**Task 1 -**

**To gain a comprehensive understanding of the factors influencing hospitalization costs a. Merge the two tables by first identifying the columns in the data tables that will help you in merging b. In both tables, add a Primary Key constraint for these columns Hint: You can remove duplicates and null values from the column and then use ALTER TABLE to add a Primary Key constraint.**

1. **Check for NULLs and Duplicates**

-- Check for NULLs

SELECT COUNT(\*) AS NullCount

FROM HospitalizationDetails

WHERE Customer\_ID IS NULL;

SELECT COUNT(\*) AS NullCount

FROM MedicalExaminations

WHERE Customer\_ID IS NULL;

-- Check for duplicates

SELECT Customer\_ID, COUNT(\*) AS freq

FROM HospitalizationDetails

GROUP BY Customer\_ID

HAVING COUNT(\*) > 1;

SELECT Customer\_ID, COUNT(\*) AS freq

FROM MedicalExaminations

GROUP BY Customer\_ID

HAVING COUNT(\*) > 1;

1. **Remove Duplicates or NULLs**

-- Remove rows with NULLs

a.

DELETE FROM HospitalizationDetails

WHERE Customer\_ID IS NULL;

b.

DELETE FROM MedicalExaminations

WHERE Customer\_ID IS NULL;

**-- Keep only one row per duplicate Customer\_ID from HospitalizationDetails table**

DELETE FROM HospitalizationDetails

WHERE Customer\_ID IN (

SELECT Customer\_ID

FROM (

SELECT Customer\_ID,

ROW\_NUMBER() OVER (PARTITION BY Customer\_ID ORDER BY Customer\_ID) AS rn

FROM HospitalizationDetails

) t

WHERE t.rn > 1

);

DELETE FROM MedicalExaminations

WHERE Customer\_ID IN (

SELECT Customer\_ID

FROM (

SELECT Customer\_ID,

ROW\_NUMBER() OVER (PARTITION BY Customer\_ID ORDER BY Customer\_ID) AS rn1

FROM MedicalExaminations

) t1

WHERE t11.n1 > 1

);

1. **Add Primary Key Constraints**

ALTER TABLE HospitalizationDetails

ADD CONSTRAINT pk\_HospitalizationDetails PRIMARY KEY (Customer\_ID);

ALTER TABLE MedicalExaminations

ADD CONSTRAINT pk\_MedicalExaminations PRIMARY KEY (Customer\_ID);

1. **Merge the Tables (INNER JOIN)**

-- Create a merged view or table

SELECT

h.Customer\_ID,

h.charges,

h.children,

h.[Hospital tier],

h.[City tier],

h.[State ID],

m.BMI,

m.HBA1C,

m.[Heart Issues],

m.[Any Transplants],

m.[Cancer history],

m.NumberOfMajorSurgeries,

m.smoker

FROM HospitalizationDetails h

JOIN MedicalExaminations m

ON h.Customer\_ID = m.Customer\_ID;

**Task 2 –**

**Retrieve information about people who are diabetic and have heart problems with their average age, the average number of dependent children, average BMI, and average hospitalization costs**

SELECT

AVG(ME.Age) AS Avg\_Age,

AVG(ME.Children) AS Avg\_Children,

AVG(ME.BMI) AS Avg\_BMI,

AVG(HD.Amount\_spent) AS Avg\_Hospitalization\_Cost

FROM

HospitalizationDetails HD

JOIN

MedicalExaminations ME

ON HD.Person\_ID = ME.Person\_ID

WHERE

ME.Diabetic = 'Yes'

AND ME.Heart\_Issues = 'Yes';

**Task 3 –**

**Find the average hospitalization cost for each hospital tier and each city level**

SELECT

Hospital\_Tier\_Label,

City\_Tier,

AVG(Amount\_spent) AS Avg\_Hospitalization\_Cost

FROM

HospitalizationDetails

GROUP BY

Hospital\_Tier\_Label,

City\_Tier

ORDER BY

Hospital\_Tier\_Label,

City\_Tier;

**Task 4 –**

**Determine the number of people who have had major surgery with a history of cancer**

SELECT

COUNT(\*) AS Number\_of\_Patients

FROM

HospitalizationDetails

WHERE

NumberOfMajorSurgeries > 0

AND Cancer\_history = 'Yes';

**Task 5 –**

**Determine the number of tier-1 hospitals in each state**

SELECT

State\_ID,

COUNT(\*) AS Tier1\_Hospital\_Count

FROM

HospitalDetails

WHERE

Hospital\_Tier\_Label = 'Tier 1'

GROUP BY

State\_ID;