**EXPERIMENT 5: Hive Queries**

Name: Sunayana M

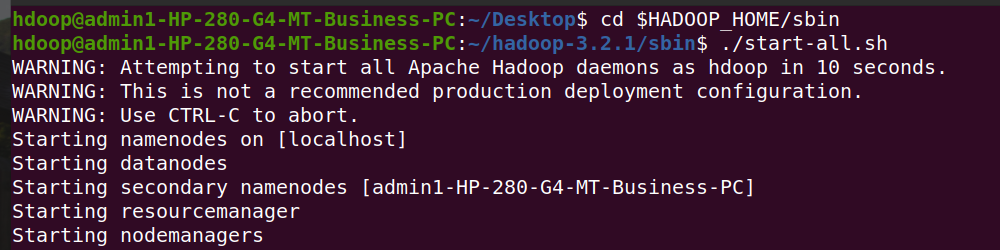
USN: 1NT19IS165

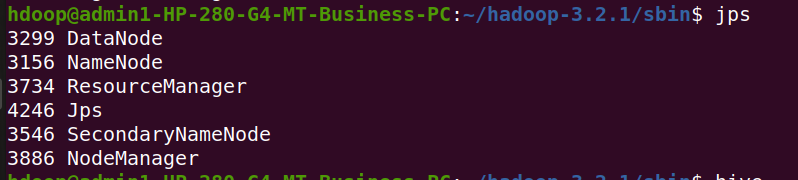
Batch : C2

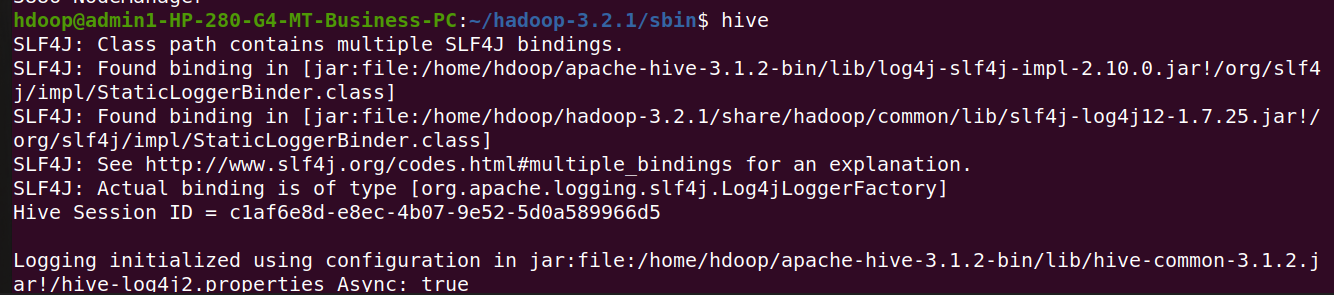
**Opening HIVE shell**

$ cd $HADOOP\_HOME/sbin

$ ./start-all.sh



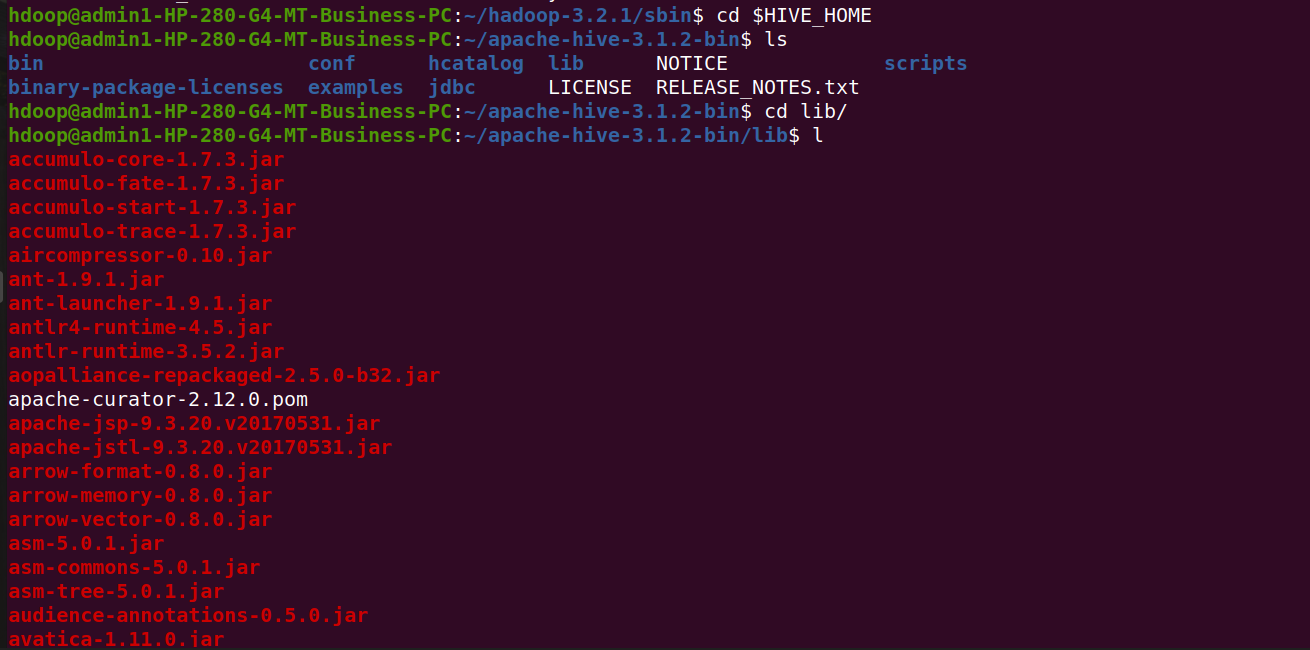
$ jps 

$ hive

$ cd $HIVE\_HOME

$ cd lib/

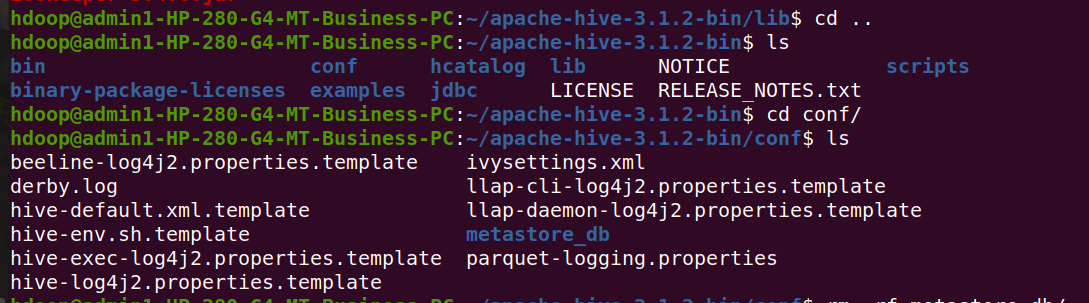
$ l



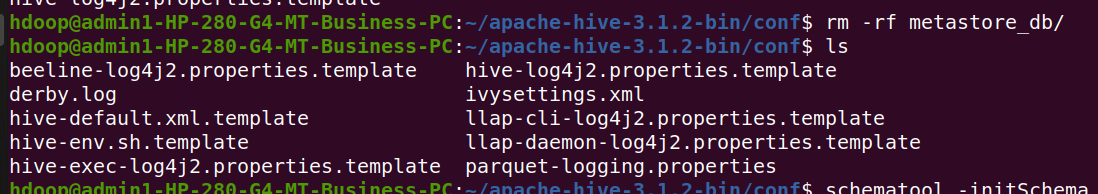
$ cd ..

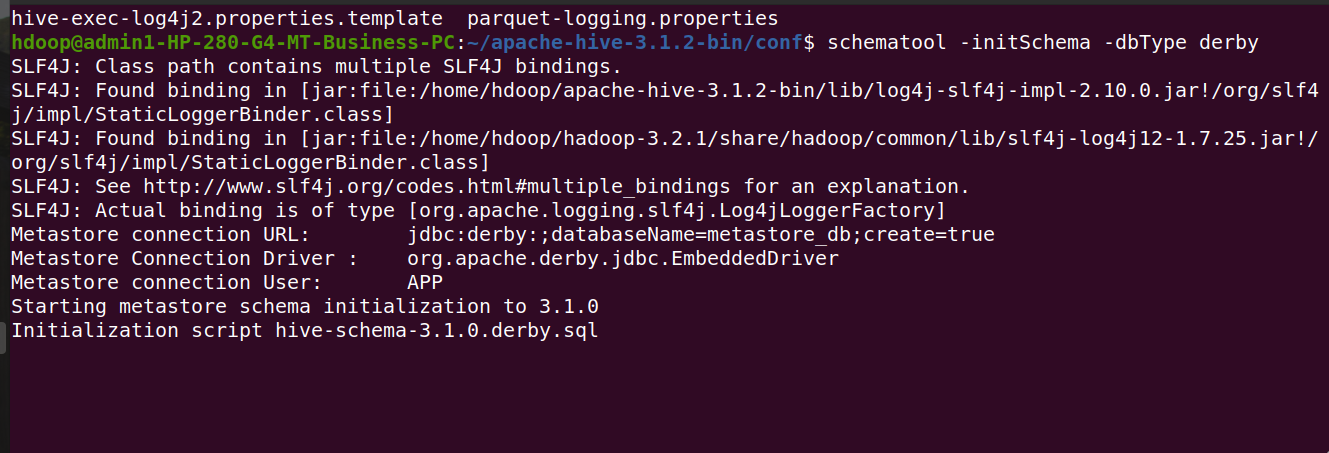
$ ls

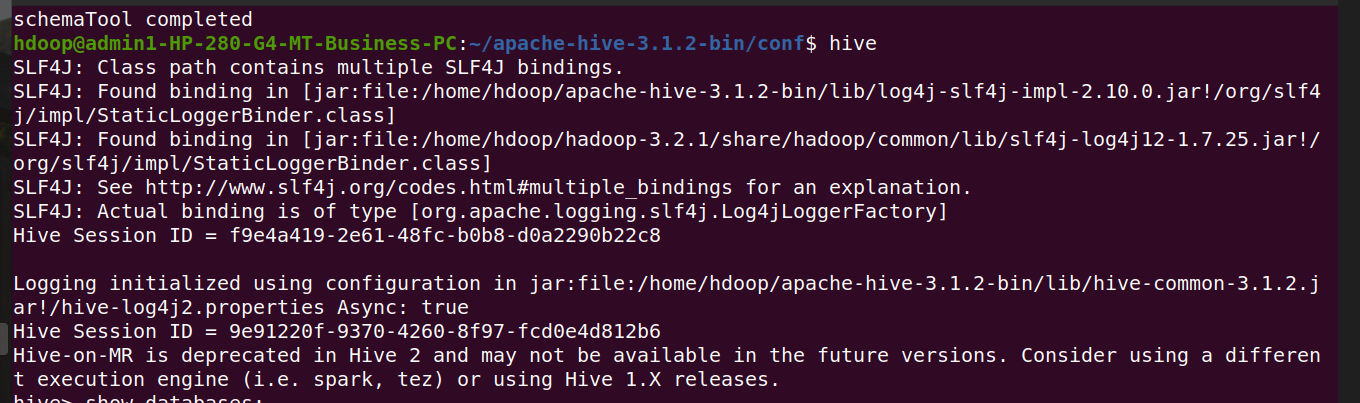
$ cd conf/

$ ls

$ rm -rf metastore\_db/

$ ls

$ schematool -initSchema -dbType derby

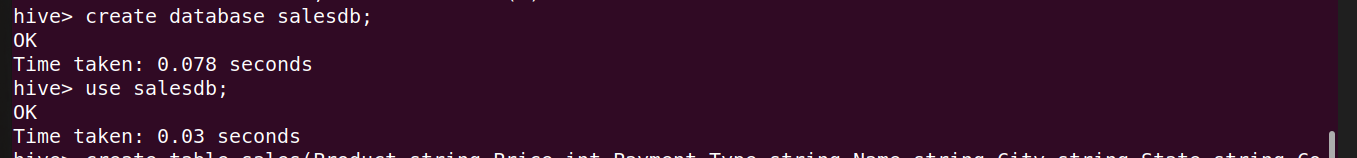
$ hive

hive> show databases;



**Create a database called salesdb.**

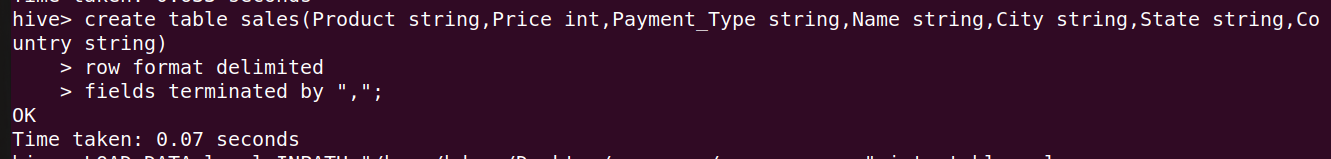
hive> create database salesdb;

hive> use salesdb; 

**Create a table under that with the heading sales.**

hive> create table sales(Product string,Price int, Payment\_type string, Name string, City string,State string,Country string)

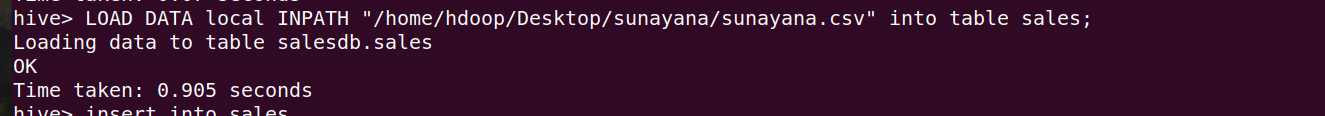
> row format delimited

> fields terminated by ","; 

**Describe the sales table.**

hive> desc sales; 

**Import the dataset from the sales.csv file.**

hive> LOAD DATA local INPATH "/home/hdoop/Desktop/1nt19is166\_supreetha/supreetha.csv" into table sales; 

**Insert 5 records using Insert command in HiveQL.**

hive> insert into sales

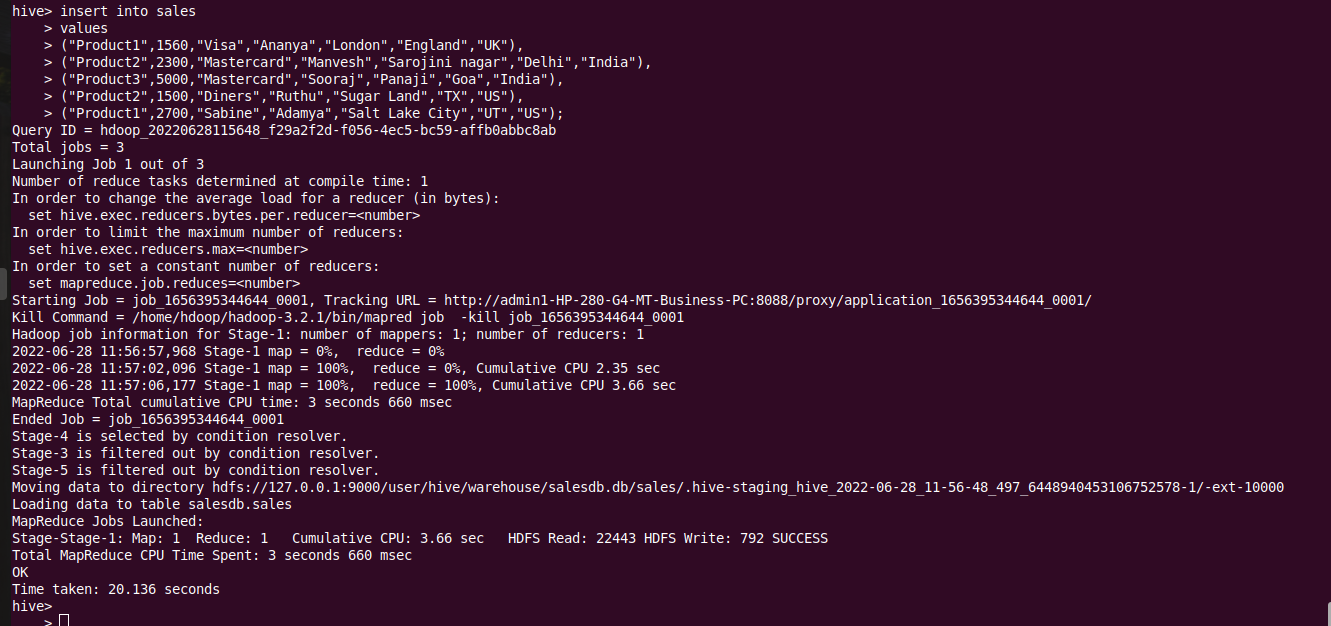
> values

> ("Product1",1560,"Visa","Ananya","London","England","UK"),

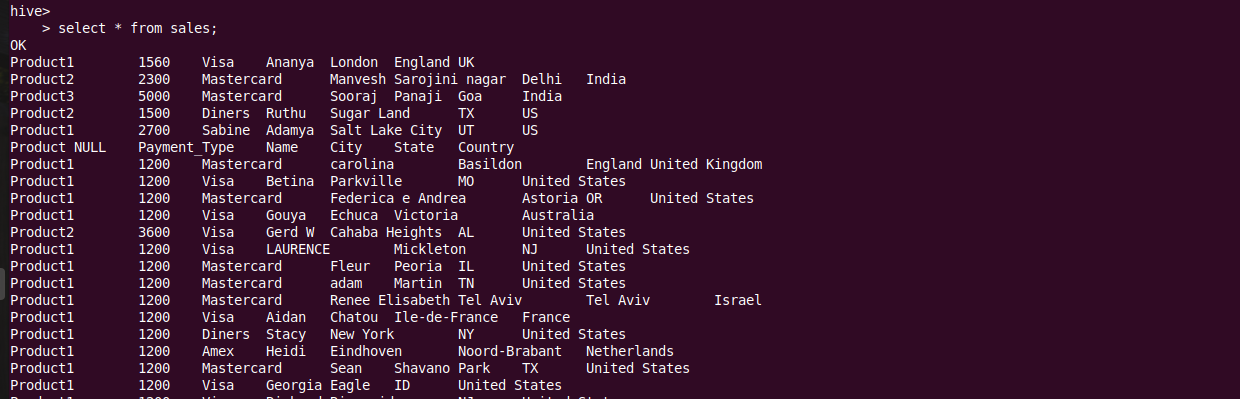
> ("Product2",2300,"Mastercard","Manvesh","Sarojini nagar","Delhi","India"),

> ("Product3",5000,"Mastercard","Sooraj","Panaji","Goa","India"),

> ("Product2",5000," Diners ","Ruthu","Sugar Land","TX","US"),

> ("Product1",2700,"Sabine","Adamya","Salt Lake City","UT","US"); 

**Display the tuples of the table.**

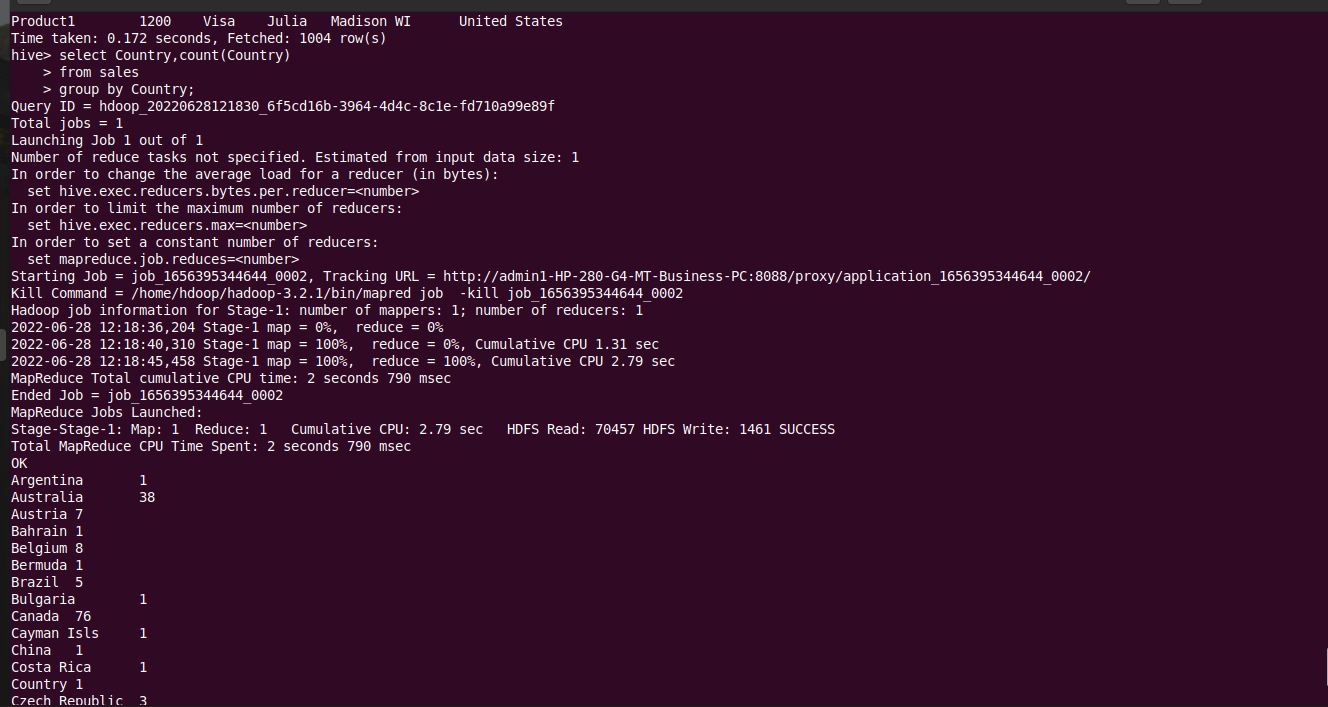
hive> select \* from sales; 

**Count the number of sales done by each country.**

hive> select Country, count(Country)

> from sales

> group by Country;

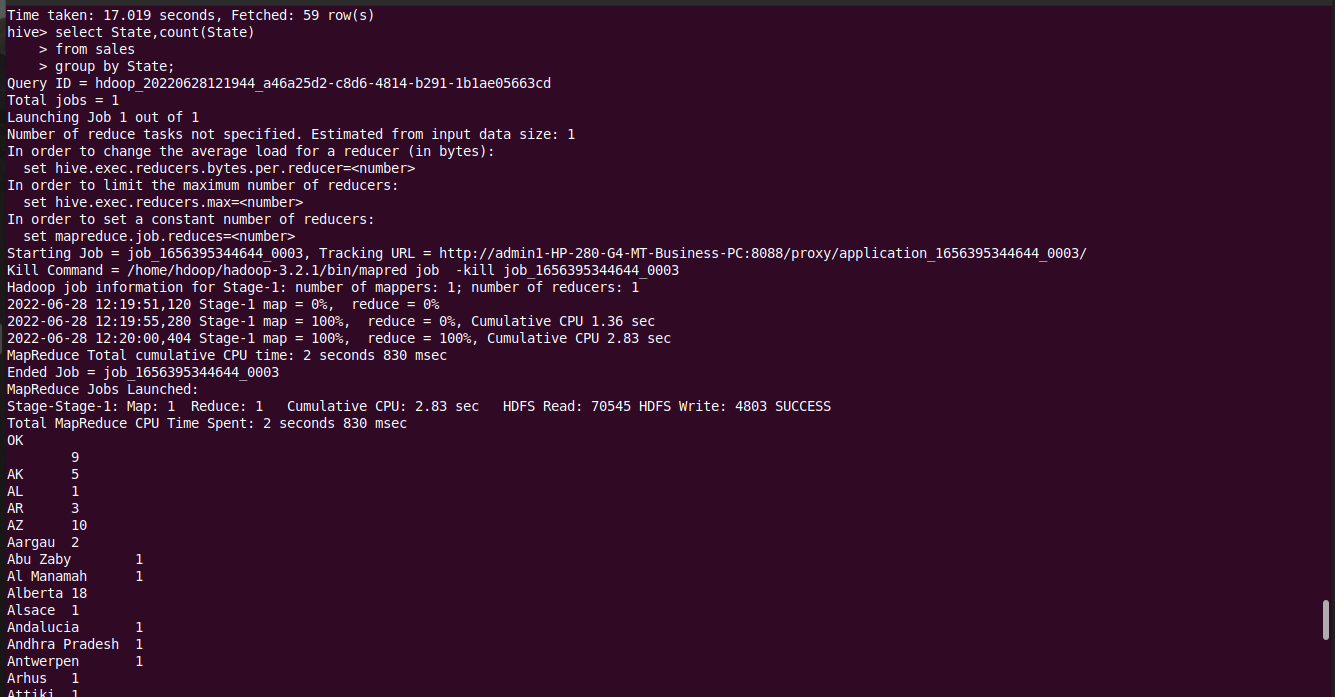


**Count the number of sales done by each state.**

hive> select State, count(State)

> from sales

> group by State;

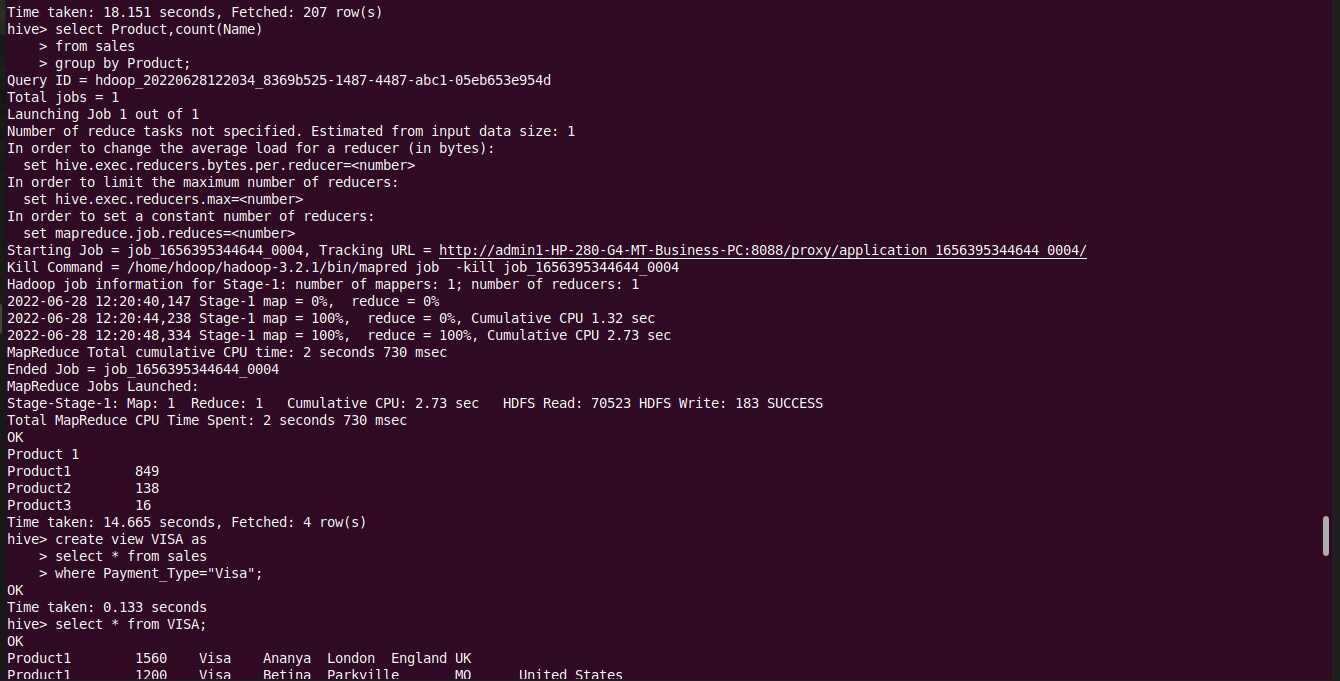


**Display (Product, name) grouped by product.**

hive> select Product, count(Name)

> from sales

> group by Product;

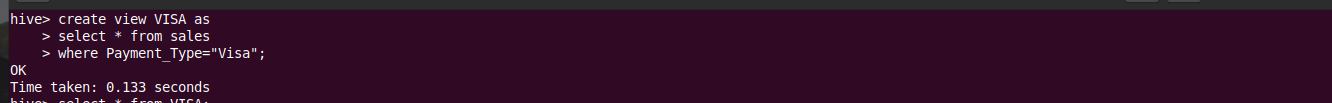


**Create separate views for VISA and Mastercard.**

For Visa:

hive> create view VISA as

> select \* from sales

> where payment\_type = "Visa"; 

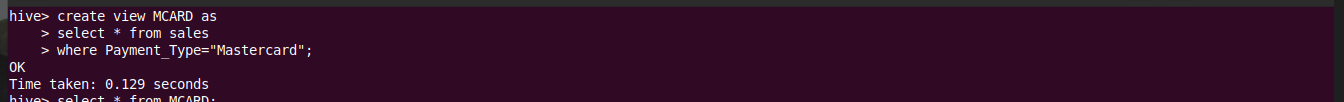
To see the View of Visa:

hive> select \* from VISA; 

For Mastercard:

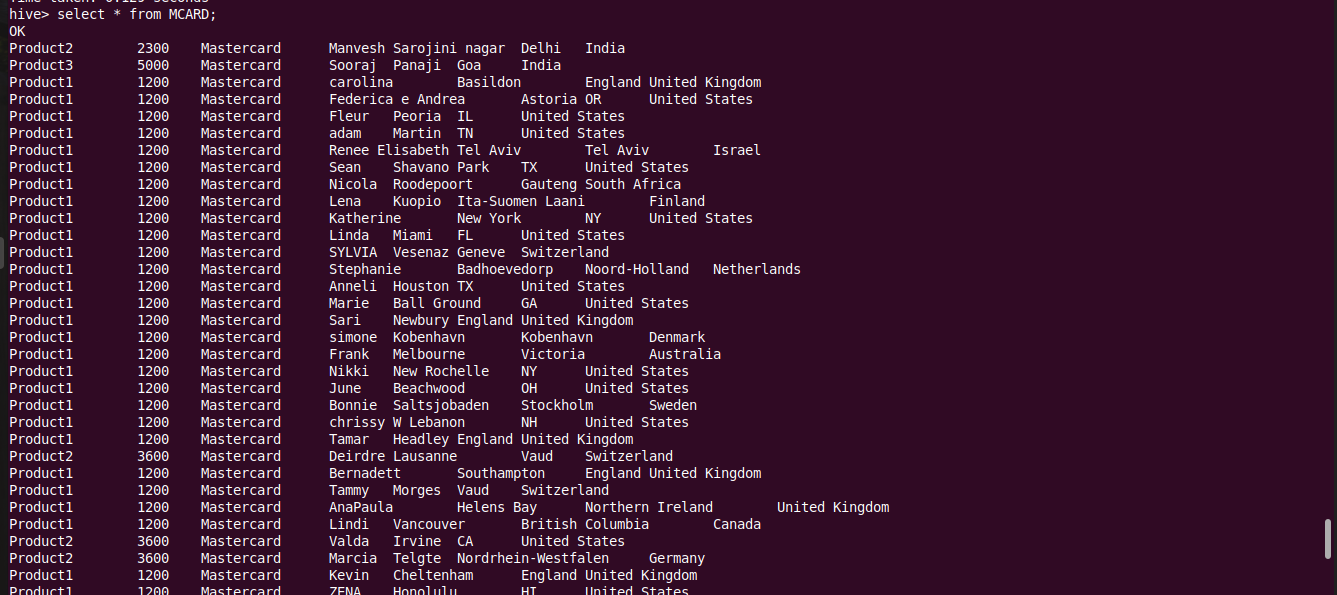
hive> create view MCARD as

> select \* from sales

> where Payment\_Type = "Mastercard"; 

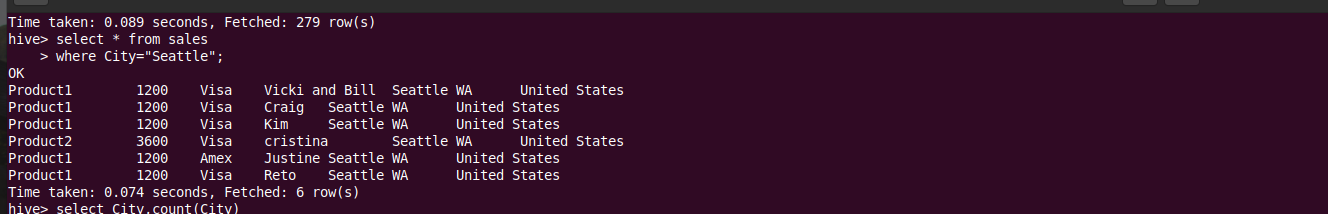
To see the View of Mastercard:

hive> select \* from MCARD;



**Show all the transactions done in Seattle.**

hive> select \* from sales

> where City="Seattle"; 

**Find the max number of transactions done within the state of Ontario.**

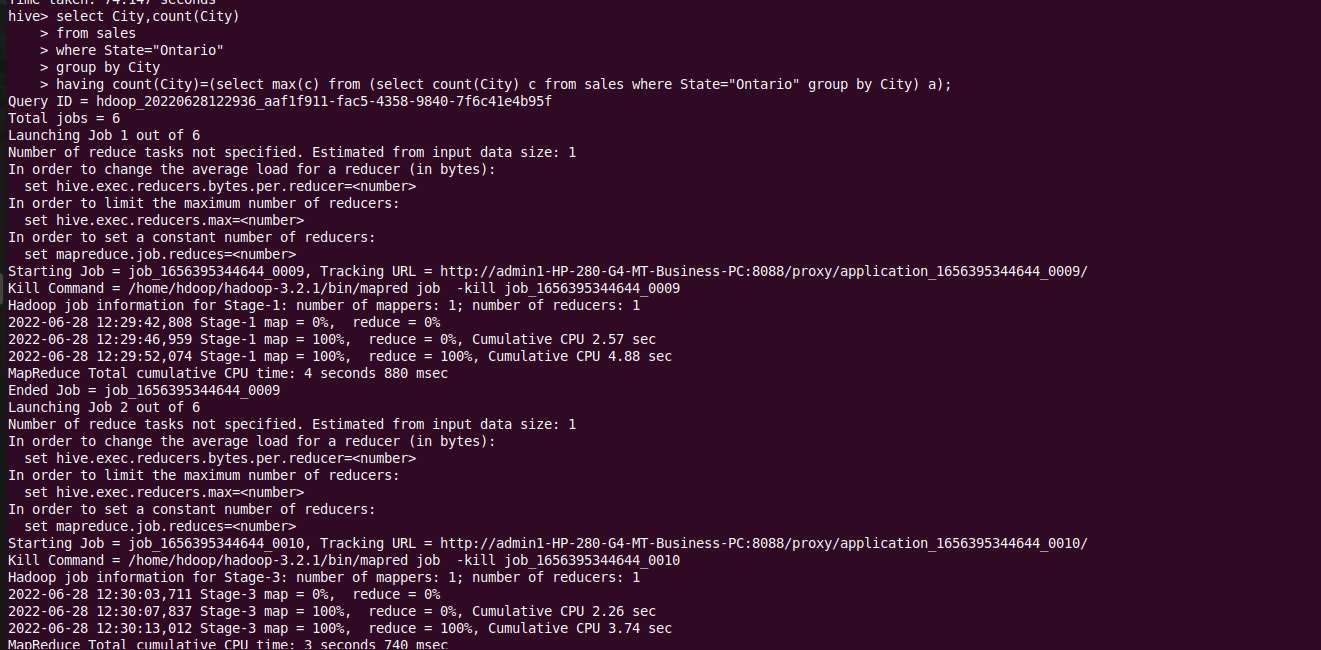
hive> select City, count(City)

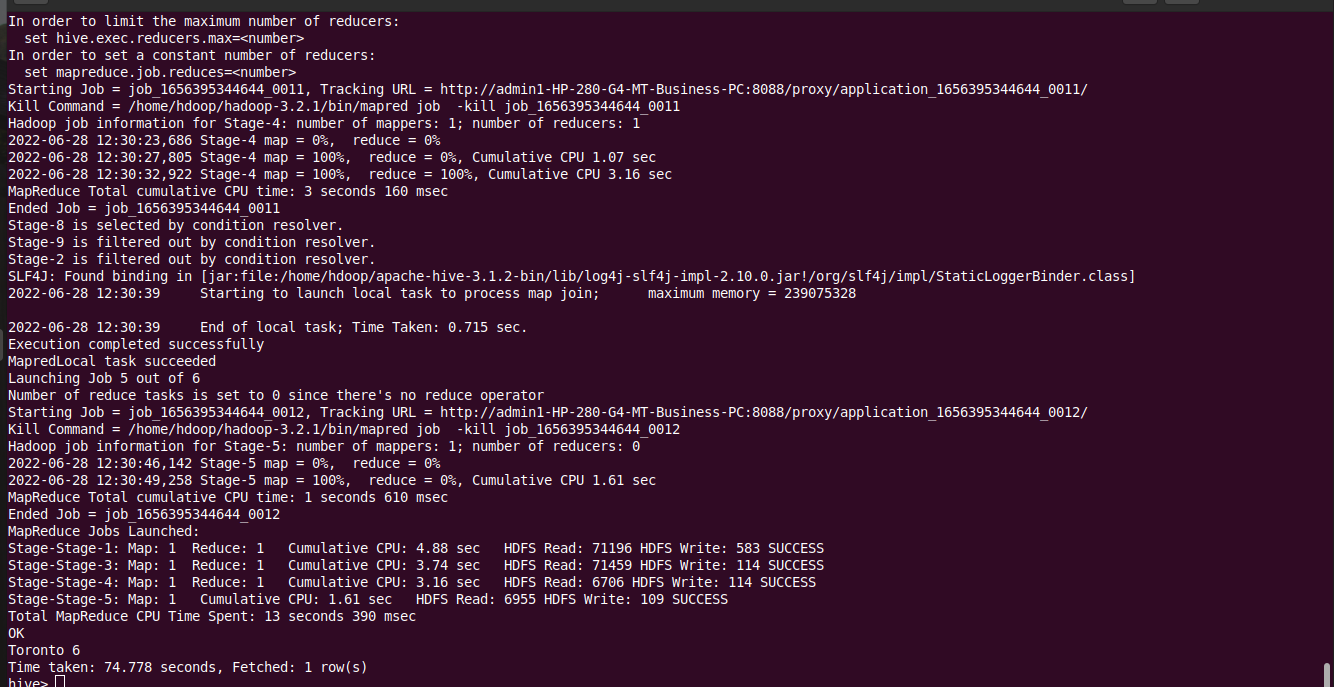
> from sales

> where State="Ontario"

> group by city

> having count(City)=(select max(c) from (select count(City) c from sales where State="Ontario" group by City) a );



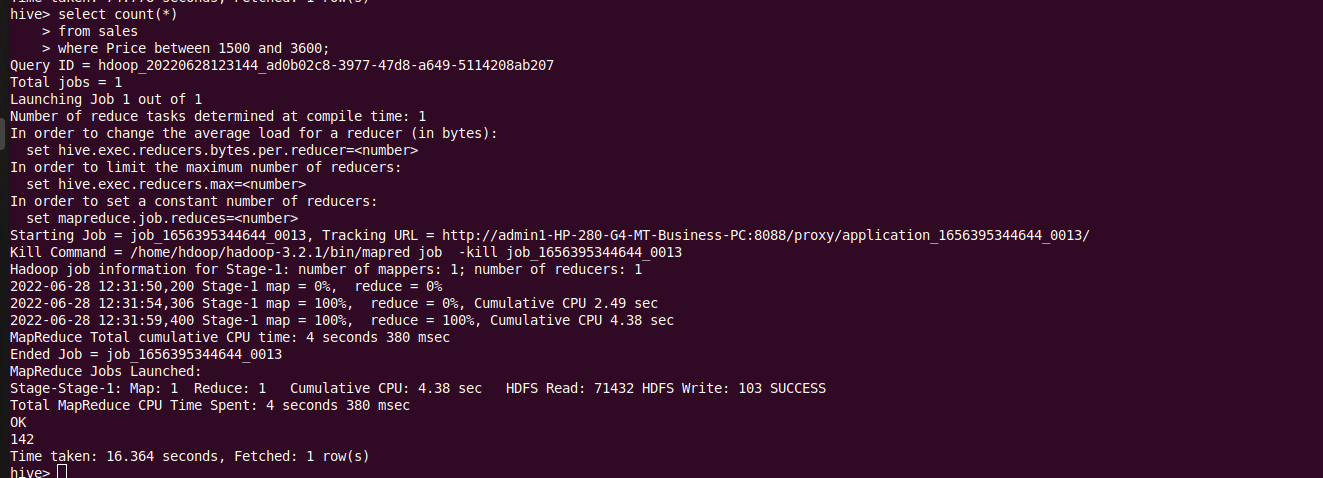


**Find the number of transactions whose price is in between 1500-3600.**

hive> select count(\*)

> from sales

> where price between 1500 and 3600;



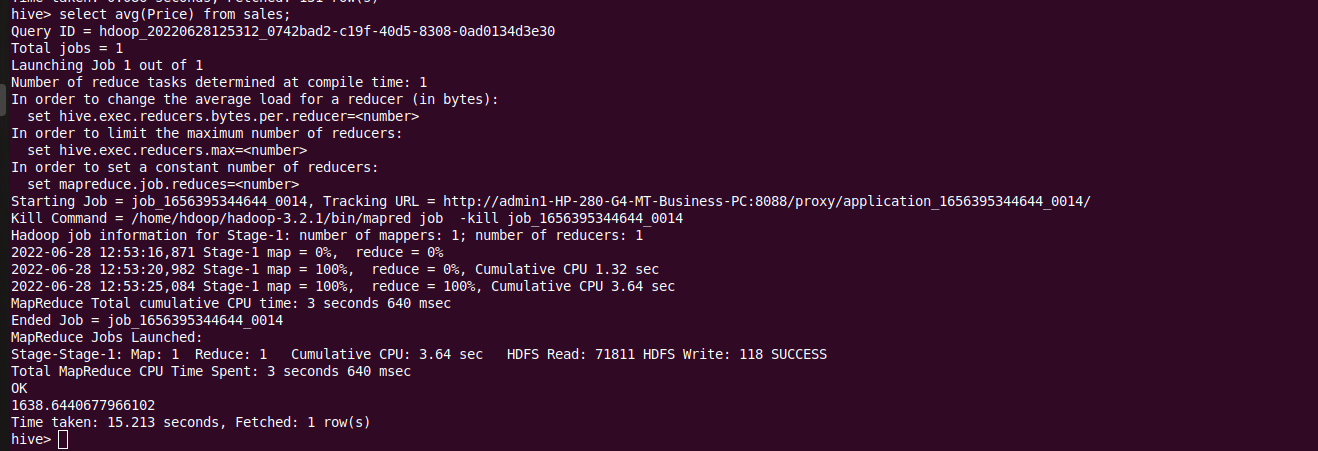
**List all the transactions done in the United States using Mastercard.**

hive> select \* from sales

> where Country="United States" and Payment\_Type="Mastercard";

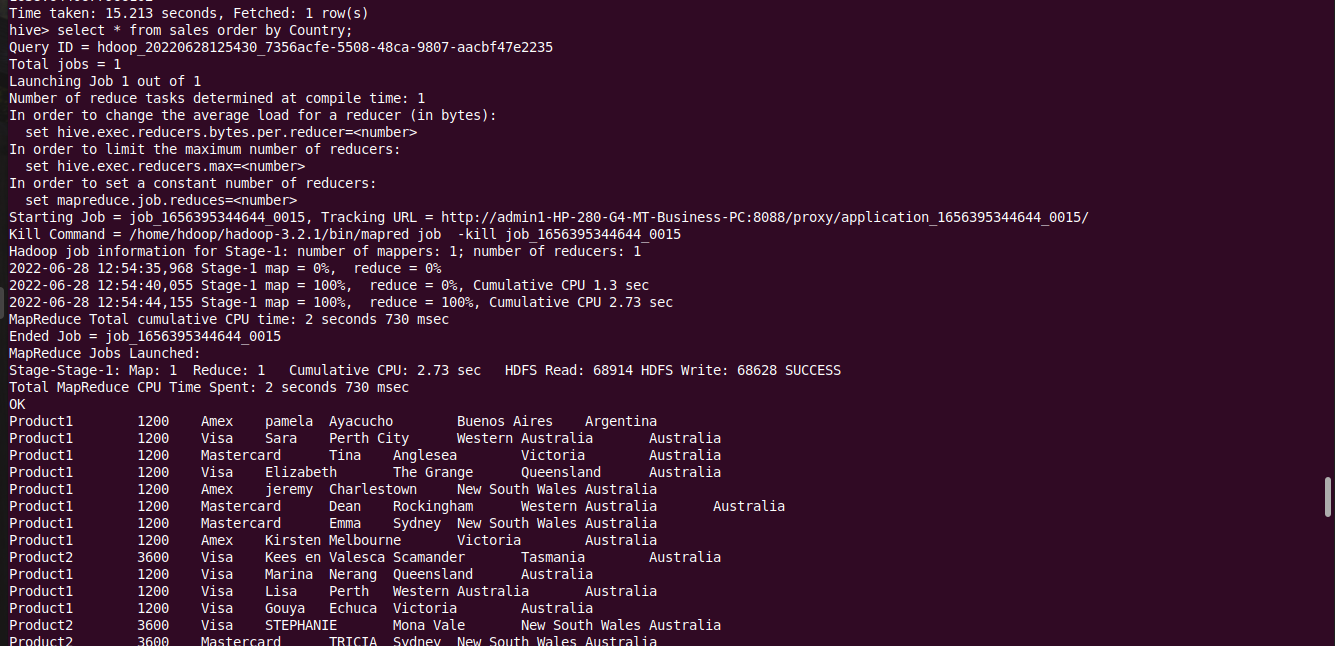


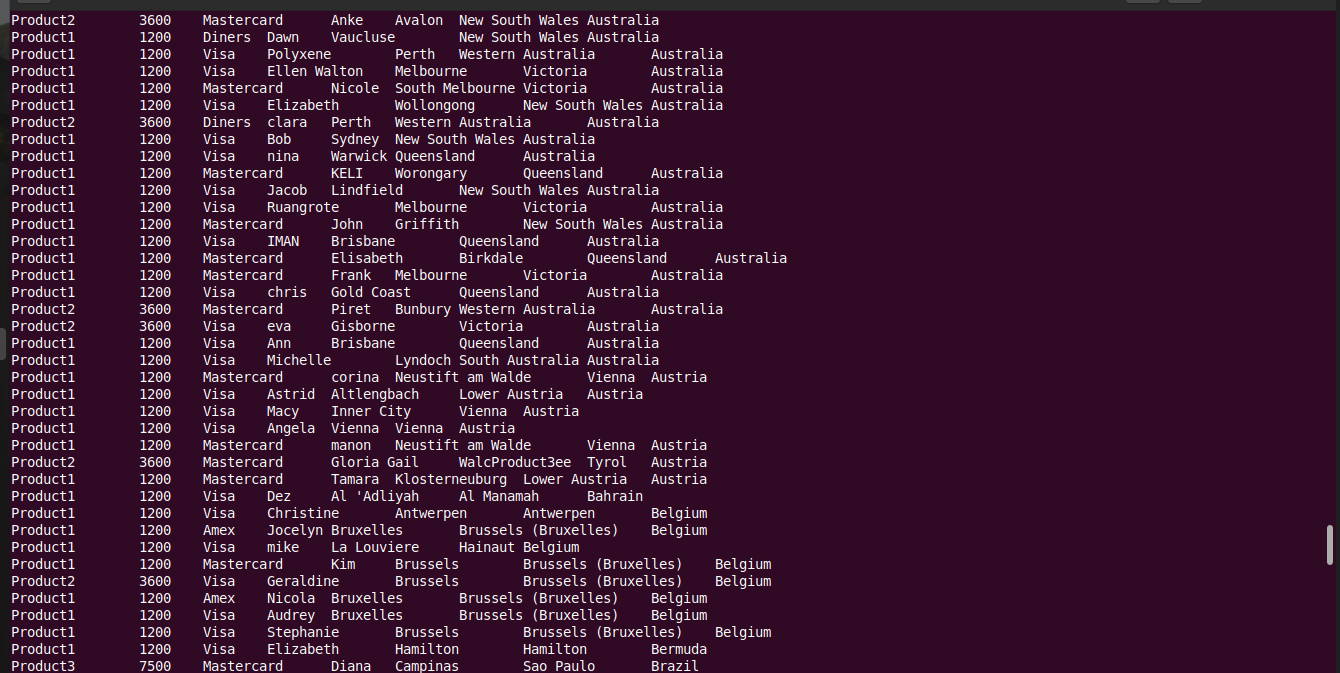
**Find the average price of all transactions.**

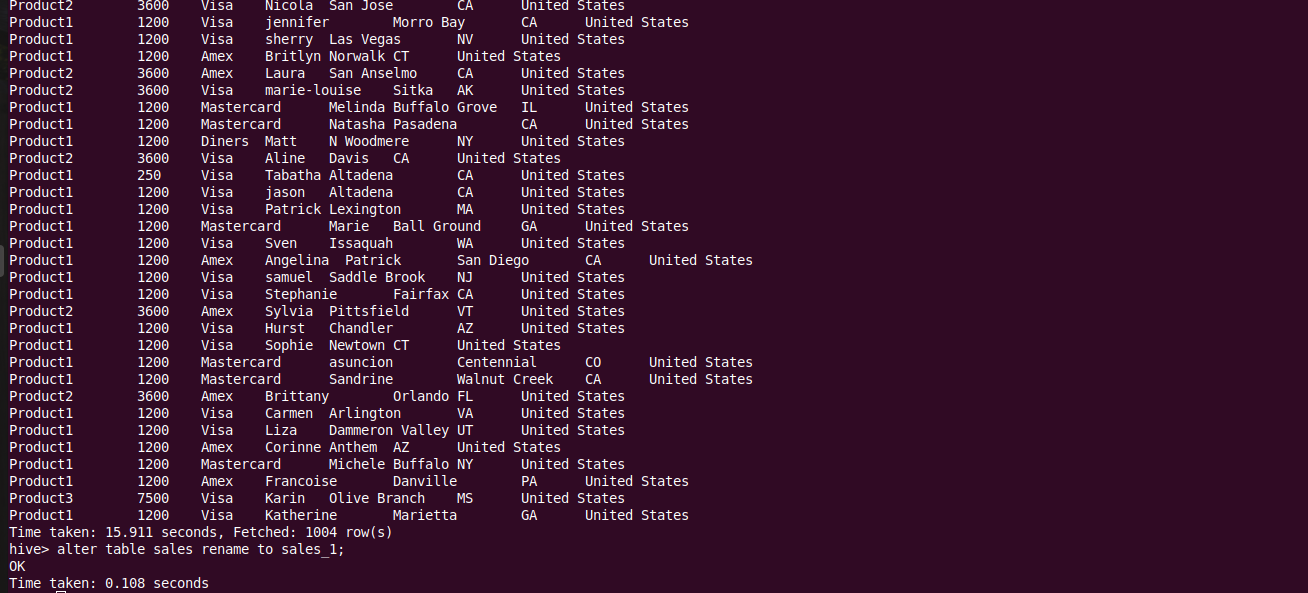
hive> select avg(Price) from sales; 

**Show the transactions in alphabetical order of their countries.**

hive> select \* from sales order by country;







**Change the table name.**

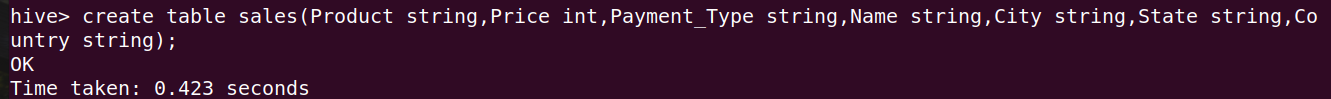
hive> alter table sales rename to sales\_1;



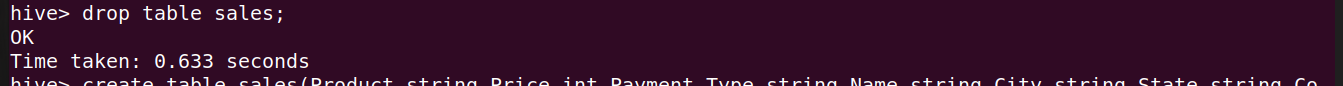
**Change any one column name in the table.**

hive> alter table sales\_1 change Name cName string; 

**Creating a demo table.**

hive> create table sales(Product string,Price int, Payment\_type string, Name string, City string,State string,Country string);

**Delete the demo table.**

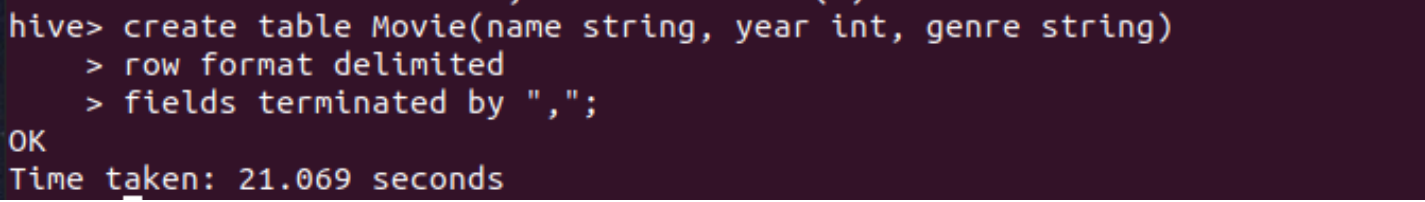
hive> drop table sales; 

**Create a table called Movie.**

hive> create table Movie(name string, year int, genre string)

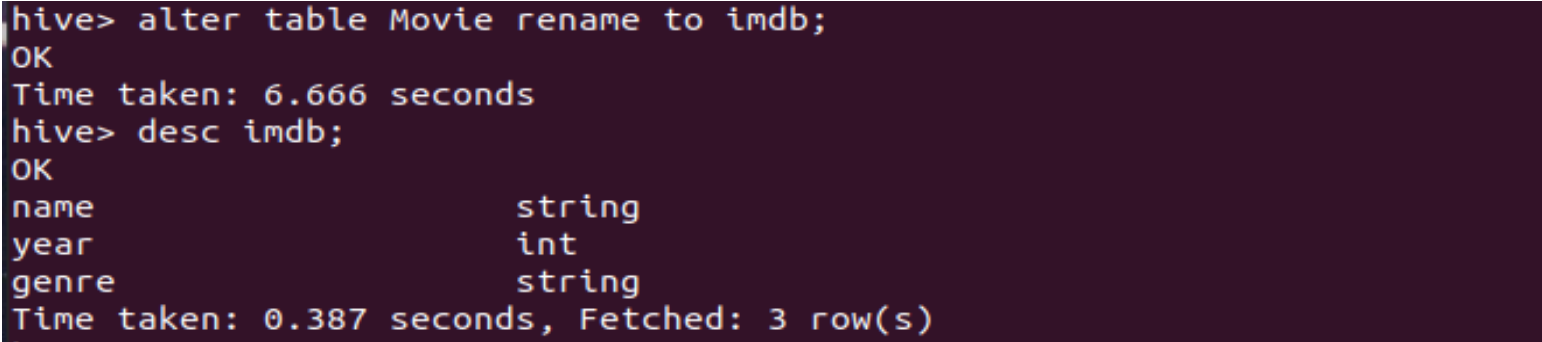
> row format delimited

> fields terminated by ",";



**Rename the Table.**

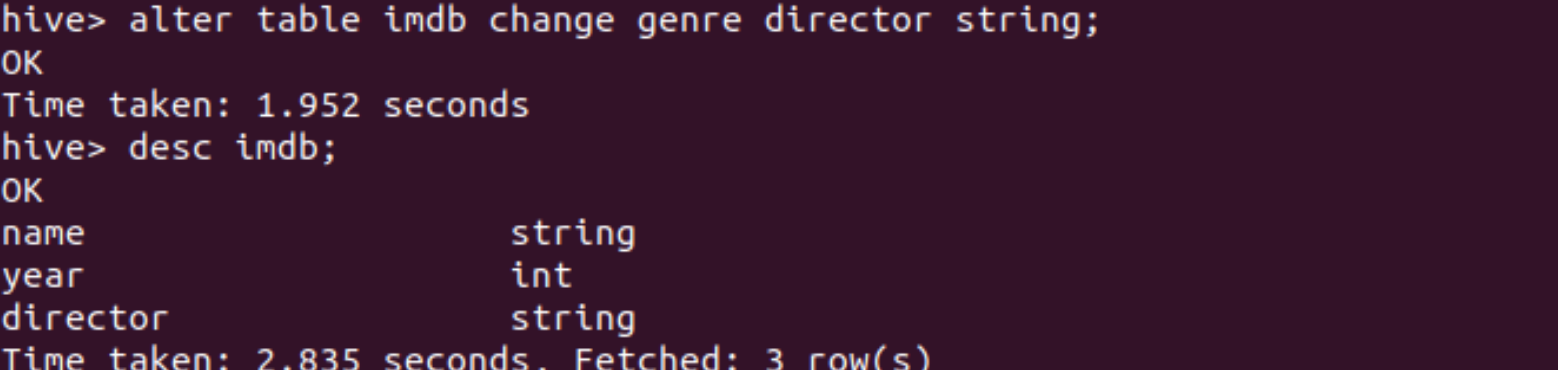
hive> alter table Movie rename to imdb;

hive> desc imdb; 

**Rename the Column.**

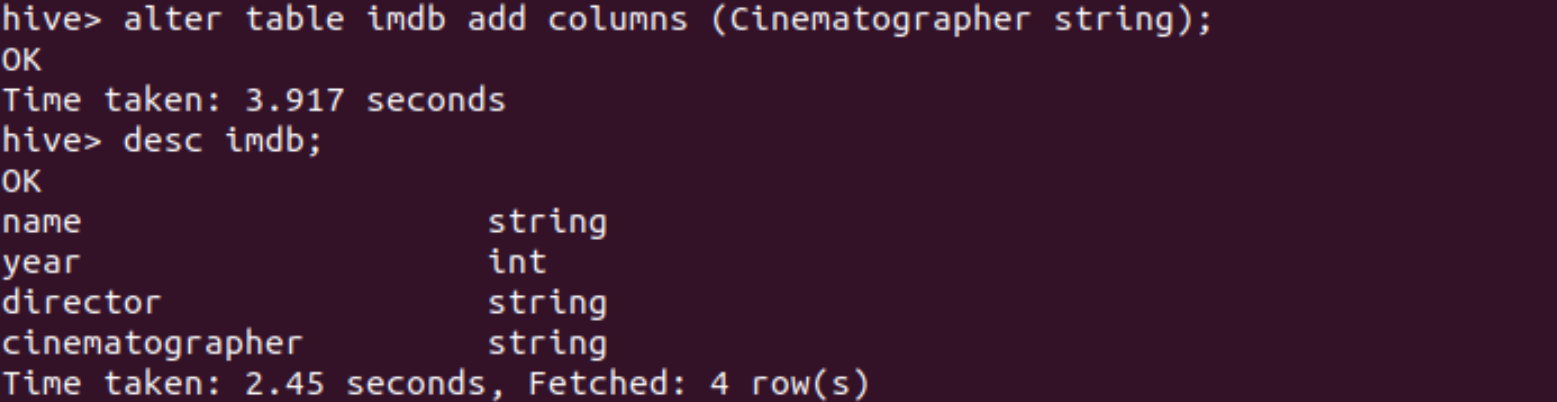
hive> alter table imdb change genre director string;

hive> desc imdb;



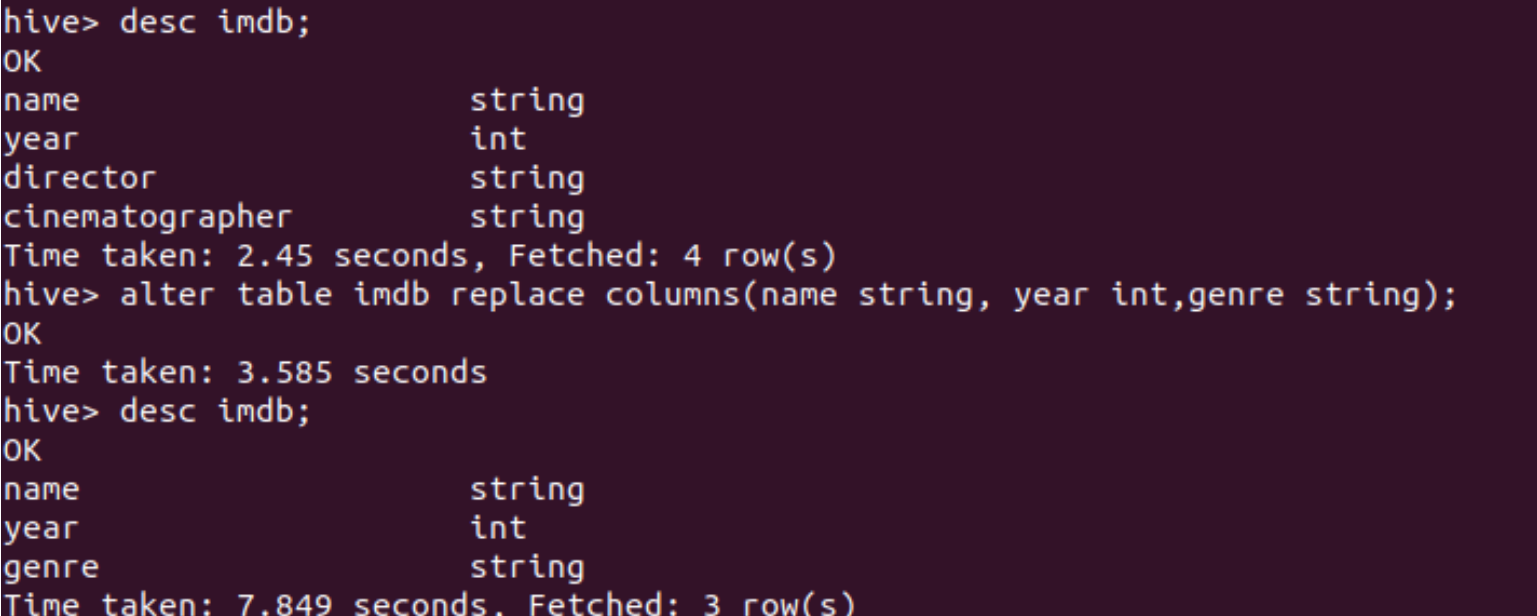
**Add a Column.**

hive> alter table imdb add columns (Cinematographer string);

hive> desc imdb; 

**Replace a Column.**

hive> alter table imdb replace columns(name string, year int,genre string);

hive> desc imdb; 

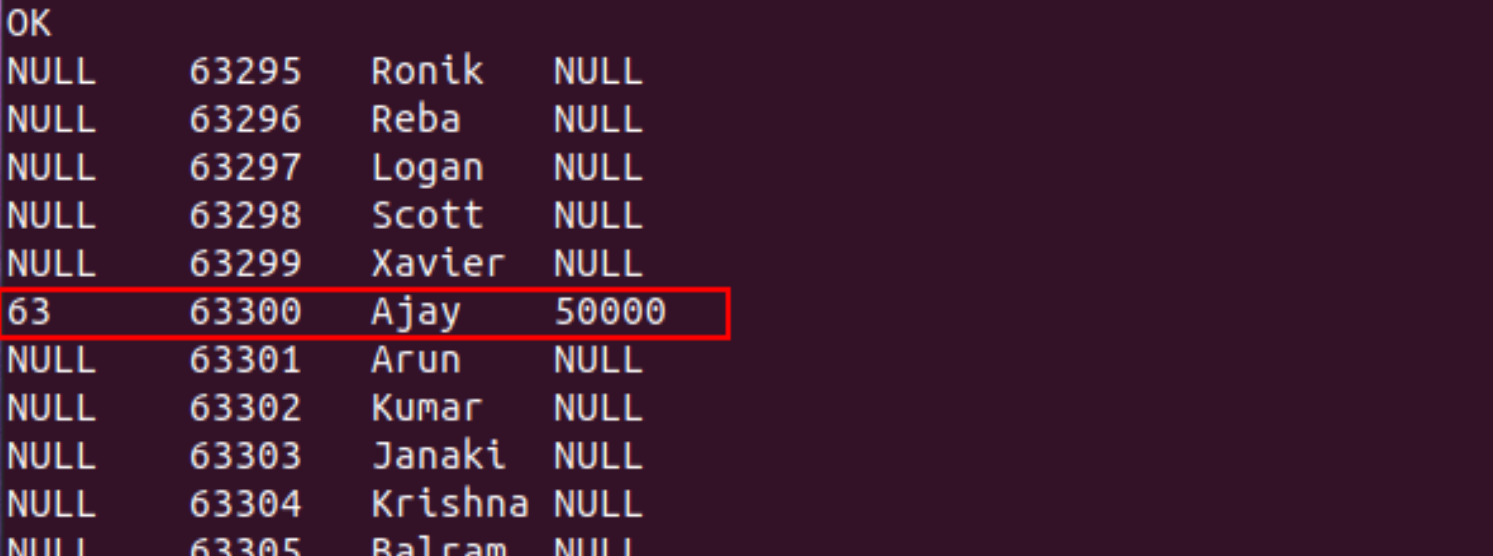
**OUTER JOIN**

hive> select o.orderid, e.ssn, e.name, o.amount

> from emp e

> FULL OUTER JOIN orders o

> ON (e.ssn = o.custssn);

Output:

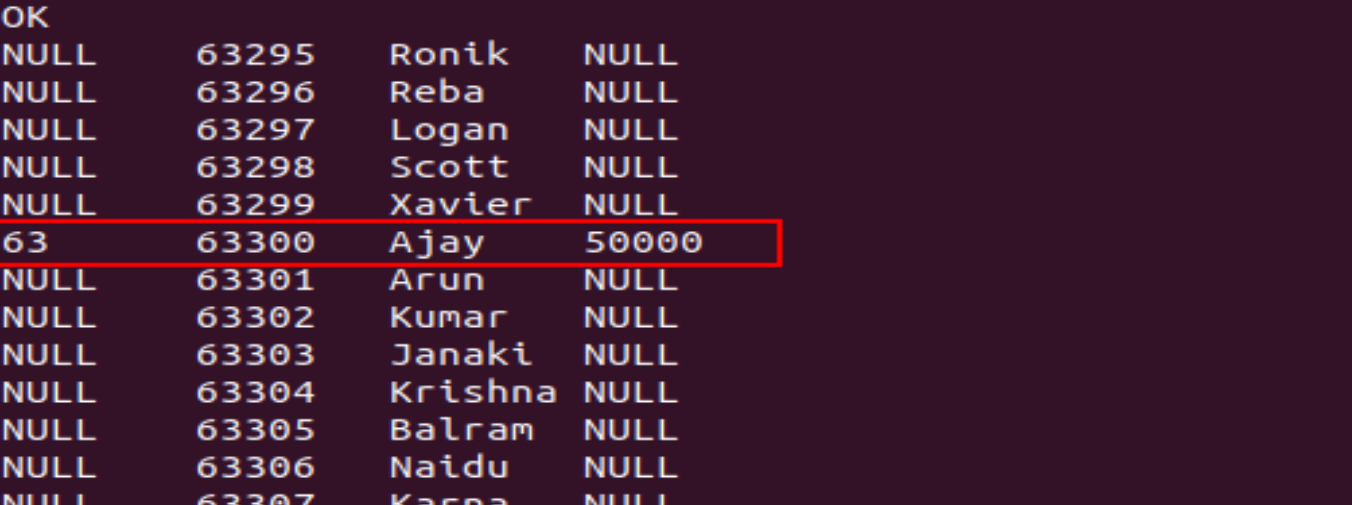
**OUTPUT LEFT OUTER JOIN**

hive> select o.orderid, e.ssn, e.name, o.amount

> from emp e

> LEFT OUTER JOIN orders o

> ON (e.ssn = o.custssn);

Output:

**OUTPUT RIGHT OUTER JOIN**

hive> select o.orderid, e.ssn, e.name, o.amount

> from emp e

> RIGHT OUTER JOIN orders o

> ON (e.ssn = o.custssn);

Output: