Session 15: Hive OperationsAssignment 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

**DATE SET DESCRIPTION**

The data set consists of the following fields.

Athlete: This field consists of the athlete name

Age: This field consists of athlete ages

Country: This fields consists of the country names which participated in Olympics

Year: This field consists of the year

Closing Date: This field consists of the closing date of ceremony

Sport: Consists of the sports name

Gold Medals: No. of Gold medals

Silver Medals: No. of Silver medals

Bronze Medals: No. of Bronze medals

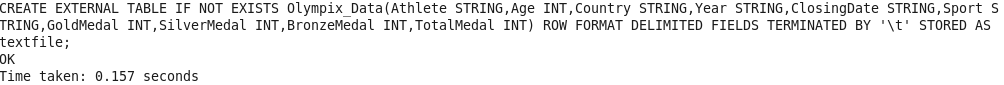
Total Medals: Consists of total no. of medals

# Problem Statement

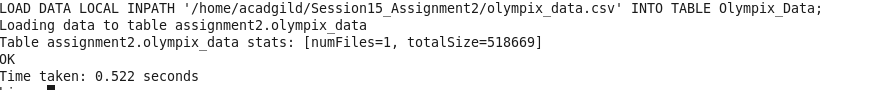
# Write a Hive program to find the number of medals won by each country in swimming.

Ans:

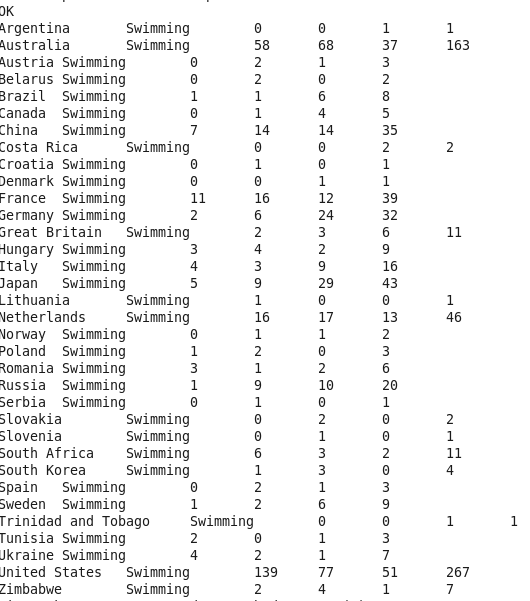
CREATE EXTERNAL TABLE IF NOT EXISTS Olympix\_Data(Athlete STRING,Age INT,Country STRING,Year STRING,ClosingDate STRING,Sport STRING,GoldMedal INT,SilverMedal INT,BronzeMedal INT,TotalMedal INT) ROW FORMAT DELIMITED FIELDS TERMINATED BY ‘\t’ STORED AS textfile;



LOAD DATA LOCAL INPATH ‘home/acadgild/Session15\_Assignmnet2/olymix\_data.csv’ INTO TABLE Olympix\_Data;



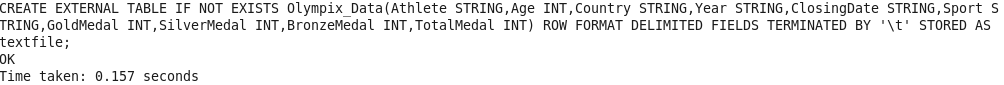
SELECT Country,Sport,SUM(GoldMedal),SUM(SilverMedal),SUM(BronzeMedal),SUM(TotalMedal) FROM Olympix\_Data WHERE Sport=’Swimming’ GROUP BY Country,Sport;



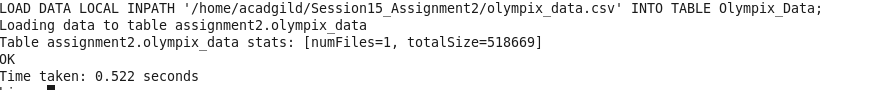
# Write a Hive program to find the number of medals that India won year wise.

**Ans:**

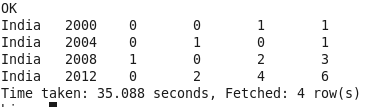
CREATE EXTERNAL TABLE IF NOT EXISTS Olympix\_Data(Athlete STRING,Age INT,Country STRING,Year STRING,ClosingDate STRING,Sport STRING,GoldMedal INT,SilverMedal INT,BronzeMedal INT,TotalMedal INT) ROW FORMAT DELIMITED FIELDS TERMINATED BY ‘\t’ STORED AS textfile;



LOAD DATA LOCAL INPATH ‘home/acadgild/Session15\_Assignmnet2/olymix\_data.csv’ INTO TABLE Olympix\_Data;



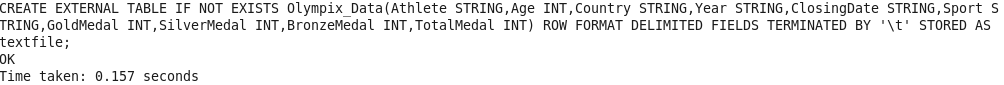
SELECT Country,Year,SUM(GoldMedal),SUM(SilverMedal),SUM(BronzeMedal),SUM(TotalMedal) FROM Olympix\_Data WHERE Country=’India’ GROUP BY Country,Year;



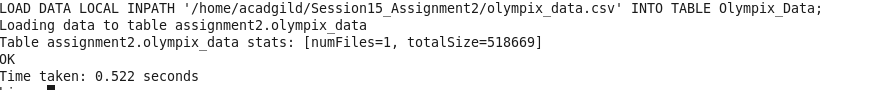
# Write a Hive Program to find the total number of medals each country won.

**Ans:**

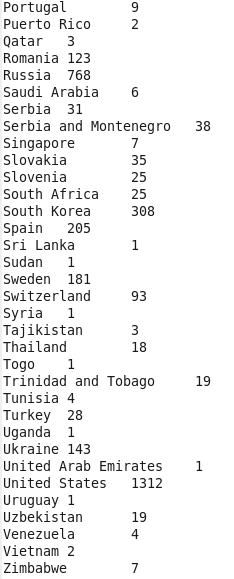
CREATE EXTERNAL TABLE IF NOT EXISTS Olympix\_Data(Athlete STRING,Age INT,Country STRING,Year STRING,ClosingDate STRING,Sport STRING,GoldMedal INT,SilverMedal INT,BronzeMedal INT,TotalMedal INT) ROW FORMAT DELIMITED FIELDS TERMINATED BY ‘\t’ STORED AS textfile;



LOAD DATA LOCAL INPATH ‘home/acadgild/Session15\_Assignmnet2/olymix\_data.csv’ INTO TABLE Olympix\_Data;



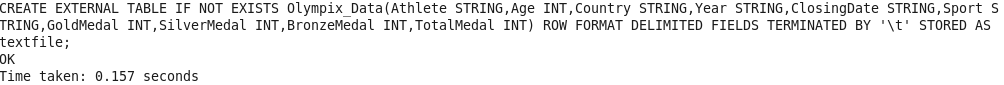
SELECT Country,SUM(TotalMedal) FROM Olympix\_Data ’ GROUP BY Country;



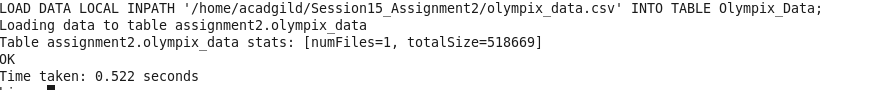
# Write a Hive program to find the number of gold medals each country won.

**Ans:**

CREATE EXTERNAL TABLE IF NOT EXISTS Olympix\_Data(Athlete STRING,Age INT,Country STRING,Year STRING,ClosingDate STRING,Sport STRING,GoldMedal INT,SilverMedal INT,BronzeMedal INT,TotalMedal INT) ROW FORMAT DELIMITED FIELDS TERMINATED BY ‘\t’ STORED AS textfile;



LOAD DATA LOCAL INPATH ‘home/acadgild/Session15\_Assignmnet2/olymix\_data.csv’ INTO TABLE Olympix\_Data;



SELECT Country,SUM(GoldMedal) FROM Olympix\_Data ’ GROUP BY Country;

