EXERCISE 6 Creating and querying tables in hive with sql

CREATE TABLE sales\_data (

order\_id INT,

product\_name STRING,

category STRING,

price FLOAT,

quantity INT,

order\_date STRING

) ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE;

LOAD DATA INPATH 'sales\_data.csv' INTO TABLE sales\_data;

-- Display all data

SELECT \* FROM sales\_data;

-- Calculate total sales per product

SELECT product\_name, SUM(price \* quantity) AS total\_salesFROM sales\_dataGROUP BY

product\_name;

-- Filter data for 'Electronics' category

SELECT \* FROM sales\_dataWHERE category = 'Electronics'

SALES AND PRODUCT DATA

**Start Hive CLI:**

bash

Copy code

hive

**Create a Database:**

sql

Copy code

CREATE DATABASE IF NOT EXISTS sales\_db;

**Use the Created Database:**

sql

Copy code

USE sales\_db;

**Create a Table for Products:**

sql

Copy code

CREATE TABLE IF NOT EXISTS products (

product\_id INT,

product\_name STRING,

category STRING,

price FLOAT

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE;

**Load Data into Products Table:**

sql

Copy code

LOAD DATA LOCAL INPATH '/path/to/product\_data.csv'

INTO TABLE products;

**Create a Table for Sales:**

sql

Copy code

CREATE TABLE IF NOT EXISTS sales (

order\_id INT,

product\_id INT,

quantity INT,

order\_date STRING

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE;

**Load Data into Sales Table:**

sql

Copy code

LOAD DATA LOCAL INPATH '/path/to/sales\_data.csv'

INTO TABLE sales;

**Query to Select All Products:**

sql

Copy code

SELECT \* FROM products;

**Query to Get Products with Price Greater than $500:**

sql

Copy code

SELECT \* FROM products

WHERE price > 500;

**Query to Count Total Sales by Product:**

sql

Copy code

SELECT p.product\_name, COUNT(s.order\_id) AS total\_sales

FROM products p

JOIN sales s ON p.product\_id = s.product\_id

GROUP BY p.product\_name;

**Query to Get Total Revenue by Product:**

sql

Copy code

SELECT p.product\_name, SUM(s.quantity \* p.price) AS total\_revenue

FROM products p

JOIN sales s ON p.product\_id = s.product\_id

GROUP BY p.product\_name;