**PROJECT REPORT**

*CAMP BI V3 Batch 1 Group 4*

**INDEX**

1. Project Description
2. Unit Test Report
3. Creation of database
4. Creation of schemas
5. Creation of table as per dataset
6. Creation of Integration Object
7. Creation of external stage for loading the data structure
8. Creation of snowpipe for autoingesting of data from S3 bucket
9. Creation of stream on the given table
10. SCD 2 operation on the consumer table

***Group Members:***

- Aryan Kaundal

- Kajol Gupta

- Saurabh Pratap Singh

- Shivam Gupta

- Shubham Dixit

- Vrishti Saxena

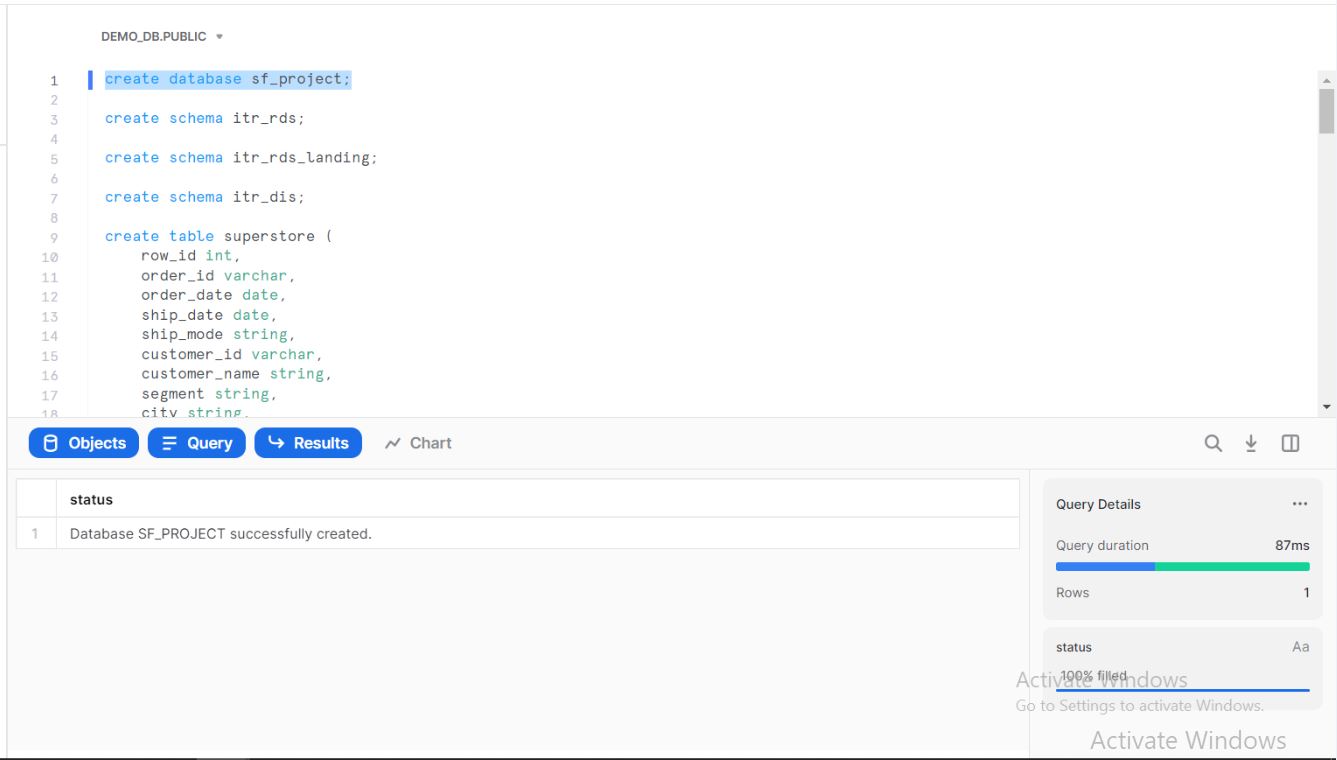
* ***PROJECT DESCRIPTION***

The project involves data ingestion and analysis from public datahub Kaggle [link](https://www.kaggle.com/datasets/vivek468/superstore-dataset-final).

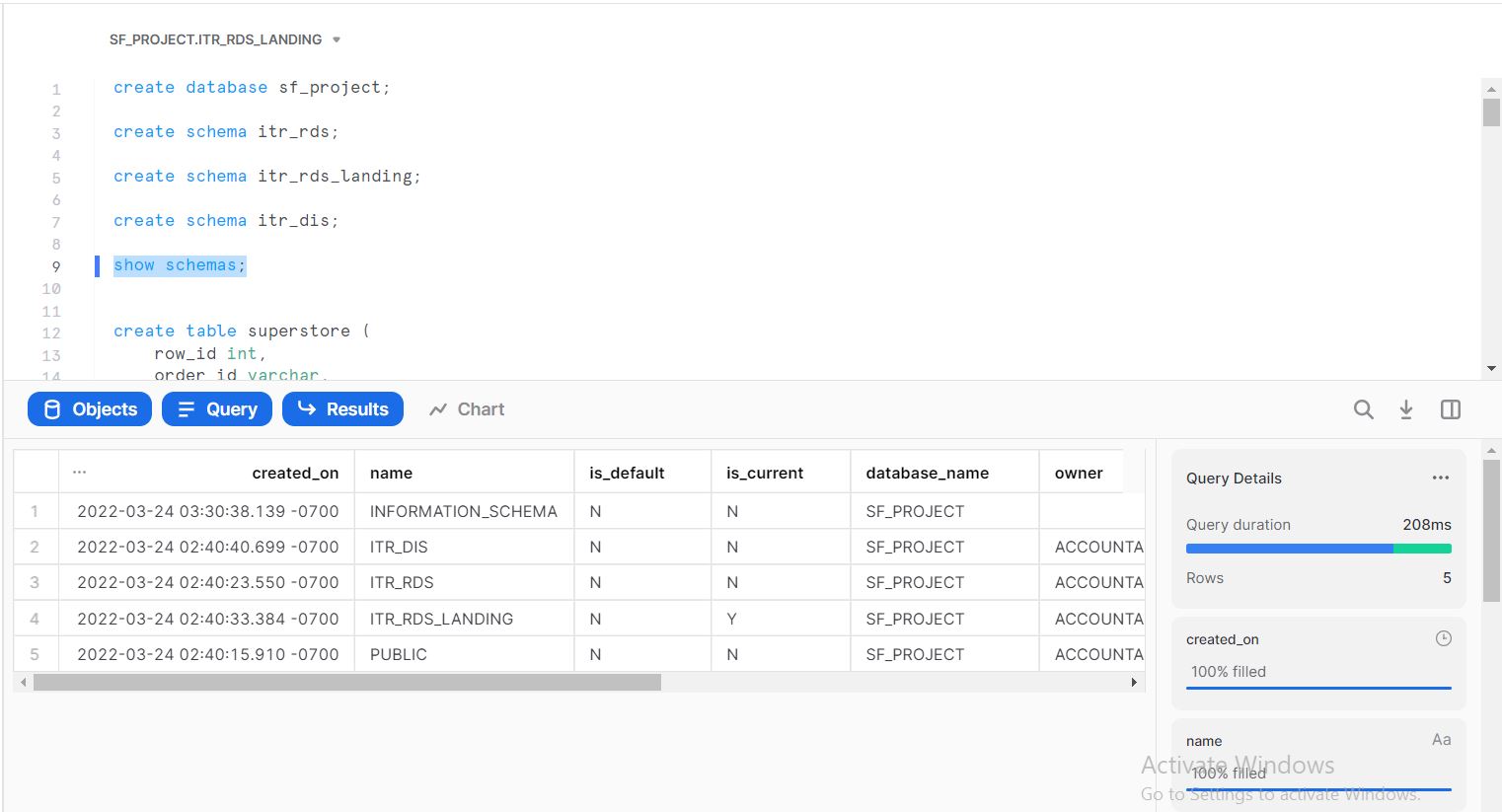
* Steps involved in performing the data ingestion:
  1. Loading data to external stage
  2. Ingesting data into the landing schema
  3. Ingesting data into the consumer table
  4. Perform analysis on the given dataset

***UNIT TEST REPORT***

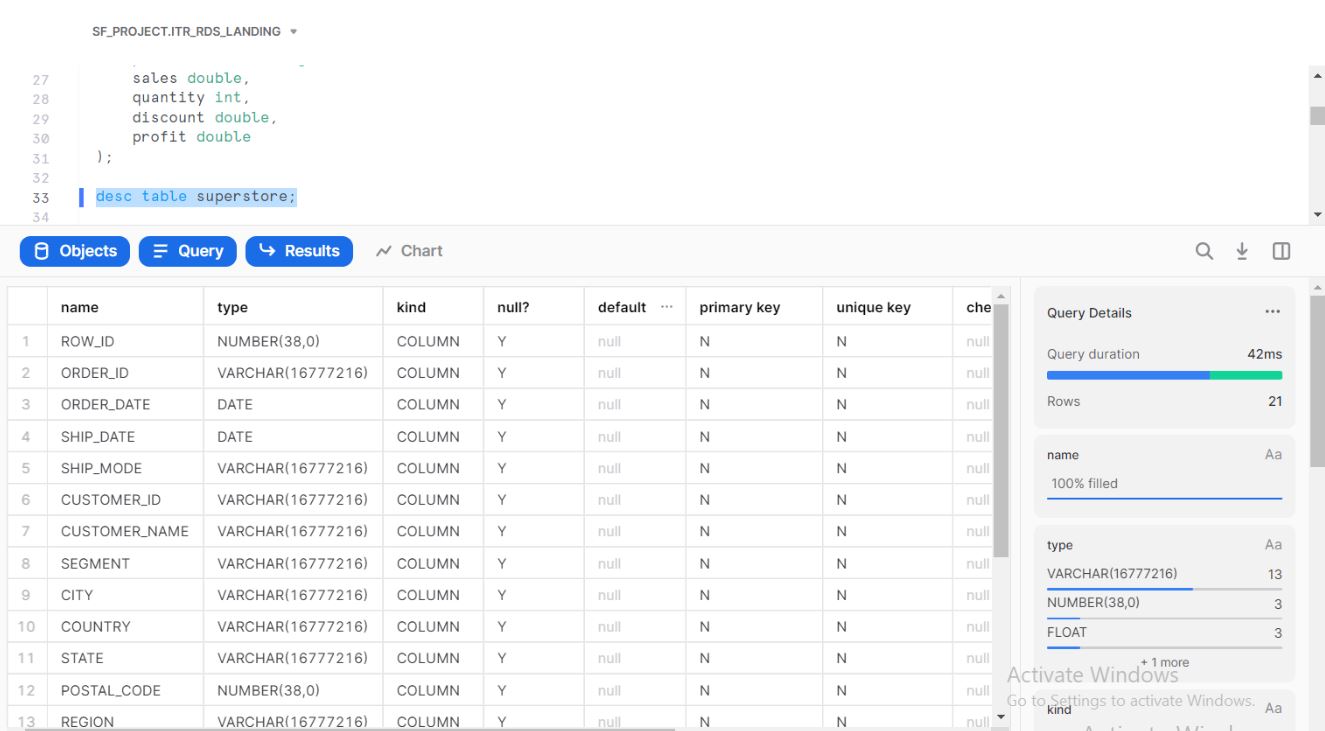
1. Created a database named SF\_PROJECT;



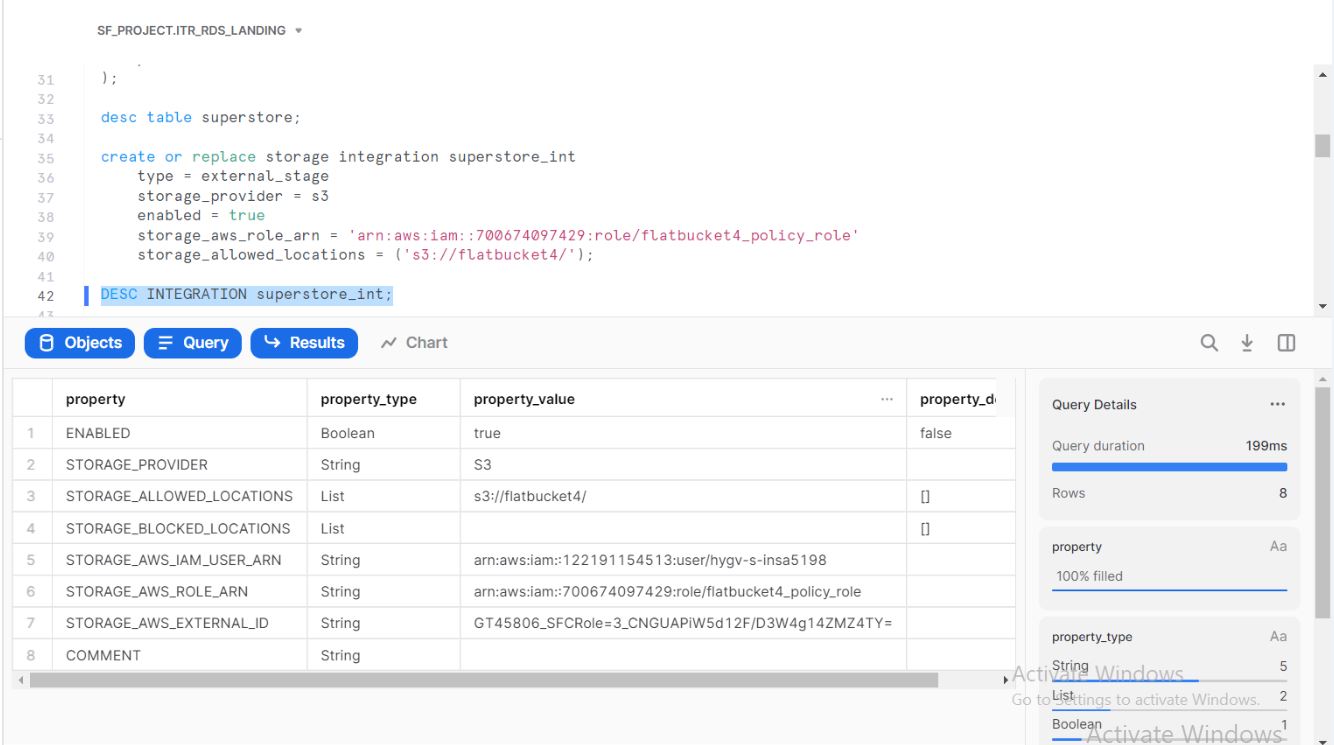
2. Created three schemas named ITR\_RDS, ITR\_RDS\_LANDING and ITR\_DIS



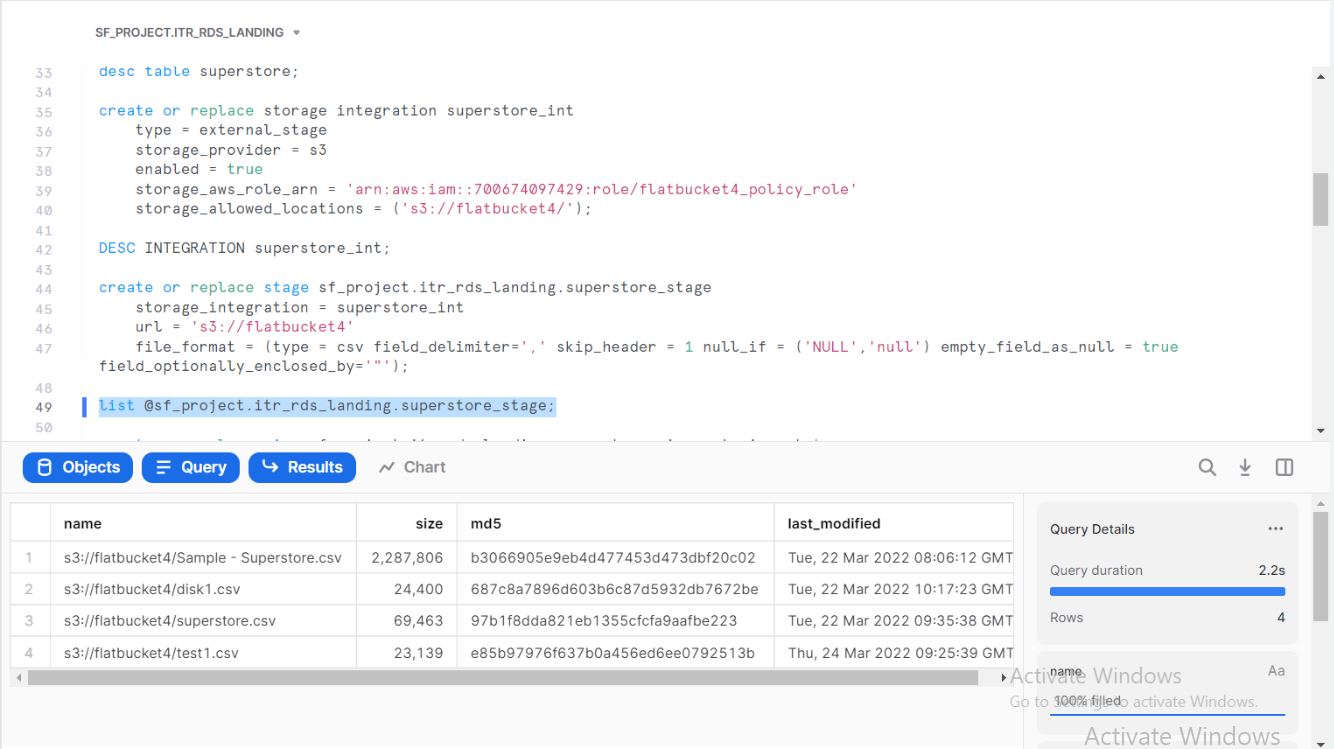
3. Created a table named SUPERTORE as per the data set



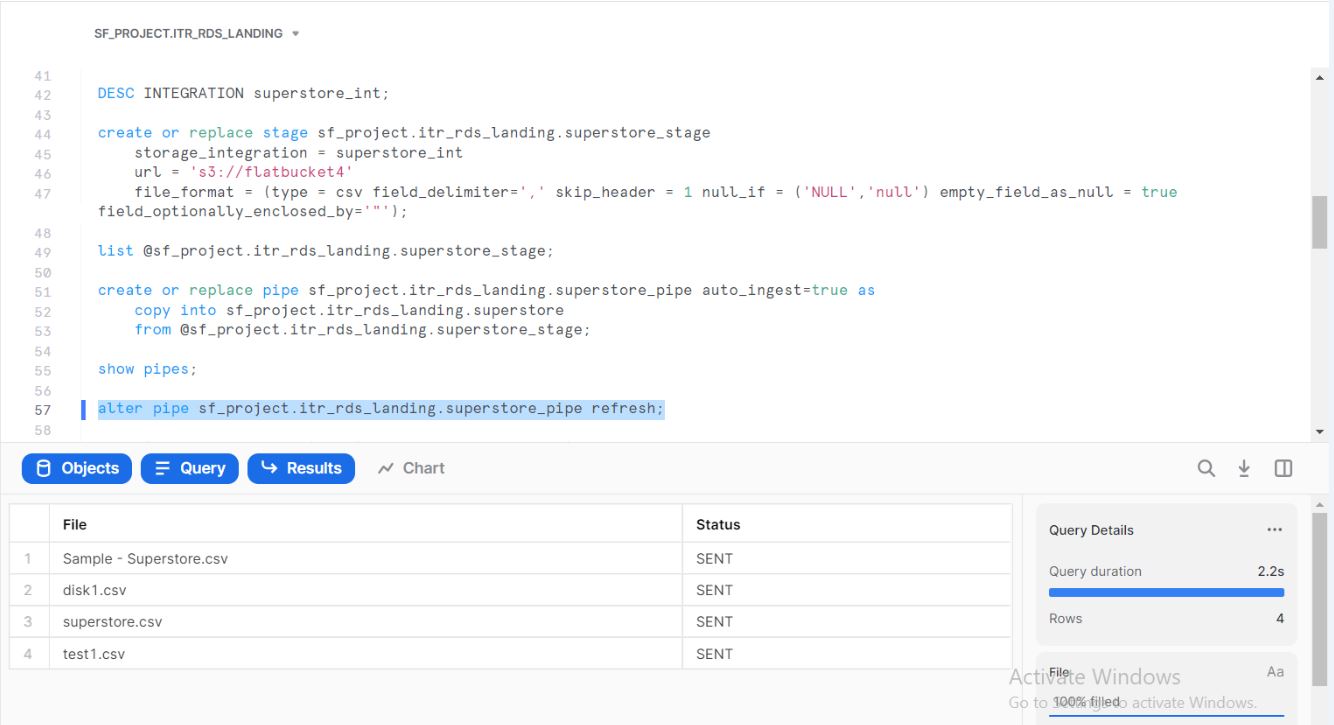
4. Created Integration object named SUPER\_STORE



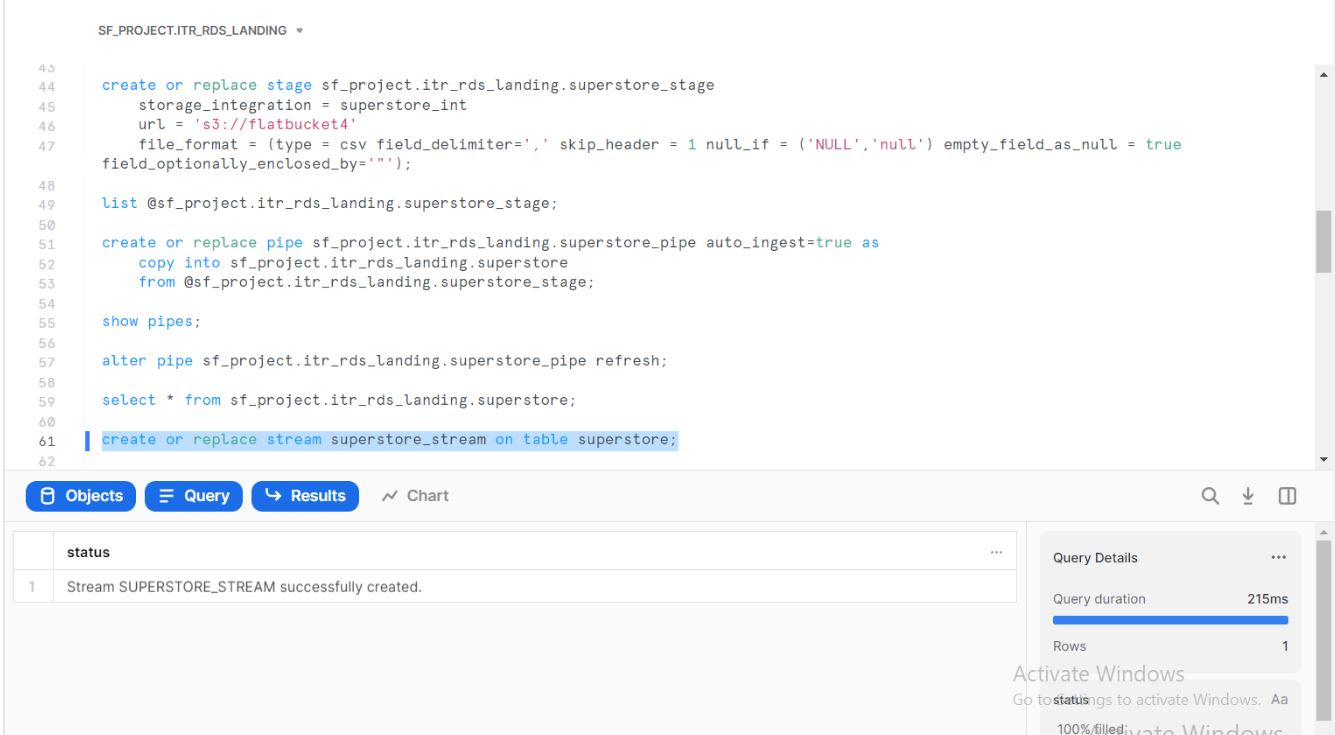
5. Created external stage named SUPERSTORE\_STAGE for loading data structures



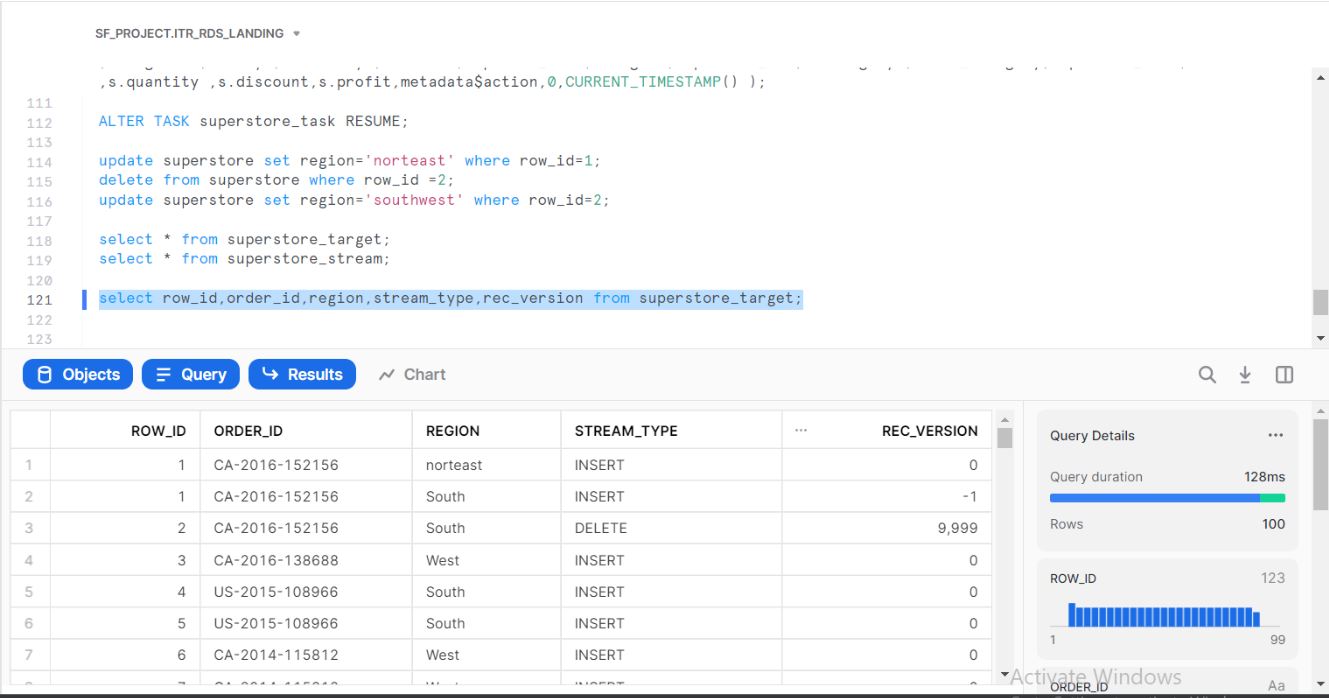
6. Created a snowpipe named SUPERSTORE\_PIPE for autoingesting the data from S3 bucket – flatbucket4



7. Created stream SUPERSTORE\_STREAM on table SUPERSTORE

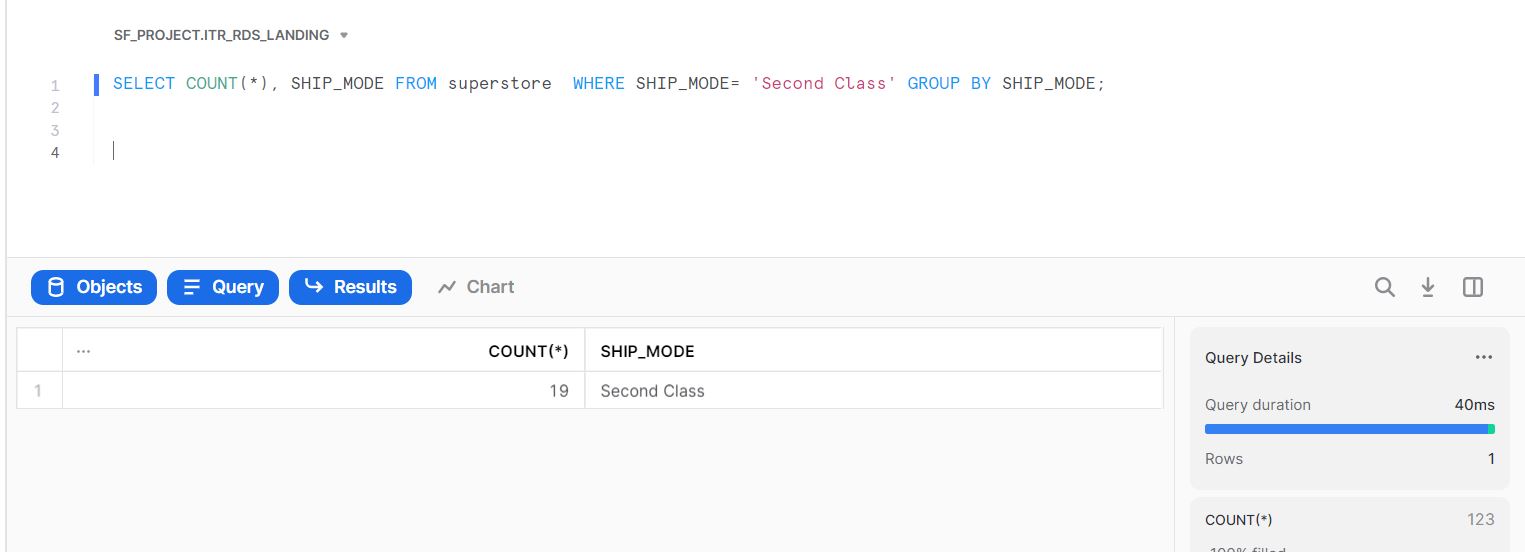


8. Performed SCD 2 operation on consumer table SUPERSTORE\_TARGET as per changes that happen in the source table SUPERSTORE



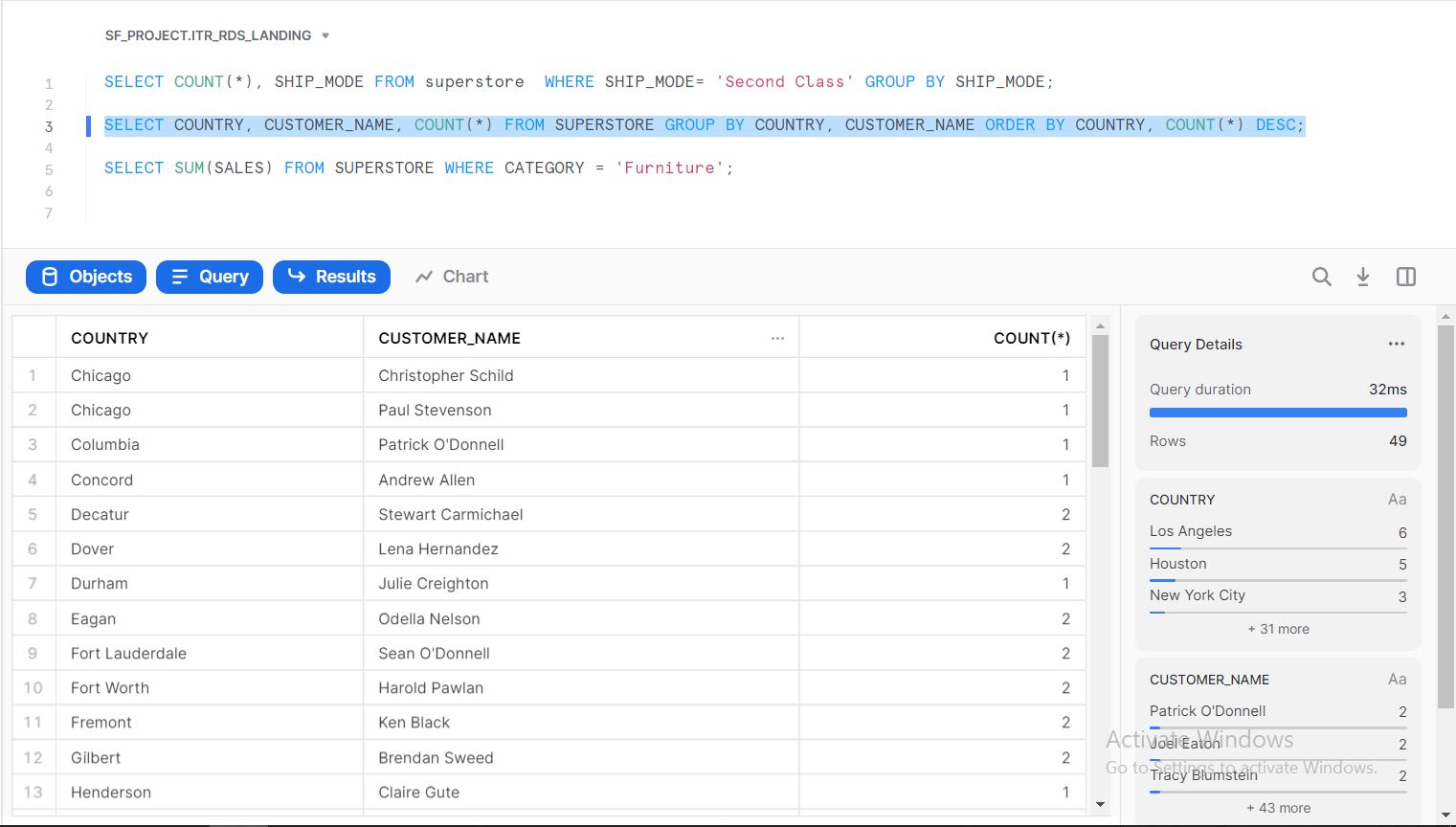
* ***DATA ANALYSIS ON THE GIVEN DATASET***

*--1. Calculate the number of Orders those with Ship Mode as ‘Second Class’*

SELECT COUNT(\*), SHIP\_MODE FROM superstore WHERE SHIP\_MODE= 'Second Class' GROUP BY SHIP\_MODE; 

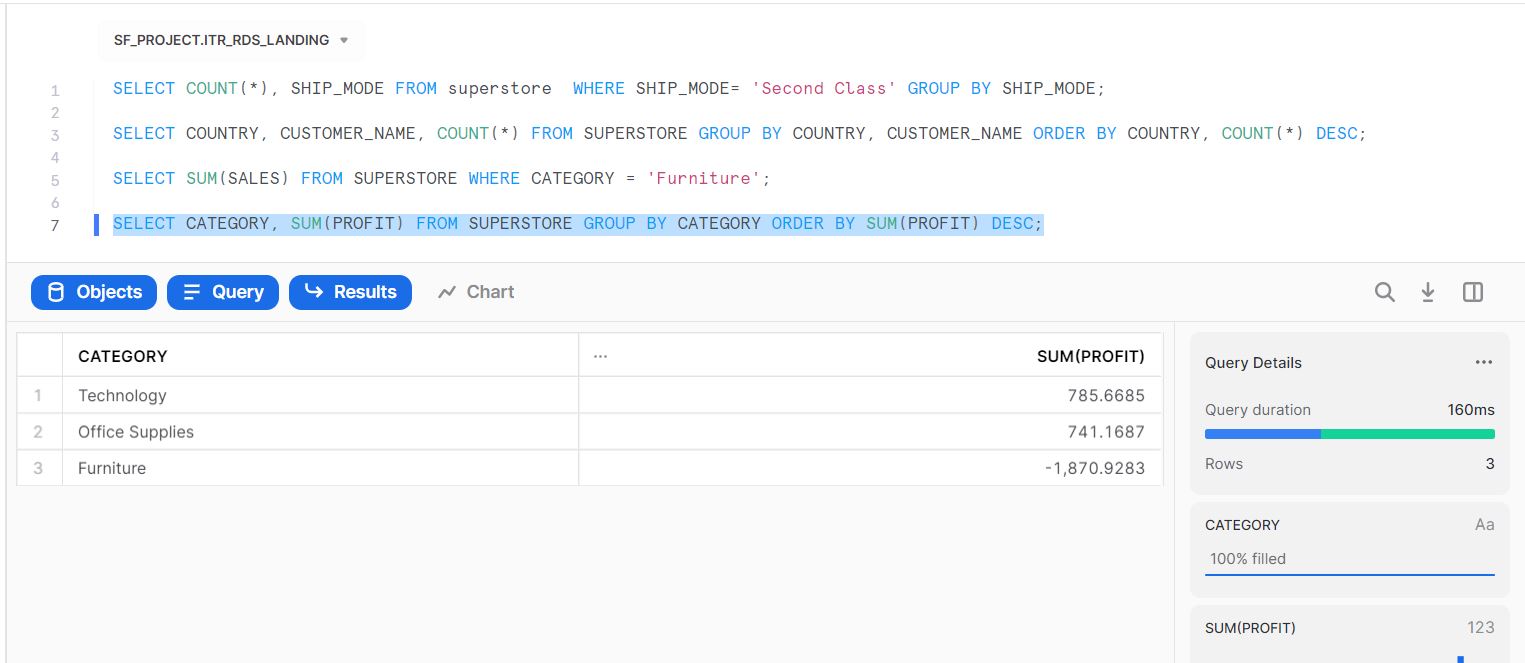
*--2. List down the most valuable customers Country wise*

SELECT COUNTRY, CUSTOMER\_NAME, COUNT() FROM SUPERSTORE GROUP BY COUNTRY, CUSTOMER\_NAME ORDER BY COUNTRY, COUNT() DESC;



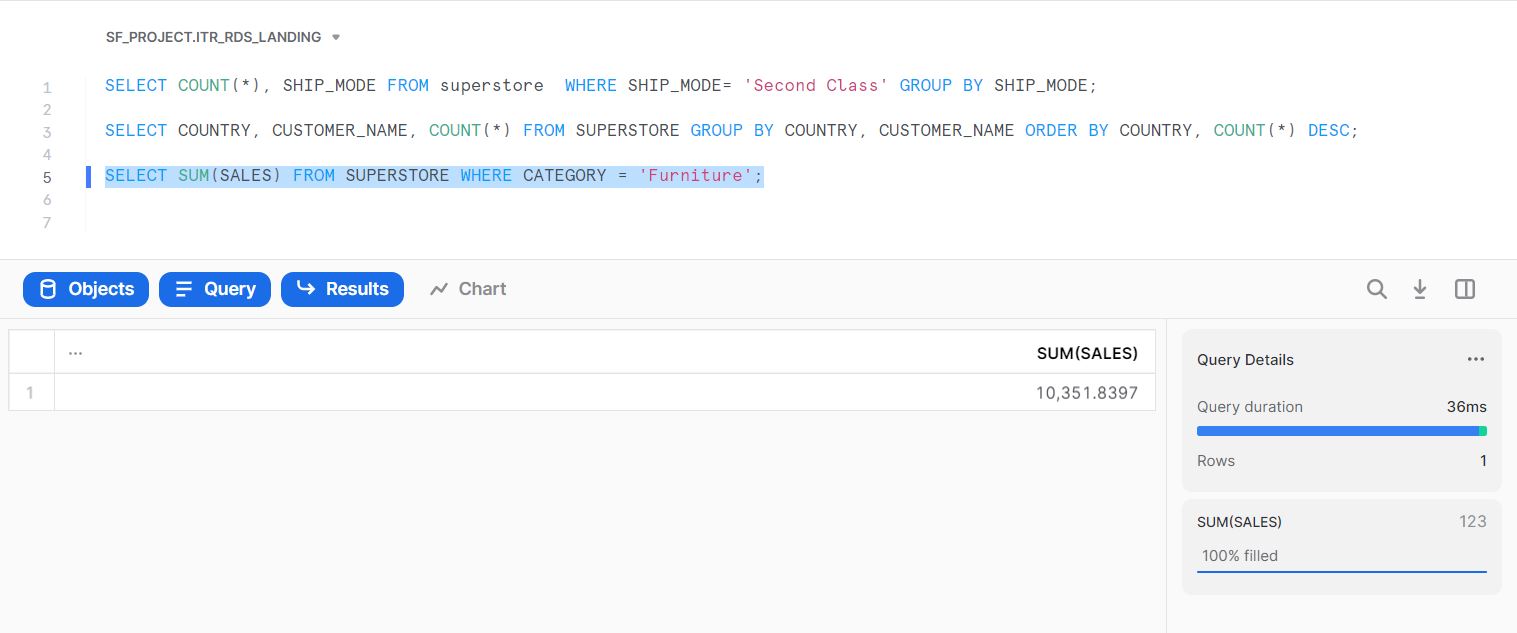
*--3. Total Sales for Category ‘Furniture’*

SELECT SUM(SALES) FROM SUPERSTORE WHERE CATEGORY = 'Furniture';



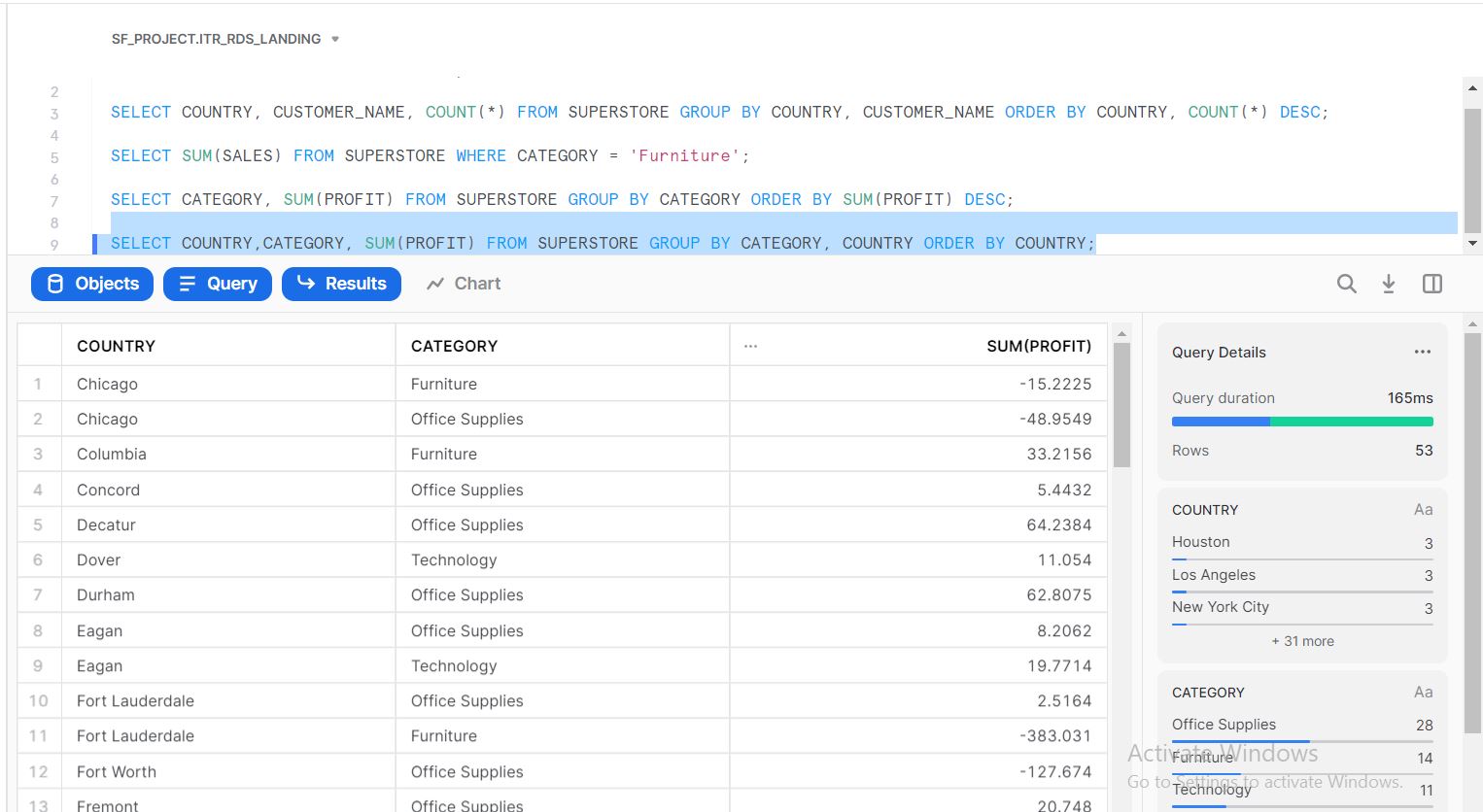
*--4. Which Product provides the maximum profit*

SELECT CATEGORY, SUM(PROFIT) FROM SUPERSTORE GROUP BY CATEGORY ORDER BY SUM(PROFIT) DESC;



*--5. Calculate the total profit made by each product category country wise*

SELECT COUNTRY,CATEGORY, SUM(PROFIT) FROM SUPERSTORE GROUP BY CATEGORY, COUNTRY ORDER BY COUNTRY;



--6. Which region of United States have majority loss?

SELECT REGION, MIN(PROFIT) FROM SUPERSTORE GROUP BY REGION;

