

## Assignment 3.2

### Associated Data Files

This Data set is about Olympics. You can download the data set from the below link:

<https://drive.google.com/open?id=0ByJLBTmJojjzV1czX3Nha0R3bTQ>

```
CREATE DATABASE olympics;
```

```
USE OLYMPICS;
```

```
hive> CREATE DATABASE olympics;
OK
Time taken: 0.191 seconds
hive> USE OLYMPICS;
OK
Time taken: 0.026 seconds
hive>
```

```
CREATE TABLE IF NOT EXISTS olympic_data (athlete String, age int, country String, the_year int,
closing_date string, sport String, gold int, silver int, bronze int, total_medals int) ROW FORMAT
DELIMITED FIELDS TERMINATED BY '\t';
```

```
LOAD DATA LOCAL INPATH '/home/acadgild/olympix_data.csv' INTO TABLE olympic_data;
```

```
hive> CREATE TABLE IF NOT EXISTS olympic_data (athlete String, age int, country String, the_year int, closing_date st
ring, sport String, gold int, silver int, bronze int, total_medals int) ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t
';
OK
Time taken: 0.141 seconds
```

```
hive> LOAD DATA LOCAL INPATH '/home/acadgild/olympix_data.csv' INTO TABLE olympic_data;
Loading data to table olympics.olympic_data
OK
Time taken: 0.741 seconds
```

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

```
SELECT country,SUM(total_medals) FROM olympic_data WHERE sport = 'Swimming' GROUP BY
country;
```

```
hive> SELECT country,SUM(total_medals) FROM olympic_data WHERE sport = 'Swimming' GROUP BY country;
```

```

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: 29.825 seconds, Fetched: 34 row(s)

```

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

SELECT the\_year, SUM(total\_medals) FROM olympic\_data WHERE country = 'India' GROUP BY the\_year;

```
hive> SELECT the_year, SUM(total_medals) FROM olympic_data WHERE country = 'India' GROUP BY the_year;
```

```

OK
2000      1
2004      1
2008      3
2012      6
Time taken: 29.513 seconds, Fetched: 4 row(s)

```

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

```
SELECT country, SUM(total_medals) FROM olympic_data GROUP BY country;
```

```
hive> SELECT country, SUM(total_medals) FROM olympic_data GROUP BY country;
```

```
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
```

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

```
SELECT country, SUM(gold) FROM olympic_data GROUP BY country;
```

```
hive> SELECT country, SUM(gold) FROM olympic_data GROUP BY country;
```

```
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
```