

Assignment 7.2

Loaded the **Olympics_data.csv** into **/user/acadgild**. Created the table **Oylympics_data** and loaded the **Olympics_data.csv** file in the table.

```
hive> create table olympics_data
> (
>   Athlete STRING,
>   Age INT,
>   Country STRING,
>   Year INT,
>   Closing_Date STRING,
>   Sport STRING,
>   Gold_Medals INT,
>   Silver_Medals INT,
>   Bronze_Medals INT,
>   Total_Medals INT
> )
> ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
OK
Time taken: 0.425 seconds
hive>
```

1. Numbers of medals won by each country for Swimming

Command:

```
hive> SELECT Country, COUNT(Total_Medals) FROM olympics_data WHERE Sport = "Swimming" GROUP BY Country;
Query ID = acadgild_20171108211616_a7104bcb-850a-4a82-8d1d-24182c0367c0
Total jobs = 1
```

Output:

```

Total MapReduce CPU Time Spent: 70.244 seconds
OK
Argentina      1
Australia      92
Austria        2
Belarus        1
Brazil         7
Canada         5
China          29
Costa Rica     1
Croatia        1
Denmark        1
France         26
Germany        27
Great Britain   9
Hungary        7
Italy          13
Japan          30
Lithuania      1
Netherlands    32
Norway         2
Poland         1
Romania        4
Russia         19
Serbia         1
Slovakia       1
Slovenia       1
South Africa   8
South Korea    2
Spain          2
Sweden         7
Trinidad and Tobago 1
Tunisia        2
Ukraine        4
United States  145
Zimbabwe       2
Time taken: 70.244 seconds
hive>

```

2. No. of medals won by India year wise.

Command :

```

SELECT Year, COUNT(Total_Medals) FROM olympics_data WHERE Country="India" GROUP BY
Year;

```

```

hive> SELECT Year, COUNT(Total Medals) FROM olympics_data WHERE Country="India" GROUP BY Year;
Query ID = acadgild_20171108212323_c8bbcdd0-7e63-4aa0-8466-6bf2314d9d6f
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_1510148864385_0007, Tracking URL = http://chemlabtest.aqs.com:8088/proxy/application_1510148864385_0007/
Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job_1510148864385_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2017-11-08 21:23:51,233 Stage-1 map = 0%, reduce = 0%
2017-11-08 21:24:05,735 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.36 sec
2017-11-08 21:24:23,237 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.82 sec
MapReduce Total cumulative CPU time: 5 seconds 820 msec
Ended Job = job_1510148864385_0007
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.82 sec HDFS Read: 518905 HDFS Write: 28 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 820 msec
OK
2000      1
2004      1
2008      3
2012      6
Time taken: 54.652 seconds, Fetched: 4 row(s)
hive>

```

3. Number of medals won by each country

Command:

```

Time taken: 54.652 seconds, Fetched: 4 row(s)
hive> SELECT Country, COUNT(Total_Medals) FROM olympics_data GROUP BY Country;

```

Output:

Total MapReduce CPU Time S	Country	Gold Medals
OK	Germany	552
	Great Britain	296
	Greece	59
Afghanistan	Grenada	1
Algeria	Guatemala	1
Argentina	Hong Kong	3
Armenia	Hungary	132
Australia	Iceland	15
Austria	India	11
Azerbaijan	Indonesia	22
Bahamas	Iran	24
Bahrain	Ireland	9
Barbados	Israel	4
Belarus	Italy	307
Belgium	Jamaica	61
Botswana	Japan	259
Brazil	Kazakhstan	42
Bulgaria	Kenya	38
Cameroon	Kuwait	2
Canada	Kyrgyzstan	3
Chile	Latvia	17
China	Lithuania	30
Chinese Taipei	Macedonia	1
Colombia	Malaysia	3
Costa Rica	Mauritius	1
Croatia	Mexico	38
Cuba	Moldova	5
Cyprus	Mongolia	10
Czech Republic	Montenegro	14
Denmark	Morocco	10
Dominican Republic	Mozambique	1
Ecuador	Netherlands	286
Egypt	New Zealand	51
Eritrea	Nigeria	39
Estonia	North Korea	21
Ethiopia	Norway	158
Finland	Panama	1
France	Paraguay	17
Gabon	Poland	73
Georgia	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: 50.582 seconds,	
	hive> █	

4. Number of Gold medals won by each country

Command:

```
hive> SELECT Country, COUNT(Gold_Medals) FROM olympics_data GROUP BY Country;
```

Output:

Total MapReduce CPU Time Sp		Germany 552		Qatar 3	
OK		Great Britain 296		Romania 97	
Afghanistan 2		Greece 59		Russia 706	
Algeria 8		Grenada 1		Saudi Arabia 6	
Argentina 141		Guatemala 1		Serbia 31	
Armenia 10		Hong Kong 3		Serbia and Montenegro 38	
Australia 524		Hungary 132		Singapore 6	
Austria 70		Iceland 15		Slovakia 33	
Azerbaijan 25		India 11		Slovenia 24	
Bahamas 23		Indonesia 22		South Africa 22	
Bahrain 1		Iran 24		South Korea 274	
Barbados 1		Ireland 9		Spain 195	
Belarus 92		Israel 4		Sri Lanka 1	
Belgium 18		Italy 307		Sudan 1	
Botswana 1		Jamaica 61		Sweden 167	
Brazil 220		Japan 259		Switzerland 87	
Bulgaria 36		Kazakhstan 42		Syria 1	
Cameroon 20		Kenya 38		Tajikistan 3	
Canada 351		Kuwait 2		Thailand 18	
Chile 20		Kyrgyzstan 3		Togo 1	
China 450		Latvia 17		Trinidad and Tobago 16	
Chinese Taipei 20		Lithuania 30		Tunisia 3	
Colombia 13		Macedonia 1		Turkey 27	
Costa Rica 1		Malaysia 3		Uganda 1	
Croatia 76		Mauritius 1		Ukraine 137	
Cuba 188		Mexico 38		United Arab Emirates 1	
Cyprus 1		Moldova 5		United States 1109	
Czech Republic 75		Mongolia 10		Uruguay 1	
Denmark 89		Montenegro 14		Uzbekistan 19	
Dominican Republic 5		Morocco 10		Venezuela 4	
Ecuador 1		Mozambique 1		Vietnam 2	
Egypt 8		Netherlands 286		Zimbabwe 2	
Eritrea 1		New Zealand 51		Time taken: 48.457 seconds,	
Estonia 16		Nigeria 39		hive>	
Ethiopia 24		North Korea 21			
Finland 112		Norway 158			
France 287		Panama 1			
Gabon 1		Paraguay 17			
Georgia 23		Poland 73			
		Portugal 9			
		Puerto Rico 2			