



HEALTH DATA ANALYSIS

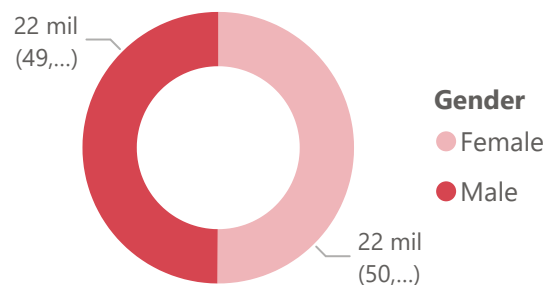
TOTAL PATIENTS TREATED

40 mil

TOTAL BILL AMOUNT

50,00 mil

COUNT BY GENDER



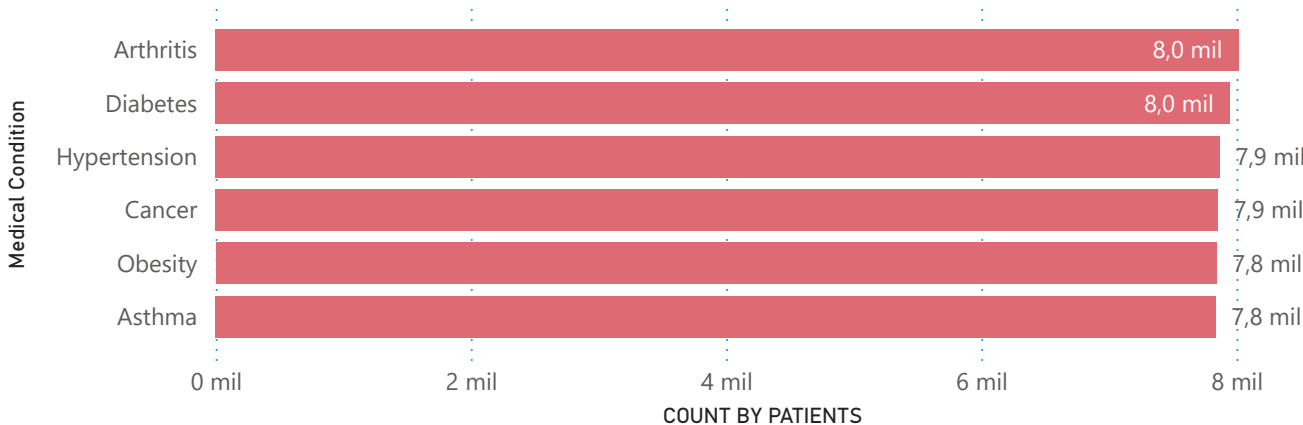
Medical Condition

- ☐ Arthritis
- ☐ Asthma
- ☐ Cancer
- ☐ Diabetes
- ☐ Hypertension
- ☐ Obesity

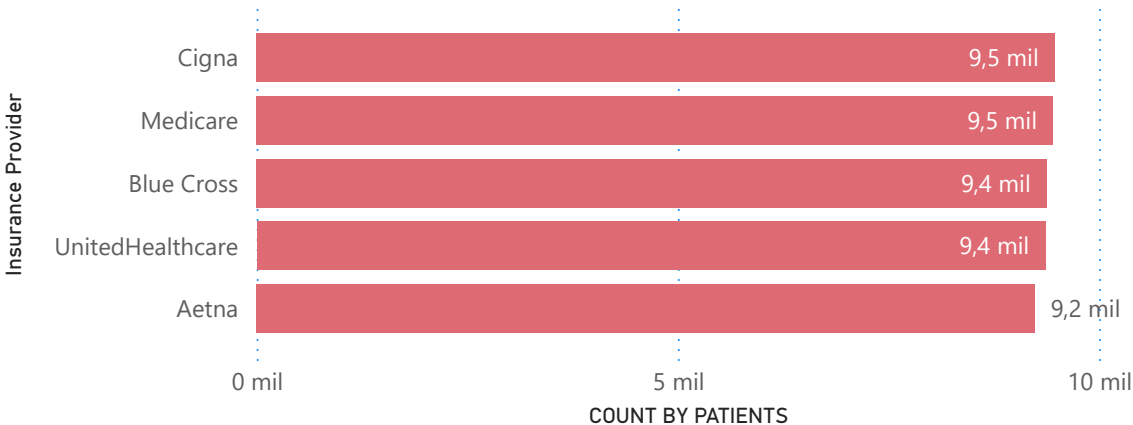
Hospital

- ☐ Abbott and Thompson, Sullivan
- ☐ Abbott Inc
- ☐ Abbott Ltd
- ☐ Abbott Moore and Williams,
- ☐ Abbott, Peters and Hoffman
- ☐ Abbott, Vazquez Bautista and
- ☐ Abbott-Castillo
- ☐ Abbott-Coleman
- ☐ Abbott-Ferrell
- ☐ Abbott-Hill

COUNT BY PATIENTS por Medical Condition



COUNT BY PATIENTS por Insurance Provider





Healthcare Data Analysis using SQL

Last Updated : 05 Feb, 2025

Healthcare data analysis plays a vital role in enhancing patient care, improving hospital efficiency and managing financial operations. By utilizing Power BI, healthcare professionals and administrators can gain valuable insights into patient demographics, medical conditions, hospital performance, and revenue trends.

In this article, we will explore a healthcare dataset and outline the step-by-step process of creating a **data analysis using SQL** then create a **Power BI dashboard** that provides essential healthcare insights. The dashboard includes key performance indicators (**KPIs**), stacked column charts, a donut chart, and interactive slicers to help stakeholders to effectively analyze and interpret healthcare data.

Healthcare Dataset

The dataset used for this analysis consists of 30,000 patient records containing detailed information on hospital admissions, treatments, and billing. The following key attributes are included in the dataset:

1. Patient Information

- **Patient_ID** – A unique identifier assigned to each patient.
- **Name** – The full name of the patient.
- **Age** – The patient's age at the time of hospital admission.
- **Gender** – The patient's gender, categorized as Male, Female, or Other.

2. Medical Information

- **Diagnosis** – The medical condition diagnosed by the doctor.
- **Treatment** – The prescribed treatment or procedure for the patient's condition.

- **Doctor** – The name of the consulting doctor treating the patient.

3. Hospital Information

- **Hospital_Name** – The name of the hospital where the patient was treated.
- **Admission_Date** – The date on which the patient was admitted to the hospital.
- **Discharge_Date** – The date on which the patient was discharged from the hospital.

4. Financial Information

- **Bill_Amount** – The total cost of treatment for the patient.

This dataset comprehensively overviews patient care, medical conditions, and hospital performance. It allows stakeholders to track trends and make data-driven decisions in the healthcare sector.

Exploratory Data Analysis Using SQL

Let's perform some **SQL** queries to get quick insights are defined below:

1. Get all patients with Kidney Disease

Retrieve the details of all patients diagnosed with Kidney Disease.

```
SELECT Patient_ID, Name, Age, Gender, Treatment, Doctor, Bill_Amount
FROM healthcare_data
WHERE Diagnosis = 'Kidney Disease';
```

Output:

	Patient_ID	Name	Age	Gender	Treatment	Doctor	Bill_Amount
▶	1001	Alice Smith	52	Male	Oxygen Therapy	Dr. Johnson	31812
	1002	John Doe	60	Male	Dialysis	Dr. Fisher	26892
	1006	Robert Brown	7	Male	Surgery	Dr. King	25859
	1011	James Taylor	98	Female	Dialysis	Dr. Johnson	5316
	1013	John Doe	98	Female	Physical Therapy	Dr. Brown	24808
	1032	Michael Wilson	1	Female	Pain Management	Dr. Fisher	4331
	1038	James Taylor	18	Male	Insulin Therapy	Dr. Johnson	4711
	1043	Michael Wilson	46	Female	Medication	Dr. Brown	47339
	1054	Emma Thomas	85	Female	Medication	Dr. Green	26766
	1064	Michael Wilson	60	Male	Oxygen Therapy	Dr. Johnson	32830
	1097	Olivia Martinez	34	Male	Radiation Therapy	Dr. Brown	30091
	1105	Emma Thomas	45	Male	Surgery	Dr. Green	31178
	1113	Olivia Martinez	68	Male	Oxygen Therapy	Dr. Evans	1951
	1123	James Taylor	34	Female	Dialysis	Dr. Adams	6191
	1147	Emma Thomas	90	Female	Physical Therapy	Dr. Harris	22931
	1170	Daniel Anderson	55	Male	Radiation Therapy	Dr. Davis	11130
	1197	Emily Davis	55	Male	Antibiotics	Dr. Davis	27862
	1198	Emma Thomas	74	Female	Pain Management	Dr. Johnson	1809
	1226	James Taylor	91	Male	Insulin Therapy	Dr. Clark	38186

output

Explanation: This query selects the relevant details of patients diagnosed with 'Kidney Disease' by filtering the rows where the Diagnosis column matches 'Kidney Disease'. The selected columns include patient ID, name, age, gender, treatment, doctor, and the bill amount.

2. Total bill amount by hospital

Calculate the total bill amount generated by each hospital.

```
SELECT Hospital_Name, SUM(Bill_Amount) AS Total_Bill
FROM healthcare_data
GROUP BY Hospital_Name;
```

Output:

	Hospital_Name	Total_Bill
▶	Lakeside Clinic	76792384
	Wellness Center	77359652
	Metro Health	77173122
	City Hospital	75215040
	Northern Health	74744981
	Sunrise Hospital	75636035
	Southview Hospital	71434260
	Central Clinic	77186535
	Greenwood Medical	75602365
	Riverbend Medical	77511715

output

Explanation: This query groups the data by Hospital_Name and calculates the total bill amount (**SUM**(Bill_Amount)) for each hospital. The result provides the sum of bill amounts for every hospital in the dataset.

3. Patients discharged after a specific date

List the patients who were discharged after January 1, 2024.

```
SELECT Patient_ID, Name, Diagnosis, Discharge_Date
FROM healthcare_data
WHERE Discharge_Date > '2024-01-01';
```

Output:

	Patient_ID	Name	Diagnosis	Discharge_Date
▶	1001	Alice Smith	Kidney Disease	2024-05-07
	1004	Emily Davis	Arthritis	2024-01-31
	1006	Robert Brown	Kidney Disease	2024-03-19
	1009	Alice Smith	Diabetes	2024-03-31
	1011	James Taylor	Kidney Disease	2024-05-22
	1015	Emily Davis	Diabetes	2024-01-24
	1021	Emily Davis	Diabetes	2024-01-31
	1022	Daniel Anderson	Heart Disease	2024-05-25
	1024	Emily Davis	Arthritis	2024-05-12
	1025	Olivia Martinez	Heart Disease	2024-03-09
	1027	Michael Wilson	Flu	2024-01-28
	1029	Robert Brown	COVID-19	2024-03-28
	1030	James Taylor	Hypertension	2024-03-05
	1035	Sophia Miller	Arthritis	2024-05-15
	1039	Alice Smith	Liver Cirrhosis	2024-01-02
	1043	Michael Wilson	Kidney Disease	2024-02-23
	1058	Emily Davis	Liver Cirrhosis	2024-03-02
	1065	Olivia Martinez	Heart Disease	2024-03-29
	1071	Michael Wilson	Diabetes	2024-01-29

output

Explanation: This query filters the data to retrieve patients whose Discharge_Date is later than January 1, 2024. The output includes patient ID, name, diagnosis, and discharge date for each relevant patient.

4. Average age of patients by treatment

Calculate the average age of patients receiving each type of treatment.

```
SELECT Treatment, AVG(Age) AS Average_Age
FROM healthcare_data
GROUP BY Treatment;
```

Output:

	Treatment	Average_Age
▶	Oxygen Therapy	50.1867
	Dialysis	50.6489
	Insulin Therapy	50.0574
	Antibiotics	51.0557
	Inhalers	51.1126
	Surgery	51.9888
	Radiation Therapy	51.0925
	Physical Therapy	50.7465
	Pain Management	50.1324
	Medication	50.9448

output

Explanation: This query groups the data by Treatment and calculates the average age (AVG(Age)) of patients receiving each type of treatment. It helps in understanding the age distribution for each treatment type.

5. Patients with the highest bill amount

Find the top 5 patients with the highest bill amounts.

```
SELECT Patient_ID, Name, Bill_Amount
FROM healthcare_data
ORDER BY Bill_Amount DESC
LIMIT 5;
```

Output:

	Patient_ID	Name	Bill_Amount
▶	18111	James Taylor	50000
	4043	James Taylor	50000
	5046	James Taylor	49998
	20465	Emily Davis	49998
	20825	John Doe	49996

output

Explanation: This query sorts the data by Bill_Amount in descending order (DESC), and limits the result to the top 5 records. This helps identify the patients who have the highest treatment costs.

Advance Data Analysis With Dashboard Creation

A well-designed [Power BI](#) dashboard helps visualize and interpret healthcare data efficiently. The dashboard consists of KPIs, stacked column charts, a donut chart, and slicers that provide actionable insights.

1. Key Performance Indicators (KPIs)

KPIs are essential for measuring hospital performance and financial health. The two primary KPIs in this dashboard are:

Total Sales Amount

- This KPI represents the total revenue generated by all hospitals.
- It is calculated by summing up the Bill_Amount across all patients.
- Monitoring total revenue helps hospital administrators track financial performance and identify revenue growth trends over time.

Total Patients

- This KPI shows the total number of patients treated.
- It is determined by counting the unique Patient_IDs in the dataset.
- Understanding patient volume is critical for resource allocation, hospital staffing, and facility expansion planning.

2. Bill Amount by Hospital Name (Stacked Column Chart)

- This visualization displays the total billing amount per hospital in a stacked column chart format.
- The X-axis represents different hospitals.
- The Y-axis shows the total revenue (sum of Bill_Amount) for each hospital.
- Each bar in the chart represents a hospital, with the height indicating the revenue generated.

This chart helps in:

- Identifying high-revenue hospitals and comparing them with lower-performing facilities.
- Understanding how different hospitals contribute to overall revenue.
- Making strategic financial decisions regarding hospital funding and operational improvements.

3. Patients by Diagnosis (Stacked Column Chart)

- This visualization shows the number of patients diagnosed with different medical conditions.
- The **X-axis** represents various diagnoses.
- The **Y-axis** shows the count of patients diagnosed with each condition.
- Each bar represents a specific medical condition, and its height indicates the number of patients affected.

This chart helps in:

- Identifying the most common medical conditions in a hospital or across multiple hospitals.
- Understanding which illnesses require more resources, specialized treatments, and trained professionals.
- Assisting in disease trend analysis and outbreak prediction.
- For example, if Hypertension and Diabetes have the highest number of patients, hospitals may focus on preventive healthcare programs to manage these conditions effectively.

4. Patients by Hospital Name (Stacked Column Chart)

- This visualization displays the distribution of patients across different hospitals using a [stacked column chart](#).
- The **X-axis** represents different hospitals.
- The **Y-axis** shows the count of patients treated at each hospital.
- Each bar represents a hospital, and its height indicates the number of patients admitted.

This chart helps in:

- Analyzing hospital capacity and patient distribution trends.
- Identifying hospitals that treat the most patients and understanding why they attract more admissions.
- Assisting in resource allocation and infrastructure development for high-demand hospitals.
- If one hospital consistently has the highest patient count, it may need more staff, beds, and facilities to manage the workload efficiently.

5. Gender Distribution (Donut Chart)

- A donut chart is used to represent the gender distribution of patients in the dataset.
- The chart is divided into three segments: Male, Female, and Other.
- Each segment represents the percentage of total patients in each gender category.

This visualization helps in:

- Understanding gender-based healthcare trends.
- Identifying whether certain medical conditions are more common in a specific gender.
- Evaluating gender inclusivity in healthcare services.
- For example, if more females are admitted for specific diagnoses, hospitals can implement targeted screening and awareness programs to address these conditions.

6. Slicers (Filters for Dynamic Analysis)

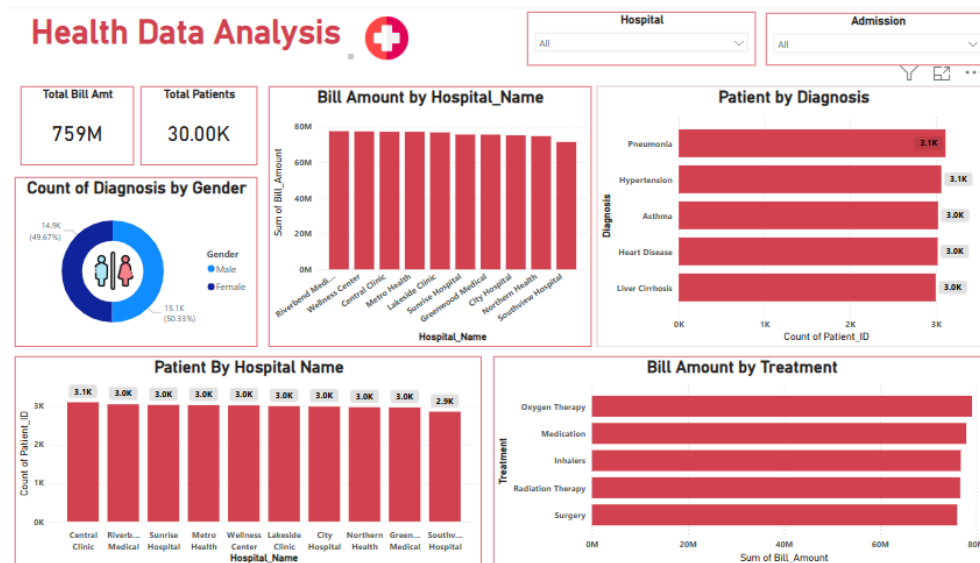
Slicers allow users to interactively filter data and explore specific insights in the dashboard. The two main slicers included are:

1. Hospital Name Slicer

- Enables users to filter the dashboard by a specific hospital.
- Helps in analyzing hospital-specific trends, financial performance, and patient demographics.

2. Admission Date Slicer

- Allows users to filter data for a specific time range.
- Useful for tracking seasonal trends in patient admissions, disease outbreaks, and financial performance.
- By using slicers, stakeholders can drill down into specific time periods or hospitals to gain granular insights into patient care and hospital operations.



Overall Dashboard

Conclusion

Overall, Power BI offers a robust platform for analyzing healthcare data by enabling stakeholders to effectively visualize key metrics and trends. The healthcare analytics dashboard discussed in this article helps hospitals, administrators, and policymakers monitor financial performance through KPIs like Total Sales Amount and Total Patients, identify high-revenue hospitals and common medical conditions using stacked column charts, track patient distribution across hospitals to optimize resource allocation, and understand gender-based healthcare patterns through a donut chart. Interactive features like slicers allow users to filter data by hospital name and admission date for deeper analysis. These insights empower healthcare professionals to enhance hospital operations, improve patient care, and drive better overall healthcare outcomes.

[Comment](#)
[More info](#)
[Advertise with us](#)
[Next Article](#)
[Data analysis using R](#)

Similar Reads

Healthcare Data Analysis using SQL

Healthcare data analysis plays a vital role in enhancing patient care, improving hospital efficiency and managing financial operations. By utilizing Power BI, healthcare professionals and administrators can gain...

7 min read

Data analysis using R

Data Analysis is a subset of data analytics, it is a process where the objective has to be made clear, collect the relevant data, preprocess the data, perform analysis(understand the data, explore insights), and then visualiz...

10 min read

Walmart Sales Data Analysis using SQL

Walmart is one of the largest retail chains globally which offers a wide range of products at competitive prices. It is known for its large network of stores and online platforms. It serves millions of customers...

8 min read

OLA Data Analysis with SQL

Have you ever thought about how ride-hailing companies manage large amounts of booking data, how to analyze customer behaviour and decide on discounts to offer? In this blog, we will do an in-depth analysis of...

10 min read

Music Store Data Analysis using SQL

Have you ever wondered how a music store keeps track of all its songs and artists? By exploring the store's dataset, we can uncover useful information, such as the total number of tracks, artists, and genre...

7 min read

Role of Big Data Analytics in Healthcare

Big data analytics in healthcare involves analyzing large data to uncover some hidden patterns and unknown correlations, market trends, customer preferences, and other useful information. In healthcare, Big data...

14 min read

SQL for Data Analysis

SQL (Structured Query Language) is an indispensable tool for data analysts, providing a powerful way to query and manipulate data stored in relational databases. With its ability to handle large datasets and...

7 min read

Data Analysis Examples

Data analysis stands as the cornerstone of informed decision-making in today's data-driven world, driving innovation and yielding actionable insights across industries. From healthcare and finance to retail and urban...

7 min read

How to Use SQL for Social Media Data Analysis.

Social media has enormous data possibilities for corporations, marketers, and researchers. SQL effectively extracts and alters data for analysis. Customer behavior and market trends were among the insights gained...

6 min read

How To Become A Healthcare Data Analyst

In today's world, the Role of a healthcare data analyst is critical in today's data-driven healthcare environment. These professionals harness statistical methods and software to process complex data sets, deriving insight...

5 min read



Corporate & Communications Address:

A-143, 7th Floor, Sovereign Corporate
Tower, Sector- 136, Noida, Uttar Pradesh
(201305)

Registered Address:

K 061, Tower K, Gulshan Vivante
Apartment, Sector 137, Noida, Gautam
Buddh Nagar, Uttar Pradesh, 201305



Advertise with us

Company

About Us
Legal
Privacy Policy
In Media
Contact Us
Advertise with us
GFG Corporate Solution

Languages

Python
Java
C++
PHP
GoLang
SQL
R Language

Placement Training Program
GeeksforGeeks Community

Android Tutorial
Tutorials Archive

DSA

Data Structures
Algorithms
DSA for Beginners
Basic DSA Problems
DSA Roadmap
Top 100 DSA Interview Problems
DSA Roadmap by Sandeep Jain
All Cheat Sheets

Data Science & ML

Data Science With Python
Data Science For Beginner
Machine Learning
ML Maths
Data Visualisation
Pandas
NumPy
NLP
Deep Learning

Web Technologies

HTML
CSS
JavaScript
TypeScript
ReactJS
NextJS
Bootstrap
Web Design

Python Tutorial

Python Programming Examples
Python Projects
Python Tkinter
Web Scraping
OpenCV Tutorial
Python Interview Question
Django

Computer Science

Operating Systems
Computer Network
Database Management System
Software Engineering
Digital Logic Design
Engineering Maths
Software Development
Software Testing

DevOps

Git
Linux
AWS
Docker
Kubernetes
Azure
GCP
DevOps Roadmap

System Design

High Level Design
Low Level Design
UML Diagrams
Interview Guide
Design Patterns
OOAD
System Design Bootcamp
Interview Questions

Interview Preparation

Competitive Programming
Top DS or Algo for CP
Company-Wise Recruitment Process
Company-Wise Preparation
Aptitude Preparation
Puzzles

School Subjects

Mathematics
Physics
Chemistry
Biology

GeeksforGeeks Videos

DSA
Python
Java
C++

Social Science
English Grammar
Commerce
World GK

Web Development
Data Science
CS Subjects

@GeeksforGeeks, Sanchhaya Education Private Limited, All rights reserved