- 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

```
create schema Project;
--Then import files
      use project;
      -- Hospitalisation table
     ALTER TABLE project. hospitalisation details'
      MODIFY 'Customer ID' VARCHAR(10) NOT NULL;
     ALTER TABLE project. hospitalisation details
      ADD PRIMARY KEY ('Customer ID');
      -- Medical Examinations table
     ALTER TABLE project. medical examinations
      MODIFY 'Customer ID' VARCHAR(10) NOT NULL;
      ALTER TABLE project. medical examinations
      ADD PRIMARY KEY (`Customer ID`);
```

```
CREATE TABLE merged data AS
SELECT
    h. Customer ID',
   h. Hospital tier',
   h. City tier',
   h. charges,
   h. 'year',
   h. children,
   h. State ID',
   m. BMI',
   m. Heart Issues ,
    m. Cancer history,
    m. NumberOfMajorSurgeries,
    m. HBA1C
FROM project. hospitalisation details h
JOIN project. medical examinations m
ON h. Customer ID = m. Customer ID;
```

- 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

```
'Assumption: Diabetes = HBA1C >= 6.5' AS note,

COUNT(*) AS people_count,

NULL AS avg_age, -- No age field present in CSVs

AVG(children) AS avg_dependent_children,

AVG(bmi) AS avg_bmi,

AVG(charges) AS avg_hospitalization_costs

FROM patient_master

WHERE hba1c >= 6.5 AND heart issues = 'Yes';
```



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

```
-- 3) Average hospitalization cost for each hospital tier AND each city level
SELECT
  hospital tier,
  city tier,
  AVG(charges) AS avg_charges,
  COUNT(*)
                 AS records
FROM patient_master
GROUP BY hospital tier, city tier
ORDER BY hospital tier, city tier;
  Result Grid
                                                   Export: Wrap Cell Content: IA
                  Filter Rows:
     hospital_tier
                  city_tier
                            avg_charges
                                           records
                            700.000000
                  tier - 3
                            29519.600814
     tier - 1
                  tier - 1
                            28788.457477
     tier - 1
                  tier - 2
                                          107
     tier - 1
                  tier - 3
                            31893.925676
                                          111
     tier - 2
                            11515.412928
                  tier - 1
                            11973.655344
     tier - 2
                  tier - 2
                            12101.225011
     tier - 2
                  tier - 3
                                          453
     tier - 3
                            770.380000
                                          1
     tier - 3
                  tier - 1
                           9775.389793
                                          242
     tier - 3
                            9283.427477
                                          222
                  tier - 2
                           9342.179912
     tier - 3
                  tier - 3
                                          228
```

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

```
COUNT(DISTINCT customer_id) AS people_with_major_surgery_and_cancer

FROM

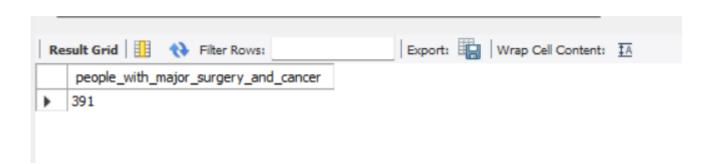
patient_master

WHERE

(number_of_major_surgeries IS NOT NULL

AND number_of_major_surgeries <> 'No major surgery')

AND cancer_history = 'Yes';
```



5. Determine the number of tier-1 hospitals in each state

```
-- 5) Number of tier-1 hospitals in each state (proxy: count tier-1 hospitalization records per state)
       The dataset has no hospital identifier, so we count tier-1 admissions per state_id.
 SELECT
   state_id,
  COUNT(*) AS tier1_records
 FROM patient master
 WHERE hospital_tier = 'tier - 1'
 GROUP BY state id
 ORDER BY state id;
Result Grid Filter Rows:
                                          Export: Wrap Cell Content: TA
  state_id
           tier1_records
          116
  R1011
  R1012
          63
  R1013
          10
  R1014
  R1015
  R1016
          8
  R1017
  R1018
  R1019
  R1023
          14
  R1024
  R 1026
```