```
hadoop@ritadmin:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [ritadmin]
Starting resourcemanager
Starting nodemanagers
hadoop@ritadmin:~$ jps
4626 SecondaryNameNode
4916 ResourceManager
5269 NodeManager
4376 DataNode
4090 NameNode
5550 Jps
hive> create database std;
OK
Time taken: 0.026 seconds
hive> use std
    > ;
OK
Time taken: 0.011 seconds
hive> CREATE TABLE student (
    >
    >   id INT,
    >
    >   name STRING,
    >
    >   age INT,
    >
    >   department STRING,
    >
    >   grade STRING
    >
    > )
    >
    > ROW FORMAT DELIMITED
    >
    > FIELDS TERMINATED BY ','
    >
    > STORED AS TEXTFILE;
OK
Time taken: 0.026 seconds
hive> desc student;
OK
id              int
name                    string
age             int
department              string
grade                   string
Time taken: 0.019 seconds, Fetched: 5 row(s)
```

```
hive> LOAD DATA INPATH '/user/hive/warehouse/student_data/students.txt' INTO TABLE
student;
Loading data to table std.student
OK
Time taken: 0.084 seconds
hive> select * from student;
OK
1       Don     10      CY      A
2       Alien   11      ISE     B
3       Bob     12      ECE     A
4       Rome    13      ME      B
5       Wick    15      CSE     A
Time taken: 0.051 seconds, Fetched: 5 row(s)
hive> create database emp;
OK
Time taken: 0.021 seconds
hive> use emp;
OK
Time taken: 0.012 seconds
hive> CREATE TABLE student (
    >   id INT,
    >
    >   name STRING,
    >
    >   age INT,
    >
    >   department STRING,
    >
    >   grade STRING
    >
    > )
    >
    > ROW FORMAT DELIMITED
    >
    > FIELDS TERMINATED BY '\t'
    >
    > STORED AS TEXTFILE;
OK
Time taken: 0.031 seconds
hive> LOAD DATA INPATH '/user/hive/warehouse/student_data/students.txt'
    > INTO TABLE student;
Loading data to table emp.student
OK
Time taken: 0.076 seconds
hive> select * from student;
OK
1       Don     10      CY      A
2       Alien   11      ISE     B
3       Bob     12      ECE     A
4       Rome    13      ME      B
5       Wick    15      CSE     A
Time taken: 0.049 seconds, Fetched: 5 row(s)
```

hive>

```
OK
id                      int
name                    string
age                     int
department              string
grade                   string
Time taken: 0.019 seconds, Fetched: 5 row(s)
hive> LOAD DATA INPATH '/user/hive/warehouse/student_data/students.txt' INTO TABLE student;
FAILED: SemanticException Line 1:17 Invalid path ''/user/hive/warehouse/student_data/students.txt'': No files matching path hdfs://localhost:9000/user/hive/warehouse/student_data/students.txt
hive> LOAD DATA INPATH '/user/hive/warehouse/student_data/students.txt' INTO TABLE student;
Loading data to table std.student
OK
Time taken: 0.084 seconds
hive> select * from student;
OK
1       Don     10      CY      A
2       Alien   11      ISE     B
3       Bob     12      ECE     A
4       Rome    13      ME      B
5       Wick    15      CSE     A
Time taken: 0.049 seconds, Fetched: 5 row(s)
hive> select * from student;
OK
1       Don     10      CY      A
2       Alien   11      ISE     B
3       Bob     12      ECE     A
4       Rome    13      ME      B
5       Wick    15      CSE     A
Time taken: 0.051 seconds, Fetched: 5 row(s)
hive> create database emp;
OK
Time taken: 0.021 seconds
hive> use database emp;
NoViableAltException(84@[])
        at org.apache.hadoop.hive.ql.parse.HiveParser_IdentifiersParser.identifier(HiveParser_IdentifiersParser.java:11910)
        at org.apache.hadoop.hive.ql.parse.HiveParser.identifier(HiveParser.java:45072)
        at org.apache.hadoop.hive.ql.parse.HiveParser.switchDatabaseStatement(HiveParser.java:6003)
        at org.apache.hadoop.hive.ql.parse.HiveParser.ddlStatement(HiveParser.java:4267)
        at org.apache.hadoop.hive.ql.parse.HiveParser.execStatement(HiveParser.java:2494)
        at org.apache.hadoop.hive.ql.parse.HiveParser.statement(HiveParser.java:1420)
        at org.apache.hadoop.hive.ql.parse.ParseDriver.parse(ParseDriver.java:220)
        at org.apache.hadoop.hive.ql.parse.ParseUtils.parse(ParseUtils.java:74)
        at org.apache.hadoop.hive.ql.parse.ParseUtils.parse(ParseUtils.java:67)
        at org.apache.hadoop.hive.ql.Driver.compile(Driver.java:616)
        at org.apache.hadoop.hive.ql.Driver.compileInternal(Driver.java:1826)
        at org.apache.hadoop.hive.ql.Driver.compileAndRespond(Driver.java:1773)
        at org.apache.hadoop.hive.ql.Driver.compileAndRespond(Driver.java:1768)
        at org.apache.hadoop.hive.ql.reexec.ReExecDriver.compileAndRespond(ReExecDriver.java:126)
        at org.apache.hadoop.hive.ql.reexec.ReExecDriver.run(ReExecDriver.java:214)
        at org.apache.hadoop.hive.cli.CliDriver.processLocalCmd(CliDriver.java:239)
        at org.apache.hadoop.hive.cli.CliDriver.processCmd(CliDriver.java:188)
        at org.apache.hadoop.hive.cli.CliDriver.processLine(CliDriver.java:402)
        at org.apache.hadoop.hive.cli.CliDriver.executeDriver(CliDriver.java:821)
        at org.apache.hadoop.hive.cli.CliDriver.run(CliDriver.java:759)
        at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:683)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:328)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:241)
FAILED: ParseException line 1:4 cannot recognize input near 'database' 'emp' '<EOF>' in switch database statement
hive> use emp;
OK
Time taken: 0.012 seconds
hive> CREATE TABLE student (
    >   id INT,
    >
    >   name STRING,
    >
    >   age INT,
    >
    >   department STRING,
    >
    >   grade STRING
    >
    > )
    >
    > ROW FORMAT DELIMITED
    >
    > FIELDS TERMINATED BY '\t'
    >
    > STORED AS TEXTFILE;
OK
Time taken: 0.031 seconds
hive> LOAD DATA INPATH '/user/hive/warehouse/student_data/students.txt'
    > INTO TABLE student
    > FIELDS TERMINATED BY '\t';
FAILED: ParseException line 3:0 missing EOF at 'FIELDS' near 'student'
hive> LOAD DATA INPATH '/user/hive/warehouse/student_data/students.txt'
    > INTO TABLE student;
FAILED: SemanticException Line 1:17 Invalid path ''/user/hive/warehouse/student_data/students.txt'': No files matching path hdfs://localhost:9000/user/hive/warehouse/student_data/students.txt
hive> LOAD DATA INPATH '/user/hive/warehouse/student_data/students.txt'
    > INTO TABLE student;
Loading data to table emp.student
OK
Time taken: 0.076 seconds
hive> select * from student;
OK
1       Don     10      CY      A
2       Alien   11      ISE     B
3       Bob     12      ECE     A
4       Rome    13      ME      B
5       Wick    15      CSE     A
Time taken: 0.049 seconds, Fetched: 5 row(s)
hive>
```

```
Stage-Stage-3: Map: 1    Cumulative CPU: 1.31 sec    HDFS Read: 9481 HDFS Write: 166 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 310 msec
OK
Alice    Engineering
Bob      HR
Charlie Marketing
Time taken: 14.214 seconds, Fetched: 3 row(s)
hive> select * from employees;
OK
1        Alice    Engineering      70000.0
2        Bob      HR        50000.0
3        Charlie Marketing          60000.0
Time taken: 0.053 seconds, Fetched: 3 row(s)
hive> select department, AVG(salary) as avg_salary
    > from employees
    > group by department;
Query ID = hadoop_20250407121049_2cb0a220-4e2b-4049-91f4-46939b401e38
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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_1744005325492_0006, Tracking URL = http://ritadmin:8088/proxy/applicatio
n_1744005325492_0006/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744005325492_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2025-04-07 12:10:55,099 Stage-1 map = 0%,  reduce = 0%
2025-04-07 12:10:58,144 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.55 sec
2025-04-07 12:11:02,194 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.95 sec
MapReduce Total cumulative CPU time: 2 seconds 950 msec
Ended Job = job_1744005325492_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.95 sec   HDFS Read: 15645 HDFS Write:
172 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 950 msec
OK
Engineering      70000.0
HR       50000.0
Marketing        60000.0
Time taken: 14.002 seconds, Fetched: 3 row(s)
hive>
```

```
Ended Job = job_1744005325492_0004
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1    Cumulative CPU: 1.31 sec    HDFS Read: 9766 HDFS Write: 87 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 310 msec
OK
Time taken: 12.494 seconds
hive> select e.name, d.dept_name
    > from employees e
    > join departments d on (e.department=d.dept_name);
Query ID = hadoop_20250407120614_046ad41d-5d70-4934-a7ac-2ee81ed58ff8
Total jobs = 1
2025-04-07 12:06:17     Dump the side-table for tag: 1 with group count: 3 into file: file:
/tmp/hadoop/6f758721-de32-4b54-a969-8a6f3c1ec871/hive_2025-04-07_12-06-14_450_3961097237014
970895-1/-local-10004/HashTable-Stage-3/MapJoin-mapfile11--.hashtable
2025-04-07 12:06:17     Uploaded 1 File to: file:/tmp/hadoop/6f758721-de32-4b54-a969-8a6f3c
1ec871/hive_2025-04-07_12-06-14_450_3961097237014970895-1/-local-10004/HashTable-Stage-3/Ma
pJoin-mapfile11--.hashtable (336 bytes)
2025-04-07 12:06:17     End of local task; Time Taken: 0.648 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1744005325492_0005, Tracking URL = http://ritadmin:8088/proxy/applicatio
n_1744005325492_0005/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744005325492_0005
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2025-04-07 12:06:23,564 Stage-3 map = 0%,  reduce = 0%
2025-04-07 12:06:27,625 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.31 sec
MapReduce Total cumulative CPU time: 1 seconds 310 msec
Ended Job = job_1744005325492_0005
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1    Cumulative CPU: 1.31 sec    HDFS Read: 9481 HDFS Write: 166 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 310 msec
OK
Alice    Engineering
Bob      HR
Charlie Marketing
Time taken: 14.214 seconds, Fetched: 3 row(s)
hive> select * from employees;
OK
1        Alice    Engineering      70000.0
2        Bob      HR        50000.0
3        Charlie Marketing          60000.0
Time taken: 0.053 seconds, Fetched: 3 row(s)
hive>
```

```
n_1744005325492_0004/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744005325492_0004
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2025-04-07 12:05:38,069 Stage-3 map = 0%,  reduce = 0%
2025-04-07 12:05:41,111 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.31 sec
MapReduce Total cumulative CPU time: 1 seconds 310 msec
Ended Job = job_1744005325492_0004
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 1.31 sec   HDFS Read: 9766 HDFS Write: 87 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 310 msec
OK
Time taken: 12.494 seconds
hive> select e.name, d.dept_name
    > from employees e
    > join departments d on (e.department=d.dept_name);
Query ID = hadoop_20250407120614_046ad41d-5d70-4934-a7ac-2ee81ed58ff8
Total jobs = 1
2025-04-07 12:06:17     Dump the side-table for tag: 1 with group count: 3 into file: file:
/tmp/hadoop/6f758721-de32-4b54-a969-8a6f3c1ec871/hive_2025-04-07_12-06-14_450_3961097237014
970895-1/-local-10004/HashTable-Stage-3/MapJoin-mapfile11--.hashtable
2025-04-07 12:06:17     Uploaded 1 File to: file:/tmp/hadoop/6f758721-de32-4b54-a969-8a6f3c
1ec871/hive_2025-04-07_12-06-14_450_3961097237014970895-1/-local-10004/HashTable-Stage-3/Ma
pJoin-mapfile11--.hashtable (336 bytes)
2025-04-07 12:06:17     End of local task; Time Taken: 0.648 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1744005325492_0005, Tracking URL = http://ritadmin:8088/proxy/applicatio
n_1744005325492_0005/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744005325492_0005
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2025-04-07 12:06:23,564 Stage-3 map = 0%,  reduce = 0%
2025-04-07 12:06:27,625 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.31 sec
MapReduce Total cumulative CPU time: 1 seconds 310 msec
Ended Job = job_1744005325492_0005
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 1.31 sec   HDFS Read: 9481 HDFS Write: 166 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 310 msec
OK
Alice    Engineering
Bob      HR
Charlie Marketing
Time taken: 14.214 seconds, Fetched: 3 row(s)
hive> 
```

```
hadoop@ritadmin:~$ hive
Hive Session ID = 6f758721-de32-4b54-a969-8a6f3c1ec871

Logging initialized using configuration in jar:file:/home/hadoop/hive/lib/hive-common-3.1.2
.jar!/hive-log4j2.properties Async: true
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Conside
r using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Hive Session ID = 74817032-b578-4a35-94b9-fe0711956935
hive> use rit
    > ;
OK
Time taken: 0.167 seconds
hive> SHOW TABLES;
OK
employees
Time taken: 0.078 seconds, Fetched: 1 row(s)
hive> SHOW TABLES LIKE 'department';
OK
Time taken: 0.032 seconds
hive> CREATE TABLE departments(
    > id INT,
    > dept_name STRING
    > )
    > ROW FORMAT DELIMITED
    > FIELDS TERMINATED BY ','
    > STORED AS TEXTFILE;
OK
Time taken: 0.21 seconds
hive> INSERT INTO TABLE departments VALUES
    > (1, 'Engineering'),
    > (2,'HR'),
    > (3,'Marketing');
Query ID = hadoop_20250407115827_2abfb02d-1e92-4472-a628-1cc11bc4c064
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_1744005325492_0003, Tracking URL = http://ritadmin:8088/proxy/applicatio
n_1744005325492_0003/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744005325492_0003
```

```
2025-04-07 12:30:40,628 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.46 sec
2025-04-07 12:30:44,683 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.66 sec
MapReduce Total cumulative CPU time: 2 seconds 660 msec
Ended Job = job_1744005325492_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.66 sec   HDFS Read: 15087 HDFS Write:
153 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 660 msec
OK
1       70000.0
1       50000.0
1       60000.0
Time taken: 12.572 seconds, Fetched: 3 row(s)
hive> select count(name)as no_employees,department, MAX(salary) max_salary
    > from employees
    > group by department;
Query ID = hadoop_20250407123103_e07f2dee-9521-467c-bd3b-6a1abf554e86
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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_1744005325492_0009, Tracking URL = http://ritadmin:8088/proxy/applicatio
n_1744005325492_0009/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744005325492_0009
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2025-04-07 12:31:08,286 Stage-1 map = 0%,  reduce = 0%
2025-04-07 12:31:12,337 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.57 sec
2025-04-07 12:31:16,384 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.7 sec
MapReduce Total cumulative CPU time: 2 seconds 700 msec
Ended Job = job_1744005325492_0009
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.7 sec   HDFS Read: 15162 HDFS Write: 1
78 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 700 msec
OK
1       Engineering     70000.0
1       HR      50000.0
1       Marketing       60000.0
Time taken: 13.986 seconds, Fetched: 3 row(s)
hive>
```

```
hadoop@ritadmin:~$ hive
Hive Session ID = 6f4c730f-df35-486a-8a76-238854cb4172

Logging initialized using configuration in jar:file:/home/hadoop/hive/lib/hive-common-3.1.2.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 different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Hive Session ID = d43c154b-98c1-4cbc-baea-f141b32df7d9
```

```
hive> select dept, count(*) from students group by dept;
Query ID = hadoop_20250409095631_3965b498-5b61-4936-9a18-81ad4aa56627
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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_1744170803527_0007, Tracking URL = http://ritadmin:8088/proxy/application_1744170803527_0007/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744170803527_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2025-04-09 09:56:36,935 Stage-1 map = 0%,  reduce = 0%
2025-04-09 09:56:39,979 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.52 sec
2025-04-09 09:56:44,032 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.7 sec
MapReduce Total cumulative CPU time: 2 seconds 700 msec
Ended Job = job_1744170803527_0007
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.7 sec   HDFS Read: 13808 HDFS Write: 218 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 700 msec
OK
Computer Science        2
Electrical Engineering  1
Electronics     1
Mechanical Engineering  1
```

```
hive> select dept, avg(age) from students group by dept;
Query ID = hadoop_20250409095708_6542f34e-a44b-46a2-8b68-b008d2c84817
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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_1744170803527_0008, Tracking URL = http://ritadmin:8088/proxy/application_1744170803527_0008/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1744170803527_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2025-04-09 09:57:13,102 Stage-1 map = 0%,  reduce = 0%
2025-04-09 09:57:16,143 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.58 sec
2025-04-09 09:57:20,188 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.02 sec
MapReduce Total cumulative CPU time: 3 seconds 20 msec
Ended Job = job_1744170803527_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.02 sec   HDFS Read: 15900 HDFS Write: 230 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 20 msec
OK
Computer Science        21.0
Electrical Engineering  21.0
Electronics      20.0
Mechanical Engineering  19.0
Time taken: 12.77 seconds, Fetched: 4 row(s)
```