**Q.1 MAP REDUCE**

hadoop jar myjar.jar AllTimeHigh CDAC/NYSE.csv exam/mr

**Q.2 HIVE**

create table customer(cust\_id int, firstname string,

> lastname string, age int, profession string)

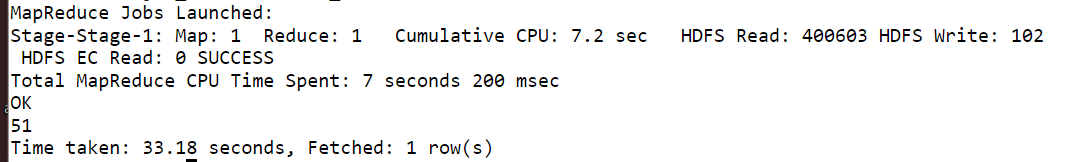
> row format delimited

> fields terminated by ','

> stored as textfile;

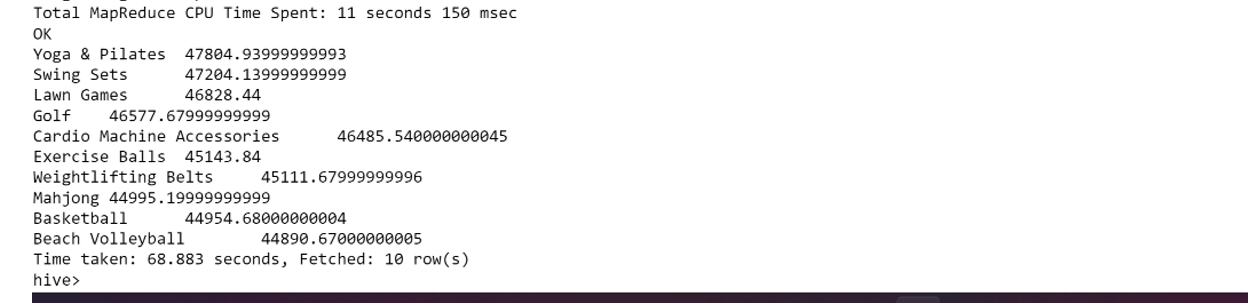
1) Write a program to find the count of customers for each profession.

== select count(distinct(profession)) from customer;



2) Write a program to find the top 10 products sales wise

== select product, sum(amount) total from tax group by product order by total desc limit 10 ;

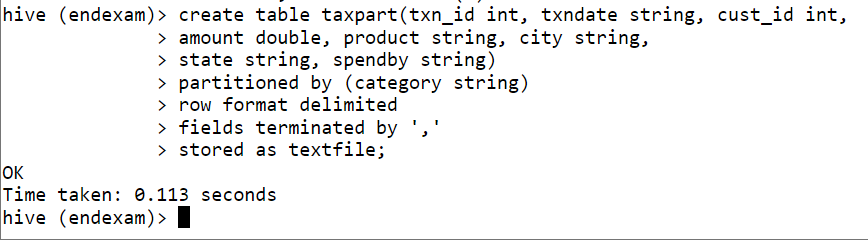


3) Write a program to create partiioned table on category

==

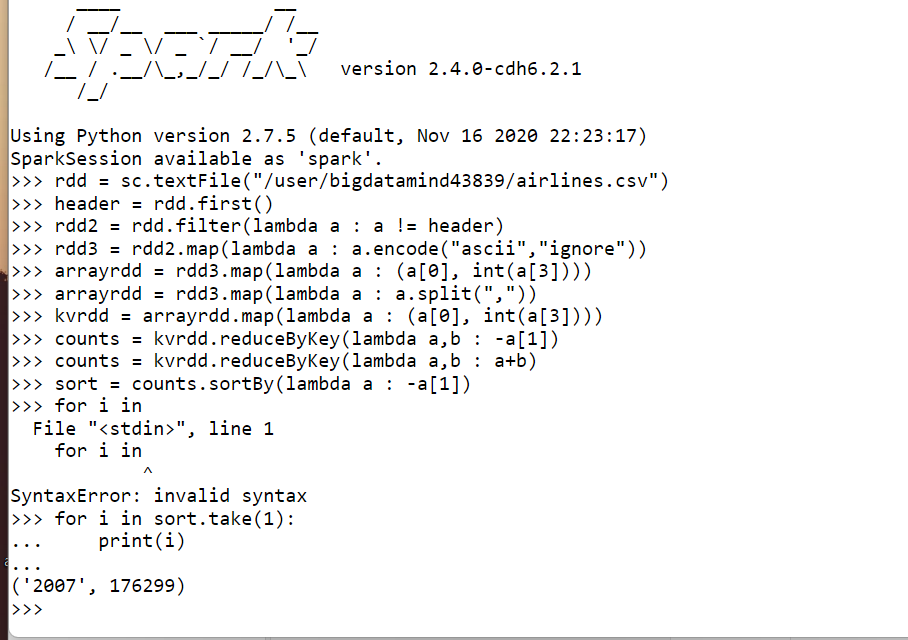
set hive.exec.dynamic.partition.mode=nonstrict;

set hive.exec.dynamic.partition = true;



**Q.3 PYSPARK**

1. What was the highest number of people travelled in which year?

****

1. Identifying the highest revenue generation for which year

=



1. Identifying the highest revenue generation for which year and quarter (Common group)

==

