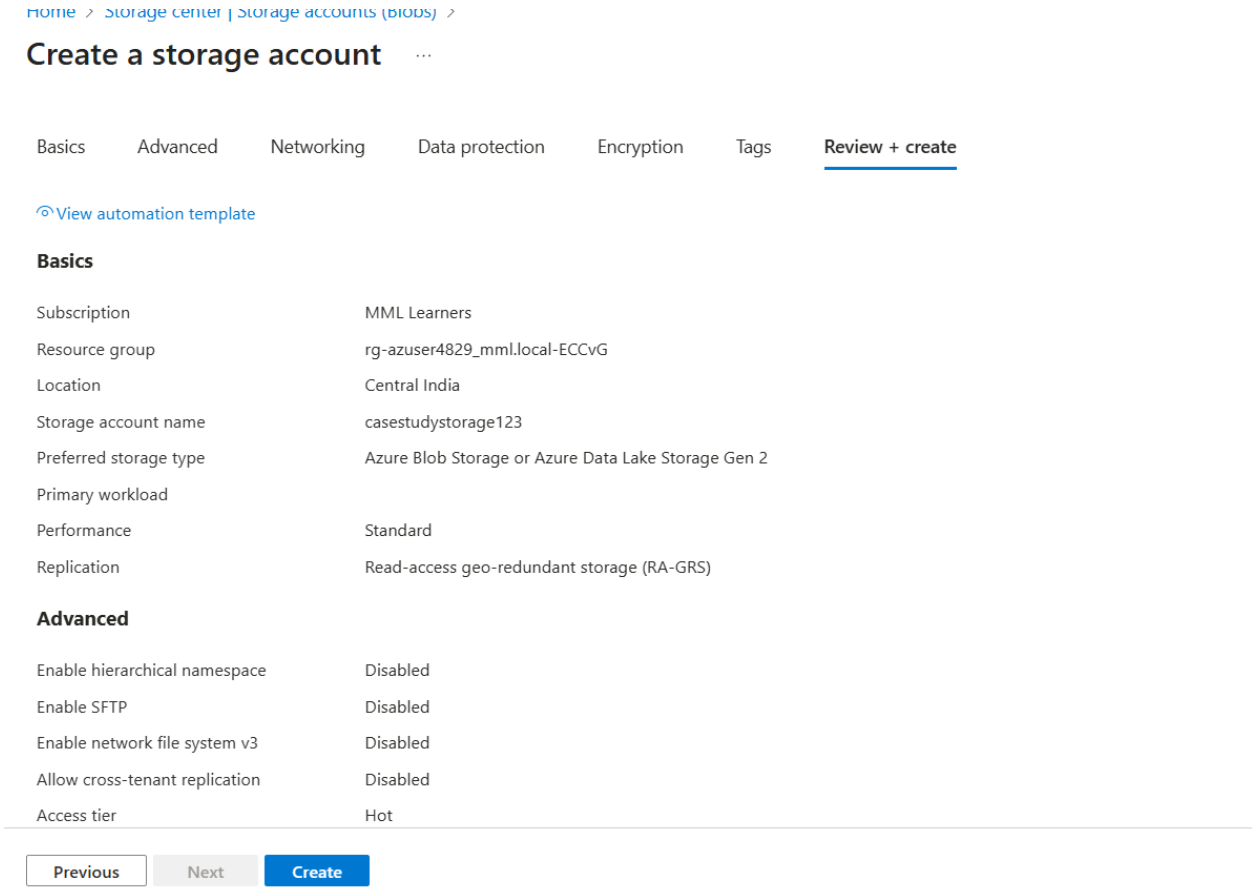
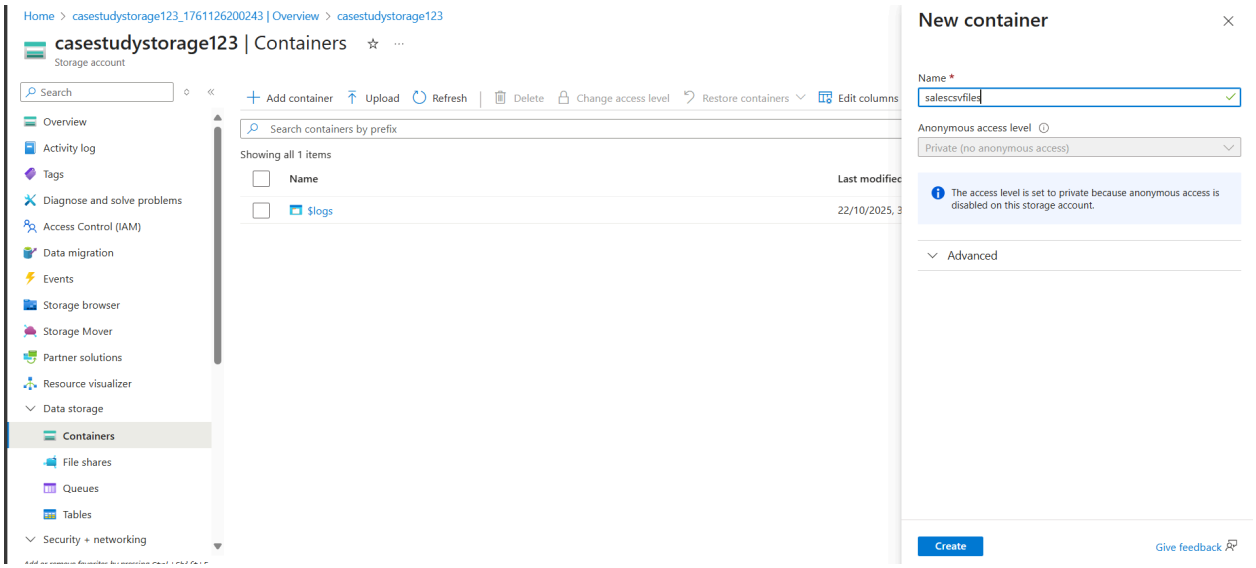
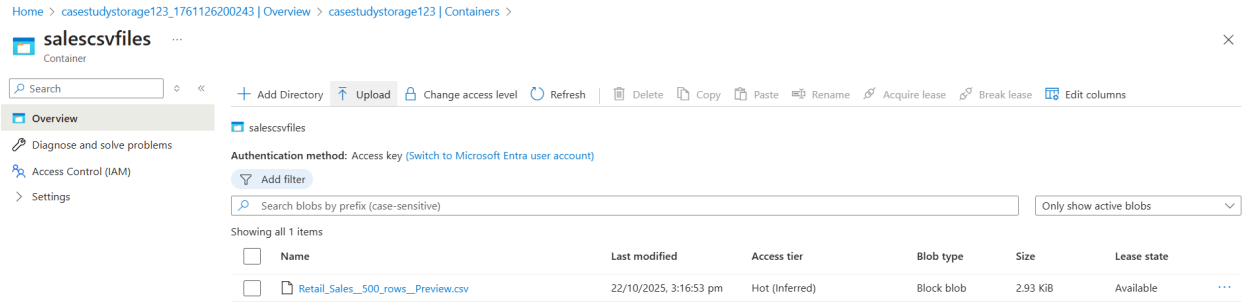
CREATING A STORAGE ACCOUNT



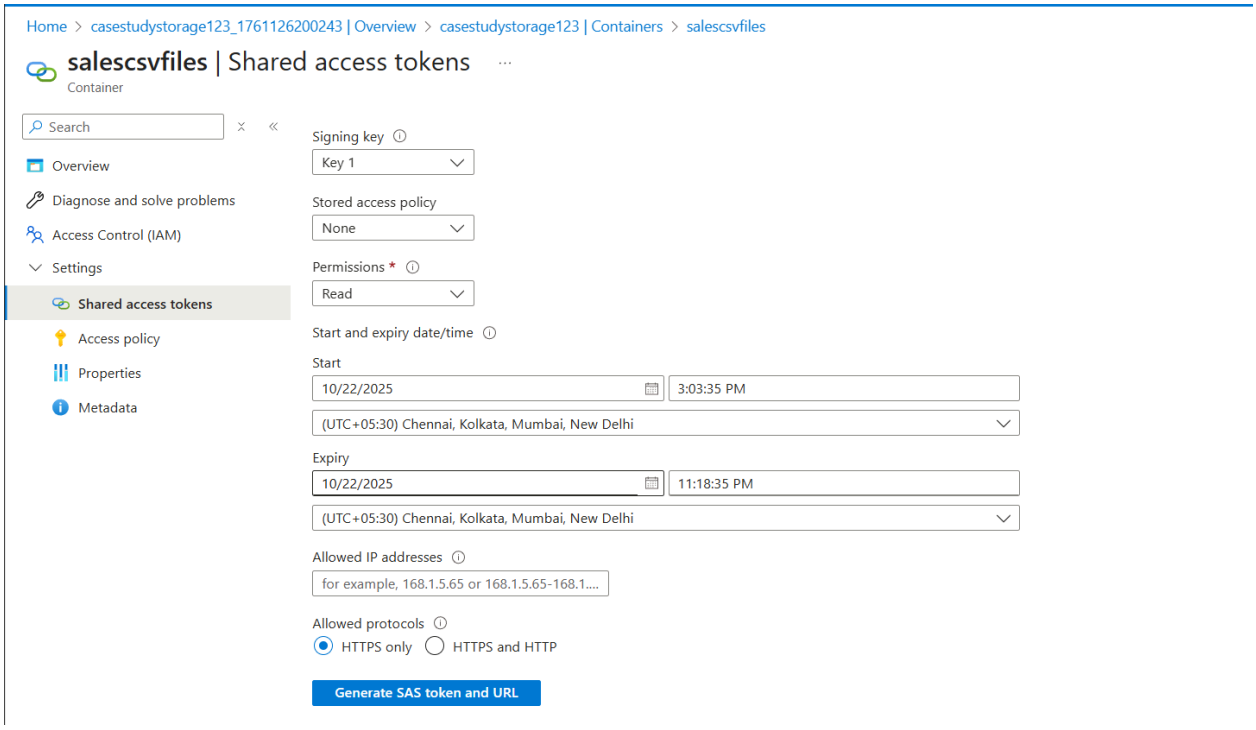
CREATING A NEW CONTAINER



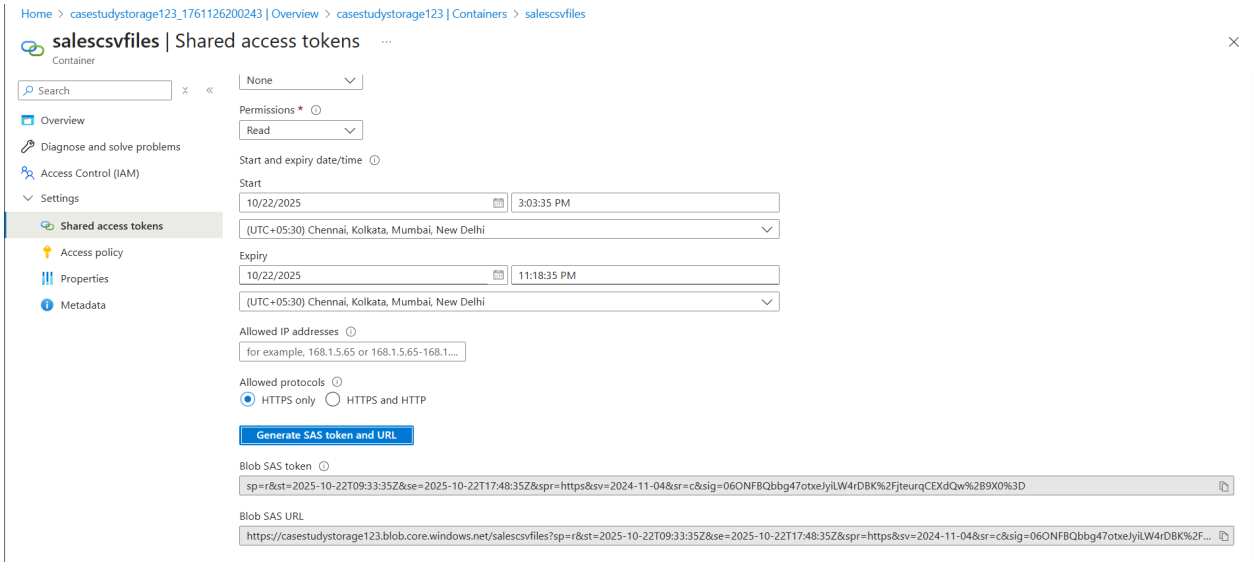
UPLOADING THE CSV FILE



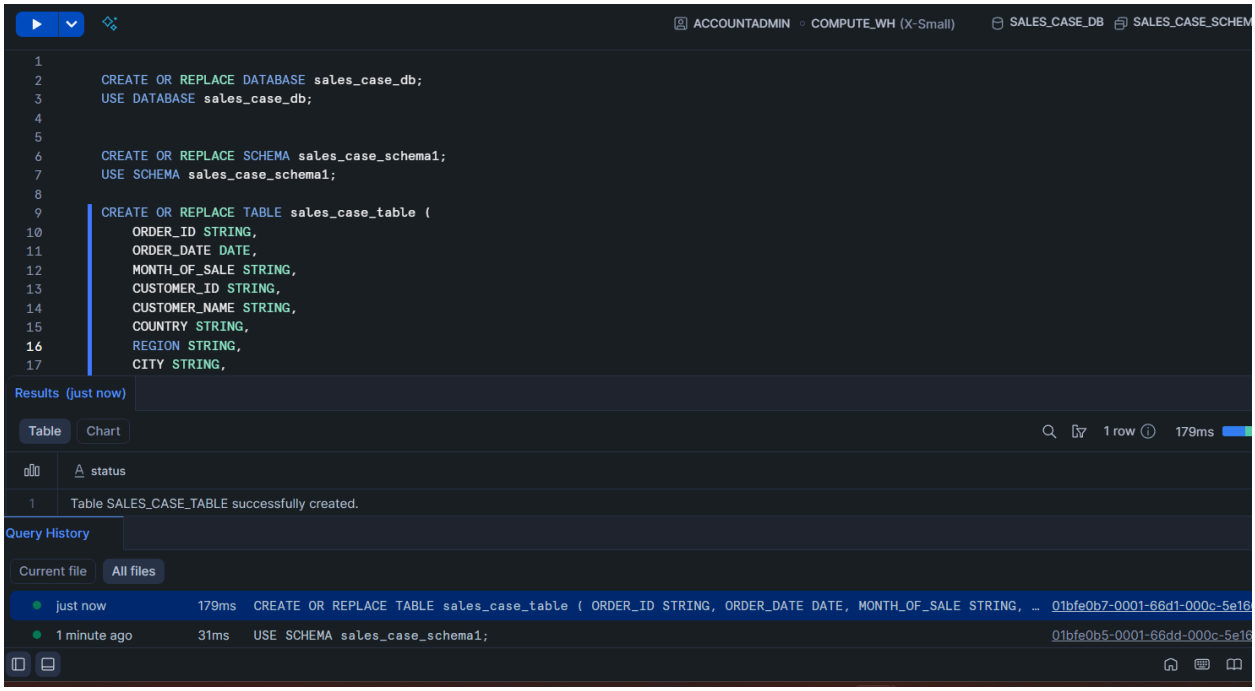
ACCESS TOKENS



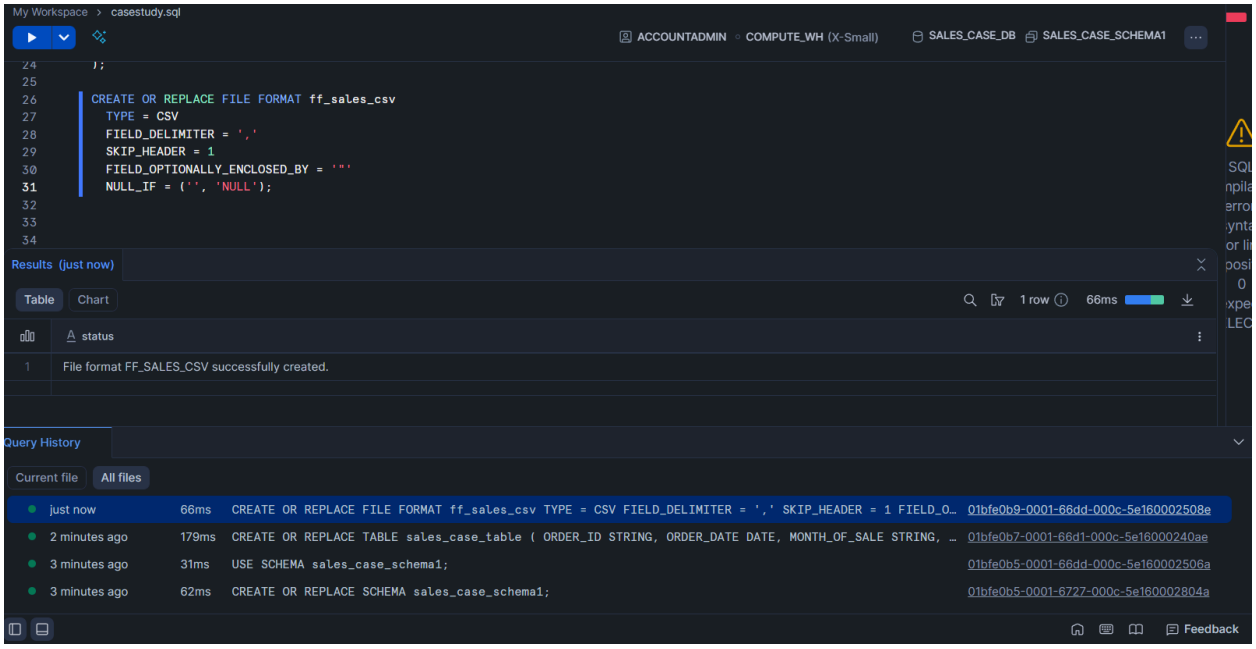
GENERATING SAS TOKENS



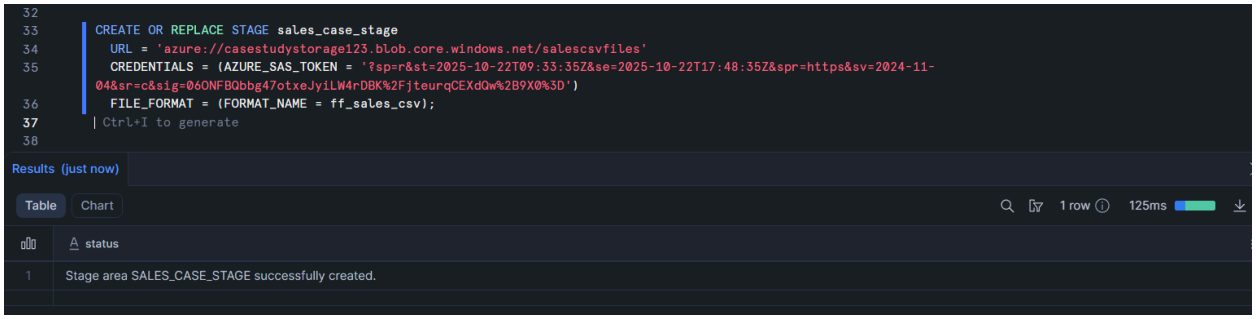
CREATING DATABASE , SCHEMA AND TABLES



FILE FORMAT

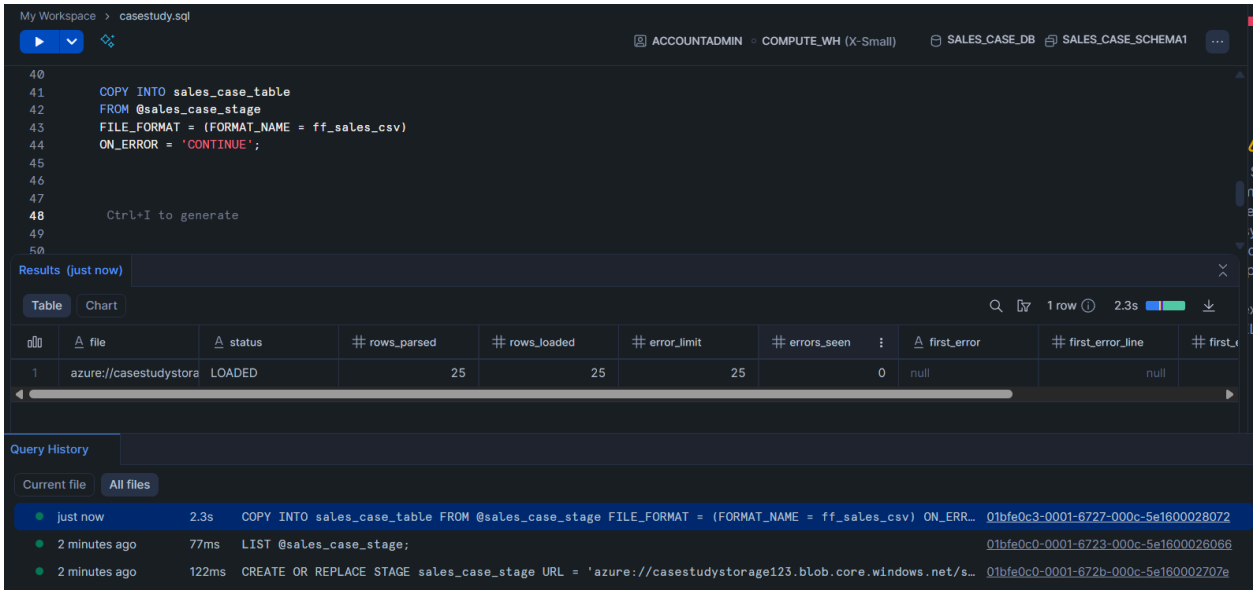


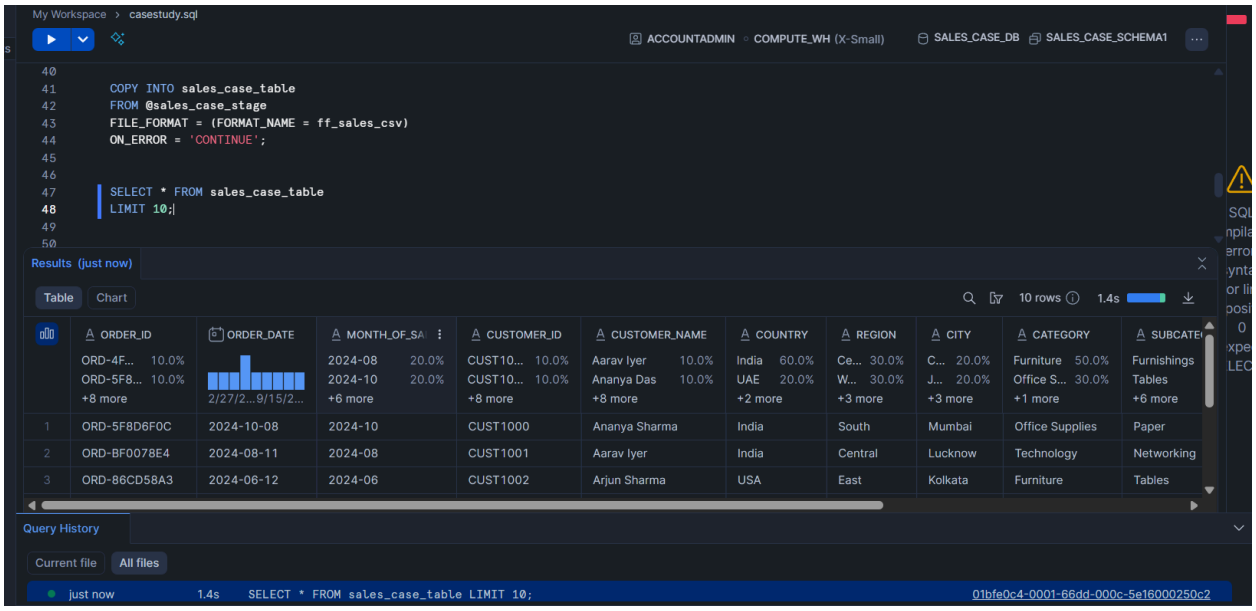
STAGING



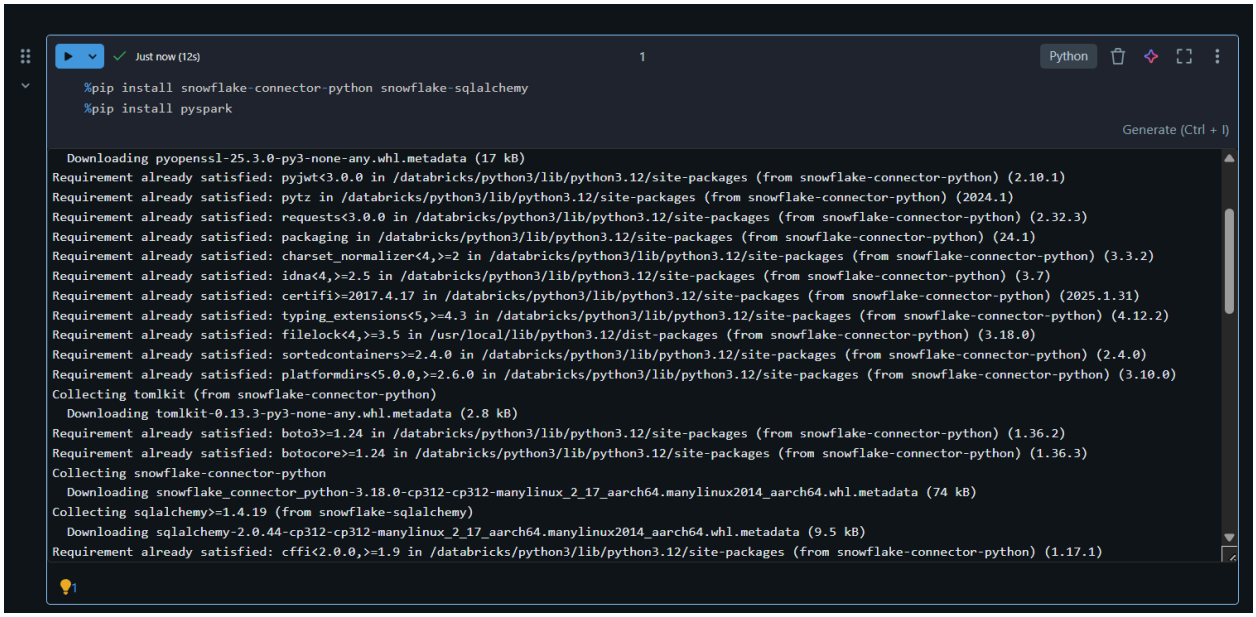
CHECKING STAGE CONTENTS

COPYING DATA FROM STAGE TO SNOWFLAKE TABLE


My Workspace › casestudy.sql
ACCOUNTADMIN · COMPUTE_WH (X-Small)
SALES_CASE_DB
SALES_CASE_SCHEMA1
COPY INTO sales_case_table
FROM @sales_case_stage
FILE_FORMAT = (FORMAT_NAME = ff_sales_csv)
ON_ERROR = 'CONTINUE';
Ctrl+I to generate
Results (just now)
Table
Chart
1 row
2.3s
file
status
rows_parsed
rows_loaded
error_limit
errors_seen
first_error
first_error_line
azure://casestudystora
LOADED
25
25
25
0
null
null
Query History
Current file
All files
just now 2.3s COPY INTO sales_case_table FROM @sales_case_stage FILE_FORMAT = (FORMAT_NAME = ff_sales_csv) ON_ERR… 01bfe0c3-0001-6727-000c-5e1600028072
2 minutes ago 77ms LIST @sales_case_stage; 01bfe0c0-0001-6723-000c-5e1600026066
2 minutes ago 122ms CREATE OR REPLACE STAGE sales_case_stage URL = 'azure://casestudystorage123.blob.core.windows.net/s… 01bfe0c0-0001-672b-000c-5e160002707e


My Workspace › casestudy.sql
ACCOUNTADMIN · COMPUTE_WH (X-Small)
SALES_CASE_DB
SALES_CASE_SCHEMA1
COPY INTO sales_case_table
FROM @sales_case_stage
FILE_FORMAT = (FORMAT_NAME = ff_sales_csv)
ON_ERROR = 'CONTINUE';
SELECT * FROM sales_case_table
LIMIT 10;
Results (just now)
Table
Chart
10 rows
1.4s
ORDER_ID
ORDER_DATE
MONTH_OF_SA
CUSTOMER_ID
CUSTOMER_NAME
COUNTRY
REGION
CITY
CATEGORY
SUBCATE
1 ORD-5F8D6F0C 2024-10-08 2024-10 CUST1000 Ananya Sharma India South Mumbai Office Supplies Paper
2 ORD-BF0078E4 2024-08-11 2024-08 CUST1001 Aarav Iyer India Central Lucknow Technology Networking
3 ORD-86CD58A3 2024-06-12 2024-06 CUST1002 Arjun Sharma USA East Kolkata Furniture Tables
Query History
Current file
All files
just now 1.4s SELECT * FROM sales_case_table LIMIT 10; 01bfe0c4-0001-66dd-000c-5e16000250c2

DATABRICKS PART



Just now (12s) 1 Python

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

Generate (Ctrl + I)

```
  Downloading pyopenssl-25.3.0-py3-none-any.whl.metadata (17 kB)
Requirement already satisfied: pyjwt<3.0.0 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (2.10.1)
Requirement already satisfied: pytz in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (2024.1)
Requirement already satisfied: requests<3.0.0 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (2.32.3)
Requirement already satisfied: packaging in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (24.1)
Requirement already satisfied: charset_normalizer<4,>=2 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (3.7)
Requirement already satisfied: certifi>=2017.4.17 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (2025.1.31)
Requirement already satisfied: typing_extensions<5,>=4.3 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (4.12.2)
Requirement already satisfied: filelock<4,>=3.5 in /usr/local/lib/python3.12/dist-packages (from snowflake-connector-python) (3.18.0)
Requirement already satisfied: sortedcontainers>=2.4.0 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (2.4.0)
Requirement already satisfied: platformdirs<5.0.0,>=2.6.0 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (3.10.0)
Collecting tomlkit (from snowflake-connector-python)
  Downloading tomlkit-0.13.3-py3-none-any.whl.metadata (2.8 kB)
Requirement already satisfied: boto3>=1.24 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (1.36.2)
Requirement already satisfied: botocore>=1.24 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (1.36.3)
Collecting snowflake-connector-python
  Downloading snowflake_connector_python-3.18.0-cp312-cp312-manylinux_2_17_aarch64.manylinux2014_aarch64.whl.metadata (74 kB)
Collecting sqlalchemy>=1.4.19 (from snowflake-sqlalchemy)
  Downloading sqlalchemy-2.0.44-cp312-cp312-manylinux_2_17_aarch64.manylinux2014_aarch64.whl.metadata (9.5 kB)
Requirement already satisfied: cffi<2.0.0,>=1.9 in /databricks/python3/lib/python3.12/site-packages (from snowflake-connector-python) (1.17.1)
```

Just now (3s) 2 Python

```
Options = {
  "sfURL": "gqggyqc-jp73258.snowflakecomputing.com",
  "sfDatabase": "SALES_CASE_DB",
  "sfSchema": "SALES_CASE_SCHEMA1",
  "sfWarehouse": "COMPUTE_WH",
  "sfRole": "ACCOUNTADMIN",
  "sfUser": "Deepta",
  "sfPassword": "#Deeptasermaraj2468"
}

df = (spark.read
      .format("snowflake")
      .options(**Options)
      .option("dbtable", "sales_case_table")
      .load())
```

df: pyspark.sql.connect.dataframe.DataFrame = [ORDER_ID: string, ORDER_DATE: date ... 12 more fields]

```python
df = (
    spark.read
    .format("snowflake")
    .options(**sfOptions)
    .option("dbtable", "sales_case_table")
    .load()
)


df.show(5)
```

See performance (1) Optimize

df: pyspark.sql.connect.dataframe.DataFrame = [ORDER_ID: string, ORDER_DATE: date ... 12 more fields]

ORDER_ID	ORDER_DATE	MONTH_OF_SALE	CUSTOMER_ID	CUSTOMER_NAME	COUNTRY	REGION	CITY	CATEGORY	SUBCATEGORY	QUANTITY	DISCOUNT	SALES	PROFIT
ORD-5F8D6F0C	2024-10-08	2024-10	CUST1000	Ananya Sharma	India	South	Mumbai	Office Supplies	Paper	9	0.0	2700.0	780.43
ORD-BF0078E4	2024-08-11	2024-08	CUST1001	Aarav Iyer	India	Central	Lucknow	Technology	Networking	4	0.15	27200.0	4135.6
ORD-86CD58A3	2024-06-12	2024-06	CUST1002	Arjun Sharma	USA	East	Kolkata	Furniture	Tables	4	0.1	31500.0	5676.96
ORD-FB0CD2D9	2024-12-18	2024-12	CUST1003	Ananya Das	India	North	Kolkata	Office Supplies	Appliances	9	0.0	36000.0	11783.22
ORD-EF35596B	2024-10-27	2024-10	CUST1004	Ishaan Bhat	UK	Central	Chennai	Furniture	Storage	4	0.0	24000.0	4189.98

only showing top 5 rows

BASIC TRANSFORMATIONS IN DATABRICKS

```python
from pyspark.sql.functions import col, round
df_transformed = df.withColumn(
    "Profit_Margin", round(col("Profit") / col("Sales") * 100, 2)
)
df_high_profit = df_transformed.filter(col("Profit_Margin") > 20)
df_high_profit.show(5)
```

See performance (1) Optimize

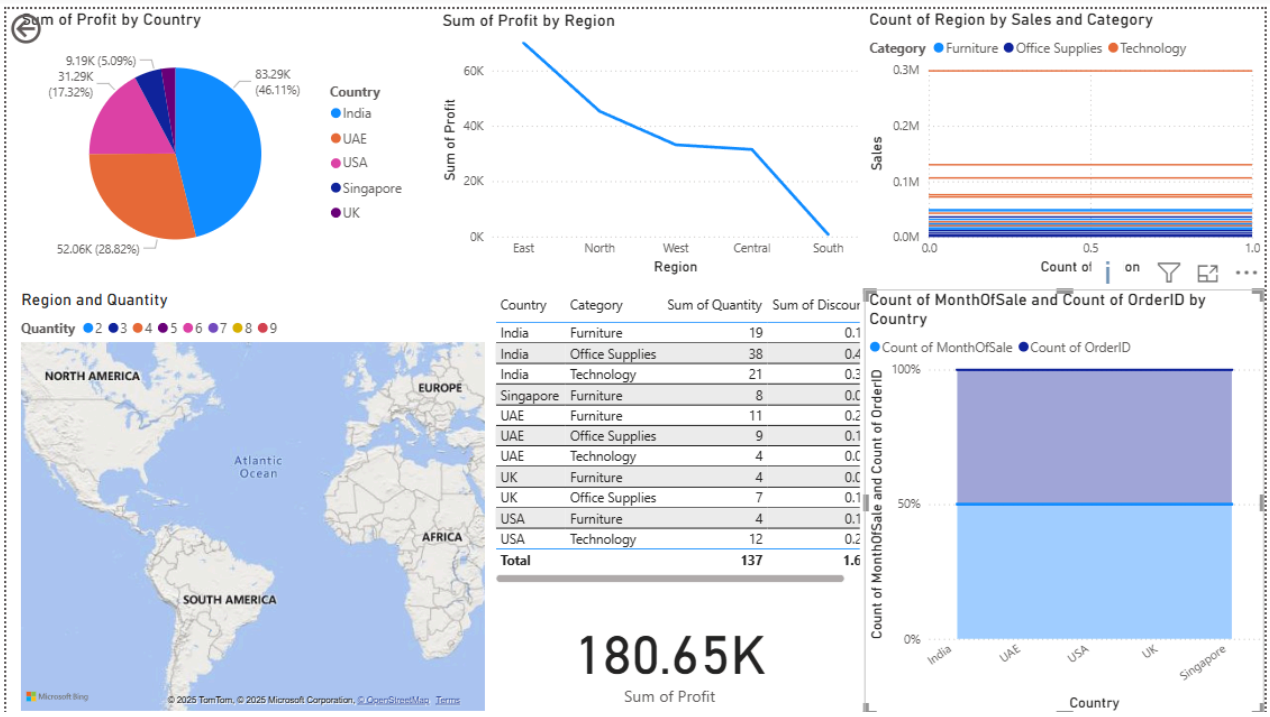
df_high_profit: pyspark.sql.connect.dataframe.DataFrame = [ORDER_ID: string, ORDER_DATE: date ... 13 more fields]

df_transformed: pyspark.sql.connect.dataframe.DataFrame = [ORDER_ID: string, ORDER_DATE: date ... 13 more fields]

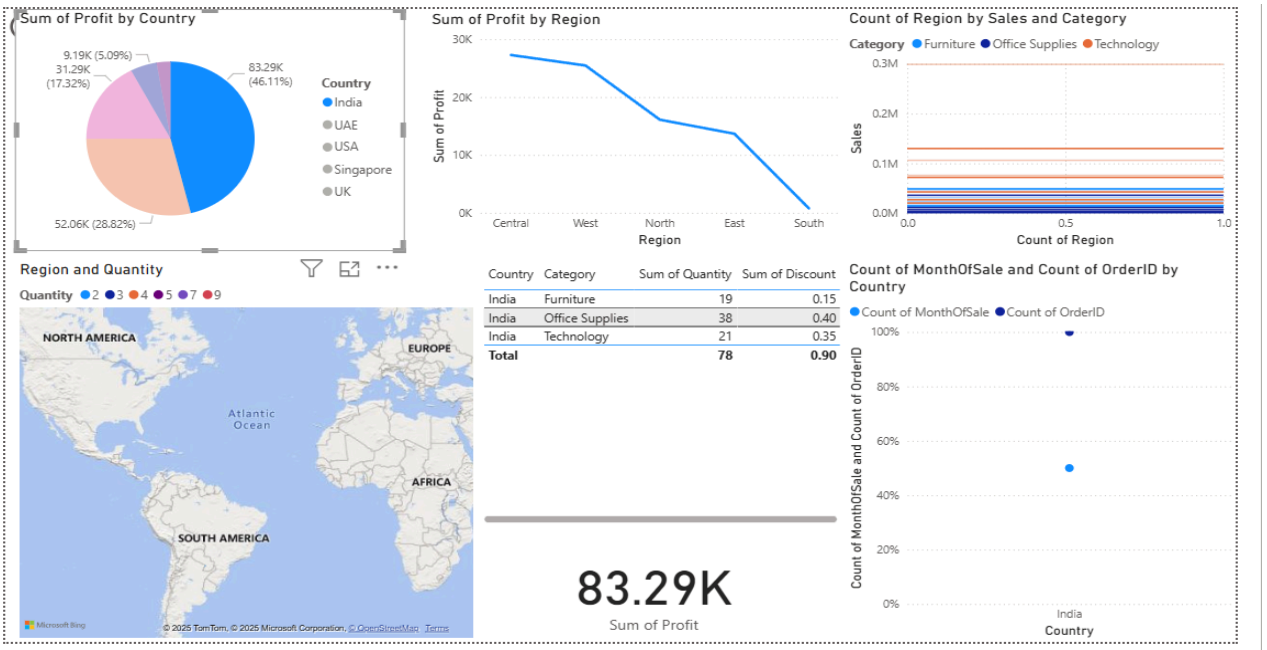
ORDER_ID	ORDER_DATE	MONTH_OF_SALE	CUSTOMER_ID	CUSTOMER_NAME	COUNTRY	REGION	CITY	CATEGORY	SUBCATEGORY	QUANTITY	DISCOUNT	SALES	PROFIT	Profit_Margin
ORD-5F8D6F0C	2024-10-08	2024-10	CUST1000	Ananya Sharma	India	South	Mumbai	Office Supplies	Paper	9	0.0	2700.0	780.43	28.9
ORD-FB0CD2D9	2024-12-18	2024-12	CUST1003	Ananya Das	India	North	Kolkata	Office Supplies	Appliances	9	0.0	36000.0	11783.22	32.73
ORD-60D1DA88	2024-08-26	2024-08	CUST1005	Neha Iyer	UAE	West	Chennai	Furniture	Furnishings	6	0.05	16387.5	3392.51	20.7
ORD-A5081404	2025-09-15	2025-09	CUST1006	Arjun Iyer	India	Central	Jaipur	Office Supplies	Binders	3	0.15	1147.5	307.47	26.79
ORD-4FCB3B05	2025-05-26	2025-05	CUST1008	Kabir Menon	India	West	Jaipur	Technology	Printers	3	0.05	42750.0	8742.33	20.45

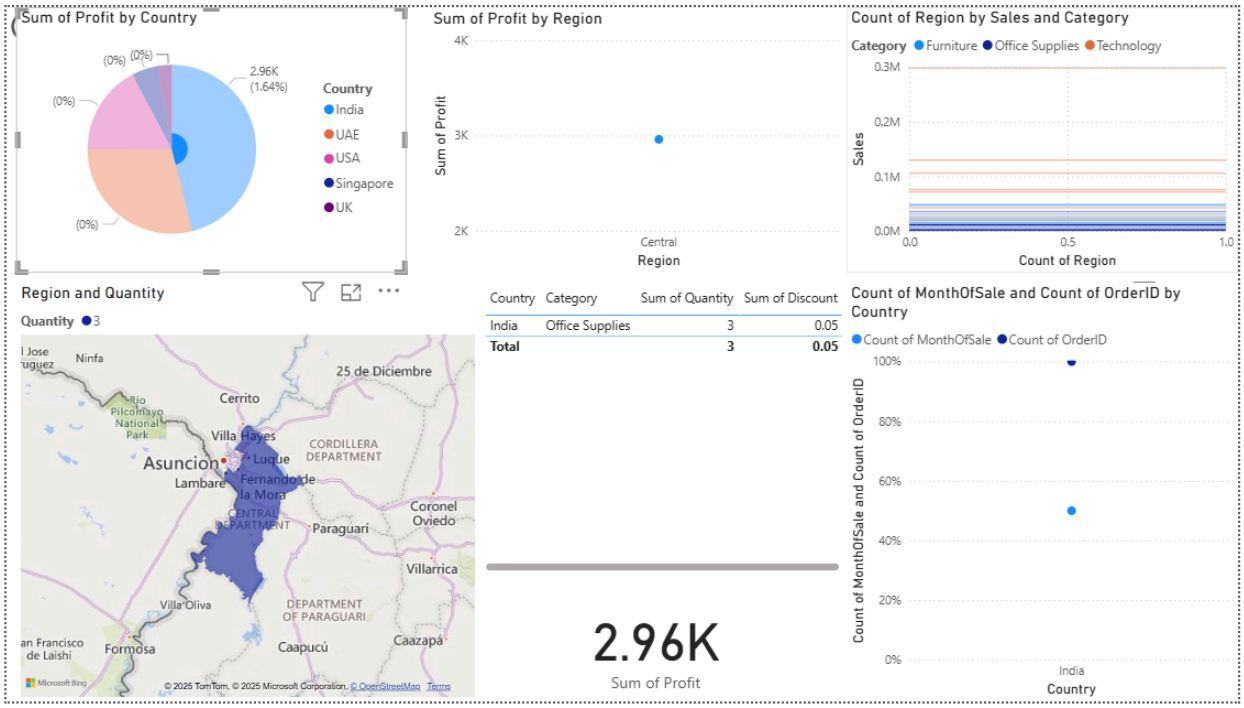
POWERBI

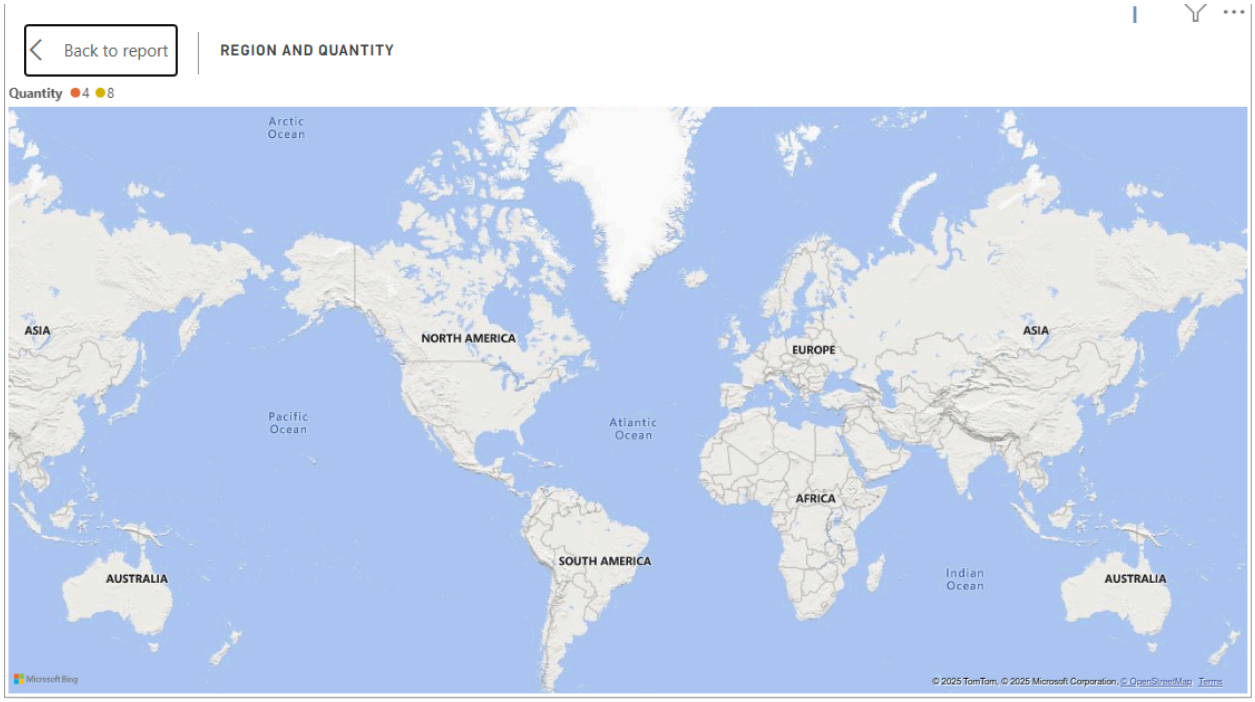
CREATING DASHBOARDS IN POWERBI, EXTRACTING THE DATAS FROM SNOWFLAKE

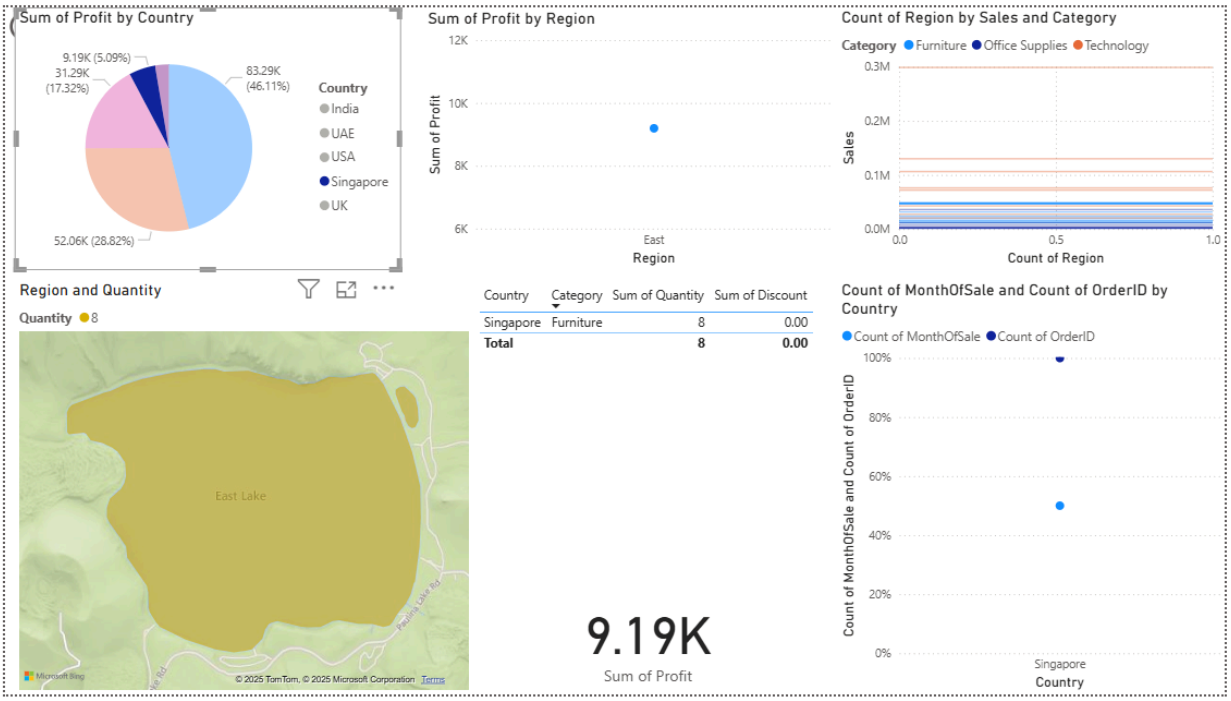


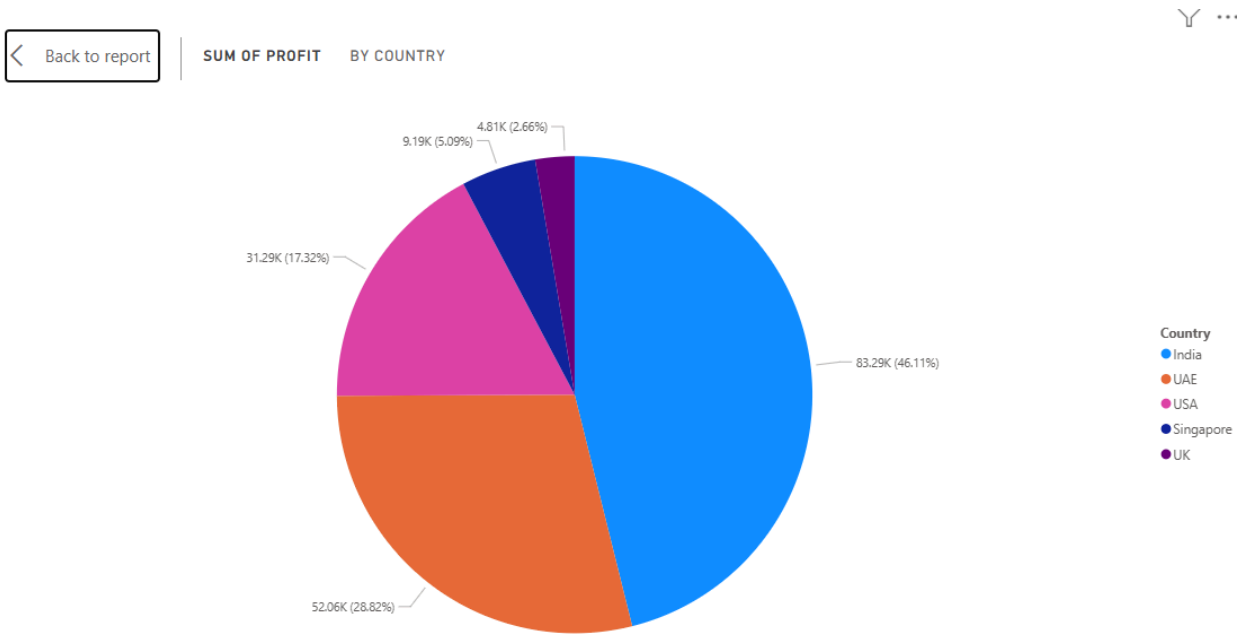
INTERACTING WITH THE DASHBOARDS




Sum of Profit by Country
(0%)
(0%)
(0%)
(0%)
2.96K
(1.64%)
Country
India
UAE
USA
Singapore
UK
Sum of Profit by Region
4K
3K
2K
Sum of Profit
Central
Region
Count of Region by Sales and Category
Category
Furniture
Office Supplies
Technology
0.3M
0.2M
0.1M
0.0M
Sales
0.0
0.5
1.0
Count of Region
Region and Quantity
Quantity
3
Ninfa
25 de Diciembre
Cerrito
Villa Hayes
Asuncion
Luque
CORDILLERA DEPARTMENT
Lambare
Fernando de la Mora
Coronel Oviedo
Paraguari
Villarrica
Villa Oliva
DEPARTMENT OF PARAGUARI
Formosa
Caapucú
Caazapá
© 2025 TomTom, © 2025 Microsoft Corporation, © OpenStreetMap Terms
Country
Category
Sum of Quantity
Sum of Discount
India
Office Supplies
3
0.05
Total
3
0.05
2.96K
Sum of Profit
Count of MonthOfSale and Count of OrderID by Country
Count of MonthOfSale
Count of OrderID
100%
80%
60%
40%
20%
0%
Count of MonthOfSale and Count of OrderID
India
Country


Back to report
REGION AND QUANTITY
Quantity
4
8
Arctic Ocean
ASIA
NORTH AMERICA
EUROPE
ASIA
Pacific Ocean
Atlantic Ocean
AFRICA
SOUTH AMERICA
AUSTRALIA
Indian Ocean
AUSTRALIA
© 2025 TomTom, © 2025 Microsoft Corporation, © OpenStreetMap Terms


Sum of Profit by Country
9.19K (5.09%)
31.29K (17.32%)
83.29K (46.11%)
52.06K (28.82%)
Country
India
UAE
USA
Singapore
UK
Sum of Profit by Region
12K
10K
8K
6K
Sum of Profit
East
Region
Count of Region by Sales and Category
Category Furniture Office Supplies Technology
0.3M
0.2M
0.1M
0.0M
Sales
0.0
0.5
1.0
Count of Region
Region and Quantity
Quantity 8
East Lake
© 2025 TomTom, © 2025 Microsoft Corporation Terms
Country Category Sum of Quantity Sum of Discount
Singapore Furniture 8 0.00
Total 8 0.00
9.19K
Sum of Profit
Count of MonthOfSale and Count of OrderID by Country
Count of MonthOfSale Count of OrderID
100%
80%
60%
40%
20%
0%
Count of MonthOfSale and Count of OrderID
Singapore
Country


Back to report
SUM OF PROFIT
BY COUNTRY
4.81K (2.66%)
9.19K (5.09%)
31.29K (17.32%)
83.29K (46.11%)
52.06K (28.82%)
Country
India
UAE
USA
Singapore
UK