

NAME:-Swapnil Ananda Badave
Roll No:-33303

Code on terminal :-

```
hive> CREATE EXTERNAL TABLE table1 (  
  > id INT,  
  > name STRING,  
  > age INT  
  > )  
  > STORED AS TEXTFILE  
  > LOCATION '/home/cloudera/Desktop/table1_data1';
```

OK

Time taken: 2.173 seconds

hive>

```
>  
[2]+ Stopped          hive  
[cloudera@quickstart Desktop]$ hive
```

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.

```
hive> CREATE EXTERNAL TABLE table2 (  
  > id INT,  
  > name STRING,  
  > age INT  
  > )  
  > STORED AS TEXTFILE  
  > LOCATION '/home/cloudera/Desktop/table1_data1';
```

OK

Time taken: 0.778 seconds

```
hive> CREATE EXTERNAL TABLE table2 (  
  > id INT,  
  > name STRING,  
  > age INT  
  > )  
  > STORED AS TEXTFILE  
  > LOCATION '/home/cloudera/Desktop/1/table1_data2';
```

FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.DDLTask.
AlreadyExistsException(message:Table table2 already exists)

```
hive> INSERT INTO TABLE table1 VALUES (1, 'Swap', 25);
```

Query ID = cloudera_20250325233232_29bacff7-055d-417c-9b42-9ef3f2d15c65

Total jobs = 3

Launching Job 1 out of 3

Number of reduce tasks is set to 0 since there's no reduce operator

Starting Job = job_1742968970641_0001, Tracking URL =

http://quickstart.cloudera:8088/proxy/application_1742968970641_0001/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1742968970641_0001

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0

2025-03-25 23:32:28,992 Stage-1 map = 0%, reduce = 0%

2025-03-25 23:32:41,285 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.84 sec

MapReduce Total cumulative CPU time: 1 seconds 840 msec

Ended Job = job_1742968970641_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:

hdfs://quickstart.cloudera:8020/home/cloudera/Desktop/table1_data1/.hive-staging_hive_2025-03-25_23-32-07_244_7349867147684068655-1/-ext-10000

Loading data to table default.table1

Table default.table1 stats: [numFiles=1, numRows=1, totalSize=10, rawDataSize=9]

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Cumulative CPU: 1.84 sec HDFS Read: 3690 HDFS Write: 80 SUCCESS

Total MapReduce CPU Time Spent: 1 seconds 840 msec

OK

Time taken: 36.078 seconds

hive> select * from table1

>

> select * from table1

> INSERT INTO TABLE table1 VALUES (2, 'Sarvesh', 30);

FAILED: ParseException line 3:0 missing EOF at 'select' near 'table1'

hive> INSERT INTO TABLE table1 VALUES (2, 'Sarvesh', 30);

Query ID = cloudera_20250325233333_8ea1ab66-dd28-4ee5-9b44-69e2cd167ac1

Total jobs = 3

Launching Job 1 out of 3

Number of reduce tasks is set to 0 since there's no reduce operator

Starting Job = job_1742968970641_0002, Tracking URL =

http://quickstart.cloudera:8088/proxy/application_1742968970641_0002/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1742968970641_0002

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0

2025-03-25 23:33:39,358 Stage-1 map = 0%, reduce = 0%

2025-03-25 23:33:50,340 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.84 sec

MapReduce Total cumulative CPU time: 1 seconds 840 msec

Ended Job = job_1742968970641_0002

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:

hdfs://quickstart.cloudera:8020/home/cloudera/Desktop/table1_data1/.hive-staging_hive_2025-03-25_23-33-28_851_5010771986281567764-1/-ext-10000

Loading data to table default.table1

Table default.table1 stats: [numFiles=2, numRows=2, totalSize=23, rawDataSize=21]

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Cumulative CPU: 1.84 sec HDFS Read: 3781 HDFS Write: 83 SUCCESS

Total MapReduce CPU Time Spent: 1 seconds 840 msec

OK

Time taken: 22.838 seconds

hive> INSERT INTO TABLE table1 VALUES (3, 'Shreyash', 35);

Query ID = cloudera_20250325233333_dcb04844-823e-4084-bc9b-0edaef18c2ff

Total jobs = 3

Launching Job 1 out of 3

Number of reduce tasks is set to 0 since there's no reduce operator

Starting Job = job_1742968970641_0003, Tracking URL =

http://quickstart.cloudera:8088/proxy/application_1742968970641_0003/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1742968970641_0003

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0

2025-03-25 23:34:09,919 Stage-1 map = 0%, reduce = 0%

2025-03-25 23:34:19,554 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.68 sec

MapReduce Total cumulative CPU time: 1 seconds 680 msec
Ended Job = job_1742968970641_0003
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:
hdfs://quickstart.cloudera:8020/home/cloudera/Desktop/table1_data1/.hive-staging_hive_2025-03-25_23-33-59_784_7161433309336397906-1/-ext-10000
Loading data to table default.table1
Table default.table1 stats: [numFiles=3, numRows=3, totalSize=37, rawDataSize=34]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 1.68 sec HDFS Read: 3782 HDFS Write: 84 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 680 msec
OK
Time taken: 21.11 seconds
hive> CREATE EXTERNAL TABLE merged_table (
 > id INT,
 > name STRING,
 > age INT
 >)
 > STORED AS TEXTFILE
 > LOCATION '/home/cloudera/Desktop/1/merged_table_data';
OK
Time taken: 0.055 seconds
hive> INSERT INTO TABLE merged_table
 > SELECT id, name, age FROM table1
 > UNION ALL
 > SELECT id, name, age FROM table2;
Query ID = cloudera_20250325233535_4888ebfd-7e69-4857-aa6a-4a5f1ac67ab0
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1742968970641_0004, Tracking URL =
http://quickstart.cloudera:8088/proxy/application_1742968970641_0004/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1742968970641_0004
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2025-03-25 23:35:44,080 Stage-1 map = 0%, reduce = 0%
2025-03-25 23:35:52,831 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.46 sec
MapReduce Total cumulative CPU time: 1 seconds 460 msec
Ended Job = job_1742968970641_0004
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:
hdfs://quickstart.cloudera:8020/home/cloudera/Desktop/1/merged_table_data/.hive-staging_hive_2025-03-25_23-35-34_000_2424193354907639986-1/-ext-10000
Loading data to table default.merged_table
Table default.merged_table stats: [numFiles=1, numRows=6, totalSize=74, rawDataSize=68]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 1.46 sec HDFS Read: 4480 HDFS Write: 150 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 460 msec
OK
Time taken: 21.231 seconds

OUTPUT:-

```
hive> SELECT * FROM merged_table;
OK
1 Swap 25
1 Swap 25
2 Sarvesh 30
2 Sarvesh 30
3 Shreyash 35
3 Shreyash 35
Time taken: 0.099 seconds, Fetched: 6 row(s)
hive> SELECT * FROM merged_table;
OK
1 Swap 25
1 Swap 25
2 Sarvesh 30
2 Sarvesh 30
3 Shreyash 35
3 Shreyash 35
Time taken: 0.086 seconds, Fetched: 6 row(s)
hive> SELECT * FROM table2;
OK
1 Swap 25
2 Sarvesh 30
3 Shreyash 35
Time taken: 0.067 seconds, Fetched: 3 row(s)
hive> SELECT * FROM table1;
OK
1 Swap 25
2 Sarvesh 30
```