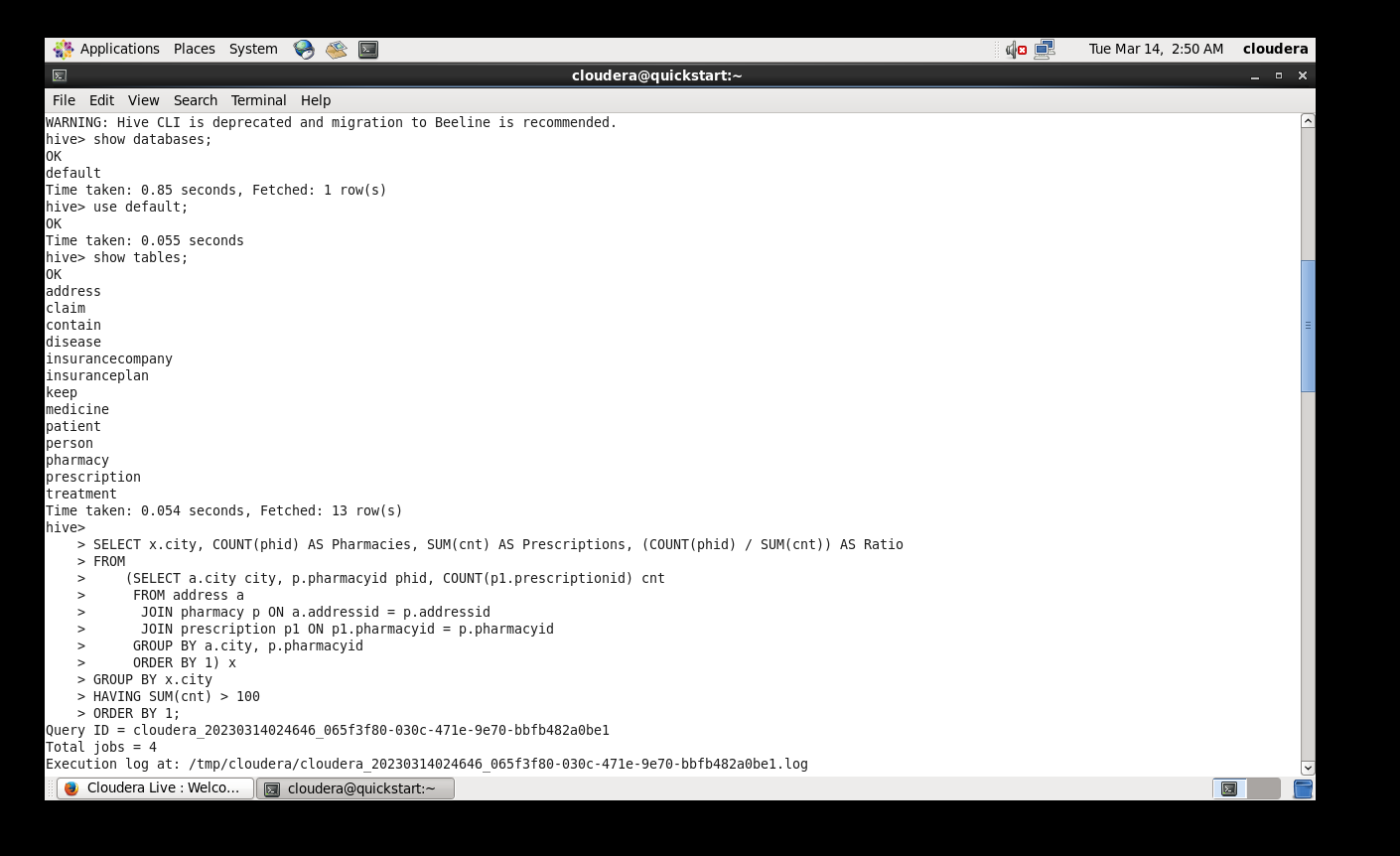
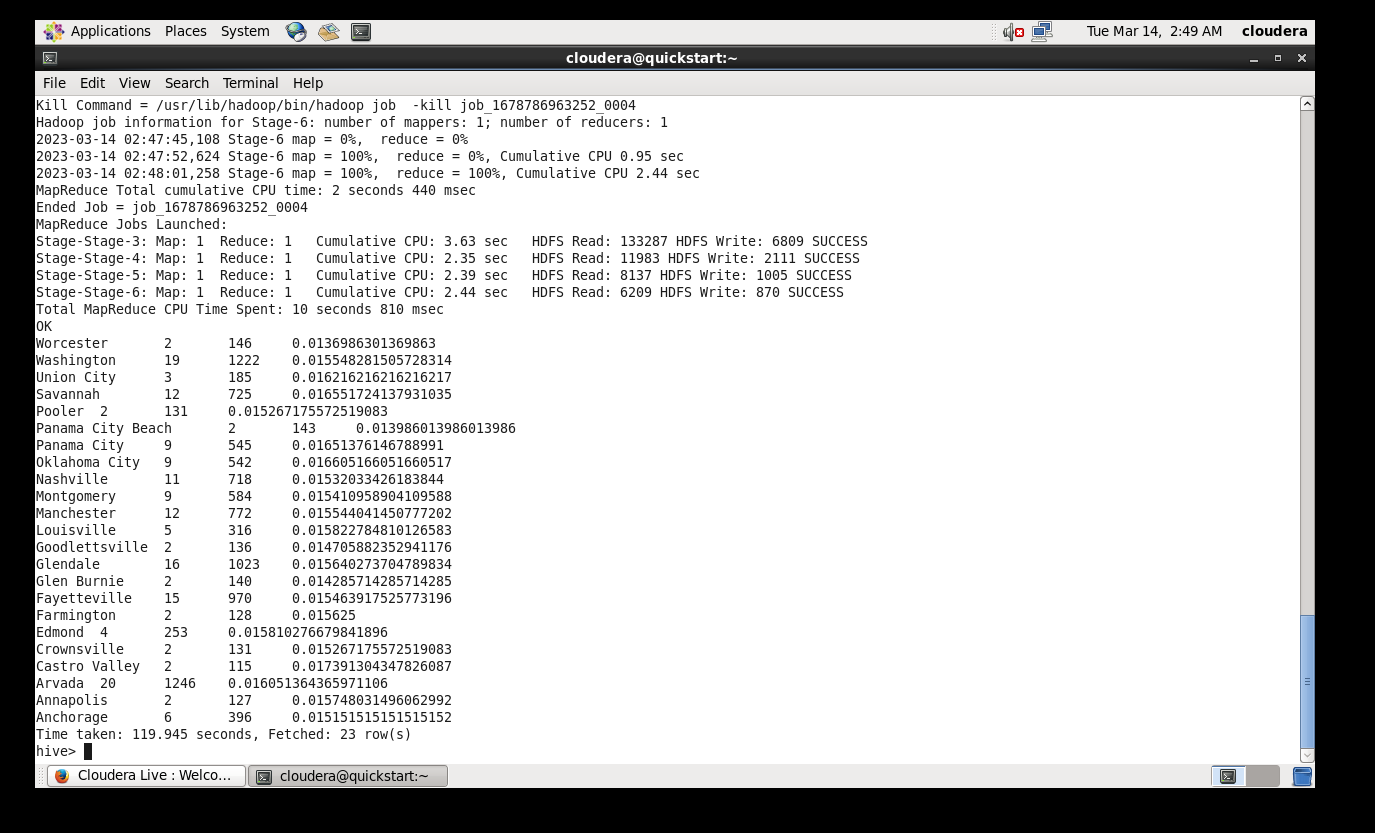
**HIVE HEALTH CARE SYSTEM**

**Problem Statement 1:** A company needs to set up 3 new pharmacies, they have come up with an idea that the pharmacy can be set up in cities where the pharmacy-to-prescription ratio is the lowest and the number of prescriptions should exceed 100. Assist the company to identify those cities where the pharmacy can be set up.





mysql>

create table p\_1(city varchar(20),Pharmacies int,Prescriptions int,Ratio double);

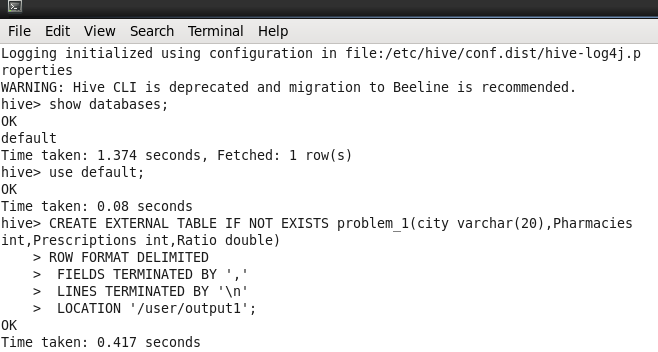
CREATE EXTERNAL TABLE IF NOT EXISTS problem\_1(city varchar(20),Pharmacies int,Prescriptions int,Ratio double)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

LOCATION '/user/output1';



hive>

INSERT OVERWRITE TABLE problem\_1 SELECT x.city, COUNT(phid) AS Pharmacies, SUM(cnt) AS Prescriptions, (COUNT(phid) / SUM(cnt)) AS Ratio

FROM

(SELECT a.city city, p.pharmacyid phid, COUNT(p1.prescriptionid) cnt

FROM address a

JOIN pharmacy p ON a.addressid = p.addressid

JOIN prescription p1 ON p1.pharmacyid = p.pharmacyid

GROUP BY a.city, p.pharmacyid

ORDER BY 1) x

GROUP BY x.city

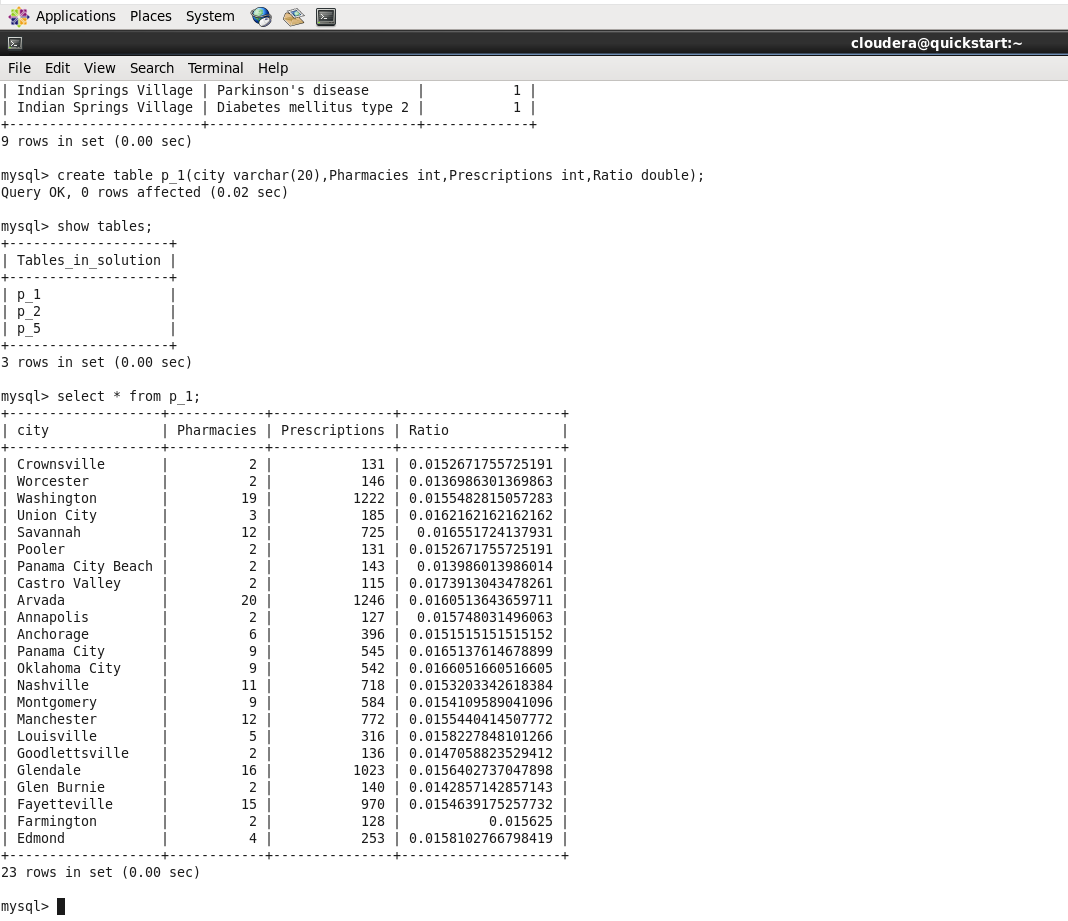
HAVING SUM(cnt) > 100

ODER BY 1;

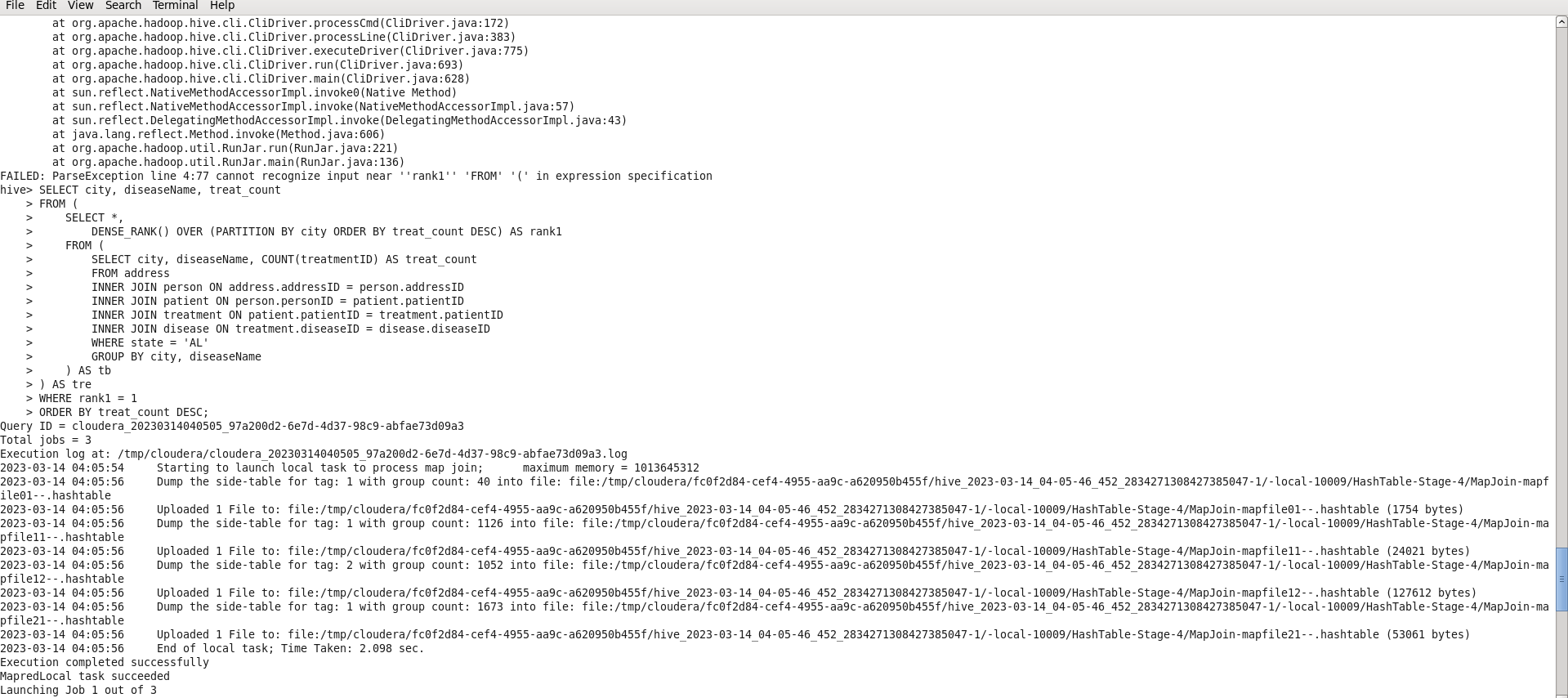


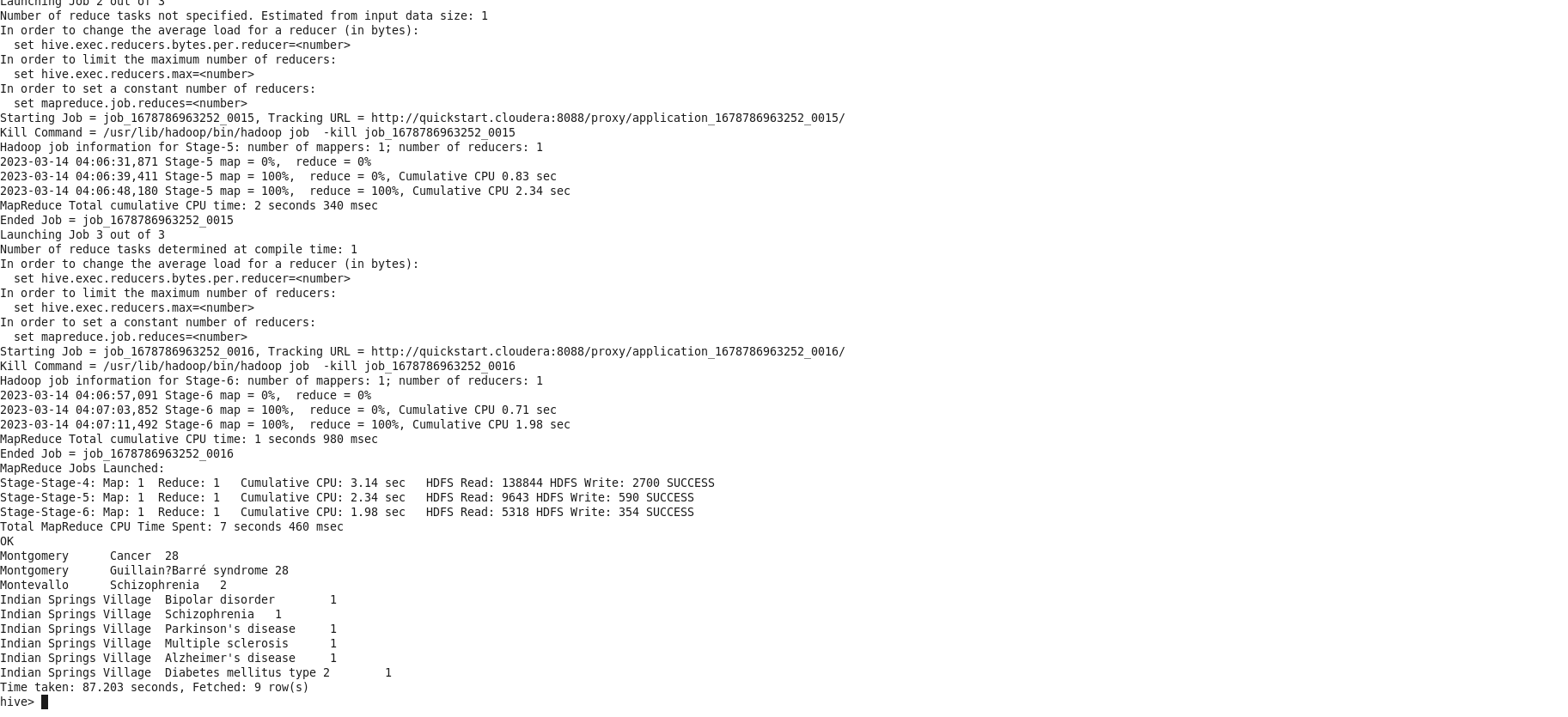
sqoop export --connect jdbc:mysql://localhost:3306/solution --username root --password cloudera --table p\_1 --export-dir /user/output1/000000\_0 --input-fields-terminated-by ',';

**Output:**



**Problem Statement 2:** The State of Alabama (AL) is trying to manage its healthcare resources more efficiently. For each city in their state, they need to identify the disease for which the maximum number of patients have gone for treatment. Assist the state for this purpose. Note: The state of Alabama is represented as AL in Address Table.





create table p\_2(city varchar(25),diseaseName varchar(25),treat\_count int);

CREATE EXTERNAL TABLE IF NOT EXISTS problem\_2 (city varchar(25),diseaseName

varchar(25),treat\_count int)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

LOCATION '/user/output2';

INSERT OVERWRITE TABLE problem\_2 SELECT city, diseaseName, treat\_count

FROM (

SELECT \*,

DENSE\_RANK() OVER (PARTITION BY city ORDER BY treat\_count DESC) AS rank1

FROM (

SELECT city, diseaseName, COUNT(treatmentID) AS treat\_count

FROM address

INNER JOIN person ON address.addressID = person.addressID

INNER JOIN patient ON person.personID = patient.patientID

INNER JOIN treatment ON patient.patientID = treatment.patientID

INNER JOIN disease ON treatment.diseaseID = disease.diseaseID

WHERE state = 'AL'

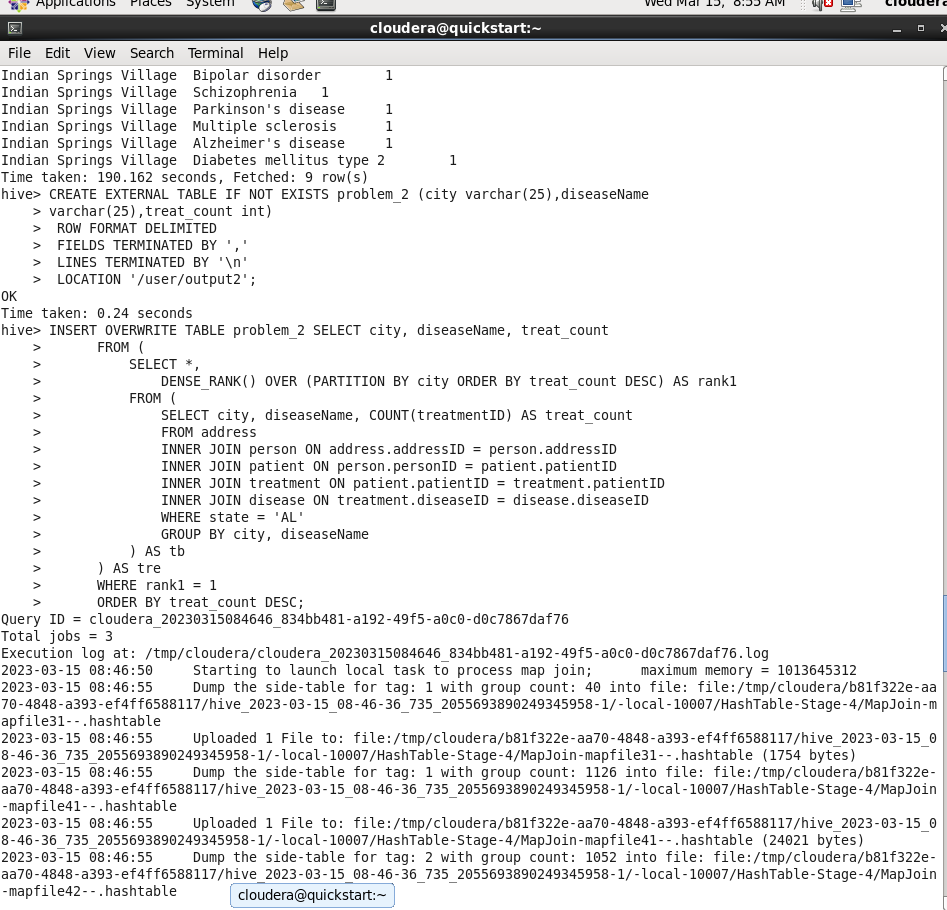
GROUP BY city, diseaseName

) AS tb

) AS tre

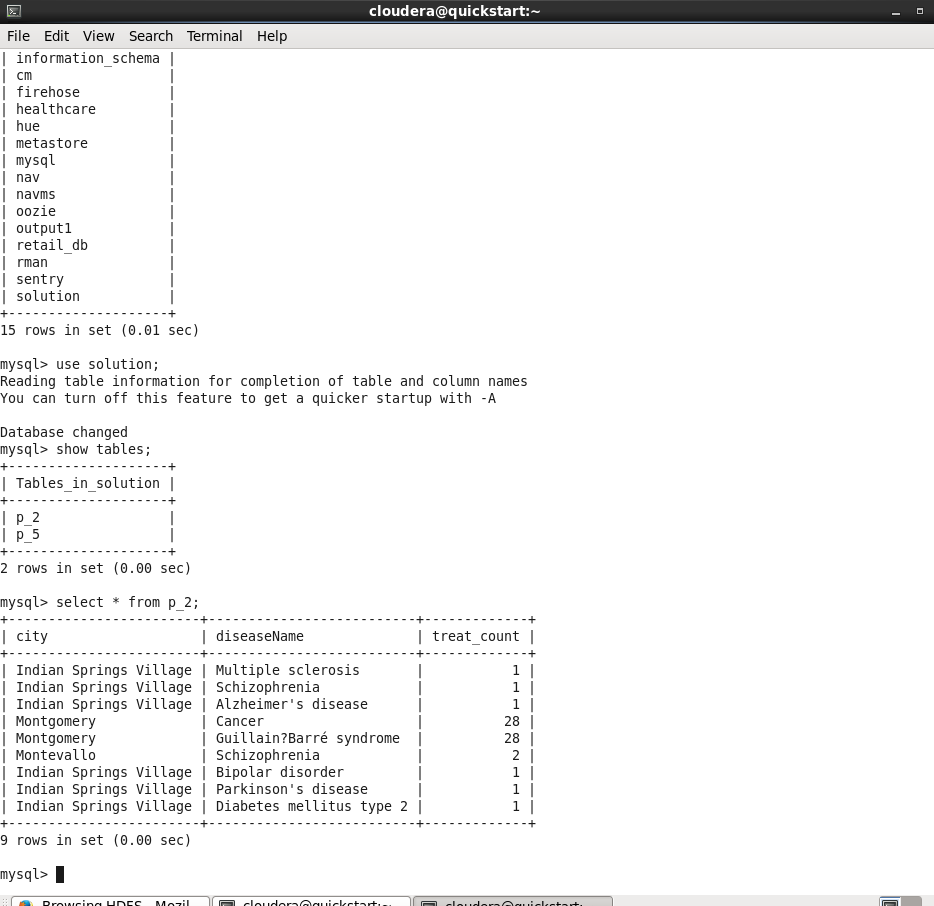
WHERE rank1 = 1

ORDER BY treat\_count DESC;



sqoop export --connect jdbc:mysql://localhost:3306/solution --username root --password cloudera --table p\_2 --export-dir /user/output2/000000\_0 --input-fields-terminated-by ',';

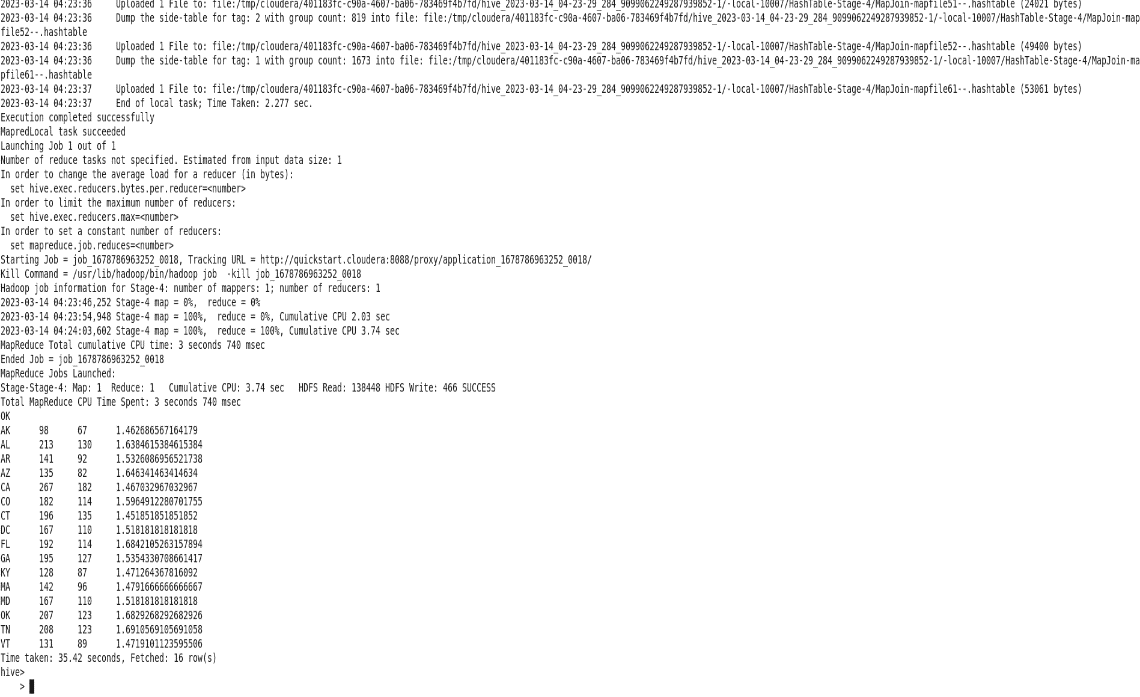
**output:**

****

**Problem Statement 3:** An Insurance company wants a state wise report of the treatments to claim ratio between 1st April 2021 and 31st March 2022 (days both included).

Assist them to create such a report.





create table p\_5(state varchar(25),treatcount int,claimcount int,ratio double);

CREATE EXTERNAL TABLE IF NOT EXISTS problem\_5 (state varchar(25),treatcount int,claimcount int,ratio double)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

LOCATION '/user/output';

INSERT OVERWRITE TABLE problem\_5 SELECT address.state, COUNT(treatment.treatmentID) AS treat\_count,

COUNT(claim.claimID) AS claim\_count,

COUNT(treatment.treatmentID) / COUNT(claim.claimID) AS ratio

FROM address

INNER JOIN person ON address.addressID = person.addressID

INNER JOIN patient ON person.personID = patient.patientID

INNER JOIN treatment ON patient.patientID = treatment.patientID

LEFT JOIN claim ON treatment.claimID = claim.claimID

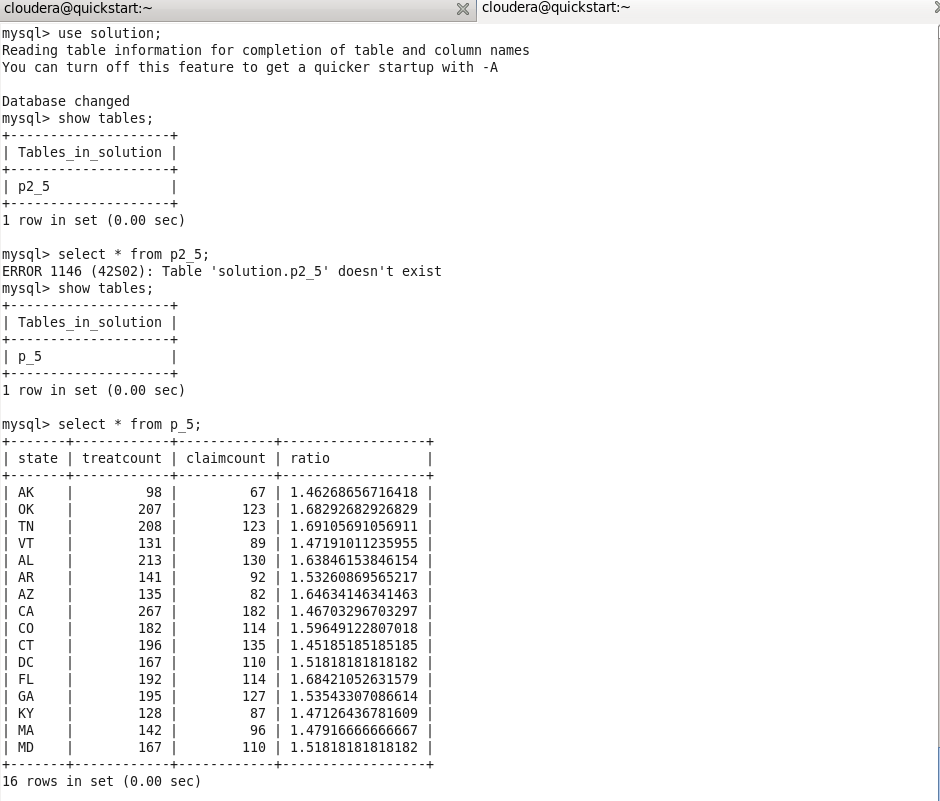
WHERE treatment.date BETWEEN '2021-04-01' AND '2022-03-31'

GROUP BY address.state;



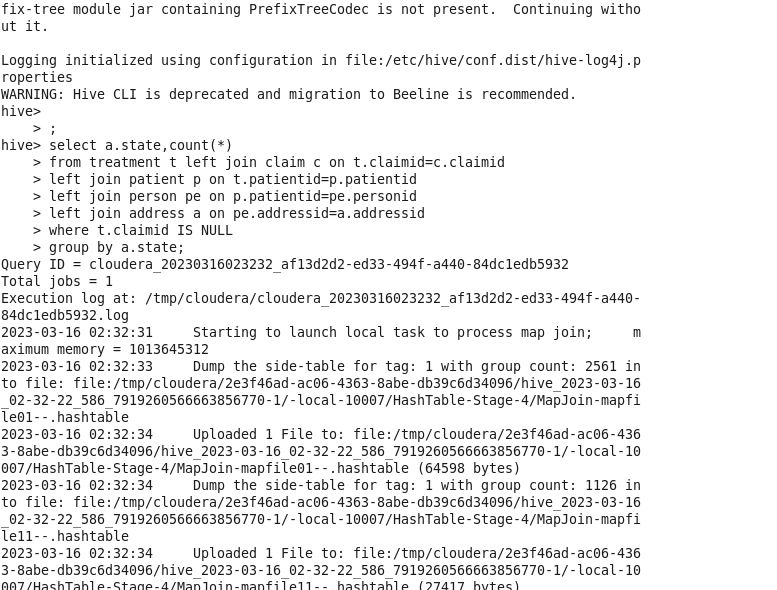
sqoop export --connect jdbc:mysql://localhost:3306/solution --username root --password cloudera --table p\_5 --export-dir /user/output/000000\_0 --input-fields-terminated-by ',';

**Output:**

****

**Problem 4:** Manish, from the healthcare department, wants to know how many registered people are registered as patients as well, in each city. Generate a report that shows each city that has 10 or more registered people belonging to it and the number of patients from that city as well as the percentage of the patient with respect to the registered people.

create table p\_6(state string, count int);

****

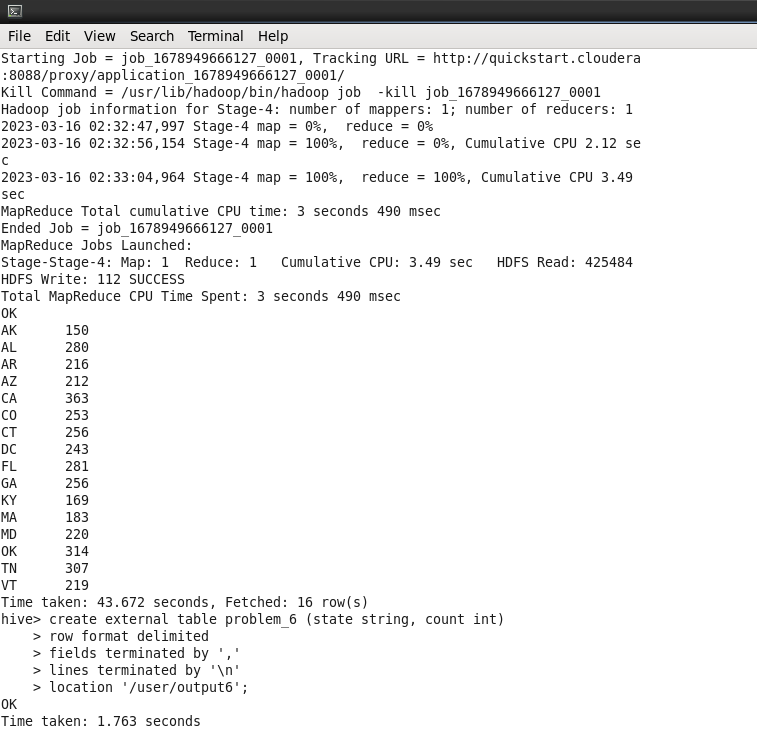
create external table problem\_6(state string, count int)

row format delimited

fields terminated by ','

lines terminated by '\n'

location '/user/output/output6';

****

INSERT OVERWRITE TABLE problem\_6 select a.state,count(\*)

from treatment t left join claim c on t.claimid=c.claimid

left join patient p on t.patientid=p.patientid

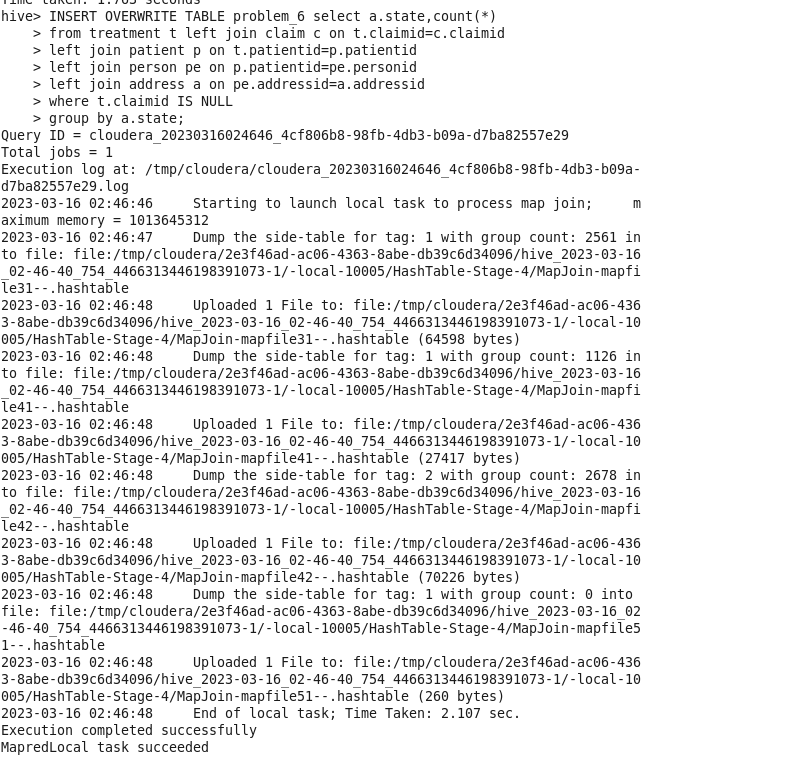
left join person pe on p.patientid=pe.personid

left join address a on pe.addressid=a.addressid

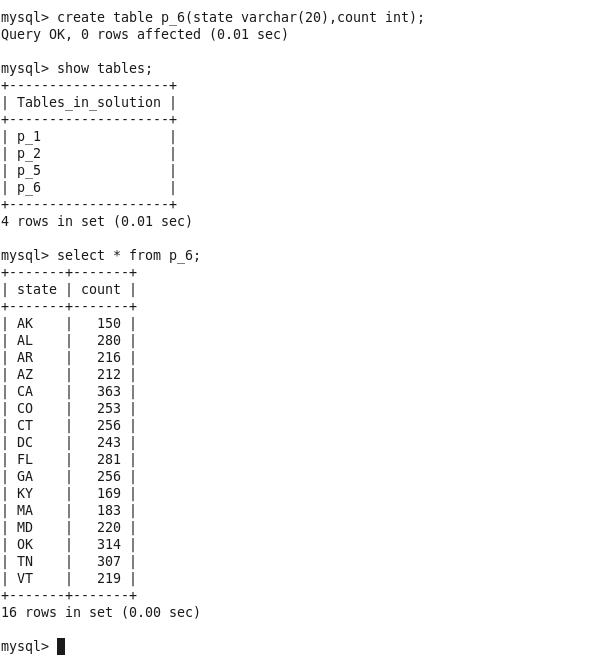
where t.claimid IS NULL

group by a.state;

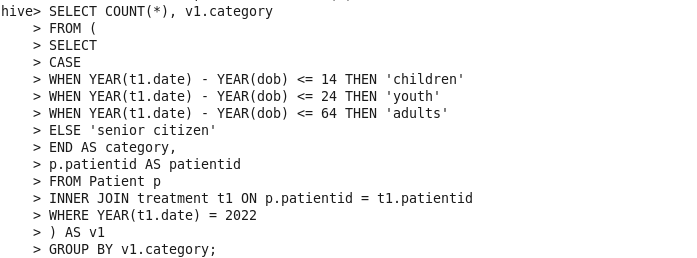
sqoop export --connect jdbc:mysql://localhost:3306/solution --username root --password cloudera --table p\_6 --export-dir /user/output6/s6\_p4/000000\_0 --input-fields-terminated-by ',';

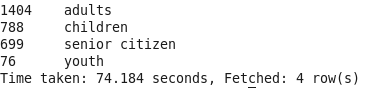


**Output:**

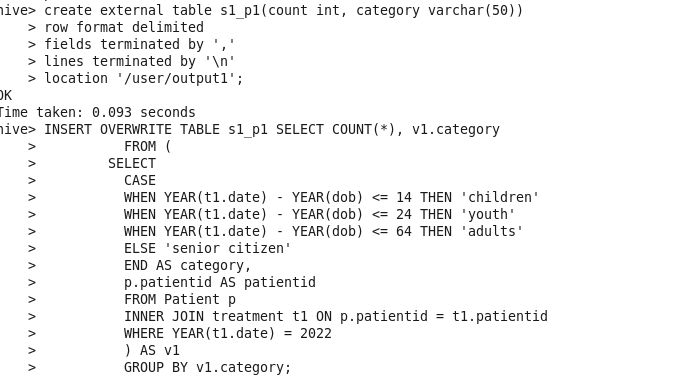
****

**Problem5:** Jimmy, from the healthcare department, has requested a report that shows how the number of treatments each age category of patients has gone through in the year 2022. The age category is as follows, Children (00-14 years), Youth (15-24 years), Adults (25-64 years), and Seniors (65 years and over).Assist Jimmy in generating the report.

****



**Creating External Table:**



**Creating a table in mysql:**

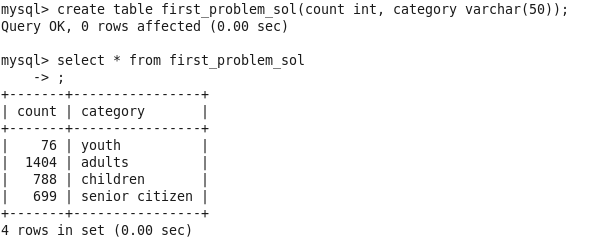
create table first\_problem\_sol(count int, category varchar(50));

**sqoop export:**

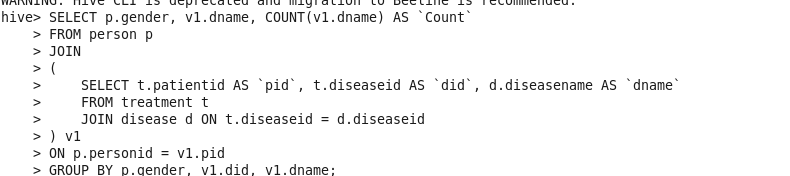
sqoop export --connect jdbc:mysql://localhost:3306/output --username root --password cloudera --table first\_problem\_sol --export-dir /user/output1/000000\_0 --input-fields-terminated-by ',';



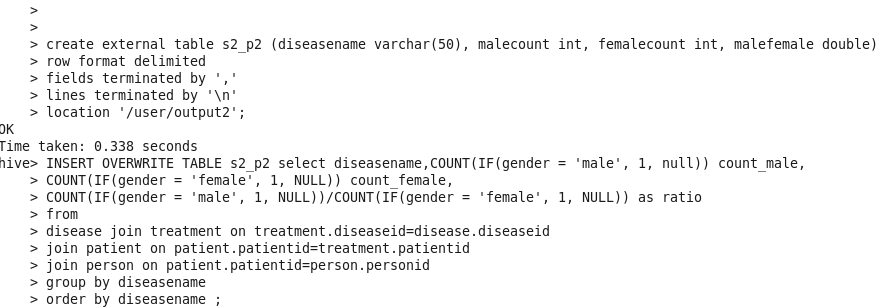
**Output:**



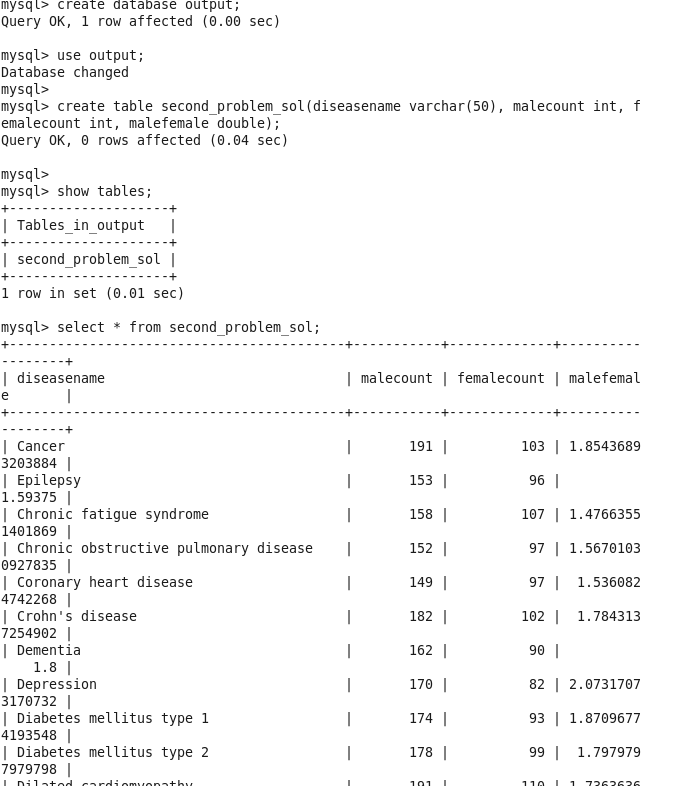
**Problem6:** Jimmy, from the healthcare department, wants to know which disease is infecting people of which gender more often. Assist Jimmy with this purpose by generating a report that shows for each disease the male-to-female ratio. Sort the data in a way that is helpful for Jimmy.

****

**Creating External Table:**

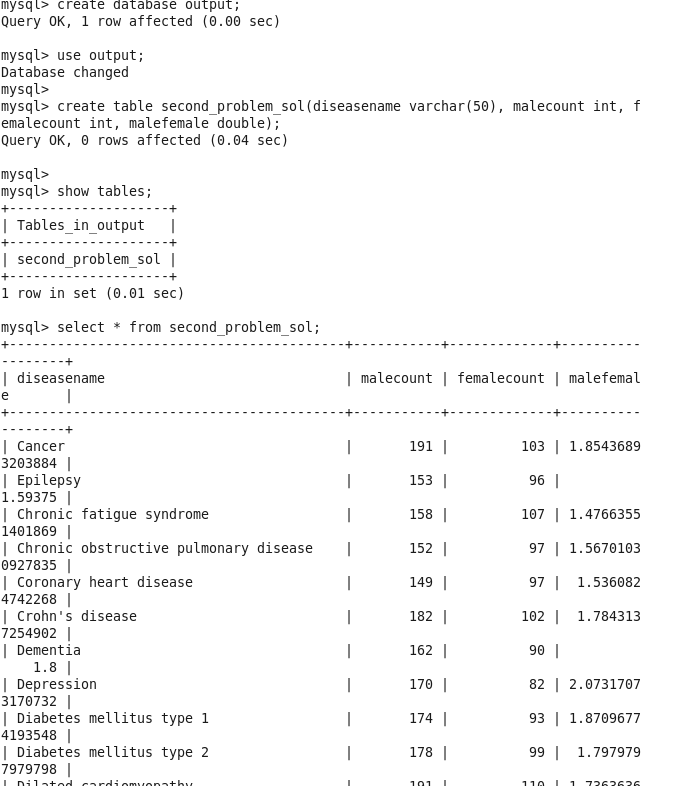
****

**Creating a table in Mysql:**

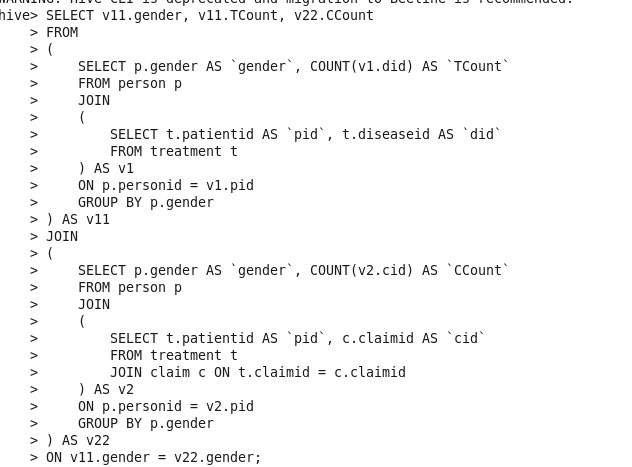
****

**Sqoop export:** sqoop export --connect jdbc:mysql://localhost:3306/output --username root --password cloudera --table second\_problem\_sol --export-dir /user/output2/000000\_0 --input-fields-terminated-by ','

**Output:**

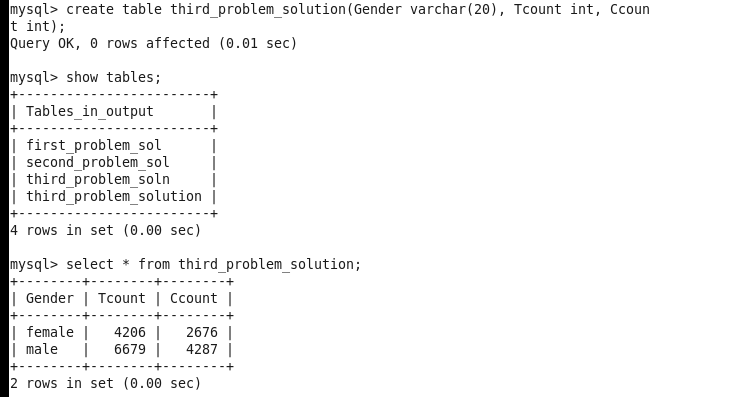


**Problem7:** Jacob, from insurance management, has noticed that insurance claims are not made for all the treatments. He also wants to figure out if the gender of the patient has any impact on the insurance claim. Assist Jacob in this situation by generating a report that finds for each gender the number of treatments, number of claims, and treatment-to-claim ratio. And notice if there is a significant difference between the treatment-to-claim ratio of male and female patients.

****

**Creating External Table:**

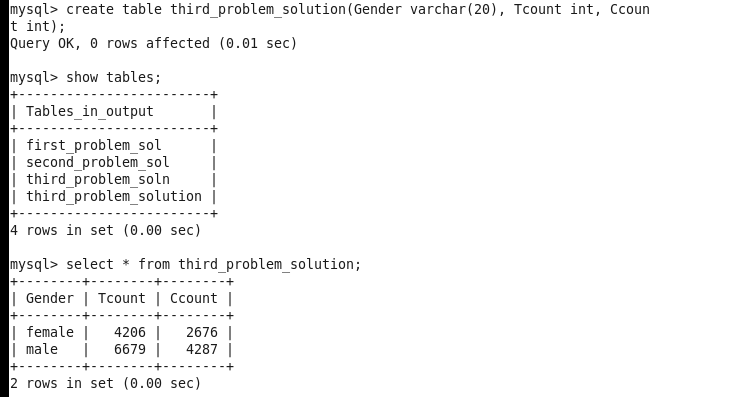


**Creating a table in MySQL:**

**Sqoop export:**



**Output:**



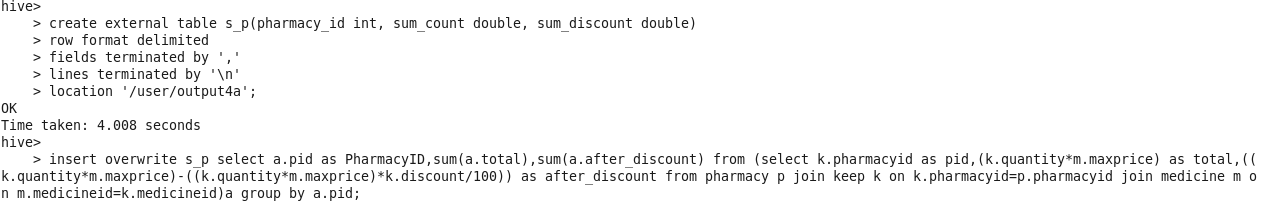
**Problem8:**The Healthcare department wants a report about the inventory of pharmacies. Generate a report on their behalf that shows how many units of medicine each pharmacy has in their inventory, the total maximum retail price of those medicines, and the total price of all the medicines after discount. Note: discount field in keep signifies the percentage of discount on the maximum price.

****

**Creating a table in Mysql:**

****

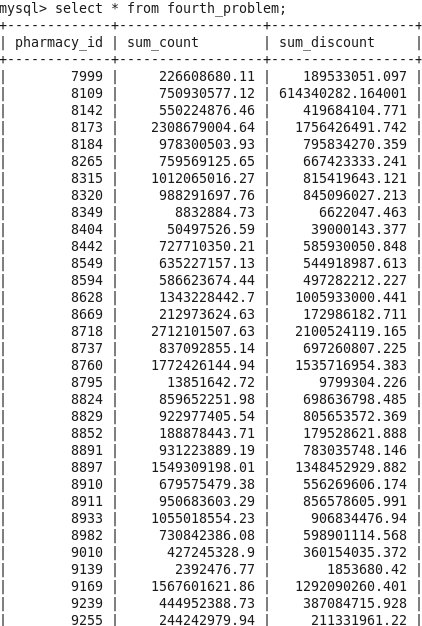
**Creating external table:**

****

**Sqoop export:**

****

**Output:**

****

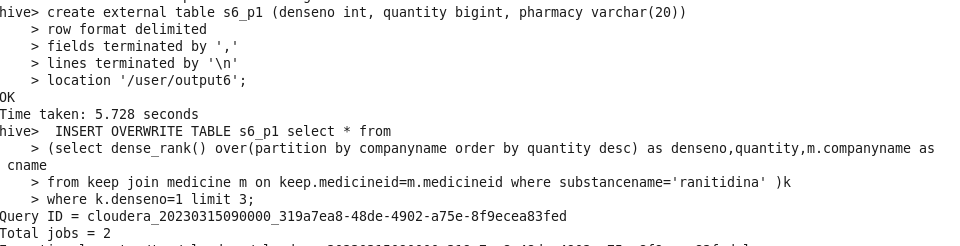
**Problem9:** It is suspected by healthcare research department that the substance “ranitidine” might be causing some side effects. Find the top 3 companies using the substance in their medicine so that they can be informed about it.

select \* from

(select dense\_rank() over(partition by companyname order by quantity desc) as denseno,quantity,m.companyname as cname

from keep join medicine m on keep.medicineid=m.medicineid where substancename='ranitidina' )k

where k.denseno=1 limit 3;

**Creating external table:**

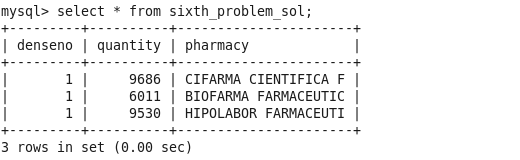
**Creating a table in mysql:**

****

**Sqoop Export:**

****

**Output:**

****