7. Create Managed table and load a csv file from local storage. Verify the location of table data in HDFS and perform query operations. Delete the table and verify the data in HDFS.

To create a managed table in Hive, load a CSV file from local storage, verify the location of the table data in HDFS, perform query operations, and then delete the table and verify the data in HDFS, follow these steps:

**Step 1: Create and Load the Managed Table**

1. **Start Hive**:

hive

**Create the Managed Table**:

CREATE TABLE employee (

id INT,

name STRING,

salary DOUBLE

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE;

**Load Data from Local Storage**: Assuming your CSV file data.csv is in the local filesystem, load it into the Hive table:

LOAD DATA LOCAL INPATH '/path/to/data.csv' INTO TABLE employee;

**Step 2: Verify the Location of Table Data in HDFS**

1. **Check the HDFS Path**: Managed tables are stored in the default location, which is usually /user/hive/warehouse/<table\_name>. You can check this using the HDFS command:

hdfs dfs -ls /user/hive/warehouse/employee

**Step 3: Perform Query Operations**

1. **Run Some Queries**

-- Select all records

SELECT \* FROM employee;

-- Count the number of records

SELECT COUNT(\*) FROM employee;

-- Calculate the total salary

SELECT SUM(salary) FROM employee;

-- Find the highest salary

SELECT MAX(salary) FROM employee;

-- Select all records

SELECT \* FROM employee;

-- Count the number of records

SELECT COUNT(\*) FROM employee;

-- Calculate the total salary

SELECT SUM(salary) FROM employee;

-- Find the highest salary

SELECT MAX(salary) FROM employee;

### Step 4: Delete the Table and Verify the Data in HDFS

1. **Drop the Table**:

DROP TABLE employee;

**Verify the Data in HDFS**: After dropping the table, the data should also be deleted from HDFS.

hdfs dfs -ls /user/hive/warehouse/employee

### Complete Commands

Here's the complete sequence of commands to run in the Hive terminal and the shell:

**Hive Terminal:**

-- Create the table

CREATE TABLE employee (

id INT,

name STRING,

salary DOUBLE

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE;

-- Load data from local storage

LOAD DATA LOCAL INPATH '/path/to/data.csv' INTO TABLE employee;

-- Run queries

SELECT \* FROM employee;

SELECT COUNT(\*) FROM employee;

SELECT SUM(salary) FROM employee;

SELECT MAX(salary) FROM employee;

-- Drop the table

DROP TABLE employee;

# Verify the HDFS path before dropping the table

hdfs dfs -ls /user/hive/warehouse/employee

# Verify the HDFS path after dropping the table

hdfs dfs -ls /user/hive/warehouse/employee

Replace /path/to/data.csv with the actual path to your CSV file. This sequence will create a managed table, load data, run queries, drop the table, and verify the data's presence in HDFS at each step.