**PROJECT**

**SQL grouping**

**Database schema:**

Diagram, schematic

Description automatically generated

**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.

WITH tmt\_cnt AS

(SELECT pnt.patientid, COUNT(t.treatmentid) AS treatment\_cnt

FROM patient pnt

LEFT JOIN treatment t USING(patientid)

GROUP BY pnt.patientid)

SELECT psn.personname, FLOOR(DATEDIFF(CURRENT\_DATE(),pnt.dob)/365) AS age, tmt\_cnt.treatment\_cnt

FROM patient pnt

INNER JOIN person psn ON pnt.patientid = psn.personid

INNER JOIN tmt\_cnt USING(patientid)

ORDER BY treatment\_cnt DESC;

**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.

WITH gender\_cnt AS

(SELECT d.diseaseid, COUNT(IF(psn.gender='male',1,NULL)) AS male\_cnt, COUNT(IF(psn.gender='female',1,NULL)) AS female\_cnt

FROM disease d

INNER JOIN treatment t USING(diseaseid)

INNER JOIN patient pnt USING(patientid)

INNER JOIN person psn ON pnt.patientid = psn.personid

WHERE YEAR(t.date) = 2021

GROUP BY d.diseaseid)

SELECT \*, male\_cnt / female\_cnt AS male\_to\_female

FROM gender\_cnt

ORDER BY male\_to\_female;

**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.

WITH tmt\_cnt AS

(SELECT d.diseasename, a.state, COUNT(t.treatmentid) AS treatment\_cnt

FROM disease d

INNER JOIN treatment t USING(diseaseid)

INNER JOIN patient pnt USING(patientid)

INNER JOIN person psn ON pnt.patientid = psn.personid

INNER JOIN address a USING(addressid)

GROUP BY d.diseasename, a.state),

tmt\_cnt\_row\_num AS

(SELECT \*, ROW\_NUMBER() OVER(PARTITION BY diseasename ORDER BY treatment\_cnt DESC) AS row\_num

FROM tmt\_cnt)

SELECT diseasename, state, treatment\_cnt

FROM tmt\_cnt\_row\_num

WHERE row\_num IN (1,2,3);

**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.

SELECT pmc.pharmacyname, d.diseasename,

COUNT(IF(YEAR(t.date)=2021, 1, NULL)) AS '2021 prescription count',

COUNT(IF(YEAR(t.date)=2022, 1, NULL)) AS '2022 prescription count'

FROM pharmacy pmc

INNER JOIN prescription psc USING(pharmacyid)

INNER JOIN treatment t USING(treatmentid)

INNER JOIN disease d USING(diseaseid)

GROUP BY pmc.pharmacyname,d.diseasename

ORDER BY pharmacyname;

**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.

WITH tmt\_cnt AS

(SELECT ic.companyname, a.state, COUNT(t.treatmentid) AS treatment\_cnt

FROM treatment t

INNER JOIN patient pnt USING(patientid)

INNER JOIN person psn ON pnt.patientid = psn.personid

INNER JOIN address a USING(addressid)

INNER JOIN claim c USING(claimid)

INNER JOIN insuranceplan ip USING(uin)

INNER JOIN insurancecompany ic USING(companyid)

GROUP BY ic.companyname, a.state),

tmt\_cnt\_row\_num AS

(SELECT \*, ROW\_NUMBER() OVER(PARTITION BY companyname ORDER BY treatment\_cnt DESC) AS row\_num

FROM tmt\_cnt)

SELECT companyname, state AS 'most targeted state'

FROM tmt\_cnt\_row\_num

WHERE row\_num = 1;