Download hive tar

Unzip tar, rename

Edit bashrc- add export HIVE\_HOME=/home/stud/Downloads/hive

start hadoop services

go to hadoop/bin and make 2 dir

hdfs dfs -mkdir -p /user/hive/warehouse

hdfs dfs -mkdir -p /tmp/hive

hdfs dfs -chmod 777 /tmp

hdfs dfs -chmod 777 /user/hive/warehouse

hdfs dfs -chmod 777 /tmp/hive

go to hive/bin

cd..(come out from hadoop bin)

cd..(come out from hadoop)

cd hive/bin

ls

./schematool -initSchema -dbtype derby

**output-**

**Metastore connection URL: jdbc:derby:;databaseName=metastore\_db;create=true**

**Metastore Connection Driver : org.apache.derby.jdbc.EmbeddedDriver**

**Metastore connection User: APP**

**Starting metastore schema initialization to 1.2.0**

**Initialization script hive-schema-1.2.0.derby.sql**

**Initialization script completed**

**schemaTool completed**

ls

start hive terminal

./hive

**1.show databases;**

**2.create database d1;**

**3.create table emp(ename string,esal int) row format delimited fields terminated by '\_'stored as textfile;**

**4.load data local inpath '/home/stud/Desktop/sample.txt' into table emp1;**

**5.create table flight(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, Origin string, Dest string, Distance int, TaxiIn int, TaxiOut int, Cancelled int, CancellationCode string, Diverted string, CarrierDelay int, WeatherDelay int, NASDelay int, SecurityDelay int, LateAircraftDelay int) row format delimited fields terminated by ',';**

**6.load data local inpath '/home/stud/Downloads/flight\_data.csv' into table flight;**

####for External Table:

**7.**Open another terminal and execute following commands to copy csv file to hdfs

**hdfs dfs -mkdir /user/hive/flight**

**hdfs dfs -put /home/stud/Downloads/flight\_ext.csv /user/hive/flight**

**hdfs dfs -ls /user/hive/flight**

8.Go to hive terminal and create external table by specifying path:

**create external table flight\_ext(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, Origin string, Dest string, Distance int, TaxiIn int, TaxiOut int, Cancelled int, CancellationCode string, Diverted string, CarrierDelay int, WeatherDelay int, NASDelay int, SecurityDelay int, LateAircraftDelay int) row format delimited fields terminated by ',' location '/user/hive/flight';**

\*\*\*Command to find the details of table in hive:

**describe formatted table\_name;**

Count the number of rows in table:

**select count(\*) from flight\_ext;**

\*\*\*\*Join tables:

**create table flight\_fare( FlightNum int, Fare int) row format delimited fields terminated by ',';**

**load data local inpath '/home/stud/Downloads/flight\_fare.csv' into table flight\_fare;**

**select flight\_csv.Origin, flight\_csv.Dest, flight\_fare.fare from flight\_csv join flight\_fare on flight\_csv.FlightNum=flight\_fare.FlightNum;**