DELTA LAKE SOLUTION USING AZURE DATABRICKS

STEP 0

INGESTION

Load CSV Archive File to Azure Databricks HBFS

STEP 1

BRONZE STEP

Load CSV Files into HBFS and Save as Delta Files

STEP 2

SILVER STEP

Cleaning the data, manipulating it to derive new fields, and dropping duplicate records

STEP 3.1

GOLD STEP

Design a Star schema based on the business requirements and optimization for operation

STEP 3.2

GOLD STEP

Load into Star Schema & optimizing the schema for query performance