Name: Renu Mariam Mathew
Student Id:16376
Hw#4

Step 1:

Create external table emp(eid int,name String,role String,salary String) row format delimited fields terminated by '\t' lines terminated by '\n';

Step 2:

LOAD DATA LOCAL INPATH 'cs570D.txt' OVERWRITE INTO TABLE EMP;

Step 1 and Step 2 combined

```
hadoop@localhost:~                                          —  ☐  ✕
OK
Time taken: 1.13 seconds
hive> Create external table emp(eid int,name String,role String,salary String)
ow format delimited fields  terminated by '\t'  lines terminated by '\n';
OK
Time taken: 1.292 seconds
hive> LOAD DATA LOCAL INPATH 'cs570D.txt' OVERWRITE INTO TABLE EMP;
Loading data to table default.emp
Table default.emp stats: [numFiles=1, numRows=0, totalSize=123, rawDataSize=0]
OK
Time taken: 4.093 seconds
hive> SELECT * FROM EMP;
OK
1201    Gopal    45000    manager
1202    Manisha 45000    reader
1203    Mast     40000    writer
1204    Kiran    40000    Hr
1205    Kranthi 30000    Op
Time taken: 0.477 seconds, Fetched: 5 row(s)
```

Step 3:

set hive.exec.dynamic.partition= true;

set hive.exec.dynamic.partition.mode=nonstrict;

```
hadoop@localhost:~
hive> set hive.exec.dynamic.partition= true;
hive> set hive.exec.dynamic.partition.mode=nonstrict;
```
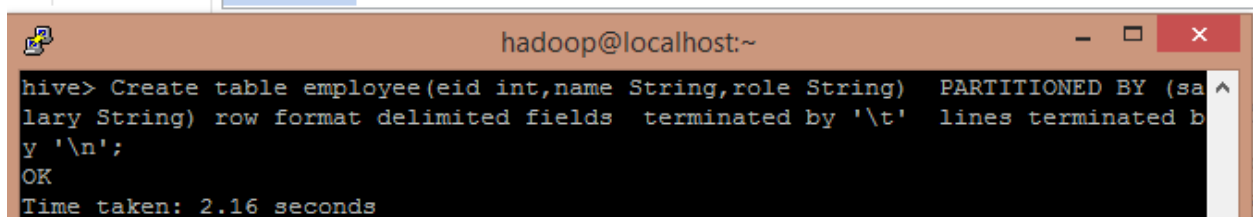
Name: Renu Mariam Mathew
Student Id:16376
Hw#4

Step 4:

Create table employee(eid int,name String,role String) PARTITIONED BY (salary String) row format delimited fields terminated by '\t' lines terminated by '\n';

Step 5:

INSERT OVERWRITE TABLE EMPLOYEE PARTITION (salary) select eid,name,salary,role from emp;

Name: Renu Mariam Mathew
Student Id:16376
Hw#4

```
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> INSERT OVERWRITE TABLE EMPLOYEE PARTITION (salary) select eid,name,salary,
role from emp;
Query ID = hadoop_20160405153737_22624c46-af17-4db0-b281-03c4d0818c1b
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_1459881998495_0006, Tracking URL = http://localhost:8088/prox
y/application_1459881998495_0006/
Kill Command = /home/hadoop/hadoop/bin/hadoop job  -kill job_1459881998495_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2016-04-05 15:38:25,360 Stage-1 map = 0%,   reduce = 0%
2016-04-05 15:39:26,172 Stage-1 map = 100%,   reduce = 0%, Cumulative CPU 5.87 se
c
MapReduce Total cumulative CPU time: 5 seconds 870 msec
Ended Job = job_1459881998495_0006
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
```

OutPut:

hdfs   dfs   –cat
/user/hive/warehouse/employee/salary=30000/000000_0

127.0.0.1:50070/explorer.html#/user/hive/warehouse/employee

Hadoop   Overview   Datanodes   Snapshot   Startup Progress   Utilities

Browse Directory

/user/hive/warehouse/employee

Permission	Owner	Group	Size	Replication	Block Size	Name
drwxrwxr-x	hadoop	supergroup	0 B	0	0 B	salary=30000
drwxrwxr-x	hadoop	supergroup	0 B	0	0 B	salary=40000
drwxrwxr-x	hadoop	supergroup	0 B	0	0 B	salary=45000

/user/hive/warehouse/employee/salary=30000

Permission	Owner	Group	Size	Replication	Block Size	Name
-rwxrwxr-x	hadoop	supergroup	16 B	1	128 MB	000000_0

Hadoop, 2014.

Snapshot   Startup Progress   Utilities

File information - 000000_0

Download

Block information --

Block ID: 107374254

Block Pool ID: BP-189

Generation Stamp: 1

Size: 16

Availability:

- localhost

000000_0

File   Edit   Format   View   Help

1205   Kranthi Op

The input file which I used is

```
1201  Gopal 45000 manager
1202  Manisha     45000 reader
1203  Mast  40000 writer
1204  Kiran 40000 Hr
1205  Kranthi     30000 Op
```