

# HIVE

Start hive:

```
hduser@localhost:~ x hduser@localhost:~ x hduser@localhost:~ x hduser@localhost:~ x
[hduser@localhost ~]$ hive --service metastore
ls: cannot access /usr/local/spark/lib/spark-assembly-*.jar: No such file or directory
Starting Hive Metastore Server
22/06/30 10:37:24 WARN conf.HiveConf: HiveConf of name hive.mapjoin.optimized.keys does not exist
22/06/30 10:37:24 WARN conf.HiveConf: HiveConf of name hive.optimize.multigroupby.common.distincts does not exist
22/06/30 10:37:24 WARN conf.HiveConf: HiveConf of name hive.mapjoin.lazy.hashtable does not exist
22/06/30 10:37:24 WARN conf.HiveConf: HiveConf of name hive.server2.thrift.http.min.worker.threads does not exist
22/06/30 10:37:24 WARN conf.HiveConf: HiveConf of name hive.server2.thrift.http.max.worker.threads does not exist
22/06/30 10:37:24 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/phoenix-4.11.0-HBase-0.98-client.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.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.slf4j.impl.Log4jLoggerFactory]
Loading class 'com.mysql.jdbc.Driver'. This is deprecated. The new driver class is 'com.mysql.cj.jdbc.Driver'. The driver is automatically registered via the SPI and manual loading of the driver class is generally unnecessary.
```

Go to hive in other terminal:

```
hduser@localhost:~ x hduser@localhost:~ x hduser@localhost:~ x hduser@localhost:~ x
[hduser@localhost ~]$ hive
ls: cannot access /usr/local/spark/lib/spark-assembly-*.jar: No such file or directory
22/06/30 10:39:07 WARN conf.HiveConf: HiveConf of name hive.mapjoin.optimized.keys does not exist
22/06/30 10:39:07 WARN conf.HiveConf: HiveConf of name hive.optimize.multigroupby.common.distincts does not exist
22/06/30 10:39:07 WARN conf.HiveConf: HiveConf of name hive.mapjoin.lazy.hashtable does not exist
22/06/30 10:39:07 WARN conf.HiveConf: HiveConf of name hive.server2.thrift.http.min.worker.threads does not exist
22/06/30 10:39:07 WARN conf.HiveConf: HiveConf of name hive.server2.thrift.http.max.worker.threads does not exist
22/06/30 10:39:07 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
Logging initialized using configuration in jar:file:/usr/local/hive/lib/hive-common-1.2.2.jar!/hive-log4j.properties
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/phoenix-4.11.0-HBase-0.98-client.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.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.slf4j.impl.Log4jLoggerFactory]
hive>
```

```
hive> show databases;
OK
cognizant
cts
custdb
default
demodb
insure
retail
retailstg
sample
Time taken: 0.93 seconds, Fetched: 9 row(s)
hive> create database sampled;
OK
Time taken: 1.092 seconds
hive> use sampled;
OK
Time taken: 0.172 seconds
hive> create table customer(
  > id bigint,
  > name string,
  > address string);
OK
Time taken: 0.518 seconds
```

```
hive> insert into customer values(1,"Alekhya","Hyderabad"),(2,"Deepti","Amaravathi"),(3,"Keerthana","Vizag"),(4,"Likhith","Nalgonda");
Query ID = hduser_20220630104907_81ab7285-a2d5-48db-9648-cab48c720c52
Total jobs = 3
```

```
hive> describe formatted customer;
OK
# col_name           data_type           comment
id                   bigint
name                 string
address              string

# Detailed Table Information
Database:            sampled
Owner:               hduser
CreateTime:          Thu Jun 30 10:42:17 IST 2022
LastAccessTime:      UNKNOWN
Protect Mode:        None
Retention:           0
Location:             hdfs://localhost:54310/user/hive/warehouse/sampledb.db/customer
Table Type:          MANAGED_TABLE
Table Parameters:
  COLUMN_STATS_ACCURATE true
  numFiles              1
  totalSize             77
  transient_lastDdlTime 1656566374
```

```
# Storage Information
SerDe Library:      org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe
InputFormat:        org.apache.hadoop.mapred.TextInputFormat
OutputFormat:       org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat
Compressed:         No
Num Buckets:        -1
Bucket Columns:     []
Sort Columns:       []
Storage Desc Params:
  serialization.format 1
Time taken: 0.153 seconds, Fetched: 31 row(s)
```

```
hive> select * from customer where address="Hyderabad";
OK
1      Alekhya Hyderabad
Time taken: 0.217 seconds, Fetched: 1 row(s)
hive>
```

Load csv file from local to hive table:

```
hive> create table if not exists orders(
  > id bigint,
  > product_id string,
  > customer_id bigint,
  > quantity int,
  > amount double)
  > row format delimited
  > fields terminated by ","
  > stored as textfile;
OK
Time taken: 0.224 seconds
hive> load data local inpath "/home/hduser/sample.csv" into table orders;
Loading data to table sampled orders
Table sampled.orders stats: [numFiles=1, totalSize=87]
OK
Time taken: 0.599 seconds
hive> select * from orders;
OK
1101    Phone    1        2        20000.0
1102    Camera   2        1        30000.0
1103    Laptop   3        6        40000.0
1104    Watch    4       17        30000.0
Time taken: 0.1 seconds, Fetched: 4 row(s)
hive>
```

Load file from hdfs to hive table:

```
[hduser@localhost ~]$ hdfs dfs -mkdir /user/hduser/hivedata
22/06/30 11:29:22 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[hduser@localhost ~]$ hdfs dfs -put sample.csv /user/hduser/hivedata
22/06/30 11:29:43 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

hive> load data inpath "/user/hduser/hivedata/sample.csv" into table orders;
Loading data to table sampled orders
Table sampled.orders stats: [numFiles=2, totalSize=174]
OK
Time taken: 0.439 seconds
hive> select * from orders;
OK
1101    Phone    1        2        20000.0
1102    Camera   2        1        30000.0
1103    Laptop   3        6        40000.0
1104    Watch    4       17        30000.0
1101    Phone    1        2        20000.0
1102    Camera   2        1        30000.0
1103    Laptop   3        6        40000.0
1104    Watch    4       17        30000.0
Time taken: 0.054 seconds, Fetched: 8 row(s)
hive>
```