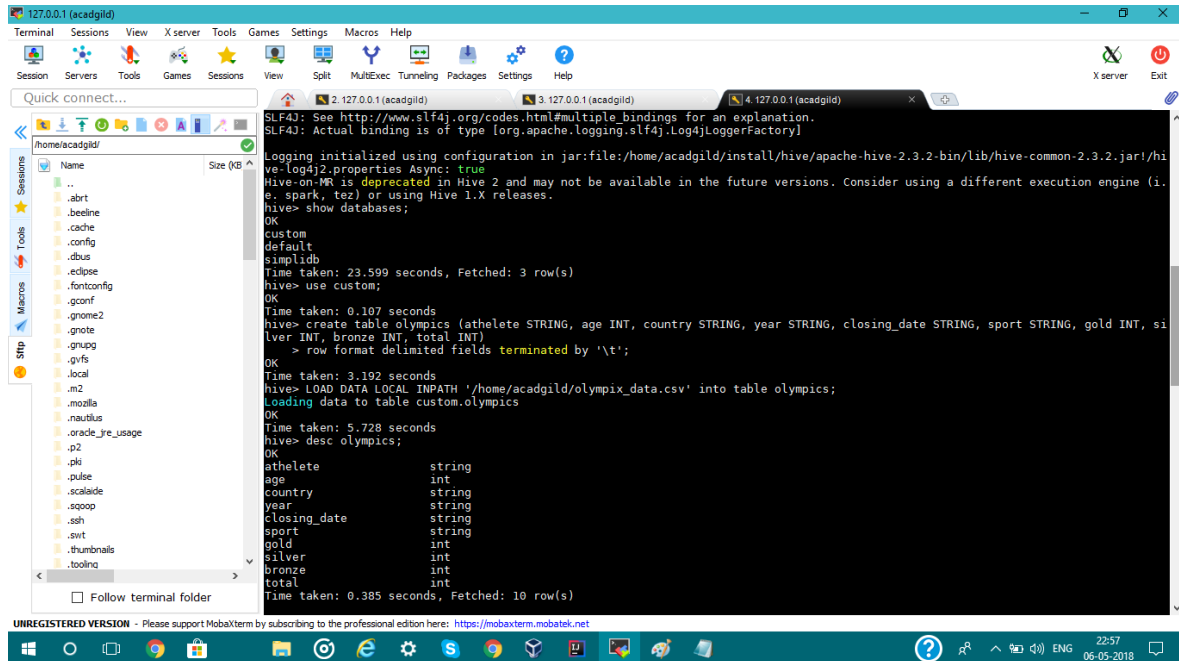


Assignment 9

Creation of table for Olympics table



```
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:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/hive-common-2.3.2.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, tez) or using Hive 1.X releases.
hive> show databases;
OK
custom
default
simplydb
Time taken: 23.599 seconds, Fetched: 3 row(s)
hive> use custom;
OK
Time taken: 0.107 seconds
hive> create table olympics (athlete STRING, age INT, country STRING, year STRING, closing_date STRING, sport STRING, gold INT, silver INT, bronze INT, total INT)
> row format delimited fields terminated by '\t';
OK
Time taken: 3.192 seconds
hive> LOAD DATA LOCAL INPATH '/home/acadgild/olympix_data.csv' into table olympics;
Loading data to table custom.olympics
OK
Time taken: 5.728 seconds
hive> desc olympics;
OK
athlete          string
age              int
country          string
year             string
closing_date     string
sport            string
gold             int
silver           int
bronze           int
total            int
Time taken: 0.385 seconds, Fetched: 10 row(s)
```

Command used: create table olympic (athlete STRING, age INT, country STRING, year STRING, closing_date STRING, sport STRING, gold INT, silver INT, bronze INT, total INT)
row format delimited fields terminated by '\t';

Task1

- Write a Hive program to find the number of medals won by each country in swimming.

Command used: select country, sum(total) from Olympics where sport='Swimming' group by country;

```

hive> select country, sum(total) from olympics where sport ='Swimming' group by country;
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 = acadgild_20180506233515_e325bf7b-98a8-4b8b-a199-43e3a88fa762
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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.reducers=<number>
Starting Job = job_1525596368858_0003, Tracking URL = http://localhost:8088/proxy/application_1525596368858_0003/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525596368858_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2018-05-06 23:36:02,927 Stage-1 map = 0%, reduce = 0%
2018-05-06 23:36:34,856 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 8.74 sec
2018-05-06 23:36:59,808 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 13.21 sec
2018-05-06 23:37:03,011 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 15.01 sec
MapReduce Total cumulative CPU time: 15 seconds 10 msec
Ended Job = job_1525596368858_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 15.01 sec HDFS Read: 528624 HDFS Write: 881 SUCCESS
Total MapReduce CPU Time Spent: 15 seconds 10 msec
OK
Argentina 1
Australia 163
Austria 3
Belarus 2
Brazil 8
Canada 5
China 35
Costa Rica 2
Croatia 1
Denmark 1
France 39
Germany 32
Great Britain 11
Hungary 9
Italy 16
Japan 43
Lithuania 1
Netherlands 46
Norway 2
Poland 3
Romania 6
Russia 20
Serbia 1
Slovakia 2
Slovenia 1
South Africa 11
South Korea 4
Spain 3
Sweden 3
Trinidad and Tobago 1
Tunisia 3
Ukraine 7
United States 267
Zimbabwe 7
Time taken: 110.52 seconds, Fetched: 34 row(s)
hive>

```

Output

```

Total MapReduce CPU Time Spent: 15 seconds 10 msec
OK
Argentina 1
Australia 163
Austria 3
Belarus 2
Brazil 8
Canada 5
China 35
Costa Rica 2
Croatia 1
Denmark 1
France 39
Germany 32
Great Britain 11
Hungary 9
Italy 16
Japan 43
Lithuania 1
Netherlands 46
Norway 2
Poland 3
Romania 6
Russia 20
Serbia 1
Slovakia 2
Slovenia 1
South Africa 11
South Korea 4
Spain 3
Sweden 3
Trinidad and Tobago 1
Tunisia 3
Ukraine 7
United States 267
Zimbabwe 7
Time taken: 110.52 seconds, Fetched: 34 row(s)
hive>

```

b. Write a Hive program to find the number of medals that India won year wise.

Command used: select year, sum (total) from Olympics where country = 'India' group by year;

The screenshot shows a MobaXterm window with a terminal session. The terminal displays the execution of a Hive query: `hive> select year, sum(total) from olympics where country='India' group by year;`. The output shows a single row for the year 2012 with a total of 6 medals. The terminal also displays various warnings and job execution details, including the number of reducers and the cumulative CPU time.

```

hive> select year, sum(total) from olympics where country='India' group by year;
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 = acadgild_20180506234850_5c24a462-31a5-4043-8da3-0a6a6000582c
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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_1525596368858_0005, Tracking URL = http://localhost:8088/proxy/application_1525596368858_0005/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525596368858_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2018-05-06 23:49:20,012 Stage-1 map = 0%, reduce = 0%
2018-05-06 23:49:46,218 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 7.21 sec
2018-05-06 23:49:47,324 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.49 sec
2018-05-06 23:50:10,838 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 12.17 sec
2018-05-06 23:50:12,491 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 13.48 sec
MapReduce Total cumulative CPU time: 13 seconds 480 msec
Ended Job = job_1525596368858_0005
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 13.48 sec HDFS Read: 528612 HDFS Write: 163 SUCCESS
Total MapReduce CPU Time Spent: 13 seconds 480 msec
OK
2000 1
2004 1
2008 3
2012 6
Time taken: 84.604 seconds, Fetched: 4 row(s)
hive>

```

c. Write a Hive Program to find the total number of medals each country won.

Command used: select country, sum (total) from Olympics group by country;

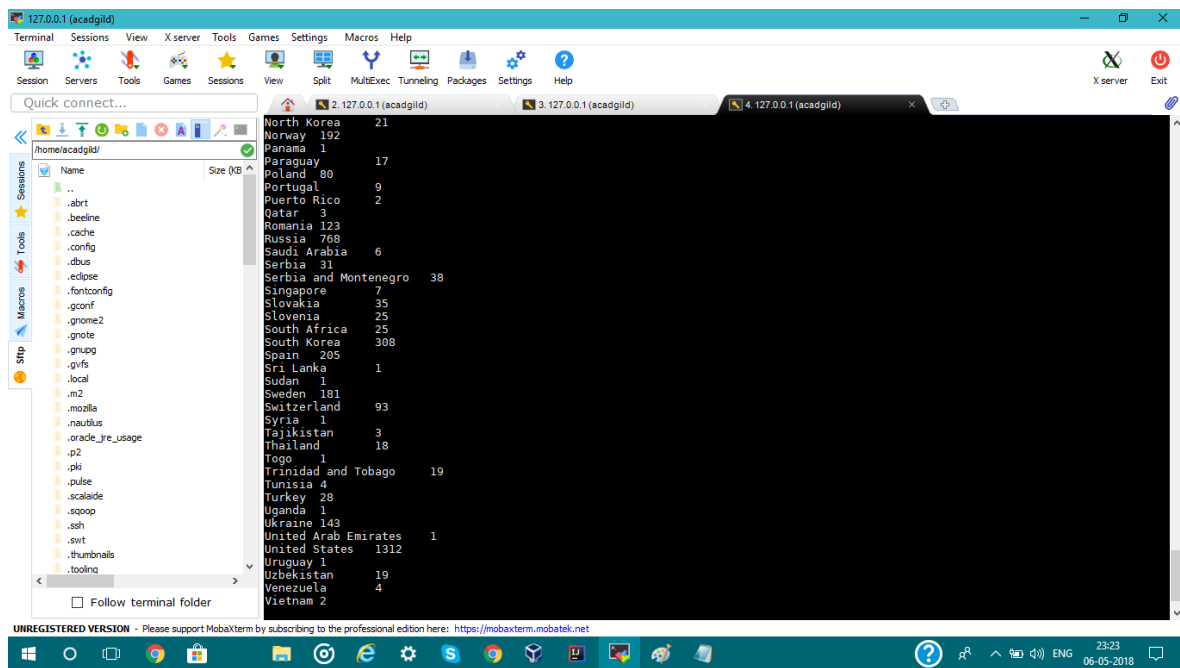
The screenshot shows a MobaXterm window with a terminal session. The terminal displays the execution of a Hive query: `hive> select country, sum(total) from olympics group by country;`. The output shows a list of countries and their total number of medals, including Afghanistan (2), Algeria (8), Argentina (141), Armenia (10), Australia (609), Austria (91), Azerbaijan (25), Bahamas (24), Bahrain (1), Barbados (1), Belarus (97), and Belgium (18). The terminal also displays various warnings and job execution details, including the number of reducers and the cumulative CPU time.

```

hive> select country, sum(total) from olympics group by country;
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 = acadgild_20180507000202_c63c5a4f-89b2-43ae-aaab-adaec94b3c9b
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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_1525596368858_0006, Tracking URL = http://localhost:8088/proxy/application_1525596368858_0006/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525596368858_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2018-05-07 00:02:30,765 Stage-1 map = 0%, reduce = 0%
2018-05-07 00:02:53,917 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.84 sec
2018-05-07 00:03:20,949 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 9.35 sec
2018-05-07 00:03:22,687 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 10.65 sec
MapReduce Total cumulative CPU time: 10 seconds 650 msec
Ended Job = job_1525596368858_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.65 sec HDFS Read: 527792 HDFS Write: 2742 SUCCESS
Total MapReduce CPU Time Spent: 10 seconds 650 msec
OK
Afghanistan 2
Algeria 8
Argentina 141
Armenia 10
Australia 609
Austria 91
Azerbaijan 25
Bahamas 24
Bahrain 1
Barbados 1
Belarus 97
Belgium 18

```

Output

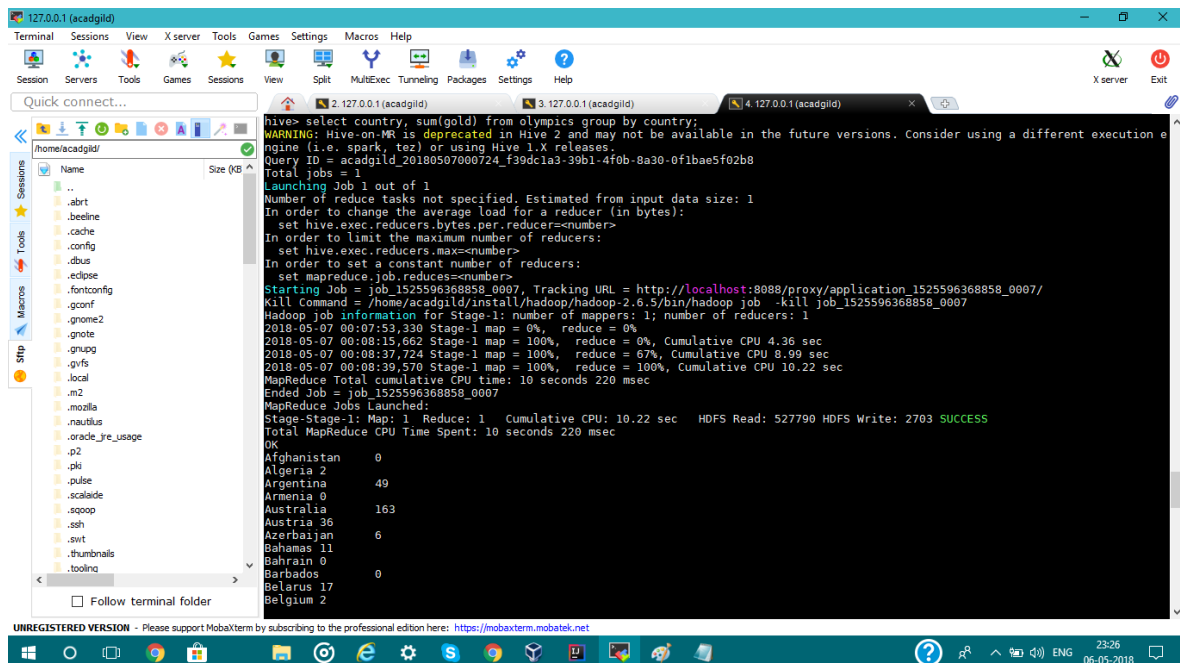


The screenshot shows the MobaXterm interface with a terminal window displaying the output of a Hive query. The query is: `select country, sum(gold) from olympics group by country;`. The output lists countries and the number of gold medals they won. The terminal window is titled '127.0.0.1 (acadgild)' and shows a file explorer on the left with the path `/home/acadgild/`. The terminal output is as follows:

```
North Korea 21
Norway 192
Panama 1
Paraguay 17
Poland 80
Portugal 9
Puerto Rico 2
Qatar 3
Romania 123
Russia 768
Saudi Arabia 6
Serbia 31
Serbia and Montenegro 38
Singapore 7
Slovakia 35
Slovenia 25
South Africa 25
South Korea 308
Spain 205
Sri Lanka 1
Sudan 1
Sweden 181
Switzerland 93
Syria 1
Tajikistan 3
Thailand 18
Togo 1
Trinidad and Tobago 19
Tunisia 4
Turkey 28
Uganda 1
Ukraine 143
United Arab Emirates 1
United States 1312
Uruguay 1
Uzbekistan 19
Venezuela 4
Vietnam 2
```

d. Write a Hive program to find the number of gold medals each country won.

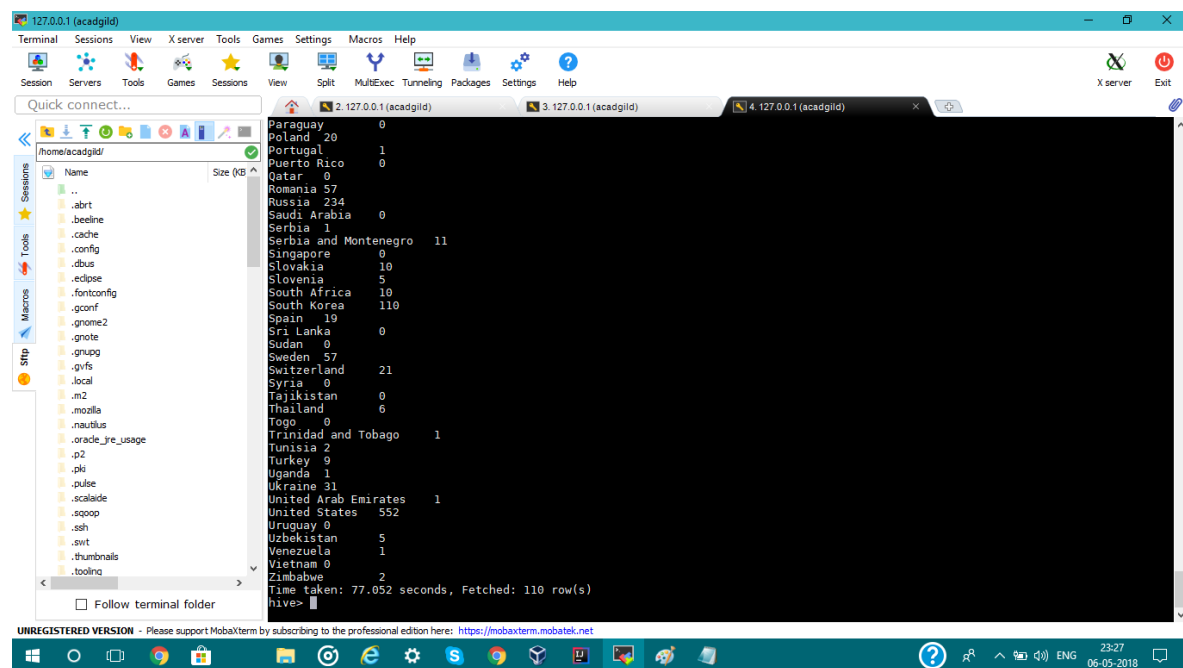
Command used: `select country, sum (gold) from Olympics group by country;`



The screenshot shows the MobaXterm interface with a terminal window displaying the execution of a Hive query. The query is: `hive> select country, sum(gold) from olympics group by country;`. The terminal output shows the execution details, including the number of jobs, the number of reducers, and the final results. The terminal window is titled '127.0.0.1 (acadgild)' and shows a file explorer on the left with the path `/home/acadgild/`. The terminal output is as follows:

```
hive> select country, sum(gold) from olympics group by country;
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 = acadgild_20180507000724_f39d1a3-39b1-4f0b-8a30-0f1bae5f02b8
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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.reducers=<number>
Starting Job = job_1525596368858_0007, Tracking URL = http://localhost:8088/proxy/application_1525596368858_0007/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525596368858_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2018-05-07 00:07:53,230 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 4.36 sec
2018-05-07 00:08:15,662 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.36 sec
2018-05-07 00:08:37,724 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 8.99 sec
2018-05-07 00:08:39,570 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 10.22 sec
MapReduce Total cumulative CPU time: 10 seconds 220 msec
Ended Job = job_1525596368858_0007
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 10.22 sec HDFS Read: 527790 HDFS Write: 2703 SUCCESS
Total MapReduce CPU Time Spent: 10 seconds 220 msec
OK
Afghanistan 0
Algeria 2
Argentina 49
Armenia 0
Australia 163
Austria 36
Azerbaijan 6
Bahamas 11
Bahrain 0
Barbados 0
Belarus 17
Belgium 2
```

Output



The screenshot shows the MobaXterm interface with a terminal window displaying a list of countries and their counts. The terminal output is as follows:

```
Paraguay 0
Poland 20
Portugal 1
Puerto Rico 0
Qatar 0
Romania 57
Russia 234
Saudi Arabia 0
Serbia 1
Serbia and Montenegro 11
Singapore 0
Slovakia 10
Slovenia 5
South Africa 10
South Korea 110
Spain 19
Sri Lanka 0
Sudan 0
Sweden 57
Switzerland 21
Syria 0
Tajikistan 0
Thailand 6
Togo 0
Trinidad and Tobago 1
Tunisia 2
Turkey 9
Uganda 1
Ukraine 31
United Arab Emirates 1
United States 552
Uruguay 0
Uzbekistan 5
Venezuela 1
Vietnam 0
Zimbabwe 2
Time taken: 77.052 seconds, Fetched: 110 row(s)
hive>
```

Task2

Write a hive UDF that implements functionality of string concat_ws(string SEP, array<string>). This UDF will accept two arguments, one string and one array of string. It will return a single string where all the elements of the array are separated by the SEP.

Dataset used: customer.txt

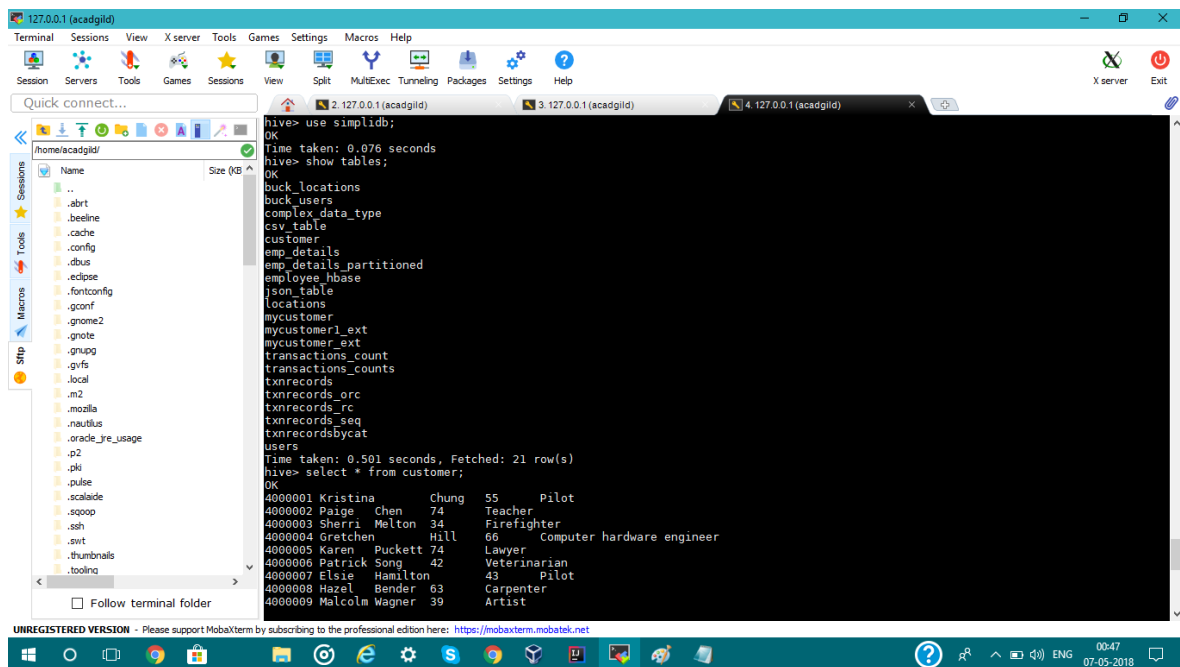
Command used: select concat_ws(' SEP ', collect_list(fname)) from customer;

Explanation: Here, the function concat_ws is used to concat the elements of an array.

Syntax: concat_ws('delimiter', array).

This function takes two arguments in case of an array: first is the delimiter for the concatenation and second is the array of elements for which the concatenation is to be done.

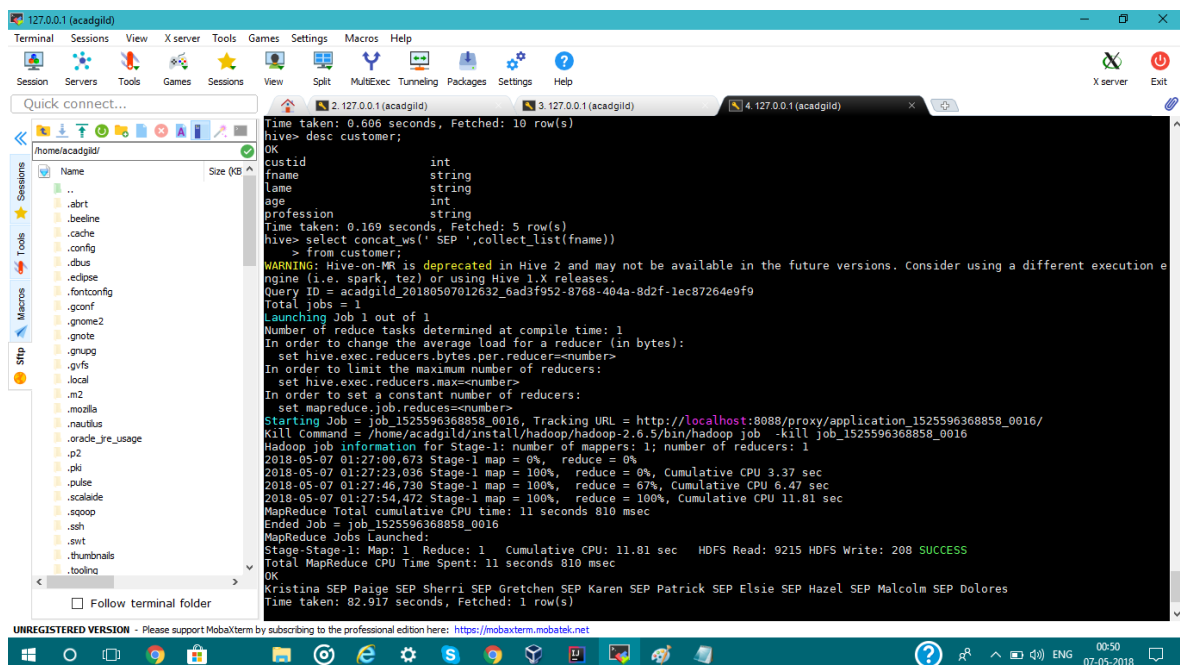
For this task, the delimiter is the string, 'SEP' with white spaces before and after; and the column for which the array is to be generated via collect_list is 'fname'



The screenshot shows a MobaXterm terminal window with the following content:

```
hive> use simplidb;
OK
Time taken: 0.076 seconds
hive> show tables;
OK
buck_locations
buck_users
complex_data_type
csv_table
customer
emp_details
emp_details_partitioned
employee_hbase
json_table
locations
mycustomer
mycustomer_ext
mycustomer_ext
transactions_count
transactions_counts
txnrecords
txnrecords_orc
txnrecords_rc
txnrecords_seq
txnrecordsbycat
users
Time taken: 0.501 seconds, Fetched: 21 row(s)
hive> select * from customer;
OK
4000001 Kristina Chung 55 Pilot
4000002 Paige Chen 74 Teacher
4000003 Sherri Melton 34 Firefighter
4000004 Gretchen Hill 66 Computer hardware engineer
4000005 Karen Puckett 74 Lawyer
4000006 Patrick Song 42 Veterinarian
4000007 Elsie Hamilton 43 Pilot
4000008 Hazel Bender 63 Carpenter
4000009 Malcolm Wagner 39 Artist
```

Output

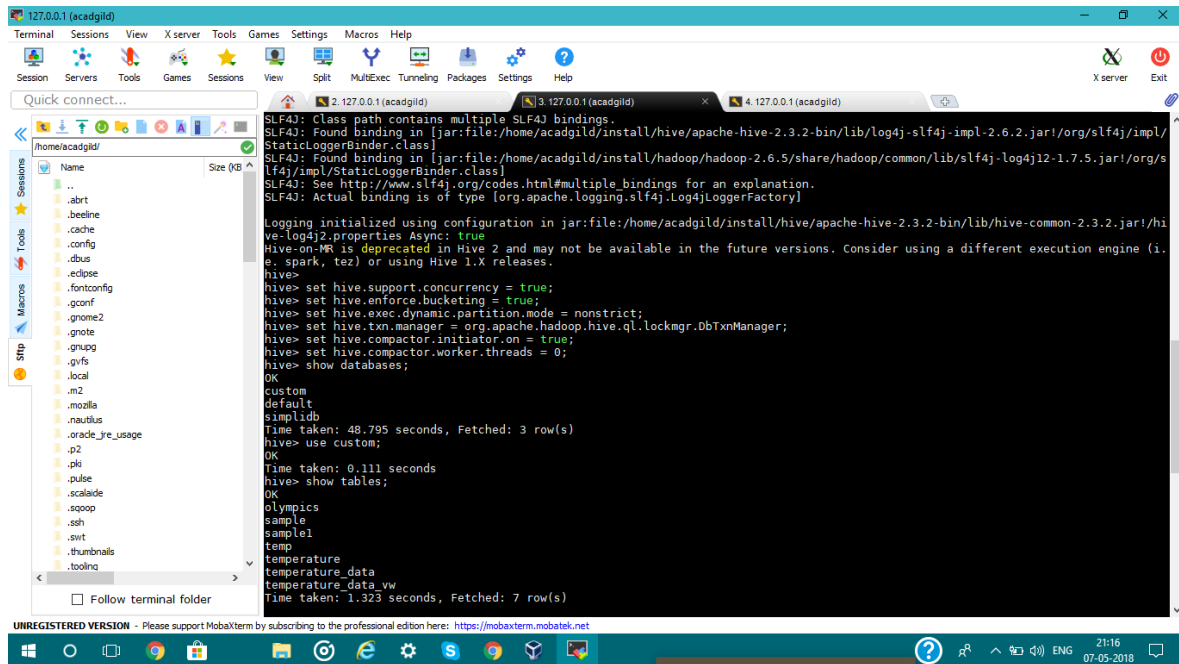


The screenshot shows a MobaXterm terminal window with the following content:

```
hive> desc customer;
OK
custid          int
fname           string
lname           string
age             int
profession      string
Time taken: 0.169 seconds, Fetched: 5 row(s)
hive> select concat_ws(' SEP ',collect_list(fname))
> from customer;
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 = acadgild_20180507012632_6ad3f952-8768-404a-8d2f-1ec87264e9f9
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.reducers=<number>
Starting Job = job_1525596368858_0016, Tracking URL = http://localhost:8088/proxy/application_1525596368858_0016/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.0.5/bin/hadoop job -kill job_1525596368858_0016
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2018-05-07 01:27:00,673 Stage-1 map = 0%, reduce = 0%
2018-05-07 01:27:23,036 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.37 sec
2018-05-07 01:27:46,730 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 6.47 sec
2018-05-07 01:27:54,472 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 11.81 sec
MapReduce Total cumulative CPU time: 11 seconds 810 msec
Ended Job = job_1525596368858_0016
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 11.81 sec HDFS Read: 9215 HDFS Write: 208 SUCCESS
Total MapReduce CPU Time Spent: 11 seconds 810 msec
OK
Kristina SEP Paige SEP Sherri SEP Gretchen SEP Karen SEP Patrick SEP Elsie SEP Hazel SEP Malcolm SEP Dolores
Time taken: 82.917 seconds, Fetched: 1 row(s)
```

Task3

a. Setting up of properties

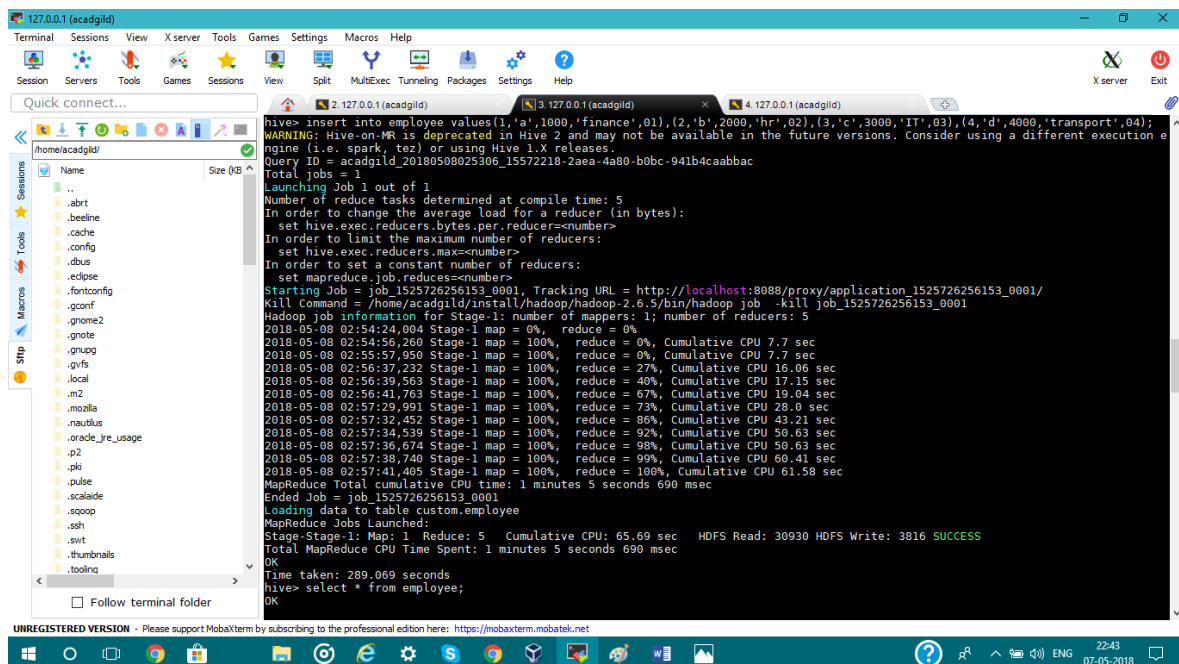


The screenshot shows a MobaXterm terminal window with a file explorer on the left displaying the directory structure of a Hive installation. The terminal output shows the following steps:

```
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.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:/home/acadgild/install/hive/apache-hive-2.3.2-bin/lib/hive-common-2.3.2.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, tez) or using Hive 1.X releases.
hive> set hive.support.concurrency = true;
hive> set hive.enforce.bucketing = true;
hive> set hive.exec.dynamic.partition.mode = nonstrict;
hive> set hive.txn.manager = org.apache.hadoop.hive ql.lockmgr.DbTxnManager;
hive> set hive.compactor.initiator.on = true;
hive> set hive.compactor.worker.threads = 0;
hive> show databases;
OK
custom
default
simp1db
Time taken: 48.795 seconds, Fetched: 3 row(s)
hive> use custom;
OK
Time taken: 0.111 seconds
hive> show tables;
OK
olympics
sample
sample1
temp
temperature
temperature_data
temperature_data_vw
Time taken: 1.323 seconds, Fetched: 7 row(s)
```

b. Creation of table that supports Hive Transaction



The screenshot shows a MobaXterm terminal window with a file explorer on the left. The terminal output shows the following steps:

```
hive> insert into employee values(1,'a',1000,'finance',01),(2,'b',2000,'hr',02),(3,'c',3000,'IT',03),(4,'d',4000,'transport',04);
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 = acadgild_20180508025306_15572218-2aea-4a80-b0bc-941b4caabbac
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 5
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.reducers=<number>
Starting Job = job_1525726256153_0001, Tracking URL = http://localhost:8080/proxy/application_1525726256153_0001/
Kill command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 5
2018-05-08 02:54:24,004 Stage-1 map = 0%, reduce = 0%
2018-05-08 02:54:56,260 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.7 sec
2018-05-08 02:55:57,950 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.7 sec
2018-05-08 02:56:37,232 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 16.06 sec
2018-05-08 02:56:39,563 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 17.15 sec
2018-05-08 02:56:41,763 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 19.04 sec
2018-05-08 02:57:29,991 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 28.0 sec
2018-05-08 02:57:32,452 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 45.21 sec
2018-05-08 02:57:34,530 Stage-1 map = 100%, reduce = 92%, Cumulative CPU 50.63 sec
2018-05-08 02:57:36,674 Stage-1 map = 100%, reduce = 98%, Cumulative CPU 50.63 sec
2018-05-08 02:57:38,740 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 60.41 sec
2018-05-08 02:57:41,405 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 61.58 sec
MapReduce Total cumulative CPU time: 1 minutes 5 seconds 690 msec
Ended Job = job_1525726256153_0001
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage1: Map: 1 Reduce: 5 Cumulative CPU: 65.69 sec HDFS Read: 30930 HDFS Write: 3816 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 5 seconds 690 msec
Time taken: 289.069 seconds
hive> select * from employee;
OK
```


The screenshot shows a MobaXterm window with a terminal and a file explorer. The terminal displays the execution of a Hive job to insert data into a table named 'employee'. The job is launched with the command: `hive> insert into employee values(1,'a',1000,'finance',01),(2,'b',2000,'hr',02),(3,'c',3000,'IT',03),(4,'d',4000,'transport',04);`. The output shows the job's progress, including the number of mappers and reducers, and the cumulative CPU time. The job is successfully completed, and the data is loaded into the table. The file explorer on the left shows the contents of the `/home/acadgild/` directory, including files like `.abrt`, `.beeline`, `.cache`, `.config`, `.dbus`, `.edpise`, `.fontconfig`, `.gnconf`, `.gnome2`, `.gnote`, `.gnupg`, `.gvfs`, `.local`, `.m2`, `.mozilla`, `.nautilus`, `.oracle_jre_usage`, `.p2`, `.plk`, `.pulse`, `.scalade`, `.sqoop`, `.ssh`, `.svt`, `.thumbnails`, and `.tooling`.

```
Number of reduce tasks determined at compile time: 5
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_1525726256153_0001, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0001/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 5
2018-05-08 02:54:24,004 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 7.7 sec
2018-05-08 02:54:56,260 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.06 sec
2018-05-08 02:55:57,950 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 17.15 sec
2018-05-08 02:56:37,232 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 19.04 sec
2018-05-08 02:56:39,563 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 19.15 sec
2018-05-08 02:57:29,991 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 28.0 sec
2018-05-08 02:57:32,452 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 43.21 sec
2018-05-08 02:57:34,539 Stage-1 map = 100%, reduce = 92%, Cumulative CPU 50.63 sec
2018-05-08 02:57:36,674 Stage-1 map = 100%, reduce = 98%, Cumulative CPU 50.63 sec
2018-05-08 02:57:38,740 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 60.41 sec
2018-05-08 02:57:41,405 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 61.58 sec
MapReduce Total cumulative CPU time: 1 minutes 5 seconds 690 msec
Ended Job = job_1525726256153_0001
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 5 Cumulative CPU: 65.69 sec HDFS Read: 30930 HDFS Write: 3816 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 5 seconds 690 msec
OK
Time taken: 289.069 seconds
hive> select * from employee;
OK
1      a      1000    finance  1
2      b      2000     hr        2
3      c      3000     IT        3
4      d      4000    transport  4
Time taken: 1.407 seconds, Fetched: 4 row(s)
hive>
```

c. Inserting data into Hive table

The screenshot shows a MobaXterm window with a terminal and a file explorer. The terminal displays the execution of a Hive job to insert data into a table named 'employee'. The job is launched with the command: `hive> insert into employee values(1,'a',1000,'finance',01),(2,'b',2000,'hr',02),(3,'c',3000,'IT',03),(4,'d',4000,'transport',04);`. The output shows the job's progress, including the number of mappers and reducers, and the cumulative CPU time. The job is successfully completed, and the data is loaded into the table. The file explorer on the left shows the contents of the `/home/acadgild/` directory, including files like `.abrt`, `.beeline`, `.cache`, `.config`, `.dbus`, `.edpise`, `.fontconfig`, `.gnconf`, `.gnome2`, `.gnote`, `.gnupg`, `.gvfs`, `.local`, `.m2`, `.mozilla`, `.nautilus`, `.oracle_jre_usage`, `.p2`, `.plk`, `.pulse`, `.scalade`, `.sqoop`, `.ssh`, `.svt`, `.thumbnails`, and `.tooling`.

```
hive> insert into employee values(1,'a',1000,'finance',01),(2,'b',2000,'hr',02),(3,'c',3000,'IT',03),(4,'d',4000,'transport',04);
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 = acadgild_20180508025306_15572218-2aea-4a80-b0bc-941b4caabbac
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 5
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_1525726256153_0001, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0001/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 5
2018-05-08 02:54:24,004 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 7.7 sec
2018-05-08 02:54:56,260 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.06 sec
2018-05-08 02:55:57,950 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 17.15 sec
2018-05-08 02:56:37,232 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 19.04 sec
2018-05-08 02:56:39,563 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 19.15 sec
2018-05-08 02:57:29,991 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 28.0 sec
2018-05-08 02:57:32,452 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 28.0 sec
2018-05-08 02:57:34,539 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 43.21 sec
2018-05-08 02:57:36,674 Stage-1 map = 100%, reduce = 92%, Cumulative CPU 50.63 sec
2018-05-08 02:57:38,740 Stage-1 map = 100%, reduce = 98%, Cumulative CPU 50.63 sec
2018-05-08 02:57:41,405 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 60.41 sec
2018-05-08 02:57:41,405 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 61.58 sec
MapReduce Total cumulative CPU time: 1 minutes 5 seconds 690 msec
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 5 Cumulative CPU: 65.69 sec HDFS Read: 30930 HDFS Write: 3816 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 5 seconds 690 msec
OK
Time taken: 289.069 seconds
hive> select * from employee;
OK
```


The screenshot shows a MobaXterm window with a terminal session. The terminal displays the execution of a Hive job. The job is identified as 'job_1525726256153_0001'. The progress bar shows the job is in progress. The terminal output includes the following information:

```
Number of reduce tasks determined at compile time: 5
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 mapreduce.job.reduces=<number>
In order to set a constant number of reducers:
  set hive.exec.reducers.max=<number>
Starting Job = job_1525726256153_0001, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0001/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 5
2018-05-08 02:54:24,004 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 7.7 sec
2018-05-08 02:54:56,260 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.06 sec
2018-05-08 02:55:57,950 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 17.15 sec
2018-05-08 02:56:37,232 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 19.04 sec
2018-05-08 02:56:39,563 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 19.04 sec
2018-05-08 02:57:29,891 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 28.0 sec
2018-05-08 02:57:32,452 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 43.21 sec
2018-05-08 02:57:34,539 Stage-1 map = 100%, reduce = 92%, Cumulative CPU 50.63 sec
2018-05-08 02:57:36,674 Stage-1 map = 100%, reduce = 98%, Cumulative CPU 50.63 sec
2018-05-08 02:57:38,740 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 60.41 sec
2018-05-08 02:57:41,405 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 61.58 sec
MapReduce Total cumulative CPU time: 1 minutes 5 seconds 690 msec
Ended Job = job_1525726256153_0001
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 5 Cumulative CPU: 65.69 sec HDFS Read: 30930 HDFS Write: 3816 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 5 seconds 690 msec
OK
Time taken: 289.069 seconds
hive> select * from employee;
OK
1 a 1000 finance 1
2 b 2000 hr 2
3 c 3000 IT 3
4 d 4000 transport 4
Time taken: 1.407 seconds, Fetched: 4 row(s)
hive>
```

d. Re-inserting of data

The screenshot shows a MobaXterm window with a terminal session. The terminal displays the execution of a Hive job. The job is identified as 'job_1525726256153_0002'. The progress bar shows the job is in progress. The terminal output includes the following information:

```
hive> insert into employee values(1,'a',1000,'finance',01),(2,'b',2000,'hr',02),(3,'c',3000,'IT',03),(4,'d',4000,'transport',04);
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 = acadgild_20180508030102_1aacf1eb-b4db-4911-b00c-a6d75a0adddc
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 5
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 mapreduce.job.reduces=<number>
In order to set a constant number of reducers:
  set hive.exec.reducers.max=<number>
Starting Job = job_1525726256153_0002, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0002/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 5
2018-05-08 03:01:41,865 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 7.76 sec
2018-05-08 03:02:10,829 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.76 sec
2018-05-08 03:03:11,264 Stage-1 map = 100%, reduce = 13%, Cumulative CPU 10.11 sec
2018-05-08 03:03:53,937 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 11.74 sec
2018-05-08 03:04:01,148 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 14.5 sec
2018-05-08 03:04:07,864 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 19.03 sec
2018-05-08 03:04:09,944 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 19.71 sec
2018-05-08 03:04:47,210 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 29.0 sec
2018-05-08 03:04:53,605 Stage-1 map = 100%, reduce = 80%, Cumulative CPU 36.54 sec
2018-05-08 03:04:55,792 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 43.51 sec
2018-05-08 03:04:59,848 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 46.51 sec
2018-05-08 03:05:02,570 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 61.74 sec
2018-05-08 03:05:09,243 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 65.68 sec
MapReduce Total cumulative CPU time: 1 minutes 5 seconds 680 msec
Ended Job = job_1525726256153_0002
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 5 Cumulative CPU: 65.68 sec HDFS Read: 30840 HDFS Write: 3820 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 5 seconds 680 msec
OK
```

The screenshot shows a MobaXterm window with a terminal session. The terminal displays the execution of a Hive job to update a table. The job is named 'job_1525726256153_0002'. The tracking URL is 'http://localhost:8088/proxy/application_1525726256153_0002/'. The kill command is '/home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0002'. The job information for Stage-1 shows 1 mapper and 5 reducers. The job is completed successfully with a cumulative CPU time of 1 minute 5 seconds 680 msec. The HDFS read and write statistics are 30840 and 3820 respectively. The job is ended with a success status.

```
set mapreduce.job.reduces=<number>
Starting Job = job_1525726256153_0002, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0002/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 5
2018-05-08 03:01:41,865 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:02:10,829 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.76 sec
2018-05-08 03:03:11,264 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.76 sec
2018-05-08 03:03:53,937 Stage-1 map = 100%, reduce = 13%, Cumulative CPU 10.11 sec
2018-05-08 03:03:56,474 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 11.74 sec
2018-05-08 03:04:01,148 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 14.5 sec
2018-05-08 03:04:07,864 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 19.03 sec
2018-05-08 03:04:09,944 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 19.71 sec
2018-05-08 03:04:47,210 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 29.0 sec
2018-05-08 03:04:53,605 Stage-1 map = 100%, reduce = 80%, Cumulative CPU 36.54 sec
2018-05-08 03:04:55,792 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 43.51 sec
2018-05-08 03:04:59,848 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 46.51 sec
2018-05-08 03:05:02,570 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 61.74 sec
2018-05-08 03:05:09,243 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 65.68 sec
MapReduce Total cumulative CPU time: 1 minutes 5 seconds 680 msec
Ended Job = job_1525726256153_0002
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 5 Cumulative CPU: 65.68 sec HDFS Read: 30840 HDFS Write: 3820 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 5 seconds 680 msec
OK
Time taken: 250.666 seconds
hive> select * from employee;
OK
1 a 1000 finance 1
1 a 1000 finance 1
2 b 2000 hr 2
2 b 2000 hr 2
3 c 3000 IT 3
3 c 3000 IT 3
4 d 4000 transport 4
4 d 4000 transport 4
Time taken: 0.961 seconds, Fetched: 8 row(s)
hive>
```

e. Updating a column (bucketed and non-bucketed)

The screenshot shows a MobaXterm window with a terminal session. The terminal displays the execution of a Hive job to update a table. The job is named 'job_1525726256153_0003'. The tracking URL is 'http://localhost:8088/proxy/application_1525726256153_0003/'. The kill command is '/home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0003'. The job information for Stage-1 shows 1 mapper and 5 reducers. The job is completed successfully with a cumulative CPU time of 1 minute 43 seconds 730 msec. The HDFS read and write statistics are 60397 and 1128 respectively. The job is ended with a success status.

```
hive> update employee set employee_id = 5 where employee_id = 4;
FAILED: SemanticException [Error 10302]: Updating values of bucketing columns is not supported. Column employee_id.
hive> update employee set dept_id = 05 where employee_id = 4;
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 = acadgild_20180508030901_d64aaf45-1ad4-47e7-875b-59719a499d2f
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 5
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_1525726256153_0003, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0003/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0003
Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 5
2018-05-08 03:09:35,553 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:10:36,998 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:11:38,016 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:12:42,003 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 56.38 sec
2018-05-08 03:13:43,007 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 56.38 sec
2018-05-08 03:14:44,968 Stage-1 map = 100%, reduce = 62%, Cumulative CPU 68.57 sec
2018-05-08 03:14:49,112 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 71.55 sec
2018-05-08 03:15:25,364 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 97.5 sec
2018-05-08 03:15:36,540 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 103.73 sec
MapReduce Total cumulative CPU time: 1 minutes 43 seconds 730 msec
Ended Job = job_1525726256153_0003
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 103.73 sec HDFS Read: 60397 HDFS Write: 1128 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 43 seconds 730 msec
OK
Time taken: 401.781 seconds
hive>
```

The screenshot shows a MobaXterm window with a terminal session on a remote host 127.0.0.1 (acadgild). The terminal displays the execution of a Hive job to insert data into a table named 'employee'. The job is configured with 5 mappers and 5 reducers. The output shows the job completed successfully with a cumulative CPU time of 103.73 seconds and 1128 HDFS writes. A file explorer on the left shows the directory structure of the remote host, including files like .abrt, .baseline, .cache, .config, .dbus, .edipose, .fontconfig, .gnome2, .gnome, .gnupg, .gvfs, .local, .m2, .mozilla, .nautilus, .oracle_jre_usage, .p2, .pk, .pulse, .scalade, .sqoop, .ssh, .svn, .thumbnails, and .tooling.

```
127.0.0.1 (acadgild)
Terminal Sessions View X server Tools Games Settings Macros Help
Session Servers Tools Games Sessions View Split MultExec Tunneling Packages Settings Help

Quick connect...
/home/acadgild/
Name Size (KB)
..
.abrt
.baseline
.cache
.config
.dbus
.edipose
.fontconfig
.gconf
.gnome2
.gnome
.gnupg
.gvfs
.local
.m2
.mozilla
.nautilus
.oracle_jre_usage
.p2
.pk
.pulse
.scalade
.sqoop
.ssh
.svn
.thumbnails
.tooling
Follow terminal folder

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_1525726256153_0003, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0003/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0003
Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 5
2018-05-08 03:09:35,553 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:10:30,998 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:11:38,816 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:12:42,003 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 56.38 sec
2018-05-08 03:13:43,007 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 56.38 sec
2018-05-08 03:14:44,968 Stage-1 map = 100%, reduce = 62%, Cumulative CPU 68.57 sec
2018-05-08 03:14:49,112 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 71.55 sec
2018-05-08 03:15:25,364 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 97.5 sec
2018-05-08 03:15:36,540 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 103.73 sec
MapReduce Total cumulative CPU time: 1 minutes 43 seconds 730 msec
Ended Job = job_1525726256153_0003
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 103.73 sec HDFS Read: 60397 HDFS Write: 1128 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 43 seconds 730 msec
OK
Time taken: 401.781 seconds
hive> select * from employee;
OK
1 a 1000 finance 1
1 a 1000 finance 1
2 b 2000 hr 2
2 b 2000 hr 2
3 c 3000 IT 3
3 c 3000 IT 3
4 d 4000 transport 5
4 d 4000 transport 5
Time taken: 1.038 seconds, Fetched: 8 row(s)
hive>
```

f. Deleting a row in Hive table

The screenshot shows a MobaXterm window with a terminal session on a remote host 127.0.0.1 (acadgild). The terminal displays the execution of a Hive job to delete a row from the 'employee' table. The job is configured with 5 mappers and 5 reducers. The output shows the job completed successfully with a cumulative CPU time of 101.06 seconds and 761 HDFS writes. A file explorer on the left shows the directory structure of the remote host, including files like .abrt, .baseline, .cache, .config, .dbus, .edipose, .fontconfig, .gnome2, .gnome, .gnupg, .gvfs, .local, .m2, .mozilla, .nautilus, .oracle_jre_usage, .p2, .pk, .pulse, .scalade, .sqoop, .ssh, .svn, .thumbnails, and .tooling.

```
127.0.0.1 (acadgild)
Terminal Sessions View X server Tools Games Settings Macros Help
Session Servers Tools Games Sessions View Split MultExec Tunneling Packages Settings Help

Quick connect...
/home/acadgild/
Name Size (KB)
..
.abrt
.baseline
.cache
.config
.dbus
.edipose
.fontconfig
.gconf
.gnome2
.gnome
.gnupg
.gvfs
.local
.m2
.mozilla
.nautilus
.oracle_jre_usage
.p2
.pk
.pulse
.scalade
.sqoop
.ssh
.svn
.thumbnails
.tooling
Follow terminal folder

hive> delete from employee where employee_id = 3;
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 = acadgild_20180508032034_b2561b7f-6c98-4de8-9a05-bf708ecc5216
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 5
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_1525726256153_0004, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0004/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0004
Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 5
2018-05-08 03:21:18,911 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:22:19,192 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:23:19,491 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:24:21,368 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 28.1 sec
2018-05-08 03:24:25,082 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 58.21 sec
2018-05-08 03:25:26,606 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 58.21 sec
2018-05-08 03:26:27,335 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 58.21 sec
2018-05-08 03:26:38,814 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 69.45 sec
2018-05-08 03:26:40,590 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 70.85 sec
2018-05-08 03:27:11,742 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 82.32 sec
2018-05-08 03:27:14,899 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 96.74 sec
2018-05-08 03:27:21,170 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 101.06 sec
MapReduce Total cumulative CPU time: 1 minutes 41 seconds 60 msec
Ended Job = job_1525726256153_0004
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 101.06 sec HDFS Read: 56275 HDFS Write: 761 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 41 seconds 60 msec
OK
Time taken: 413.803 seconds
hive>
```

127.0.0.1 (acadgild)

Terminal Sessions View X server Tools Games Settings Macros Help

Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help

Quick connect...

Sessions

Tools

Macros

Stop

/home/acadgild/

Name Size (KB)

- ..
- .abrt
- .beeline
- .cache
- .config
- .dbus
- .edipoe
- .fontconfig
- .gconf
- .gnome2
- .gnote
- .gnupg
- .gvfs
- .local
- .m2
- .mozilla
- .nautilus
- .oracle_jre_usage
- .p2
- .pki
- .pulse
- .scalade
- .sftp
- .ssh
- .swt
- .thumbnails
- .tooling

☐ Follow terminal folder

2. 127.0.0.1 (acadgild)

3. 127.0.0.1 (acadgild)

4. 127.0.0.1 (acadgild)

```
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_1525726256153_0004, Tracking URL = http://localhost:8088/proxy/application_1525726256153_0004/
Kill Command = /home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1525726256153_0004
Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 5
2018-05-08 03:21:18,911 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:22:19,192 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:23:19,491 Stage-1 map = 0%, reduce = 0%
2018-05-08 03:24:21,368 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 28.1 sec
2018-05-08 03:24:25,082 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 58.21 sec
2018-05-08 03:25:26,606 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 58.21 sec
2018-05-08 03:26:27,335 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 58.21 sec
2018-05-08 03:26:38,814 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 69.45 sec
2018-05-08 03:27:11,742 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 70.85 sec
2018-05-08 03:27:14,899 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 82.32 sec
2018-05-08 03:27:21,170 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 96.74 sec
2018-05-08 03:27:21,170 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 101.06 sec
MapReduce Total cumulative CPU time: 1 minutes 41 seconds 60 msec
Ended Job = job_1525726256153_0004
Loading data to table custom.employee
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 101.06 sec HDFS Read: 56275 HDFS Write: 761 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 41 seconds 60 msec
OK
Time taken: 413.803 seconds
hive> select * from employee;
OK
1 a 1000 finance 1
1 a 1000 finance 1
2 b 2000 hr 2
2 b 2000 hr 2
4 d 4000 transport 5
4 d 4000 transport 5
Time taken: 1.12 seconds, Fetched: 6 row(s)
hive>
```

UNREGISTERED VERSION - Please support MobaXterm by subscribing to the professional edition here: <https://mobaxterm.mobatek.net>

22:00 07-05-2018