EXP: 5 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.
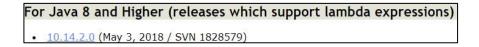
PROCEDURE:

Hive Download and installation:

1. Hive Installation setup:

- Download and install Apache Derby version 10.14.2.0:

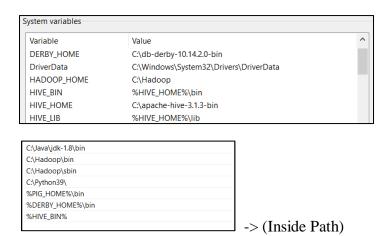https://db.apache.org/derby/derby_downloads.html#For+Java+8+and+Higher

For Java 8 and Higher (releases which support lambda expressions)
- 10.14.2.0 (May 3, 2018 / SVN 1828579)

-Download and install Apache Hive version 3.1.3:

https://downloads.apache.org/hive/hive-3.1.3/

downloads.apache.org/hive/hive-3.1.3/

Index of /hive/hive-3.1.3

Name	Last modified	Size	Description
Parent Directory		-	
apache-hive-3.1.3-bin.tar.gz	2022-04-08 17:42	312M	

2. Add environment variables:

Environment variables > System variables > Add the below paths

System variables

Variable	Value
DERBY_HOME	C:\db-derby-10.14.2.0-bin
DriverData	C:\Windows\System32\Drivers\DriverData
HADOOP_HOME	C:\Hadoop
HIVE_BIN	%HIVE_HOME%\bin
HIVE_HOME	C:\apache-hive-3.1.3-bin
HIVE_LIB	%HIVE_HOME%\lib

```
C:\Java\jdk-1.8\bin
C:\Hadoop\bin
C:\Hadoop\sbin
C:\Python39\
%PIG_HOME%\bin
%DERBY_HOME%\bin
%HIVE_BIN%
```
-> (Inside Path)

3. Copy Derby libraries

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

4. 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

5. Starting Hadoop Services

Open PowerShell as administrator and go to Hadoop sbin directory and start hadoop services using the following commands:

start-dfs.cmd

start-yarn.cmd

```
PS C:\WINDOWS\system32> cd C:\hadoop\sbin
PS C:\hadoop\sbin> start-dfs.cmd
PS C:\hadoop\sbin> start-yarn.cmd
starting yarn daemons
PS C:\hadoop\sbin> jps
12608 NameNode
17136 Jps
13044 ResourceManager
8260 NodeManager
312 DataNode
PS C:\hadoop\sbin>
```

6. Derby Network Server:

Open another PowerShell window and run the following command to open Derby:

StartNetworkServer -h 0.0.0.0

```
PS C:\WINDOWS\system32> StartNetworkServer -h 0.0.0.0
Fri Sep 27 19:57:13 IST 2024 : Security manager installed using the Basic server security policy.
Fri Sep 27 19:57:13 IST 2024 : Apache Derby Network Server - 10.14.2.0 - (1828579) started and ready to accept connectio
ns on port 1527
```

Go to first PowerShell window and check whether NetworkServerControl is running.

```
PS C:\hadoop\sbin> jps
12608 NameNode
13044 ResourceManager
14724 Jps
8260 NodeManager
312 DataNode
21148 NetworkServerControl
PS C:\hadoop\sbin>
```

7. Starting Apache Hive:

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

> hive --service schematool -dbType derby –initSchema

```
PS D:\hive\bin> hive --service schematool -dbType derby -initSchema
2024-09-27 20:05:28,491 INFO conf.HiveConf: Found configuration file file:/D:/hive/conf/hive-site.xml
2024-09-27 20:05:28,929 INFO tools.HiveSchemaHelper: Metastore connection URL:    jdbc:derby://localhost:1527/m
b;create=true
Metastore connection URL:        jdbc:derby://localhost:1527/metastore_db;create=true
2024-09-27 20:05:28,929 INFO tools.HiveSchemaHelper: Metastore Connection Driver :      org.apache.derby.jdbc
ver
Metastore Connection Driver :     org.apache.derby.jdbc.ClientDriver
2024-09-27 20:05:28,930 INFO tools.HiveSchemaHelper: Metastore connection User:  APP
Metastore connection User:        APP
Starting metastore schema initialization to 3.1.0
Initialization script hive-schema-3.1.0.derby.sql
```

8. **Open Hive shell by typing:**

```
Hive\bin>hive
4-08-30T23:28:56,167 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Found configuration file file:/C:/Hive/conf/hive-site.xml
e Session ID = 1c22e2df-c29b-4d8f-b0b3-63904207bbe2

ging initialized using configuration in jar:file:/C:/Hive/lib/hive-common-3.1.3.jar!/hive-log4j2.properties Async: true
4-08-30T23:28:58,690 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Created HDFS directory: /tmp/hive/user/1c22e2df-c29b-4d8f-b0b3-63904207bbe2
4-08-30T23:28:58,699 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Created local directory: C:/Users/user/AppData/Local/Temp/user/1c22e2df-c29b-4d8f-
-63904207bbe2
4-08-30T23:28:58,704 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Created HDFS directory: /tmp/hive/user/1c22e2df-c29b-4d8f-b0b3-63904207bbe2/_tmp_s
db
4-08-30T23:28:58,729 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
4-08-30T23:28:58,729 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
4-08-30 23:29:00,913 INFO DataNucleus.Persistence: Property datanucleus.cache.level2 unknown - will be ignored
4-08-30 23:29:03,726 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:03,730 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:03,731 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:03,731 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:03,732 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:03,732 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:06,794 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:06,813 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:06,819 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:06,820 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:06,820 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
4-08-30 23:29:06,821 WARN DataNucleus.MetaData: Metadata has jdbc-type of null yet this is not valid. Ignored
e Session ID = af275d1f-8ef5-4bbc-baa3-d570d2222985
e-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X rel
.
4-08-30T23:29:12,251 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] CliDriver - Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions
nsider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
4-08-30T23:29:12,273 INFO [pool-9-thread-1] org.apache.hadoop.hive.ql.session.SessionState - Created HDFS directory: /tmp/hive/user/af275d1f-8ef5-4bbc-baa3-d570d222
4-08-30T23:29:12,277 INFO [pool-9-thread-1] org.apache.hadoop.hive.ql.session.SessionState - Created local directory: C:/Users/user/AppData/Local/Temp/user/af275d1f
-4bbc-baa3-d570d2222985
4-08-30T23:29:12,284 INFO [pool-9-thread-1] org.apache.hadoop.hive.ql.session.SessionState - Created HDFS directory: /tmp/hive/user/af275d1f-8ef5-4bbc-baa3-d570d222
_tmp_space.db
```

> 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:

> CREATE DATABASE financials;

```
hive> CREATE DATABASE financials;
2024-08-30T23:35:44,067 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:35:44,068 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
OK
Time taken: 0.046 seconds
2024-08-30T23:35:44,119 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:35:44,119 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

2. Verify the database is present:

SHOW DATABASES;

```
hive> show databases;
2024-09-08 19:28:14,778 INFO conf.HiveConf: Using the default value passed in for log id: 168ffab1-50a4-42fa-b931-df2184794e2e
2024-09-08 19:28:14,966 INFO ql.Driver: Compiling command(queryId=user_20240908192814_5a0bb98c-9d8d-41ec-9ae0-4b59b629930a): show databases
2024-09-08 19:28:15,646 INFO ql.Driver: Concurrency mode is disabled, not creating a lock manager
2024-09-08 19:28:15,687 INFO ql.Driver: Semantic Analysis Completed (retrial = false)
2024-09-08 19:28:15,816 INFO ql.Driver: Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:database_name, type:string, comment:from deserializer)], properties:null)
2024-09-08 19:28:15,996 INFO exec.ListSinkOperator: Initializing operator LIST_SINK[0]
2024-09-08 19:28:16,017 INFO ql.Driver: Completed compiling command(queryId=user_20240908192814_5a0bb98c-9d8d-41ec-9ae0-4b59b629930a); Time taken: 1.137 seconds
2024-09-08 19:28:16,017 INFO reexec.ReExecDriver: Execution #1 of query
2024-09-08 19:28:16,020 INFO ql.Driver: Concurrency mode is disabled, not creating a lock manager
2024-09-08 19:28:16,020 INFO ql.Driver: Executing command(queryId=user_20240908192814_5a0bb98c-9d8d-41ec-9ae0-4b59b629930a): show databases
2024-09-08 19:28:16,047 INFO ql.Driver: Starting task [Stage-0:DDL] in serial mode          .
2024-09-08 19:28:16,050 INFO metastore.HiveMetaStore: 0: get_databases: @hive#
2024-09-08 19:28:16,050 INFO HiveMetaStore.audit: ugi=user       ip=unknown-ip-addr      cmd=get_databases: @hive#
2024-09-08 19:28:16,053 INFO exec.DDLTask: results : 3
2024-09-08 19:28:16,120 INFO ql.Driver: Completed executing command(queryId=user_20240908192814_5a0bb98c-9d8d-41ec-9ae0-4b59b629930a); Time taken: 0.1 seconds
OK
2024-09-08 19:28:16,122 INFO ql.Driver: OK
2024-09-08 19:28:16,125 INFO ql.Driver: Concurrency mode is disabled, not creating a lock manager
2024-09-08 19:28:16,146 INFO Configuration.deprecation: mapred.input.dir is deprecated. Instead, use mapreduce.input.fileinputformat.inputdir
2024-09-08 19:28:16,216 INFO mapred.FileInputFormat: Total input files to process : 1
2024-09-08 19:28:16,268 INFO exec.ListSinkOperator: RECORDS_OUT_INTERMEDIATE:0, RECORDS_OUT_OPERATOR_LIST_SINK_0:3,
default
financials
test
Time taken: 1.25 seconds, Fetched: 3 row(s)
2024-09-08 19:28:16,288 INFO CliDriver: Time taken: 1.25 seconds, Fetched: 3 row(s)
2024-09-08 19:28:16,289 INFO conf.HiveConf: Using the default value passed in for log id: 168ffab1-50a4-42fa-b931-df2184794e2e
2024-09-08 19:28:16,291 INFO session.SessionState: Resetting thread name to  main
```

3. **Switch to the new database:**

Switch to the newly created database:

USE financials;

```
                                                                                                    Resetting thread name to  main
hive> use financials;
2024-08-30T23:36:00,862 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:36:00,864 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
OK
Time taken: 0.048 seconds
2024-08-30T23:36:00,917 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:36:00,918 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

Create a Table in Hive:

Create a simple table in your database:

CREATE TABLE students_table (id INT, name STRING);

```
hive> CREATE TABLE finance_table(id INT, name STRING);
2024-08-30T23:36:32,350 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:36:32,351 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
OK
Time taken: 0.078 seconds
2024-08-30T23:36:32,439 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:36:32,440 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

Insert values into the table:

You can insert sample data into the table:

INSERT INTO finance_table VALUES (1, 'Alice'), (2,'Bob'), (3, 'Charlie');

```
hive> INSERT INTO finance_table VALUES (1,'Alice'), (2,'Bob'), (3,'Charlie');
2024-08-30T23:37:36,149 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:37:36,149 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
2024-08-30T23:37:41,265 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.common.FileUtils - Creating directory if it doesn't exist: hdfs://localh
ost:9000/user/hive/warehouse/financials.db/finance_table/.hive-staging_hive_2024-08-30_23-37-36_169_2114510206597313358-1
2024-08-30T23:37:41,461 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.common.FileUtils - Creating directory if it doesn't exist: hdfs://localh
ost:9000/tmp/hive/user/1c22e2df-c29b-4d8f-b0b3-63904207bbe2/hive_2024-08-30_23-37-41_328_1237454592308593528-1/-mr-10000/.hive-staging_hive_2024-08-30_23-37-41_328_1237
454592308593528-1
Query ID = user_20240830233736_01a5737e-9654-4013-bcdd-9071c1f42090
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>
```

```
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>
2024-08-30T23:37:42,285 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.conf.Configuration.deprecation - mapred.submit.replication is deprecated. Ins
tead, use mapreduce.client.submit.file.replication
2024-08-30T23:37:49,235 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.conf.Configuration.deprecation - yarn.resourcemanager.system-metrics-publishe
r.enabled is deprecated. Instead, use yarn.system-metrics-publisher.enabled
2024-08-30T23:37:49,791 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.conf.Configuration - resource-types.xml not found
Starting Job = job_1725040228625_0001, Tracking URL = http://DESKTOP-DKU83QM:8088/proxy/application_1725040228625_0001/
Kill Command = C:\hadoop\bin\mapred job  -kill job_1725040228625_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2024-08-30 23:38:30,410 Stage-1 map = 0%,  reduce = 0%
2024-08-30 23:38:51,438 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 6.092 sec
2024-08-30 23:39:16,522 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 11.215 sec
MapReduce Total cumulative CPU time: 11 seconds 215 msec
Ended Job = job_1725040228625_0001
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to directory hdfs://localhost:9000/user/hive/warehouse/financials.db/finance_table/.hive-staging_hive_2024-08-30_23-37-36_169_2114510206597313358-1/-ext-100
00
Loading data to table financials.finance_table
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 11.215 sec   HDFS Read: 15666 HDFS Write: 291 SUCCESS
Total MapReduce CPU Time Spent: 11 seconds 215 msec
OK
Time taken: 107.363 seconds
2024-08-30T23:39:23,519 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:39:23,519 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

Query your data:

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

CREATE VIEW myview AS SELECT name, id FROM finance_table;

```
hive> CREATE VIEW myview AS SELECT name,id FROM finance_table;
2024-08-30T23:40:04,569 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:40:04,570 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
OK
Time taken: 0.242 seconds
2024-08-30T23:40:04,817 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:40:04,817 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

List Hive Tables and Data:

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

SHOW TABLES;

```
hive> show tables;
2024-09-08 19:51:33,595 INFO conf.HiveConf: Using the default value passed in for log id: 168ffab1-50a4-42fa-b931-df2184794e2e
2024-09-08 19:51:33,597 INFO session.SessionState: Updating thread name to 168ffab1-50a4-42fa-b931-df2184794e2e main
2024-09-08 19:51:33,602 INFO ql.Driver: Compiling command(queryId=user_20240908195133_3f65224b-258c-4f26-9e5d-977da5b8b6f0): show tables
2024-09-08 19:51:33,625 INFO ql.Driver: Concurrency mode is disabled, not creating a lock manager
2024-09-08 19:51:33,631 INFO metastore.HiveMetaStore: 0: get_database: @hive#financials
2024-09-08 19:51:33,632 INFO HiveMetaStore.audit: ugi=user        ip=unknown-ip-addr        cmd=get_database: @hive#financials
2024-09-08 19:51:33,635 INFO ql.Driver: Semantic Analysis Completed (retrial = false)
2024-09-08 19:51:33,640 INFO ql.Driver: Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:tab_name, type:string, comment:from deserializer)], properties:null)
2024-09-08 19:51:33,642 INFO exec.ListSinkOperator: Initializing operator LIST_SINK[0]
2024-09-08 19:51:33,643 INFO ql.Driver: Completed compiling command(queryId=user_20240908195133_3f65224b-258c-4f26-9e5d-977da5b8b6f0); Time taken: 0.041 seconds
2024-09-08 19:51:33,644 INFO reexec.ReExecDriver: Execution #1 of query
2024-09-08 19:51:33,644 INFO ql.Driver: Concurrency mode is disabled, not creating a lock manager
2024-09-08 19:51:33,645 INFO ql.Driver: Executing command(queryId=user_20240908195133_3f65224b-258c-4f26-9e5d-977da5b8b6f0): show tables
2024-09-08 19:51:33,647 INFO ql.Driver: Starting task [Stage-0:DDL] in serial mode
2024-09-08 19:51:33,648 INFO metastore.HiveMetaStore: 0: get_database: @hive#financials
2024-09-08 19:51:33,648 INFO HiveMetaStore.audit: ugi=user        ip=unknown-ip-addr        cmd=get_database: @hive#financials
2024-09-08 19:51:33,654 INFO metastore.HiveMetaStore: 0: get_tables: db=@hive#financials pat=.*
2024-09-08 19:51:33,654 INFO HiveMetaStore.audit: ugi=user        ip=unknown-ip-addr        cmd=get_tables: db=@hive#financials pat=.*
2024-09-08 19:51:33,675 INFO ql.Driver: Completed executing command(queryId=user_20240908195133_3f65224b-258c-4f26-9e5d-977da5b8b6f0); Time taken: 0.03 seconds
OK
2024-09-08 19:51:33,676 INFO ql.Driver: OK
2024-09-08 19:51:33,677 INFO ql.Driver: Concurrency mode is disabled, not creating a lock manager
2024-09-08 19:51:33,684 INFO mapred.FileInputFormat: Total input files to process : 1
2024-09-08 19:51:33,693 INFO exec.ListSinkOperator: RECORDS_OUT_INTERMEDIATE:0, RECORDS_OUT_OPERATOR_LIST_SINK_0:2,
finance_table
myview
Time taken: 0.076 seconds, Fetched: 2 row(s)
2024-09-08 19:51:33,700 INFO CliDriver: Time taken: 0.076 seconds, Fetched: 2 row(s)
2024-09-08 19:51:33,703 INFO conf.HiveConf: Using the default value passed in for log id: 168ffab1-50a4-42fa-b931-df2184794e2e
2024-09-08 19:51:33,704 INFO session.SessionState: Resetting thread name to  main
```

To show table column names and data types, run:

DESCRIBE finance_table;

```
hive> DESCRIBE finance_table;
2024-08-30T23:40:37,803 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:40:37,804 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
OK
id                      int
name                    string
Time taken: 0.058 seconds, Fetched: 2 row(s)
2024-08-30T23:40:37,876 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:40:37,876 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

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

SELECT*FROM myview;

```
hive> SELECT*FROM myview;
2024-08-30T23:40:19,505 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:40:19,505 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
2024-08-30T23:40:19,681 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.common.FileUtils - Creating directory if it doesn't exist: hdfs://localh
ost:9000/tmp/hive/user/1c22e2df-c29b-4d8f-b0b3-63904207bbe2/hive_2024-08-30_23-40-19_523_4767235147659174033-1/-mr-10001/.hive-staging_hive_2024-08-30_23-40-19_523_4767
235147659174033-1
OK
Alice   1
Bob     2
Charlie 3
Time taken: 0.217 seconds, Fetched: 3 row(s)
2024-08-30T23:40:19,750 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:40:19,750 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

Alter a Table:

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

ALTER TABLE finance_table ADD COLUMNS (age INT);

```
hive> ALTER TABLE finance_table ADD COLUMNS (age INT);
2024-08-30T23:41:04,648 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:41:04,648 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
OK
Time taken: 0.158 seconds
2024-08-30T23:41:04,808 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:41:04,808 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
hive> DESCRIBE finance_table;
2024-08-30T23:41:48,873 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:41:48,873 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
OK
id                      int
name                    string
age                     int
Time taken: 0.048 seconds, Fetched: 3 row(s)
2024-08-30T23:41:48,939 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c
29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:41:48,939 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Resetting thread name to  main
```

Quit Hive:

To exit the Hive CLI, simply type:

quit;

```
hive> quit;
2024-08-30T23:41:53,103 INFO [main] org.apache.hadoop.hive.conf.HiveConf - Using the default value passed in for log id: 1c22e2df-c29b-4d8f-b0b3-63904207bbe2
2024-08-30T23:41:53,104 INFO [main] org.apache.hadoop.hive.ql.session.SessionState - Updating thread name to 1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main
2024-08-30T23:41:53,119 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Deleted directory: /tmp/hive/user/1c22e2df-c29
b-4d8f-b0b3-63904207bbe2 on fs with scheme hdfs
2024-08-30T23:41:53,120 INFO [1c22e2df-c29b-4d8f-b0b3-63904207bbe2 main] org.apache.hadoop.hive.ql.session.SessionState - Deleted directory: C:/Users/user/AppData/Local
/Temp/user/1c22e2df-c29b-4d8f-b0b3-63904207bbe2 on fs with scheme file

C:\Hive\bin>_
```

RESULT:

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