* Scoop import command –

sqoop import-all-tables --connect jdbc:mysql://localhost:3306/healthdb --username root --password cloudera --hive-import --m 1



1. **Find the number of male & female disease count and male to female ratio per disease**

* Creating an external table in hive,  
    
  create external table query\_1

(

diseaseName string,

male int,

female int,

ratio double

)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n';

* Inserting data into query\_1,  
    
  insert into table query\_1

select \*, round((male/female),2) as m\_f\_ratio from

(

select d.diseaseName,

sum(case when gender = 'male' then 1 else 0 end ) as male,

sum(case when gender = 'female' then 1 else 0 end ) as female

from person p inner join treatment t

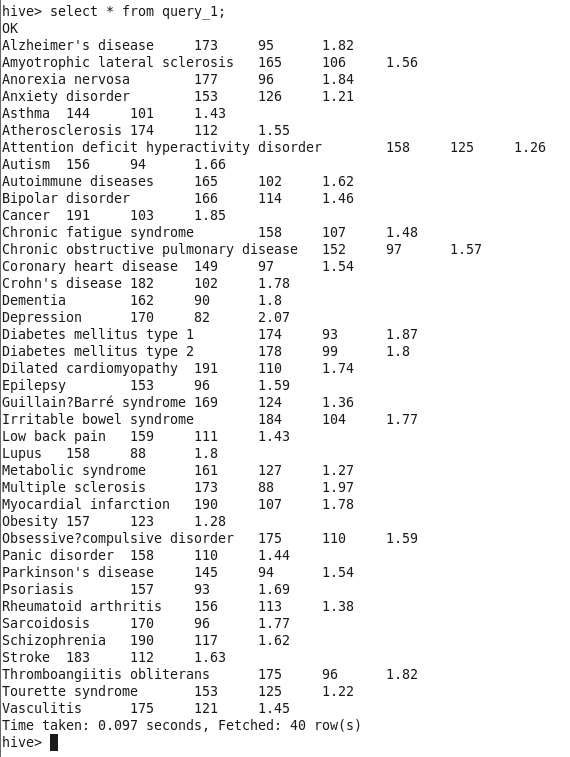
on p.personid = t.patientid

inner join disease d

on t.diseaseID = d.diseaseID

group by d.diseaseName

) a;



* Creating table in mysql

create table query\_1

(

diseaseName varchar(15),

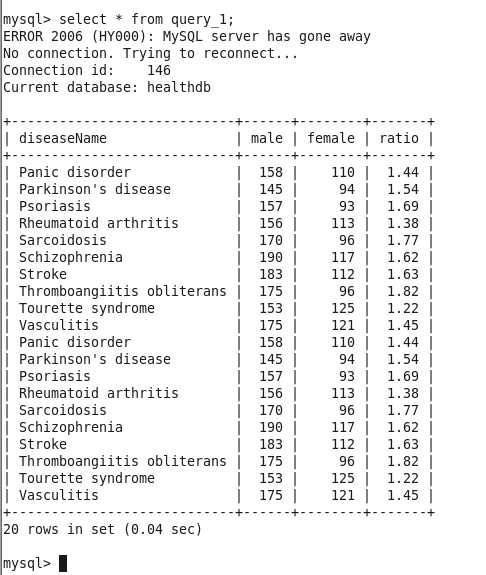
male int,

female int,

ratio double

);

* sqoop export --connect jdbc:mysql://localhost:3306/healthdb --username root --password cloudera --table query\_1 --export-dir /user/hive/warehouse/query\_1 --input-fields-terminated-by ','



1. **Find out total number of treatment , number of claims and male to female treatment to claim ratio**

* Creating an external table,

create external table query\_2(

gender string,

total\_treatments int,

total\_claims int,

ratio double)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n';

* Inserting data,

insert into table query\_2

select \*, round(total\_treatments/ total\_claims, 2) as ratio from

(

select p.gender, count(t.treatmentID) as total\_treatments,

count(c.claimID) as total\_claims from

person p inner join treatment t

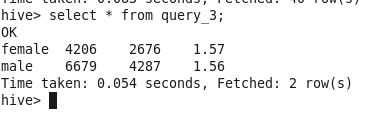
on p.personID = t.patientID

left join claim c

on t.claimID = c.claimID

group by p.gender

)a;



* Creating table in mysql,

create table query\_3(

gender varchar(10),

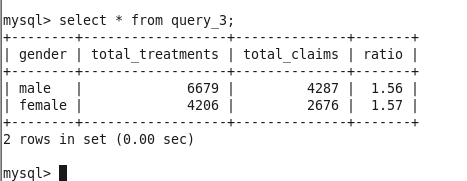
total\_treatments int,

total\_claims int,

ratio double);

* Scoop export to mysql,

sqoop export --connect jdbc:mysql://localhost:3306/healthdb --username root --password cloudera --table query\_3 --export-dir /user/hive/warehouse/query\_3 --input-fields-terminated-by ','



1. **Show the prescription Id, the Total Quantity of all the medicines in that**

**prescription, and the Quantity tag for all the prescriptions issued by**

**'Ally Scripts'.**

* Creating external table in hive,

create external table query\_4 (prescriptionID double, qty int, tag string) row format delimited fields terminated by ',' lines terminated by '\n';

* Inserting data,

Insert into table query\_4

select pr.prescriptionID, sum(c.quantity),

case

when sum(c.quantity)<20 then 'low quantity'

when sum(quantity)>=20 and sum(c.quantity)<50 then 'medium quantity'

when sum(c.quantity)>=50 then 'high quantity'

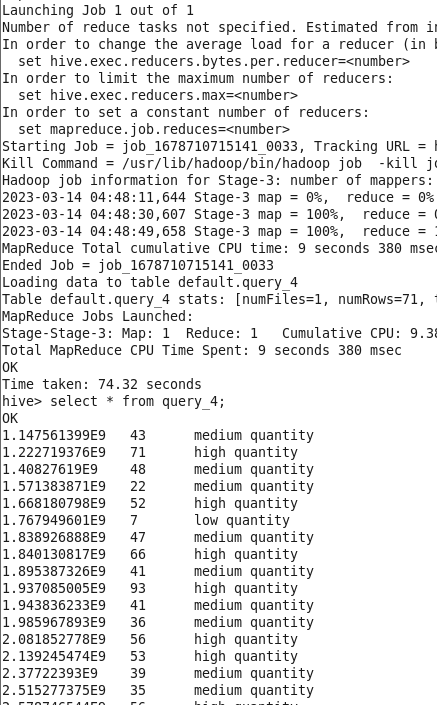
end Tag

from pharmacy p inner join prescription pr on p.pharmacyID = pr.pharmacyID

inner join contain c on c.prescriptionID=pr.prescriptionID

where p.pharmacyName = 'Ally Scripts'

group by pr.prescriptionID;

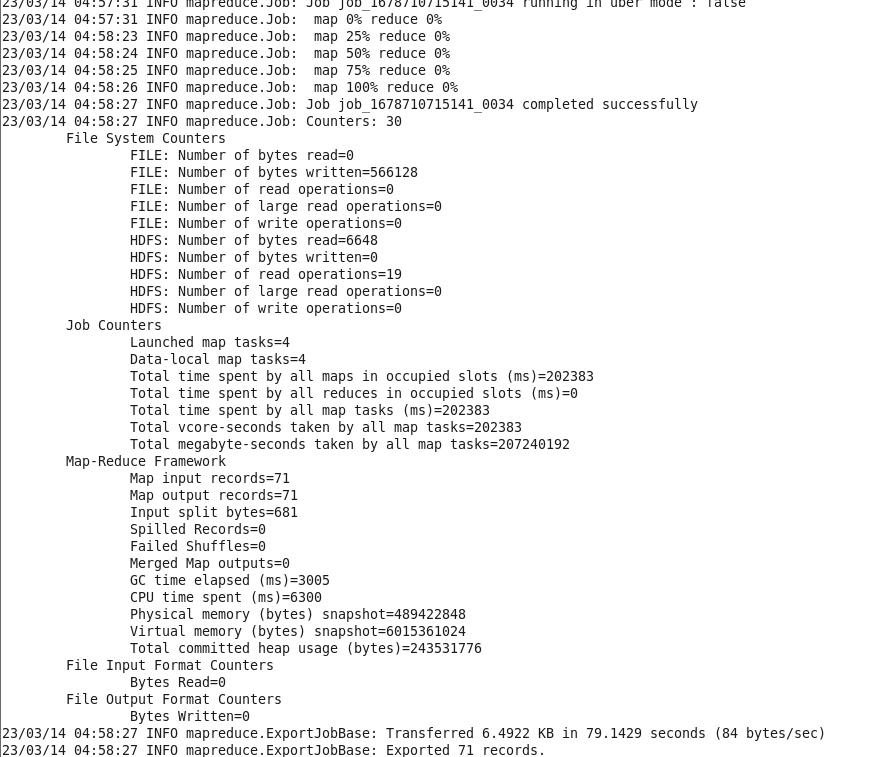


* Creating table in mysql,

create table query\_4 (prescriptionID double, qty int, tag varchar(20)) ;

* Scoop export data to the mysql,

sqoop export --connect jdbc:mysql://localhost:3306/healthdb --username root --password cloudera --table query\_4 --export-dir /user/hive/warehouse/query\_4 --input-fields-terminated-by ','



1. **Some complaints have been lodged by patients that they have been prescribed hospital-exclusive medicine that they can’t find elsewhere and facing problems due to that. Joshua, from the pharmacy management, wants to get a report of which pharmacies have prescribed hospital-exclusive medicines the most in the years 2021 and 2022. Assist Joshua to generate the report so that the pharmacies who prescribe hospital-exclusive medicine more often are advised to avoid such practice if possible.**

* Creating an external table,

create external table query\_5 (name string, year int, hospitalExclusive int) row format delimited fields terminated by ',' lines terminated by '\n';

* Inserting data,

insert into table query\_5

select p.pharmacyName, year(t.date) year, count(m.medicineID) Hospital\_Exclusive

from medicine m inner join contain c on m.medicineID = c.medicineID

inner join prescription pr on pr.prescriptionID = c.prescriptionID

inner join pharmacy p on p.pharmacyID = pr.pharmacyID

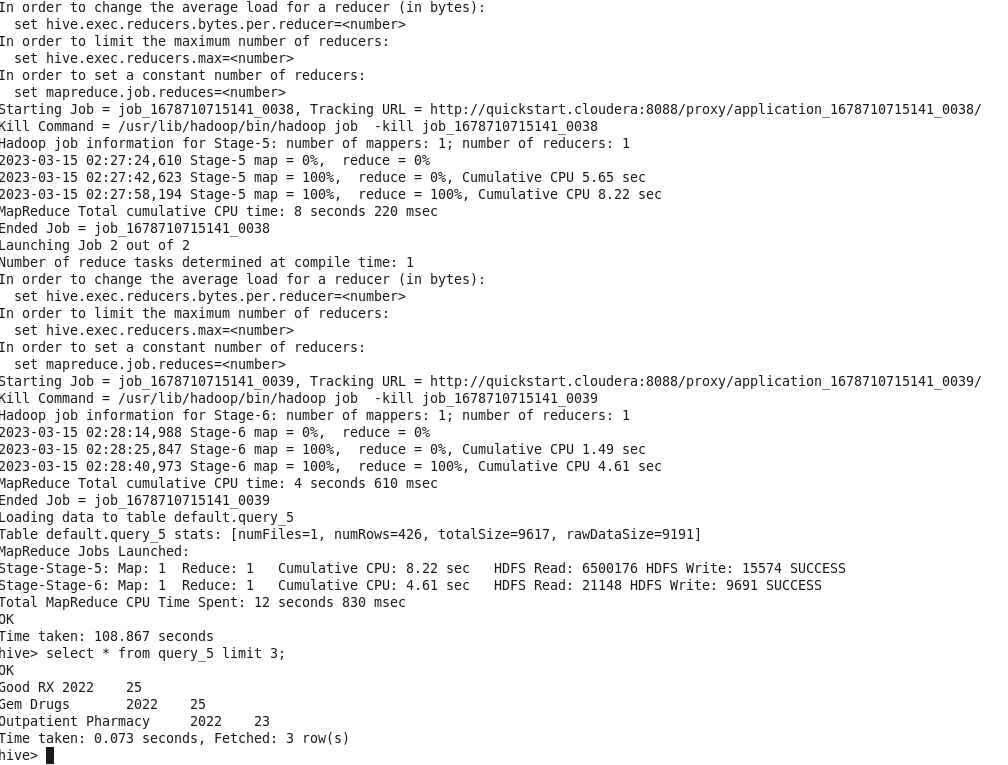
inner join treatment t on pr.treatmentID = t.treatmentID

where m.hospitalExclusive='S'

and year(t.date) in ('2021','2022')

group by p.pharmacyName, year(t.date)

order by Hospital\_Exclusive desc;



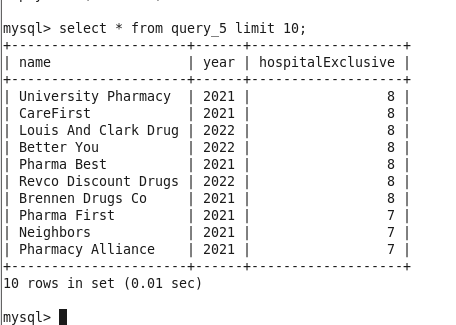
* Creating mysql table,

create table query\_5 (name varchar(40), year int, hospitalExclusive int) ;

* Scoop export to the table,

sqoop export --connect jdbc:mysql://localhost:3306/healthdb --username root --password cloudera --table query\_5 --export-dir /user/hive/warehouse/query\_5 --input-fields-terminated-by ','

;





* Create external table,

create external table query\_6 (state string, diseaseID int, treat\_count int) row format delimited fields terminated by ',' lines terminated by '\n';

* Insertion of data,

with cte as

(

select ad.state state,t.diseaseID diseaseID, count(t.treatmentID) as treat\_count

from address ad inner join person p

on ad.addressID = p.addressID

inner join (select \* from treatment\_part\_bkt where year = 2022) t

on p.personID = t.patientID

group by ad.state,t.diseaseID

),

cte\_2 as

(

select \*, dense\_rank() over(partition by state order by treat\_count desc) dn\_desc,

dense\_rank() over(partition by state order by treat\_count ) dn\_asc from cte

)

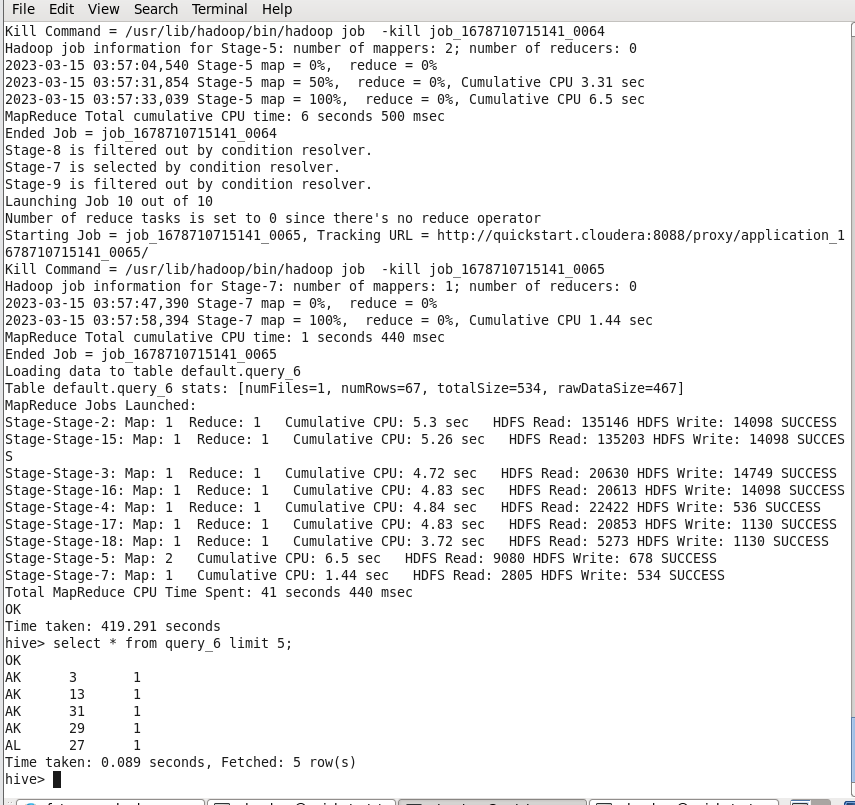
insert into table query\_6

select state,diseaseID,treat\_count from cte\_2 where dn\_desc = 1

union all

select state,diseaseID,treat\_count from cte\_2 where dn\_asc = 1

order by state;



* Create mysql table,

Create table query\_6 (state varchar(30), diseaseID int, treat\_count int);

* Sqoop export data form hive external table to mysql,

sqoop export --connect jdbc:mysql://localhost:3306/healthdb --username root --password cloudera --table query\_6 --export-dir /user/hive/warehouse/query\_6 --input-fields-terminated-by ',' ;

