# Create tables in Hive and write queries to access the data in the table

### AIM:

To create tables in Hive and write queries to access the data in the table.

### **PROCEDURE:**

- 1. Install the Apache Derby.
- 2. Download Apache Hive binaries and set the environment variables.
- 2a. Copy Derby libraries

we should go to the Derby libraries directory (db-derby-10.14.2.0\lib) and copy all \*.jar files. Then, we should paste them within the Hive libraries directory.

2b. Configuring hive-site.xml and Hive's Bin folder

Refer following link to download the file. Also download the guava file. Put hive-site.xml file to hive's conf location and replace hive's current guava file with this one in lib location. Also download the bin folder from link and replace the existing hive's bin folder.

https://drive.google.com/drive/folders/1r1OTDNcw8aF74V21k8GlD2IUGMJ\_v11o?usp=drive\_link

3. Starting Hadoop Services.

```
C:\Windows\System32>start-all.cmd
This script is Deprecated. Instead use start-dfs.cmd and start-yarn.cmd
starting yarn daemons
C:\Windows\System32>jps
10068 NodeManager
30616 Jps
21308 DataNode
5612 ResourceManager
5836 NameNode
```

4.Start Derby Network Server: Run the following command in separate window to open Derby

```
C:\Windows\System32>StartNetworkServer -h 0.0.0.0
Mon Sep 09 21:06:55 IST 2024 : Security manager installed using the Basic server security policy.
Mon Sep 09 21:06:55 IST 2024 : Apache Derby Network Server - 10.12.1.1 - (1704137) started and ready to accept connections on port 1527
```

5. Starting Apache Hive: Go to Apache Hive's bin location with cd command and run the following command

hive --service schematool -dbType derby -initSchema

6. Open Hive shell by typing:

hive

## Creating tables and accessing its data:

1. Create database name joshin and use that database to create table. Create a table named students with columns for student name, department, and age.

## **CREATE DATABASE joshin;**

USE joshin;

**CREATE TABLE students (student\_name STRING,department STRING,age INT);** 

2. Insert data into the students table.

INSERT INTO TABLE students VALUES ('Joshin', 'Computer Science', 20);

3. View the structure of the students table using the DESCRIBE command:

### **DESCRIBE** students:

4. List all tables in the current database:

## **SHOW TABLES;**

5. View the contents of the students table using a SELECT query:

## **SELECT \* FROM students;**

### **OUTPUT:**

### 1.Creating Database:

```
hive> CREATE DATABASE joshin;
2024-09-09721:14:11,259 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186060da36
2024-09-09721:14:11,259 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to cc9ab347-f559-4102-b401-da186060da36 main
2024-09-09721:14:11,389 WARN [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.ql.session.SessionState - METASTORE_FILTER_HOOK will be igno
red, since hive.security.authorization.manager is set to instance of HiveAuthorizerFactory.
OK
Time taken: 0.265 seconds
2024-09-09721:14:11,528 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id
cc9ab347-f559-4102-b401-da186060da36
2024-09-09721:14:11,528 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to main
```

```
hive> SHOW DATABASES;
2024-09-09T21:14:25,449 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186060da3 6
2024-09-09T21:14:25,449 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to cc9ab347-f559-4102-b401-da186060da36 main OK 2024-09-09T21:14:25,680 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.conf.Configuration.deprecation - mapred.input.dir is deprecated. Instead, use mapreduce.input.fileinputformat.inputdir default joshin my.database Time taken: 0.22 seconds, Fetched: 3 row(s) 2024-09-09T21:14:25,789 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: c203b347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.conf.HiveConf - Resetting thread name to main
```

#### 2.Use created Database to create table:

have USE Joshin;
2024-09-09721:14:36,499 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186060da3
6
2024-09-09721:14:36,499 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to cc9ab347-f559-4102-b401-da186060da36 main
0K
Time taken: 0.03 seconds
2024-09-09721:14:36,532 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id
cc9ab347-f559-4102-b401-da186060da36
2024-09-09721:14:36,532 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to main

#### 3. Creating table to store data:

hive CREATE TABLE students (student name STRING department STRING age INT);
2024-09-09712:18:17, 439 INFO [main] org. apache hadoop hive con-fiseConfor - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186960da36
2024-09-09721:18:17, 439 INFO [main] org. apache hadoop hive con-fiseConfor - Using thread name to cc9ab347-f559-4102-b401-da186060da36 main
ok
Time taken: 0.669 seconds
2024-09-09721:18:18,110 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org. apache.hadoop.hive.conf. HiveConf - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186060da36
600436
2024-09-09712:18:18,110 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org. apache.hadoop.hive.ql.session.SessionState - Resetting thread name to main

#### 4. Inserting data into the table:

sives INSERT INTO TABLE students VALUES ('Joshin', 'Computer Science', 20);

2024-09-0971:19:43, 43 INFO [main] org. aparche haddoop hive conf HiveCoff - Using the default value passed in for log id: cc0ab347-f559-4102-b401-da186060da36

2024-09-0971:19:43, 455 INFO [main] org. aparche haddoop hive conf HiveCoff - Updating thread name to cc0ab347-f559-4102-b401-da186060da36 main]

2024-09-0971:19:46, 501 INFO [cc2ab347-f559-4102-b40]-da186060da36 main] org. aparche haddoop hive.common.FileUtils - Creating directory if it doesn't exist: hdfs://localhost:9000/user/hive/mrehouse/joshin.db/students/.hive-staging.hive\_staging.hive\_2024-09-09\_21-19-43\_480\_722993039303565443-1

2024-09-0971:19:46,674 INFO [cc2ab347-f559-4102-b40]-da186060436 main] org. aparche haddoop.hive.common.FileUtils - Creating directory if it doesn't exist: hdfs://localhost:9000/tmp/hive/ass/cc2ab347-f559-4102-b401-da186060436/hive\_2024-09-09\_21-19-46\_336\_744273267980252980-1/-mr-10000/.hive-staging\_hive\_2024-09-09\_21-19-46\_336\_744273267980252980-1/-mr-10000/.hive-staging\_hive\_2024-09-09\_21-19-46\_336\_744273267980252980-1

#### 5. Viewing the created table:

hive> DESCRIBE students;

2024-09-0971:22:56,638 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: cc0ab347-f559-4102-b401-da186060da36
2024-09-09721:22:56,638 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to cc9ab347-f559-4102-b401-da186060da36 main

Of the string department string int interpretation of the string interpretation of the

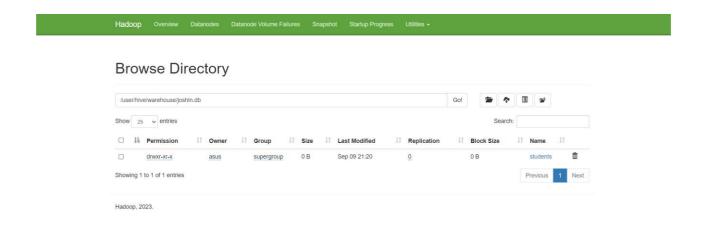
Nave> Select: F MON students;
2024-09-09721:23:19,237 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186060da36
2024-09-09721:23:19,237 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.common.Ficlutisr - Creating directory if it doesn't exist: hdfs://localhost:9000/tmp/hive/as
us/cc9ab347-f559-4102-b401-da186060da36/hive\_2024-09-09\_21-23-19\_255\_2005033459224140111-1/-mr-10001/.hive-staging\_hive\_2024-09-09\_21-23-19\_255\_2005033459224140111-1

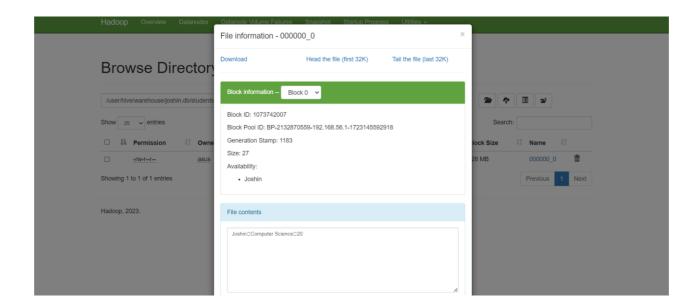
Now Joshin Computer Science 20
Time talen: 0.216 seconds, Fetched: 1 row(s)
2024-09-09721:23:19,484 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.comf.HiveConf - Using the default value passed in for log id: cc9ab347-f559-4102-b401-da186060da36
2024-09-09721:23:19,484 INFO [cc9ab347-f559-4102-b401-da186060da36 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to main

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities •

## Browse Directory







# **RESULT:**

Thus to create tables in Hive and write queries to access the data in the table is completed successfully.