

---

## Task 2: Insert data into the table

There are two tables(ie: Mascot, College). The below are the rows that are inserted using HIVE queries.

Table 1: Mascot

### Step 1: Inserting Values into the Table

```
hive> insert into mascot values (101, 'Pistol Pete');
```

```
hive> insert into mascot values (104, 'WuShock');
```

```
hive> insert into mascot values (102, 'Brutus Buckeye');
```

```
hive> insert into mascot values (103, 'Raider Red');
```

```
hive> insert into mascot values (105, 'Beaver');
```

```
hive> insert into mascot values (106, 'Sparty');
```

### Output:

```
Query ID = anagave_20210503225055_7b5ef215-e208-48d5-8671-a3481d24689c
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_1620056543801_0056, Tracking URL = http://hadoop-nn001.cs.okstate.edu:8088/proxy/application\_1620056543801\_0056/
Kill Command = /usr/local/hadoop-3.3.0/bin/mapred job -kill job_1620056543801_0056
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-05-03 22:51:06,815 Stage-1 map = 0%, reduce = 0%
2021-05-03 22:51:12,116 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 6.44 sec
2021-05-03 22:51:19,457 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 10.49 sec
MapReduce Total cumulative CPU time: 10 seconds 490 msec
Ended Job = job_1620056543801_0056
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to directory hdfs://hadoop-nn001.cs.okstate.edu:9000/user/anagave/achyu/mascot/.hive-staging\_hive\_2021-05-03\_22-50-55\_271\_6303030171381478245-1/-ext-10000
Loading data to table achyu.mascot
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.49 sec HDFS Read: 15626 HDFS Write: 253
SUCCESS
Total MapReduce CPU Time Spent: 10 seconds 490 msec
OK
Time taken: 26.068 seconds
```

## Step 2: Displaying the inserted Values

```
hive> select * from mascot;  
OK  
101      Pistol Pete  
102      Brutus Buckeye  
103      Raider Red  
104      WuShock  
105      Beaver  
106      Sparty  
Time taken: 0.19 seconds, Fetched: 6 row(s)
```

## Table 2: College

### Step 1: Inserting Values into the Table

```
hive> INSERT INTO TABLE College PARTITION(Country ='US') Select 103, 'Ariana', ar  
ay( 1215678903, 1217504890) , named_struct('city','Lubbock', 'state','Texas' )  
( 'OS' , 2, 'DS' , 2) from mascot limit 1  
> ;
```

```
hive> INSERT INTO TABLE College PARTITION(Country ='US') Select 102, 'Lee', ar  
1215678903, 1217504890) , named_struct('city','Columbus', 'state','Ohio' ),MA  
S' , 2, 'FLT' , 2) from mascot limit 1;
```

```
hive> INSERT INTO TABLE College PARTITION(Country ='US') Select 106, 'Tina', a  
( 1215678903, 1217504890) , named_struct('city','East Lansing', 'state','Michi  
) ,MAP('COA' , 2, 'DS' , 2) from mascot limit 1;
```

```
hive> INSERT INTO TABLE College PARTITION(Country = 'US') Select 101, 'Kim', a  
( 2127893564, 2121123546) , named_struct('city','Stillwater', 'state','Oklahom  
,MAP('COA' , 2, 'OS' , 2) from mascot limit 1;
```

```
hive> INSERT INTO TABLE College PARTITION(Country = 'US') Select 105, 'Lee', a  
( 212128978, 212976542) , named_struct('city','PASADENA', 'state','California'  
PC('FLT' , 2, 'DS' , 2) from mascot limit 1  
> ;
```

```
hive> INSERT INTO TABLE College PARTITION(Country='US') Select 104, 'John', ar  
212120201, 212456392) , named_struct('city','Wichita', 'state','Kansas' ),MAP  
' , 2, 'DS' , 2) from mascot limit 1;
```

### Output:

```
Query ID = anagave_20210503231200_99aad534-2bf0-4f66-9367-0199379765dd  
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\_1620056543801\_0073, Tracking URL = [http://hadoop-nn001.cs.okstate.edu:8088/proxy/application\\_1620056543801\\_0073/](http://hadoop-nn001.cs.okstate.edu:8088/proxy/application_1620056543801_0073/)

Kill Command = /usr/local/hadoop-3.3.0/bin/mapred job -kill job\_1620056543801\_0073

Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 1

2021-05-03 23:12:06,906 Stage-1 map = 0%, reduce = 0%

2021-05-03 23:12:12,265 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 5.29 sec

2021-05-03 23:12:13,314 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 23.96 sec

2021-05-03 23:12:16,462 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 27.8 sec

MapReduce Total cumulative CPU time: 27 seconds 800 msec

Ended Job = job\_1620056543801\_0073

Loading data to table achyu.college partition (country=US)

MapReduce Jobs Launched:

Stage-Stage-1: Map: 5 Reduce: 1 Cumulative CPU: 27.8 sec HDFS Read: 42629 HDFS

Write: 134 SUCCESS

Total MapReduce CPU Time Spent: 27 seconds 800 msec

OK

Time taken: 18.527 seconds

## Step 2: Displaying the inserted values

```
hive> select * from college;
OK
103 Ariana [1215678903,1217504890] {"city":"Lubbock","state":"Texas"} {"OS":2,"DS":2} US
102 Lee [1215678903,1217504890] {"city":"Columbus","state":"Ohio"} {"DS":2,"FLT":2} US
106 Tina [1215678903,1217504890] {"city":"East Lansing","state":"Michigan"} {"COA":2,"DS":2} US
101 Kim [2127893564,2121123546] {"city":"Stillwater","state":"Oklahoma"} {"COA":2,"OS":2} US
105 Lee [212128978,212976542] {"city":"PASADENA","state":"California"} {"FLT":2,"DS":2} US
104 John [212120201,212456392] {"city":"Wichita","state":"Kansas"} {"OS":2,"DS":2} US
Time taken: 0.153 seconds, Fetched: 6 row(s)
```