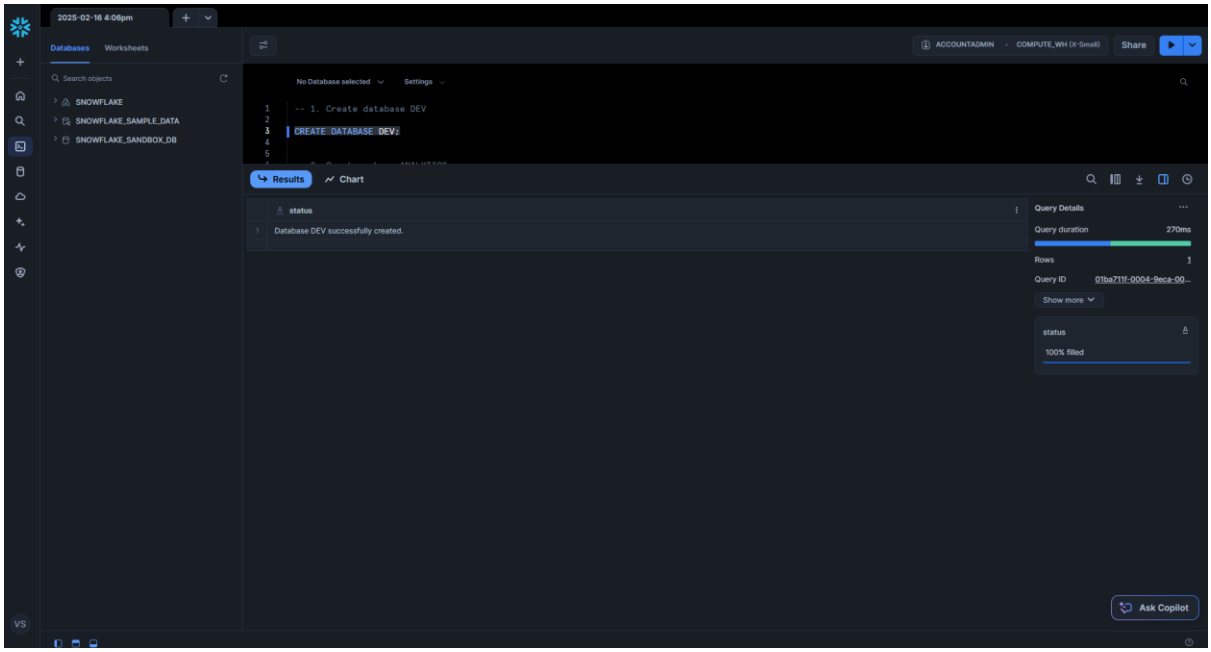
DATA 226- DATAWAREHOUSE

Homework 3

Name : Vimalanandhan Sivanandham

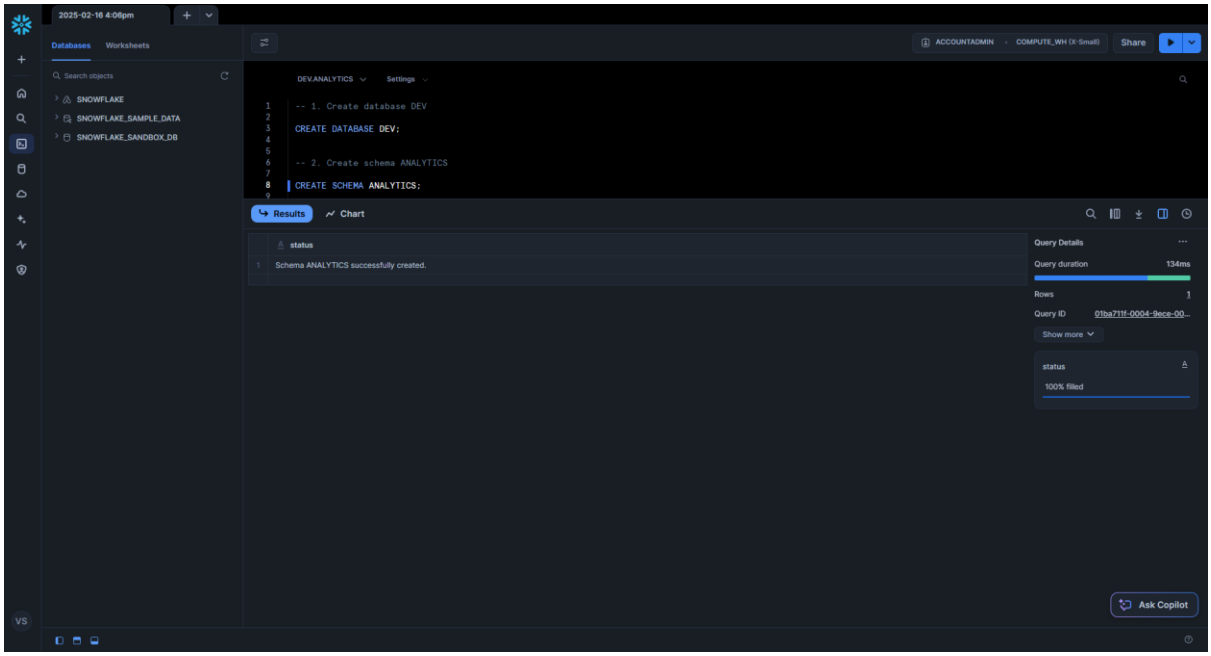
SJSU ID: 017596436

1. (+1) Create database DEV and schema ANALYTICS



-- 1. Create database DEV

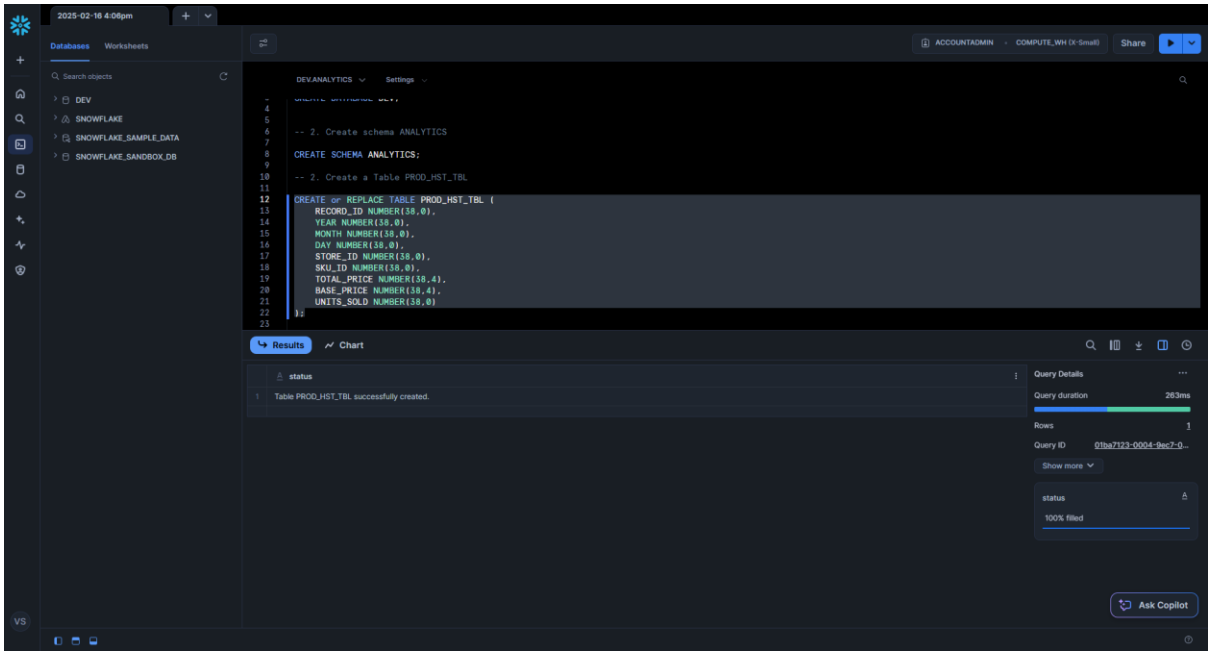
CREATE DATABASE DEV;



-- 2. Create schema ANALYTICS

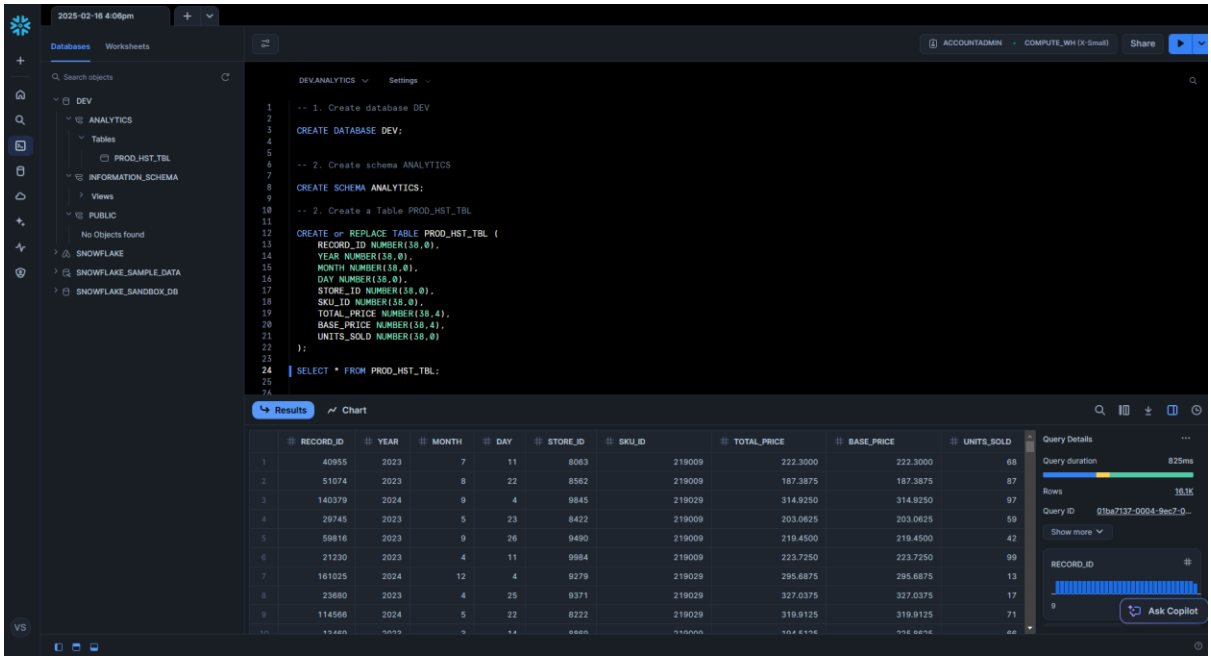
CREATE SCHEMA ANALYTICS;

2. (+2) Create a Table PROD_HST_TBL



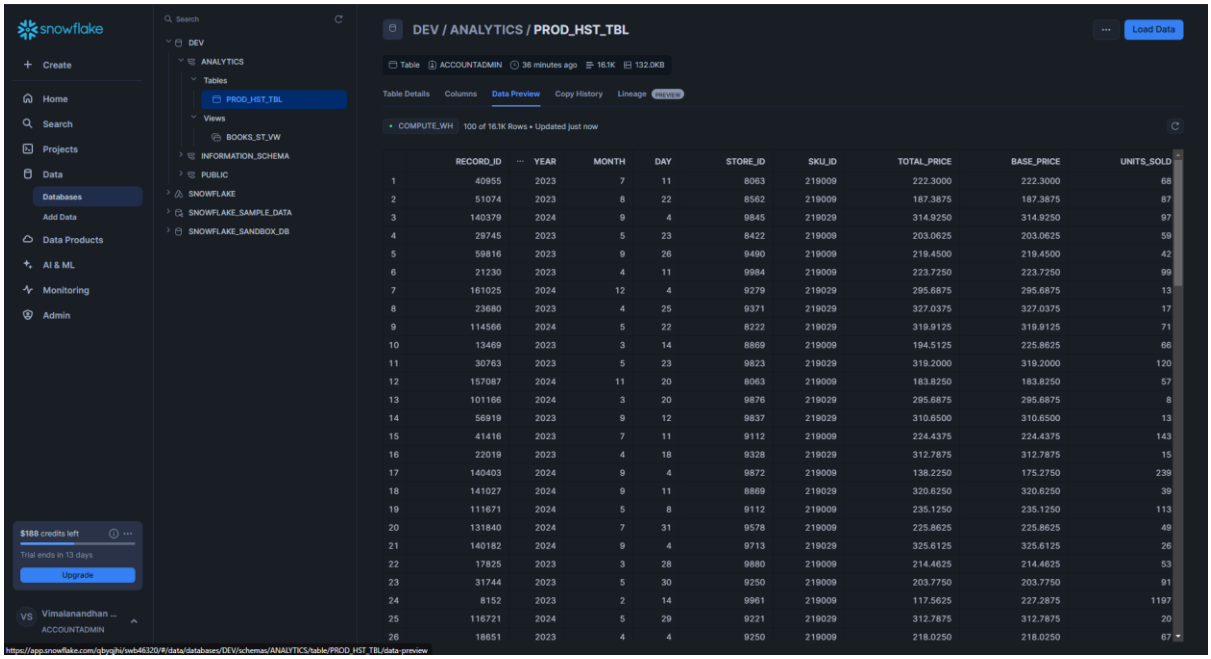
-- 2. Create a Table PROD_HST_TBL

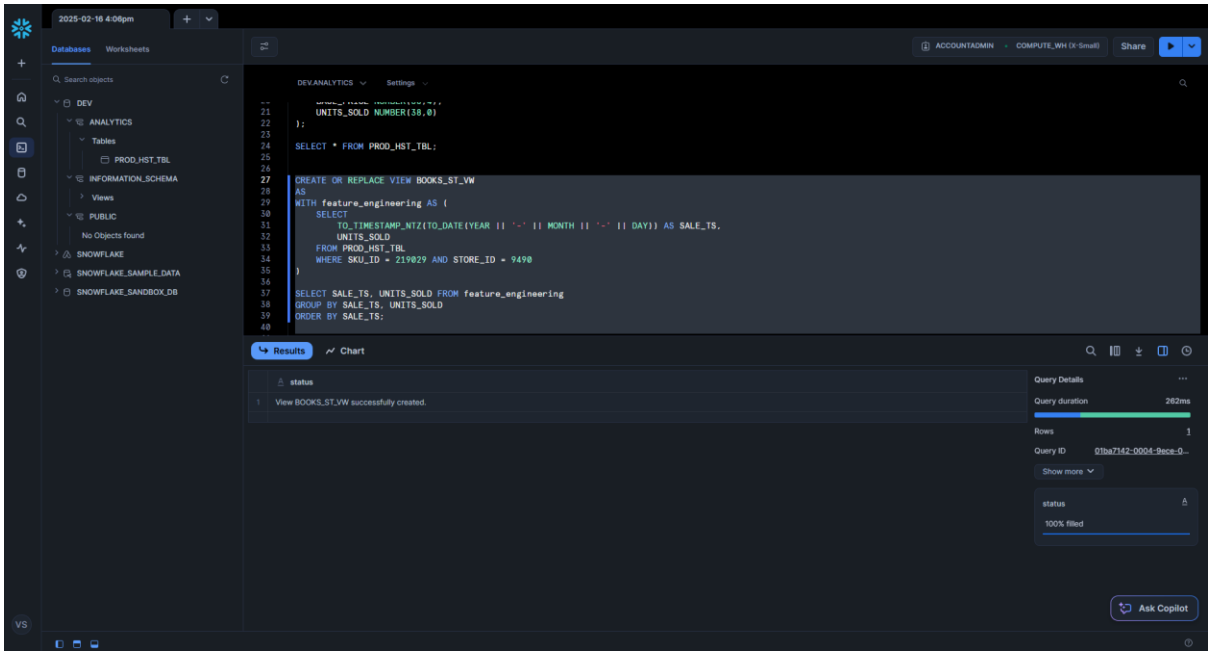
```
CREATE or REPLACE TABLE PROD_HST_TBL (
   RECORD_ID NUMBER(38,0),
   YEAR NUMBER(38,0),
   MONTH NUMBER(38,0),
   DAY NUMBER(38,0),
   STORE_ID NUMBER(38,0),
   SKU_ID NUMBER(38,0),
   TOTAL_PRICE NUMBER(38,4),
   BASE_PRICE NUMBER(38,4),
   UNITS_SOLD NUMBER(38,0)
);
```

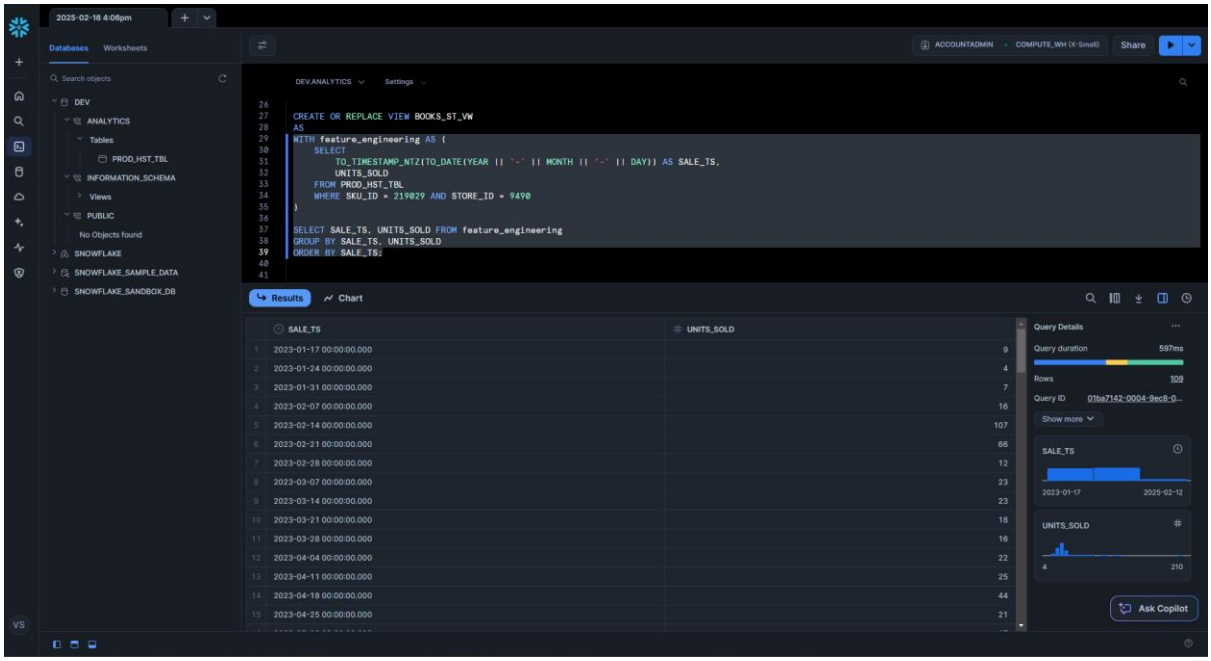


SELECT * FROM PROD_HST_TBL;

3. (+3) Create a view to forecast SKU ‘219029’ of STORE ‘9490’





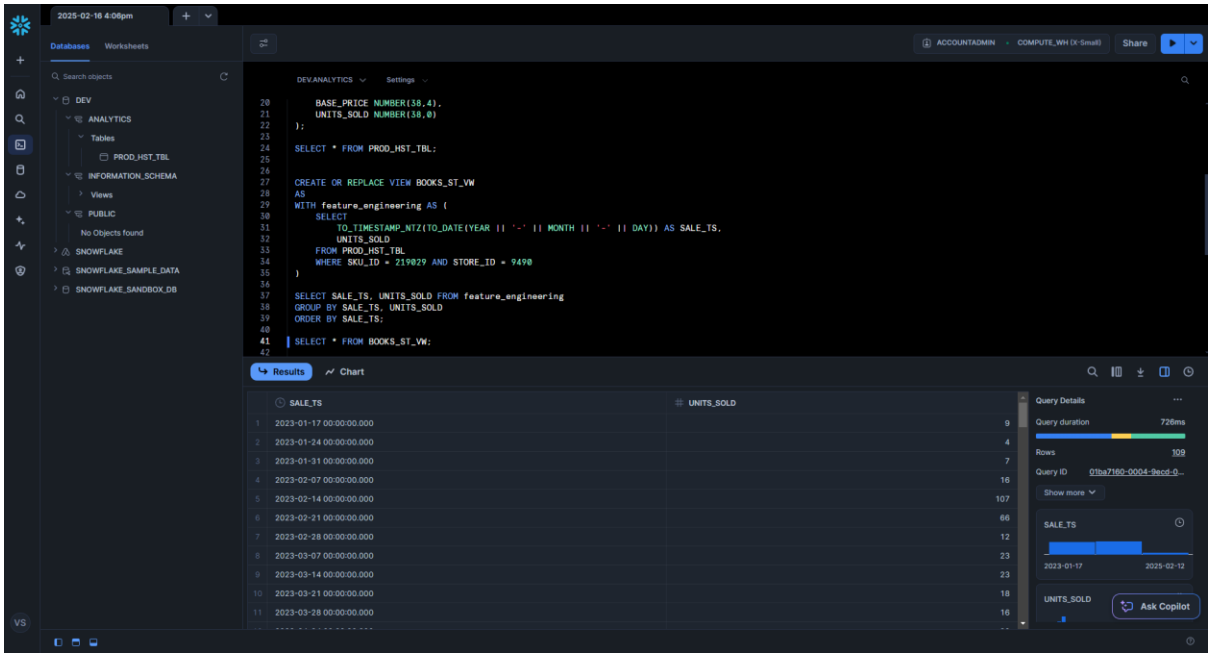


-- 3. Create a view to forecast SKU '219029' of STORE '9490'

```
CREATE OR REPLACE VIEW BOOKS_ST_VW
AS
WITH feature_engineering AS (
  SELECT
    TO_TIMESTAMP_NTZ(TO_DATE(YEAR || '-' || MONTH || '-' || DAY)) AS SALE_TS,
    UNITS_SOLD
```

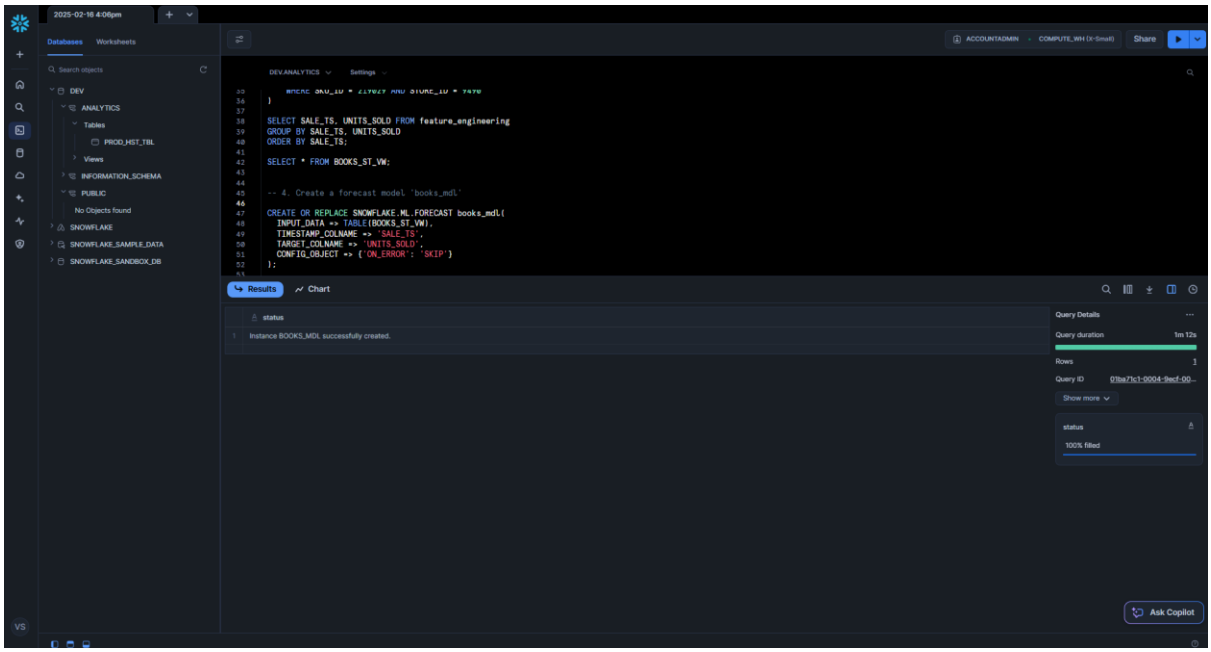
```
   FROM PROD_HST_TBL
   WHERE SKU_ID = 219029 AND STORE_ID = 9490
)
```

```
SELECT SALE_TS, UNITS_SOLD FROM feature_engineering
GROUP BY SALE_TS, UNITS_SOLD
ORDER BY SALE_TS;
```




```
SELECT * FROM BOOKS_ST_VW;
```

4. (+3) Create a forecast model ‘books_mdl’



-- 4. Create a forecast model 'books_mdl'

CREATE OR REPLACE SNOWFLAKE.ML.FORECAST books_mdl(

INPUT_DATA => TABLE(BOOKS_ST_VW),

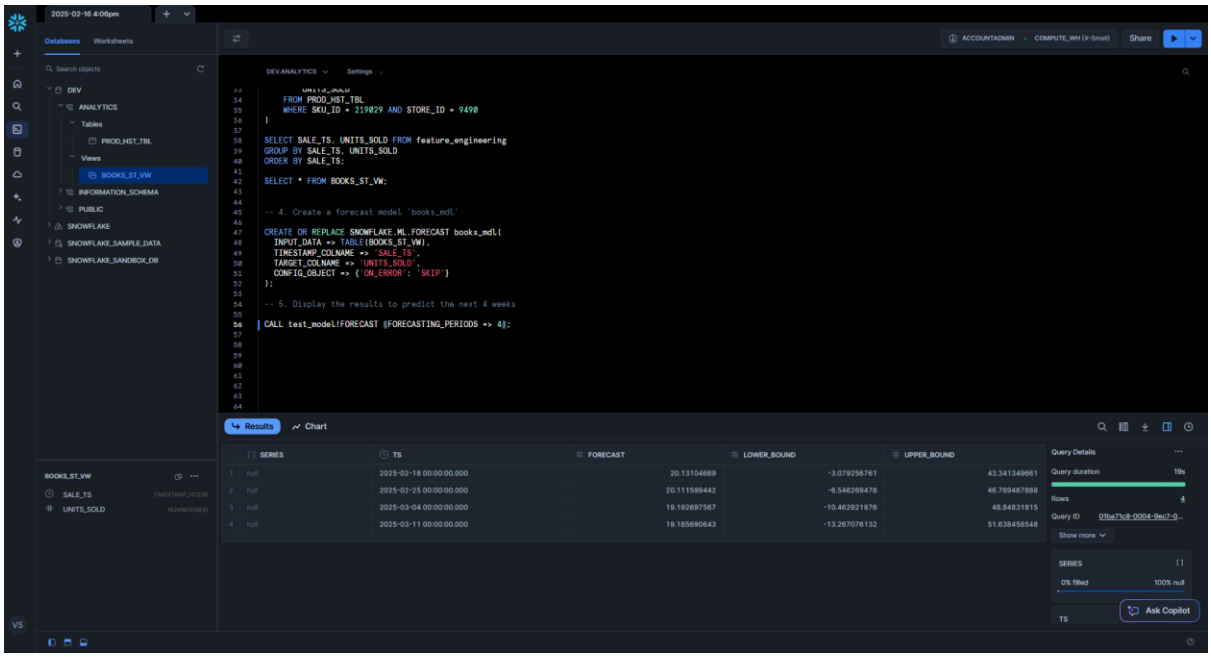
TIMESTAMP_COLNAME => 'SALE_TS',

TARGET_COLNAME => 'UNITS_SOLD',

CONFIG_OBJECT => {'ON_ERROR': 'SKIP'}

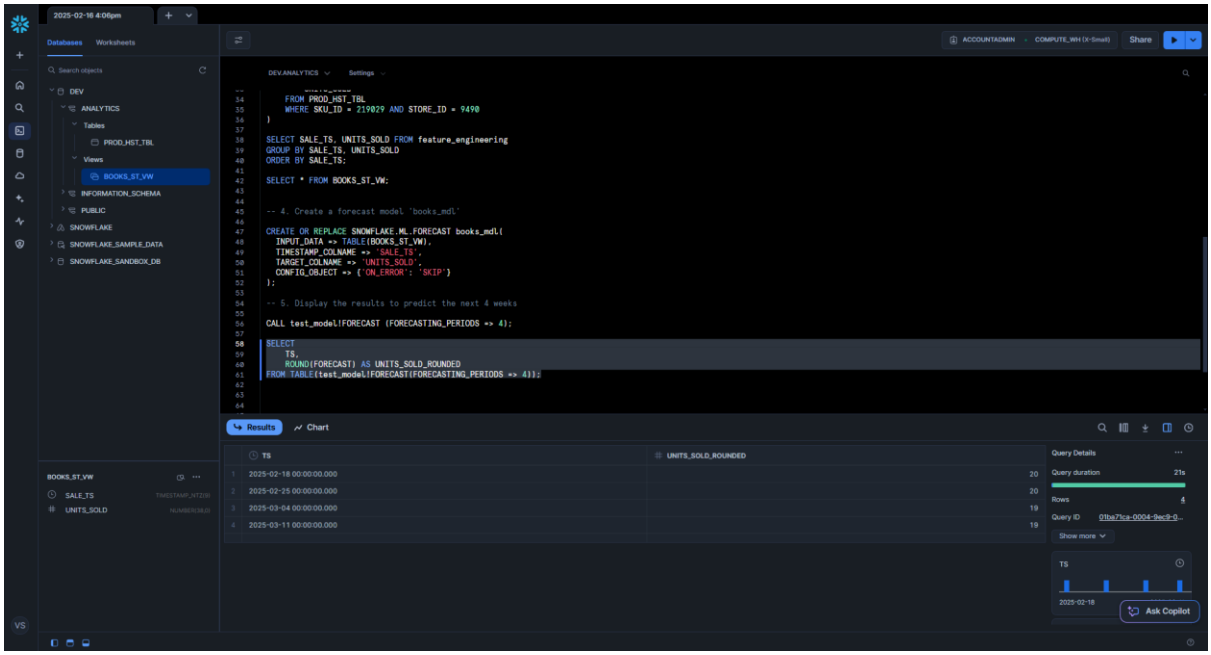
);

5. (+1) Display the Results to predict next 4 weeks.



-- 5. Display the results to predict the next 4 weeks

CALL test_model!FORECAST (FORECASTING_PERIODS => 4);



```
SELECT
  TS,
  ROUND(FORECAST) AS UNITS_SOLD_ROUNDED
FROM TABLE(test_model!FORECAST(FORECASTING_PERIODS => 4));
```

6. (+3) Explain your understanding about the Forecasting Process.

This assignment's forecasting procedure entails building a structured pipeline in Snowflake to project future sales for a certain SKU and retailer. In order to store historical sales data, the PROD_HST_TBL table must be created after the DEV database and ANALYTICS schema have been set up. Data is then explicitly filtered for SKU '219029' at shop '9490' using a view that has been developed. Building a forecasting model (books_mdl) with Snowflake ML is the main step in the forecasting process. This model uses historical sales data (BOOKS_ST_VW), finds trends, and forecasts future sales based on patterns in the SALE_TS (timestamp) and UNITS_SOLD (goal variable) columns. Presenting forecasts for the upcoming four weeks, which include information on anticipated sales, is the last phase.

In order to ensure real-time updates, the pipeline diagram illustrates how data is imported and converted through the RAW, CURATION, and ANALYTICS layers, and then saved in a dynamic table. The forecasts are more accurate and useful since this methodical methodology guarantees that only pertinent and clean data is used.

(+1) For following the format

Followed the process.