# PROJECT

# HEALTHCARE DATA INSIGHTS

**Name:** Harish Kumar Sahu

**Name of Project:** Health Care Data Insights

**Summary:**


- I analyzed a dataset of **55,000 US patients**, which included details like **Patient Name**, **Age**, **Gender**, **Blood Group**, and **Medical Conditions**, etc. Using **SQL** scripts, I gained a deeper understanding of the **healthcare patterns** within the data.

- The **list of tasks** I performed can be essential for **The Hospital's Chief medical director and the managers** to get insights about a particular health issue or a particular age group that is suffering from a common medical condition.


```
use healthcare;
CREATE TABLE us_healthcare (
    P_No INT NOT NULL,
    Patient_Name VARCHAR(255) NOT NULL,
    Age INT NOT NULL,
    Gender VARCHAR(255) NOT NULL,
    Blood_type VARCHAR(255) NOT NULL,
    Medical_Condition VARCHAR(255) NOT NULL,
    Date_of_Admission DATE NOT NULL,
    Doctor VARCHAR(255) NOT NULL,
    Hospital VARCHAR(255) NOT NULL,
    Insurance_Provider VARCHAR(255) NOT NULL,
    Billing_Amount FLOAT NOT NULL,
    Room_Number INT NOT NULL,
    Admission_type VARCHAR(255) NOT NULL,
    Discharge_Date DATE NOT NULL,
    Medications VARCHAR(255) NOT NULL,

    );
show databases;
show tables;
select * from us_healthcare ;
select count(*) from us_healthcare;
```

| Database |
|---|
| healthcare |
| information_schema |
| mysql |
| performance_schema |
| sys |

Result Grid | Filter Rows: | Export:

| Tables_in_healthcare |
|---|
| us_healthcare |

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

| P_No | Patient_Name | Age | Gender | Blood_type | Medical_Condition | Date_of_Admission | Doctor | Hospital | Insurance_Provider | Billing_Amount |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Bobby JacksOn | 30 | Male | B- | Cancer | 2024-01-31 | Matthew Smith | Sons and Miller | Blue Cross | 18856.3 |
| 2 | LesLie TErRy | 62 | Male | A+ | Obesity | 2019-08-20 | Samantha Davies | Kim Inc | Medicare | 33643.3 |
| 3 | DaNnY sMitH | 76 | Female | A- | Obesity | 2022-09-22 | Tiffany Mitchell | Cook PLC | Aetna | 27955.1 |
| 4 | andrEw waTtS | 28 | Female | O+ | Diabetes | 2020-11-18 | Kevin Wells | Hernandez Rogers and Vang, | Medicare | 37909.8 |
| 5 | adrIENNE bEll | 43 | Female | AB+ | Cancer | 2022-09-19 | Kathleen Hanna | White-White | Aetna | 14238.3 |
| 6 | EMILY JOHNSOn | 36 | Male | A+ | Asthma | 2023-12-20 | Taylor Newton | Nunez-Humphrey | UnitedHealthcare | 48145.1 |
| 7 | edwArD EDWaRDs | 21 | Female | AB- | Diabetes | 2020-11-03 | Kelly Olson | Group Middleton | Medicare | 19580.9 |
| 8 | CHrisTInA MARtinez | 20 | Female | A+ | Cancer | 2021-12-28 | Suzanne Thomas | Powell Robinson and Valdez, | Cigna | 45820.5 |
| 9 | JASmINe aGuIlaR | 82 | Male | AB+ | Asthma | 2020-07-01 | Daniel Ferguson | Sons Rich and | Cigna | 50119.2 |
| 10 | ChRISTopher BerG | 58 | Female | AB- | Cancer | 2021-05-23 | Heather Day | Padilla-Walker | UnitedHealthcare | 19784.6 |
| 11 | mIchElLe daniELs | 72 | Male | O+ | Cancer | 2020-04-19 | John Duncan | Schaefer-Porter | Medicare | 12576.8 |
| 12 | aaRon MARtiNeZ | 38 | Female | A- | Hypertension | 2023-08-13 | Douglas Mayo | Lyons-Blair | Medicare | 7999.59 |
| 13 | connOR HANsEn | 75 | Female | A+ | Diabetes | 2019-12-12 | Kenneth Fletcher | Powers Miller, and Flores | Cigna | 43282.3 |

us_healthcare 3 | Read Only

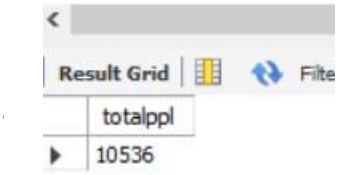
Result Grid | Filter Rows:

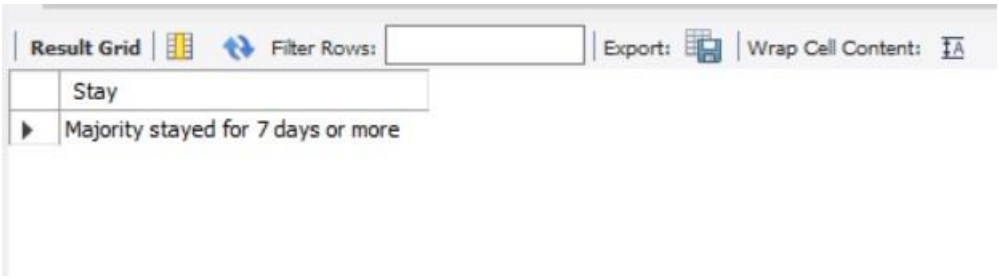
| count(*) |
|---|
| 52798 |

1. How many stayed in the hospital for less than 7 days? Did the majority stay in the hospital for less than seven days?

```
SELECT COUNT(*) AS totalppl
FROM us_healthcare
WHERE DATEDIFF(Discharge_Date,Date_of_Admission ) < 7;
```

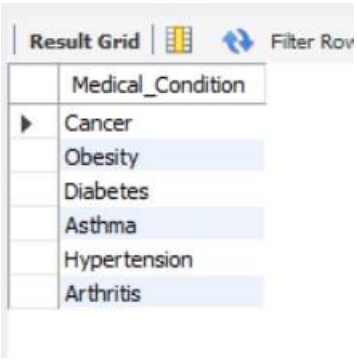
```
SELECT
    CASE
        WHEN COUNT(*) > (SELECT COUNT(P_No) / 2 FROM us_healthcare)
        THEN 'Majority stayed for less than 7 days'
        ELSE 'Majority stayed for 7 days or more'
    END AS Stay
FROM us_healthcare
WHERE DATEDIFF(Discharge_Date,Date_of_Admission) < 7;
```

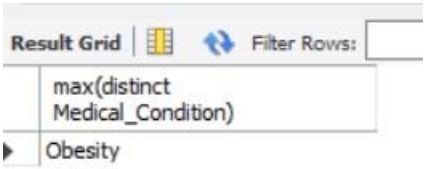



2. How many types of health problems are there, and what was the most common among them?

```
SELECT distinct Medical_Condition FROM us_healthcare ;
SELECT count(distinct Medical_Condition) FROM us_healthcare ;
SELECT max(distinct Medical_Condition) FROM us_healthcare ;
```



3. How many patients received care under each hospital, and which medical hospitals have a count of more than 35 patients?

```
SELECT max(distinct Insurance_Provider) FROM us_healthcare ;
SELECT
    Insurance_Provider,
    COUNT(*) AS Insproviders
FROM
    us_healthcare
GROUP BY
      Insurance_Provider
HAVING COUNT(*) > 10000 ;
```

Result Grid | Filter Rows:

| count(distinct Hospital) |
|---|
| 39876 |

Result Grid | Filter Rows: | Export:

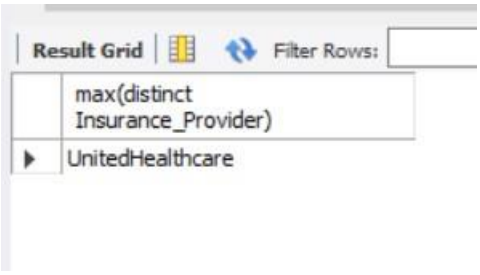
| Hospital | NumberOfPatients |
|---|---|
| Abbott and Thompson, Sullivan | 1 |
| Abbott Inc | 1 |
| Abbott Ltd | 1 |
| Abbott Moore and Williams, | 1 |
| Abbott-Castillo | 1 |
| Abbott-Coleman | 1 |
| Abbott-Ferrell | 1 |
| Abbott-Hill | 1 |
| Abbott-Jones | 1 |
| Abbott-Martin | 2 |
| Abbott-Rios | 1 |
| Abbott-Wilson | 1 |
| Abbott, Peters and Hoffman | 2 |
| Abbott, Vazquez Bautista and | 1 |

Result Grid |

| Hospital |
|---|
| Johnson PLC |
| LLC Smith |
| Ltd Smith |
| Smith Ltd |

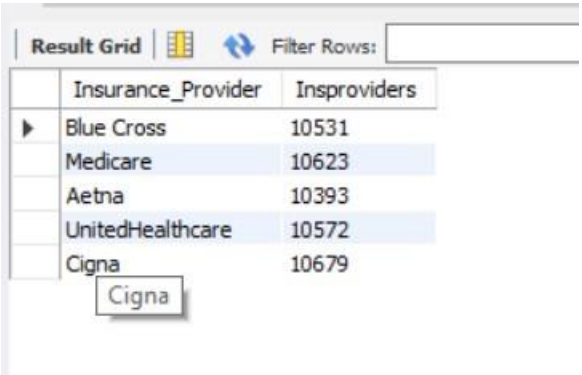
4. Which Insurance Provider has the max patients enrolled under them and which insurance providers has more than 10000 patients.

```
SELECT max(distinct Insurance_Provider) FROM us_healthcare ;
SELECT
    Insurance_Provider,
    COUNT(*) AS Insproviders
FROM
    us_healthcare
GROUP BY
      Insurance_Provider
HAVING COUNT(*) > 10000 ;
```



Result Grid | Filter Rows:

| max(distinct Insurance_Provider) |
|---|
| UnitedHealthcare |



Result Grid | Filter Rows:

| Insurance_Provider | Insproviders |
|---|---|
| Blue Cross | 10531 |
| Medicare | 10623 |
| Aetna | 10393 |
| UnitedHealthcare | 10572 |
| Cigna | 10679 |

Cigna

5. Which doctor treated the most number of patients?

```
SELECT count(distinct Doctor) FROM us_healthcare ;
SELECT max(distinct Doctor) FROM us_healthcare ;
SELECT
    Doctor,
    COUNT(*) AS Treated
FROM
    us_healthcare
GROUP BY
    Doctor
ORDER BY
    Treated DESC
LIMIT 1;
```

Result Grid | Filter Rows:

| Doctor | Treated |
|---|---|
| Michael Smith | 26 |

6. Was there any correlation between medical condition and the number of days in hospital stay?

```
SELECT
    Medical_Condition,
    AVG(DATEDIFF(Discharge_Date,Date_of_Admission)) AS AvgStayDays
FROM
    us_healthcare
GROUP BY
    Medical_Condition
order by
    AvgStayDays desc;
```

Result Grid | Filter Rows:

| Medical_Condition | AvgStayDays |
|---|---|
| Asthma | 15.6766 |
| Arthritis | 15.5542 |
| Cancer | 15.5111 |
| Obesity | 15.4581 |
| Hypertension | 15.4556 |
| Diabetes | 15.4186 |

7. Find different type of medications provided and order by most used.

```
SELECT
    DISTINCT Medications,
    COUNT(*) AS UsageCount
FROM
    us_healthcare
GROUP BY
    Medications
ORDER BY
    UsageCount DESC;
```

Result Grid | Filter Rows:

| Medications | UsageCount |
|---|---|
| Lipitor | 10600 |
| Aspirin | 10582 |
| Ibuprofen | 10566 |
| Penicillin | 10526 |
| Paracetamol | 10524 |

8. Which patients came into the hospital with an emergency) and stayed for less than the average time and give a total count in the end?

```sql
WITH AvgStay AS (
    SELECT
        AVG(DATEDIFF(Discharge_Date, Date_of_Admission)) AS AvgStayDays
    FROM
        us_healthcare
)
SELECT
    P_No,
    Patient_Name,
    Gender,
    Age,
    Admission_Type,
    DATEDIFF(Discharge_Date, Date_of_Admission) AS StayDays
FROM
    us_healthcare
WHERE
    Admission_Type = 'emergency'
    AND DATEDIFF(Discharge_Date, Date_of_Admission) < (SELECT AvgStayDays FROM AvgStay);
```

```sql
WITH AvgStay AS (
    SELECT
        AVG(DATEDIFF(Discharge_Date, Date_of_Admission)) AS AvgStayDays
    FROM
        us_healthcare
)
SELECT
    COUNT(*) AS NumberOfPatients
FROM
    us_healthcare
WHERE
    Admission_Type = 'emergency'
    AND DATEDIFF(Discharge_Date, Date_of_Admission) < (SELECT AvgStayDays FROM AvgStay);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| P_No | Patient_Name | Gender | Age | Admission_Type | StayDays |
|---|---|---|---|---|---|
| 2 | LesLie TErRy | Male | 62 | Emergency | 6 |
| 3 | DaNnY sMitH | Female | 76 | Emergency | 15 |
| 7 | edwArD EDWaRDs | Female | 21 | Emergency | 12 |
| 8 | CHrisTInA MARtinez | Female | 20 | Emergency | 10 |
| 28 | mr. KenNEth MoORE | Female | 34 | Emergency | 9 |
| 29 | MaRy hUNter | Female | 38 | Emergency | 13 |
| 56 | mR. DAVID pIERce Md | Female | 73 | Emergency | 2 |
| 57 | beThaNY MoOrE | Female | 55 | Emergency | 12 |
| 58 | MIChael MIllS | Male | 23 | Emergency | 4 |
| 63 | tRAvIs carTeR | Male | 18 | Emergency | 12 |
| 69 | MicHAEl MillEr | Male | 31 | Emergency | 4 |
| 72 | JOSe LopEz | Male | 18 | Emergency | 6 |
| 73 | kEVIn SiMmoNs Jr. | Male | 24 | Emergency | 5 |
| 78 | anRil SANTIAnO | Female | 33 | Emergency | 5 |

Result Grid | Filter Rows:

| NumberOfPatients |
|---|
| 8591 |

9. Categorize different age groups of patients along with their gender in a single query.

```
SELECT
    CASE
        WHEN Age BETWEEN 0 AND 3 THEN 'Infants'
        WHEN Age BETWEEN 4 AND 12 THEN 'School-aged Child'
        WHEN Age BETWEEN 13 AND 19 THEN 'Teenagers'
        WHEN Age BETWEEN 20 AND 40 THEN 'Adults'
        WHEN Age BETWEEN 41 AND 60 THEN 'Aged'
        ELSE 'Senior Citizens'
    END AS AgeGroup,
    Gender,
    COUNT(*) AS TotalCount
FROM us_healthcare
GROUP BY
    AgeGroup, Gender
ORDER BY
    AgeGroup ASC, Gender ASC;
```

Result Grid | Filter Rows:

| AgeGroup | Gender | TotalCount |
|---|---|---|
| Adults | Female | 8072 |
| Adults | Male | 8054 |
| Aged | Female | 7807 |
| Aged | Male | |
| Senior Citizens | Female | |
| Senior Citizens | Male | 9658 |
| Teenagers | Female | 811 |
| Teenagers | Male | 770 |

Male

10. List billing amount 1 to 10000 as low,10001 to 40000 as middling and above 40000 as high and give the count of people.

```
SELECT
    CASE
        WHEN Billing_Amount BETWEEN 1 AND 10000 THEN 'Low'
        WHEN Billing_Amount BETWEEN 10001 AND 40000 THEN 'Middling'
         ELSE 'High'
    END AS BillingCategory,
    COUNT(*) AS TotalCount
FROM
    us_healthcare
GROUP BY
    BillingCategory
ORDER BY
    BillingCategory ASC;
```

Result Grid | Filter Rows:

| BillingCategory | TotalCount |
|---|---|
| High | 10992 |
| Low | 9628 |
| Middling | 32178 |

11. Categorize patients based on blood group and distribute no of patients.

```
SELECT
    Blood_Type,
    COUNT(*) AS NumberOfPatients
FROM
    us_healthcare
GROUP BY
    Blood_Type
ORDER BY
    NumberOfPatients DESC;
```

Result Grid | Filter Rows:

| Blood_Type | NumberOfPatients |
|---|---|
| AB+ | 6625 |
| A+ | 6614 |
| A- | 6611 |
| AB- | 6605 |
| B+ | 6602 |
| O+ | 6600 |
| B- | 6584 |
| O- | 6557 |