**Assignment 7.2**

hive> **create table olympic**

(athlete string,age int,country string,year string,c\_date string,sports string,gold\_medal int,silver\_medal int,bronze\_medal int,total\_medal int)

row format delimited fields terminated by '\t';

OK

Time taken: 0.557 seconds

hive>

hive> **LOAD DATA LOCAL INPATH '/home/acadgild/hive/olympix\_data.csv'INTO TABLE olympic;**

Loading data to table b1.olympic

Table b1.olympic stats: [numFiles=1, totalSize=518669]

OK

Time taken: 0.763 seconds

hive>

-----------------------------------------------------------------------------------

**--Problem 1**

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

**hive> SELECT country,SUM(total\_medal) FROM olympic where sports='Swimming' group by country ;**

Query ID = acadgild\_20171117222222\_0d6ccf4d-07d7-454c-b02d-267c3194a285

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\_1505655198300\_0016, Tracking URL = http://localhost:8088/proxy/application\_1505655198300\_0016/

Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job\_1505655198300\_0016

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

2017-11-17 22:22:44,570 Stage-1 map = 0%, reduce = 0%

2017-11-17 22:23:05,845 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.42 sec

2017-11-17 22:23:21,705 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.39 sec

MapReduce Total cumulative CPU time: 7 seconds 390 msec

Ended Job = job\_1505655198300\_0016

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.39 sec HDFS Read: 518895 HDFS Write: 386 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 390 msec

OK

**country \_c1**

**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 9**

**Trinidad and Tobago 1**

**Tunisia 3**

**Ukraine 7**

**United States 267**

**Zimbabwe 7**

**Time taken: 61.647 seconds, Fetched: 34 row(s)**

**hive>**

**---------------------------------------------------------------------------------**

**--problem 2**

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

**hive> SELECT year,SUM(total\_medal) FROM olympic where country='India' group by year ;**

Query ID = acadgild\_20171117223333\_69f0598b-9cfd-4ded-a907-9db3b47f3b3c

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\_1505655198300\_0017, Tracking URL = http://localhost:8088/proxy/application\_1505655198300\_0017/

Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job\_1505655198300\_0017

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

2017-11-17 22:33:44,818 Stage-1 map = 0%, reduce = 0%

2017-11-17 22:34:04,912 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.01 sec

2017-11-17 22:34:25,817 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.79 sec

MapReduce Total cumulative CPU time: 7 seconds 790 msec

Ended Job = job\_1505655198300\_0017

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.79 sec HDFS Read: 518895 HDFS Write: 28 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 790 msec

OK

year \_c1

**2000 1**

**2004 1**

**2008 3**

**2012 6**

Time taken: 77.374 seconds, Fetched: 4 row(s)

hive>

----------------------------------------------------------------------------------

**--problem 3**

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

**hive> SELECT country,SUM(total\_medal) FROM olympic group by country;**

Query ID = acadgild\_20171117224747\_eea4fbbf-6db9-4936-b413-28c5ea69d96c

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\_1505655198300\_0018, Tracking URL = http://localhost:8088/proxy/application\_1505655198300\_0018/

Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job\_1505655198300\_0018

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

2017-11-17 22:47:35,138 Stage-1 map = 0%, reduce = 0%

2017-11-17 22:47:52,733 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.96 sec

2017-11-17 22:48:15,419 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.71 sec

MapReduce Total cumulative CPU time: 6 seconds 710 msec

Ended Job = job\_1505655198300\_0018

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.71 sec HDFS Read: 518895 HDFS Write: 1315 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 710 msec

OK

**country \_c1**

**Afghanistan 2**

**Algeria 8**

**Argentina 141**

**Armenia 10**

**Australia 609**

**Austria 91**

**Azerbaijan 25**

**Bahamas 24**

**Bahrain 1**

**Barbados 1**

**Belarus 97**

**Belgium 18**

**Botswana 1**

**Brazil 221**

**Bulgaria 41**

**Cameroon 20**

**Canada 370**

**Chile 22**

**China 530**

**Chinese Taipei 20**

**Colombia 13**

**Costa Rica 2**

**Croatia 81**

**Cuba 188**

**Cyprus 1**

**Czech Republic 81**

**Denmark 89**

**Dominican Republic 5**

**Ecuador 1**

**Egypt 8**

**Eritrea 1**

**Estonia 18**

**Ethiopia 29**

**Finland 118**

**France 318**

**Gabon 1**

**Georgia 23**

**Germany 629**

**Great Britain 322**

**Greece 59**

**Grenada 1**

**Guatemala 1**

**Hong Kong 3**

**Hungary 145**

**Iceland 15**

**India 11**

**Indonesia 22**

**Iran 24**

**Ireland 9**

**Israel 4**

**Italy 331**

**Jamaica 80**

**Japan 282**

**Kazakhstan 42**

**Kenya 39**

**Kuwait 2**

**Kyrgyzstan 3**

**Latvia 17**

**Lithuania 30**

**Macedonia 1**

**Malaysia 3**

**Mauritius 1**

**Mexico 38**

**Moldova 5**

**Mongolia 10**

**Montenegro 14**

**Morocco 11**

**Mozambique 1**

**Netherlands 318**

**New Zealand 52**

**Nigeria 39**

**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**

**Zimbabwe 7**

Time taken: 71.84 seconds, Fetched: 110 row(s)

hive>

---------------------------------------------------------------------------------

**--problem 4**

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

hive> SELECT country,SUM(gold\_medal) FROM olympic group by country;

Query ID = acadgild\_20171117225252\_d8a6504d-52bf-4f66-af78-2c0901a46455

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\_1505655198300\_0019, Tracking URL = http://localhost:8088/proxy/application\_1505655198300\_0019/

Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job\_1505655198300\_0019

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

2017-11-17 22:53:22,053 Stage-1 map = 0%, reduce = 0%

2017-11-17 22:53:39,633 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.26 sec

2017-11-17 22:53:58,356 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.6 sec

MapReduce Total cumulative CPU time: 6 seconds 600 msec

Ended Job = job\_1505655198300\_0019

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.6 sec HDFS Read: 518895 HDFS Write: 1276 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 600 msec

OK

country \_c1

**Afghanistan 0**

**Algeria 2**

**Argentina 49**

**Armenia 0**

**Australia 163**

**Austria 36**

**Azerbaijan 6**

**Bahamas 11**

**Bahrain 0**

**Barbados 0**

**Belarus 17**

**Belgium 2**

**Botswana 0**

**Brazil 46**

**Bulgaria 8**

**Cameroon 20**

**Canada 168**

**Chile 3**

**China 234**

**Chinese Taipei 2**

**Colombia 2**

**Costa Rica 0**

**Croatia 35**

**Cuba 57**

**Cyprus 0**

**Czech Republic 14**

**Denmark 46**

**Dominican Republic 3**

**Ecuador 0**

**Egypt 1**

**Eritrea 0**

**Estonia 6**

**Ethiopia 13**

**Finland 11**

**France 108**

**Gabon 0**

**Georgia 6**

**Germany 223**

**Great Britain 124**

**Greece 12**

**Grenada 1**

**Guatemala 0**

**Hong Kong 0**

**Hungary 77**

**Iceland 0**

**India 1**

**Indonesia 5**

**Iran 10**

**Ireland 1**

**Israel 1**

**Italy 86**

**Jamaica 24**

**Japan 57**

**Kazakhstan 13**

**Kenya 11**

**Kuwait 0**

**Kyrgyzstan 0**

**Latvia 3**

**Lithuania 5**

**Macedonia 0**

**Malaysia 0**

**Mauritius 0**

**Mexico 19**

**Moldova 0**

**Mongolia 2**

**Montenegro 0**

**Morocco 2**

**Mozambique 1**

**Netherlands 101**

**New Zealand 18**

**Nigeria 6**

**North Korea 6**

**Norway 97**

**Panama 1**

**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: 76.665 seconds, Fetched: 110 row(s)

hive>