**//Checking databases**

hive> show databases;

OK

custom

default

demo

Time taken: 0.124 seconds, Fetched: 3 row(s)

**// using custom schema**

hive> use custom;

OK

Time taken: 0.034 seconds

**// creating table olympix\_data**

hive> create table olympix\_data

> (

> athlete string, age int, country string, year string, closingDate string, sport string, goldMedals int,

> silverMedals int, bronzeMedals int, totalMedals int)

> row format delimited

> fields terminated by '\t';

OK

Time taken: 0.086 seconds

**// Loading data into olympix\_data table**

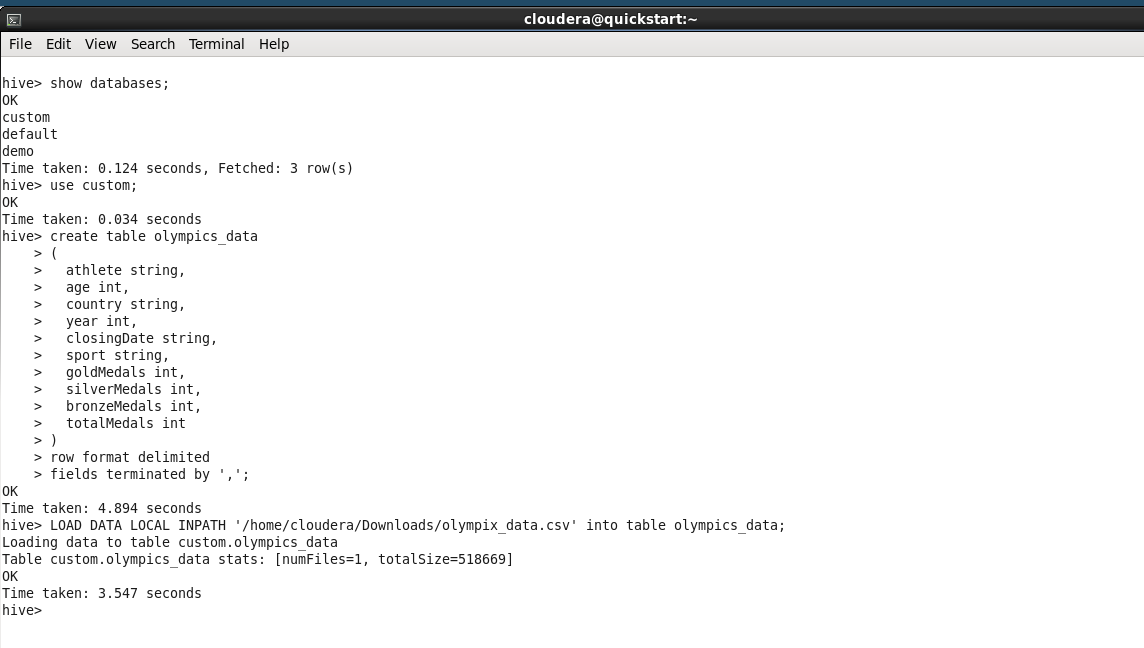
hive> LOAD DATA LOCAL INPATH '/home/cloudera/Downloads/olympix\_data.csv' into table olympix\_data;

Loading data to table custom.olympics\_data

Table custom.olympics\_data stats: [numFiles=1, totalSize=883926]

OK

Time taken: 0.419 seconds



**1. Fetching and displaying total medals won in swimming by each country**

hive> select country, sum(totalMedals) from olympix\_data where sport = 'Swimming' group by country;

Query ID = cloudera\_20171222213030\_75b645e5-4860-4292-bd67-786a040ec89a

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\_1513997930705\_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application\_1513997930705\_0002/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job\_1513997930705\_0002

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

2017-12-22 21:31:07,770 Stage-1 map = 0%, reduce = 0%

2017-12-22 21:31:41,574 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.0 sec

2017-12-22 21:31:52,680 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.89 sec

MapReduce Total cumulative CPU time: 4 seconds 890 msec

Ended Job = job\_1513997930705\_0002

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.89 sec HDFS Read: 518958 HDFS Write: 386 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 890 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

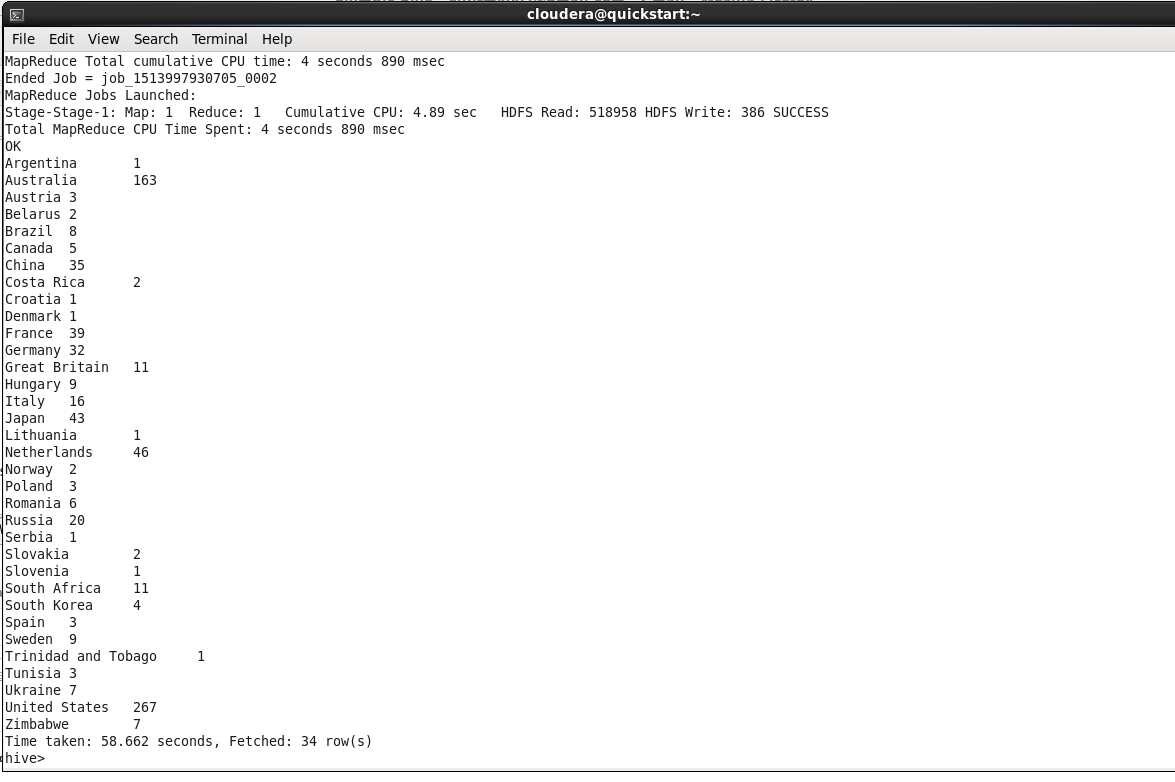
Tunisia 3

Ukraine 7

United States 267

Zimbabwe 7

Time taken: 58.662 seconds, Fetched: 34 row(s)



**2. Fetching and displaying total medalswon by India every year**

hive> select country, year, sum(totalMedals) from olympix\_data where country = 'India' group by year,country;

Query ID = cloudera\_20171222213939\_b8909223-9f01-4956-b6b3-a55b08ac7ad3

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\_1513997930705\_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application\_1513997930705\_0003/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job\_1513997930705\_0003

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

2017-12-22 21:39:22,289 Stage-1 map = 0%, reduce = 0%

2017-12-22 21:40:52,006 Stage-1 map = 0%, reduce = 0%

2017-12-22 21:42:18,516 Stage-1 map = 0%, reduce = 0%

2017-12-22 21:42:25,237 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.65 sec

2017-12-22 21:42:43,216 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.81 sec

MapReduce Total cumulative CPU time: 6 seconds 810 msec

Ended Job = job\_1513997930705\_0003

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.81 sec HDFS Read: 519465 HDFS Write: 52 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 810 msec

OK

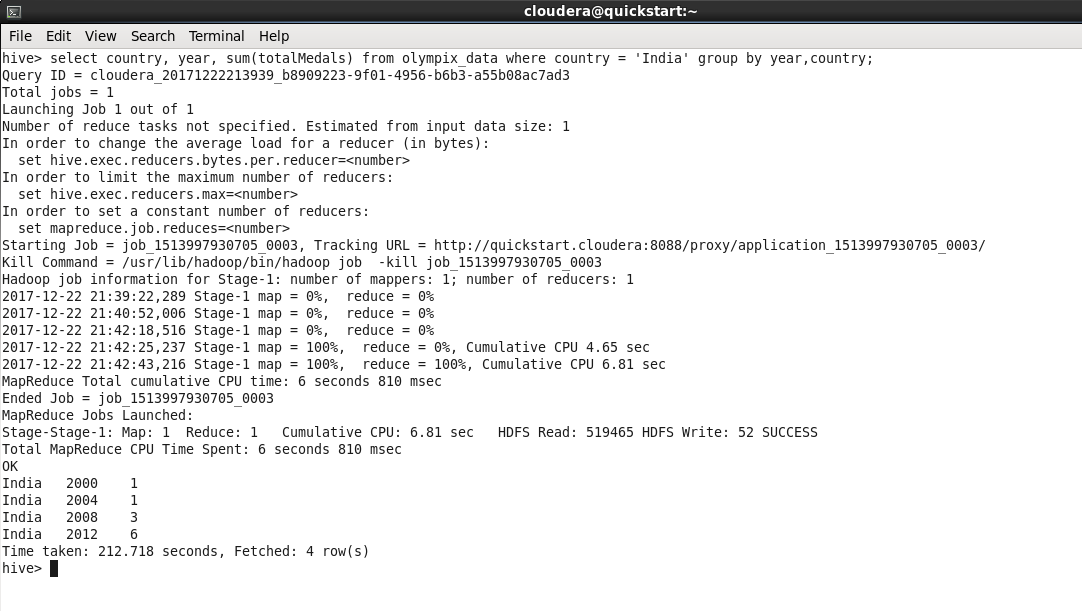
India 2000 1

India 2004 1

India 2008 3

India 2012 6

Time taken: 212.718 seconds, Fetched: 4 row(s)



3. **Fetching and displaying total medals won by each country**

hive> select country, sum(totalMedals) from olympix\_data group by country;

Query ID = cloudera\_20171222214444\_c0baddf3-2fde-4d91-81e2-c333d37bcf43

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\_1513997930705\_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application\_1513997930705\_0004/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job\_1513997930705\_0004

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

2017-12-22 21:45:04,692 Stage-1 map = 0%, reduce = 0%

2017-12-22 21:45:17,198 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.31 sec

2017-12-22 21:45:28,346 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.1 sec

MapReduce Total cumulative CPU time: 4 seconds 100 msec

Ended Job = job\_1513997930705\_0004

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.1 sec HDFS Read: 518165 HDFS Write: 1315 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 100 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

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

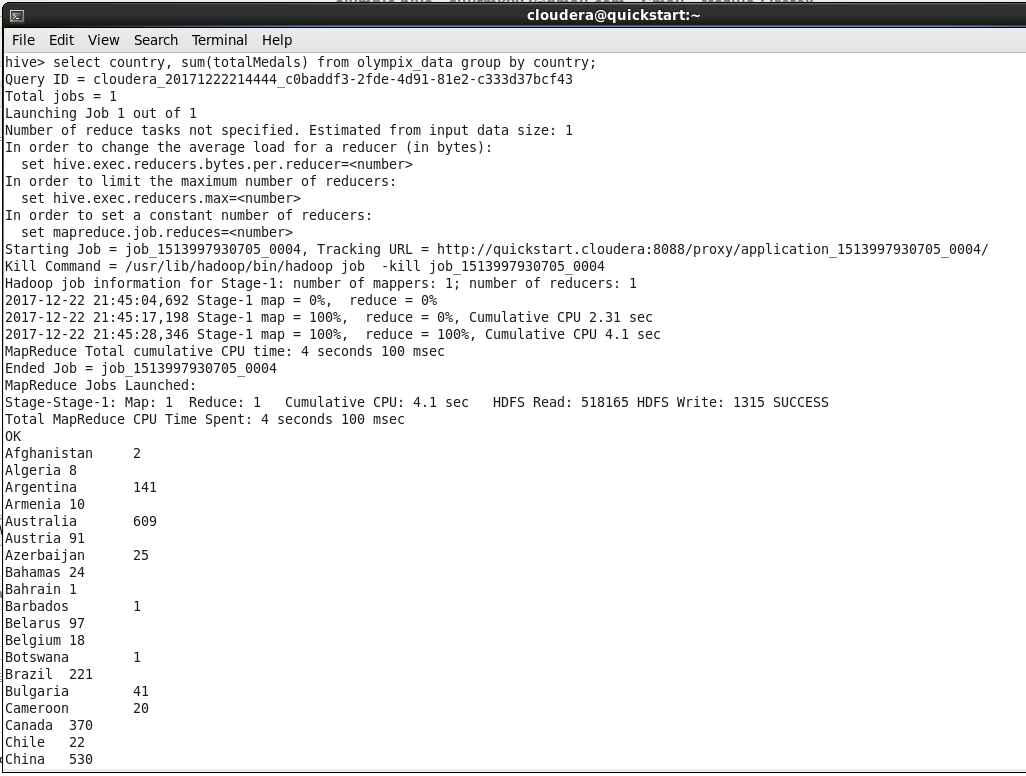
Uzbekistan 19

Venezuela 4

Vietnam 2

Zimbabwe 7

Time taken: 36.633 seconds, Fetched: 110 row(s)



4. **Fetching and displaying total gold medals won by each country**

hive> select country, sum(goldMedals) from olympix\_data group by country;

Query ID = cloudera\_20171222214646\_a7bf323d-6c30-4a0f-8dbb-95fcb408e1aa

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\_1513997930705\_0005, Tracking URL = http://quickstart.cloudera:8088/proxy/application\_1513997930705\_0005/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job\_1513997930705\_0005

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

2017-12-22 21:46:53,215 Stage-1 map = 0%, reduce = 0%

2017-12-22 21:47:15,650 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.45 sec

2017-12-22 21:47:25,570 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.22 sec

MapReduce Total cumulative CPU time: 4 seconds 220 msec

Ended Job = job\_1513997930705\_0005

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.22 sec HDFS Read: 518164 HDFS Write: 1276 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 220 msec

OK

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: 45.904 seconds, Fetched: 110 row(s)

hive>

