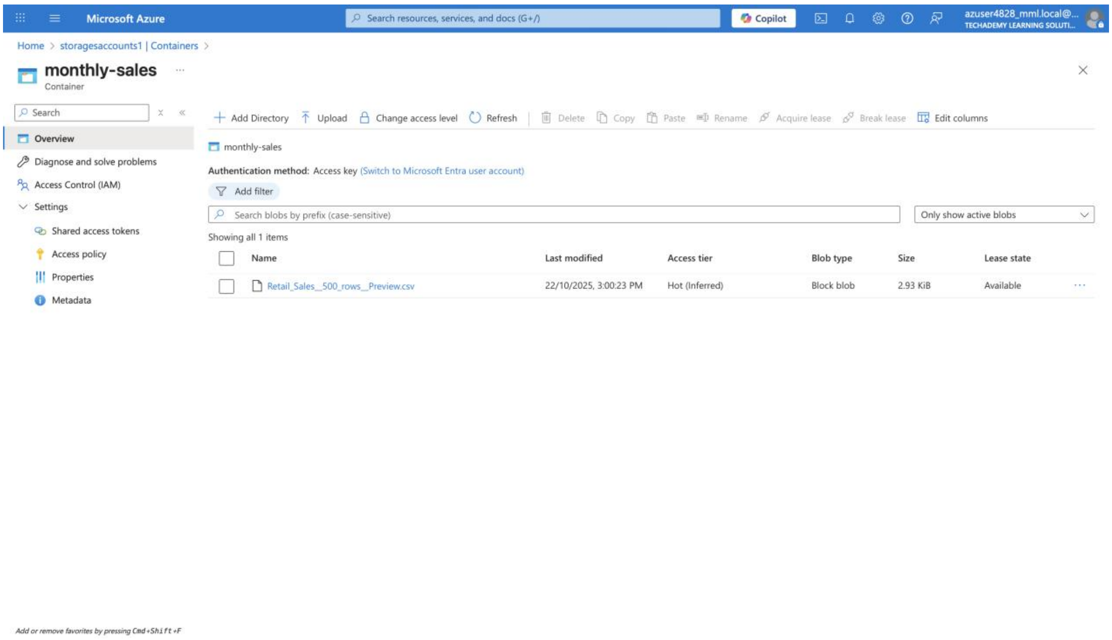


Showing all 2 items

	Name	Last modified	Anonymous access level	Lease state
□	$logs	22/10/2025, 2:56:19 PM	Private	Available
☑	monthly-sales	22/10/2025, 2:59:12 PM	Private	Available

Upload CSV Files:

- Upload monthly CSVs (e.g., Retail_Sales__500_rows__Preview.csv).



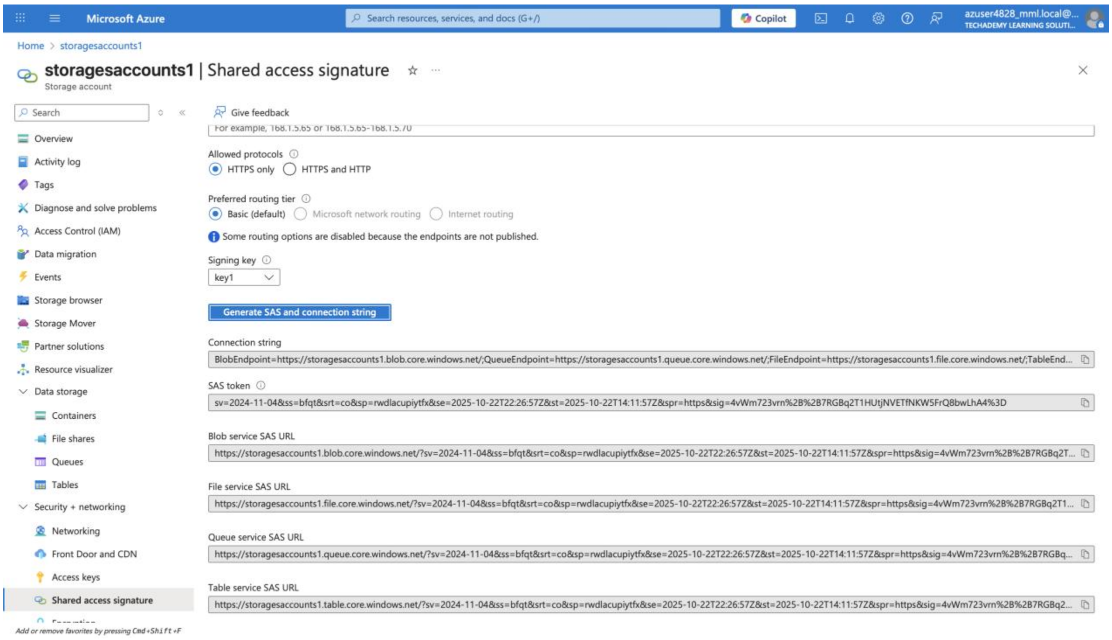
Authentication method: Access key (Switch to Microsoft Entra user account)

Showing all 1 items

	Name	Last modified	Access tier	Blob type	Size	Lease state
□	Retail_Sales__500_rows__Preview.csv	22/10/2025, 3:00:23 PM	Hot (Inferred)	Block blob	2.93 KiB	Available

Generate SAS Token for Access:

- Click on the file or container → Generate SAS.
- Set permissions (Read, List, Write if needed).
- Set expiry date/time.
- Copy the full URL including SAS token for later use.



storagesaccounts1 | Shared access signature

Storage account

For example, 168.1.5.65 or 168.1.5.65-168.1.5.70

Allowed protocols
● HTTPS only ○ HTTPS and HTTP

Preferred routing tier
● Basic (default) ○ Microsoft network routing ○ Internet routing

Some routing options are disabled because the endpoints are not published.

Signing key
key1

Generate SAS and connection string

Connection string
```
BlobEndpoint=https://storagesaccounts1.blob.core.windows.net/;QueueEndpoint=https://storagesaccounts1.queue.core.windows.net/;FileEndpoint=https://storagesaccounts1.file.core.windows.net/;TableEnd...
```

SAS token
```
sv=2024-11-04&ss=bfqt&srt=co&sp=rwdlacupiytfx&se=2025-10-22T22:26:57Z&st=2025-10-22T14:11:57Z&spr=https&sig=4vWm723vrn%2B%2B7RGBq2T1HUtjNVETfNKW5FrQ8bwLhA4%3D
```

Blob service SAS URL
```
https://storagesaccounts1.blob.core.windows.net/?sv=2024-11-04&ss=bfqt&srt=co&sp=rwdlacupiytfx&se=2025-10-22T22:26:57Z&st=2025-10-22T14:11:57Z&spr=https&sig=4vWm723vrn%2B%2B7RGBq2T...
```

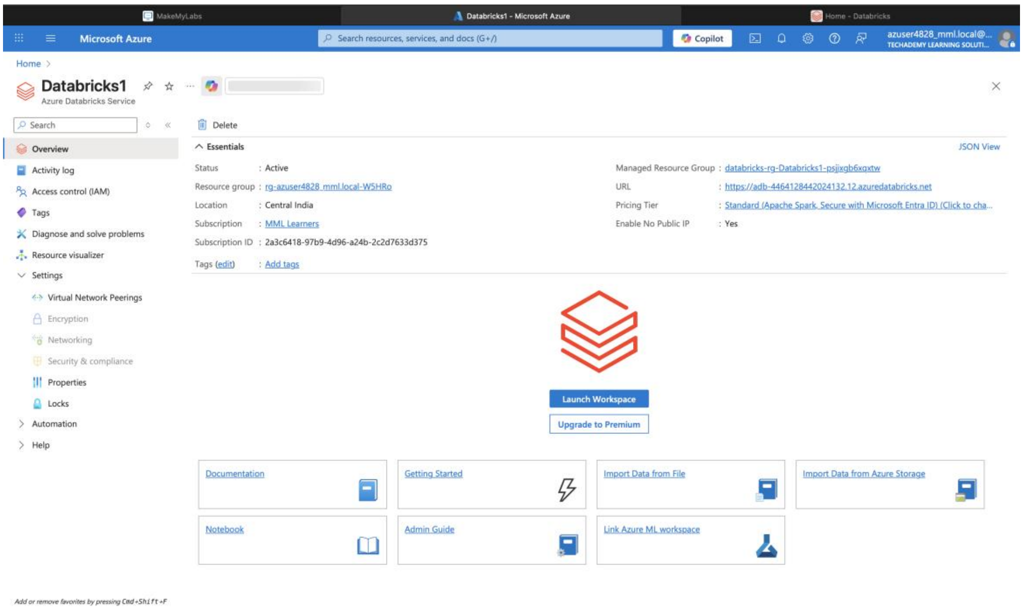
File service SAS URL
```
https://storagesaccounts1.file.core.windows.net/?sv=2024-11-04&ss=bfqt&srt=co&sp=rwdlacupiytfx&se=2025-10-22T22:26:57Z&st=2025-10-22T14:11:57Z&spr=https&sig=4vWm723vrn%2B%2B7RGBq2T1...
```

Queue service SAS URL
```
https://storagesaccounts1.queue.core.windows.net/?sv=2024-11-04&ss=bfqt&srt=co&sp=rwdlacupiytfx&se=2025-10-22T22:26:57Z&st=2025-10-22T14:11:57Z&spr=https&sig=4vWm723vrn%2B%2B7RGBq...
```

Table service SAS URL
```
https://storagesaccounts1.table.core.windows.net/?sv=2024-11-04&ss=bfqt&srt=co&sp=rwdlacupiytfx&se=2025-10-22T22:26:57Z&st=2025-10-22T14:11:57Z&spr=https&sig=4vWm723vrn%2B%2B7RGBq2...
```

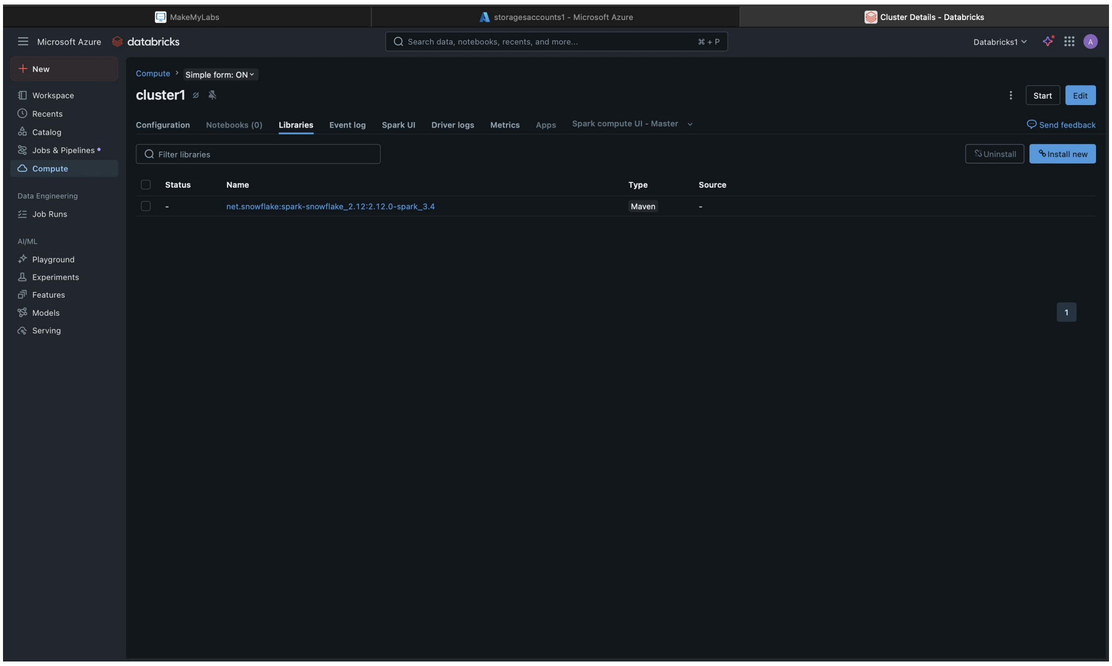
Step 2: Azure Databricks Setup:

- Create a Databricks Workspace (if not already available).



Status	: Active	Managed Resource Group	: databricks-rg-Databricks1-psjjxgb6xqxtw
Resource group	: rg-azuser4828_mml.local-W5HRo	URL	: https://adb-4464128442024132.12.azuredatabricks.net
Location	: Central India	Pricing Tier	: Standard (Apache Spark, Secure with Microsoft Entra ID) (Click to cha...
Subscription	: MML Learners	Enable No Public IP	: Yes
Subscription ID	: 2a3c6418-97b9-4d96-a24b-2c2d7633d375		
Tags (edit)	: Add tags		

- Launch a Cluster with appropriate size.



□	Status	Name	Type	Source
□	-	net.snowflake:spark-snowflake_2.12:2.12.0-spark_3.4	Maven	-

- Install Required Libraries.

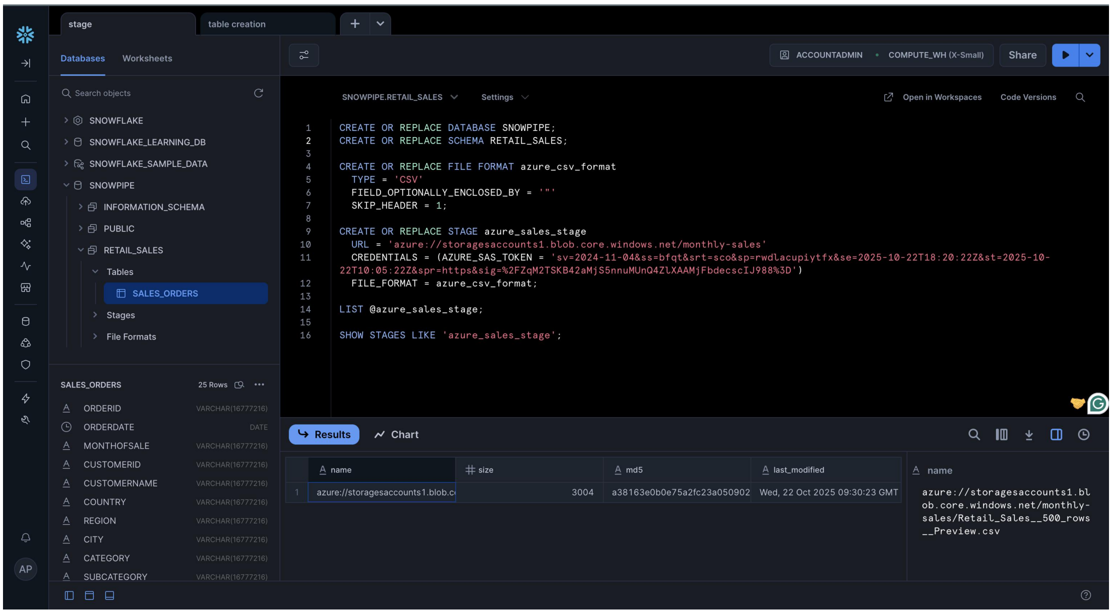
Step 2: External stage creation:

```
CREATE OR REPLACE DATABASE SNOWPIPE;
CREATE OR REPLACE SCHEMA RETAIL_SALES;

CREATE OR REPLACE FILE FORMAT azure_csv_format
  TYPE = 'CSV'
  FIELD_OPTIONALLY_ENCLOSED_BY = '"'
  SKIP_HEADER = 1;

CREATE OR REPLACE STAGE azure_sales_stage
```

```
          URL = 'azure://storagesaccounts1.blob.core.windows.net/monthly-sales'
          CREDENTIALS = (AZURE_SAS_TOKEN = 'sv=2024-11-04&ss=bfqt&srt=sco&sp=rwdlacupiytfx&se=2025-10-22T18:20:22Z&st=2025-10-22T10:05:22Z&spr=https&sig=%2FZqM2TSKB42aMjS5nnuMUnQ4ZlXAAMjFbdecscIJ988%3D')
          FILE_FORMAT = azure_csv_format;
```

Step 3.1: Verification:

```
LIST @azure_sales_stage;

SHOW STAGES LIKE 'azure_sales_stage';
```



SALES_ORDERS 25 Rows

ORDERID	VARCHAR(16777216)
ORDERDATE	DATE
MONTHOFSALE	VARCHAR(16777216)
CUSTOMERID	VARCHAR(16777216)
CUSTOMERNAME	VARCHAR(16777216)
COUNTRY	VARCHAR(16777216)
REGION	VARCHAR(16777216)
CITY	VARCHAR(16777216)
CATEGORY	VARCHAR(16777216)
SUBCATEGORY	VARCHAR(16777216)

```
CREATE OR REPLACE DATABASE SNOWPIPE;
CREATE OR REPLACE SCHEMA RETAIL_SALES;

CREATE OR REPLACE FILE FORMAT azure_csv_format
  TYPE = 'CSV'
  FIELD_OPTIONALLY_ENCLOSED_BY = '"'
  SKIP_HEADER = 1;

CREATE OR REPLACE STAGE azure_sales_stage
  URL = 'azure://storagesaccounts1.blob.core.windows.net/monthly-sales'
  CREDENTIALS = (AZURE_SAS_TOKEN = 'sv=2024-11-04&ss=bfqt&srt=sco&sp=rwdlacupiytfx&se=2025-10-22T18:20:22Z&st=2025-10-22T10:05:22Z&spr=https&sig=%2FZqM2TSKB42aMjS5nnuMUnQ4ZlXAAMjFbdecscIJ988%3D')
  FILE_FORMAT = azure_csv_format;

LIST @azure_sales_stage;

SHOW STAGES LIKE 'azure_sales_stage';
```

	name	size	md5	last_modified
1	azure://storagesaccounts1.blob.c	3004	a38163e0b0e75a2fc23a050902	Wed, 22 Oct 2025 09:30:23 GMT

name

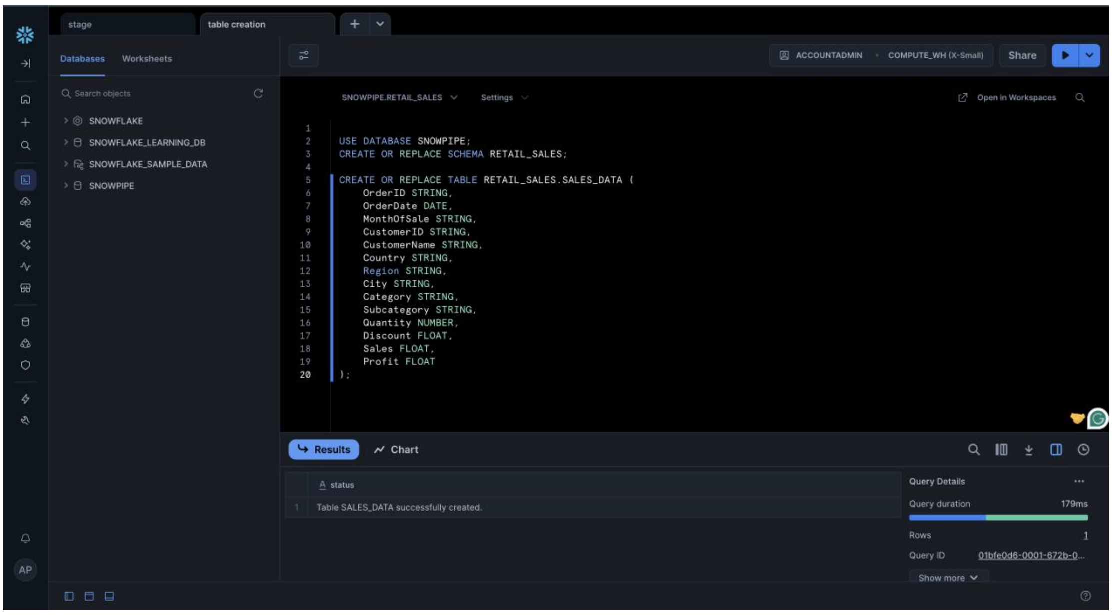
```
azure://storagesaccounts1.bl
ob.core.windows.net/monthly-
sales/Retail_Sales__500_rows
__Preview.csv
```

Step 4: Table creation in snowlake:

```
          USE DATABASE SNOWPIPE;
          USE SCHEMA RETAIL_SALES;

          CREATE OR REPLACE TABLE
RETAIL_SALES.SALES_ORDERS (
             OrderID STRING,
             OrderDate DATE,
             MonthOfSale STRING,
             CustomerID STRING,
             CustomerName STRING,
             Country STRING,
             Region STRING,
             City STRING,
             Category STRING,
             Subcategory STRING,
             Quantity NUMBER,
             Discount FLOAT,
             Sales FLOAT,
             Profit FLOAT);
```

Step 4.1: Data Loading:

```
COPY INTO RETAIL_SALES.SALES_ORDERS
FROM @azure_sales_stage
FILE_FORMAT = (FORMAT_NAME = azure_csv_format)
ON_ERROR = 'CONTINUE';
```



```
USE DATABASE SNOWPIPE;
CREATE OR REPLACE SCHEMA RETAIL_SALES;

CREATE OR REPLACE TABLE RETAIL_SALES.SALES_DATA (
    OrderID STRING,
    OrderDate DATE,
    MonthOfSale STRING,
    CustomerID STRING,
    CustomerName STRING,
    Country STRING,
    Region STRING,
    City STRING,
    Category STRING,
    Subcategory STRING,
    Quantity NUMBER,
    Discount FLOAT,
    Sales FLOAT,
    Profit FLOAT
);
```

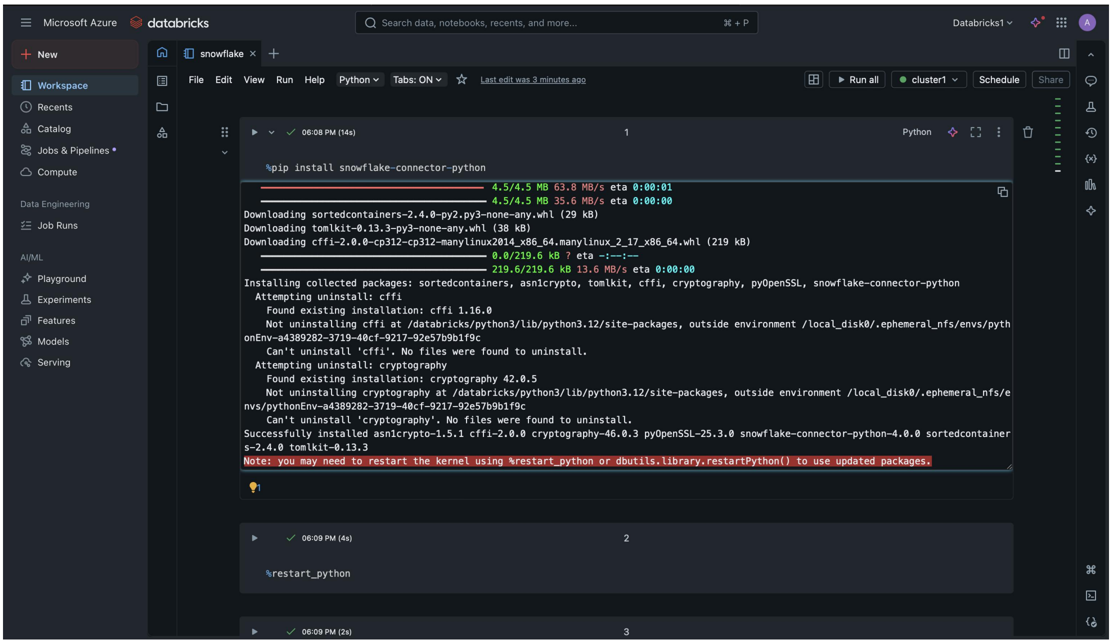
	status
1	Table SALES_DATA successfully created.

Query duration 179ms

Rows 1

Query ID 01bfe0d6-0001-672b-0...

Step 5: Snowpark - Data Transformation:

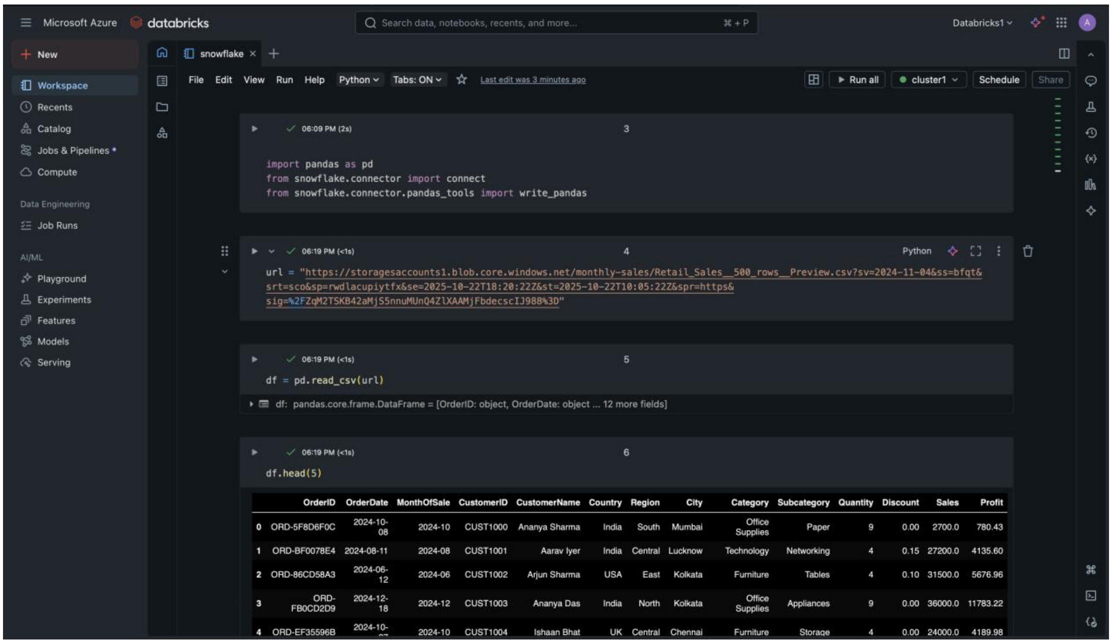


```
%pip install snowflake-connector-python
```

```
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 4.5/4.5 MB 63.8 MB/s eta 0:00:01
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 4.5/4.5 MB 35.6 MB/s eta 0:00:00
Downloading sortedcontainers-2.4.0-py2.py3-none-any.whl (29 kB)
Downloading tomlkit-0.13.3-py3-none-any.whl (38 kB)
Downloading cffi-2.0.0-cp312-cp312-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (219 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 0.0/219.6 kB ? eta -:--:--
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 219.6/219.6 kB 13.6 MB/s eta 0:00:00
Installing collected packages: sortedcontainers, asn1crypto, tomlkit, cffi, cryptography, pyOpenSSL, snowflake-connector-python
  Attempting uninstall: cffi
    Found existing installation: cffi 1.16.0
    Not uninstalling cffi at /databricks/python3/lib/python3.12/site-packages, outside environment /local_disk0/.ephemeral_nfs/envs/pyth
onEnv-a4389282-3719-40cf-9217-92e57b9b1f9c
    Can't uninstall 'cffi'. No files were found to uninstall.
  Attempting uninstall: cryptography
    Found existing installation: cryptography 42.0.5
    Not uninstalling cryptography at /databricks/python3/lib/python3.12/site-packages, outside environment /local_disk0/.ephemeral_nfs/e
nvs/pythonEnv-a4389282-3719-40cf-9217-92e57b9b1f9c
    Can't uninstall 'cryptography'. No files were found to uninstall.
Successfully installed asn1crypto-1.5.1 cffi-2.0.0 cryptography-46.0.3 pyOpenSSL-25.3.0 snowflake-connector-python-4.0.0 sortedcontainer
s-2.4.0 tomlkit-0.13.3
Note: you may need to restart the kernel using %restart_python or dbutils.library.restartPython() to use updated packages.
```

```
%restart_python
```

Step 6: Connect Storage account with databricks:



06:09 PM (2s) 3

```
import pandas as pd
from snowflake.connector import connect
from snowflake.connector.pandas_tools import write_pandas
```

06:19 PM (<1s) 4 Python

```
url = "https://storagesaccounts1.blob.core.windows.net/monthly-sales/Retail_Sales__500_rows__Preview.csv?sv=2024-11-04&ss=bfqt&
srt=sco&sp=rwdlacupiytfx&se=2025-10-22T18:20:22Z&st=2025-10-22T10:05:22Z&spr=https&
sig=%2FZqM2TSKB42aMjS5nnuMUnQ4ZlXAAMjFbdecscIJ988%3D"
```

06:19 PM (<1s) 5

```
df = pd.read_csv(url)
```

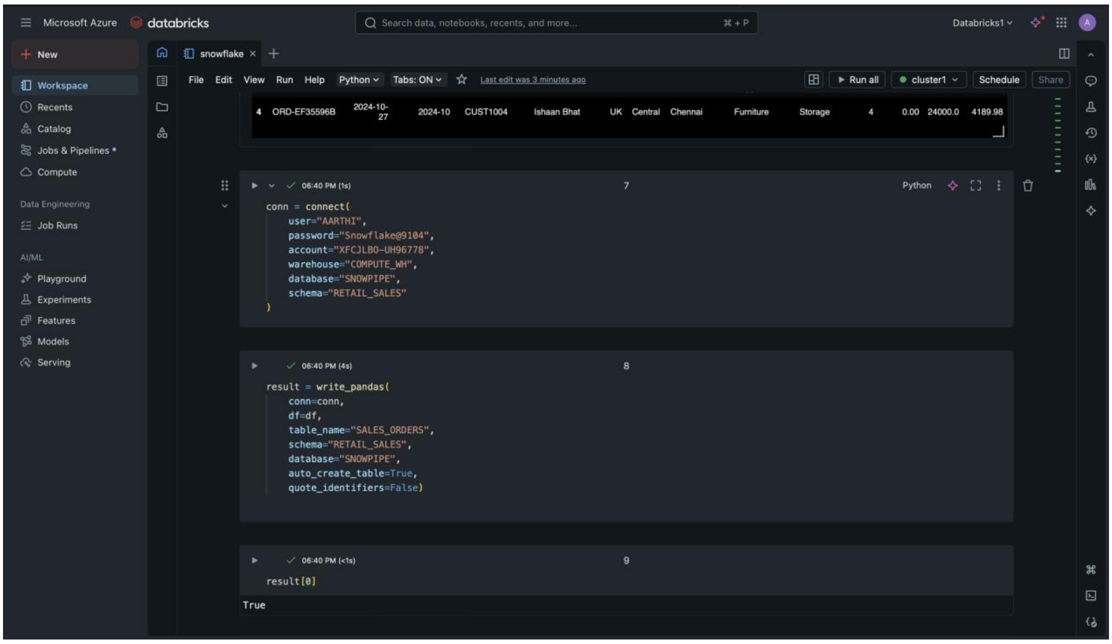
df: pandas.core.frame.DataFrame = [OrderID: object, OrderDate: object ... 12 more fields]

06:19 PM (<1s) 6

```
df.head(5)
```

	OrderID	OrderDate	MonthOfSale	CustomerID	CustomerName	Country	Region	City	Category	Subcategory	Quantity	Discount	Sales	Profit
0	ORD-5F8D6F0C	2024-10-08	2024-10	CUST1000	Ananya Sharma	India	South	Mumbai	Office Supplies	Paper	9	0.00	2700.0	780.43
1	ORD-BF0078E4	2024-08-11	2024-08	CUST1001	Aarav Iyer	India	Central	Lucknow	Technology	Networking	4	0.15	27200.0	4135.60
2	ORD-86CD58A3	2024-06-12	2024-06	CUST1002	Arjun Sharma	USA	East	Kolkata	Furniture	Tables	4	0.10	31500.0	5676.96
3	ORD-FB0CD2D9	2024-12-18	2024-12	CUST1003	Ananya Das	India	North	Kolkata	Office Supplies	Appliances	9	0.00	36000.0	11783.22
4	ORD-EF35596B	2024-10-	2024-10	CUST1004	Ishaan Bhat	UK	Central	Chennai	Furniture	Storage	4	0.00	24000.0	4189.98

Step 7: Snowflake connection:



4	ORD-EF35596B	2024-10-27	2024-10	CUST1004	Ishaan Bhat	UK	Central	Chennai	Furniture	Storage	4	0.00	24000.0	4189.98

06:40 PM (1s) 7 Python

```
conn = connect(
    user="AARTHI",
    password="Snowflake@9104",
    account="XFCJLBO-UH96778",
    warehouse="COMPUTE_WH",
    database="SNOWPIPE",
    schema="RETAIL_SALES"
)
```

06:40 PM (4s) 8

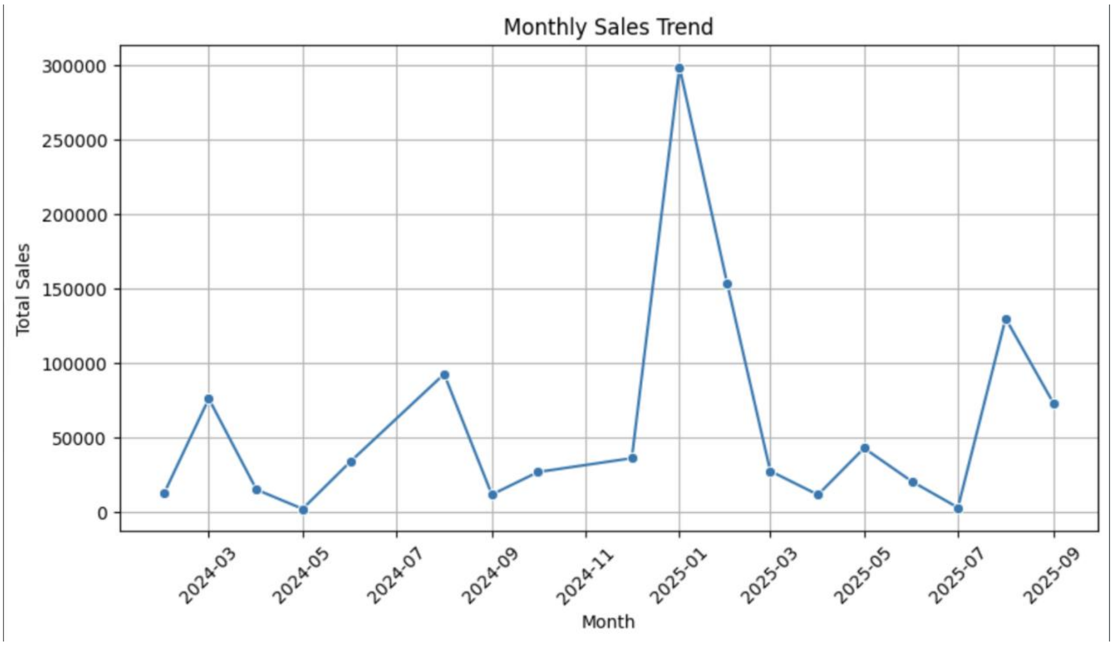
```
result = write_pandas(
    conn=conn,
    df=df,
    table_name="SALES_ORDERS",
    schema="RETAIL_SALES",
    database="SNOWPIPE",
    auto_create_table=True,
    quote_identifiers=False)
```

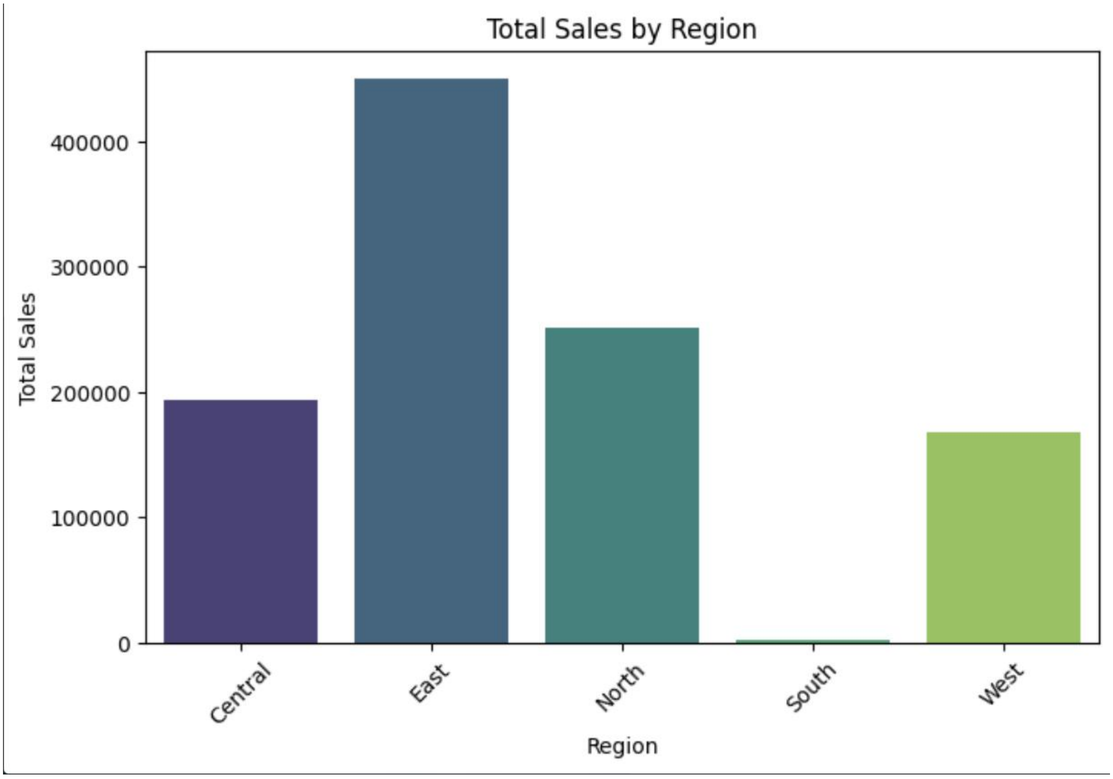
06:40 PM (<1s) 9

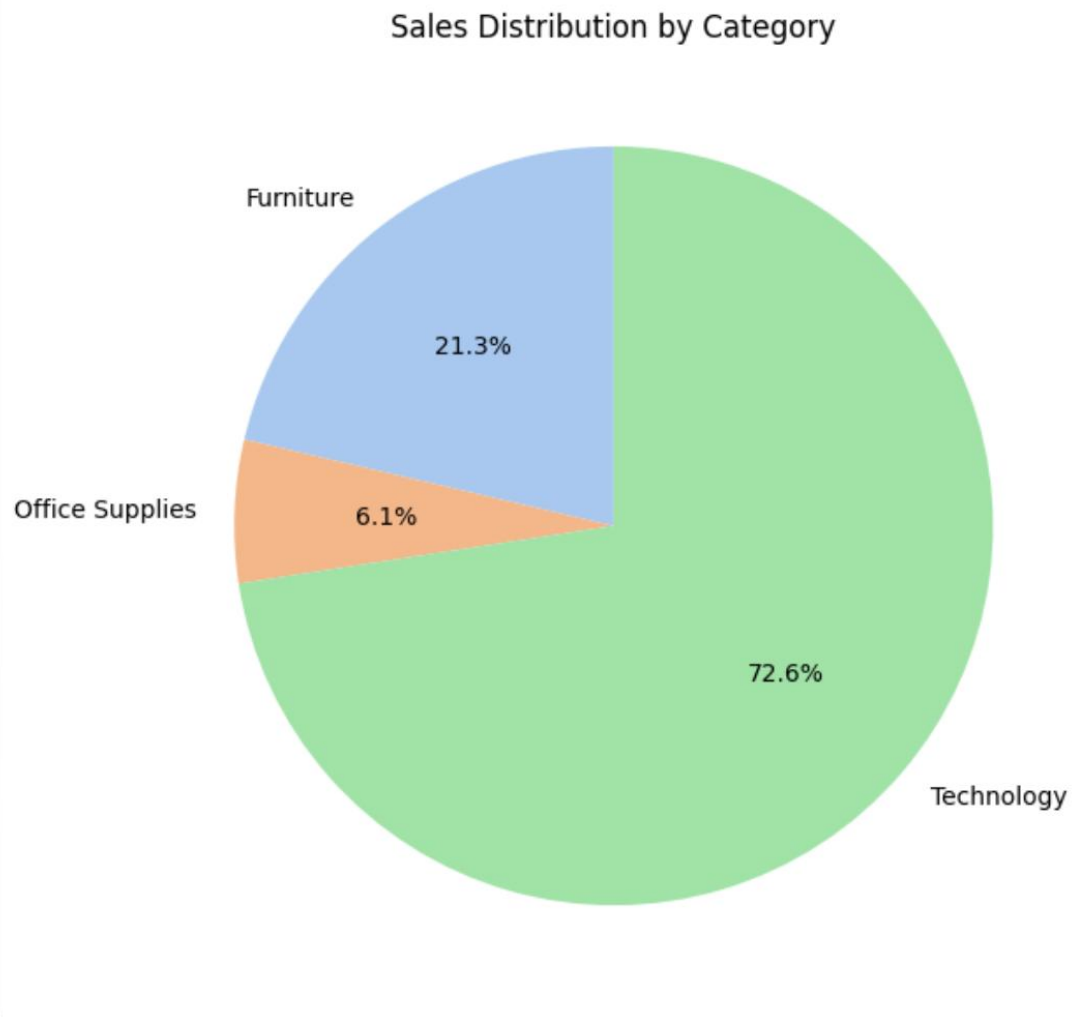
```
result[0]
```

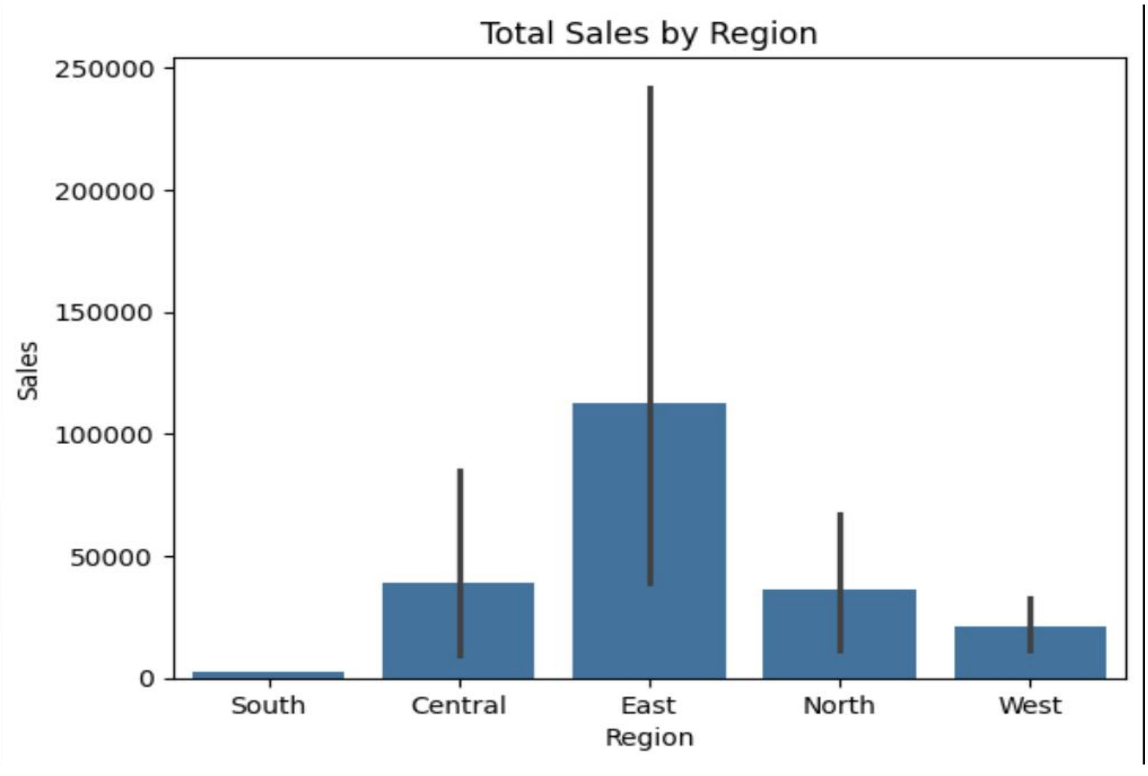
```
True
```

Step 8: Visualization:


Monthly Sales Trend
Total Sales
300000
250000
200000
150000
100000
50000
0
2024-03
2024-05
2024-07
2024-09
2024-11
2025-01
2025-03
2025-05
2025-07
2025-09
Month


Total Sales by Region
Total Sales
400000
300000
200000
100000
0
Central
East
North
South
West
Region


Sales Distribution by Category
Furniture
21.3%
Office Supplies
6.1%
72.6%
Technology


Total Sales by Region
250000
200000
150000
100000
50000
0
Sales
South
Central
East
North
West
Region