Exp no: 5a

Design and test various schema models to optimize data storage and retrieval using Hive

Aim:

To Design and test various schema models to optimize data storage and retrieval Using

Hbase.

Procedure:

Step 1: Start Hive

Open a terminal and start Hive by running:

$hive

Step 2: Create a Database

Create a new database in Hive:

hive>CREATE DATABASE financials;

```
hive> CREATE DATABASE financials;
OK
Time taken: 0.41 seconds
```

Step 3: Use the Database:

Switch to the newly created database:

hive>use financials

```
hive> use financials;
OK
Time taken: 0.031 seconds
```

Step 4: Create a Table:

Create a simple table in your database:

hive>CREATE TABLE finance_table(id INT, name STRING);

```
hive> CREATE TABLE finance_table (id INT, name STRING);
OK
Time taken: 0.552 seconds
```

Step 5: Load Sample Data:

You can insert sample data into the table:

hive>INSERT INTO finance_tableVALUES (1, 'Alice'), (2, 'Bob'), (3, 'Charlie');

```
hive> INSERT INTO finance_table VALUES (1, 'Alice'), (2,'Bob'), (3,'Charlie');
Query ID = haresh_20240912212701_f51e1fcd-d01a-4b83-8d6e-bf0388e6e714
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1726155698102_0001, Tracking URL = http://fedora:8088/proxy
2_0001/
```

Step 6: Query Your Data

Use SQL-like queries to retrieve data from your table:

hive>CREATE VIEW myview AS SELECT name, id FROM finance_table;

```
hive> CREATE VIEW myview AS SELECT name, id FROM finance_table;
OK
Time taken: 0.256 seconds
```

Step 7: View the data:

To see the data in the view, you would need to query the view

hive>SELECT*FROM myview;

```
hive> SELECT * FROM myview;
OK
Alice   1
Bob     2
Charlie 3
Time taken: 0.213 seconds, Fetched: 3 row(s)
```

Step 8: Describe a Table:

You can describe the structure of a table using the DESCRIBE command:

hive>DESCRIBE finance_table;

```
hive> DESCRIBE finance_table;
OK
id                      int
name                    string
Time taken: 0.075 seconds, Fetched: 2 row(s)
```

Step 9: Alter a Table:

You can alter the table structure by adding a new column:

hive>ALTER TABLE finance_table ADD COLUMNS (age INT);

```
hive> ALTER TABLE finance_table ADD COLUMNS (age INT);
OK
Time taken: 0.126 seconds
hive> quit;
```

Step 10: Quit Hive:

To exit the Hive CLI, simply type:

hive>quit;

Result:

Thus, the usage of various commands in Hive has been successfully completed