ACD\_BDD2.3\_Session\_7\_Assignment\_2

Hive assignment 7.2

**hive> CREATE TABLE olympic( athelete string,age int,country string,year int,closing\_date date,**

**> sports string,gold int,silver int,bronze int,tot\_medal int)**

**> row format delimited fields**

**> terminated by '\t'**

**> ;**

OK

Time taken: 1.122 seconds

**hive> load data local inpath '/home/acadgild/hive\_folder/olympix\_data.csv' into table olympic;**

Loading data to table default.olympic

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

OK

Time taken: 3.77 seconds

hive> select athelete from olympic limit 2;

OK

Michael Phelps

Michael Phelps

Time taken: 0.989 seconds, Fetched: 2 row(s)

**hive> select \* from olympic limit 2;**

**OK**

**Michael Phelps 23 United States 2008 NULL Swimming 8 0 0 8**

**Michael Phelps 19 United States 2004 NULL Swimming 6 0** 2 8

Time taken: 1.169 seconds, Fetched: 2 row(s)

Problem Statement

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

**hive> select country ,sum(tot\_medal) from olympic where sports = 'Swimming' group by country;**

Starting Job = job\_1505564459575\_0006, Tracking URL = http://localhost:8088/proxy/application\_1505564459575\_0006/

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 129.25 sec HDFS Read: 518889 HDFS Write: 386 SUCCESS

Total MapReduce CPU Time Spent: 2 minutes 9 seconds 250 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 9**

**Trinidad and Tobago 1**

**Tunisia 3**

**Ukraine 7**

**United States 267**

**Zimbabwe 7**

**Time taken: 853.077 seconds, Fetched: 34 row(s)**

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

**hive> select year, sum(tot\_medal) from olympic where country = 'India' group by year;**

Starting Job = job\_1505564459575\_0007, Tracking URL = http://localhost:8088/proxy/application\_1505564459575\_0007/

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

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 127.04 sec HDFS Read: 518889 HDFS Write: 28 SUCCESS

Total MapReduce CPU Time Spent: 2 minutes 7 seconds 40 msec

OK

**2000 1**

**2004 1**

**2008 3**

**2012 6**

**Time taken: 886.466 seconds, Fetched: 4 row(s)**

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

**hive> select country,sum(tot\_medal) from olympic group by country;**

Starting Job = job\_1505564459575\_0008, Tracking URL = http://localhost:8088/proxy/application\_1505564459575\_0008/

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 104.26 sec HDFS Read: 518889 HDFS Write: 1315 SUCCESS

Total MapReduce CPU Time Spent: 1 minutes 44 seconds 260 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**

**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**

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

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

**hive> select country,count(gold) from olympic group by country;**

Starting Job = job\_1505564459575\_0009, Tracking URL = http://localhost:8088/proxy/application\_1505564459575\_0009/

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 94.51 sec HDFS Read: 518889 HDFS Write: 1314 SUCCESS

**Total MapReduce CPU Time Spent: 1 minutes 34 seconds 510 msec**

**OK**

**Afghanistan 2**

**Algeria 8**

**Argentina 141**

**Armenia 10**

**Australia 524**

**Austria 70**

**Azerbaijan 25**

**Bahamas 23**

**Bahrain 1**

**Barbados 1**

**Belarus 92**

**Belgium 18**

**Botswana 1**

**Brazil 220**

**Bulgaria 36**

**Cameroon 20**

**Canada 351**

**Chile 20**

**China 450**

**Chinese Taipei 20**

**Colombia 13**

**Costa Rica 1**

**Croatia 76**

**Cuba 188**

**Cyprus 1**

**Czech Republic 75**

**Denmark 89**

**Dominican Republic 5**

**Ecuador 1**

**Egypt 8**

**Eritrea 1**

**Estonia 16**

**Ethiopia 24**

**Finland 112**

**France 287**

**Gabon 1**

**Georgia 23**

**Germany 552**

**Great Britain 296**

**Greece 59**

**Grenada 1**

**Guatemala 1**

**Hong Kong 3**

**Hungary 132**

**Iceland 15**

**India 11**

**Indonesia 22**

**Iran 24**

**Ireland 9**

**Israel 4**

**Italy 307**

**Jamaica 61**

**Japan 259**

**Kazakhstan 42**

**Kenya 38**

**Kuwait 2**

**Kyrgyzstan 3**

**Latvia 17**

**Lithuania 30**

**Macedonia 1**

**Malaysia 3**

**Mauritius 1**

**Mexico 38**

**Moldova 5**

**Mongolia 10**

**Montenegro 14**

**Morocco 10**

**Mozambique 1**

**Netherlands 286**

**New Zealand 51**

**Nigeria 39**

**North Korea 21**

**Norway 158**

**Panama 1**

**Paraguay 17**

**Poland 73**

**Portugal 9**

**Puerto Rico 2**

**Qatar 3**

**Romania 97**

**Russia 706**

**Saudi Arabia 6**

**Serbia 31**

**Serbia and Montenegro 38**

**Singapore 6**

**Slovakia 33**

**Slovenia 24**

**South Africa 22**

**South Korea 274**

**Spain 195**

**Sri Lanka 1**

**Sudan 1**

**Sweden 167**

**Switzerland 87**

**Syria 1**

**Tajikistan 3**

**Thailand 18**

**Togo 1**

**Trinidad and Tobago 16**

**Tunisia 3**

**Turkey 27**

**Uganda 1**

**Ukraine 137**

**United Arab Emirates 1**

**United States 1109**

**Uruguay 1**

**Uzbekistan 19**

**Venezuela 4**

**Vietnam 2**

**Zimbabwe 2**

**Time taken: 809.55 seconds, Fetched: 110 row(s)**