

Snowflake Final Project

Data Load unload

- 1) Download and save sales data file on your local machine
- 2) Create **SALES_DATA** table in Snowflake (same structure as sales data csv file)
- 3) Load csv data into snowflake **SALES_DATA** table
- 4) Verify total record count (note down for future reference)
- 5) Unload **SALES_DATA** table into seven CSV files partition by **REGION** column
- 6) Create Snowflake internal stage and load these seven csv file there
- 7) Truncate **SALSE_DATA** table

Transformation

- 8) Create table **AGG_SALES_BY_COUNTRY** with column REGION,COUNTRY,TOTAL_PROFIT
- 9) Load seven source csv files into **AGG_SALES_BY_COUNTRY** table using COPY command (only one row for each country should be present along with aggregated Profit)
- 10) Load these seven csv files into **SALSE_DATA** table but with Order Priority column transformation as per below
(L = Low, M = Medium, H = High, C = Critical)

Clustering

- 11) Write query to get total number of orders placed in ASIA region
- 12) Analyze query pattern and note down execution time
- 13) Cluster **SALSE_DATA** table and invalidate the cash
- 14) Again use same query as step 11, analyze query pattern and execution time

Tasks

- 15) Create snowflake task to load **SALES_DATA** table every hour except Sunday
- 16) Create snowflake task to load **AGG_SALES_BY_COUNTRY** at 4am IST twice a week on Monday and Thursday

Data Sharing

- 17) Create two readers accounts **READER_IND** and **READER_GERMANY**
- 18) Share table **SALES_DATA** with these two readers accounts in such a way **READER_IND** should be able to view sales only for India and same with **READER_GERMANY**