# **Objective Questions**

**Q1. In analyzing the hospital dataset with Power BI, ensure data cleaning to address inconsistencies and missing values before further analysis.**

**Ans.** Before loading the dataset into Power BI, we can use the Power Query Editor for necessary data transformations.

* In the Hospital ER dataset, approximately **72%** of the “patient\_sat\_score” column contains **null values.**
* Ideally, addressing null and blank values should involve revisiting the data source for additional insights or acquiring a more reliable dataset.
* The most practical approach for a numerical dataset like “**patient\_sat\_score**” is to replace null values with a representative value, such as the average.
* In this instance, the null entries in the “patient\_sat\_score” column have been substituted with an **average value** of **5.**

**Q2. Assess the Average Waiting Time: Analyze the patient wait times to identify the average duration a patient spends before receiving care.**

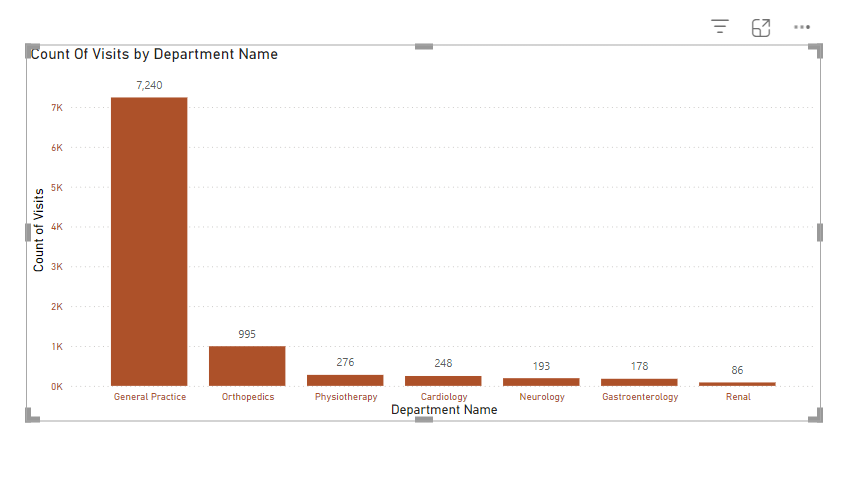
**Ans.**  I have used a DAX Measure to find the average waiting time for patients.





* The above visual shows that the average wait time for patients before receiving care is **35.26 minutes**.

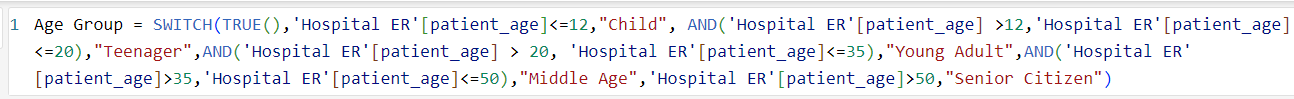
**Q3. Visits by Department Referral: Calculate the total number of visits to each department based on referrals to understand which departments are most frequently visited.**

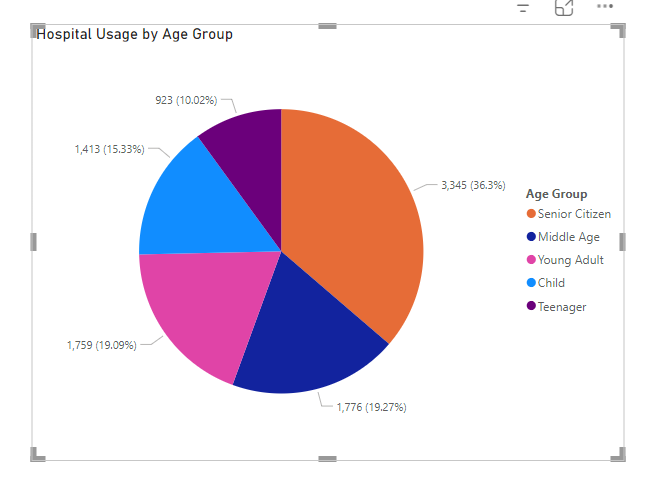
**Ans.**  I have used a bar chart in Power BI to calculate the number of visits per department like below:

* The chart above reveals that **General Practice** is the most frequently visited department, with approximately **7,240 visits**.
* This is followed by **Orthopedics**, **Physiotherapy**, and **Cardiology**, which recorded **995**, **276**, and **248 visits**, respectively.
* On the other hand, the **Renal** department received the fewest visits, with only **86 visits**, trailing behind **Gastroenterology** and **Neurology**, which saw **178** and **193 visits**, respectively.

**Q4. Patient Visits by Age Group: Segregate patient visits according to different age groups to see which demographics utilize healthcare services the most.**

**Ans.**

* I have created a new column, “Age Group,” to categorize patients into five distinct age groups.
* The categories are: “Child” (Age < 12), “Teenager” (12 < Age ≤ 20), “Young Adult” (20 < Age ≤ 35), “Middle Age” (36 < Age ≤ 50), and “Senior Citizen” (Age > 50).
* This new column was added using the “Add Column” functionality with the following DAX formula.



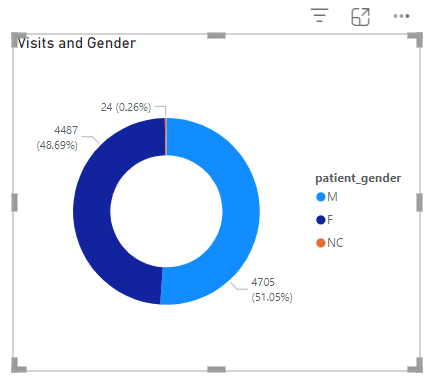
* We can observe in the above pie diagram that **Senior Citizens** utilize this hospital the most with around 3,345 patients or **36.3%** of the total patient in the hospital.
* This is followed closely by the **Middle Age** and **Young Adult** group with 1,776 people or **19.27%** and 1,759 or **19.09%** of people.
* The least usage is done by teenagers with only 923 teenage patients which is **10.02%** and then Children with 1413 total children or **15.33%.**

**Q5. Were there any Null values in the data? What would be the best way to handle these Null values and which approach have you opted for?**

**Ans.**

* The **“patient\_sat\_score”** column contained **72% blank values**, which could have a considerable impact on the accuracy of our data analysis.
* Ideally, addressing this issue involves revisiting the data source to retrieve the missing information.
* Additionally, improving the data collection process can help reduce or eliminate the occurrence of null values in the future.
* In this case, I first converted the blank entries in the column to **“null”** for consistency.
* Once all blank values were standardized as null, I replaced them with the column’s average value.
* As a result, the missing **“patient\_sat\_score”** values were filled with the average score of the column.

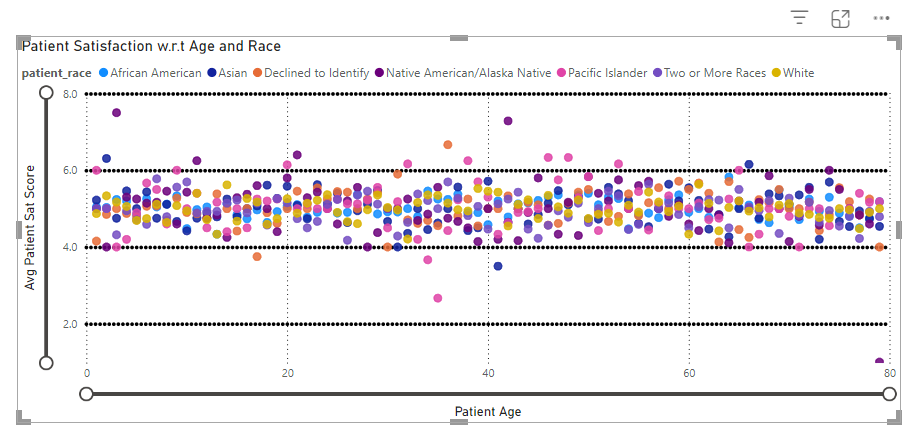
**Q6. Is there any relation between the number of visits and the Gender of the patients?**



**Ans.**

* The donut chart above shows that **Male** and **Female** patients have a nearly equal share of visits.
* Male patients account for **4,705 visits** (51.05%), while female patients contribute **4,487 visits** (48.69%) of the total.
* A small proportion of patients, categorized as **“Non-Confirmed (NC)”** because they either chose not to disclose their gender or do not align with the two primary gender categories, represent **0.26%** of the visits, totaling **26 visits**.

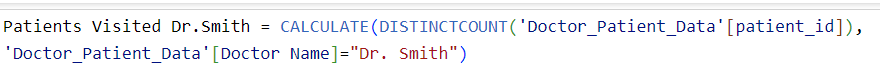
**Q7. Average Satisfaction by Demographics: Determine the relationship between patient satisfaction scores, their age groups, and racial backgrounds to pinpoint areas for improvement in patient experience.**

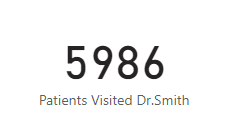
**Ans.**  I have used a scatter plot to observe a relation between patient satisfaction score, their age and racial background.

* The patient **satisfaction scores** consistently range between **4.0** and **6.0** across all age groups.
* Similarly, patients from various racial backgrounds also maintain satisfaction scores within this range, regardless of age.
* This indicates that satisfaction scores remain unaffected by variations in age or racial background.

**Q8. Using ‘Calculate’ and a row iteration DAX function calculate the total number of patients who have visited Dr. Smith**

**Ans.** I have used the COUNTDISTINCT() DAX function inside CALCULATE() and given the filter as *“Doctor Name = “Dr. Smith”*

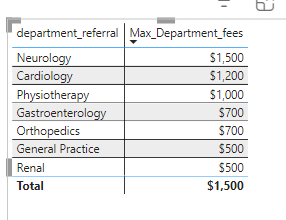


* We can visualize the above DAX measures output with help of a card visual
* We can see that 5,986 patients have visited Dr. Smith in our dataset.

**Q9. Which department is charging the highest appointment fees in general? Use an aggregation DAX function to solve this question.**

**Ans.** I have used the following aggregation DAX function to calculate the highest appointment fees.



I have further used a visual matrix and added departments to see maximum appointment fees for each department.



* From the matrix above, we can see that the highest appointment fee of **$1,500** is charged by the **Neurology Department.**
* The Cardiology Department follows in second place with $1,200, while Physiotherapy ranks third with $1,000.
* **Renal** and **General Practice** departments charge the lowest fees at **$500**, slightly below Gastroenterology and Orthopedics, which charge $700.

**Q10. Create a tabular visualization in the Report view which consists of Month-wise total visits in the hospital. Add a third column in the table that consists of the previous month’s total visits for each month’s row. Also, include a column that states whether the visits in a month are greater than that of the previous month's visits.**

**Ans.**

* We can use a tabular visualization alongside calculations for visual function to resolve the above query.
* First, we will create an additional column to extract months from the date column. We can use the following DAX query to add a new calculated column in our model.

**Month = MONTH('Hospital ER'[date])**

* Then we can use “Calculations for Visual” on our table and do the following calculations to add a column with running total and trend of visits per month.

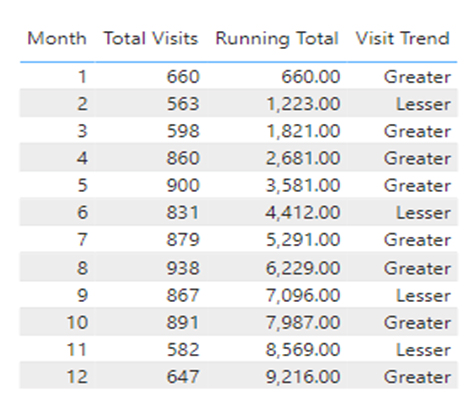
**Running Total = RUNNINGSUM([Total Visits])**

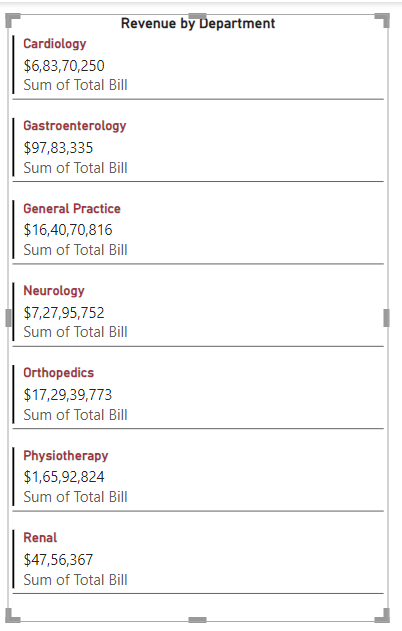
* For adding a table to show visit trends, we will first create a hidden calculation that subtracts the current total visit value from the previous months total visit value.

**Versus previous = [Total Visits] - PREVIOUS([Total Visits])**

* We can use the above hidden calculation to create our Visit Trends column in the below DAX query.

**Visit Trend = SWITCH(TRUE(),[Versus previous]>0,"Greater",[Versus previous]<0,"Lesser")**



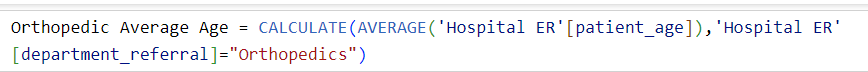
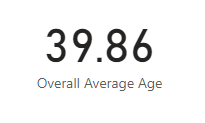
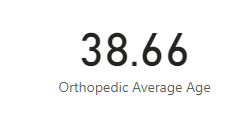
**Q11. The hospital's managing director seeks to evaluate the revenue of each department to understand how much revenue is generated by each.**

**Ans.**

* The multi-row cards visual highlights the revenue contributions of various hospital departments.
* **Cardiology** has generated revenue amounting to **$68,370,250.**
* **Gastroenterology** and **General Practice have** contributed **$9,783,335** and **$164,070,816,** respectively**.**
* **Renal** and **Physiotherapy** have brought in **$4,756,367** and **$16,592,824**, while **Orthopedics** and **Neurology** have revenues of **$172,939,773** and **$72,795,752**, respectively.

**Q12. Calculate the average age of the patients who visit the Orthopedics department. Will the approach used to calculate this metric be different if the requirement had been all departments’ average age?**

**Ans.**

* For calculating the average age of patients who visit the Orthopedics Department we make use of the CALCULATE() DAX function along with AVERAGE() and filter on the Department.
* However, if we calculate the overall average age including all the departments then we need not use the CALCULATE() DAX function and use the AVERAGE() on age directly.
* Below I have used the card visual for visualizing both the measures.
* We can see that the average age for patients in the Orthopedics Department is 38.66 meanwhile the overall average is 39.86

**Q13. Were there any data format issues in the data, and if there were/are how you handle them?**

**Ans.**

* There were data format issues in the Date column of the "Hospital ER" table, as well as in the Appointment Fees and Total Bill columns in the "Doctor Patient Data" table.
* To resolve this, we selected the respective columns and updated their Data Type through the "Structures" tab in Table View.
* For the Appointment Fees and Total Bill columns, we specifically changed the data type from Whole Number to Currency.
* Considering that Columbia is located in the US, we assumed the monetary values in the dataset are in USD and formatted both columns accordingly to reflect USD currency.

**Q14. When we add a column in Power Query what’s the code that comes in M language in the formula bar? What do you know about M-query?**

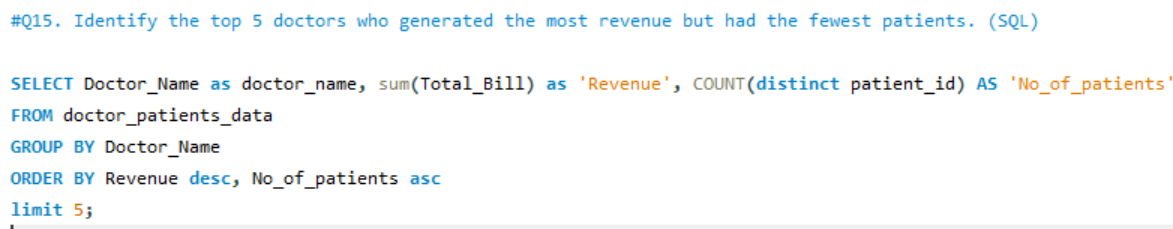
**Ans:**

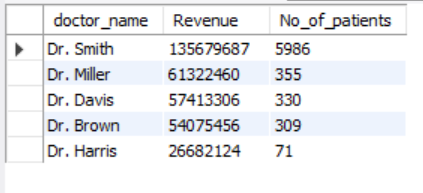
* The M-query language is used to manipulate tables and perform transformations on datasets in Power BI.
* Whenever columns are added, removed, or modified in a dataset or table, the corresponding M-query is generated and displayed in the formula bar.
* The Power Query UI automatically creates M-query code as you perform actions on your data.
* For example, when you add a column in Power Query UI, the following M-query code may appear in the formula bar:

**= Table.AddColumn(#"Replaced Value", "Full Name", each Text.Combine({[patient\_first\_initial], " ", [patient\_last\_name]}), type text)**

* In this example, the M-query is used to merge the patient\_first\_initial and patient\_last\_name columns into a new column called Full Name. This was achieved using the Column from Examples feature in the Power Query UI.

**Q15. Identify the top 5 doctors who generated the most revenue but had the fewest patients. (SQL)**

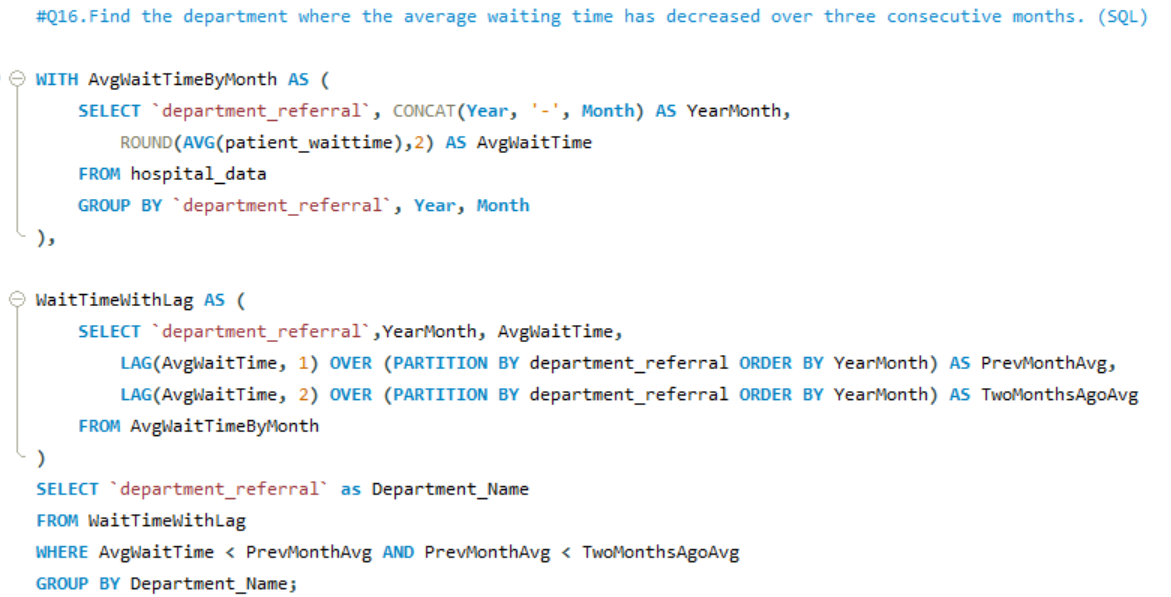
**Ans:**

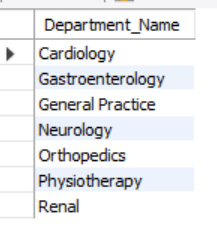


**Output:**

* **Dr. Harris** earned **26,682,124** from just **71 patients**, showing the highest revenue per patient among the group.
* **Dr. Brown** generated **54,075,456** revenue from **309 patients**, placing them among the top for efficiency.
* **Dr. Davis** earned **57,413,306** from **330 patients**, maintaining a high revenue-to-patient ratio.
* **Dr. Miller** generated revenue of **61,322,460** from just **355 patients**, suggesting a high revenue per patient and significant efficiency.
* **Dr. Smith** has the highest revenue of **135,679,687** and managed **5,986 patients**, indicating efficient revenue generation with a relatively high patient count.

**Q16. Find the department where the average waiting time has decreased over three consecutive months. (SQL).**

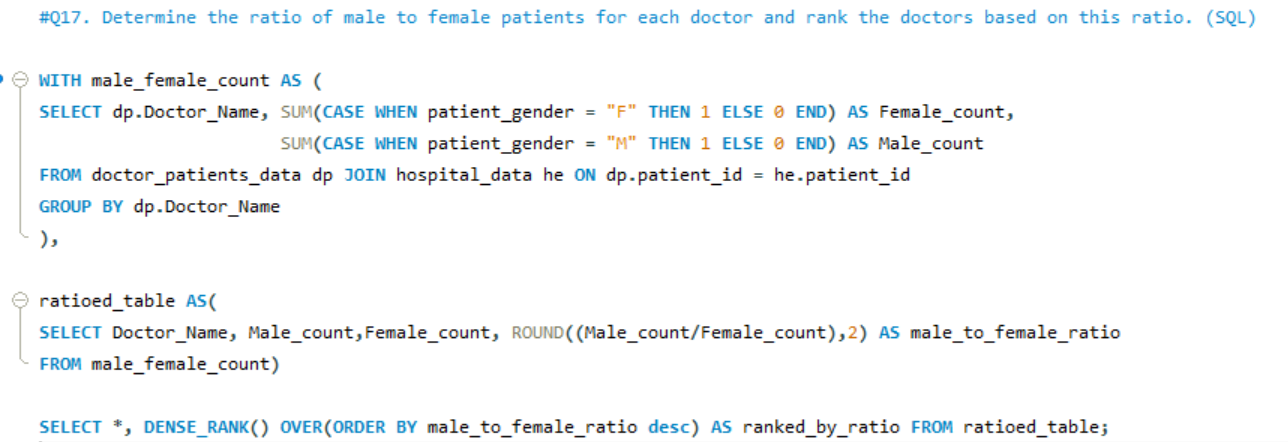
**Ans:** I have used the below query to identify the same. The departments with consistently decreasing average waiting times over three consecutive months were identified. This indicates improved efficiency and better patient flow management, contributing to enhanced service quality.

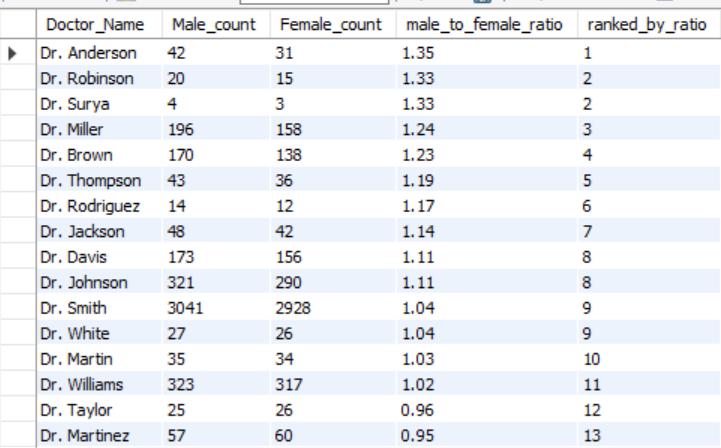


**Output:**

**Q17.Determine the ratio of male to female patients for each doctor and rank the doctors based on this ratio. (SQL)**

**Ans.** I have used the below query to identify the same. The ratio of male to female patients for each doctor is calculated and used to rank doctors in descending order based on this ratio. This analysis provides insights into patient demographics for each doctor, helping to identify trends or biases in patient distribution.

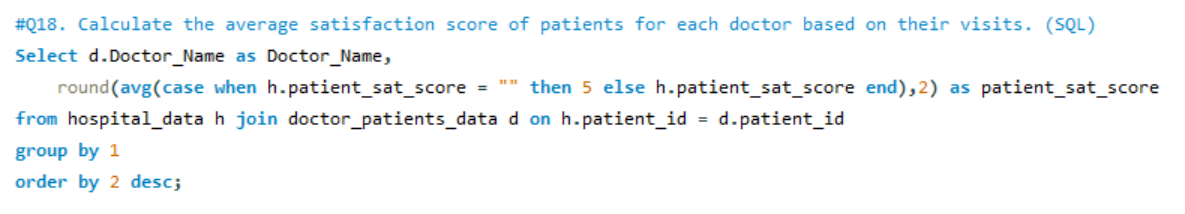


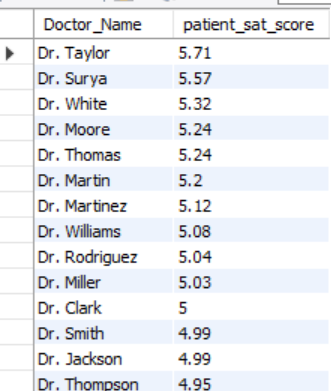


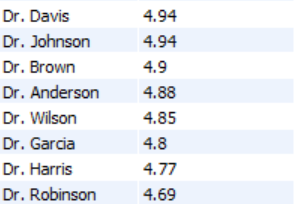
**Output:**

**Q18. Calculate the average satisfaction score of patients for each doctor based on their visits. (SQL)**

**Ans.** I have used the following query for the above need. The average patient satisfaction score for each doctor is calculated, considering visits where missing scores are replaced with a default value of 5. This provides an overall measure of how patients rate their experience with each doctor, ranked by the highest satisfaction scores.

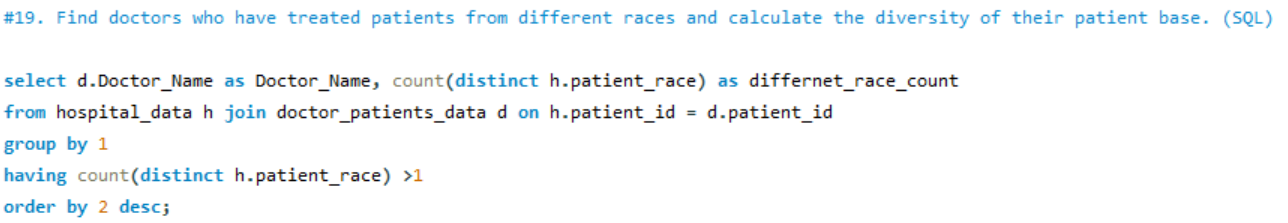


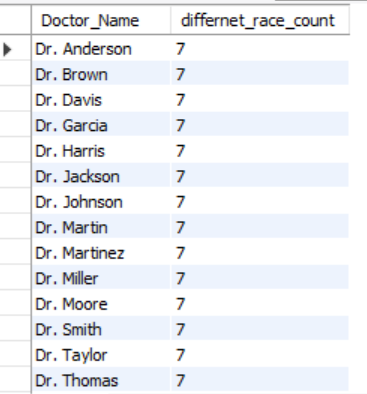


**Output:**

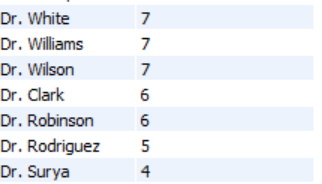
**Q19. Find doctors who have treated patients from different races and calculate the diversity of their patient base. (SQL)**

**Ans.** The ratio of total bills generated by male patients to those generated by female patients is calculated for each department. This analysis offers insights into the financial contributions of each gender, enabling the identification of spending trends across various departments.



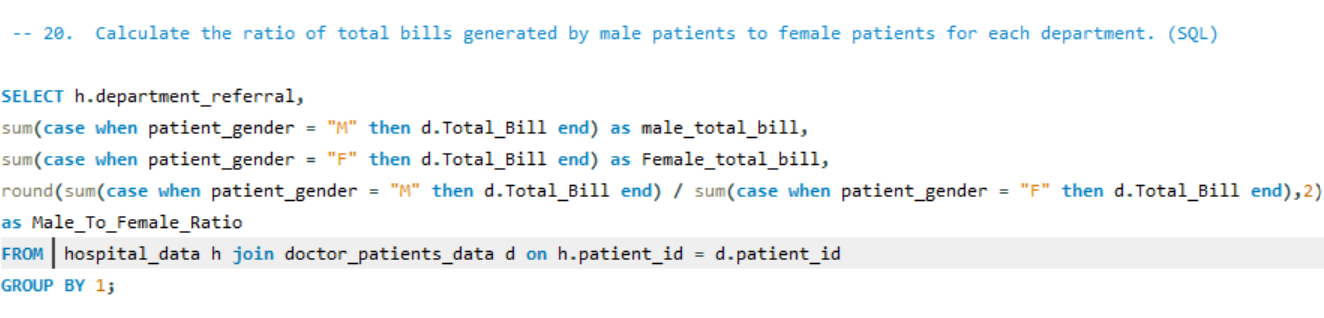


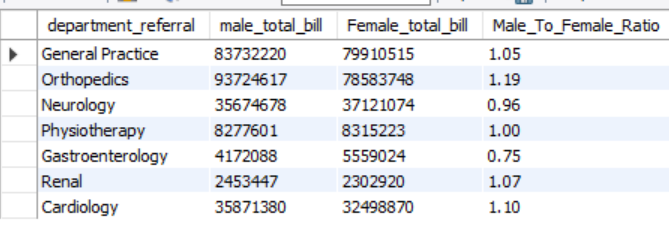
**Output:**



**Q20. Calculate the ratio of total bills generated by male patients to female patients for each department. (SQL)**

**Ans.** The ratio of total bills generated by male patients to female patients is calculated for each department. This provides insights into the financial contribution from each gender, helping to identify trends in patient spending across different departments.

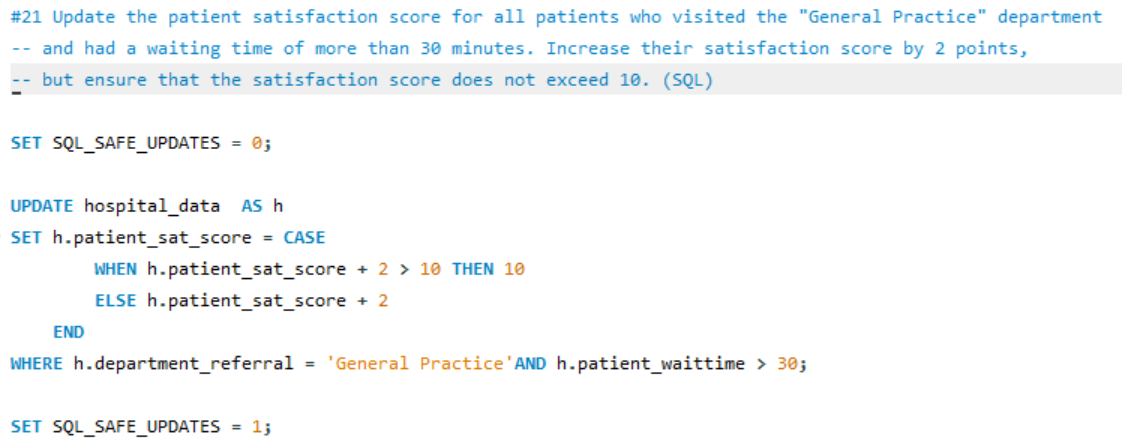




**Output:**

**Q21. Update the patient satisfaction score for all patients who visited the "General Practice" department and had a waiting time of more than 30 minutes. Increase their satisfaction score by 2 points, but ensure that the satisfaction score does not exceed 10. (SQL)**

**Ans:** In the "General Practice" department, patients with waiting times longer than 30 minutes have their satisfaction scores increased by 2 points. If the adjusted score surpasses 10, it is limited to a maximum of 10. This approach ensures fairness while keeping scores within the allowed range.



# 

# 

# **Subjective Questions**

**Q1. What is the relation between patient wait time and satisfaction scores?**

**Ans.**

* A quick measure was created in Power BI, with department\_referral set as the category, patient\_sat\_score as Measure1, and patient\_waittime as Measure2.
* The DAX formula was automatically generated to compute the correlation between the two measures.
* A Gauge Chart was used to visualize the correlation, with the range configured from 0 to 2. The observed correlation value was 1.
* In Power BI, a quick measure was created using **department\_referral** as the category, **patient\_sat\_score** as Measure1, and **patient\_waittime** as Measure2.
* A DAX formula was automatically generated to calculate the correlation between these two measures.
* The correlation value was visualized using a Gauge Chart, with the range set from 0 to 2. The resulting correlation value was 1.
* **Insight:** The analysis indicated a moderate positive correlation of 1, suggesting that patient wait times influence satisfaction scores, though not significantly.  
  Departments that minimize wait times or effectively manage patient expectations may achieve higher satisfaction levels. However, factors such as service quality and communication likely have a substantial impact on shaping patient perceptions.
* **DAX Formula:**

Correlation of patient\_sat\_score and patient\_waittime =

VAR \_\_CORRELATION\_TABLE = VALUES('Hospital ER'[department\_referral])

VAR \_\_COUNT =

COUNTX(

KEEPFILTERS(\_\_CORRELATION\_TABLE),

CALCULATE(

COUNTA('Hospital ER'[patient\_sat\_score])

\* SUM('Hospital ER'[patient\_waittime])

)

)

VAR \_\_SUM\_X =

SUMX(

KEEPFILTERS(\_\_CORRELATION\_TABLE),

CALCULATE(COUNTA('Hospital ER'[patient\_sat\_score]))

)

VAR \_\_SUM\_Y =

SUMX(

KEEPFILTERS(\_\_CORRELATION\_TABLE),

CALCULATE(SUM('Hospital ER'[patient\_waittime]))

)

VAR \_\_SUM\_XY =

SUMX(

KEEPFILTERS(\_\_CORRELATION\_TABLE),

CALCULATE(

COUNTA('Hospital ER'[patient\_sat\_score])

\* SUM('Hospital ER'[patient\_waittime]) \* 1.

)

)

VAR \_\_SUM\_X2 =

SUMX(

KEEPFILTERS(\_\_CORRELATION\_TABLE),

CALCULATE(COUNTA('Hospital ER'[patient\_sat\_score]) ^ 2)

)

VAR \_\_SUM\_Y2 =

SUMX(

KEEPFILTERS(\_\_CORRELATION\_TABLE),

CALCULATE(SUM('Hospital ER'[patient\_waittime]) ^ 2)

)

RETURN

DIVIDE(

\_\_COUNT \* \_\_SUM\_XY - \_\_SUM\_X \* \_\_SUM\_Y \* 1.,

SQRT(

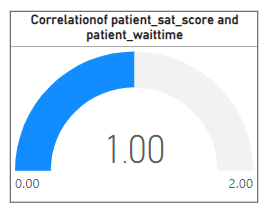
(\_\_COUNT \* \_\_SUM\_X2 - \_\_SUM\_X ^ 2)

\* (\_\_COUNT \* \_\_SUM\_Y2 - \_\_SUM\_Y ^ 2)

)

)

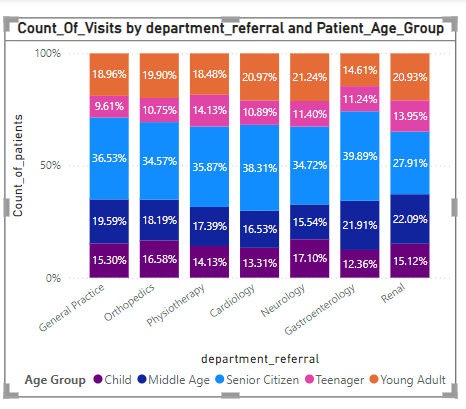
* **Result :** This analysis highlights the need to balance wait times and other quality-of-service factors to improve overall patient satisfaction.

**Visualization:** 

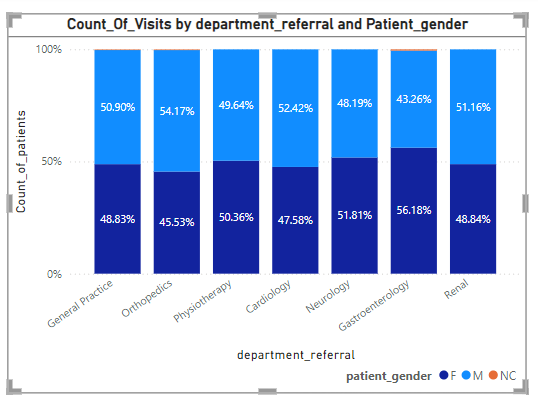
**Q2. How do patient demographics affect the frequency of visits to different departments?**

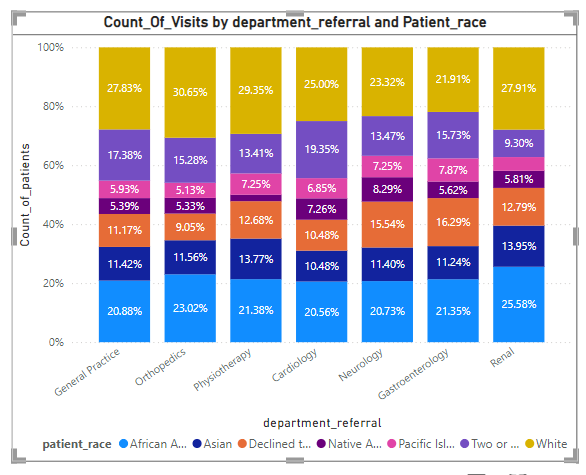
**Ans.**

* Prepare the data with columns for Department, Patient Count, and demographics **(Age Group, Race, Gender)**.
* Created a **Stacked Column Chart** with Department on the x-axis and Patient Count on the y-axis.
* This analysis aims to examine the impact of patient demographics on visit frequencies across various departments. A Stacked Column Chart is utilized to visualize how different demographic groups contribute to the total patient count in each department.
* This approach provides insights into whether specific departments are favored by certain demographic groups and identifies potential patterns or disparities in departmental usage based on patient characteristics.
* **Insight:** Demographics do not significantly influence the number of visitors to various departments, as the number of visits is primarily determined by the patient’s health condition and the hospital's location. The demographics of each department are not directly impacted.



**Visualizations:**





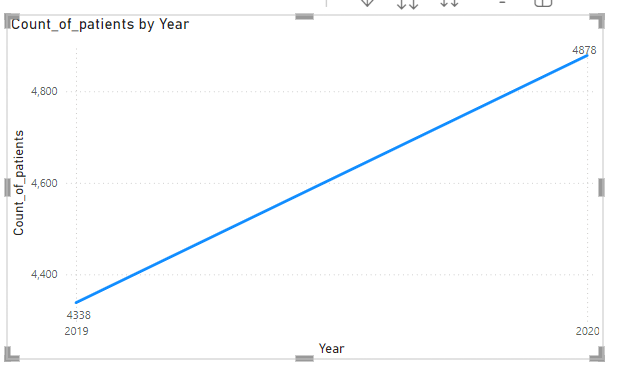
**Q3. Is there a noticeable trend in the volume of patient visits throughout the year?**

**Ans.**

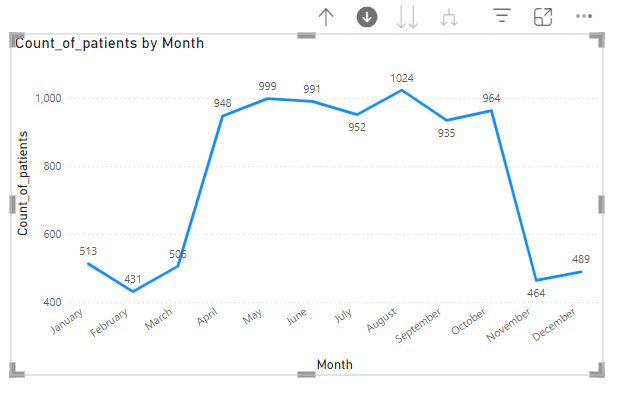
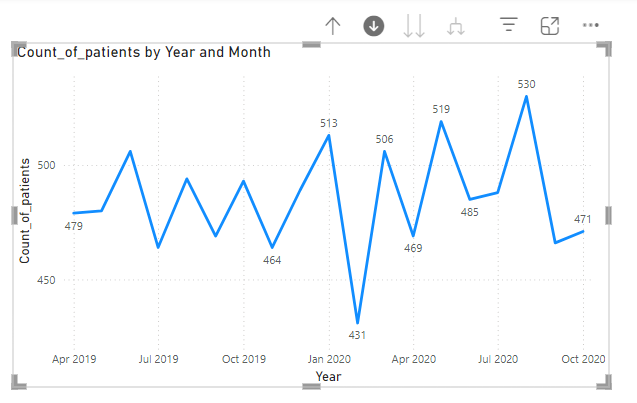
* Utilized a **Line Chart** to display the trend in patient visit volumes over time.
* Incorporate a hierarchy of **Year and Month** Name on the **x-axis** to enable drill-up and drill-down functionality, allowing analysis at various time granularities.
* Placed the **Count\_of\_Patients** measure on the y-axis to represent the monthly patient count effectively.
* Identify trends in patient visit volumes over the course of the year.
* A Line Chart is used with a **Year** and **Month Name** hierarchy.
* The **Count\_of\_Patients** measure tracks how patient visit frequency changes over time.
* The hierarchy enables analysis at both high-level (yearly) and detailed (monthly) granularities.
* This visualization highlights potential seasonal trends or fluctuations in patient visits.

### **Insight :**In 2019, there were a total of 4,338 patient visits, which rose to 4,878 in 2020—a **7.5% increase.** This growth highlights a potential shift in healthcare demand, possibly influenced by factors such as heightened health awareness, the introduction of new medical services, or external events like the pandemic.

* **Result:**The analysis of patient visit trends over the years reveals a noticeable upward trend in 2020 compared to 2019.



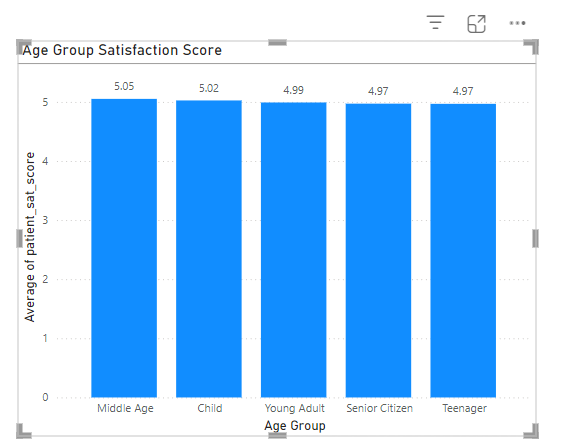
**Visualizations:**



**Q4. Which age groups report the highest and lowest satisfaction scores?**

**Ans.**

* **Satisfaction Range:** Scores across age groups range between 4.97 and 5.05**.**
* **Highest Satisfaction:** The Middle age group reported the highest score of **5.05.**
* **Lowest Satisfaction:** The teenager and Senior Citizen age group had the lowest score of **4.97.**
* This indicates potential areas for improvement in services for **senior and teenager patients.**

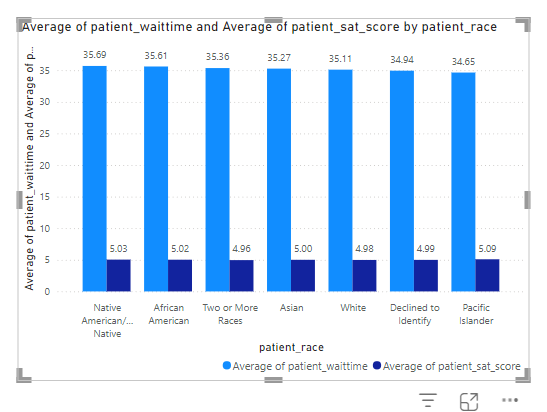


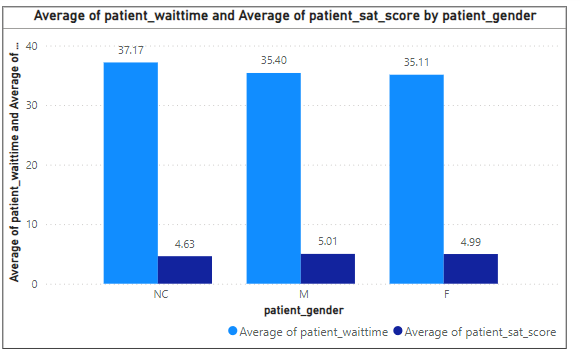
**Visualization:**

**Q5. Say someone outside of the hospital claims that there is racial or gender-based discrimination in the hospital, how will you identify whether the claim was right or not?**

**Ans.**

* Use a **Clustered Column Chart** , the X-axis contains **race** and **gender** in 2 different visuals.
* Plot the **average waiting time** and **average satisfaction score** on the Y-axis for each group.
* This chart enables a comparison of how various **race** and **gender** groups experience their visits, focusing on waiting time and satisfaction scores.
* A **Clustered Column Chart** is used to examine potential racial or gender-based disparities in the hospital.
* **Insights:** By comparing metrics such as waiting times and satisfaction scores, we can assess whether any demographic group faces significantly longer waits or lower satisfaction, which may suggest possible discriminatory practices.
* The analysis of average waiting times and patient satisfaction scores reveals no indication of racial or gender discrimination in the hospital.  
  The data demonstrates consistent and fair treatment across all race and gender groups, with no significant differences in waiting times or satisfaction levels.
* **Result :** The analysis reveals that waiting times and satisfaction scores are similar across race and gender groups, suggesting that the hospital does **not exhibit any discriminatory practices.**

**Visualizations:**



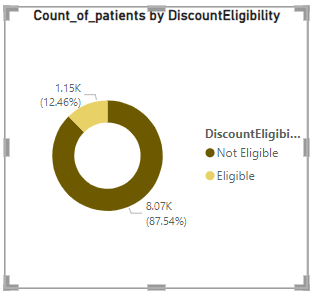
**Q6. The hospital management intends to offer discounts to patients. How should these offers/discounts be assigned to patients, on what basis, and why?**

**Ans.**

* To identify patients eligible for discounts, I created a DAX formula that checks two conditions:
* A **Total Bill** of 10,000 or more.
* A **Satisfaction Score** of 6 or higher.
* Based on these criteria, I developed a **DiscountEligibility** column that categorizes patients as either "Eligible" or "Not Eligible." I then visualized the outcomes in a donut chart, displaying the count of eligible versus non-eligible patients.
* The hospital management intended to offer discounts to high-spending and satisfied patients. I used a DAX formula to categorize patients based on their total bill and satisfaction score, creating a field to reflect whether a patient qualifies for the discount.
* **The formula used is:**

DiscountEligibility = IF( SUM(Doctor\_Patient\_Data[Total Bill]) >= 10000 && 'Hospital ER'[patient\_sat\_score] >= 6, "Eligible", "Not Eligible")

* **Insight:** Out of the total 9,216 patients, **12.46% were eligible for the discount, and 87.54% were not.**
* **Result:** The donut chart shows that a small proportion (12.46%) of patients are eligible for the discount, while the majority (87.54%) do not meet the eligibility criteria.

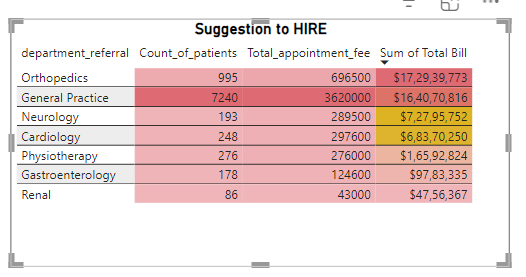
**Visualization:**

**Q7.The hospital has a budget to hire 2-3 new doctors. They have asked for your suggestions on which departments they should hire.**

**Ans.**

* I evaluated the hospital's departments using key metrics such as **Count of Patients**, **Sum of Appointment Fees**, and **Sum of Total Bills** to recommend which departments should consider hiring 2-3 additional doctors. This analysis was presented through a table visualization.
* Departments with higher patient volumes, appointment fees, and total bills were given priority.
* **Orthopedics** had the highest patient count (**995**), appointment fees (**696,500**), and total bill (**172,939,773**).
* **General Practice** had a large patient count (**7,240**) and generated significant revenue.
* **Neurology** had fewer patients (**193**) but substantial appointment fees and total bills.
* **Insight: Orthopedics** and **General Practice** have the highest patient load and revenue, making them top choices for new hires.
* **Neurology,** despite lower patient count, still shows significant demand.
* **Result:** I suggest hiring new doctors for **Orthopedics, General Practice, and Neurology** based on their high patient volume and revenue generation.

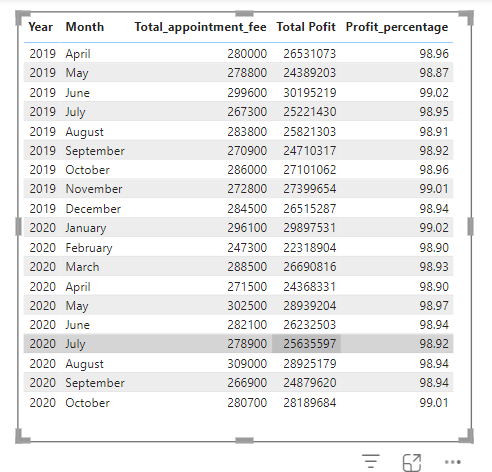
**Visualization:**



**Q8. Is the hospital profitable? How will you determine the profitability?**

**Ans.**

* To assess the hospital's profitability, I created a visual table with **Year** and **Month** as columns.
* I calculated key metrics such as **total revenue**, **total appointment fees**, **total profit**, and **profit percentage** using DAX measures.
* These metrics provided insights into whether the hospital is operating profitably.
* **Total Appointment Fee**: The sum of appointment fees for each patient.
* **Total Revenue**: The total of the bill amounts for all patients.
* **Total Profit**: Calculated by subtracting the total appointment fee from the total revenue.
* **Profit Percentage**: Determined by dividing the total profit by total revenue and multiplying by 100 to express it as a percentage.
* The **Profit Percentage** remained consistently high, around **98.96%** and **98.87%**, indicating a strong profit margin.
* **Insight:** The profit margin remains consistently high, well above 90%, which indicates the hospital is highly profitable.
* The hospital generates substantial revenue compared to its costs, with only a small portion of the revenue going toward appointment fees**.**
* **Result:** Based on the profit percentage consistently being above 98%, it is clear that the hospital is highly profitable.

**Visualization:**

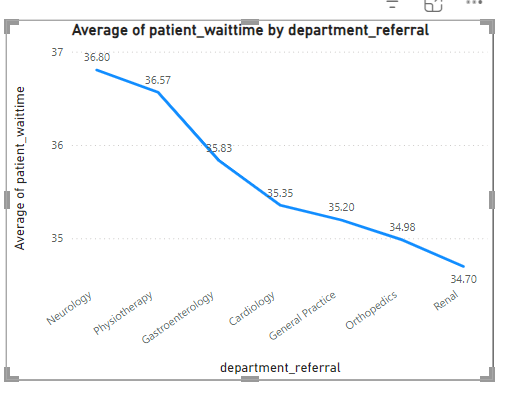
**Q9. Any Department for which the waiting time is oddly large?**

**Ans.**

* To identify departments with unusually high waiting times, I created a line chart with **Department** on the x-axis and **Average Waiting Time** on the y-axis. This provided a clear comparison of waiting times across all departments.
* The line chart revealed variations in waiting times between departments. Upon analysis, I noted that the **Neurology** department had the highest average waiting time of **36.80 minutes**, significantly higher than other departments.

### **Insight:**The **Neurology** department's elevated waiting time may suggest potential issues like inadequate staffing, high patient demand, or scheduling inefficiencies.

### **Result:**The **Neurology** department was found to have the highest average waiting time of **36.80 minutes**, notably exceeding the waiting times of other departments.



**Visualization:**

**Q10. Come up with strategies to provide discounts to the patients.**

**Ans.**

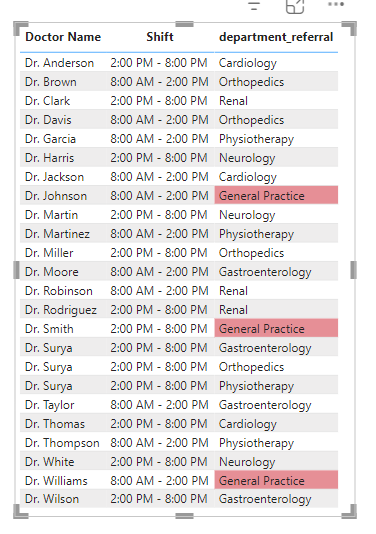
* Offer discounts to patients whose total bill exceeds ₹10,000 to incentivize higher spenders.
* Provide discounts to patients with a satisfaction score of 6 or above to foster loyalty and encourage positive feedback.
* Introduce graduated discount tiers based on the total bill, such as 5% for bills between ₹10,000 and ₹20,000, and 10% for amounts over ₹20,000.
* Award compensatory discounts to patients facing exceptionally long waiting times to enhance their overall experience.
* Implement seasonal or promotional discounts to attract more patients during periods of lower demand.

**Q11. Say you need to align the doctors of the “General Practice” department to work in one of the two shifts, how will you identify what will these two shifts' timings be, and how will you divide the doctors in these two shifts? And also will this 2 shift policy be helpful for the hospital?**

**Ans.**

* A table visualization was created displaying Doctor Name and Department. DAX was used to assign doctors to two shifts: "**8:00 AM - 2:00 PM**" and "**2:00 PM - 8:00 PM**". Conditional formatting was applied to highlight the General Practice department.
* To allocate doctors to shifts, the following **DAX formula** was implemented:

Shift = IF(MOD(RANKX(ALL('Doctor\_Patient\_data'), 'Doctor\_Patient\_Data'[Doctor Name], , ASC), 2) = 0,"8:00 AM - 2:00 PM","2:00 PM - 8:00 PM")

* This formula ranks doctors by name and alternates the shift assignment using the MOD function.
* Dr. Johnson and Dr. Williams were assigned to the morning shift (8:00 AM - 2:00 PM) due to higher patient traffic, while Dr. Smith was assigned to the evening shift (2:00 PM - 8:00 PM) to handle lower patient volume.
* **Result:** The two-shift policy allows for even distribution of workload, ensures better patient care, and reduces wait times, particularly in the General Practice department.

**Visualization:**

**Q12.What do you understand by PowerBI gateway? What are its use cases?**

**Ans:**

### **Power BI Gateway:**

A Power BI Gateway is a crucial tool that facilitates secure data transfer between on-premises data sources and the Power BI cloud service. It enables the seamless refresh and access of locally stored data within reports and dashboards, ensuring up-to-date information is available for analysis.

### **Use Cases:**

1. **Data Refresh:**The gateway allows for the automatic refresh of on-premises data, ensuring that reports and dashboards are always up to date with the latest information.
2. **Secure Data Transfer:**It ensures secure, encrypted transfer of data between on-premises sources (such as SQL Server, Oracle, or Excel) and the Power BI cloud service, safeguarding sensitive information.
3. **Access to On-Premises Data:**The gateway allows users to bring in data from on-premises sources into Power BI, making it possible to analyze local data alongside cloud data.
4. **Hybrid Data Integration:**It enables organizations to combine on-premises and cloud data sources into unified reports, allowing for a holistic view of all relevant information.
5. **DirectQuery:**Power BI Gateway supports DirectQuery, enabling real-time queries of live data sources, providing instant access to the most current data without the need for data import.
6. **Scheduled Data Sync:**Beyond real-time data access, the gateway can be configured for scheduled data synchronization, allowing for controlled and predictable data refreshes at specified intervals.
7. **Support for Large Scale Organizations:**In large enterprises, the Power BI Gateway ensures that multiple users across different departments can access consistent and secure data, promoting collaboration and reducing data silos.

### **Types of Gateways:**

* **Personal Gateway**:  
  Ideal for individual users, this gateway supports refreshing personal data and is typically used for smaller-scale projects.
* **Enterprise Gateway**:  
  Designed for organizations, the enterprise gateway supports multiple users and a wide range of data sources, making it suitable for large-scale deployments and cross-department collaboration.

In summary, the Power BI Gateway is an essential component for businesses that rely on a mix of on-premises and cloud data, enabling secure, real-time, and consistent reporting across the organization.

**Q13 .How would you approach this problem, if the objective and subjective questions weren't given?**

**Ans:**

* **Import Data into Power BI**:  
  Begin the analysis by importing the dataset into Power BI for further processing.
* **Data Transformation in Power Query**:  
  Use **Power Query** to clean, filter, merge, and reshape the data, ensuring it is structured properly for analysis.
* **Data Cleaning**:  
  Address any inconsistencies, handle missing values, correct data types, and standardize the dataset format to ensure its quality and consistency.
* **Identify Key Performance Indicators (KPIs)**:  
  Determine the most relevant KPIs that align with the objectives of the analysis, which will guide the insights and decision-making process.
* **KPI Analysis**:  
  Dive deeper into each KPI by exploring trends, relationships, and patterns in the data, aiming to derive actionable insights.
* **Create Visual Reports and Dashboards**:  
  Develop comprehensive reports and dashboards using various visualizations such as **bar charts**, **line graphs**, and **tables** to present the findings clearly.
* **Add Interactivity**:  
  Include **interactive slicers** and filters in the reports, allowing users to explore the data dynamically across different dimensions.
* **Review and Finalize the Report**:  
  Conduct a thorough review of the final report to ensure accuracy, clarity, and relevance. Provide actionable recommendations based on the insights gathered throughout the analysis.

### **Strategy and Recommendations:**

1. **Maximizing Revenue Generation**:  
   To enhance revenue, I would request additional data such as the **hourly rates** for different hospital room categories, **machine operation costs** in each room, **medicine costs**, and **doctor consultation fees**. This would provide a more detailed understanding of the cost structure and help in optimizing pricing strategies.
2. **Insights for Hiring New Employees**:  
   In addition to focusing on doctors, I would also consider the **nursing staff** and **support staff** in the analysis. A holistic approach to staffing across all roles would ensure the hospital operates efficiently and can meet patient demands effectively.
3. **Workload Distribution and Staffing Decisions**:  
   By gaining a broader understanding of each department's performance and resource requirements, we can better manage workload distribution and make informed decisions on **hiring** or **firing** employees, ensuring departments are appropriately staffed according to the hospital’s needs.
4. **Patient Discounts Strategy**:  
   My strategy for patient discounts would remain focused on the **age group** that generates the most revenue. Offering targeted discounts can incentivize repeat business and enhance patient loyalty.
5. **Long-Term Benefits of Age-Specific Discounts**:  
   In addition to boosting **immediate cash flow**, focusing on a specific age group could lead to the establishment of a **specialty department** tailored to their needs. This would not only increase the hospital's appeal to that demographic but also attract more high-revenue patients, contributing to sustainable growth.

**Q14. Can you analyze and write the type of relationship between the doctor id and department, is it one-to-one?**

**Ans:**

* Each **doctor\_id** is linked to a specific department.
* A **department** can have multiple **doctors** assigned to it.
* From the **doctors' perspective**, it is a **one-to-many relationship** because multiple doctors can belong to the same department.
* From the **department's perspective**, it is a **many-to-one relationship** because each department can have several doctors, but each doctor is assigned to only one department.
* The relationship between the “Doctor Patient Data” table and “Hospital ER” table is a one-to-one relationship.
* The primary key for “Hospital ER” is *patient\_id* and foreign key for “Doctor Patient Data” table is *patient\_id*.
* For each *patient\_id* in the “Hospital ER” table there is only one entry in the “Doctor Patient Data” table.
* There is only one relationship established between these tables and it is active by default.
* The cross filter between these tables is bi-directional.

