

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

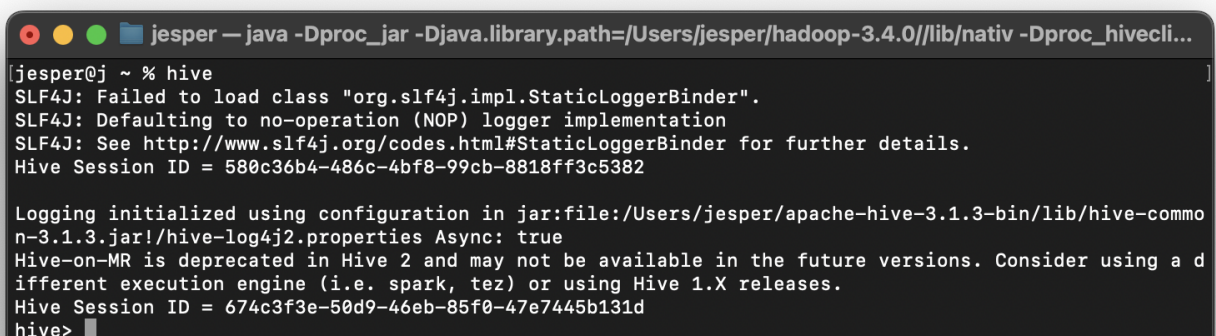
Aim:

To create tables in hive and write queries to access the data in the table using Apache Hive and Hadoop.

Procedure:

1. Start Hadoop using start-all.sh command
2. Run the SQL Server in the Settings in Mac.
3. Run hive using the command ./hive
4. Create a database and use the database.
5. Create a table and define its structure.
6. Import the data into the table from HDFS.
7. Query the data using SQL commands.

Output:



```
jesper — java -Dproc_jar -Djava.library.path=/Users/jesper/hadoop-3.4.0/lib/nativ -Dproc_hivecli...
[jesper@j ~ % hive
SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.
Hive Session ID = 580c36b4-486c-4bf8-99cb-8818ff3c5382

Logging initialized using configuration in jar:file:/Users/jesper/apache-hive-3.1.3-bin/lib/hive-commo
n-3.1.3.jar!/hive-log4j2.properties Async: true
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a d
ifferent execution engine (i.e. spark, tez) or using Hive 1.X releases.
Hive Session ID = 674c3f3e-50d9-46eb-85f0-47e7445b131d
hive>
```

```
jesper — java -Dproc_jar -Djava.library.path=/Users/jesper/hadoop-3.4.0/lib/nativ -Dproc_hiv...
default
Time taken: 0.578 seconds, Fetched: 1 row(s)
[hive> create table student(rollno int, name varchar(20), mark int);
OK
Time taken: 0.57 seconds
```

```
jesper — java -Dproc_jar -Djava.library.path=/Users/jesper/hadoop-3.4.0/lib/nativ -Dproc_hiv...
[hive> insert into student values (
[ > 101, 'joe', 99),
[ > (102, 'mike', 45),
[ > (103, 'jake', 80),
[ > (104, 'jack', 60);
Query ID = jesper_20240905094644_2be800b2-de27-4d6b-9532-142979f8b1d5
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_1725509520955_0001, Tracking URL = http://localhost:8088/proxy/application_17255
09520955_0001/
Kill Command = /Users/jesper/hadoop-3.4.0/bin/mapred job -kill job_1725509520955_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2024-09-05 09:46:56,970 Stage-1 map = 0%, reduce = 0%
2024-09-05 09:47:03,211 Stage-1 map = 100%, reduce = 0%
2024-09-05 09:47:07,343 Stage-1 map = 100%, reduce = 100%
Ended Job = job_1725509520955_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://0.0.0.0:9000/user/hive/warehouse/student/.hive-staging_hive_2024-09
-05_09-46-44_400_3461119791769901288-1/-ext-10000
Loading data to table default.student
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 HDFS Read: 17186 HDFS Write: 379 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
Time taken: 26.763 seconds
hive> █
```

```
jesper — java -Dproc_jar -Djava.library.path=/Users/jesper/hadoop-3.4.0/lib/nativ -Dproc_hiv...
[hive> select * from student;
OK
101      joe      99
102      mike     45
103      jake     80
104      jack     60
Time taken: 0.135 seconds, Fetched: 4 row(s)
hive> █
```

Result:

Thus the Installation, Configuration and run Hadoop and HDFS is successfully executed.