Week 7 Assignment

Customer Churn model dataset

Dataset: https://www.kaggle.com/datasets/kukuroo3/churn-model-data-set-competition-form

Project Architecture
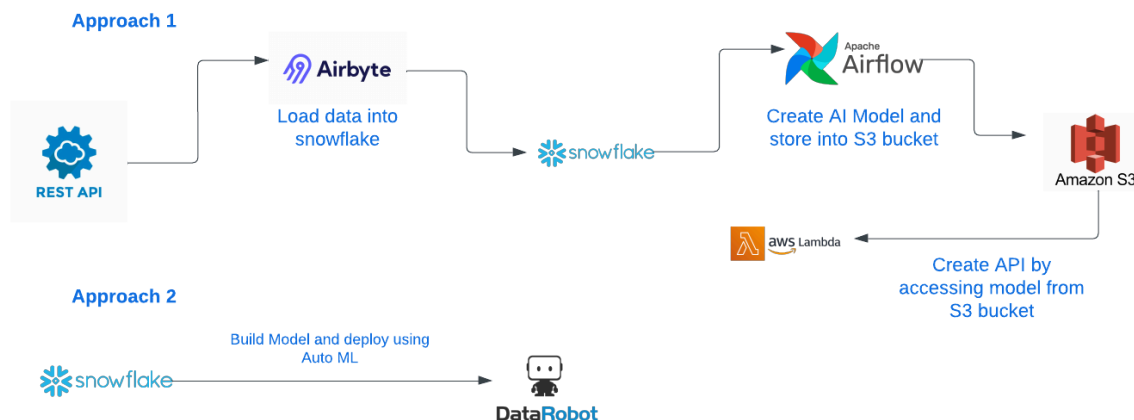
ELT Process Performed in the Demo:

Step 1: Download data csv file from github
Step 2: Convert CSV into JSON using python code. Python code is present into github folder
Step 3: Create Mock API (https://mocki.io/fake-json-api)
Step 4: Create Trial account in AirByte (https://cloud.airbyte.com/signup)
Step 5: Create Trial account in Snowflake from here (https://signup.snowflake.com/)
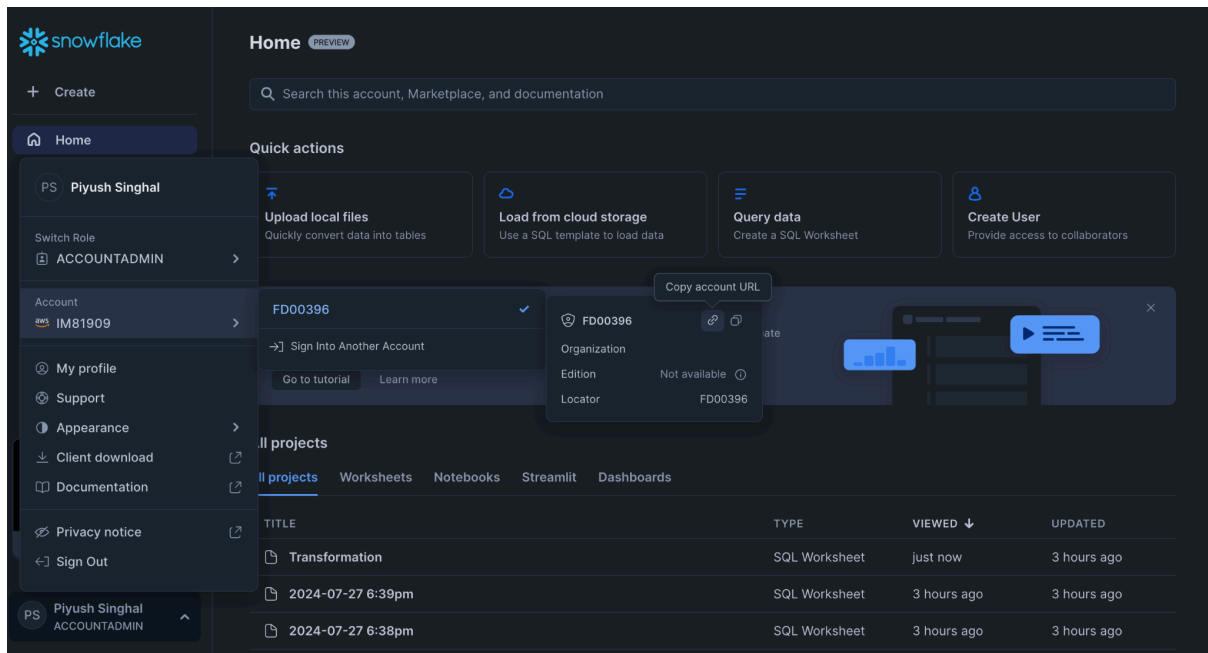1. Go to Snowflake and create Database named: demo_ai101
Step 6: Create ELT Pipeline into Airbyte (**Extract and Transform**)
1. Setup Source
    a. Select -> File (CSV, JSON, Excel, Feather, Parquet)
    b. File format: JSON
    c. Storage Provide: HTTPS: Public Web
    d. API URL
2. Setup Destination
    a. Name: Snowflake
    b. Host: {accountname}.{aws_location}.aws.snowflakecomputing.com
       Eg: fd00396.ap-south-1.aws.snowflakecomputing.com

      c. Role: ACCOUNTADMIN

      d. Warehouse: COMPUTE_WH

      e. Database: demo_ai101

      f. Default Schema: public

      g. Username: <Snowflake Username>

      h. Authorization Type: Username and Password

      i. Password: <Snowflake Password>

3. Setup Schedule

4. Run the Pipeline

5. Validate data into Snowflake

      a. Tables created by Airbyte

      b. Preview the data

      c. Run the SQL or create view on the table (**Transformation Layer**)

            Eg:

            CREATE VIEW TRANSFORMATION_VIEW AS

            (Select GENDER,RACE, concat(CHANGE, ' TEMP') as

            transformed_column from DEMO_AI101.public."HOSPITAL API" limit

            10);

Step 7: Download Tableau Desktop (https://www.tableau.com/products/desktop/download)

Step 8. Connect Tableau with Snowflake

Step 9: Try to Analyse data using Tableau

Step 10: Create a basic dashboard and publish on Tableau Public

-