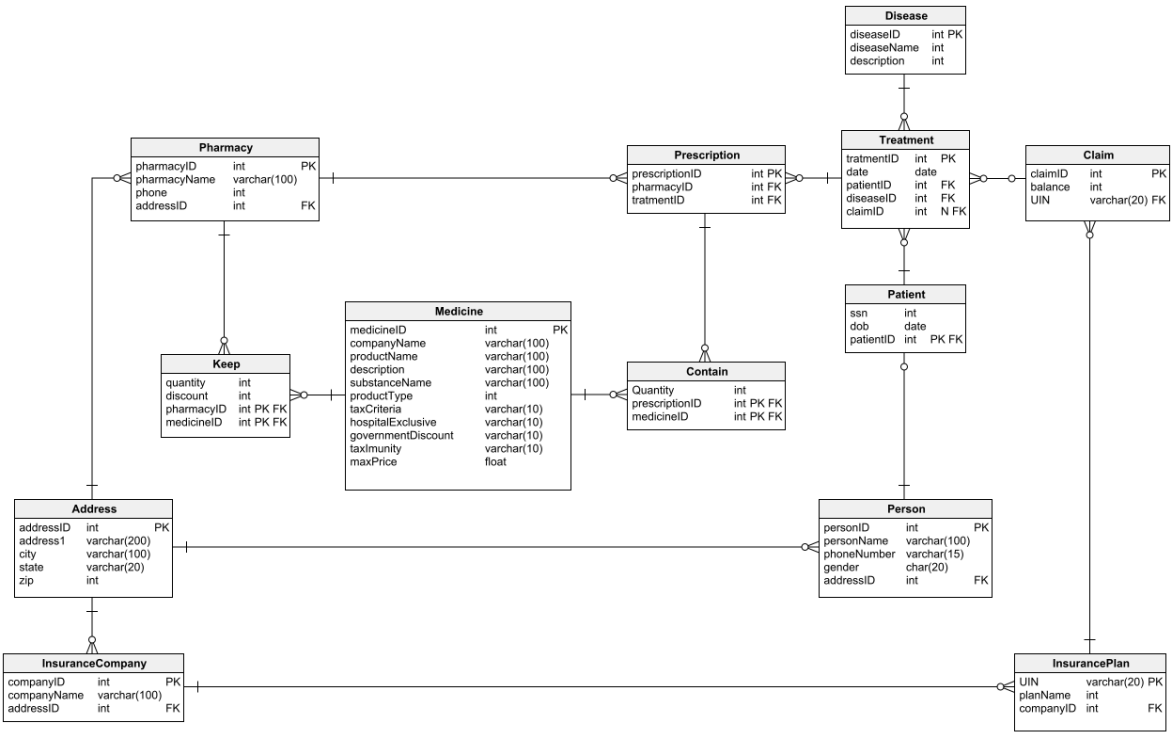
PROJECT

Database schema:



Problem Statement 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.

Problem Statement 2: Insurance companies want to assess the performance of their insurance plans. Generate a report that shows each insurance plan, the company that issues the plan, and the number of treatments the plan was claimed for.

Problem Statement 3: Insurance companies want to assess the performance of their insurance plans. Generate a report that shows each insurance company's name with their most and least claimed insurance plans.

Problem Statement 4: The healthcare department wants a state-wise health report to assess which state requires more attention in the healthcare sector. Generate a report for them that shows the

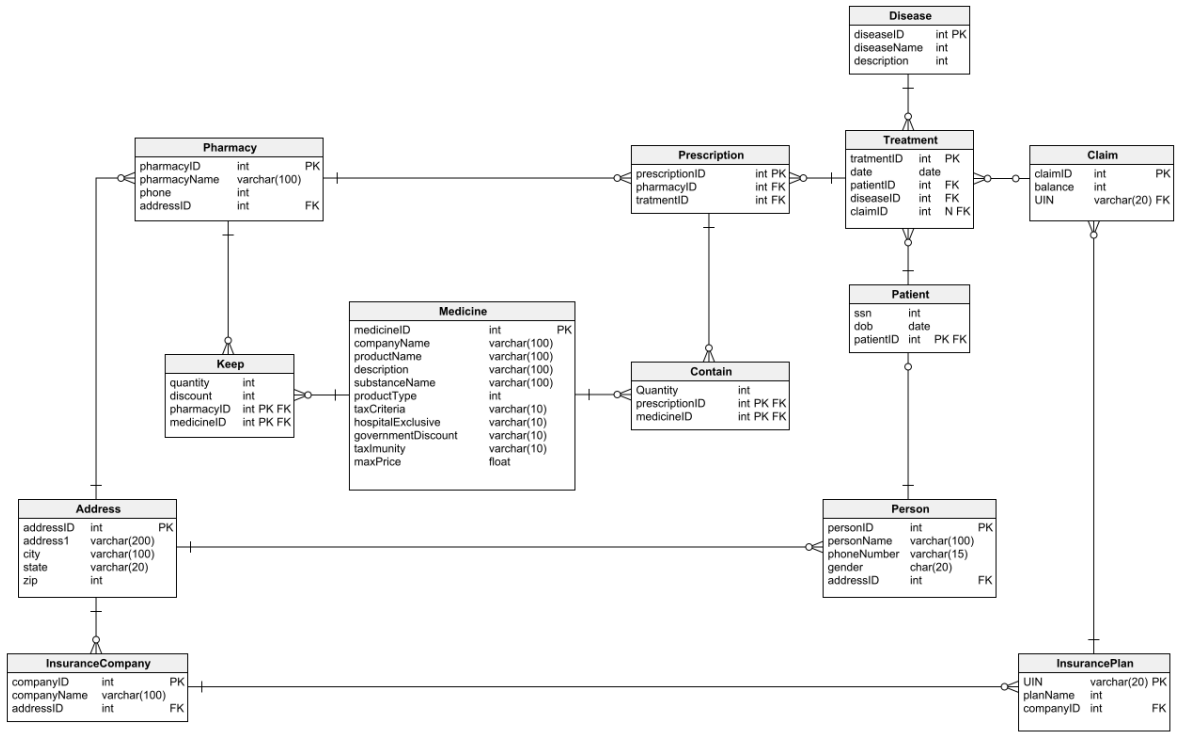
state name, number of registered people in the state, number of registered patients in the state, and the people-to-patient ratio. sort the data by people-to-patient ratio.

Problem Statement 5: Jhonny, from the finance department of Arizona(AZ), has requested a report that lists the total quantity of medicine each pharmacy in his state has prescribed that falls under Tax criteria I for treatments that took place in 2021. Assist Jhonny in generating the report.

PROJECT

SQL rollup

Database schema:



Problem Statement 1:

Brian, the healthcare department, has requested for a report that shows for each state how many people underwent treatment for the disease “Autism”. He expects the report to show the data for each state as well as each gender and for each state and gender combination.

Prepare a report for Brian for his requirement.

Problem Statement 2:

Insurance companies want to evaluate the performance of different insurance plans they offer.

Generate a report that shows each insurance plan, the company that issues the plan, and the number of treatments the plan was claimed

for. The report would be more relevant if the data compares the performance for different years(2020, 2021 and 2022) and if the report also includes the total number of claims in the different years, as well as the total number of claims for each plan in all 3 years combined.

Problem Statement 3:

Sarah, from the healthcare department, is trying to understand if some diseases are spreading in a particular region. Assist Sarah by creating a report which shows each state the number of the most and least treated diseases by the patients of that state in the year 2022. It would be helpful for Sarah if the aggregation for the different combinations is found as well. Assist Sarah to create this report.

Problem Statement 4:

Jackson has requested a detailed pharmacy report that shows each pharmacy name, and how many prescriptions they have prescribed for each disease in the year 2022, along with this Jackson also needs to view how many prescriptions were prescribed by each pharmacy, and the total number prescriptions were prescribed for each disease.

Assist Jackson to create this report.

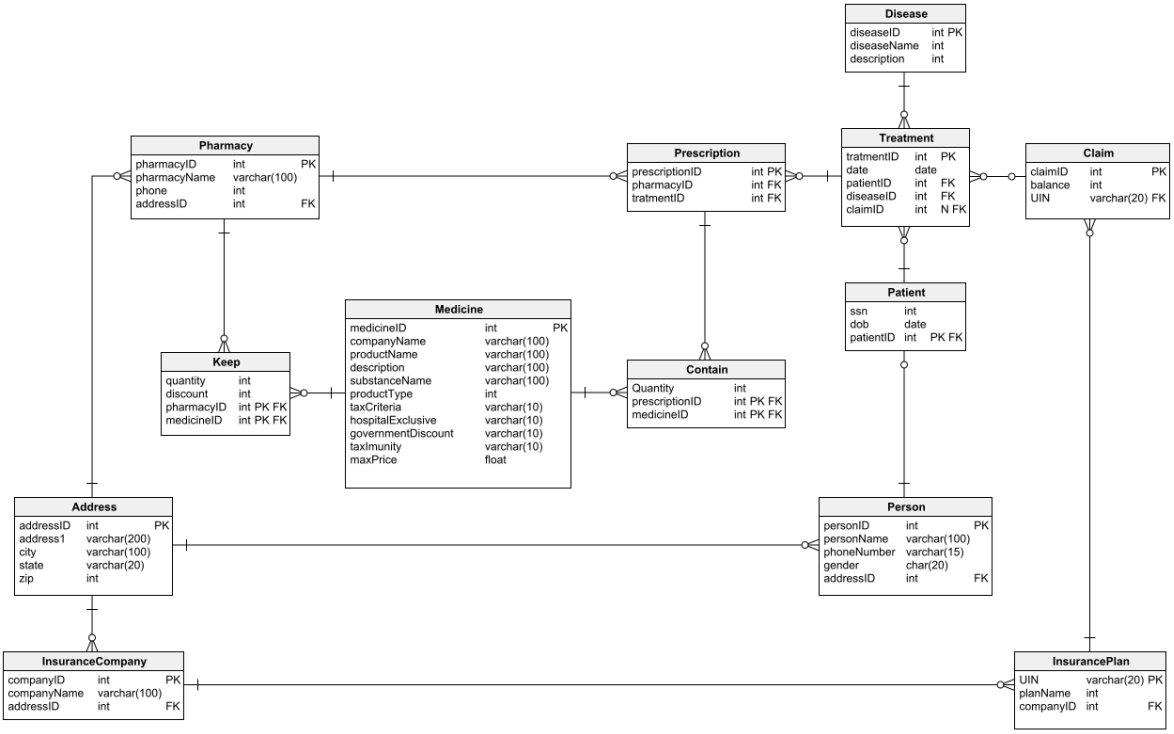
Problem Statement 5:

Praveen has requested for a report that finds for every disease how many males and females underwent treatment for each in the year 2022. It would be helpful for Praveen if the aggregation for the different combinations is found as well.

Assist Praveen to create this report.

PROJECT

Database schema:



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.

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.

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.

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.

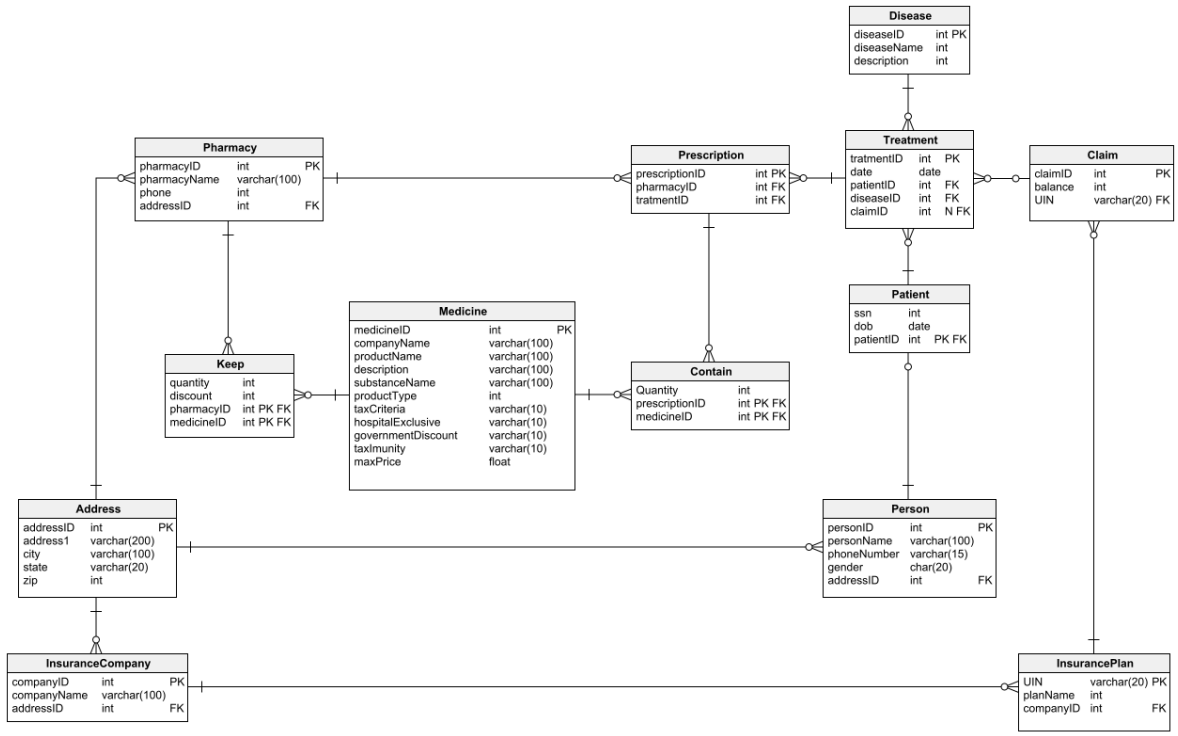
Note: discount field in keep signifies the percentage of discount on the maximum price.

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.

PROJECT

SQL grouping

Database schema:



Problem Statement 1:

The healthcare department wants a pharmacy report on the percentage of hospital-exclusive medicine prescribed in the year 2022.

Assist the healthcare department to view for each pharmacy, the pharmacy id, pharmacy name, total quantity of medicine prescribed in 2022, total quantity of hospital-exclusive medicine prescribed by the pharmacy in 2022, and the percentage of hospital-exclusive medicine to the total medicine prescribed in 2022.

Order the result in descending order of the percentage found.

Problem Statement 2:

Sarah, from the healthcare department, has noticed many people do not claim insurance for their treatment. She has requested a state-wise report of the percentage of treatments that took place without claiming insurance. Assist Sarah by creating a report as per her requirement.

Problem Statement 3:

Sarah, from the healthcare department, is trying to understand if some diseases are spreading in a particular region. Assist Sarah by creating a report which shows for each state, the number of the most and least treated diseases by the patients of that state in the year 2022.

Problem Statement 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.

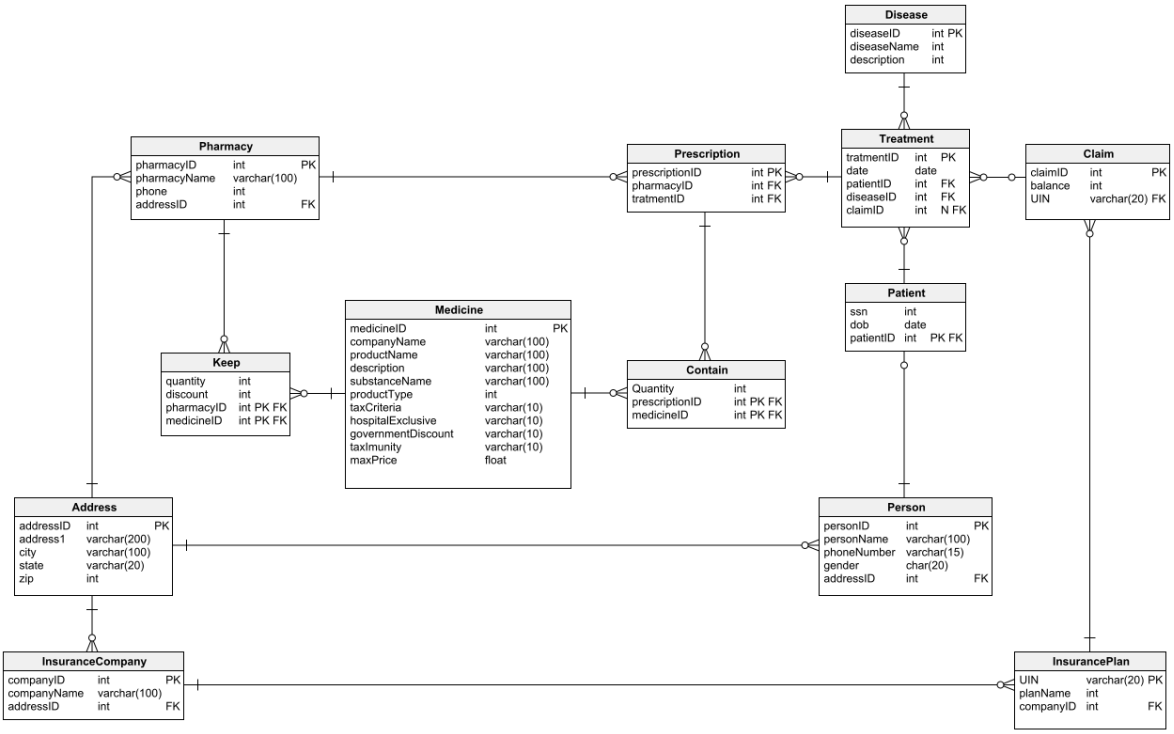
Problem Statement 5:

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.

PROJECT

SQL Stored Routines

Database schema:



Problem Statement 1:

The healthcare department has requested a system to analyze the performance of insurance companies and their plan.

For this purpose, create a stored procedure that returns the performance of different insurance plans of an insurance company. When passed the insurance company ID the procedure should generate and return all the insurance plan names the provided company issues, the number of treatments the plan was claimed for, and the name of the disease the plan was claimed for the most. The plans which are claimed more are expected to appear above the plans that are claimed less.

Problem Statement 2:

It was reported by some unverified sources that some pharmacies are more popular for certain diseases. The healthcare department wants to check the validity of this report.

Create a stored procedure that takes a disease name as a parameter and would return the top 3 pharmacies the patients are preferring for the treatment of that disease in 2021 as well as for 2022.

Check if there are common pharmacies in the top 3 list for a disease, in the years 2021 and the year 2022.

Call the stored procedure by passing the values “Asthma” and “Psoriasis” as disease names and draw a conclusion from the result.

Problem Statement 3:

Jacob, as a business strategist, wants to figure out if a state is appropriate for setting up an insurance company or not.

Write a stored procedure that finds the num_patients, num_insurance_companies, and insurance_patient_ratio, the stored procedure should also find the avg_insurance_patient_ratio and if the insurance_patient_ratio of the given state is less than the avg_insurance_patient_ratio then it Recommendation section can have the value “Recommended” otherwise the value can be “Not Recommended”.

Description of the terms used:

num_patients: number of registered patients in the given state

num_insurance_companies: The number of registered insurance companies in the given state

insurance_patient_ratio: The ratio of registered patients and the number of insurance companies in the given state

avg_insurance_patient_ratio: The average of the ratio of registered patients and the number of insurance for all the states.

Problem Statement 4:

Currently, the data from every state is not in the database, The management has decided to add the data from other states and cities as well. It is felt by the management that it would be helpful if the date and time were to be stored whenever new city or state data is inserted.

The management has sent a requirement to create a PlacesAdded table if it doesn’t already exist, that has four attributes. placeID, placeName, placeType, and timeAdded.

Description

placeID: This is the primary key, it should be auto-incremented starting from 1

placeName: This is the name of the place which is added for the first time

placeType: This is the type of place that is added for the first time. The value can either be ‘city’ or ‘state’

timeAdded: This is the date and time when the new place is added

You have been given the responsibility to create a system that satisfies the requirements of the management. Whenever some data is inserted in the Address table that has a new city or state name, the PlacesAdded table should be updated with relevant data.

Problem Statement 5:

Some pharmacies suspect there is some discrepancy in their inventory management. The quantity in the ‘Keep’ is updated regularly and there is no record of it. They have requested to create a system that keeps track of all the transactions whenever the quantity of the inventory is updated.

You have been given the responsibility to create a system that automatically updates a Keep_Log table which has the following fields:

id: It is a unique field that starts with 1 and increments by 1 for each new entry

medicineID: It is the medicineID of the medicine for which the quantity is updated.

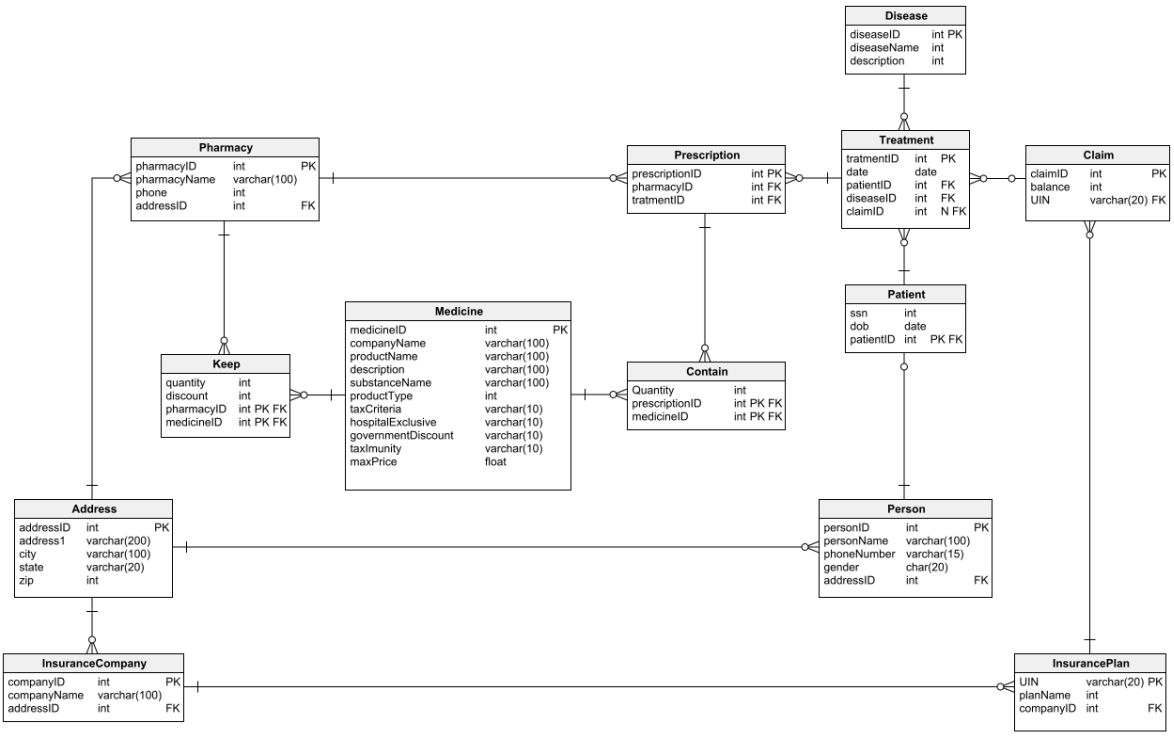
quantity: The quantity of medicine which is to be added. If the quantity is reduced then the number can be negative.

For example: If in Keep the old quantity was 700 and the new quantity to be updated is 1000, then in Keep_Log the quantity should be 300.

Example 2: If in Keep the old quantity was 700 and the new quantity to be updated is 100, then in Keep_Log the quantity should be -600.

PROJECT

Database schema:



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.

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.

Problem Statement 3: The healthcare department needs a report about insurance plans. The report is required to include the insurance plan, which was claimed the most and least for each disease. Assist to create such a report.

Problem Statement 4: The Healthcare department wants to know which disease is most likely to infect multiple people in the same household. For each disease find the number of households that has more than one patient with the same disease.

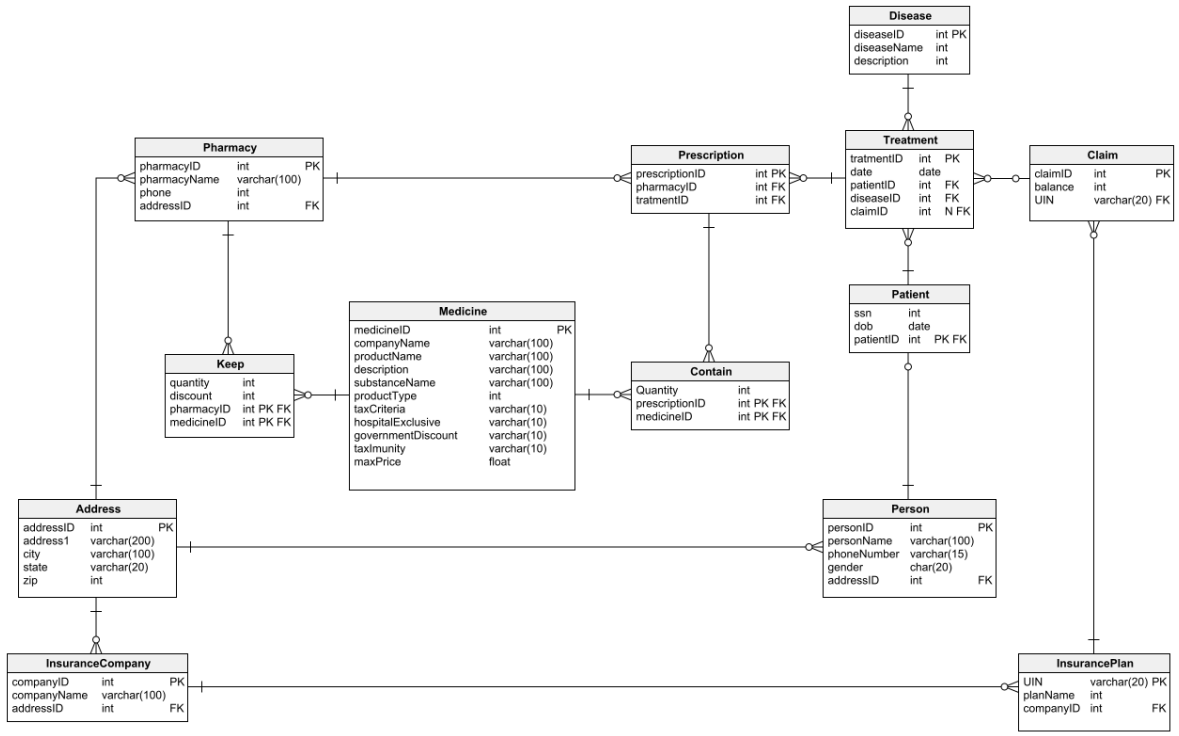
Note: 2 people are considered to be in the same household if they have the same address.

Problem Statement 5: 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.

PROJECT

SQL If-else

Database schema:



Problem Statement 1:

Insurance companies want to know if a disease is claimed higher or lower than average. Write a stored procedure that returns “claimed higher than average” or “claimed lower than average” when the diseaseID is passed to it.

Hint: Find average number of insurance claims for all the diseases. If the number of claims for the passed disease is higher than the average return “claimed higher than average” otherwise “claimed lower than average”.

Problem Statement 2:

Joseph from Healthcare department has requested for an application which helps him get genderwise report for any disease.

Write a stored procedure when passed a disease_id returns 4 columns, **disease_name, number_of_male_treated, number_of_female_treated, more_treated_gender**

Where, **more_treated_gender** is either ‘male’ or ‘female’ based on which gender underwent more often for the disease, if the number is same for both the genders, the value should be ‘same’.

Problem Statement 3:

The insurance companies want a report on the claims of different insurance plans.

Write a query that finds the top 3 most and top 3 least claimed insurance plans.

The query is expected to return the insurance plan name, the insurance company name which has that plan, and whether the plan is the most claimed or least claimed.

Problem Statement 4:

The healthcare department wants to know which category of patients is being affected the most by each disease.

Assist the department in creating a report regarding this.

Provided the healthcare department has categorized the patients into the following category.

YoungMale: Born on or after 1st Jan 2005 and gender male.

YoungFemale: Born on or after 1st Jan 2005 and gender female.

AdultMale: Born before 1st Jan 2005 but on or after 1st Jan 1985 and gender male.

AdultFemale: Born before 1st Jan 2005 but on or after 1st Jan 1985 and gender female.

MidAgeMale: Born before 1st Jan 1985 but on or after 1st Jan 1970 and gender male.

MidAgeFemale: Born before 1st Jan 1985 but on or after 1st Jan 1970 and gender female.

ElderMale: Born before 1st Jan 1970, and gender male.

ElderFemale: Born before 1st Jan 1970, and gender female.

Problem Statement 5:

Anna wants a report on the pricing of the medicine. She wants a list of the most expensive and most affordable medicines only.

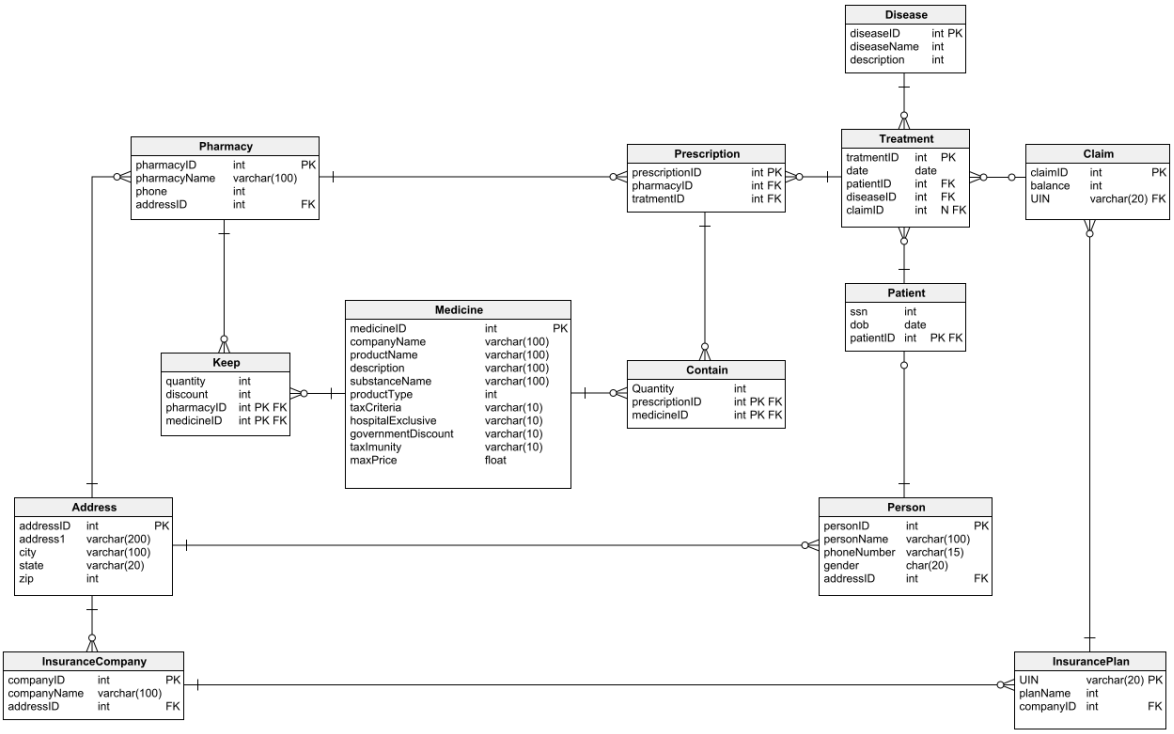
Assist anna by creating a report of all the medicines which are pricey and affordable, listing the companyName, productName, description, maxPrice, and the price category of each. Sort the list in descending order of the maxPrice.

Note: A medicine is considered to be “**pricey**” if the max price exceeds 1000 and “**affordable**” if the price is under 5. Write a query to find

PROJECT

SQL Query Optimization

Database schema:



The healthcare department attempting to use the resources more efficiently. It already has some queries that are being used for different purposes. The management suspects that these queries might not be efficient so they have requested to optimize the existing queries wherever necessary.

Given are some queries written in SQL server which may be optimized if necessary.

Query 1:

```
-- For each age(in years), how many patients have gone for treatment?
SELECT DATEDIFF(hour, dob , GETDATE())/8766 AS age, count(*) AS numTreatments
FROM Person
JOIN Patient ON Patient.patientID = Person.personID
JOIN Treatment ON Treatment.patientID = Patient.patientID
group by DATEDIFF(hour, dob , GETDATE())/8766
order by numTreatments desc;
```

Query 2:

```
-- For each city, Find the number of registered people, number of pharmacies, and
number of insurance companies.

drop table if exists T1;
drop table if exists T2;
drop table if exists T3;

select Address.city, count(Pharmacy.pharmacyID) as numPharmacy
into T1
from Pharmacy right join Address on Pharmacy.addressID = Address.addressID
group by city
order by count(Pharmacy.pharmacyID) desc;

select Address.city, count(InsuranceCompany.companyID) as numInsuranceCompany
into T2
from InsuranceCompany right join Address on InsuranceCompany.addressID =
Address.addressID
group by city
order by count(InsuranceCompany.companyID) desc;

select Address.city, count(Person.personID) as numRegisteredPeople
into T3
from Person right join Address on Person.addressID = Address.addressID
group by city
order by count(Person.personID) desc;

select T1.city, T3.numRegisteredPeople, T2.numInsuranceCompany, T1.numPharmacy
from T1, T2, T3
where T1.city = T2.city and T2.city = T3.city
order by numRegisteredPeople desc;
```

Query 3:

```
-- Total quantity of medicine for each prescription prescribed by Ally Scripts
-- If the total quantity of medicine is less than 20 tag it as "Low Quantity".
-- If the total quantity of medicine is from 20 to 49 (both numbers including) tag it
as "Medium Quantity".
-- If the quantity is more than equal to 50 then tag it as "High quantity".

select
C.prescriptionID, sum(quantity) as totalQuantity,
CASE WHEN sum(quantity) < 20 THEN 'Low Quantity'
WHEN sum(quantity) < 50 THEN 'Medium Quantity'
ELSE 'High Quantity' END AS Tag

FROM Contain C
JOIN Prescription P
on P.prescriptionID = C.prescriptionID
JOIN Pharmacy on Pharmacy.pharmacyID = P.pharmacyID
where Pharmacy.pharmacyName = 'Ally Scripts'
group by C.prescriptionID;
```

Query 4:

```
-- The total quantity of medicine in a prescription is the sum of the quantity of all
the medicines in the prescription.
-- Select the prescriptions for which the total quantity of medicine exceeds
-- the avg of the total quantity of medicines for all the prescriptions.

drop table if exists T1;


select Pharmacy.pharmacyID, Prescription.prescriptionID, sum(quantity) as
totalQuantity
into T1
from Pharmacy
join Prescription on Pharmacy.pharmacyID = Prescription.pharmacyID
join Contain on Contain.prescriptionID = Prescription.prescriptionID
join Medicine on Medicine.medicineID = Contain.medicineID
join Treatment on Treatment.treatmentID = Prescription.treatmentID
where YEAR(date) = 2022
group by Pharmacy.pharmacyID, Prescription.prescriptionID
order by Pharmacy.pharmacyID, Prescription.prescriptionID;


select * from T1
where totalQuantity > (select avg(totalQuantity) from T1);
```

Query 5:

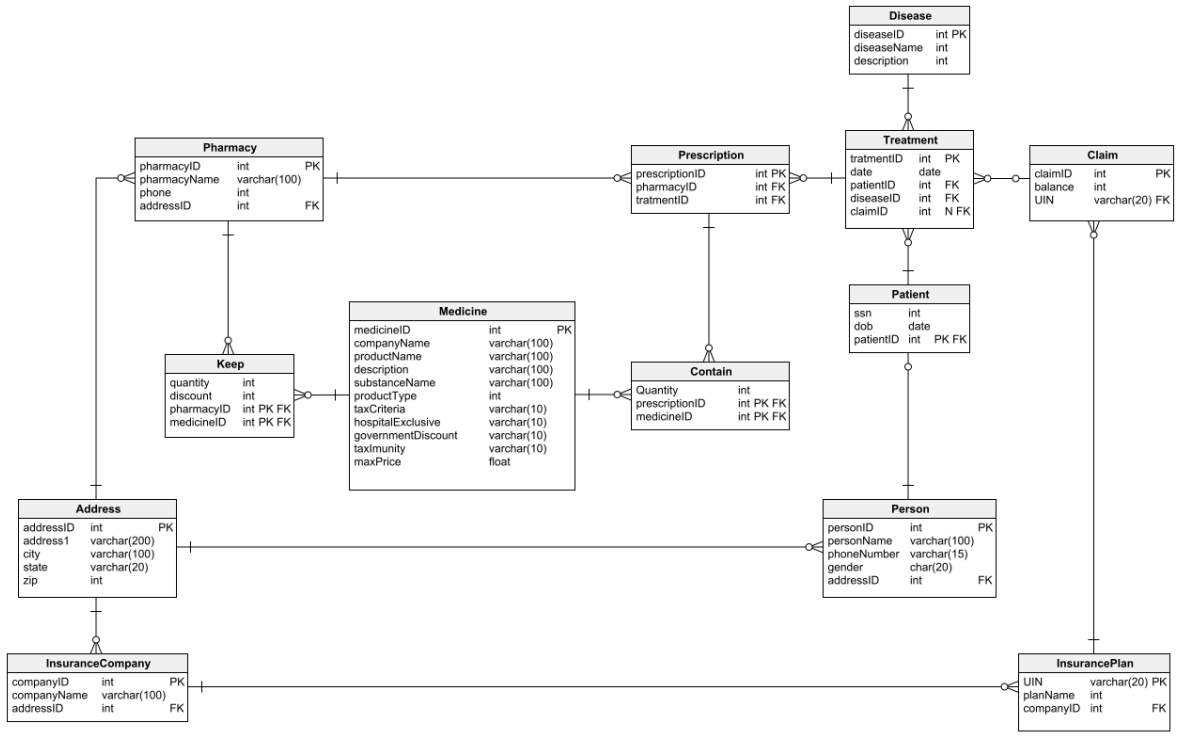
```
-- Select every disease that has 'p' in its name, and
-- the number of times an insurance claim was made for each of them.

SELECT Disease.diseaseName, COUNT(*) as numClaims
FROM Disease
JOIN Treatment ON Disease.diseaseID = Treatment.diseaseID
JOIN Claim On Treatment.claimID = Claim.claimID
WHERE diseaseName IN (SELECT diseaseName from Disease where diseaseName LIKE '%p%')
GROUP BY diseaseName;
```

PROJECT

SQL grouping

Database schema:



Problem Statement 1:

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.

Problem Statement 2:

Bharat is researching the impact of gender on different diseases, He wants to analyze if a certain disease is more likely to infect a certain gender or not.

Help Bharat analyze this by creating a report showing for every disease how many males and females underwent treatment for each in the year 2021. It would also be helpful for Bharat if the male-to-female ratio is also shown.

Problem Statement 3:

Kelly, from the Fortis Hospital management, has requested a report that shows for each disease, the top 3 cities that had the most number treatment for that disease.

Generate a report for Kelly’s requirement.

Problem Statement 4:

Brooke is trying to figure out if patients with a particular disease are preferring some pharmacies over others or not, For this purpose, she has requested a detailed pharmacy report that shows each pharmacy name, and how many prescriptions they have prescribed for each disease in 2021 and 2022, She expects the number of prescriptions prescribed in 2021 and 2022 be displayed in two separate columns.

Write a query for Brooke’s requirement.

Problem Statement 5:

Walde, from Rock tower insurance, has sent a requirement for a report that presents which insurance company is targeting the patients of which state the most.

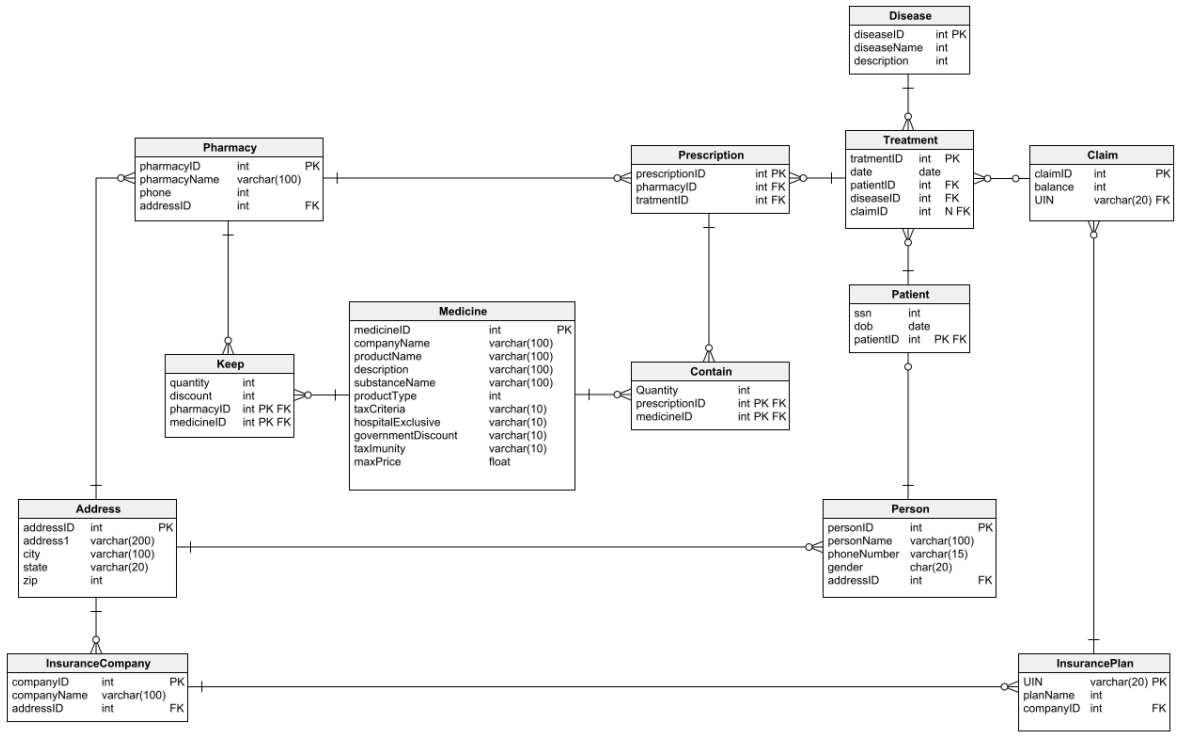
Write a query for Walde that fulfills the requirement of Walde.

Note: We can assume that the insurance company is targeting a region more if the patients of that region are claiming more insurance of that company.

PROJECT

CASE expression

Database schema:



Problem Statement 1:

"HealthDirect" pharmacy finds it difficult to deal with the product type of medicine being displayed in numerical form, they want the product type in words. Also, they want to filter the medicines based on tax criteria.

Display only the medicines of product categories 1, 2, and 3 for medicines that come under tax category I and medicines of product categories 4, 5, and 6 for medicines that come under tax category II.

Write a SQL query to solve this problem.

ProductType numerical form and ProductType in words are given by

1 - Generic,

2 - Patent,

3 - Reference,

4 - Similar,

5 - New,

6 - Specific,

7 - Biological,

8 – Dinamized

3 random rows and the column names of the Medicine table are given for reference.

Medicine (medicineID, companyName, productName, description, substanceName, productType, taxCriteria, hospitalExclusive, governmentDiscount, taxImunity, maxPrice)

12	LIBRA COMERCIO DE PRODUTOS FARMACEUTICOS LTDA	OXALIPLATINA	100 MG PO LIOFILIZADO FR/AMP X 1000 MG	NC/NI	1	I	N	N	N	2373.63
13	LIBRA COMERCIO DE PRODUTOS FARMACEUTICOS LTDA	SULBACTAM SODICO + AMPICILINA SODICA	1 G + 2 G CT FR AMP VD INC	NC/NI	4	II	N	N	N	29.59
14	LIBRA COMERCIO DE PRODUTOS FARMACEUTICOS LTDA	PACLITAXEL	6 MG/ML SOL INJ CT FR/AMP X 50 ML	NC/NI	1	I	N	N	N	4122.12

Problem Statement 2:

'Ally Scripts' pharmacy company wants to find out the quantity of medicine prescribed in each of its prescriptions.

Write a query that finds the sum of the quantity of all the medicines in a prescription and if the total quantity of medicine is less than 20 tag it as “low quantity”. If the quantity of medicine is from 20 to 49 (both numbers including) tag it as “medium quantity“ and if the quantity is more than equal to 50 then tag it as “high quantity”.

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'.

3 rows from the resultant table may be as follows:

prescriptionID	***totalQuantity***	***Tag***
1147561399	*43*	*Medium Quantity*
1222719376	*71*	*High Quantity*
1408276190	*48*	*Medium Quantity*

Problem Statement 3:

In the Inventory of a pharmacy 'Spot Rx' the quantity of medicine is considered ‘HIGH QUANTITY’ when the quantity exceeds 7500 and ‘LOW QUANTITY’ when the quantity falls short of 1000. The discount is considered “HIGH” if the discount rate on a product is 30% or higher, and the discount is considered “NONE” when the discount rate on a product is 0%.

'Spot Rx' needs to find all the Low quantity products with high discounts and all the high-quantity products with no discount so they can adjust the discount rate according to the demand.

Write a query for the pharmacy listing all the necessary details relevant to the given requirement.

*Hint: Inventory is reflected in the **Keep** table.*

Problem Statement 4:

Mack, From HealthDirect Pharmacy, wants to get a list of all the affordable and costly, hospital-exclusive medicines in the database. Where affordable medicines are the medicines that have a maximum price of less than 50% of the avg maximum price of all the medicines in the database, and costly medicines are the medicines that have a maximum price of more than double the avg maximum price of all the medicines in the database. Mack wants clear text next to each medicine name to be displayed that identifies the medicine as affordable or costly. The medicines that do not fall under either of the two categories need not be displayed.

Write a SQL query for Mack for this requirement.

Problem Statement 5:

The healthcare department wants to categorize the patients into the following category.

YoungMale: Born on or after 1st Jan 2005 and gender male.

YoungFemale: Born on or after 1st Jan 2005 and gender female.

AdultMale: Born before 1st Jan 2005 but on or after 1st Jan 1985 and gender male.

AdultFemale: Born before 1st Jan 2005 but on or after 1st Jan 1985 and gender female.

MidAgeMale: Born before 1st Jan 1985 but on or after 1st Jan 1970 and gender male.

MidAgeFemale: Born before 1st Jan 1985 but on or after 1st Jan 1970 and gender female.

ElderMale: Born before 1st Jan 1970, and gender male.

ElderFemale: Born before 1st Jan 1970, and gender female.

Write a SQL query to list all the patient name, gender, dob, and their category.