

## Task 1

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

```

hive> select country, sum(total) from olympic where sport="Swimming" group by country;
WARNING: Hive-on-HR 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 = acedgild_20180520220327_000001795-7704-4e30-bf4c-2270a2054200
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_1526792975751_0000, Tracking URL = http://localhost:8080/proxy/application_1526792975751_0000/
Kill Command = /home/acedgild/install/hadoop/hadoop-2.6.0/bin/hadoop job -kill job_1526792975751_0000
Hadoop job information for Stage-1: number of mappers: 1, number of reducers: 1
2018-05-20 22:03:40,330 Stage-1 map = 0%, reduce = 0%
2018-05-20 22:03:54,912 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.95 sec
2018-05-20 22:04:00,450 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.09 sec
MapReduce Total cumulative CPU time: 5 seconds 90 msec
Ended Job = job_1526792975751_0000
HadoopJob Launched:
Stage-Stage 1: Map: 1 Reducers: 1 Cumulative CPU: 5.09 sec HDFS Read: 528574 HDFS Write: 831 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 90 msec
OK
Argentina      1
Australia      153
Austria        3
Belarus        2
Brazil         8
Canada         5
China          30
Costa Rica     2
Croatia        1
Denmark        1
France         39
Germany        32
Great Britain  11
Hungary        9

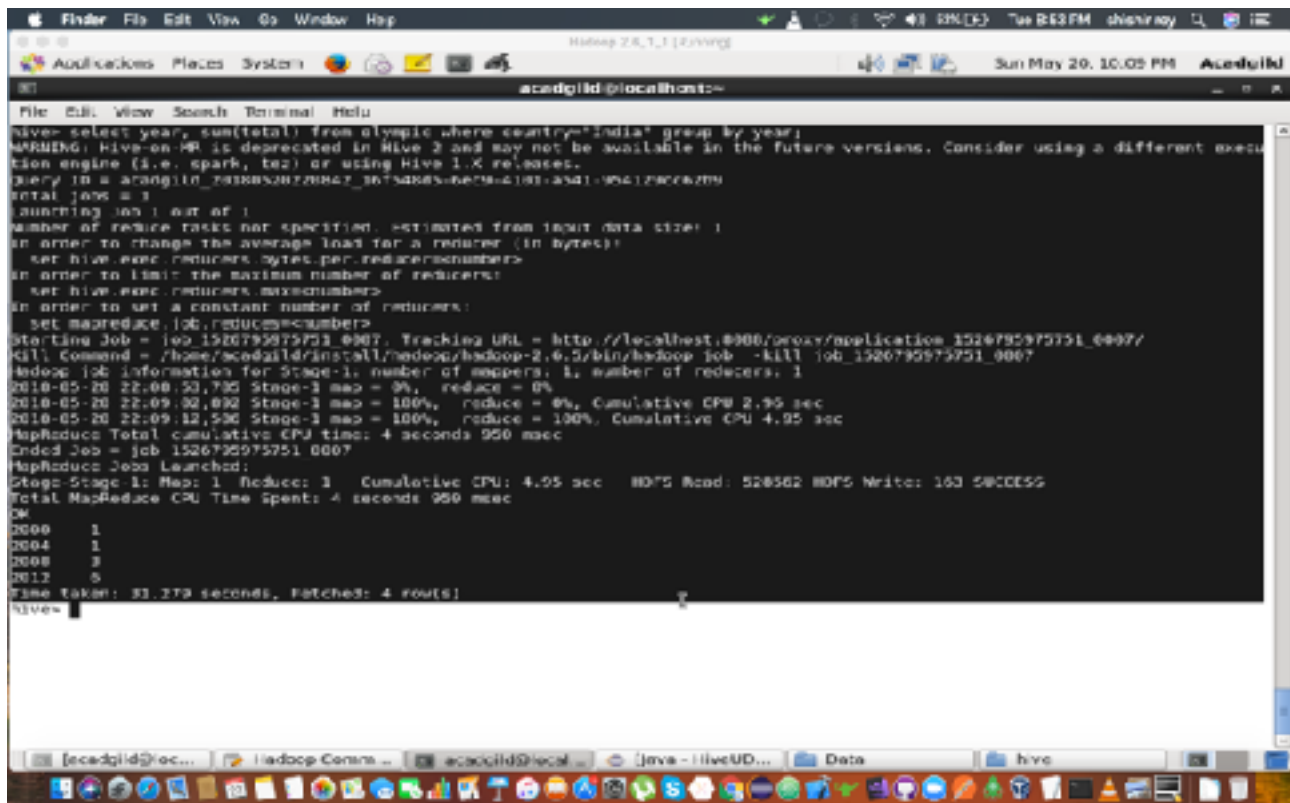
```

```

Stage-Stage 1: Map: 1 Reducers: 1 Cumulative CPU: 5.09 sec HDFS Read: 528574 HDFS Write: 831 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 90 msec
OK
Argentina      1
Australia      153
Austria        3
Belarus        2
Brazil         8
Canada         5
China          30
Costa Rica     2
Croatia        1
Denmark        1
France         39
Germany        32
Great Britain  11
Hungary        9
Italy          16
Japan          43
Lithuania      1
Netherlands    45
Norway         2
Poland         3
Romania        6
Russia         20
Serbia         1
Slovakia       2
Slovenia       1
South Africa   11
South Korea    4
Spain          5
Sweden         9
Trinidad and Tobago 1
Tunisia        3
Ukraine        7
United States  257
Zimbabwe       7
Time taken: 30.88 seconds, Fetched: 34 row(s)
hive>

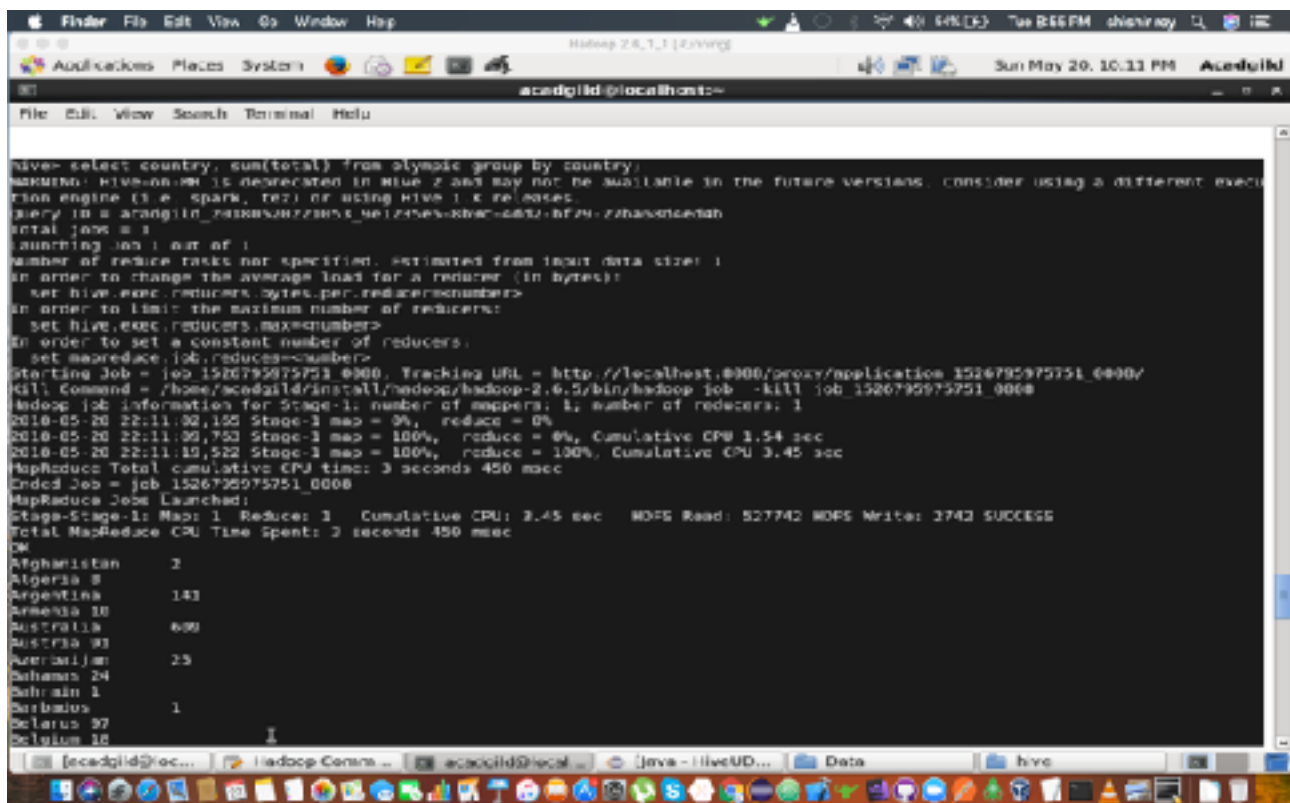
```

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



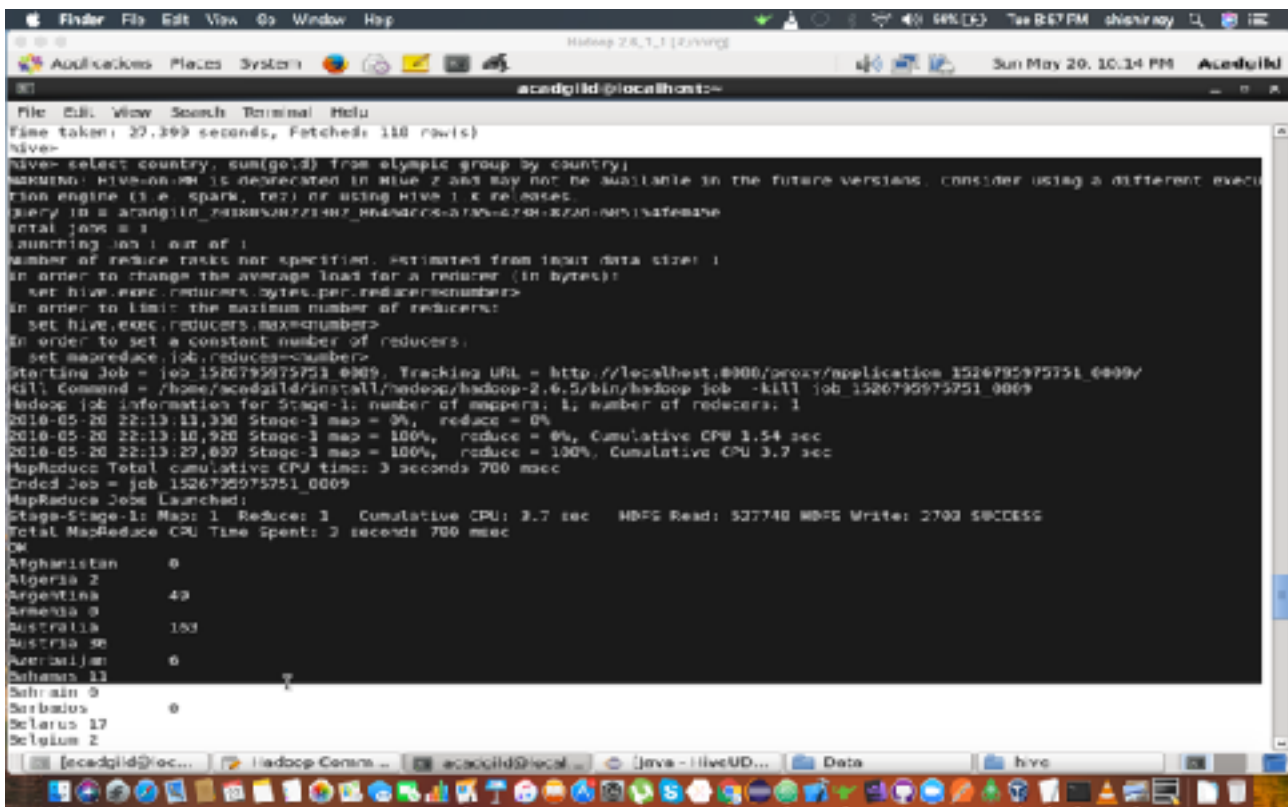
```
hive> select year, sum(total) from olympic where country="India" group by year;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
query id = acedgild_20180520220047_1675485-bc9a-4103-aa01-00412ac0a200
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_1526792975751_0007, Tracking URL = http://localhost:8080/proxy/application_1526792975751_0007/
Kill Command = /home/acedgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526792975751_0007
Hadoop job information for Stage-1: number of mappers: 1, number of reducers: 1
2018-05-20 22:00:20,730 Stage-1 map = 0%, reduce = 0%
2018-05-20 22:00:32,000 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.95 sec
2018-05-20 22:00:32,530 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.95 sec
MapReduce Total cumulative CPU time: 4 seconds 950 msec
Ended Job = job_1526792975751_0007
MapReduce Jobs Launched:
  Stage-1: Map: 1 Reduc: 1 Cumulative CPU: 4.95 sec  HDFS Read: 528562 HDFS Write: 163 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 950 msec
OK
2008      1
2004      1
2008      3
2012      6
Time taken: 33.273 seconds, Fetched: 4 row(s)
```

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



```
hive> select country, sum(total) from olympic 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 = acedgild_20180520221107_1675485-bc9a-4aa2-b779-27ba9dceda00
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_1526792975751_0008, Tracking URL = http://localhost:8080/proxy/application_1526792975751_0008/
Kill Command = /home/acedgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526792975751_0008
Hadoop job information for Stage-1: number of mappers: 1, number of reducers: 1
2018-05-20 22:11:02,130 Stage-1 map = 0%, reduce = 0%
2018-05-20 22:11:09,730 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.54 sec
2018-05-20 22:11:10,522 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.45 sec
MapReduce Total cumulative CPU time: 3 seconds 450 msec
Ended Job = job_1526792975751_0008
MapReduce Jobs Launched:
  Stage-1: Map: 1 Reduc: 1 Cumulative CPU: 3.45 sec  HDFS Read: 527742 HDFS Write: 2743 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 450 msec
OK
Afghanistan      2
Algeria           8
Argentina         143
Armenia           10
Australia         609
Austria          93
Azerbaijan       23
Bahamas          24
Bahrain          1
Belarus           1
Belgium          97
Belgium          16
```

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



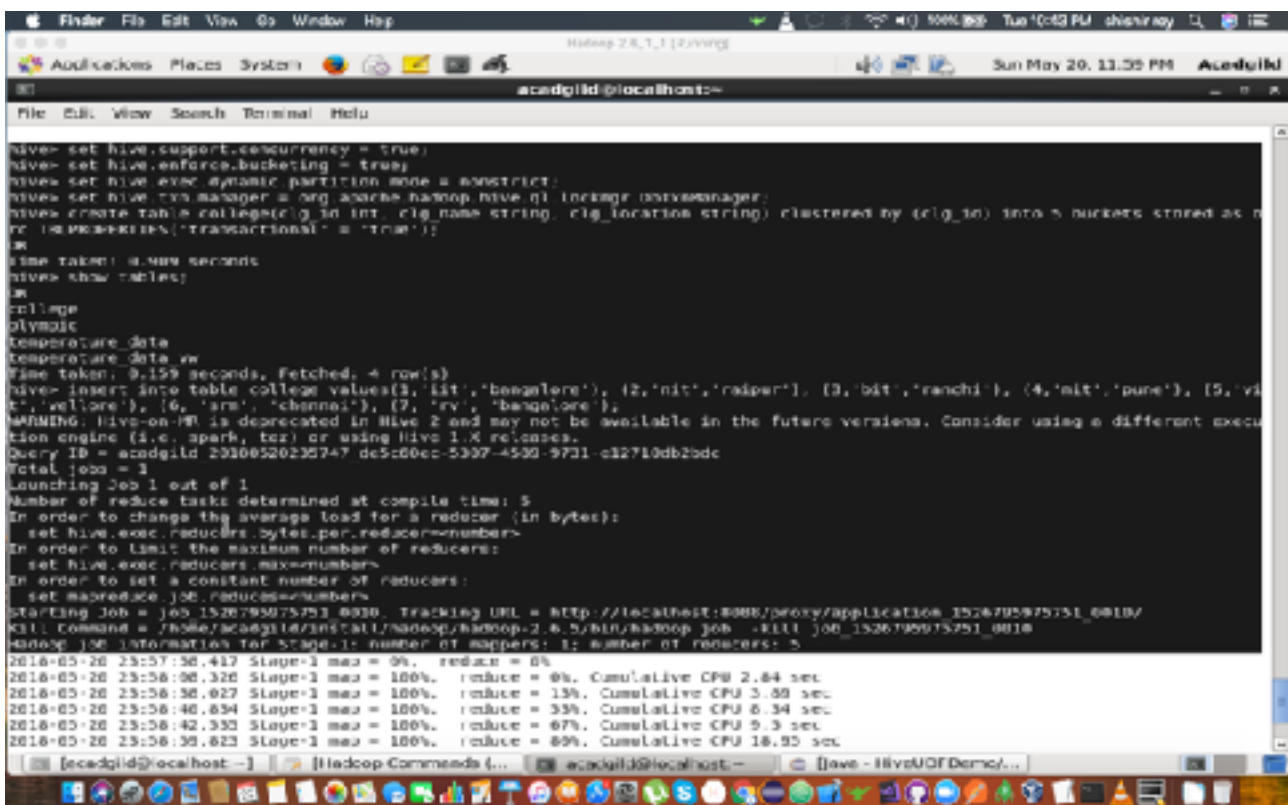
```
acacgild@localhost:~$ hive
Time taken: 27.399 seconds, Fetched: 110 row(s)
hive>
hive> select country, sum(gold) from olympic group by country;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.x releases.
Query ID = acacgild_20180520221407_0040accc-0700-0700-0000-000000000000
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_1526795975751_0009, Tracking URL = http://localhost:8080/proxy/application_1526795975751_0009/
Kill Command = /home/acacgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526795975751_0009
Hadoop job information for Stage-1: number of mappers: 1, number of reducers: 1
2018-05-20 22:13:11,330 Stage-1 map = 0%, reduce = 0%
2018-05-20 22:13:10,920 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.54 sec
2018-05-20 22:13:27,007 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.7 sec
HadoopMapReduce Total cumulative CPU time: 3 seconds 700 msec
Ended Job = job_1526795975751_0009
HadoopMapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reducer: 1 Cumulative CPU: 3.7 sec HDFS Read: 527740 HDFS Write: 2700 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 700 msec
OK
Afghanistan      0
Algeria          2
Argentina        49
Armenia          0
Australia        163
Austria          38
Azerbaijan       6
Bahamas          11
Bahrain          0
Barbados         0
Belarus          17
Belgium          2
```

## Task 2

Link: <https://acacgild.com/blog/transactions-in-hive/>

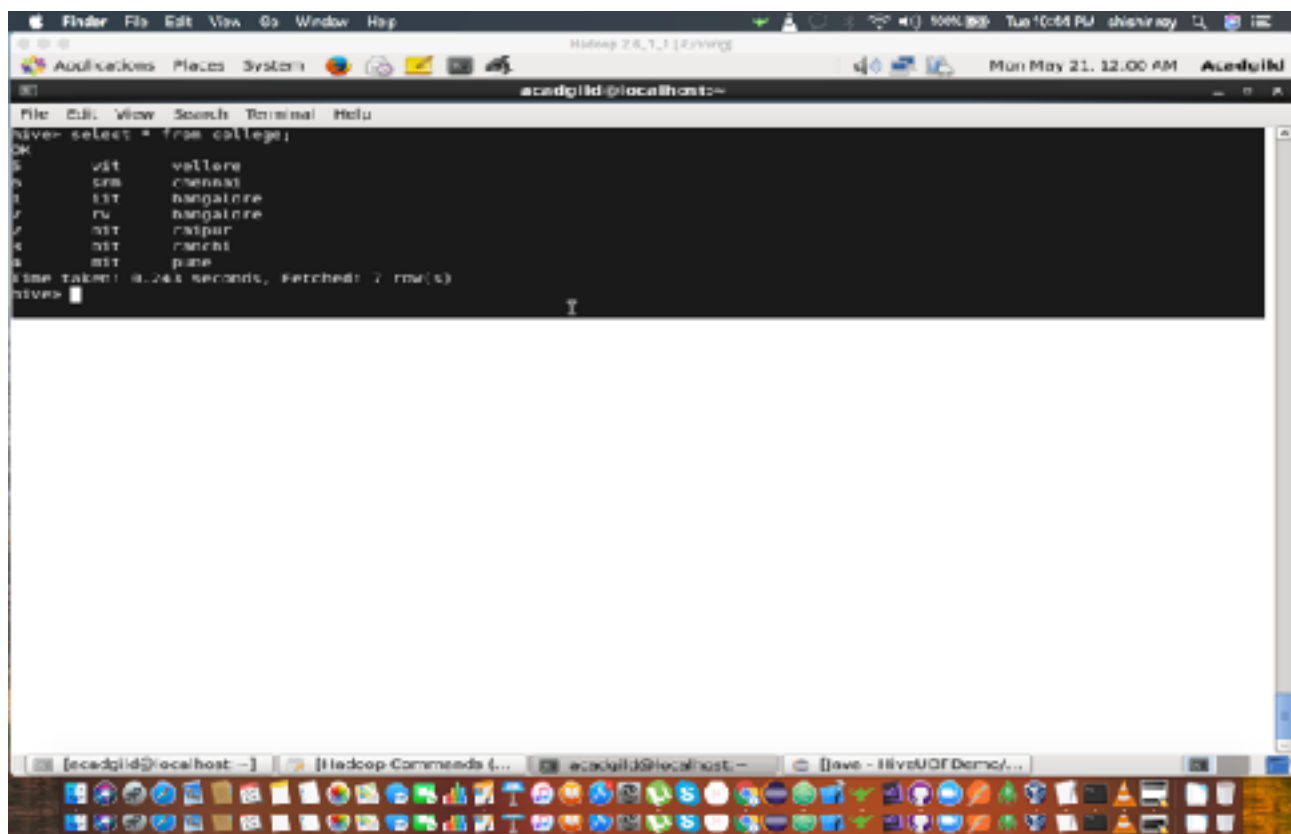
Refer the above given link for transactions in Hive and implement the operations given in the blog using your own sample data set and send us the screenshot.

Hive Properties with transaction, creation of table and insert value in hive table



```
acacgild@localhost:~$ hive
Time taken: 0.499 seconds
hive> show tables;
OK
college
olympic
temperature_data
temperature_data_mv
Time taken: 0.159 seconds, Fetched: 4 row(s)
hive> insert into table college values(1, 'bangalore'), (2, 'nit', 'raipet'), (3, 'bit', 'ranchi'), (4, 'nit', 'pune'), (5, 'vit', 'vellore'), (6, 'ara', 'chennai'), (7, 'rv', 'bangalore');
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 = acacgild_201805202235747_d5c5d0ec-5307-4530-9731-c32710db2bdc
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_1526795975751_0010, Tracking URL = http://localhost:8080/proxy/application_1526795975751_0010/
Kill Command = /home/acacgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526795975751_0010
Hadoop job information for Stage-1: number of mappers: 1, number of reducers: 5
2018-05-20 23:57:36,417 Stage-1 map = 0%, reduce = 0%
2018-05-20 23:58:06,326 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.84 sec
2018-05-20 23:58:06,027 Stage-1 map = 100%, reduce = 13%, Cumulative CPU 3.09 sec
2018-05-20 23:58:46,834 Stage-1 map = 100%, reduce = 33%, Cumulative CPU 6.34 sec
2018-05-20 23:58:42,332 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 9.3 sec
2018-05-20 23:58:39,623 Stage-1 map = 100%, reduce = 89%, Cumulative CPU 18.95 sec
```

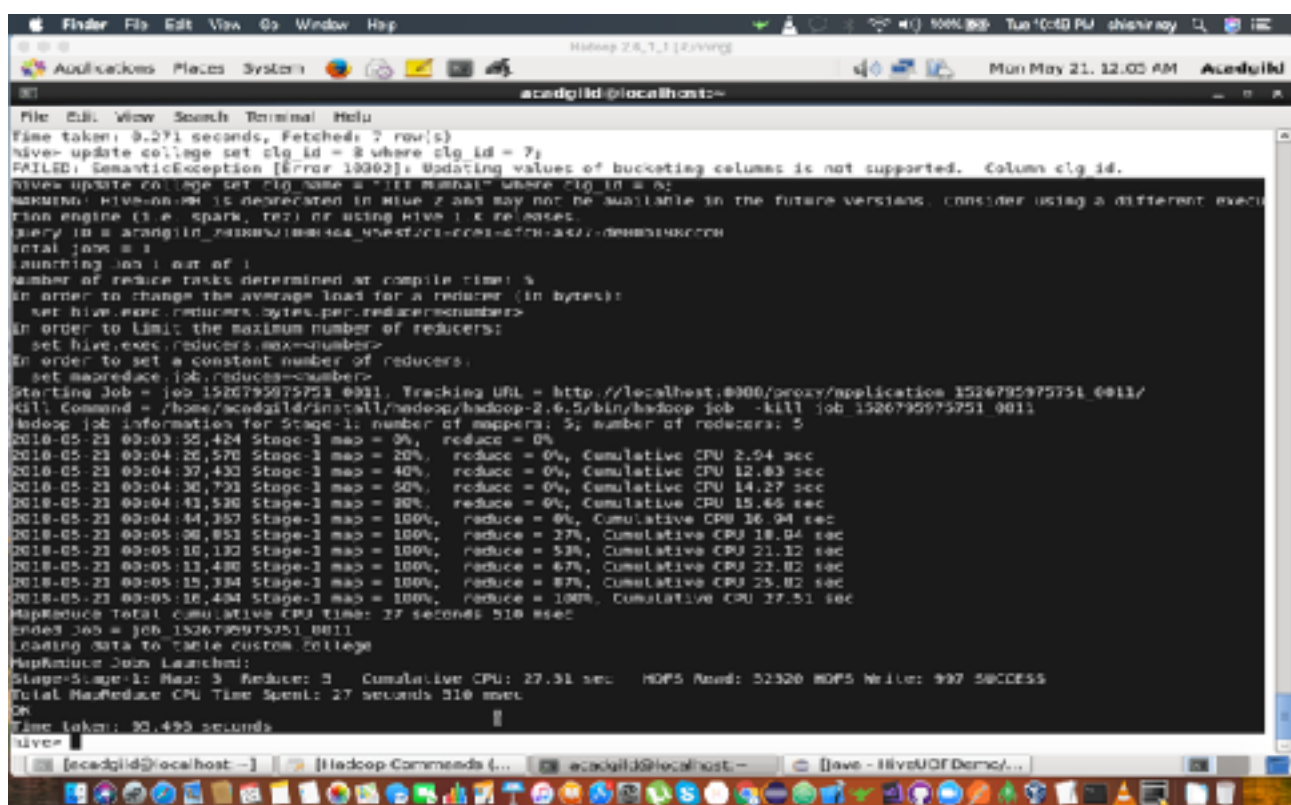
Select value in table



A terminal window titled 'accdgild@localhost:~' shows a Hive query execution. The command is 'hive> select \* from college;'. The output is a table with 7 rows and 3 columns: 'colg\_id', 'colg\_name', and 'colg\_desc'. The rows are: (1, 'vit', 'vallure'), (2, 'srn', 'chennai'), (3, 'ist', 'bangalore'), (4, 'ru', 'bangalore'), (5, 'nit', 'raipur'), (6, 'nit', 'raichur'), and (7, 'nit', 'pune'). Below the table, it says 'Time taken: 0.241 seconds, Fetched: 7 row(s)'. The prompt 'hive>' is visible at the bottom.

```
hive> select * from college;
OK
colg_id  colg_name  colg_desc
1        vit        vallure
2        srn        chennai
3        ist        bangalore
4        ru         bangalore
5        nit        raipur
6        nit        raichur
7        nit        pune
Time taken: 0.241 seconds, Fetched: 7 row(s)
hive>
```

Update table with bucketing columns



A terminal window titled 'accdgild@localhost:~' shows a Hive update command and its execution details. The command is 'hive> update college set colg\_id = 8 where colg\_id = 7;'. The output shows a 'FAILED: SemanticException [Error 10002]: Updating values of bucketing columns is not supported. Column colg\_id.' followed by a message about the deprecated 'hive-noset' option. Below this, it shows the execution details for the job, including the number of mappers and reducers, and the cumulative CPU time. The prompt 'hive>' is visible at the bottom.

```
hive> update college set colg_id = 8 where colg_id = 7;
FAILED: SemanticException [Error 10002]: Updating values of bucketing columns is not supported. Column colg_id.
hive> update college set colg_name = "nit mumbai" where colg_id = 6;
WARNING: Hive-noset 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 = accdgild_201805210004_000000_000000_000000_000000_000000
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_1526795975751_0011, Tracking URL = http://localhost:8080/proxy/application_1526795975751_0011/
Kill Command = /home/accdgild/install/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526795975751_0011
Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 5
2018-05-21 00:03:55,424 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 2.94 sec
2018-05-21 00:04:20,570 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 12.83 sec
2018-05-21 00:04:37,433 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 14.27 sec
2018-05-21 00:04:30,723 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 15.45 sec
2018-05-21 00:04:41,520 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 15.45 sec
2018-05-21 00:04:44,357 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.94 sec
2018-05-21 00:05:06,853 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 18.94 sec
2018-05-21 00:05:10,132 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 21.12 sec
2018-05-21 00:05:11,480 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 22.82 sec
2018-05-21 00:05:15,334 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 25.82 sec
2018-05-21 00:05:16,434 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 27.51 sec
MapReduce Total cumulative CPU time: 27 seconds 510 msec
Ended job = job_1526795975751_0011
Loading data to table custom.college
MapReduce Job Launches:
Stage-1: Map: 5 Reducers: 5 Cumulative CPU: 27.51 sec HDFS Read: 52520 HDFS Write: 937 SUCCESS
Total MapReduce CPU Time Spent: 27 seconds 510 msec
OK
Time taken: 30.493 seconds
hive>
```



```
File Edit View Search Terminal Help
Hadoop 2.6.1 [Ctrl+G]
Applications Places System
acacgild@localhost:~$
hive> select * from college;
OK
+----+-----+
| s   | vit   | vellore |
| s   | csm   | chennai |
| s   | iit   | bangalore |
| s   | ru    | bangalore |
| s   | nit   | raipur   |
| s   | nit   | ranchi   |
| s   | nit   | pune     |
+----+-----+
Time taken: 0.271 seconds, Fetched: 7 row(s)
hive> update college set cld_id = s where cld_id = 2;
FAILED: SemanticException [Error 10002]: updating values of bucketing columns is not supported. column cld_id.
hive> update college set cld_name = "iit Mumbai" where cld_id = 0;
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 = acacgild_20100521000344_95e3f2c1-ccc3-4fc0-a377-de000190ccc0
Total jobs = 3
Launching Job 1 out of 3
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_1526795975751_0011, Tracking URL = http://localhost:8080/proxy/application_1526795975751_0011/
Kill Command = /home/acacgild/instal/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526795975751_0011
Hadoop job information for stage-1: number of mappers: 5; number of reducers: 5
2010-05-23 00:03:55,424 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 2.94 sec
2010-05-23 00:04:20,570 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 12.83 sec
2010-05-23 00:04:37,483 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 14.27 sec
2010-05-23 00:04:38,793 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 14.27 sec
```

Select table with updating table

```
File Edit View Search Terminal Help
Hadoop 2.6.1 [Ctrl+G]
Applications Places System
acacgild@localhost:~$
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_1526795975751_0011, Tracking URL = http://localhost:8080/proxy/application_1526795975751_0011/
Kill Command = /home/acacgild/instal/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526795975751_0011
Hadoop job information for stage-1: number of mappers: 5; number of reducers: 5
2010-05-23 00:04:38,793 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 14.27 sec
2010-05-23 00:04:43,530 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 15.69 sec
2010-05-23 00:04:44,357 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.94 sec
2010-05-23 00:05:06,053 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 18.94 sec
2010-05-23 00:05:10,132 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 21.12 sec
2010-05-23 00:05:11,430 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 22.02 sec
2010-05-23 00:05:15,334 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 25.02 sec
2010-05-23 00:05:16,494 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 27.51 sec
MapReduce Total cumulative CPU time: 27 seconds 510 msec
Ended Job = job_1526795975751_0011
Loading data to table custom.college
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reducers: 5 Cumulative CPU: 27.51 sec HDFS Read: 53520 HDFS Write: 997 SUCCESS
Total MapReduce CPU Time Spent: 27 seconds 510 msec
OK
Time taken: 26.495 seconds
hive> select * from college;
OK
+----+-----+
| s   | vit   | vellore |
| s   | iit Mumbai | chennai |
| s   | iit   | bangalore |
| s   | ru    | bangalore |
| s   | nit   | raipur   |
| s   | nit   | ranchi   |
| s   | nit   | pune     |
+----+-----+
Time taken: 0.28 seconds, Fetched: 7 row(s)
hive>
```

## Delete data from table

```
File Edit View Go Window Help
Hadoop 2.6.1 [Living]
Acadgild Places System
acadgild@localhost:~$
File Edit View Search Terminal Help

Hives delete from college where <id> is <id>
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_hives21000276_eb1d4ee5-b7b6-4aa-wf0/-aa1c0bf0a5
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_1526795975751_0012, Tracking URL = http://localhost:8000/proxy/application-1526795975751_0012/
Hive Command = /home/acadgild/installs/hadoop/hadoop-2.6.5/bin/hadoop job -kill job_1526795975751_0012
Hadoop job information for stage=1: number of mappers: 5; number of reducers: 5
2018-05-23 09:00:30,577 Stage-1 map = 0%, reduce = 0%
2018-05-23 09:00:42,636 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 5.57 sec
2018-05-23 09:00:45,497 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 9.08 sec
2018-05-23 09:00:48,025 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 12.15 sec
2018-05-23 09:00:50,710 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 15.23 sec
2018-05-23 09:00:53,987 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 18.30 sec
2018-05-23 09:00:57,256 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 19.06 sec
2018-05-23 09:01:00,890 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 19.86 sec
2018-05-23 09:01:04,472 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 21.48 sec
2018-05-23 09:01:07,534 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 22.02 sec
MapReduce Total cumulative CPU time: 22 seconds 920 msec
Ended job = job_1526795975751_0012
Loading data to table custom.college
MapReduce jobs Launched:
Stage=Stage-1: Map: 5 Reduce: 5 Cumulative CPU: 22.92 sec HDFS Read: 50577 HDFS Write: 747 SUCCESS
Total MapReduce CPU Time Spent: 22 seconds 920 msec
OK
Time taken: 56.779 seconds
hive select * from college;
OK
[acadgild@localhost ~] [Hadoop Commands (...)] [acadgild@localhost ~] [Java - HiveUDF Demo (...)]
```

The screenshot shows a terminal window with the following content:

```

File Edit View Search Terminal Help
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_1526795975751_0012, tracking url = http://localhost:8080/proxy/application_1526795975751_0012/
kill command = /home/scadgild@scadgild:~/hadoop/bin/hadoop-2.6.0/bin/hadoop job -kill job_1526795975751_0012
Hadoop job information for stage-1: number of mappers: 5; number of reducers: 5
2016-05-21 09:08:04,572 Stage-1 map = 0%, reduce = 0%
2016-05-21 09:08:42,630 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 3.57 sec
2016-05-21 09:08:45,497 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 6.08 sec
2016-05-21 09:08:46,095 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 9.15 sec
2016-05-21 09:08:56,710 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 12.35 sec
2016-05-21 09:09:51,937 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 22.09 sec
2016-05-21 09:09:17,256 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 16.96 sec
2016-05-21 09:09:19,030 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 18.06 sec
2016-05-21 09:09:22,472 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 21.40 sec
2016-05-21 09:09:23,594 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 22.92 sec
MapReduce Total cumulative CPU time: 22 seconds 920 msec
Ended Job = job_1526795975751_0012
Loading data to table custom.college
MapReduce Jobs Launched:
Stage-Stage-1: Map: 5 Reduc: 5 Cumulative CPU: 22.92 sec HDFS Read: 54677 HDFS Write: 747 SUCCESS
Total MapReduce CPU Time Spent: 22 seconds 920 msec
OK
Time taken: 38.773 seconds
hive> select * from college;
OK
 0      EIT Mumbai Chennai
 1      EIT Bangalore
 2      RV Bangalore
 3      MIT Manipal
 4      MIT Ranchi
 5      MIT Pune
Time taken: 9.223 seconds, Fetched: 6 row(s)
hive>

```

