CREATE DATABASE IF NOT EXISTS olympics ;

CREATE TABLE IF NOT EXISTS olympic\_winners

(

athlete\_name string,

age int,

country string,

year int,

closing\_date string,

sport string,

gold\_medals int,

silver\_medals int,

bronze\_medals int,

total\_medals\_won int

)

row format delimited fields terminated by '\t' ;

LOAD DATA LOCAL INPATH "/home/cloudera/chhaya/hive/olympix\_data.csv" INTO TABLE olympic\_winners;

SELECT COUNT(\*) FROM olympic\_winners ;

Partitioning the table with column country name---

CREATE TABLE IF NOT EXISTS olympics\_partitioned

(

athlete\_name string,

age int,

year int,

closing\_date string,

sport string,

gold\_medals int,

silver\_medals int,

bronze\_medals int,

total\_medals\_won int

) partitioned by (country string)

;

set hive.exec.dynamic.partition.mode = nonstrict;

INSERT OVERWRITE TABLE olympics\_partitioned

partition(country = 'India')

SELECT athlete\_name,

age,

year int,

closing\_date,

sport,

gold\_medals,

silver\_medals,

bronze\_medals,

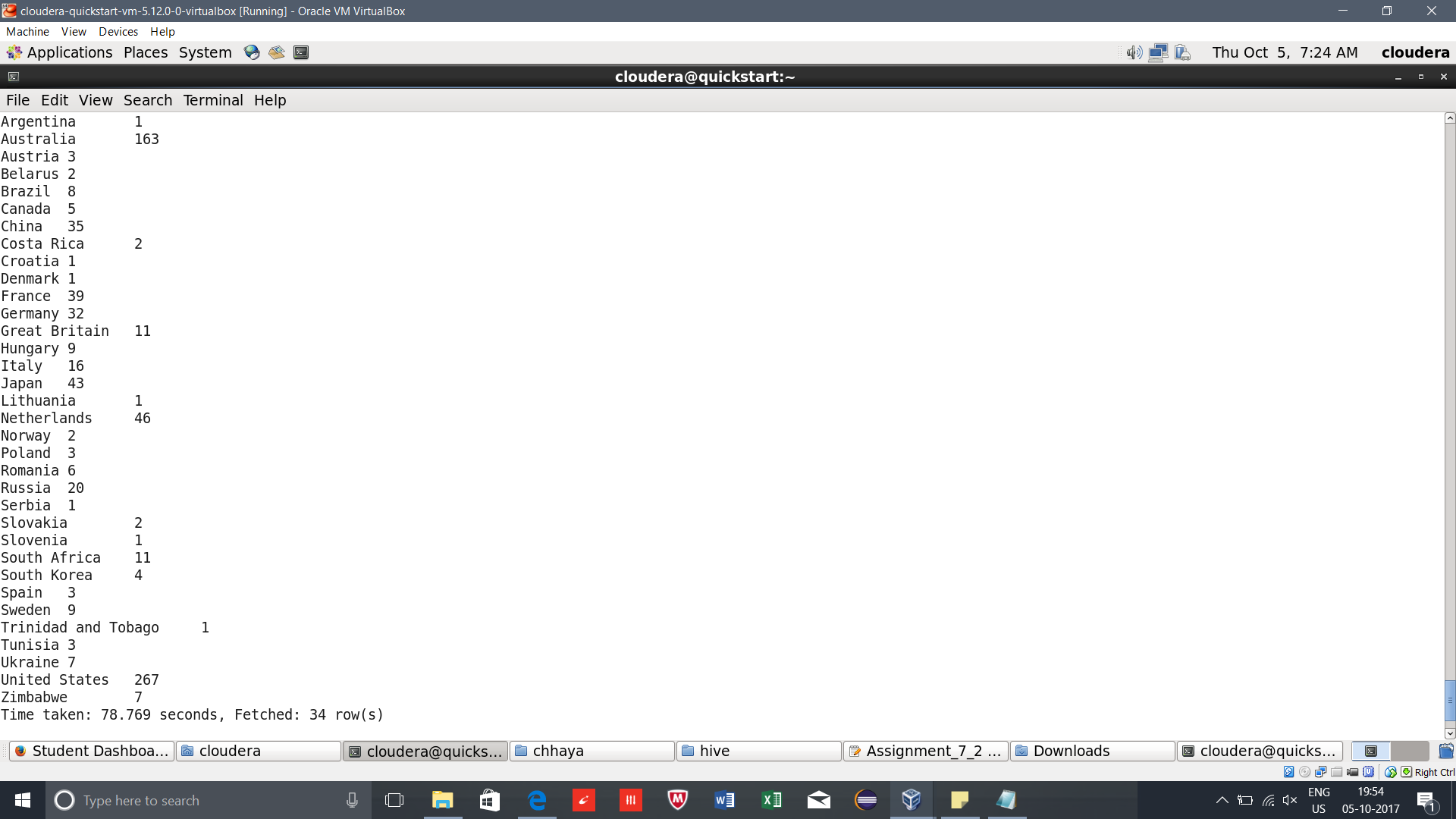
total\_medals\_won

FROM olympic\_winners where country = 'India';

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

hive (olympics)> SELECT country, SUM(total\_medals\_won) FROM olympic\_winners WHERE sport = 'Swimming' GROUP BY country ORDER BY country;

**Output 7.2.1:**



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

hive (olympics)> SELECT year,sum(total\_medals\_won) from olympics\_partitioned where country = 'India' group by year order by year;

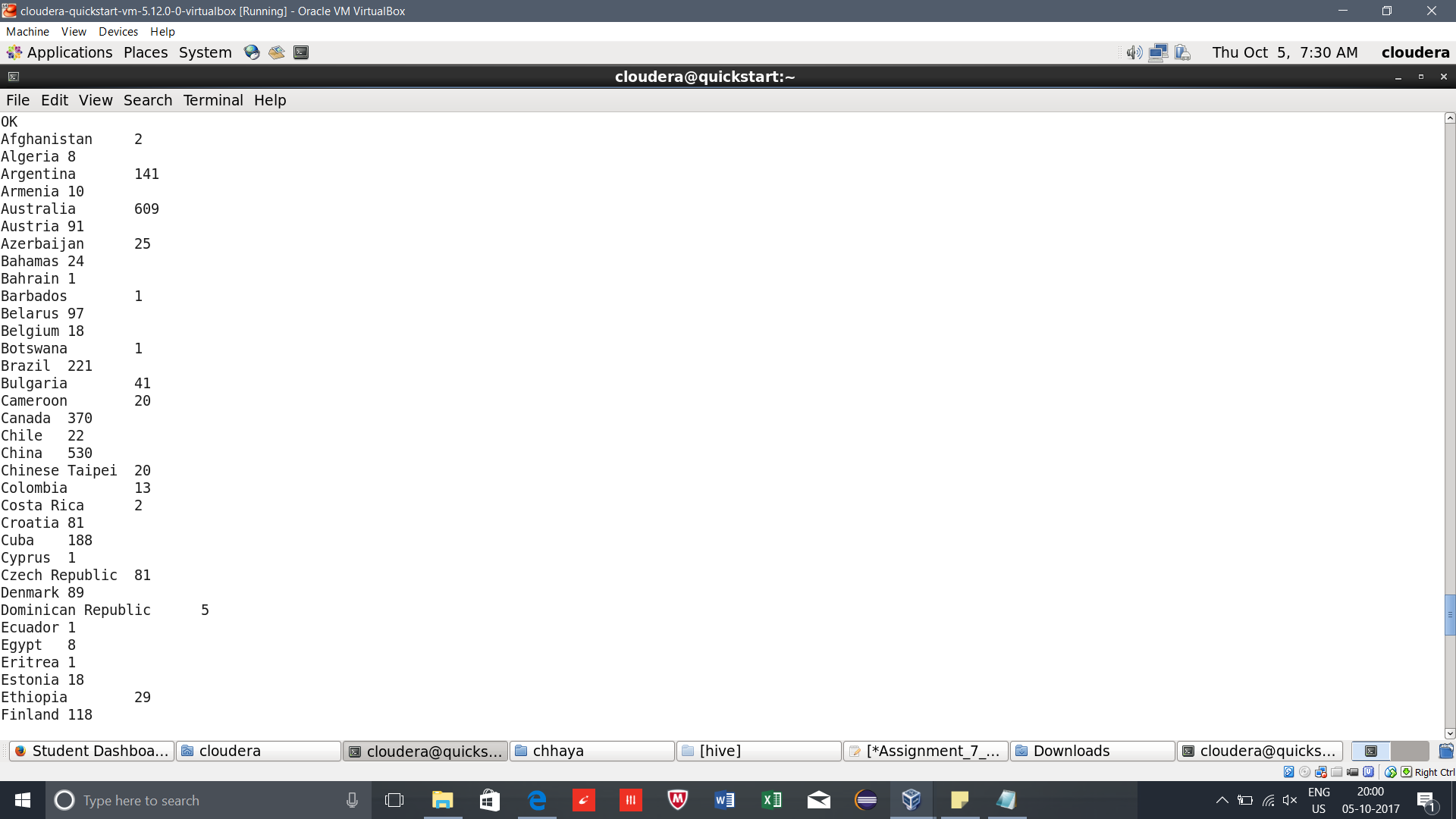
**Output 7.2.2:**

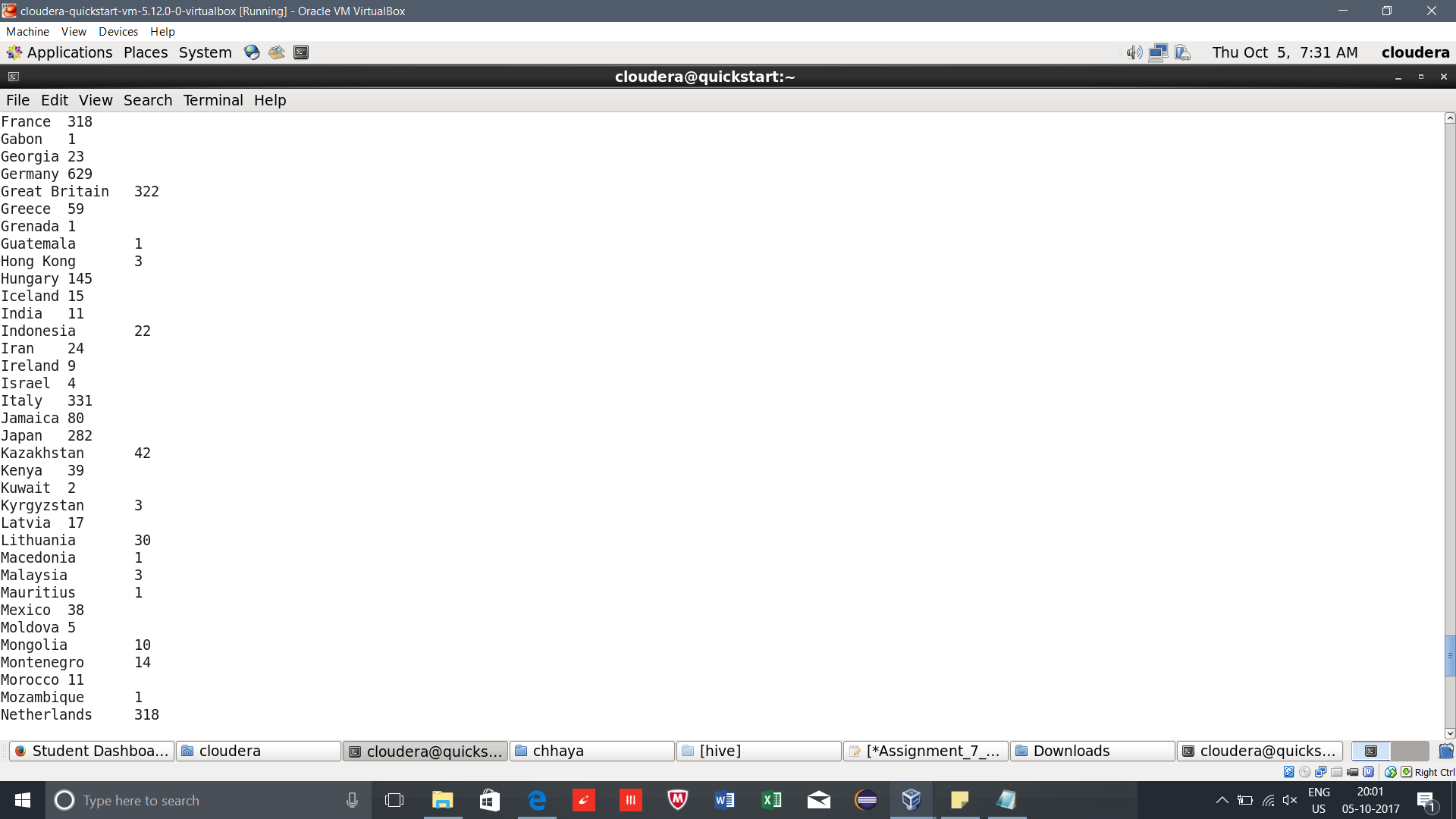


**Problem Statement 3:3. Write a Hive Program to find the total number of medals each country won.**

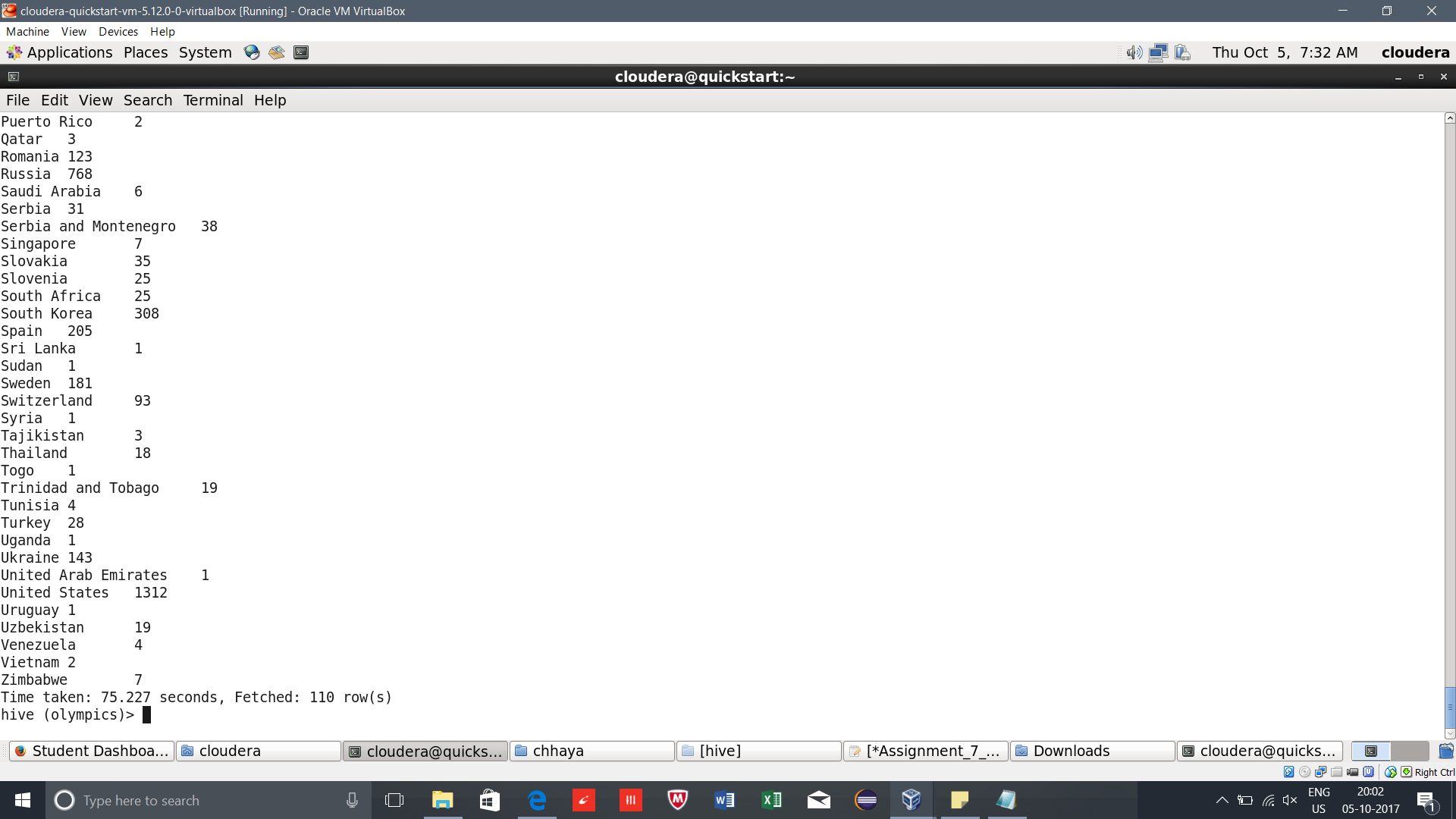
hive (olympics)> SELECT country,sum(total\_medals\_won) from olympic\_winners group by country order by country;

**Output 7.2.3**









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

hive (olympics)> SELECT country, SUM(gold\_medals) from olympic\_winners group by country order by country;

**Output 7.2.4:**

