

## 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:

#### Installation of Hive:

##### 1. Installing Apache Derby

Install Apache Derby 10.14.2.0

[https://db.apache.org/derby/derby\\_downloads.html#For+Java+8+and+Higher](https://db.apache.org/derby/derby_downloads.html#For+Java+8+and+Higher)

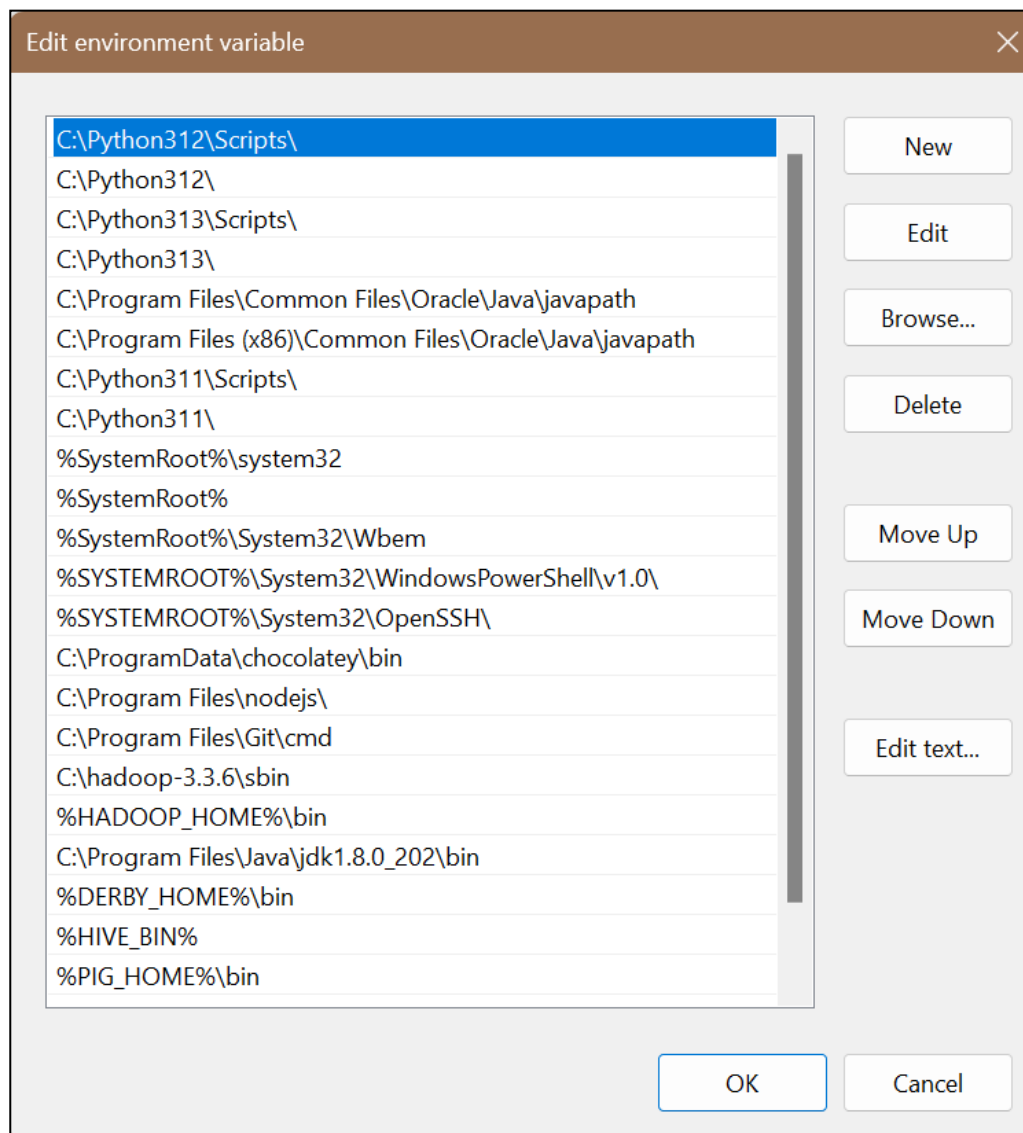
##### 2. Downloading Apache Hive binaries

<https://downloads.apache.org/hive/hive-3.1.2/>

##### 3. Setting environment variables

Variable	Value
HADOOP_HOME	C:\hadoop-3.3.6
HIVE_BIN	%HIVE_HOME%\bin
HIVE_HOME	C:\apache-hive-3.1.3-bin
HIVE_LIB	%HIVE_HOME%\lib

DERBY_HOME	C:\db-derby-10.14.2.0-bin
------------	---------------------------



### 3.1. 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.

### 3.2. 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://1drv.ms/f/s!ArSg3Xpur4Grmw0SDqW0g44T7HYU?e=wDsoBn>

#### 4. 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>
```

#### 5. Derby Network Server: Run the following command in separate window to open Derby

```
C:\Windows\System32>StartNetworkServer -h 0.0.0.0
Fri Sep 13 08:40:27 IST 2024 : Security manager installed using the Basic server security policy.
Fri Sep 13 08:40:27 IST 2024 : Apache Derby Network Server - 10.14.2.0 - (1828579) started and ready to accept connections on port 1527
```

#### 6. Starting Apache Hive

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

#### 7. Open Hive shell by typing:

```
C:\hive\bin>hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/C:/hive/lib/log4j-slf4j-impl-2.17.1.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/J:/SOFTWARE/hadoop-3.3.6/share/hadoop/common/lib/slf4j-log4j12.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2024-09-13T08:45:58,809 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Found configuration file at C:/hive/conf/hive-site.xml
2024-09-13T08:45:59,130 WARN [main] org.apache.hadoop.hive.conf.HiveConf - HiveConf configuration file C:/hive/conf/hive-site.xml is not found
2024-09-13T08:46:10,000 INFO [pool-9-thread-1] org.apache.hadoop.hive.ql.exec.DDLTask - Creating database 'hive'
2024-09-13T08:46:10,004 INFO [pool-9-thread-1] org.apache.hadoop.hive.ql.exec.DDLTask - Database 'hive' created successfully
2024-09-13T08:46:10,011 INFO [pool-9-thread-1] org.apache.hadoop.hive.ql.exec.DDLTask - Database 'hive' created successfully
hive>
```

#### Create a Database:

Start by creating a database. Open the Hive CLI and follow the steps below:

1. Use the **CREATE DATABASE** statement to create a new database:

```
hive> CREATE DATABASE mydatabase;
2024-08-31T12:19:59,567 INFO [9816a416-cf
e default value passed in for log id: 981
2024-08-31T12:20:01,052 WARN [9816a416-cf
- METASTORE_FILTER_HOOK will be ignored,
actory.
OK
Time taken: 1.048 seconds
```

2. Verify the database is present:

```
hive> SHOW DATABASES;
2024-09-13T08:58:21,243 INFO [main] org.apache.
6388abfb-ae9d-47d6-b952-e8c51a48d2f6
2024-09-13T08:58:21,244 INFO [main] org.apache.
9d-47d6-b952-e8c51a48d2f6 main
OK
2024-09-13T08:58:21,848 INFO [6388abfb-ae9d-47d
n - mapred.input.dir is deprecated. Instead, us
default
mydatabase
```

3. Switch to the new database:

```
hive> USE default;
2024-09-13T08:59:30,880 INFO [main] org.apache.hadoop.hive
6388abfb-ae9d-47d6-b952-e8c51a48d2f6
2024-09-13T08:59:30,880 INFO [main] org.apache.hadoop.hive
9d-47d6-b952-e8c51a48d2f6 main
OK
Time taken: 0.054 seconds
2024-09-13T08:59:30,944 INFO [6388abfb-ae9d-47d6-b952-e8c5
```

### Create a Table in Hive

```
CREATE TABLE students (name STRING, roll INT, dept STRING);
```

### List Hive Tables and Data

To show all tables in a selected database, use the following statement:

```
SHOW TABLES;
```

To show table column names and data types, run:

```
DESC students;
```

To display table data, use a **SELECT** statement. For example, to select everything in a table, run:

```
SELECT * FROM students;
```

### OUTPUT:

```
hive> INSERT INTO TABLE mytable VALUES (1, 'John Doe', 30);
```

```
hive> SELECT * FROM mytable;
2024-08-31T12:23:32,793 INFO [main] org.apache.ha
9816a416-cfec-42c1-ac09-42e0f81079f2
2024-08-31T12:23:32,794 INFO [main] org.apache.ha
ec-42c1-ac09-42e0f81079f2 main
2024-08-31T12:23:33,166 INFO [9816a416-cfec-42c1-
ing directory if it doesn't exist: hdfs://localho
8-31_12-23-32_823_243596335034078621-1/-mr-10001/
OK
1      John Doe      30
2      Jane Smith    25
Time taken: 0.442 seconds, Fetched: 2 row(s)
```

### RESULT:

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