**Problem Statement 1:**  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.

**Query:**

hive>create view p11 as select e.category, count(\*) as count from (select (case when DATEDIFF("2022-12-01",p.dob) / 365.25 <=14 then "children"

> when DATEDIFF("2022-12-01",p.dob) / 365.25 <=24 then "youth"

> when datediff("2022-12-01",p.dob) <= 64 then "Adults"

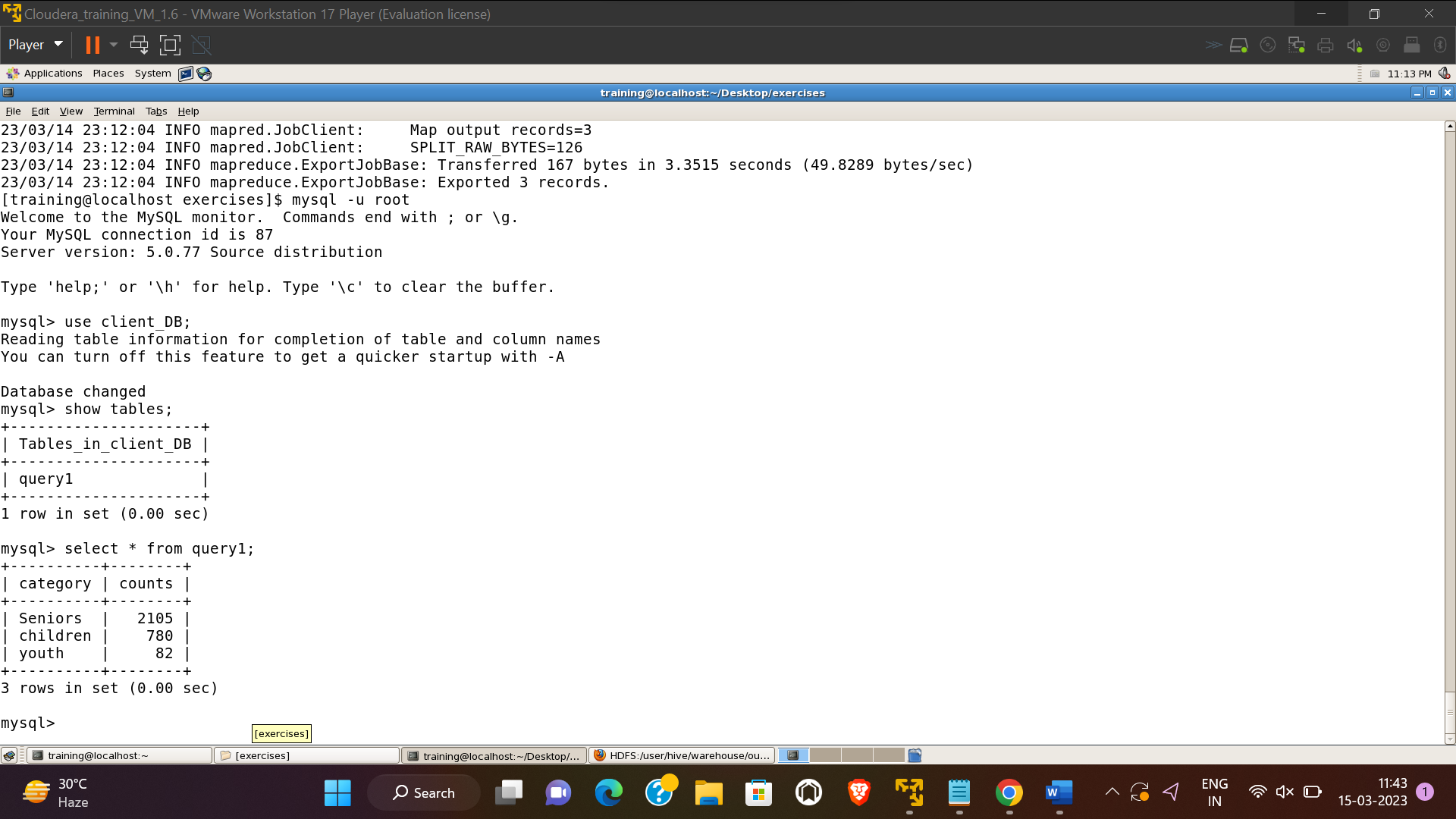
> else "Seniors"

> end) as category from treatment t join patient p on t.patientID=p.patientID where year(t.`date`)=2022) e group by e.category;

**External Table:**

> create external table out1(category varchar(10), count int);

Created a database and table in sql named client\_DB and exported the out1 to the table query1:



**Problem Statement 2:**  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.

**Hive Query:**

create view v1as select d.diseasename, p.gender, count(\*) as cnt from disease d join treatment t on d.diseaseid = t.diseaseid join person p on p.personid = t.patientid group by d.diseasename, p.gender order by cnt;

create view v2 as select a.d diseasename, b.cnt male, a.cnt female from (select diseasename d , cnt from v1 where gender = 'female')a join (select diseasename d, cnt from v1 where gender='male')b on a.d=b.d;

**External Table:**

hive> create external table out2(diseasename string, male int,female int);

hive> insert overwrite table out2 select \* from v2;

**Creating empty table in Mysql:**

mysql> create table query2(diseasename varchar(50),male int, female int);

Graphical user interface, text, application, Word

Description automatically generated

Problem Statement 3: 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.

Hive Query:

create view v3 as select p.gender, count(t.treatmentid), count(t.claimid), count(t.treatmentid)/count(t.claimid) from treatment t join person p on t.patientid=p.personid group by p.gender;

External Table:

create external table out4(Gender string, treatments int, claims int, ratio double);

insert overwrite table out4 select \* from v3;

sqoop export --connect jdbc:mysql://localhost:3306/client\_DB --username root --table query3 --export-dir /user/hive/warehouse/out4/000000\_0 --input-fields-terminated-by '\0001'

Graphical user interface, text, application

Description automatically generated

Problem Statement 4: 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.

Hive Query:

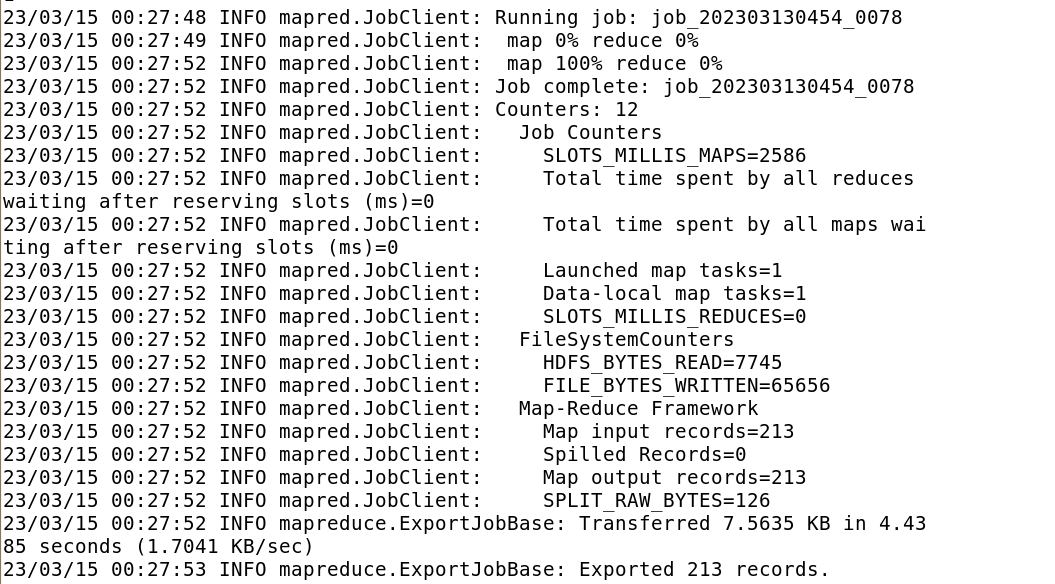
create view v4 as select p.pharmacyid, round(sum( k.quantity\*m.maxPrice),5), round(sum((k.quantity\*m.maxPrice)-(k.quantity\*m.maxPrice\*k.discount/100)),5) from pharmacy p join keep k on p.pharmacyID=k.pharmacyID join medicine m on k.medicineID=m.medicineID group by p.pharmacyid;

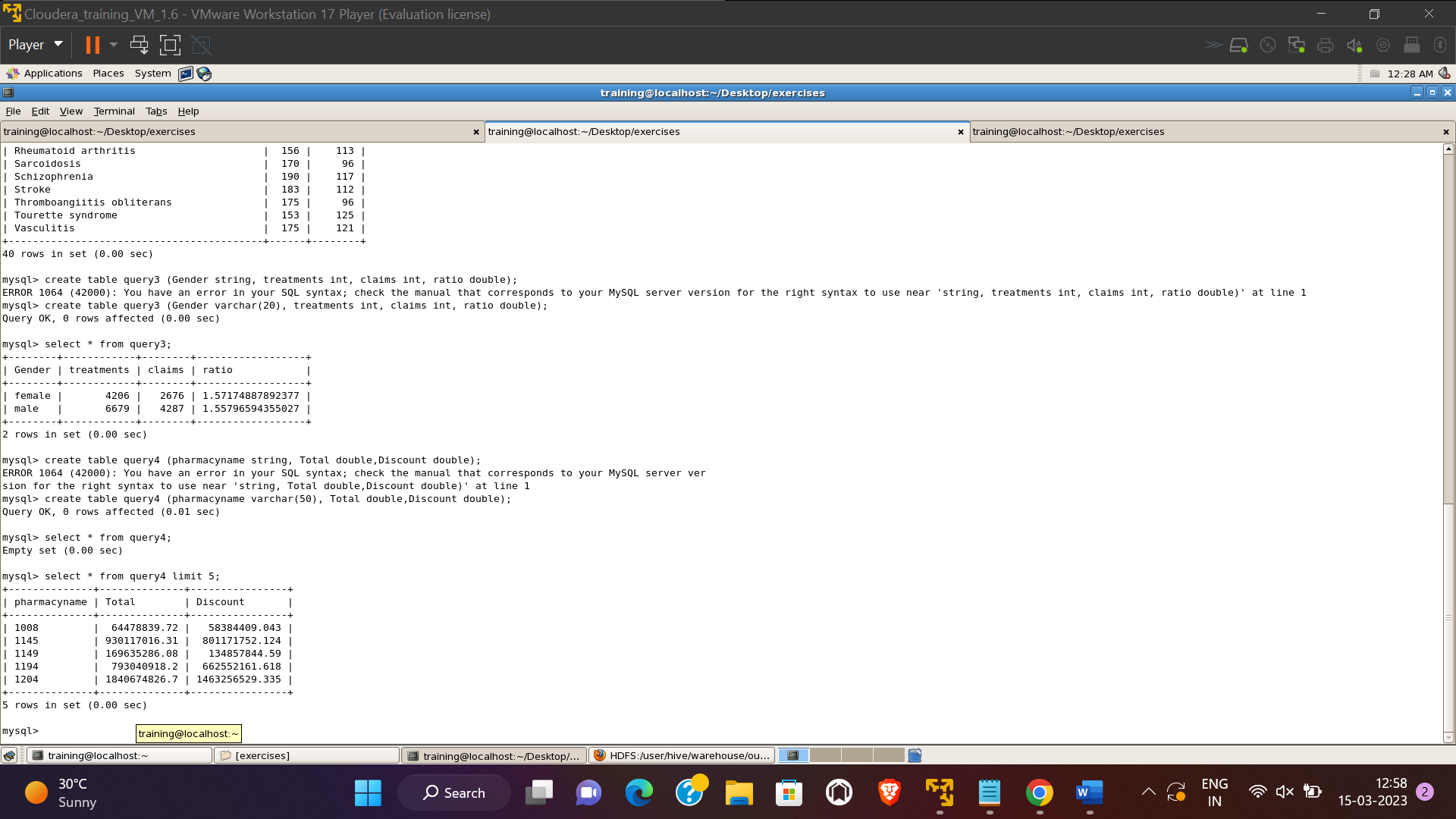
External Table:

create external table out5(pharmacyname string, Total double,Discount double);

insert overwrite table out5 select \* from v4;

sqoop export --connect jdbc:mysql://localhost:3306/client\_DB --username root --table query4 --export-dir /user/hive/warehouse/out5/000000\_0 --input-fields-terminated-by '\0001'





Problem Statement 5: The healthcare department suspects that some pharmacies prescribe more medicines than others in a single prescription, for them, generate a report that finds for each pharmacy the maximum, minimum and average number of medicines prescribed in their prescriptions.

Hive Query:

create view v5 as select pharmacyname, max(cnt) `max`, min(cnt) `min`,avg(cnt) `avg` from (select p1.pharmacyname, p.prescriptionid, sum(c.quantity) cnt from prescription p join pharmacy p1 on p1.pharmacyid = p.pharmacyid join contain c on c.prescriptionid = p.prescriptionid group by p1.pharmacyname,p.prescriptionid)a group by pharmacyname;

External Table:

create external table out6 (pharmacyname string,`max` int, `min` int,avg double);

insert overwrite table out6 select \* from v5;

sqoop export --connect jdbc:mysql://localhost:3306/client\_DB --username root --table query5 --export-dir /user/hive/warehouse/out6/000000\_0 --input-fields-terminated-by '\0001'

Text

Description automatically generated

Graphical user interface, text, application, Word

Description automatically generated

Problem Statement 6:

Johansson is trying to prepare a report on patients who have gone through treatments more than once. Help Johansson prepare a report that shows the patient's name, the number of treatments they have undergone, and their age,Sort the data in a way that the patients who have undergone more treatments appear on top.

create view v6 as SELECT P.PERSONNAME as PERSONNAME,X.CNT as TREATMENTCOUNT,cast(datediff('2023-03-14',PA.DOB)/365 as int) as AGE FROM

(select T.PATIENTID as PATIENTID,COUNT(t.TREATMENTID) as CNT FROM TREATMENT T join PATIENT P on P.PATIENTID=T.PATIENTID

GROUP BY T.PATIENTID HAVING COUNT(t.TREATMENTID)>1 ORDER BY CNT)X

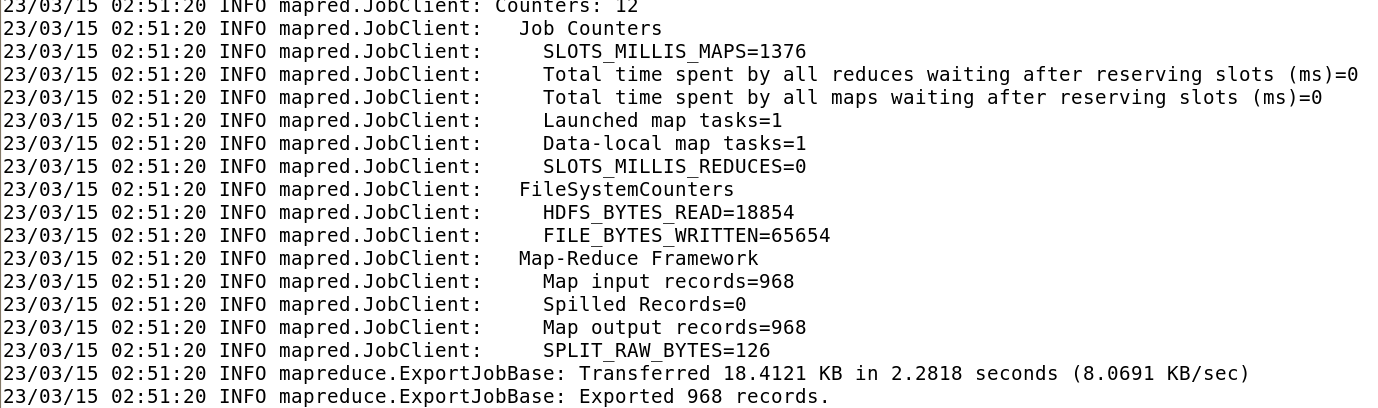
join Patient PA on PA.PATIENTID=X.PATIENTID join Person P on P.PERSONID=PA.PATIENTID ORDER BY TREATMENTCOUNT DESC;

creating external table:

create external table out7(personname string,Tcount int,age int);

insert overwrite table out7 select \* from v6;

sqoop export --connect jdbc:mysql://localhost:3306/client\_DB --username root --table query6 --export-dir /user/hive/warehouse/out7/000000\_0 --input-fields-terminated-by '\0001'



A picture containing text

Description automatically generated

Problem statement 7:

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 IF NOT EXISTS address\_PART (addressid int, address1 String, city String,zip int)

COMMENT 'address\_PART details'

PARTITIONED BY (state String)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

STORED AS TEXTFILE;

create view V7 as select a.city,d.diseasename,count(t.patientid) as counTT from address\_part a join person p on p.addressid=a.addressid join treatment t on t.patientID=p.personid join disease d on t.diseaseid=d.diseaseid where a.state='AL' group by a.city,d.diseasename;

insert overwrite table out8 select a.c,a.d,a.co from (select city as c,diseasename as d,counTT as co,ROW\_NUMBER() OVER(partition by city order by counTT desc) as rn from p6) as a where a.rn=1;

create table query7(city varchar(50),diseasename varchar(50),counTT numeric(10));

exporting to client database:

sqoop export --connect jdbc:mysql://localhost:3306/client\_DB --username root --table query7 --export-dir /user/hive/warehouse/out8/000000\_0 --input-fields-terminated-by '\0001'

Text, letter

Description automatically generated

Problem statement 8:

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.

Create view v8 as select ph.pharmacyid,count(c.medicineid) as counTT from treatment t join Prescription ph on t.treatmentid=ph.treatmentid join contain c on c.prescriptionid=ph.prescriptionid join medicine m on c.medicineid=m.medicineid where m.hospitalexclusive ='S' and year(t.date) in (2021,2022) group by ph.pharmacyid order by counTT desc;

create table query8 (pharmacyid int, count int);

insert overwrite table out9 as select \* from v8;

exporting to client database:

sqoop export --connect jdbc:mysql://localhost:3306/client\_DB --username root --table query8 --export-dir /user/hive/warehouse/out9/000000\_0 --input-fields-terminated-by '\0001'

Text, letter

Description automatically generated