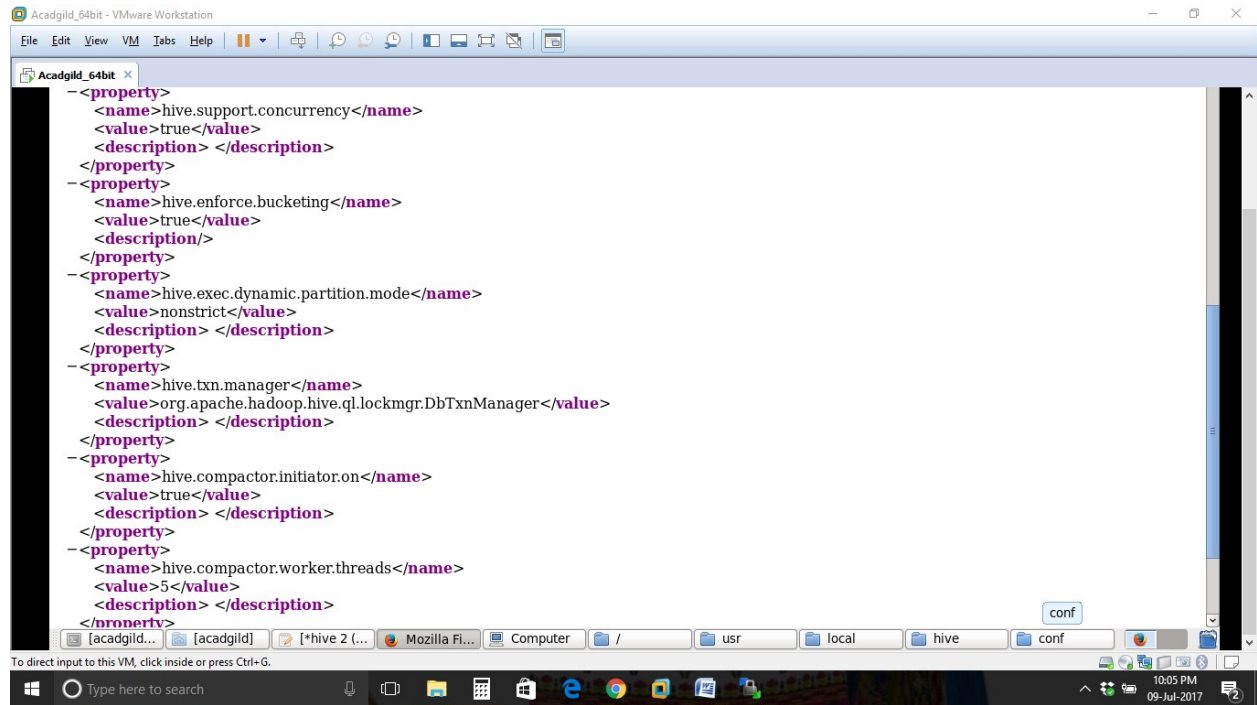


Assignment 16.3

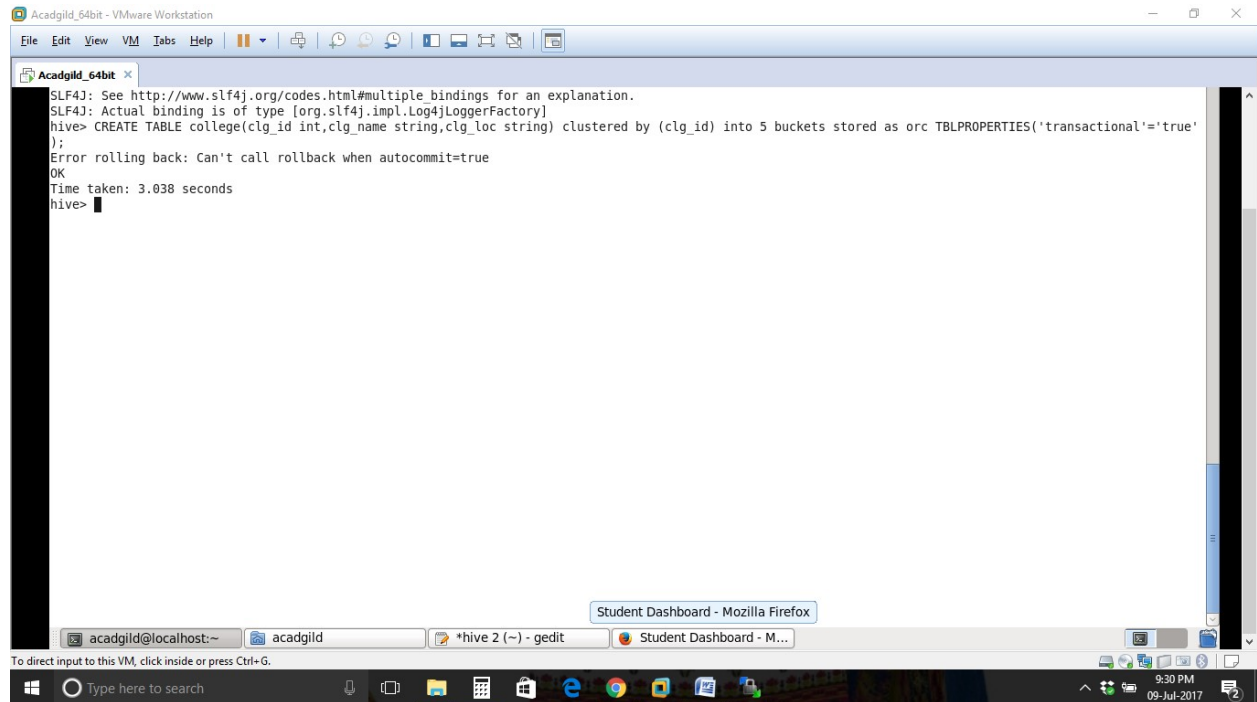
Transactions in Hive

ACID Properties Set in hive-site.xml file



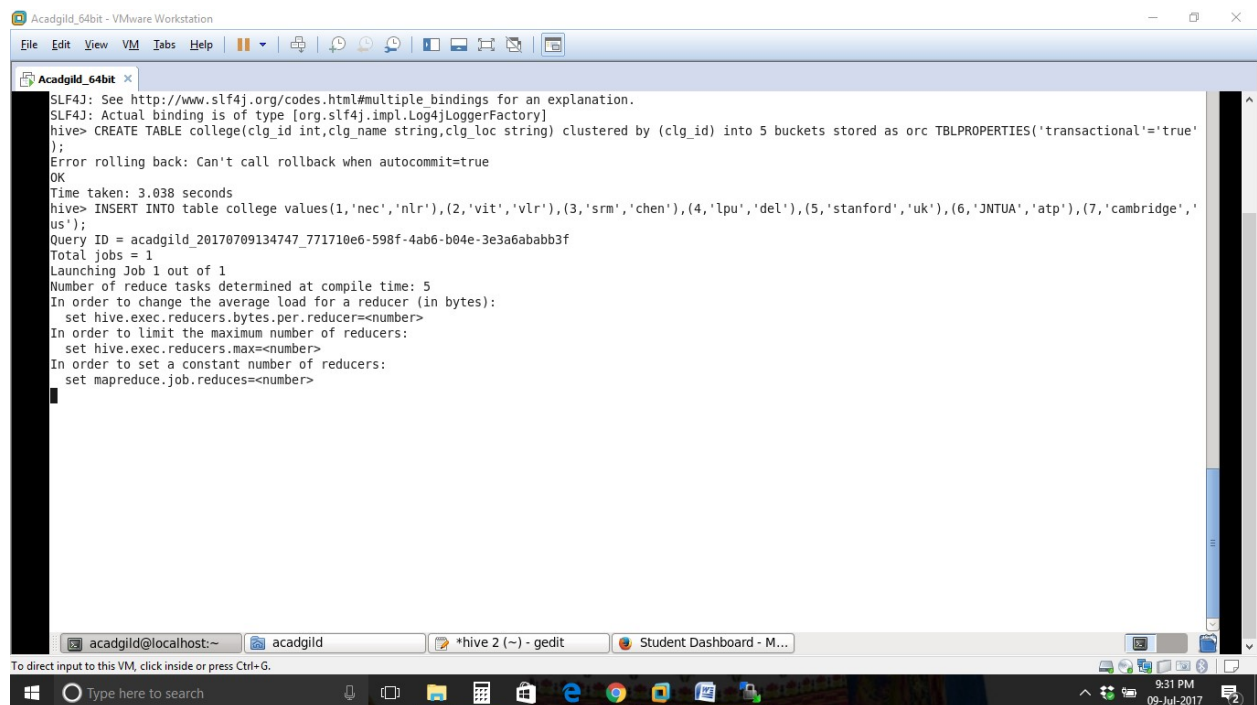
```
<?xml version="1.0"?>
<configuration>
  <property>
    <name>hive.support.concurrency</name>
    <value>true</value>
    <description> </description>
  </property>
  <property>
    <name>hive.enforce.bucketing</name>
    <value>true</value>
    <description> </description>
  </property>
  <property>
    <name>hive.exec.dynamic.partition.mode</name>
    <value>nonstrict</value>
    <description> </description>
  </property>
  <property>
    <name>hive.txn.manager</name>
    <value>org.apache.hadoop.hive.ql.lockmgr.DbTxnManager</value>
    <description> </description>
  </property>
  <property>
    <name>hive.compactor.initiator.on</name>
    <value>true</value>
    <description> </description>
  </property>
  <property>
    <name>hive.compactor.worker.threads</name>
    <value>5</value>
    <description> </description>
  </property>
</configuration>
```

CREATE TABLE college(clg_id int,clg_name string,clg_loc string) clustered by (clg_id) into 5 buckets stored as orc TBLPROPERTIES('transactional'='true');



```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
hive> CREATE TABLE college(clg_id int,clg_name string,clg_loc string) clustered by (clg_id) into 5 buckets stored as orc TBLPROPERTIES('transactional'='true');
Error rolling back: Can't call rollback when autoccommit=true
OK
Time taken: 3.038 seconds
hive>
```

INSERT INTO table college values (1,'nec','nlr'), (2,'vit','vlr'), (3,'srm','chen'), (4,'lpu','del'), (5,'stanford','uk'), (6,'JNTUA','atp'),(7,'cambridge','us');



```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
hive> CREATE TABLE college(clg_id int,clg_name string,clg_loc string) clustered by (clg_id) into 5 buckets stored as orc TBLPROPERTIES('transactional'='true');
Error rolling back: Can't call rollback when autoccommit=true
OK
Time taken: 3.038 seconds
hive> INSERT INTO table college values (1,'nec','nlr'), (2,'vit','vlr'), (3,'srm','chen'), (4,'lpu','del'), (5,'stanford','uk'), (6,'JNTUA','atp'), (7,'cambridge','us');
Query ID = acadgild_20170709134747_771710e6-598f-4ab6-b04e-3e3a6ababb3f
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>
hive>
```

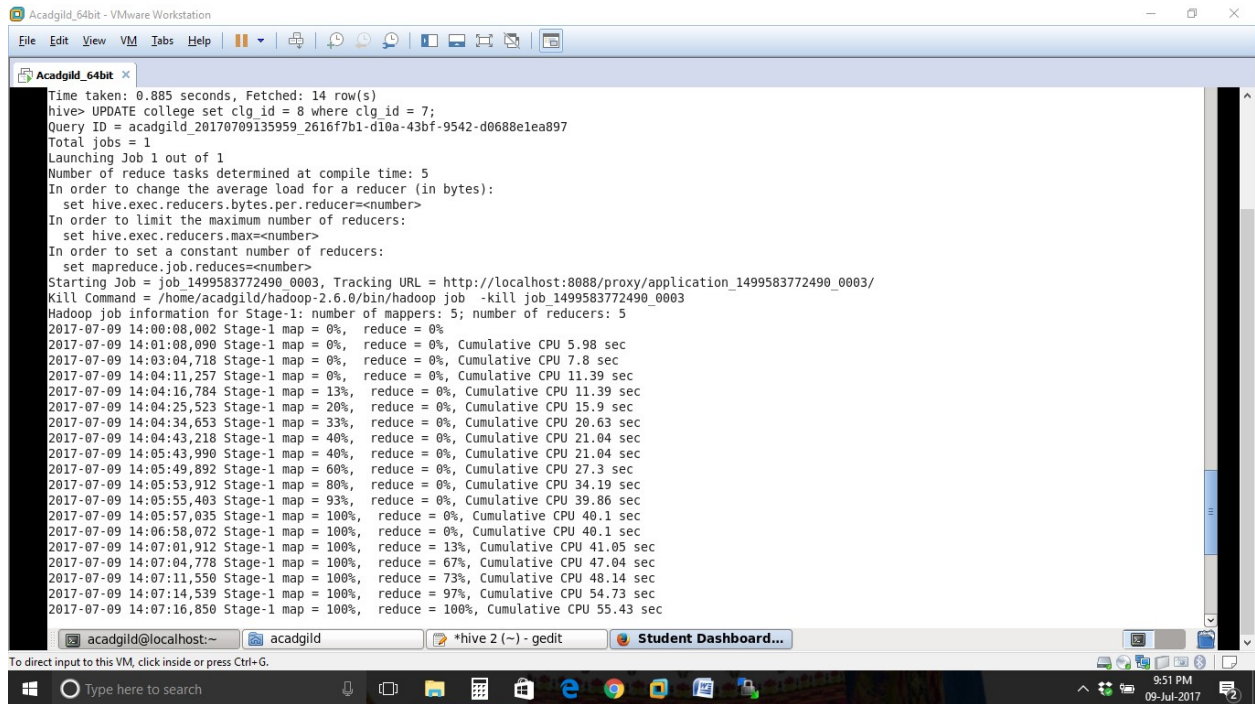
```
INSERT INTO table college values (1,'nec','nlr'), (2,'vit','vlr'), (3,'srm','chen'), (4,'lpu','del'),
(5,'stanford','uk'),(6,'JNTUA','atp'),(7,'cambridge','us');
```

```
us');
Query ID = acadgild_20170709134747_771710e6-598f-4ab6-b04e-3e3a6ababb3f
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_1499583772490_0001, Tracking URL = http://localhost:8088/proxy/application_1499583772490_0001/
Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job_1499583772490_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 5
2017-07-09 13:49:21,482 Stage-1 map = 0%, reduce = 0%
2017-07-09 13:49:58,320 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.58 sec
2017-07-09 13:50:59,167 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 7.58 sec
2017-07-09 13:51:04,194 Stage-1 map = 100%, reduce = 43%, Cumulative CPU 11.43 sec
2017-07-09 13:51:05,988 Stage-1 map = 100%, reduce = 56%, Cumulative CPU 11.68 sec
2017-07-09 13:51:07,619 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 13.31 sec
2017-07-09 13:51:24,003 Stage-1 map = 100%, reduce = 89%, Cumulative CPU 22.88 sec
2017-07-09 13:51:27,047 Stage-1 map = 100%, reduce = 92%, Cumulative CPU 25.03 sec
2017-07-09 13:51:45,817 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 32.59 sec
2017-07-09 13:51:47,132 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 35.54 sec
MapReduce Total cumulative CPU time: 35 seconds 540 msec
Ended Job = job_1499583772490_0001
Loading data to table default.college
Table default.college stats: [numFiles=5, numRows=7, totalSize=3495, rawDataSize=0]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 5 Cumulative CPU: 35.54 sec HDFS Read: 358 HDFS Write: 3870 SUCCESS
Total MapReduce CPU Time Spent: 35 seconds 540 msec
OK
Time taken: 267.226 seconds
hive>
```

Select * from college;

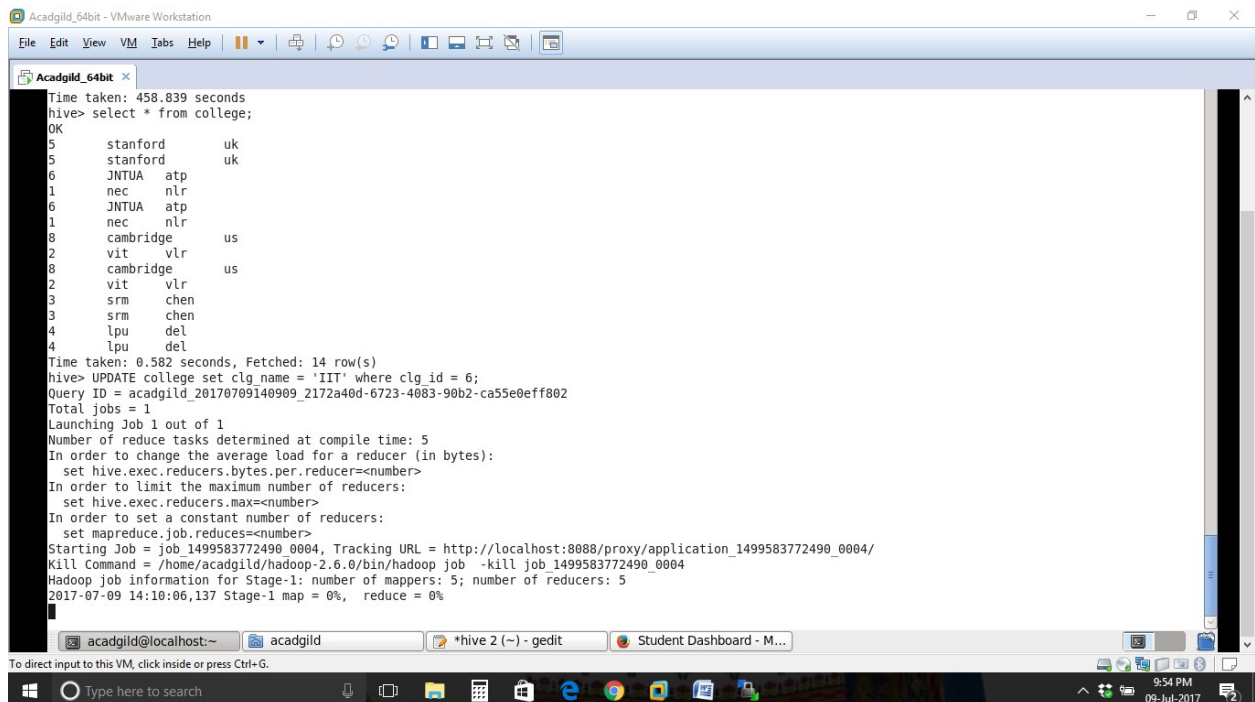
```
2017-07-09 13:54:24,774 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 11.35 sec
2017-07-09 13:54:29,360 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 13.82 sec
2017-07-09 13:54:33,634 Stage-1 map = 100%, reduce = 80%, Cumulative CPU 15.96 sec
2017-07-09 13:54:35,180 Stage-1 map = 100%, reduce = 84%, Cumulative CPU 16.98 sec
2017-07-09 13:54:36,634 Stage-1 map = 100%, reduce = 90%, Cumulative CPU 19.08 sec
2017-07-09 13:54:37,984 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 19.75 sec
2017-07-09 13:54:40,658 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 22.4 sec
MapReduce Total cumulative CPU time: 22 seconds 400 msec
Ended Job = job_1499583772490_0002
Loading data to table default.college
Table default.college stats: [numFiles=10, numRows=14, totalSize=6994, rawDataSize=0]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 5 Cumulative CPU: 22.77 sec HDFS Read: 358 HDFS Write: 3874 SUCCESS
Total MapReduce CPU Time Spent: 22 seconds 770 msec
OK
Time taken: 125.987 seconds
hive> select * from college;
OK
5      stanford      uk
5      stanford      uk
6      JNTUA        atp
1      nec           nlr
6      JNTUA        atp
1      nec           nlr
7      cambridge    us
2      vit           vlr
7      cambridge    us
2      vit           vlr
3      srm          chen
3      srm          chen
4      lpu          del
4      lpu          del
Time taken: 0.885 seconds, Fetched: 14 row(s)
hive>
```

UPDATE college set clg_id = 8 where clg_id = 7;



```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
Time taken: 0.885 seconds, Fetched: 14 row(s)
hive> UPDATE college set clg_id = 8 where clg_id = 7;
Query ID = acadgild_20170709135959_2616f7b1-d10a-43bf-9542-d0688e1ea897
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_1499583772490_0003, Tracking URL = http://localhost:8088/proxy/application_1499583772490_0003/
Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job_1499583772490_0003
Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 5
2017-07-09 14:00:08,002 Stage-1 map = 0%, reduce = 0%
2017-07-09 14:01:08,090 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 5.98 sec
2017-07-09 14:03:04,718 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 7.8 sec
2017-07-09 14:04:11,257 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 11.39 sec
2017-07-09 14:04:16,784 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 11.39 sec
2017-07-09 14:04:25,523 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 15.9 sec
2017-07-09 14:04:34,653 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 20.63 sec
2017-07-09 14:04:43,218 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 21.04 sec
2017-07-09 14:05:43,990 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 21.04 sec
2017-07-09 14:05:49,892 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 27.3 sec
2017-07-09 14:05:53,912 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 34.19 sec
2017-07-09 14:05:55,403 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 39.86 sec
2017-07-09 14:05:57,035 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 40.1 sec
2017-07-09 14:06:58,072 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 40.1 sec
2017-07-09 14:07:01,912 Stage-1 map = 100%, reduce = 13%, Cumulative CPU 41.05 sec
2017-07-09 14:07:04,778 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 47.04 sec
2017-07-09 14:07:11,550 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 48.14 sec
2017-07-09 14:07:14,539 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 54.73 sec
2017-07-09 14:07:16,850 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 55.43 sec
acadgild@localhost:~ acadgild *hive 2 (~) - gedit Student Dashboard...
```

UPDATE college set clg_name = 'IIT' where clg_id = 6;



```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
Time taken: 458.839 seconds
hive> select * from college;
OK
5      stanford      uk
5      stanford      uk
6      JNTUA      atp
1      nec      nlr
6      JNTUA      atp
1      nec      nlr
8      cambridge      us
2      vit      vlr
8      cambridge      us
2      vit      vlr
3      srm      chen
3      srm      chen
4      lpu      del
4      lpu      del
Time taken: 0.582 seconds, Fetched: 14 row(s)
hive> UPDATE college set clg_name = 'IIT' where clg_id = 6;
Query ID = acadgild_20170709140909_2172a40d-6723-4083-90b2-ca55e0eff802
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_1499583772490_0004, Tracking URL = http://localhost:8088/proxy/application_1499583772490_0004/
Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job_1499583772490_0004
Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 5
2017-07-09 14:10:06,137 Stage-1 map = 0%, reduce = 0%
acadgild@localhost:~ acadgild *hive 2 (~) - gedit Student Dashboard - M...
```

```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
2017-07-09 14:13:31,746 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 26.11 sec
2017-07-09 14:13:34,566 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 28.75 sec
2017-07-09 14:13:35,914 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 30.14 sec
2017-07-09 14:13:37,271 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 31.62 sec
2017-07-09 14:13:42,927 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 32.64 sec
2017-07-09 14:13:44,342 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 37.93 sec
2017-07-09 14:13:46,845 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 39.54 sec
MapReduce Total cumulative CPU time: 39 seconds 540 msec
Ended Job = job_1499583772490_0004
Loading data to table default.college
Table default.college stats: [numFiles=12, numRows=14, totalSize=8446, rawDataSize=0]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 39.54 sec HDFS Read: 10723 HDFS Write: 943 SUCCESS
Total MapReduce CPU Time Spent: 39 seconds 540 msec
OK
Time taken: 251.355 seconds
hive> select * from college;
OK
5      stanford      uk
5      stanford      uk
6      IIT      atp
1      nec      nlr
6      IIT      atp
1      nec      nlr
8      cambridge    us
2      vit      vlr
8      cambridge    us
2      vit      vlr
3      srm      chen
3      srm      chen
4      lpu      del
4      lpu      del
Time taken: 0.482 seconds, Fetched: 14 row(s)
hive>
```

acadgild@localhost:~ acadgild *hive 2 (~) - gedit Student Dashboard...

To direct input to this VM, click inside or press Ctrl+G.

Type here to search 9:57 PM 09-Jul-2017

delete from college where clg_id=5;

```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 39.54 sec HDFS Read: 10723 HDFS Write: 943 SUCCESS
Total MapReduce CPU Time Spent: 39 seconds 540 msec
OK
Time taken: 251.355 seconds
hive> select * from college;
OK
5      stanford      uk
5      stanford      uk
6      IIT      atp
1      nec      nlr
6      IIT      atp
1      nec      nlr
8      cambridge    us
2      vit      vlr
8      cambridge    us
2      vit      vlr
3      srm      chen
3      srm      chen
4      lpu      del
4      lpu      del
Time taken: 0.482 seconds, Fetched: 14 row(s)
hive> delete from college where clg_id=5;
Query ID = acadgild_20170709141515_f21ccfa1-0d19-4a3f-a1fe-006f4584dc80
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>
```

acadgild@localhost:~ acadgild *hive 2 (~) - gedit Student Dashboard - Mozilla Firefox Student Dashboard - M...

To direct input to this VM, click inside or press Ctrl+G.

Type here to search 9:58 PM 09-Jul-2017

Acadgild_64bit - VMware Workstation

File Edit View VM Tabs Help

Acadgild_64bit

```
2017-07-09 14:16:44,085 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 9.76 sec
2017-07-09 14:16:53,234 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 15.89 sec
2017-07-09 14:17:05,867 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 20.75 sec
2017-07-09 14:18:05,990 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 20.75 sec
2017-07-09 14:18:17,211 Stage-1 map = 100%, reduce = 13%, Cumulative CPU 21.67 sec
2017-07-09 14:18:20,037 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 27.77 sec
2017-07-09 14:18:26,796 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 31.8 sec
2017-07-09 14:18:29,576 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 33.25 sec
2017-07-09 14:18:32,046 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 36.01 sec
MapReduce Total cumulative CPU time: 36 seconds 10 msec
Ended Job = job_1499583772490_0005
Loading data to table default.college
Table default.college stats: [numFiles=13, numRows=12, totalSize=8944, rawDataSize=0]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 36.01 sec HDFS Read: 11664 HDFS Write: 737 SUCCESS
Total MapReduce CPU Time Spent: 36 seconds 10 msec
OK
Time taken: 209.918 seconds
hive> select * from college;
OK
6      IIT      atp
1      nec      nlr
6      IIT      atp
1      nec      nlr
8      cambridge  us
2      vit      vlr
8      cambridge  us
2      vit      vlr
3      srm      chen
3      srm      chen
4      lpu      del
4      lpu      del
Time taken: 4.801 seconds, Fetched: 12 row(s)
hive>
```

acadgild@localhost:~ acadgild *hive 2 (~) - gedit Student Dashboard - M...

To direct input to this VM, click inside or press Ctrl+G.

Type here to search

10:02 PM 09-Jul-2017