Outputs :

```
hdp123@stud-OptiPlex-3060:/usr/local/hive/bin$ ./hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.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]

Logging initialized using configuration in jar:file:/usr/local/hive/lib/hive-common-2.3.9.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, t
ez) or using Hive 1.X releases.
hive> show databases;
OK
d1
default
Time taken: 4.003 seconds, Fetched: 2 row(s)
hive> create database d157;
OK
Time taken: 0.198 seconds
hive> show databases
    > hdp123@stud-OptiPlex-3060:/usr/local/hive/bin$
hdp123@stud-OptiPlex-3060:/usr/local/hive/bin$ ./hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.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]

Logging initialized using configuration in jar:file:/usr/local/hive/lib/hive-common-2.3.9.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, t
ez) or using Hive 1.X releases.
hive> show databases;
OK
d1
d157
default
```

```
hive> use d157;
OK
Time taken: 0.015 seconds
hive> create table flight(Year int, Month int, DayofMonth int, DayOfWeek int, DepTime  int, CRSDepTime int, ArrTime int, CRSArrTime int, Uniqu
eCarrier string, FlightNum int, TailNum string, ActualElapsedTime int, CRSElapsedTime int, AirTime int, ArrDelay int, DepDelay int, Origin str
ing, Dest string, Distance int, TaxiIn int, TaxiOut int, Cancelled int, CancellationCode string, Diverted string, CarrierDelay int, WeatherDel
ay int, NASDelay int,  SecurityDelay int, LateAircraftDelay int) row format delimited fields terminated by ',';
OK
Time taken: 0.854 seconds
hive> show tables;
OK
flight
Time taken: 0.035 seconds, Fetched: 1 row(s)
hive> load data local inpath '/home/stud/Downloads/flight_data.csv' into table flight;
Loading data to table d157.flight
OK
Time taken: 1.057 seconds
hive> create table flight_fare( FlightNum int,  Fare int) row format delimited fields terminated by ',';
OK
Time taken: 0.077 seconds
hive> show tables;
OK
flight
flight_fare
Time taken: 0.021 seconds, Fetched: 2 row(s)
hive> load data local inpath '/home/stud/Downloads/flight_fare.csv' into table flight_fare;
Loading data to table d157.flight_fare
OK
Time taken: 0.637 seconds
hive> select flight.Origin, flight.Dest, flight_fare.fare from flight join flight_fare on flight.FlightNum=flight_fare.FlightNum;
WARNING: 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.
```

```
hive> select flight.Origin, flight.Dest, flight_fare.fare from flight join flight_fare on flight.FlightNum=flight_fare.FlightNum;
WARNING: 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.
Query ID = hdp123_20230503153024_10d2b95e-6f57-44a3-9ba4-ec5a14582ca7
Total jobs = 1
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.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]
2023-05-03 15:30:30     Starting to launch local task to process map join;      maximum memory = 477626368
2023-05-03 15:30:30     Dump the side-table for tag: 1 with group count: 47 into file: file:/tmp/hdp123/170595e9-39e7-4c42-bd8d-55926809ca68/h
ive_2023-05-03_15-30-24_454_8304525915922042190-1/-local-10004/HashTable-Stage-3/MapJoin-mapfile01--.hashtable
2023-05-03 15:30:30     Uploaded 1 File to: file:/tmp/hdp123/170595e9-39e7-4c42-bd8d-55926809ca68/hive_2023-05-03_15-30-24_454_830452591592204
2190-1/-local-10004/HashTable-Stage-3/MapJoin-mapfile01--.hashtable (1357 bytes)
2023-05-03 15:30:30     End of local task; Time Taken: 0.734 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_1683107607090_0001, Tracking URL = http://stud-OptiPlex-3060:8088/proxy/application_1683107607090_0001/
Kill Command = /usr/local/hadoop/bin/hadoop job  -kill job_1683107607090_0001
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2023-05-03 15:30:38,061 Stage-3 map = 0%,  reduce = 0%
2023-05-03 15:30:42,150 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.93 sec
MapReduce Total cumulative CPU time: 1 seconds 930 msec
Ended Job = job_1683107607090_0001
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 1.93 sec   HDFS Read: 13124 HDFS Write: 1317 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 930 msec
OK
SEA     SJC     5181
SEA     PSP     7905
SAN     SEA     8781
```

```
SEA     PSP     7905
SAN     SEA     8781
SEA     GEG     7852
TUS     SEA     9009
LAX     SEA     6250
LAX     SEA     8535
ANC     PDX     7141
LAS     SEA     9184
SJC     SEA     7997
SEA     DEN     9245
ANC     SEA     5897
SEA     TUS     8321
SEA     SAN     8446
SEA     SAN     9338
SEA     ANC     5785
SEA     LAS     5807
SEA     LAX     6358
SEA     PSP     5434
LAX     SEA     9479
LAX     SEA     7158
ANC     SEA     6297
DFW     SEA     5715
SEA     LAX     8082
DEN     SEA     6825
SFO     SEA     9186
GEG     SEA     7127
ANC     FAI     6404
SMF     SEA     7503
SEA     PSP     7247
SAN     PDX     6992
LAX     SEA     9505
DEN     SEA     9582
SFO     PSP     7465
SEA     SNA     9400
```

```
stud@stud-OptiPlex-3060:~$ su hdp123
Password:
hdp123@stud-OptiPlex-3060:/home/stud$ cd
hdp123@stud-OptiPlex-3060:~$ hdfs dfs -mkdir /user/hive/flight157
23/05/03 15:31:40 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where appli
cable
hdp123@stud-OptiPlex-3060:~$ hdfs dfs -put /home/stud/Downloads/flight_ext.csv /user/hive/flight157
23/05/03 15:32:14 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where appli
cable
hdp123@stud-OptiPlex-3060:~$ hdfs dfs -ls /user/hive/flight157
23/05/03 15:32:25 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where appli
cable
Found 1 items
-rw-r--r--   1 hdp123 supergroup       3842 2023-05-03 15:32 /user/hive/flight157/flight_ext.csv
```

```
Time taken: 0.024 seconds, Fetched: 2 row(s)
hive> create external table flight_ext(Year int, Month int, DayofMonth int, DayOfWeek int, DepTime  int, CRSDepTime int, ArrTime int, CRSArrTi
me int, UniqueCarrier string, FlightNum int, TailNum string, ActualElapsedTime int, CRSElapsedTime int, AirTime int, ArrDelay int, DepDelay in
t, Origin string, Dest string, Distance int, TaxiIn int, TaxiOut int, Cancelled int, CancellationCode string, Diverted string, CarrierDelay in
t, WeatherDelay int, NASDelay int,  SecurityDelay int, LateAircraftDelay int) row format delimited fields terminated by ',' location '/user/hi
ve/flight157';
OK
Time taken: 0.301 seconds
hive> show tables;
OK
flight
flight_ext
flight_fare
Time taken: 0.023 seconds, Fetched: 3 row(s)
hive> describe formatted table_name;
FAILED: SemanticException [Error 10001]: Table not found table_name
hive> describe formatted flight_ext;
OK
# col_name              data_type               comment

year                    int
month                   int
dayofmonth              int
dayofweek               int
deptime                 int
crsdeptime              int
arrtime                 int
crsarrtime              int
uniquecarrier           string
flightnum               int
tailnum                 string
actualelapsedtime       int
crselapsedtime          int
airtime                 int
arrdelay                int
depdelay                int
```

```
hive> select count(*) from flight_ext;
WARNING: 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.
Query ID = hdp123_20230503153750_23481d31-4331-46a5-ba59-32d08a50fa2f
Total jobs = 1
Launching Job 1 out of 1
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_1683107607090_0002, Tracking URL = http://stud-OptiPlex-3060:8088/proxy/application_1683107607090_0002/
Kill Command = /usr/local/hadoop/bin/hadoop job  -kill job_1683107607090_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-05-03 15:37:57,500 Stage-1 map = 0%,  reduce = 0%
2023-05-03 15:38:01,593 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.21 sec
2023-05-03 15:38:05,650 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.53 sec
MapReduce Total cumulative CPU time: 2 seconds 530 msec
Ended Job = job_1683107607090_0002
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.53 sec   HDFS Read: 15231 HDFS Write: 102 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 530 msec
OK
39
Time taken: 15.783 seconds, Fetched: 1 row(s)
hive> select avg(depdelay) from flight;
```

```
hive> select avg(depdelay) from flight;
WARNING: 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.
Query ID = hdp123_20230503153850_bbfc3872-552b-46ab-a05b-dd088e62676e
Total jobs = 1
Launching Job 1 out of 1
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_1683107607090_0003, Tracking URL = http://stud-OptiPlex-3060:8088/proxy/application_1683107607090_0003/
Kill Command = /usr/local/hadoop/bin/hadoop job  -kill job_1683107607090_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-05-03 15:38:56,319 Stage-1 map = 0%,  reduce = 0%
2023-05-03 15:38:59,381 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.16 sec
2023-05-03 15:39:03,438 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.45 sec
MapReduce Total cumulative CPU time: 2 seconds 450 msec
Ended Job = job_1683107607090_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.45 sec   HDFS Read: 16653 HDFS Write: 118 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 450 msec
OK
10.021739130434783
Time taken: 14.047 seconds, Fetched: 1 row(s)
hive>
```