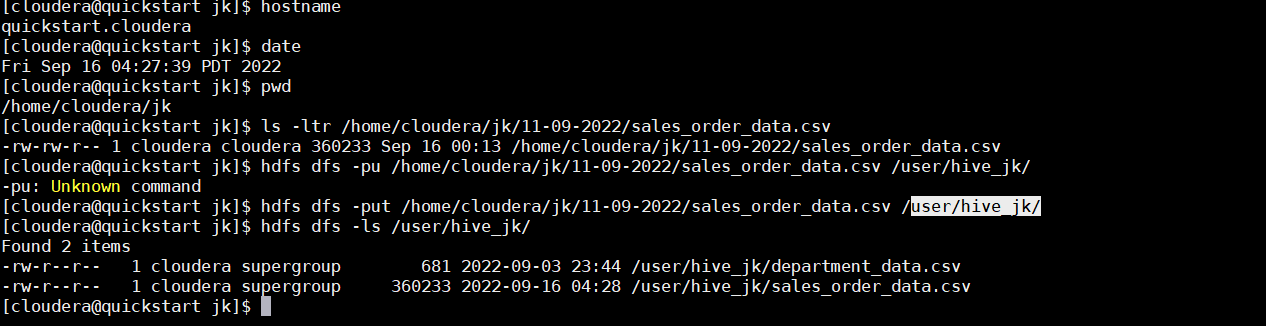
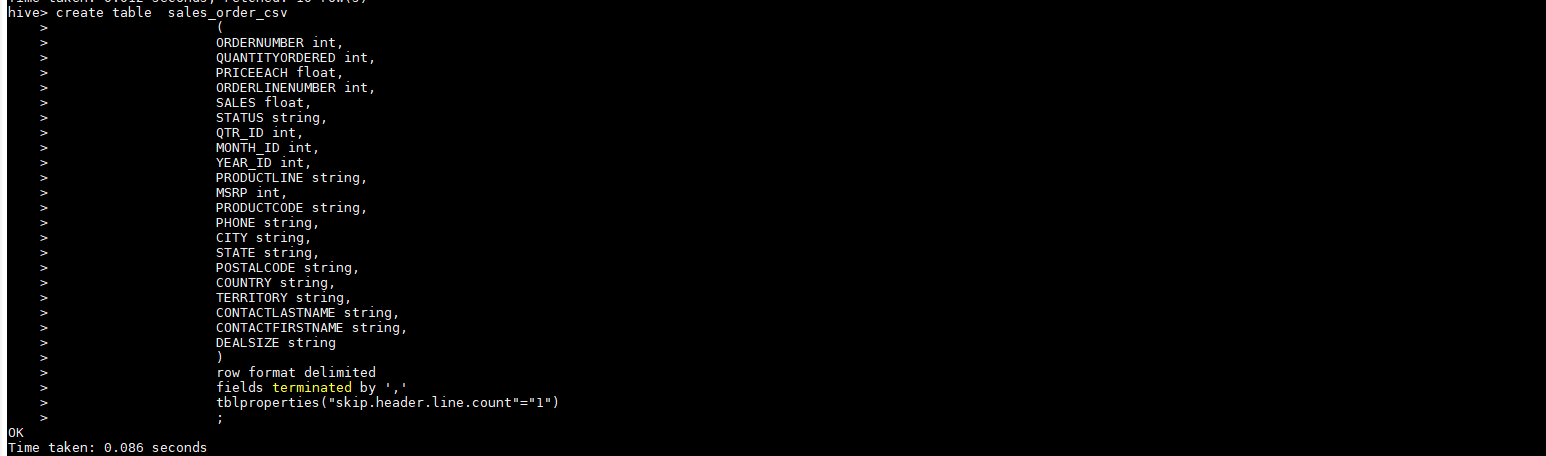
1. Download vechile sales data -> https://github.com/shashank-mishra219/Hive-Class/blob/main/sales\_order\_data.csv

2. Store raw data into hdfs location

Ans JK:

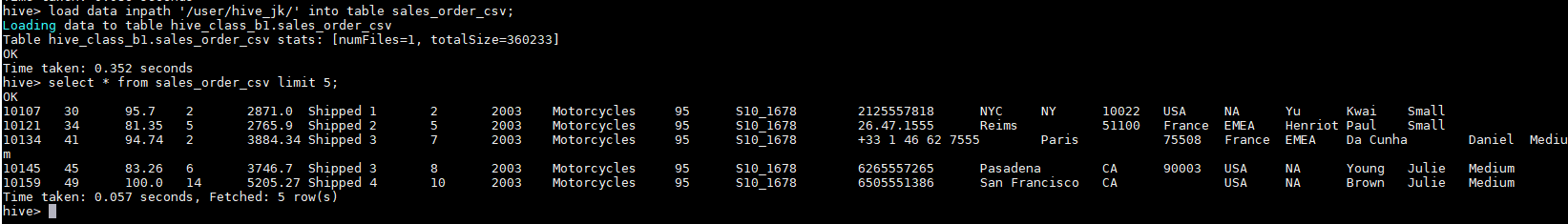


3. Create a internal hive table "sales\_order\_csv" which will store csv data sales\_order\_csv .. make sure to skip header row while creating table.



4. Load data from hdfs path into "sales\_order\_csv"

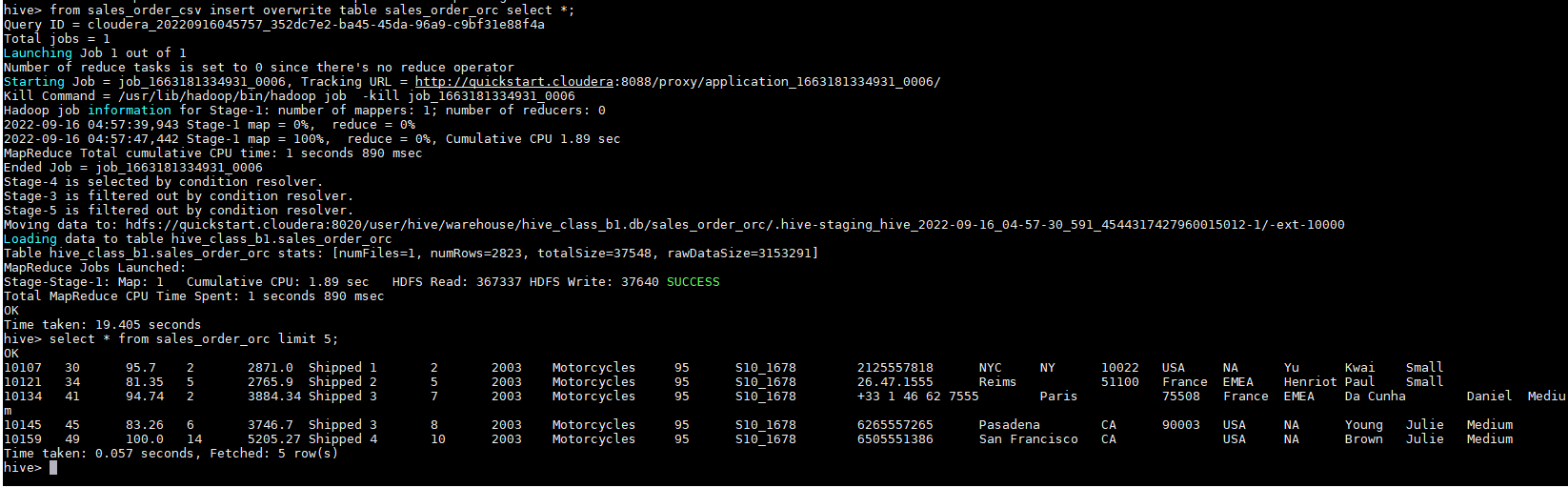
*Here I faced one issue , I learned that only the original data csv file can be placed here , if more csv files even though different they are in format the data can’t be loaded properly , it was dumping null values and all the csv files are disappeared (only original expected) from hdfs location as well.*



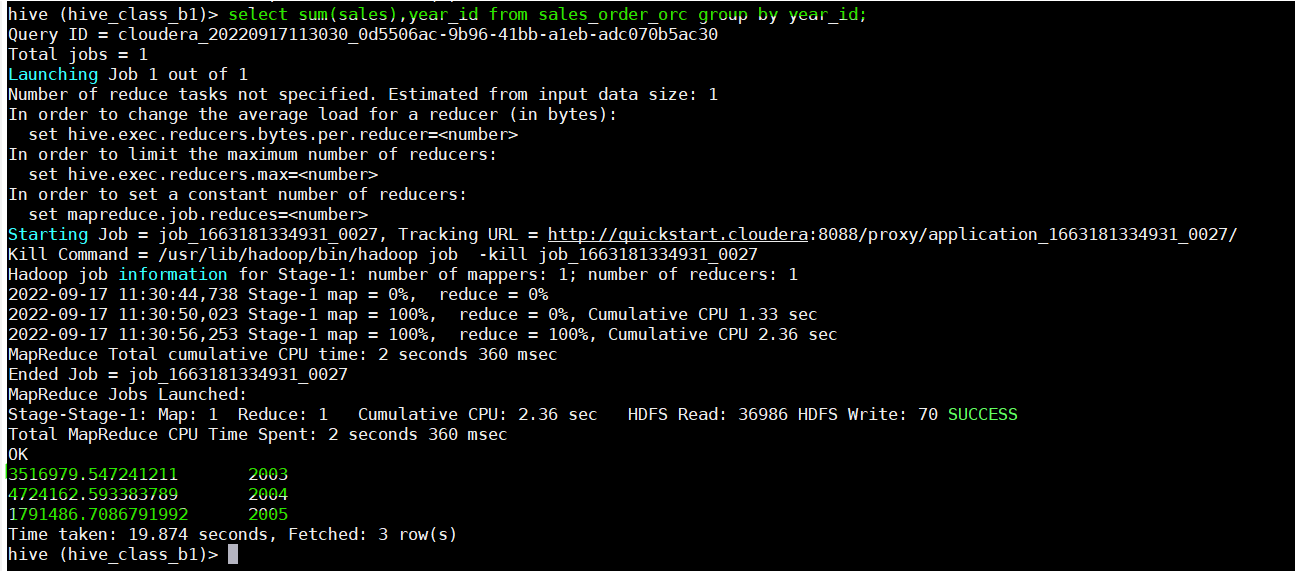
5. Create an internal hive table which will store data in ORC format "sales\_order\_orc"



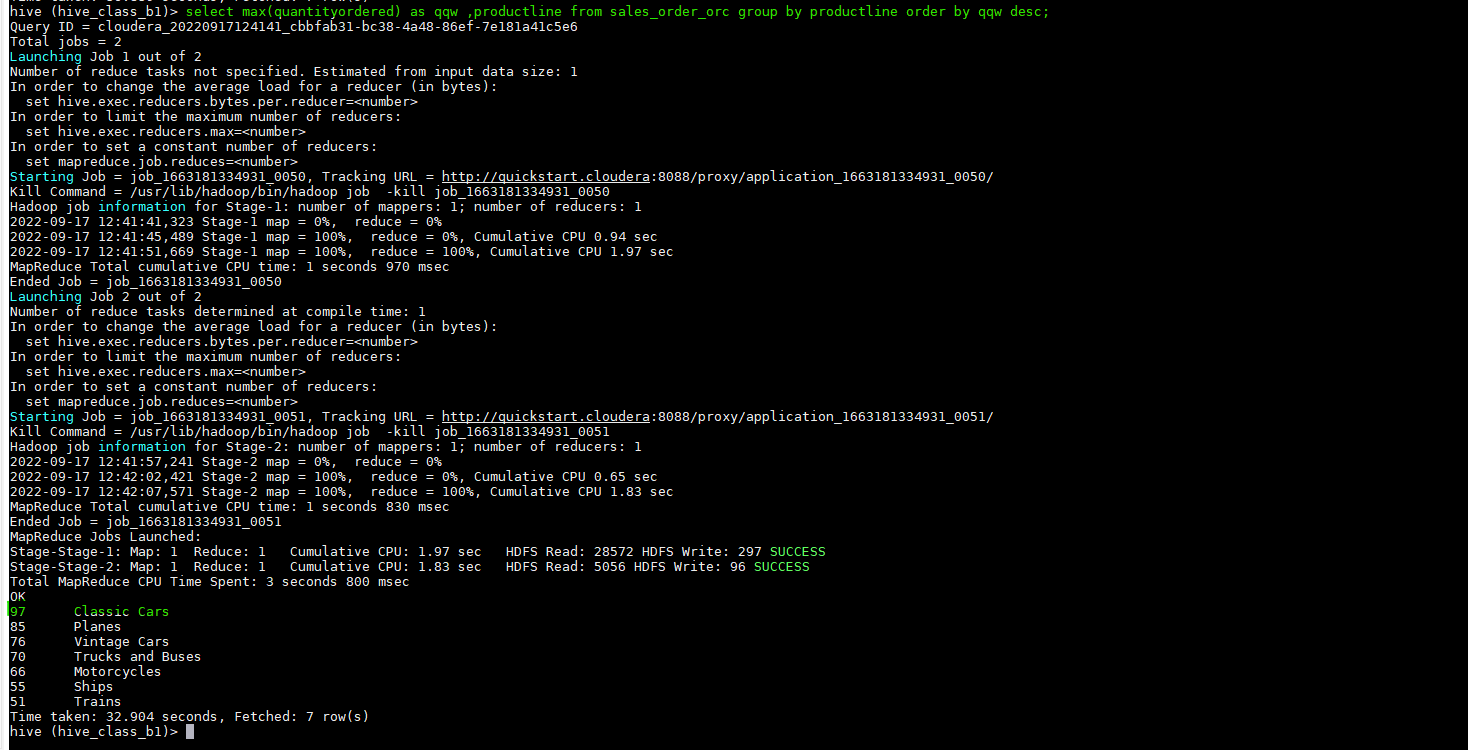
6. Load data from "sales\_order\_csv" into "sales\_order\_orc"



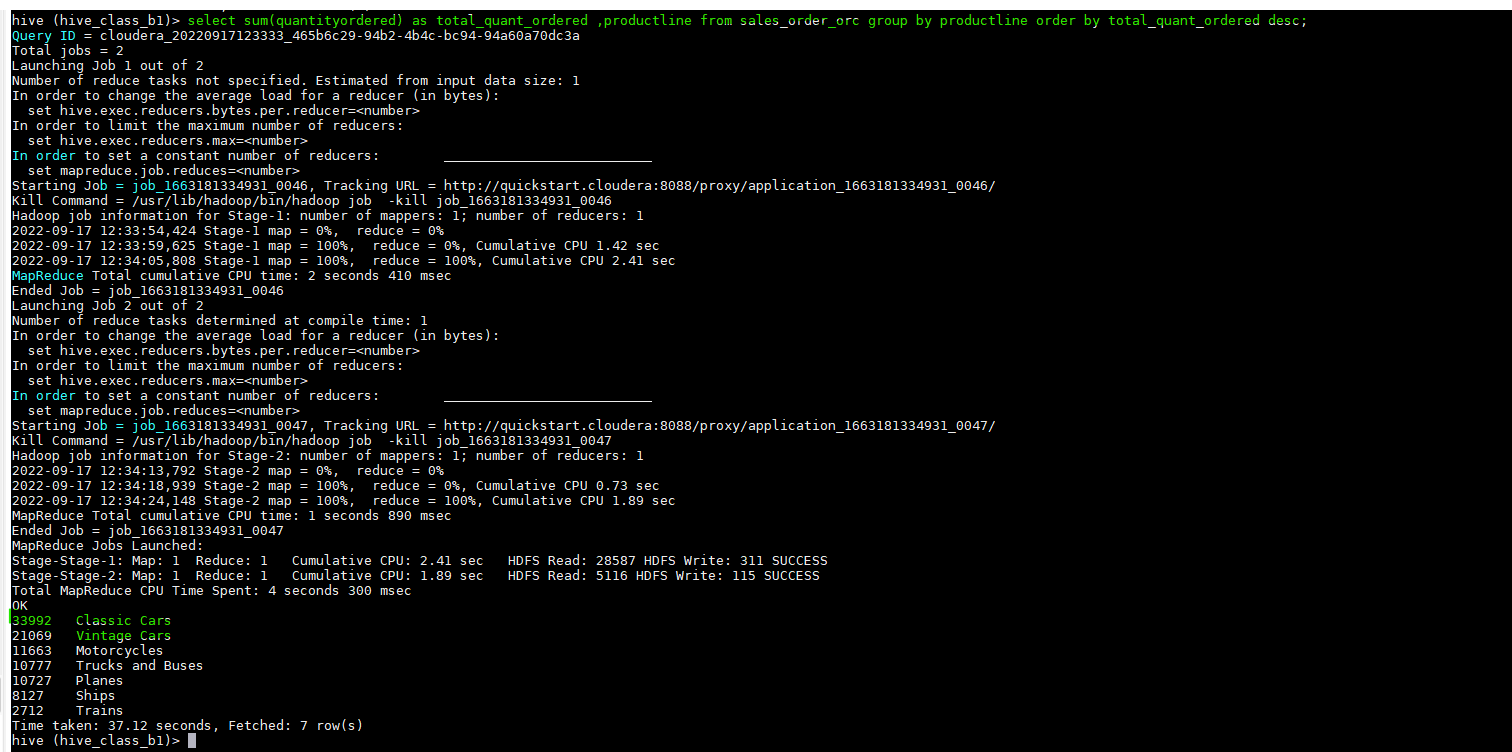
1. Calculate total sales per year



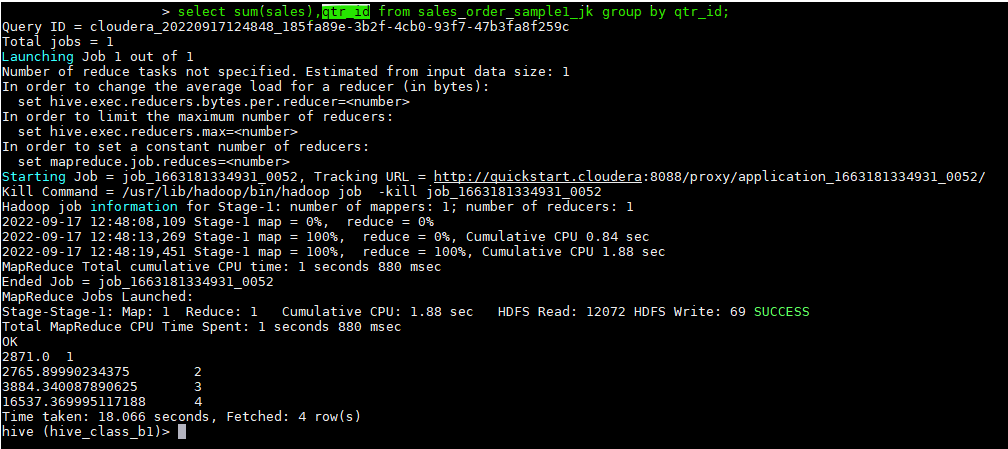
1. Find a product for which maximum orders were placed.



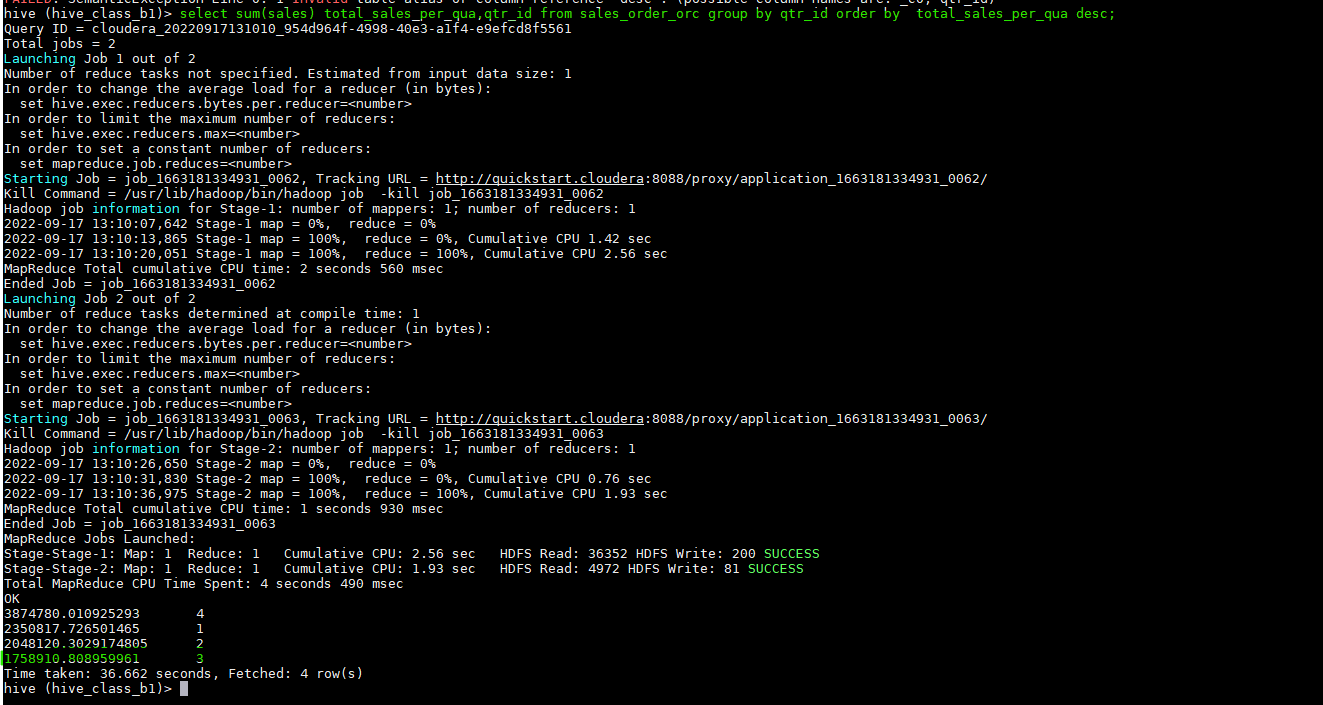
b.1 Finding the product which is ordered highest (my own interest )



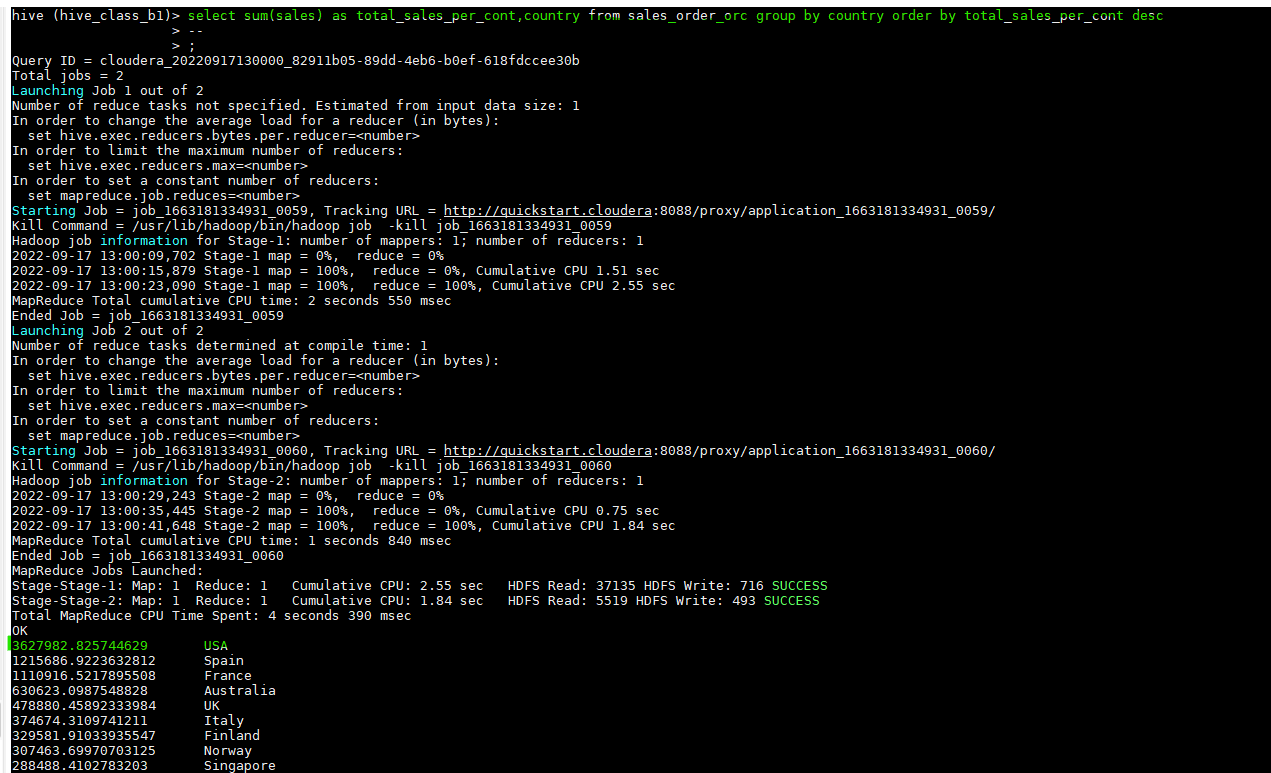
1. Calculate the total sales for each quarter.



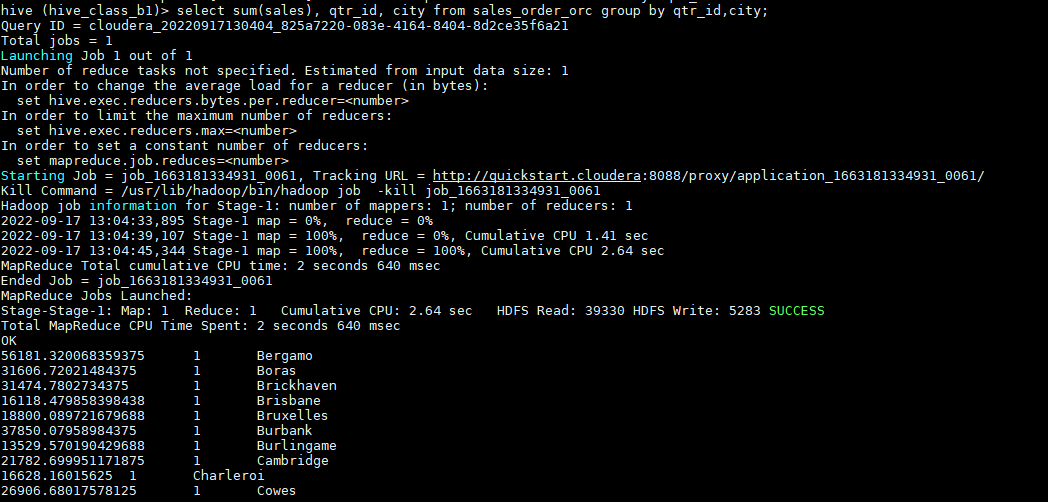
1. In which quarter sales was minimum



1. In which country sales was maximum and in which country sales was minimum



1. Calculate quartelry sales for each city



1. Find a month for each year in which maximum number of quantities were sold

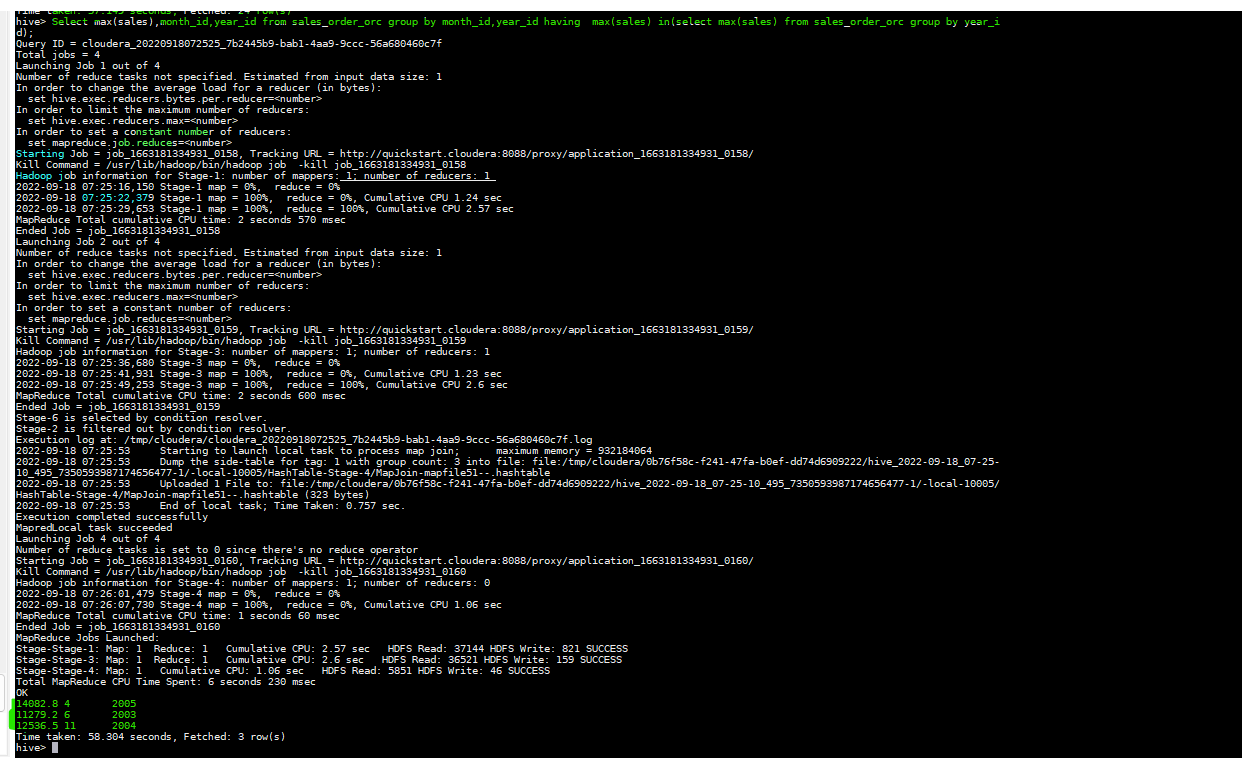
Initial thinking.

Select max(sales) as maximum\_sale,year\_id,month\_id from sales\_order\_orc group by month\_id, year\_id order by year\_id,maximum\_sale desc;

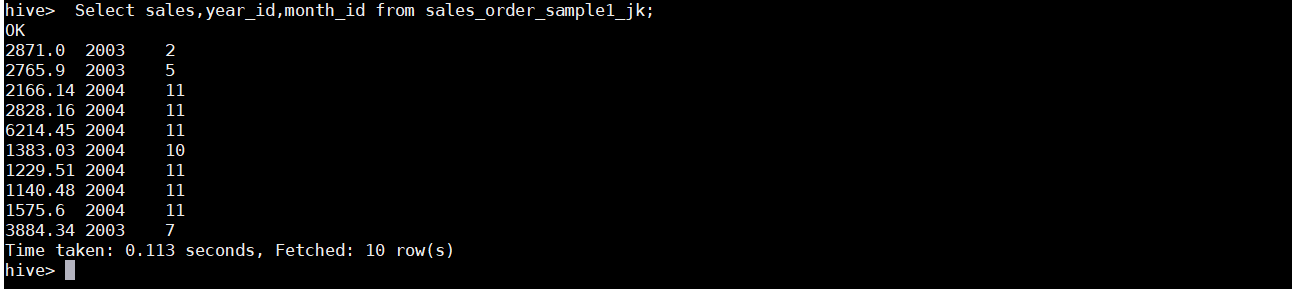
Final Query.

Select max(sales),month\_id,year\_id from sales\_order\_orc group by month\_id,year\_id having max(sales) in(select max(sales) from sales\_order\_orc group by year\_id);

Result:



Sample data for test my query (full details of sales, year, month):



Test result(Finding the month for each year in which maximum sales happened)

