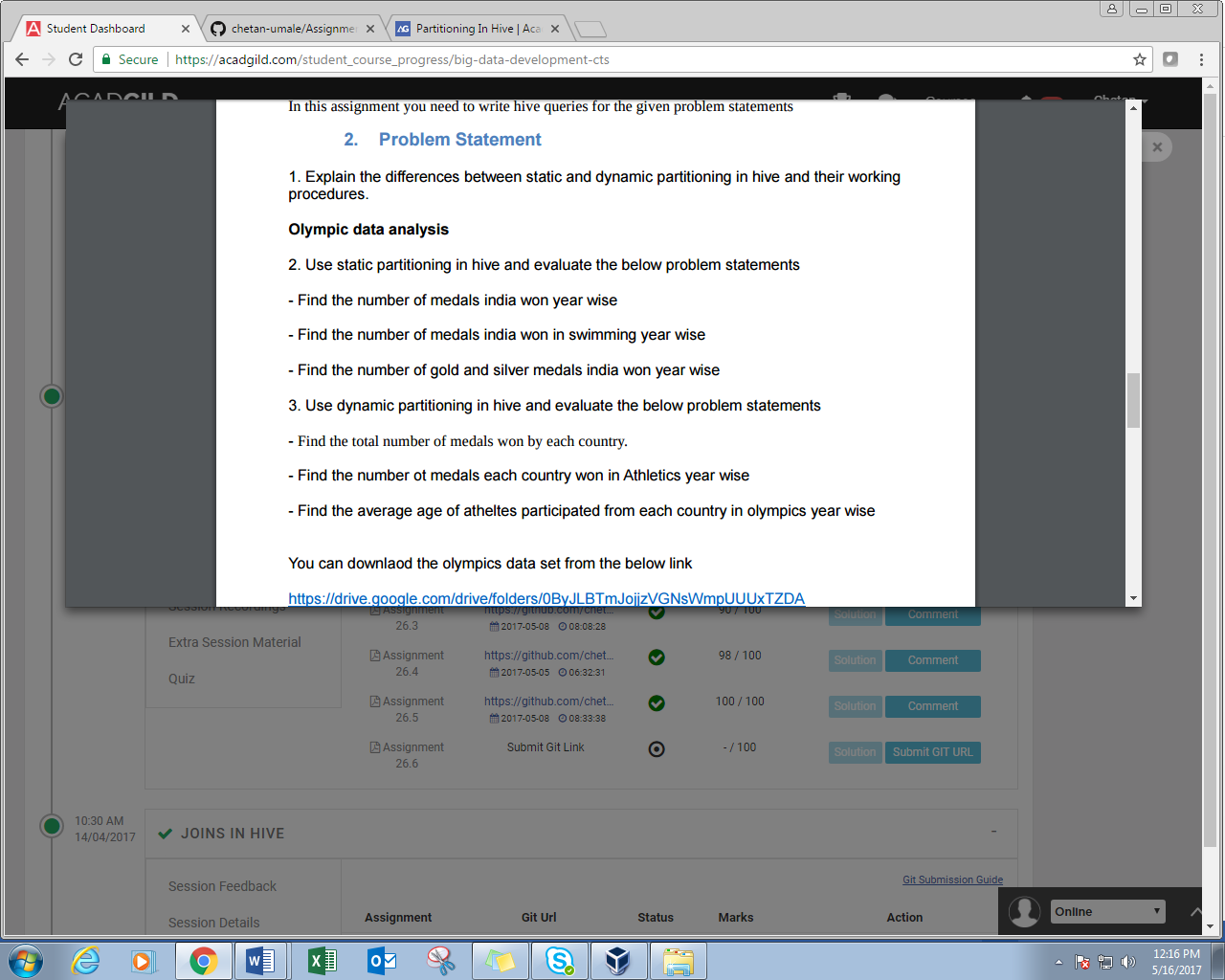
**PROBLEM STATEMENT:**



**Q1-**

**SOLUTION:**

1. Hive organizes tables into partitions. It is a way of dividing a table into related parts based on the values of partitioned columns such as date, city, and department.

2. Using partition, it is easy to query a portion of the data.

3. Tables or partitions are sub-divided into **buckets,** to provide extra structure to the data that may be used for more efficient querying.

4. Bucketing works based on the value of hash function of some column of a table.

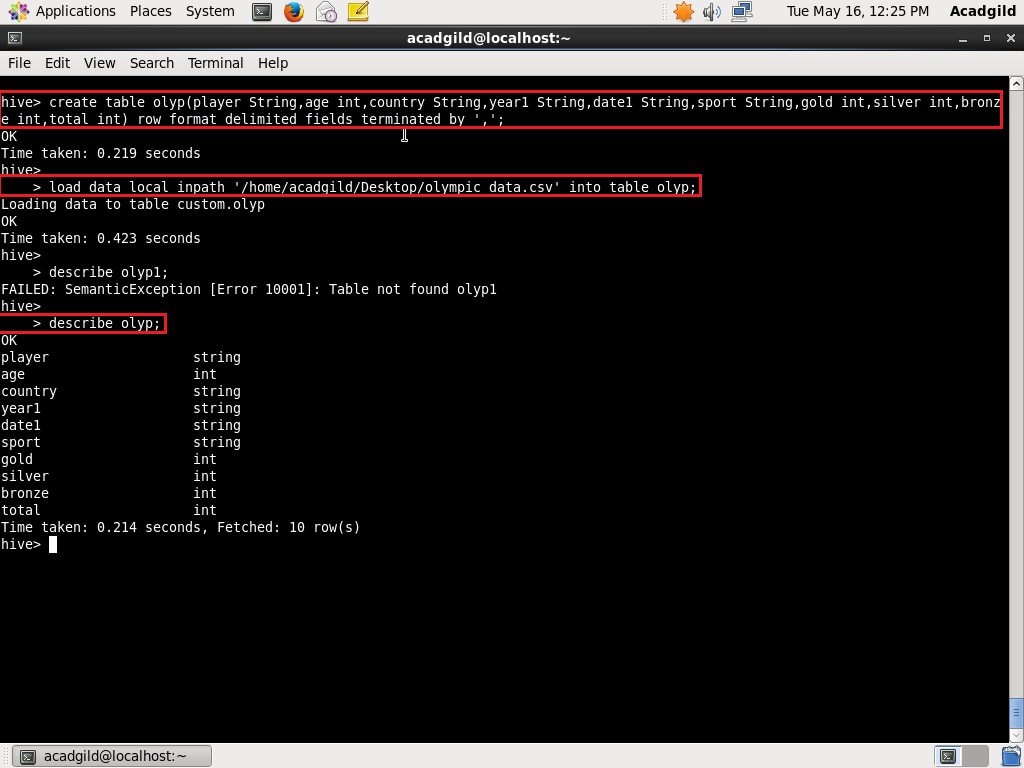
5. For example, a table named **Tab1** contains employee data such as id, name, dept, and yoj (i.e., year of joining). Suppose you need to retrieve the details of all employees who joined in 2012. A query searches the whole table for the required information. However, if you partition the employee data with the year and store it in a separate file, it reduces the query processing time.

|  |  |
| --- | --- |
| **STATIC PARTITIONING** | **DYNAMIC PARTITIONING** |
| Insert input data files individually into a partition table is Static Partition. | Single insert to partition table is known as dynamic partition. |
| Static Partition saves your time in loading data compared to dynamic partition | Dynamic Partition takes more time in loading data compared to static partition. |
| We can alter the partition in static partition. | We can’t perform alter on Dynamic partition. |
| Static partition is in Strict Mode. | If you want to use Dynamic partition in hive then mode is in non strict mode. |

**Q2-**

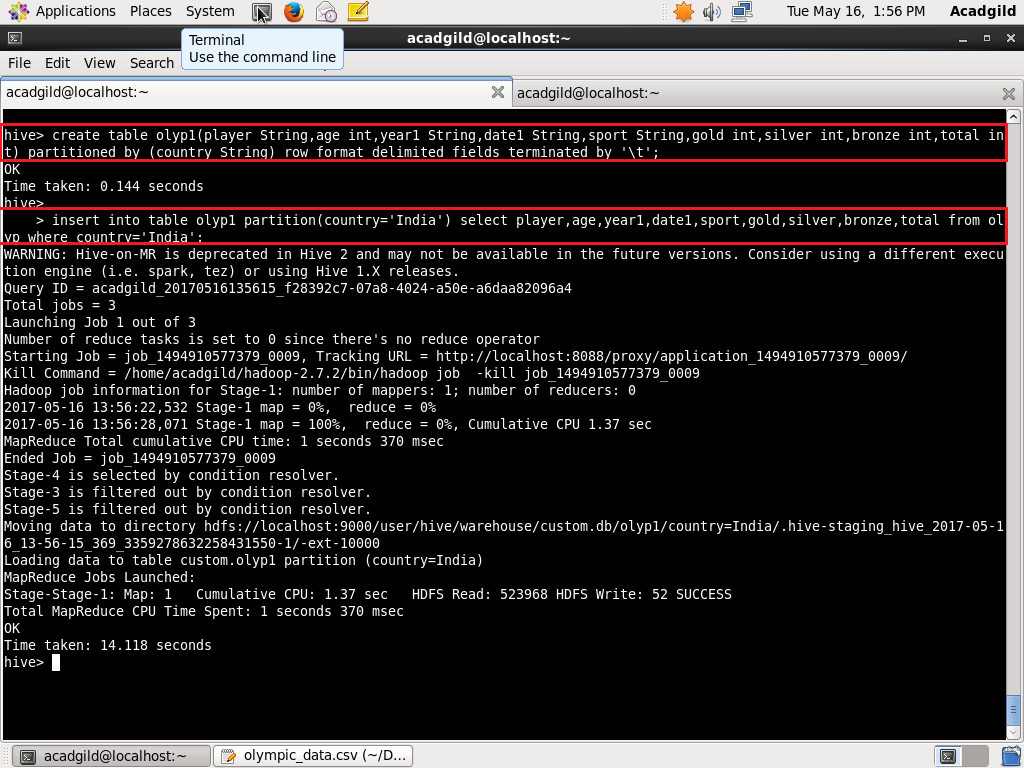
**SOLUTION:**

//initially a table olyp is created which contains the entire dataset.This table is used as a source to insert data into other tables

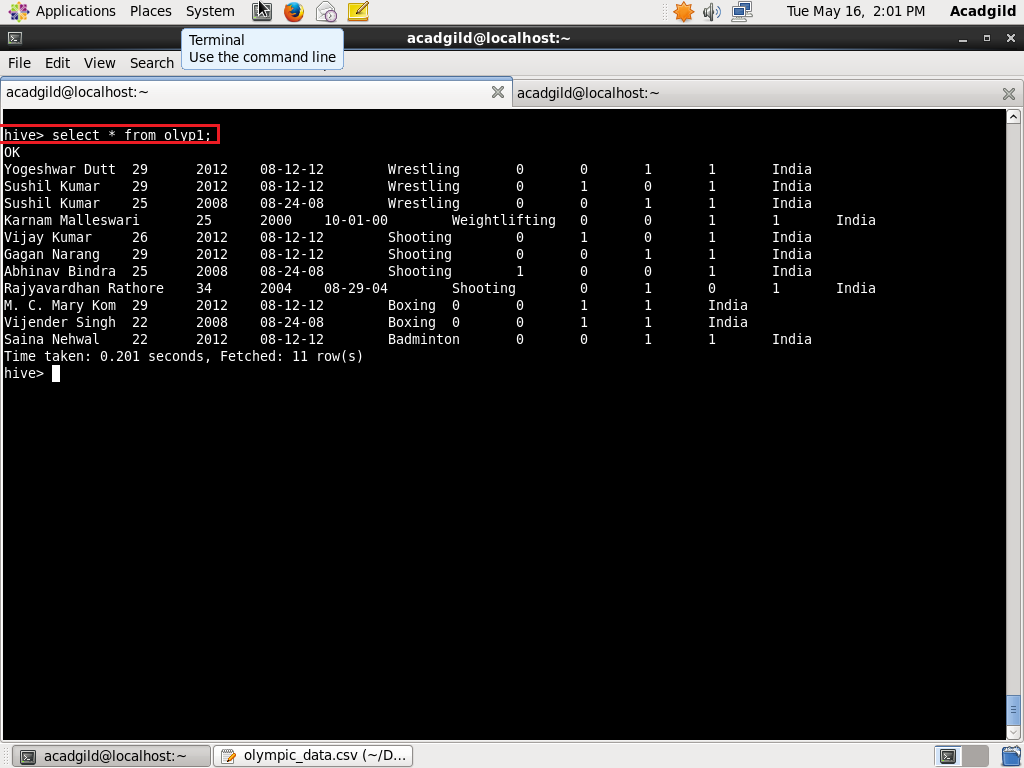


1. Find the number of medals india won year wise.

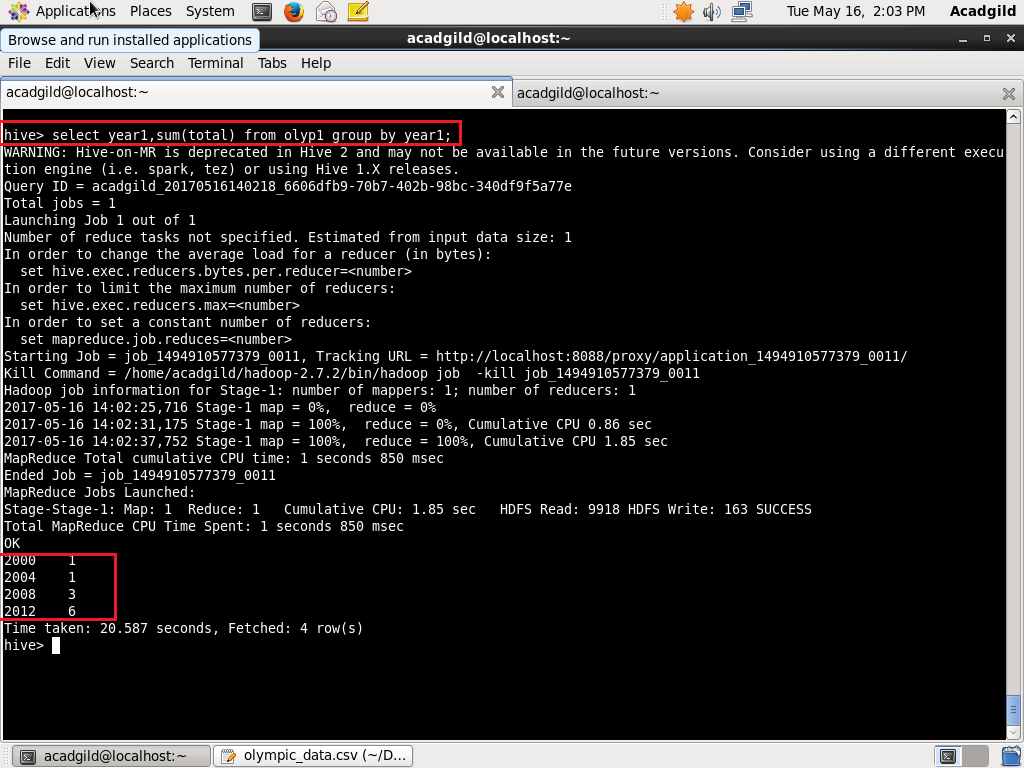
//creating partitioned table



//displaying contents of partitioned table

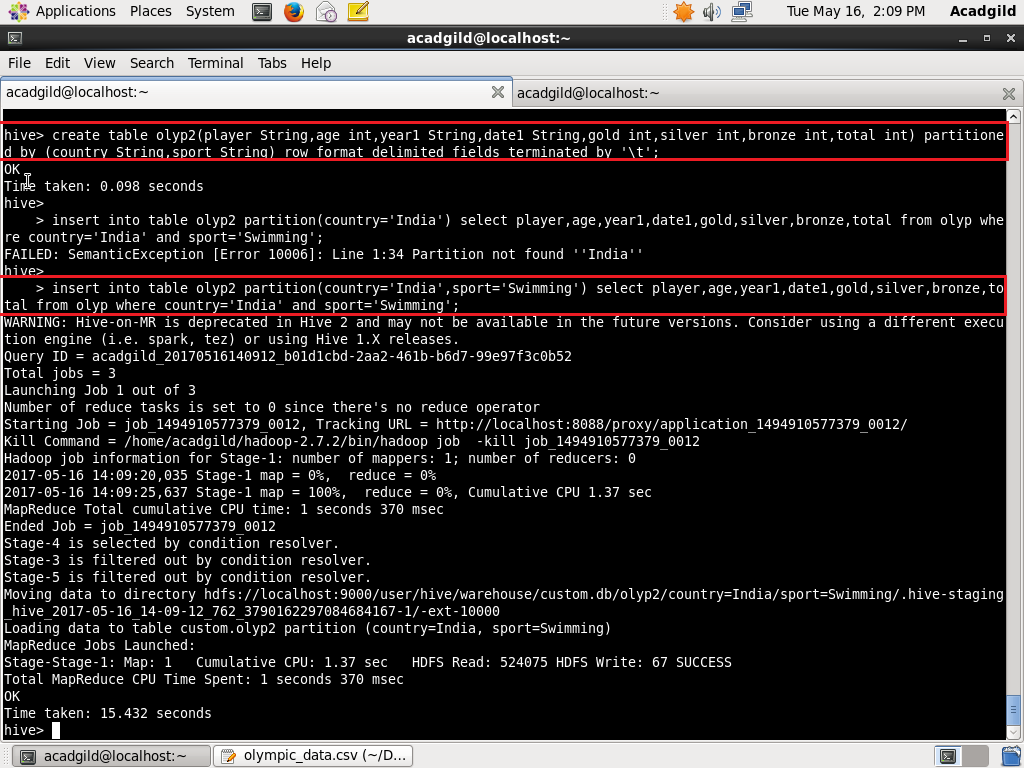


//final output

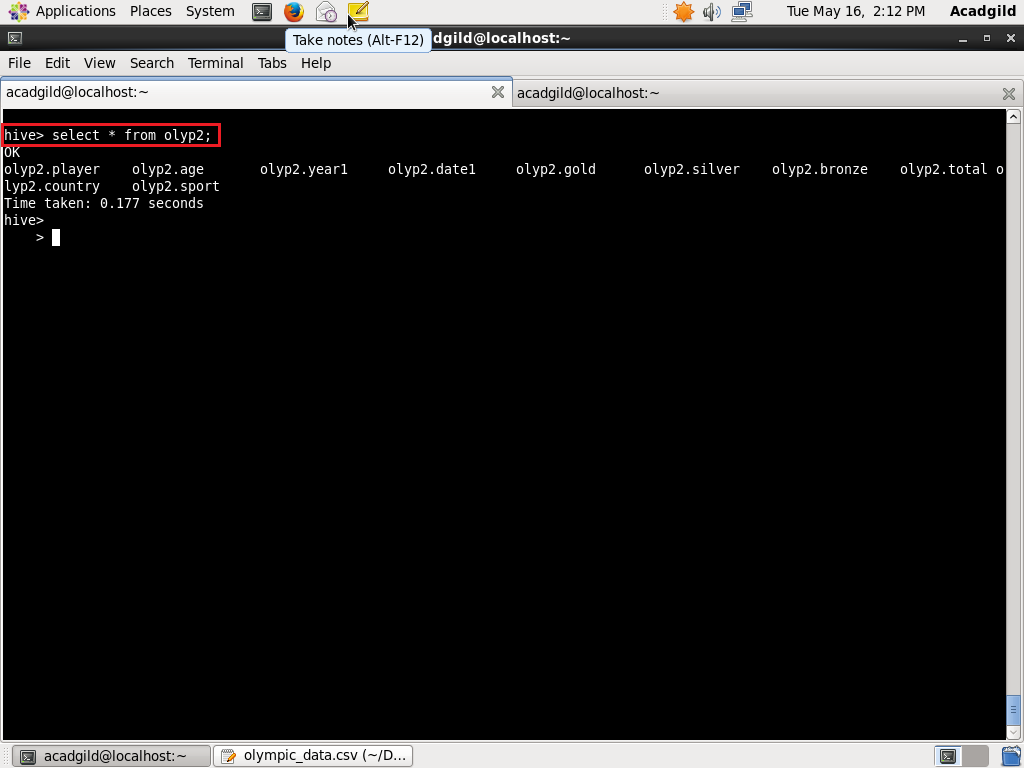


2. Find the number of medals india won in swimming year wise

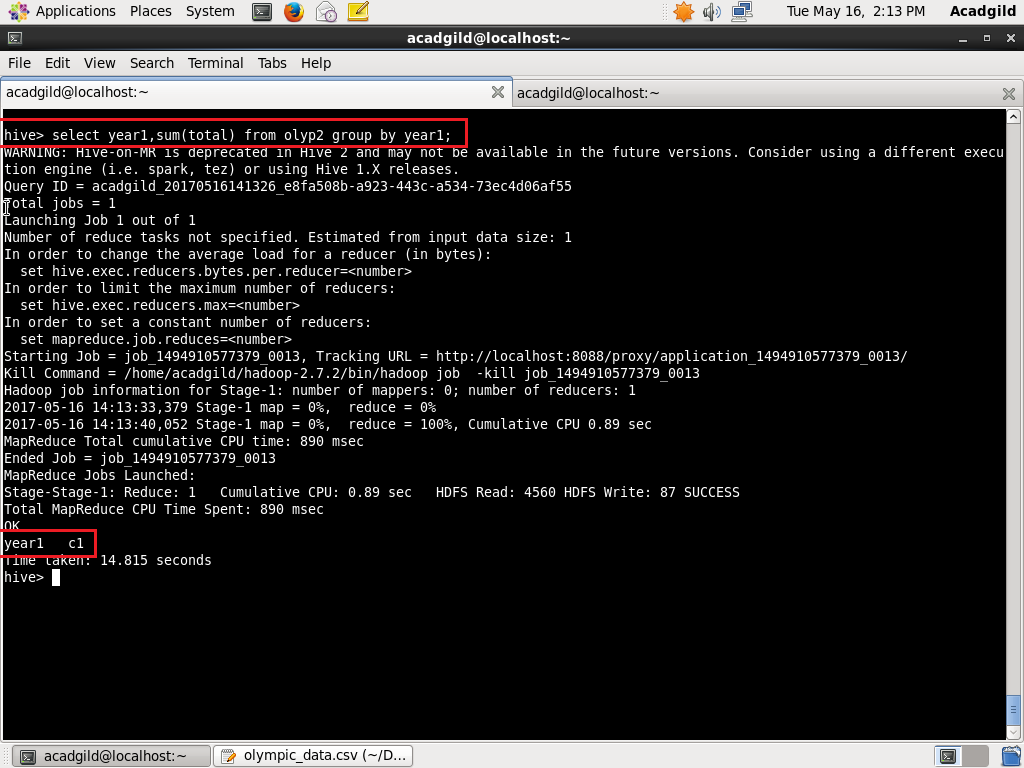
//creating partitioned table



//displaying contents of partitioned table

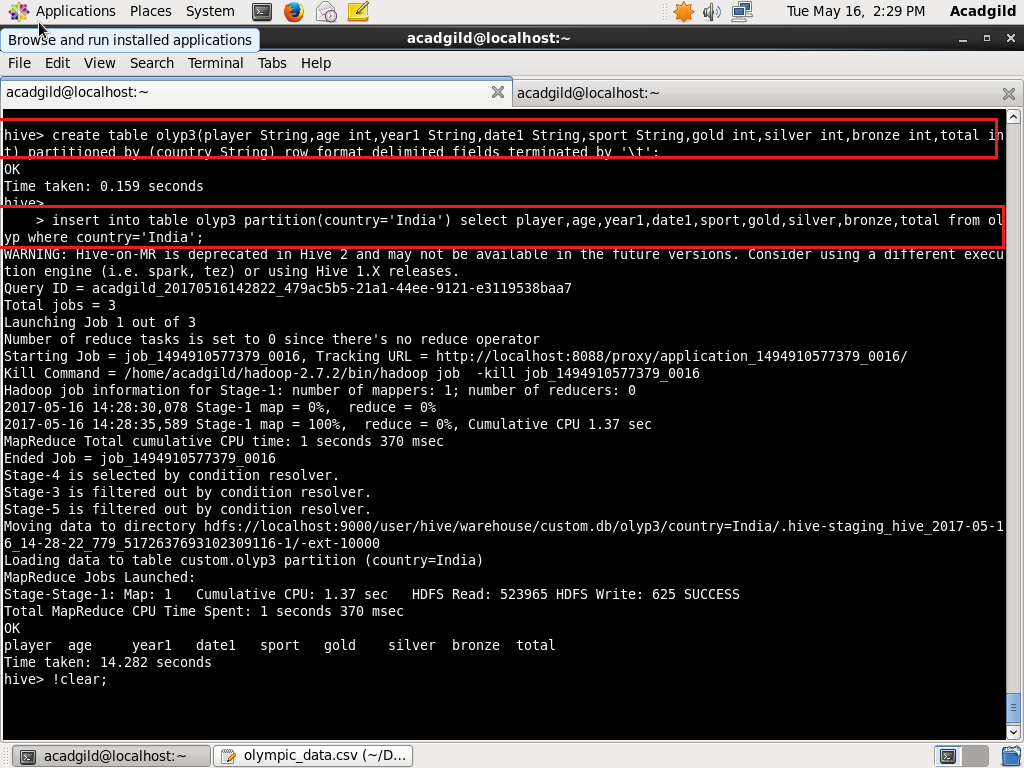


//final output

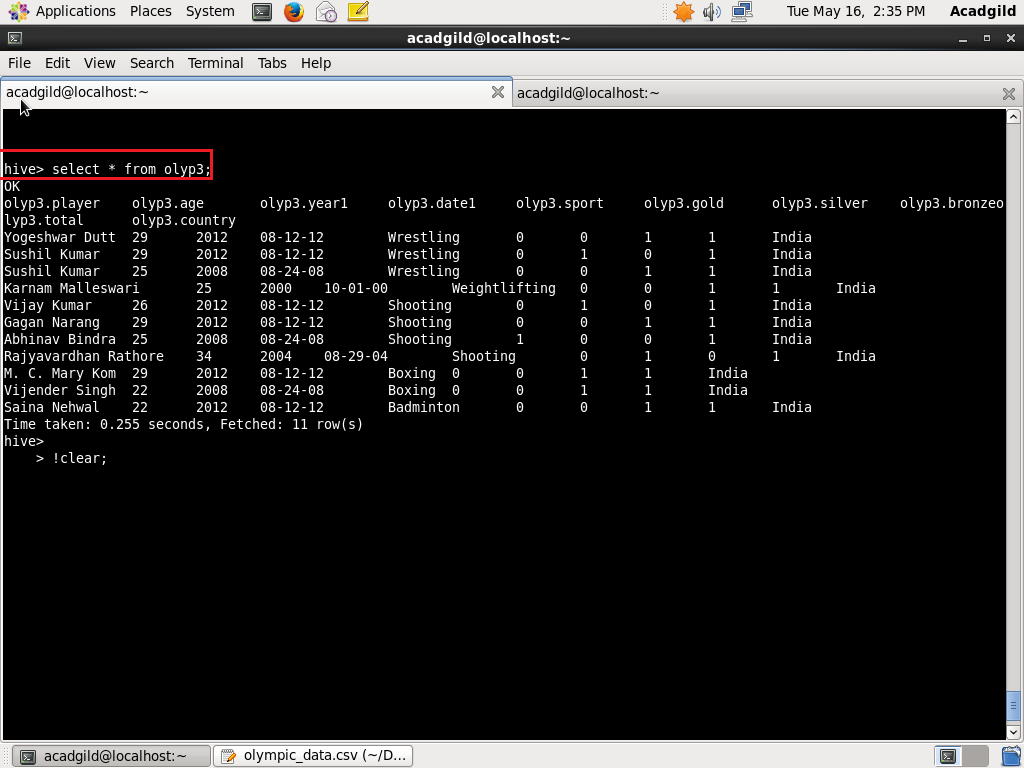


3. Find the number of gold and silver medals india won year wise

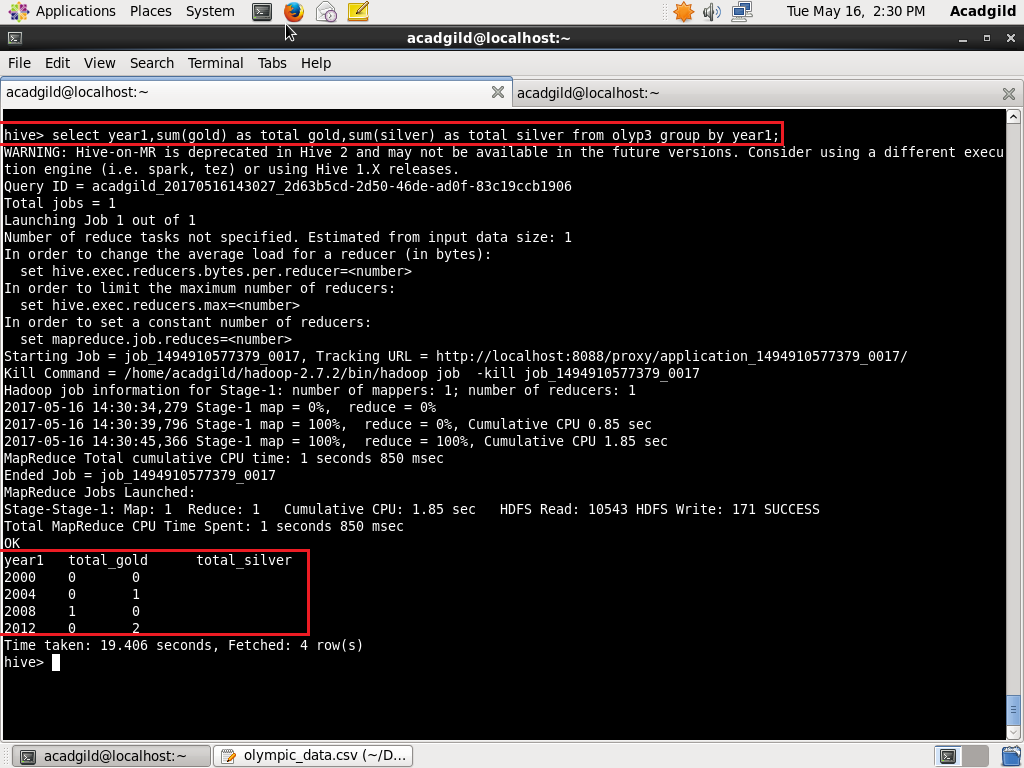
//creating partitioned table



//displaying the contents of partitioned table



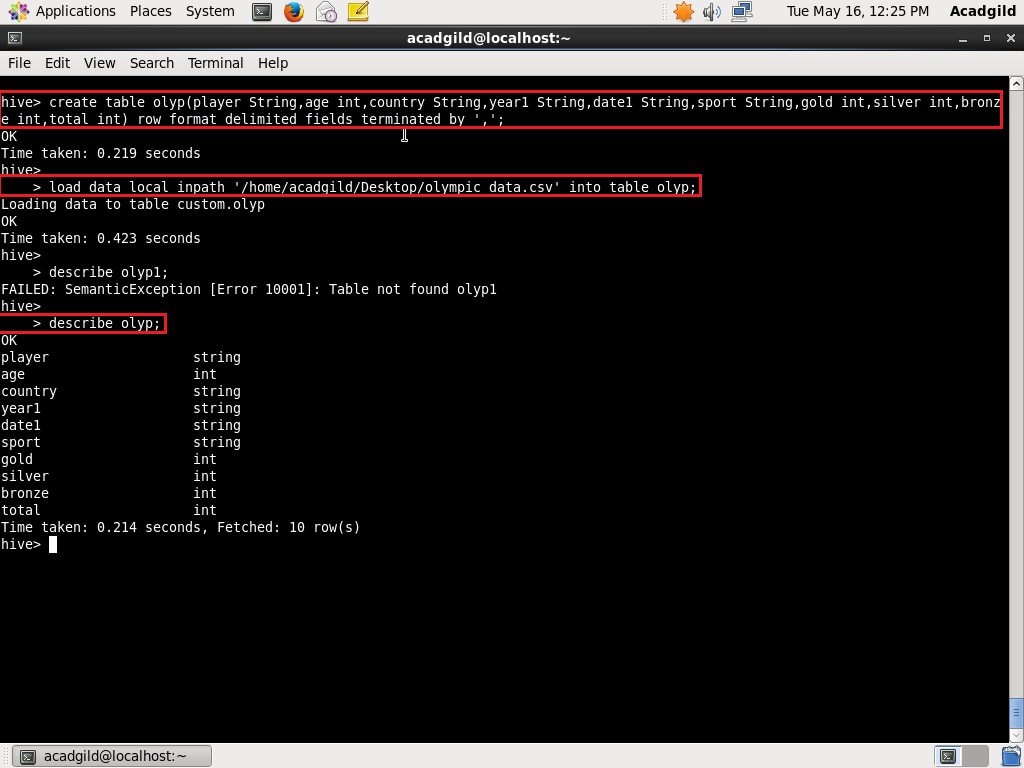
//final output



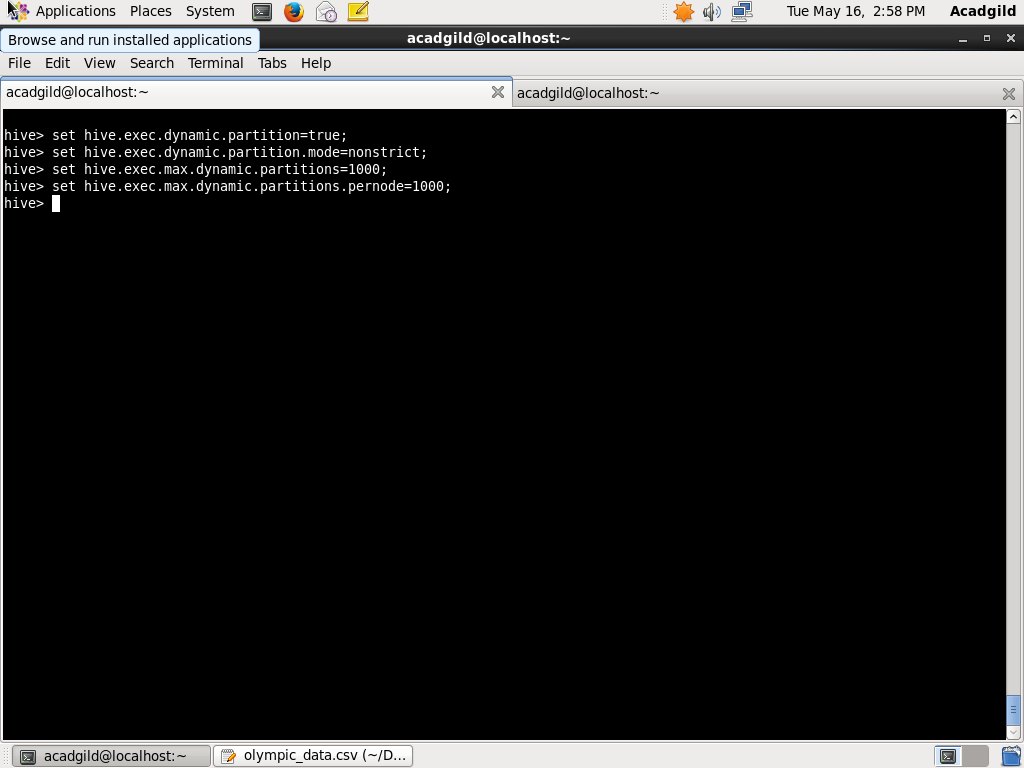
**Q3-**

**SOLUTION:**

//initially a table olyp is created which contains the entire dataset.This table is used as a source to insert data into other tables

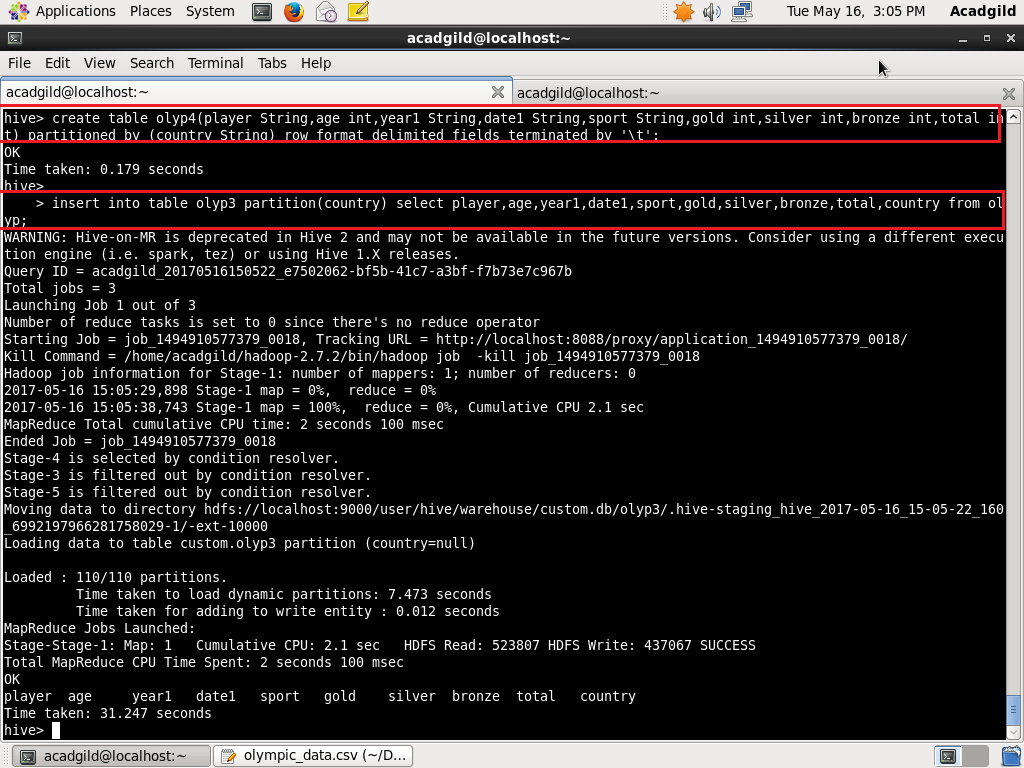


//setting properties to enable dynamic partitions

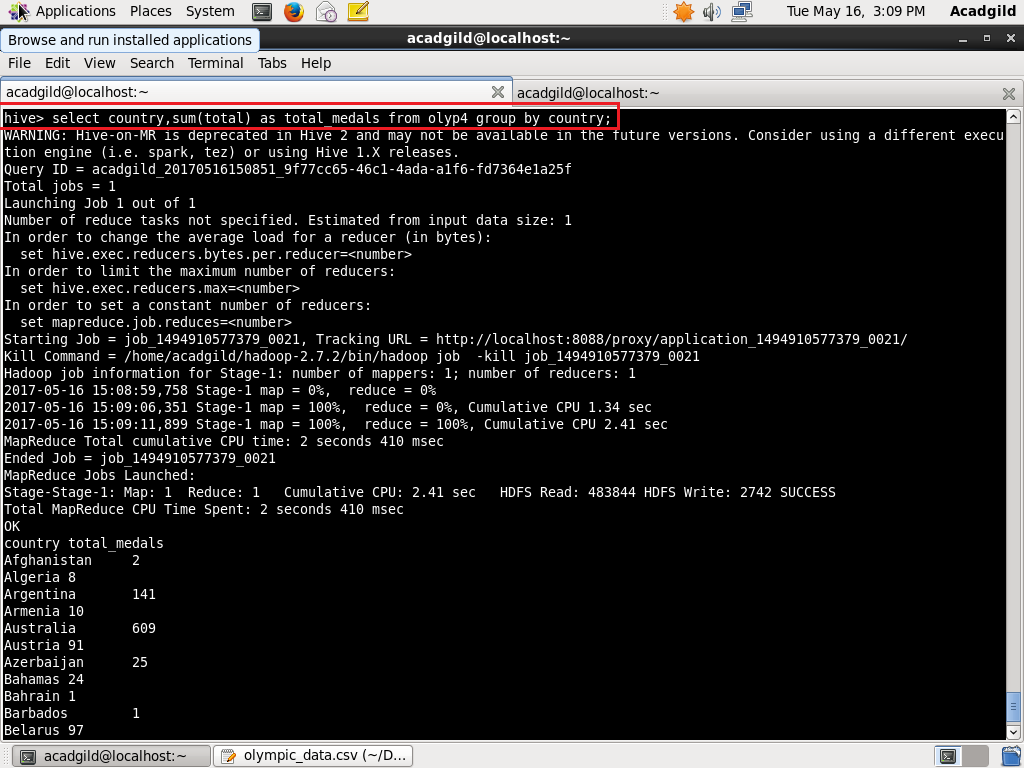


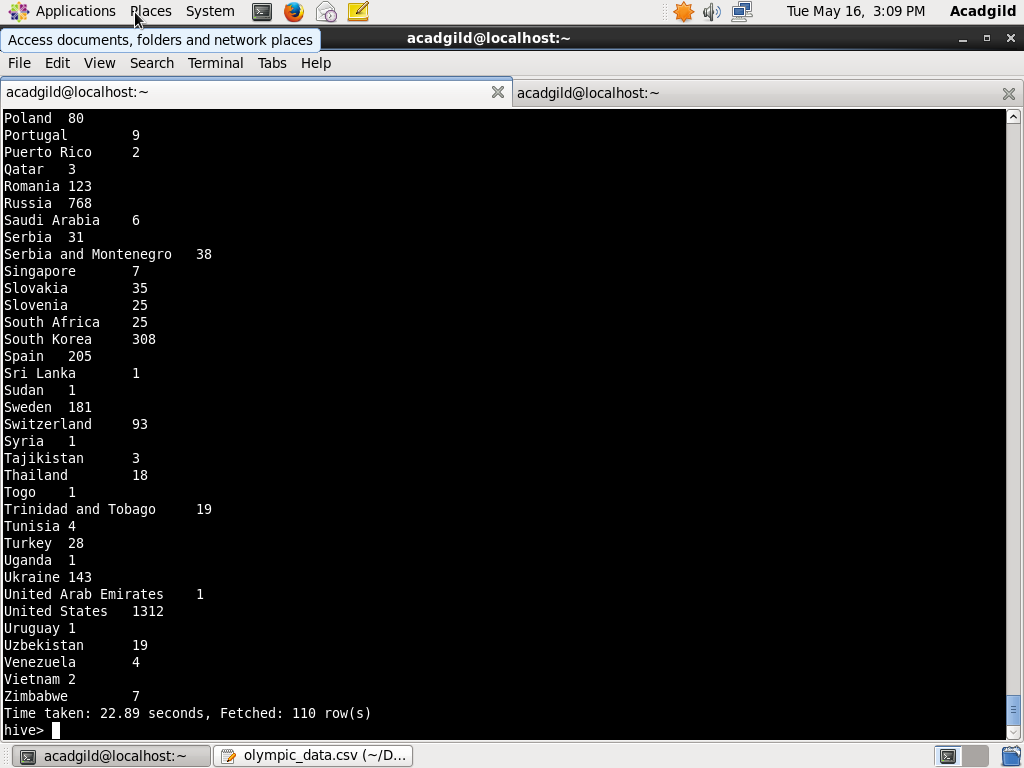
1. Find the total number of medals won by each country.

//creating partitioned table



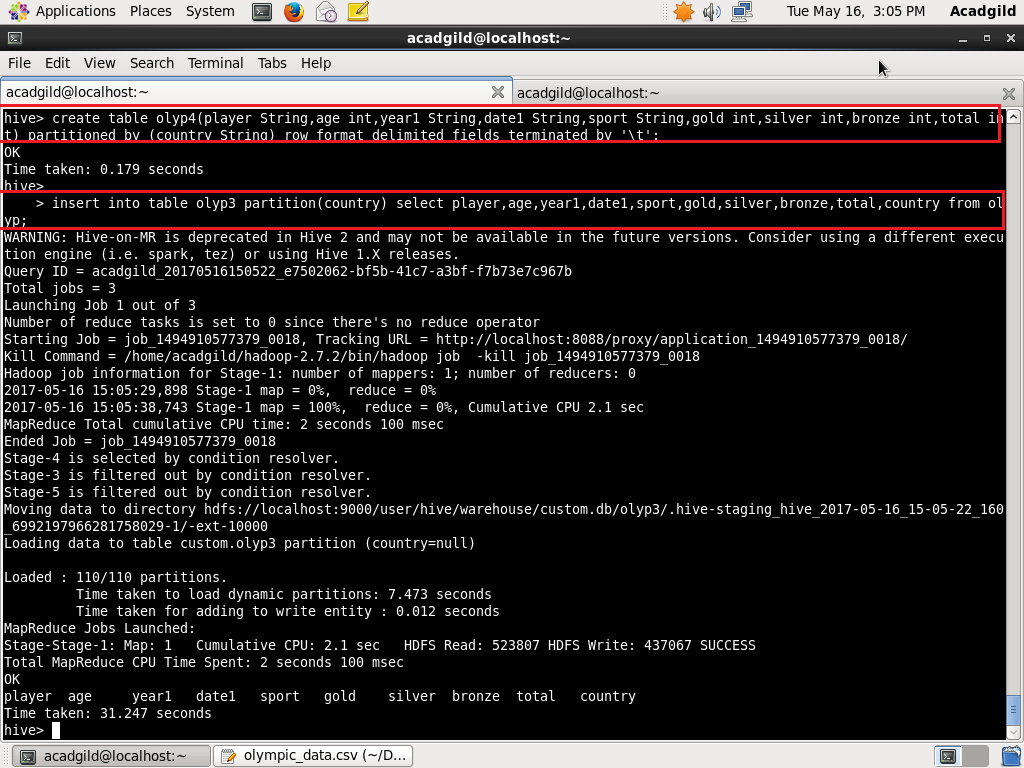
//final output



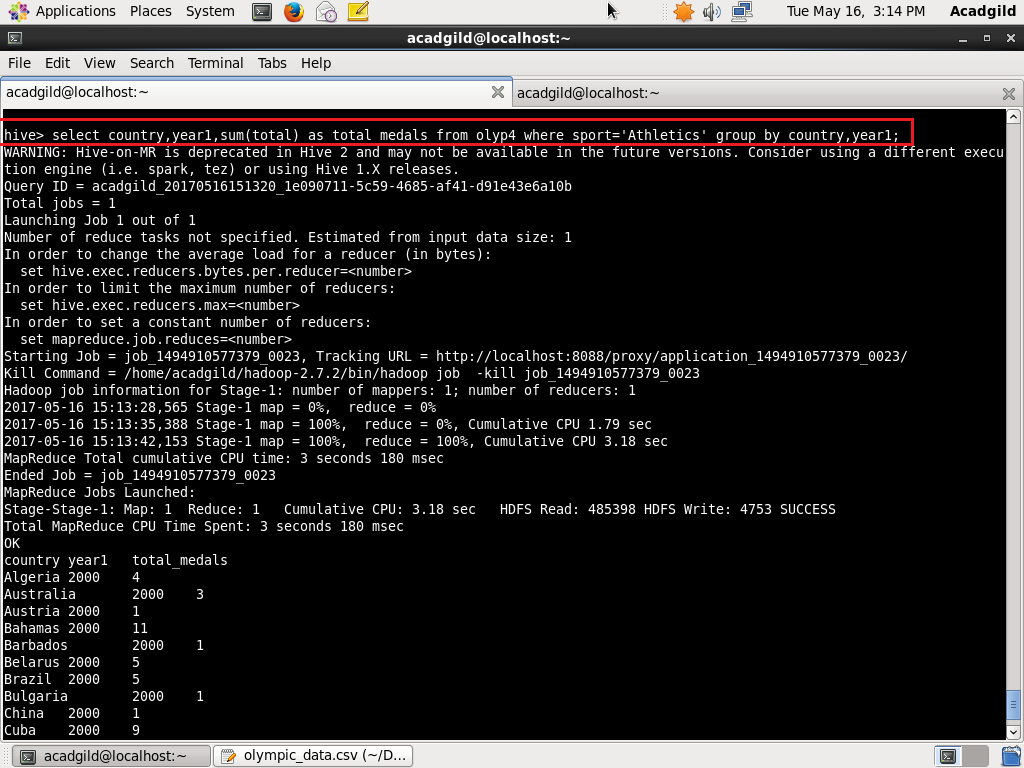


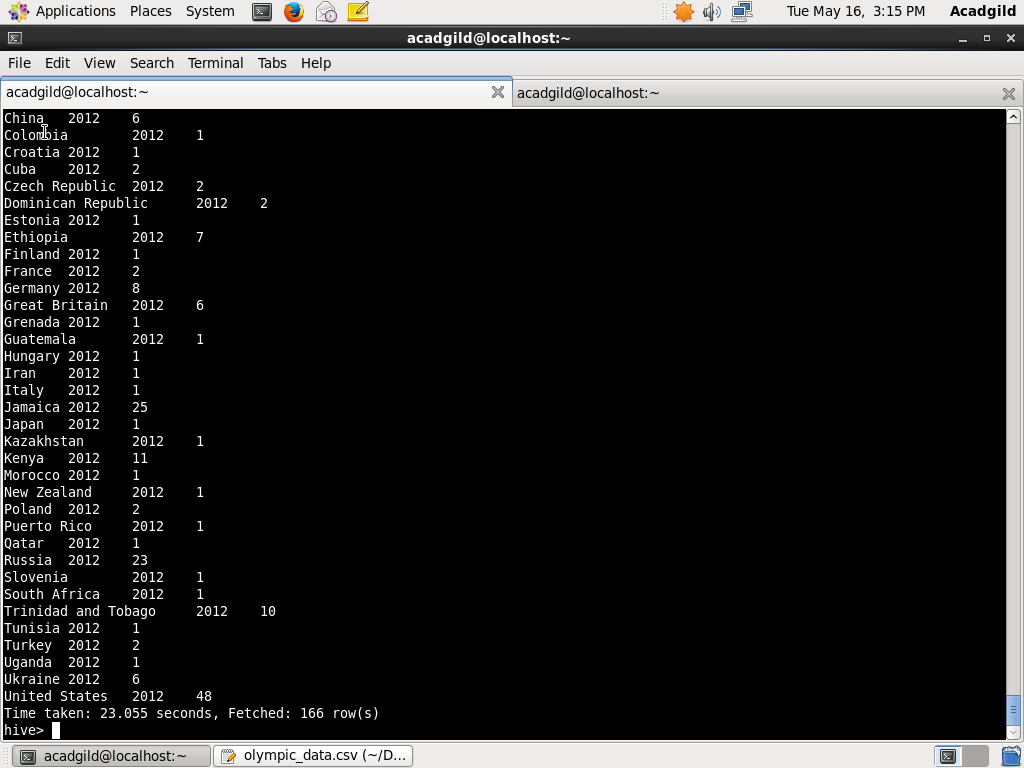
2. Find the number ot medals each country won in Athletics year wise

//creating partitioned table

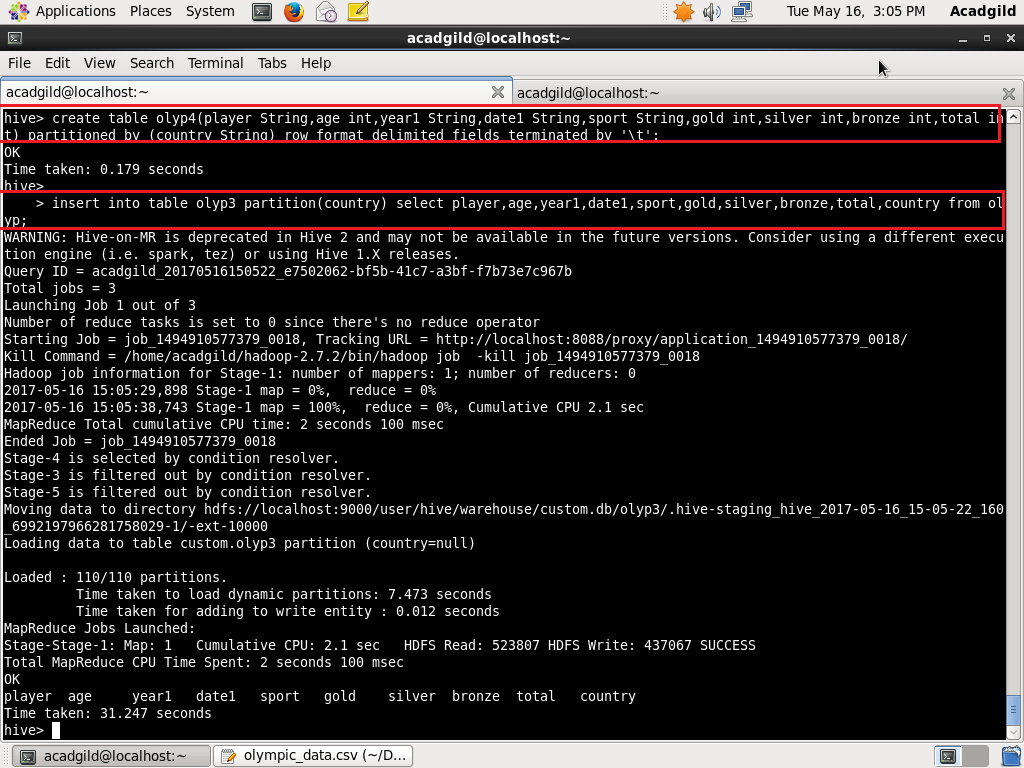


//final output





3. Find the average age of atheltes participated from each country in olympics year wise//creating partitioned table



//final output

