**Had to change the column names in the table as it contains reserved keyword.**

ALTER TABLE healthcare\_dataset CHANGE `Blood Type` `Blood\_Group` VARCHAR(4);

ALTER TABLE healthcare\_dataset CHANGE `Medical Condition` `Diagnosis` VARCHAR(120);

ALTER TABLE healthcare\_dataset CHANGE `Insurance Provider` Insurance\_Provider VARCHAR(120);

ALTER TABLE healthcare\_dataset CHANGE `Billing Amount` Billing\_Amount INT;

ALTER TABLE healthcare\_dataset CHANGE `Room Number` Room\_Number INT;

ALTER TABLE healthcare\_dataset CHANGE `Admission Type` Admission\_Type VARCHAR (30);

ALTER TABLE healthcare\_dataset CHANGE `Test Results` Test\_Results VARCHAR (30);

ALTER TABLE healthcare\_dataset ADD COLUMN Admitted DATE;

UPDATE healthcare\_dataset SET Admitted = STR\_TO\_DATE(`Date of Admission`, '%m/%d/%Y');

SELECT `Date of Admission`, Admitted FROM healthcare\_dataset;

ALTER TABLE healthcare\_dataset DROP COLUMN `Date of Admission`;

ALTER TABLE healthcare\_dataset ADD COLUMN Dischared DATE;

ALTER TABLE healthcare\_dataset CHANGE Dischared Discharged DATE;

UPDATE healthcare\_dataset SET Discharged = STR\_TO\_DATE(`Discharge Date`, '%m/%d/%Y');

SELECT `Discharge Date`, Discharged FROM healthcare\_dataset;

ALTER TABLE healthcare\_dataset DROP COLUMN `Discharge Date`;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Males & Females in the dataset**

**SELECT Gender, COUNT(\*) AS count FROM healthcare\_dataset GROUP BY Gender;**

Gender count

Male 27774

Female 27726

**Gender differentiation in Admission type & Diagnosis**

**SELECT Gender, Admission\_Type, Diagnosis, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Gender, Admission\_Type, Diagnosis ORDER BY most\_visited DESC LIMIT 5;**

# Gender Admission\_Type Diagnosis most\_visited

Male Urgent Diabetes 1637

Male Urgent Asthma 1620

Female Elective Hypertension 1613

Male Elective Hypertension 1608

Female Urgent Diabetes 1592

**Who visited emergency care most?**

SELECT Gender, COUNT(\*) AS most\_visited FROM healthcare\_dataset WHERE Admission\_Type = 'Emergency' GROUP BY Gender ORDER BY most\_visited DESC LIMIT 3;

# Gender most\_visited

Female 9244

Male 9025

**Why they visited emergency care the most?**

**SELECT Gender, Diagnosis, COUNT(\*) AS most\_visited FROM healthcare\_dataset WHERE Admission\_Type = 'Emergency' GROUP BY Gender, Diagnosis ORDER BY most\_visited DESC LIMIT 3;**

# Gender Diagnosis most\_visited

Female Arthritis 1573

Male Obesity 1568

Female Cancer 1561

**Highest visits in a year**

**SELECT Admitted AS highest\_visits, COUNT(\*) AS highest\_count FROM healthcare\_dataset GROUP BY Admitted ORDER BY Highest\_count DESC LIMIT 5;**

# highest\_visits highest\_count

2024-03-16 50

2022-07-24 49

2020-10-22 49

2020-04-03 48

2021-12-28 48

**Number of visits by Blood Group**

**SELECT Blood\_Group, COUNT(\*) AS count FROM healthcare\_dataset GROUP BY Blood\_Group;**

# Blood\_Group count

O- 6877

O+ 6917

B- 6944

B+ 6945

AB- 6945

AB+ 6947

A- 6969

A+ 6956

**Which Blood Group has highest visits based on the diagnosis & admission type?**

**SELECT Blood\_Group, Admission\_Type, Diagnosis, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Blood\_Group, Admission\_Type, Diagnosis ORDER BY most\_visited DESC LIMIT 10;**

# Blood\_Group Admission\_Type Diagnosis most\_visited

A+ Urgent Diabetes 435

O+ Emergency Arthritis 426

B+ Emergency Diabetes 425

A+ Elective Cancer 423

O- Elective Hypertension 422

AB+ Elective Hypertension 421

AB- Elective Hypertension 420

A+ Emergency Obesity 419

B- Emergency Obesity 417

1. Elective Asthma 416

**Which medical condition had highest visits?**

**SELECT Diagnosis, COUNT(\*) AS count FROM healthcare\_dataset GROUP BY Diagnosis;**

# Diagnosis count

Cancer 9227

Obesity 9231

Diabetes 9304

Asthma 9185

Hypertension 9245

Arthritis 9308

**Highest visits based on the Insurance Provider**

**SELECT Insurance\_Provider, COUNT(\*) AS count FROM healthcare\_dataset GROUP BY Insurance\_Provider;**

# Insurance\_Provider count

Aetna 10913

Blue Cross 11059

UnitedHealthcare 11125

Medicare 11154

Cigna 11249

**Higest visits based on the admission type**

**SELECT Admission\_Type, COUNT(\*) AS count FROM healthcare\_dataset GROUP BY Admission\_Type;**

# Admission\_Type count

Urgent 18576

Emergency 18269

Elective 18655

**Add a column to calculate number of days each patient stayed in the hospital**

**ALTER TABLE healthcare\_dataset ADD COLUMN days\_stayed INT;**

**UPDATE healthcare\_dataset SET days\_stayed = DATEDIFF(Discharged, Admitted);**

**SELECT `Name`, days\_stayed FROM healthcare\_dataset;**

**To calculate total days stayed in the hospital for each Admission type**

**SELECT Admission\_Type, SUM(days\_stayed) AS total\_days\_stayed FROM healthcare\_dataset GROUP BY Admission\_Type ORDER BY total\_days\_stayed DESC LIMIT 3;**

# Admission\_Type total\_days\_stayed

Elective 289625

Urgent 286219

Emergency 284906

**No of days stayed based on each insurance provider**

SELECT Insurance\_Provider, SUM(days\_stayed) AS total\_days\_stayed FROM healthcare\_dataset GROUP BY Insurance\_Provider ORDER BY total\_days\_stayed DESC LIMIT 5;

# Insurance\_Provider total\_days\_stayed

Medicare 174371

Cigna 174201

UnitedHealthcare 171941

Blue Cross 171659

Aetna 168578

**Most prescribed/used Medication**

SELECT Medication, COUNT(\*) AS most\_prescribed FROM healthcare\_dataset GROUP BY Medication ORDER BY most\_prescribed DESC LIMIT 5;

# Medication most\_prescribed

Lipitor 11140

Ibuprofen 11127

Aspirin 11094

Paracetamol 11071

Penicillin 11068

**Where Lipitor was given the most?**

**SELECT Admission\_Type, COUNT(\*) AS Lipitor\_count FROM healthcare\_dataset WHERE Medication = 'Lipitor' GROUP BY Admission\_Type ORDER BY Lipitor\_count DESC LIMIT 3;**

# Admission\_Type Lipitor\_count

Urgent 3807

Elective 3667

Emergency 3666

**Where Ibuprofen was given the most?**

**SELECT Admission\_Type, COUNT(\*) AS Ibuprofen\_count FROM healthcare\_dataset WHERE Medication = 'Ibuprofen' GROUP BY Admission\_Type ORDER BY Ibuprofen\_count DESC LIMIT 3;**

# Admission\_Type Ibuprofen\_count

Urgent 3739

Elective 3734

Emergency 3654

**Where Aspirin was given the most?**

**SELECT Admission\_Type, COUNT(\*) AS Aspirin\_count FROM healthcare\_dataset WHERE Medication = 'Aspirin' GROUP BY Admission\_Type ORDER BY Aspirin\_count DESC LIMIT 3;**

# Admission\_Type Aspirin\_count

Urgent 3686

Emergency 3624

Elective 3784

**Where Penicillin was given the most?**

**SELECT Admission\_Type, COUNT(\*) AS Penicillin\_count FROM healthcare\_dataset WHERE Medication = 'Penicillin' GROUP BY Admission\_Type ORDER BY Penicillin\_count DESC LIMIT 3;**

# Admission\_Type Penicillin\_count

Elective 3733

Urgent 3684

Emergency 3651

**Which Blood Group Received Lipitor the most?**

**SELECT Blood\_Group, COUNT(\*) AS Lipitor\_count FROM healthcare\_dataset WHERE Medication = 'Lipitor' GROUP BY Blood\_Group ORDER BY Lipitor\_count DESC LIMIT 3;**

# Blood\_Group Lipitor\_count

B+ 1450

A- 1402

AB+ 1402

A+ 1385

B- 1382

O- 1379

O+ 1378

AB- 1362

**Most days stayed in the hospital based on the Diagnosis**

**SELECT Diagnosis, COUNT(\*) AS most\_days\_stayed FROM healthcare\_dataset GROUP BY Diagnosis ORDER BY most\_days\_stayed DESC LIMIT 5;**

# Diagnosis most\_days\_stayed

Arthritis 9308

Diabetes 9304

Hypertension 9245

Obesity 9231

Cancer 9227

**Average Amount Billed for each Admission Type**

**SELECT Admission\_Type, AVG(Billing\_Amount) AS Average\_Amount\_Billed FROM healthcare\_dataset GROUP BY Admission\_Type ORDER BY Average\_Amount\_Billed DESC LIMIT 3;**

# Admission\_Type Average\_Amount\_Billed

Elective 25602.2215

Urgent 25517.3655

Emergency 25497.3971

**Average Amount Billed for each Diagnosis**

**SELECT Diagnosis, AVG(Billing\_Amount) AS Average\_Amount\_Billed FROM healthcare\_dataset GROUP BY Diagnosis ORDER BY Average\_Amount\_Billed DESC LIMIT 5;**

# Diagnosis Average\_Amount\_Billed

Obesity 25805.9695

Diabetes 25638.4004

Asthma 25635.2515

Arthritis 25497.3207

Hypertension 25497.1017

**Average Amount Billed for each Provider**

**SELECT Insurance\_Provider, AVG(Billing\_Amount) AS Average\_Amount\_Billed FROM healthcare\_dataset GROUP BY Insurance\_Provider ORDER BY Average\_Amount\_Billed DESC LIMIT 5;**

# Insurance\_Provider Average\_Amount\_Billed

Medicare 25615.9959

Blue Cross 25613.0115

Aetna 25553.2905

Cigna 25525.7673

UnitedHealthcare 25389.1636

**Average amount billed for number of days stayed**

**SELECT days\_stayed, AVG(Billing\_Amount) AS Average\_Amount\_Billed FROM healthcare\_dataset GROUP BY days\_stayed ORDER BY Average\_Amount\_Billed DESC LIMIT 5;**

# days\_stayed Average\_Amount\_Billed

14 26286.2787

8 26123.3824

4 26005.4799

3 25963.8235

6 25939.1209

**So, there is no related between the amount billed and the days stayed in the hospital.**

**Emergency admission type based on Insurance Provider**

**SELECT Insurance\_Provider, COUNT(\*) AS Emergency\_count FROM healthcare\_dataset WHERE Admission\_Type = 'Emergency' GROUP BY Insurance\_Provider ORDER BY Emergency\_count DESC LIMIT 5;**

# Insurance\_Provider Emergency\_count

Cigna 3677

Medicare 3675

Aetna 3675

Blue Cross 3629

UnitedHealthcare 3613

**Urgent care admission type based on Insurance Provider**

**SELECT Insurance\_Provider, COUNT(\*) AS Urgent\_count FROM healthcare\_dataset WHERE Admission\_Type = 'Urgent' GROUP BY Insurance\_Provider ORDER BY Urgent\_count DESC LIMIT 5;**

# Insurance\_Provider Urgent\_count

Cigna 3863

Medicare 3800

UnitedHealthcare 3746

Blue Cross 3683

Aetna 3484

**Elective care admission type based on Insurance Provider**

**SELECT Insurance\_Provider, COUNT(\*) AS Elective\_count FROM healthcare\_dataset WHERE Admission\_Type = 'Elective' GROUP BY Insurance\_Provider ORDER BY Elective\_count DESC LIMIT 5;**

# Insurance\_Provider Elective\_count

UnitedHealthcare 3766

Aetna 3754

Blue Cross 3747

Cigna 3709

Medicare 3679

**Most days stayed in the hospital based on the insurance provider**

**SELECT Insurance\_Provider, COUNT(\*) AS most\_days\_stayed FROM healthcare\_dataset GROUP BY Insurance\_Provider ORDER BY most\_days\_stayed DESC LIMIT 5;**

# Insurance\_Provider most\_days\_stayed

Cigna 11249

Medicare 11154

UnitedHealthcare 11125

Blue Cross 11059

Aetna 10913

**Doctor with highest number of visits**

**SELECT Doctor, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Doctor ORDER BY most\_visited DESC LIMIT 5;**

# Doctor most\_visited

Michael Smith 27

John Smith 22

Robert Smith 22

Michael Johnson 20

James Smith 20

**SELECT Doctor, Admission\_Type, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Doctor, Admission\_Type ORDER BY most\_visited DESC LIMIT 5;**

# Doctor Admission\_Type most\_visited

David Smith Urgent 10

Michael Smith Emergency 10

John Smith Emergency 9

Michael Smith Elective 9

Christopher Smith Elective 9

**Most no. of visits based on Diagnosis**

**SELECT Diagnosis, Admission\_Type, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Diagnosis, Admission\_Type ORDER BY most\_visited DESC LIMIT 5;**

# Diagnosis Admission\_Type most\_visited

Diabetes Urgent 3229

Hypertension Elective 3221

Cancer Elective 3148

Obesity Emergency 3126

Arthritis Urgent 3117

**People with Obesity visited Emergency care the most. This might be an alarming factor.**

**Age vs Admission Type**

**SELECT Age, Admission\_Type, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Age , Admission\_Type ORDER BY most\_visited DESC LIMIT 5;**

# Age Admission\_Type most\_visited

38 Elective 336

66 Urgent 317

69 Elective 313

57 Elective 309

58 Elective 309

**Most visited Hospital**

**SELECT Hospital, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Hospital ORDER BY most\_visited DESC LIMIT 5;**

# Hospital most\_visited

LLC Smith 44

Ltd Smith 39

Johnson PLC 38

Smith Ltd 37

Smith Group 36

**Most visited hospital based on Diagnosis**

**SELECT Diagnosis, Hospital, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Diagnosis, Hospital ORDER BY most\_visited DESC LIMIT 5;**

# Diagnosis Hospital most\_visited

Arthritis Smith PLC 12

Arthritis Smith LLC 11

Hypertension LLC Smith 11

Diabetes LLC Smith 11

Obesity Ltd Smith 11

**Highest visits based on Admission Type & Diagnosis**

**SELECT Diagnosis, Hospital, Admission\_Type, COUNT(\*) AS most\_visited FROM healthcare\_dataset GROUP BY Diagnosis, Hospital, Admission\_Type ORDER BY most\_visited DESC LIMIT 5;**

# Diagnosis Hospital Admission\_Type most\_visited

Cancer Sons and Brown Elective 8

Arthritis Smith PLC Elective 7

Diabetes PLC Brown Urgent 7

Arthritis Smith LLC Elective 6

Hypertension Johnson Group Urgent 6