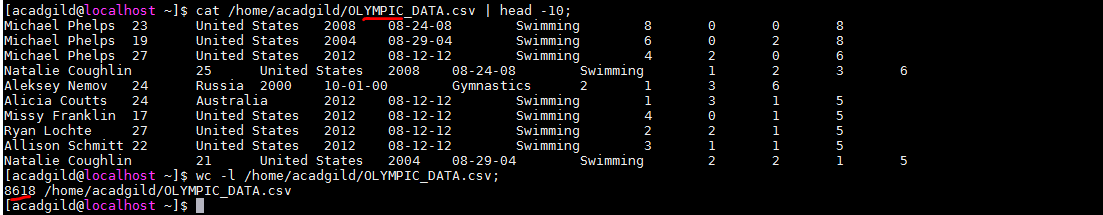
**ACD\_BDDOF\_Session\_3\_Assignment\_2\_Main**

**Problem Statement:**

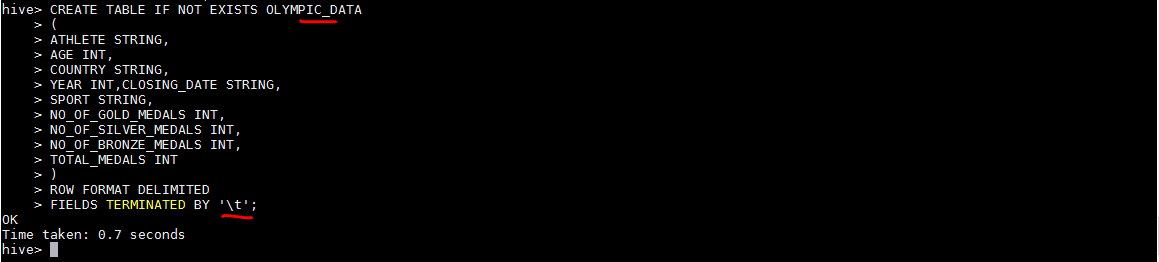
**Input dataset:**

File is tab separated “\t”.



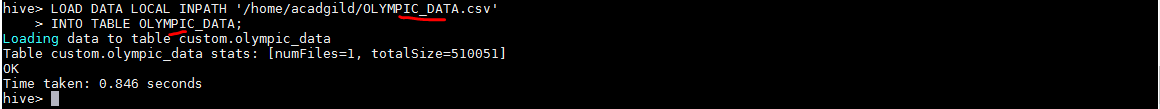
**Table creation:**

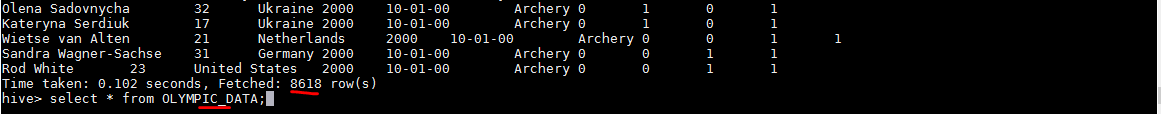
Created a table “OLYMPIC\_DATA” with fields terminated by ‘\t’. IF NOT EXISTS will suppress warnings if table already exists else table will be created.



**Data load:**

All the records from input dataset is loaded to the table.

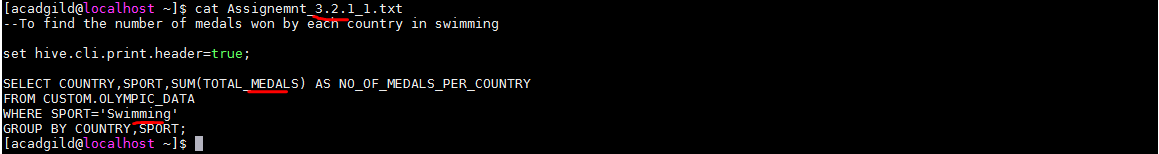




1. Write a Hive program to find the number of medals won by each country in swimming.

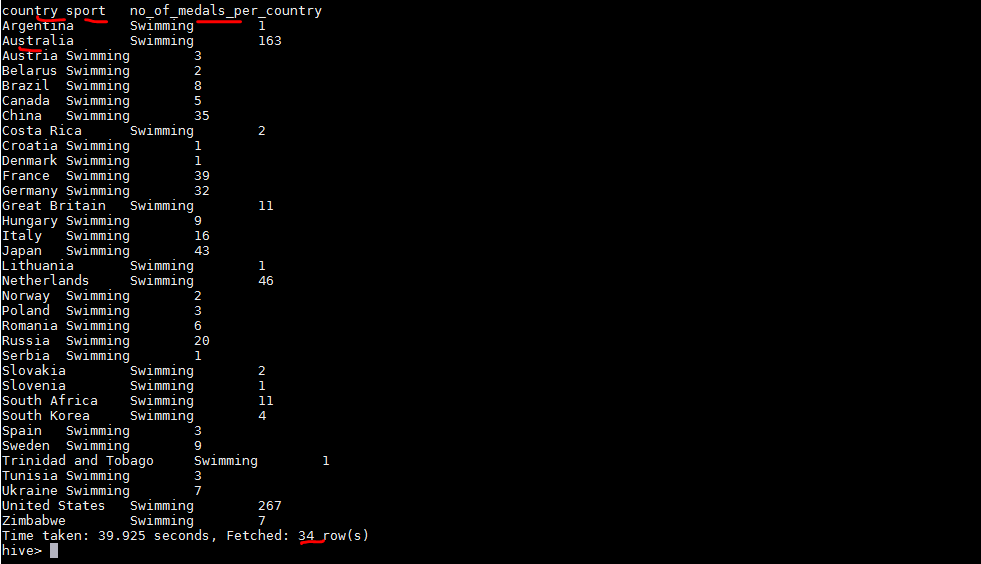
**Hive Program:**

Grouped by COUNTRY,SPORT and got SUM(TOTAL\_MEDALS) for records having SPORT as “Swimming”.



**Hive program execution:**

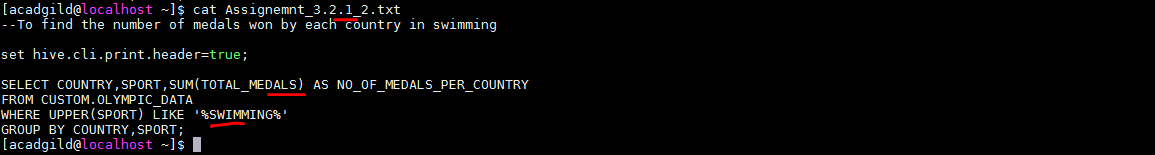
**Output:**

Displays Country, Sport(swimming),no of total medals for each country.

(OR)

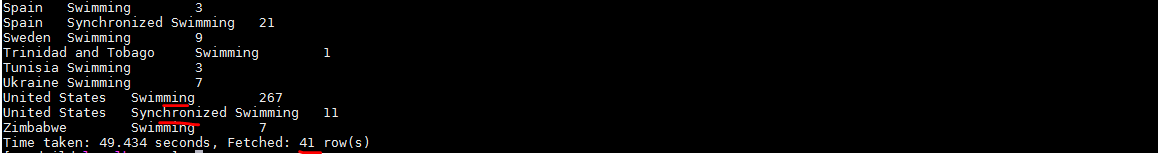
**Alternative:**

There is other sport called Synchronized Swimming. To get both swimming **g**rouped by COUNTRY,SPORT and got SUM(TOTAL\_MEDALS) for records having SPORT like “%Swimming%”- will get both Synchronized Swimming & Swimming .



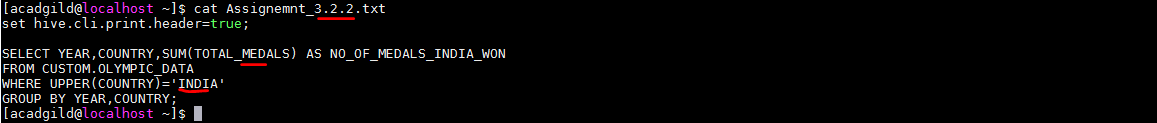


**Output:**



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

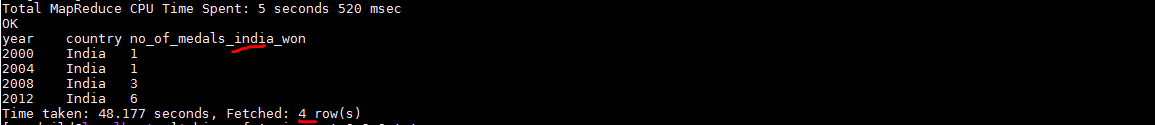
Grouped by COUNTRY,YEAR and got SUM(TOTAL\_MEDALS) for records having COUNTRY as ‘INDIA’.



**Hive Program execution:**



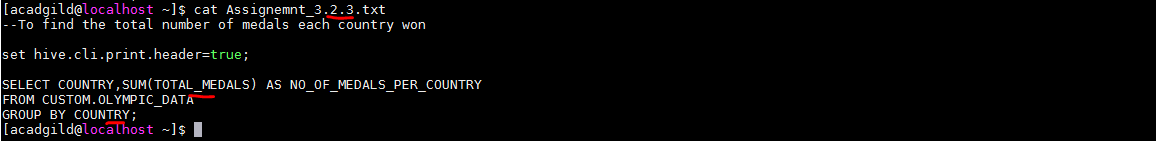
**Output:**

Displays the SUM(TOTAL\_MEDALS) across different YEAR,COUNTRY(INDIA).

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

**Hive Program:**

Grouped by COUNTRY and got SUM(TOTAL\_MEDALS) across all records.

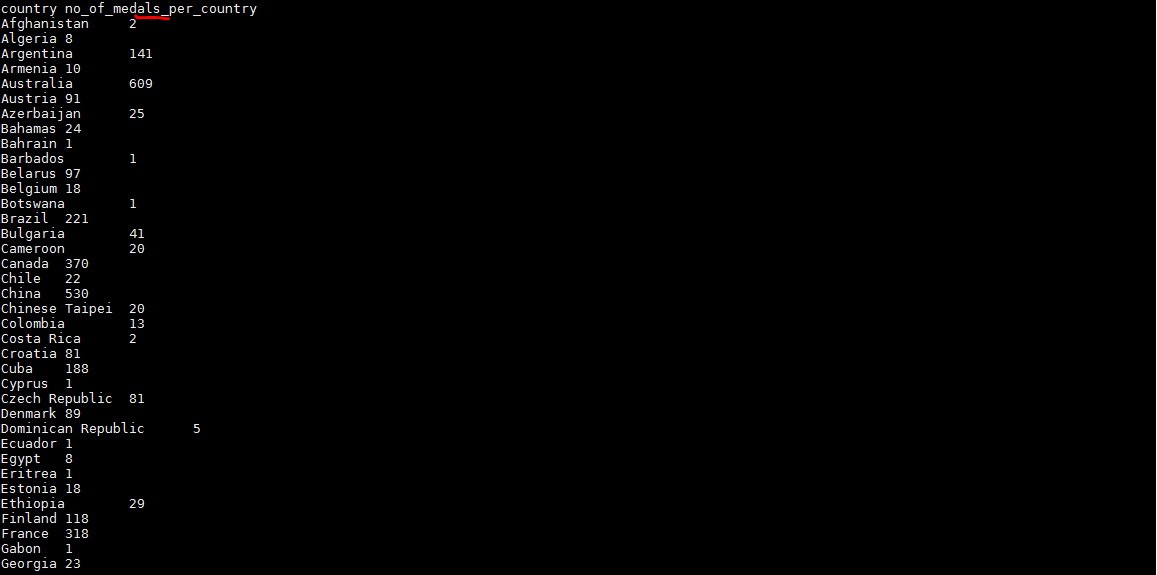


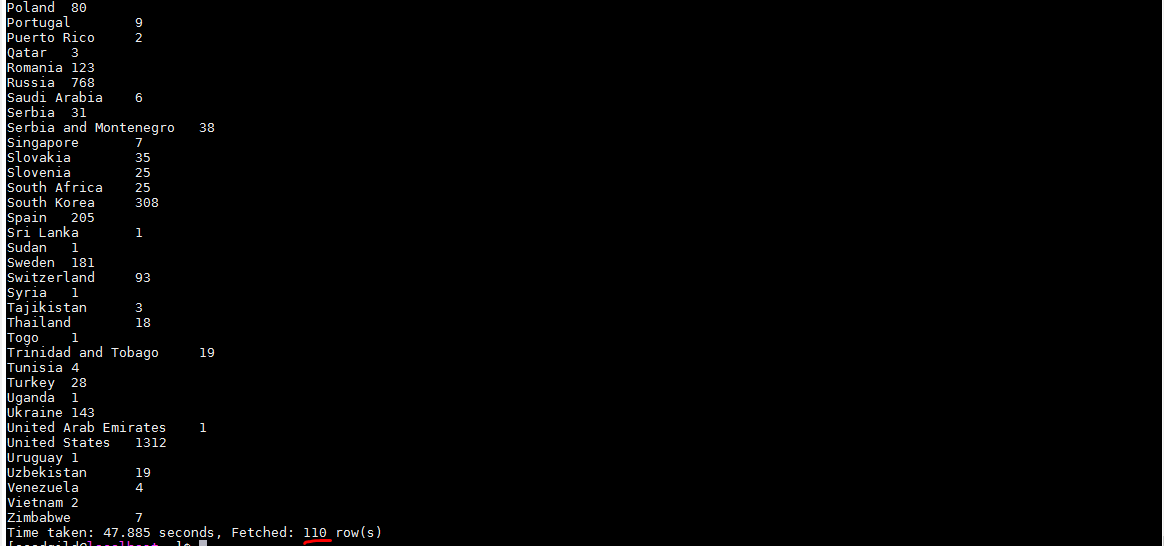
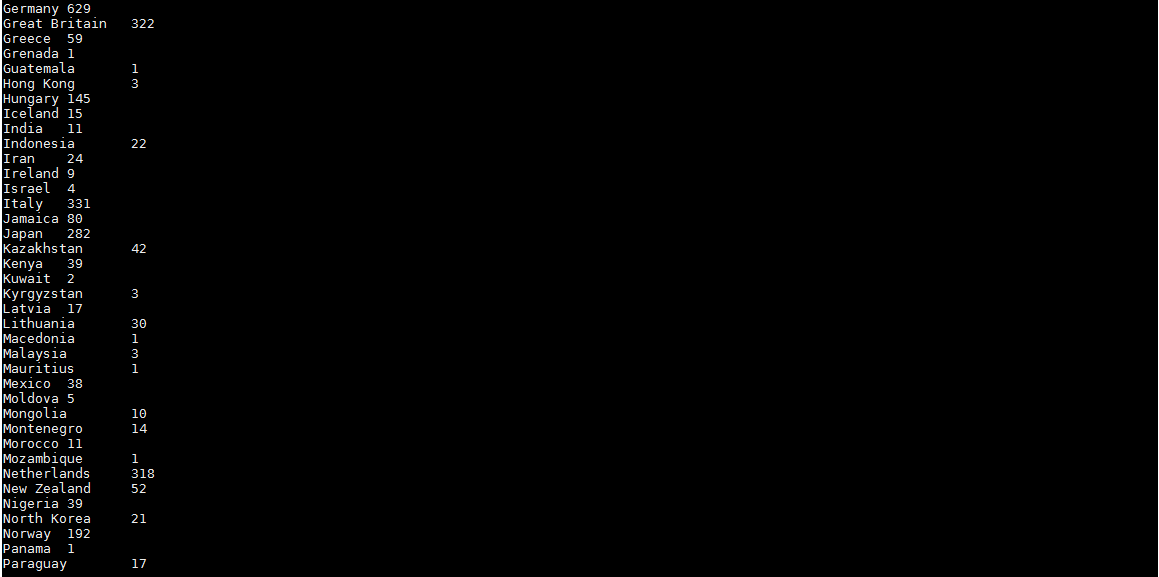
**Hive Program execution:**



**Output:**

Displays the SUM(TOTAL\_MEDALS) for every country.

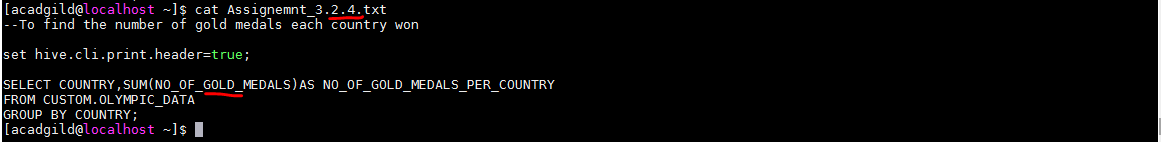




4.Write a Hive program to find the number of gold medals each country won.

**Hive Program:**

Grouped by COUNTRY and got SUM(GOLD\_MEDALS) .



**Hive program execution:**



**Output:**

Displays the SUM(GOLD\_MEDALS) for every country.