**Objective Questions**

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

**Answer:** To prepare the hospital dataset for accurate analysis in Power BI, it's important to clean the data first. Here's how this was done:

* **Identifying Missing or Inconsistent Values:**

I used the Power Query Editor in Power BI to examine the dataset for any missing (null) values or irregular data. During this process, I noticed that the patient\_sat\_score column contained several blank entries.

* **Handling Missing Data:**

To prevent errors during analysis caused by missing data, I chose to replace the blank values. Specifically, I filled the empty entries in the patient\_sat\_score column with the average of the available scores. This approach helps preserve the overall accuracy of the dataset while minimizing potential bias.

* **DAX Formula:**

patient\_updated\_sat\_score=If(ISBLANK('Hospital ER'[patient\_sat\_score]), AVERAGE('Hospital ER'[patient\_sat\_score]), 'Hospital ER' [patient\_sat\_score])

This formula checks if the value is blank. If it is, it replaces it with the average of the column. If not, it keeps the original score.

* **Result of the Cleaning Process:**

After applying these steps, the data became complete and consistent. This makes sure that any visualizations or analysis built on this data will be trustworthy and meaningful.

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

**Answer: Average Waiting Time: 35.26 min.**

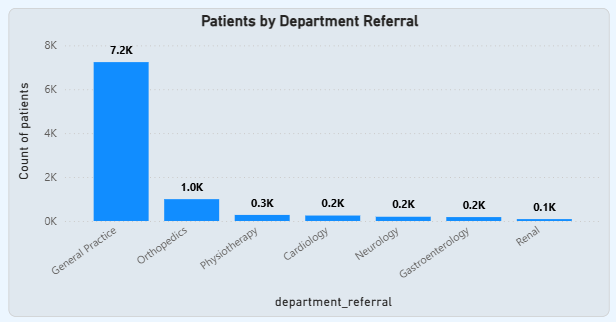
* **DAX Formula:** Average Waiting Time = AVERAGE('Hospital ER'[patient\_waittime])

****

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

**Answer:**

* **DAX Formula:** Count of patients = COUNT('Hospital ER'[patient\_id])

****

**Insight:**

* General Practice receives the vast majority of patient referrals (7.2K).
* Orthopedics is the distant second most common referral department (1.0K).
* All other specialized departments have significantly fewer referrals (0.1K-0.3K).
* Renal department has the lowest number of referrals at 0.1K.

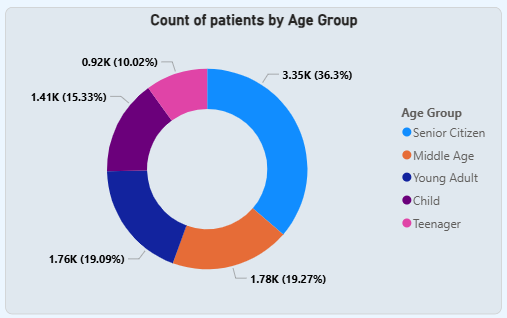
1. **Patient Visits by Age Group: Segregate patient visits according to different age groups to see which demographics utilize healthcare services the most.**

**Answer:**

* Create a new column named “Age Group” to classify patients based on their age.
* **DAX Formula:** Age Group = IF('Hospital ER'[patient\_age] <= 12,  "Child",IF('Hospital ER'[patient\_age] > 12 && 'Hospital ER'[patient\_age] <= 20,  "Teenager",  IF('Hospital ER'[patient\_age] > 20 && 'Hospital ER'[patient\_age] <= 35, "Young Adult", IF( 'Hospital ER'[patient\_age] > 35 && 'Hospital ER'[patient\_age] <= 50, "Middle Age",  "Senior Citizen" ))))

The age groups are defined as follows:

* **Child: Age ≤ 12**
* **Teenager: 13 ≤ Age ≤ 20**
* **Young Adult: 21 ≤ Age ≤ 35**
* **Middle Age: 36 ≤ Age ≤ 50**
* **Senior Citizen: Age > 50**

****

**Insight:**

* The largest age group among patients is "Senior Citizen," accounting for over a third (36.3%) of all patients.
* "Middle Age" and "Young Adult" groups represent a significant portion of patients, with each making up over 19% of the total.
* "Teenagers" and "Children" constitute the smallest patient groups, collectively making up approximately 25% of the total patient count.
* There are considerably more senior citizen patients (3.35K) than patients in any other age group, highlighting a potential focus area for healthcare services.

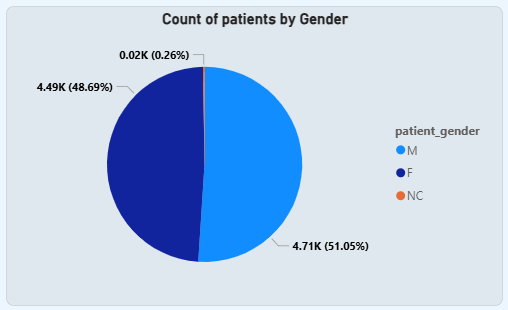
1. **Were there any Null values in the data? What would be the best way to handle these Null values and which approach have you opted for?**

**Answer:**

* **Presence of Null Values:** Yes, the dataset contains null (missing) values.
* **Identifying Missing Data:** Used Power Query Editor to inspect the dataset for any null or inconsistent entries.
* **Handling Missing Data:** Replaced null values—such as those in the patient\_sat\_score column—with the column's average to preserve data integrity and completeness.
* **DAX Formula:** patient\_updated\_sat\_score = If(ISBLANK('Hospital ER'[patient\_sat\_score]),AVERAGE('Hospital ER'[patient\_sat\_score]),'Hospital ER'[patient\_sat\_score])

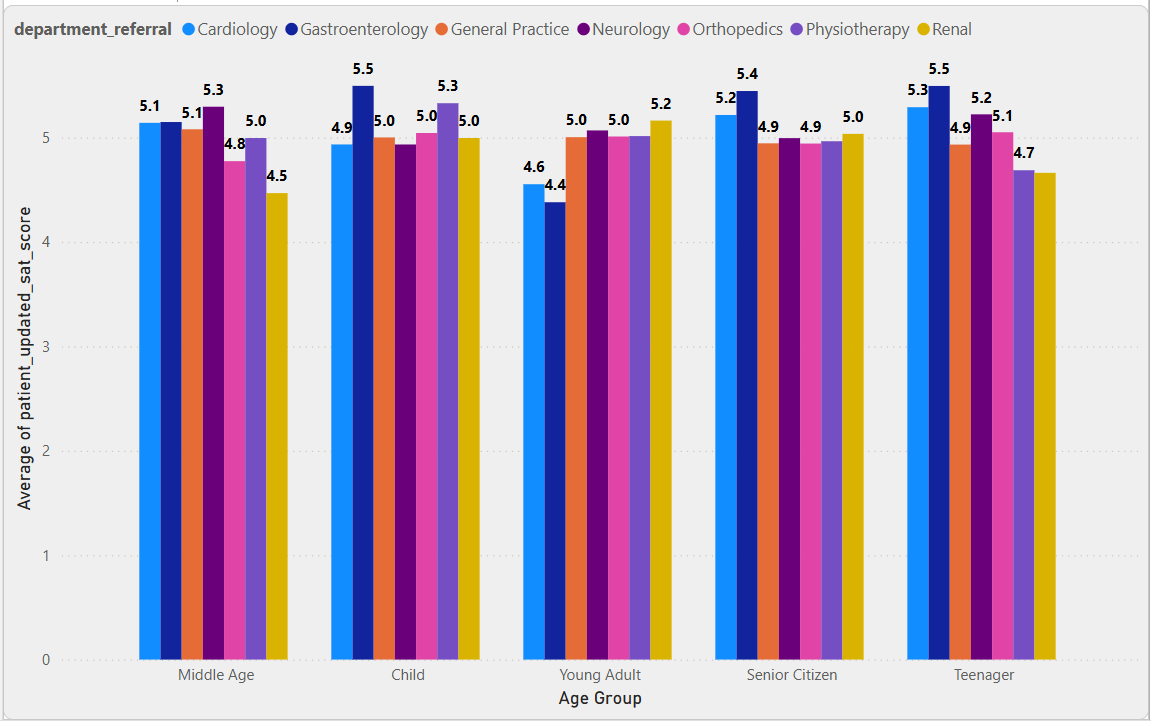
1. **Is there any relation between the number of visits and the Gender of the patients?**

**Answer:** It shows that visits are almost equally shared between male and female patients, with males making up a slightly larger share. The "NC" group is negligible, representing just 0.26% of the total.



1. **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.**

**Answer:**

****

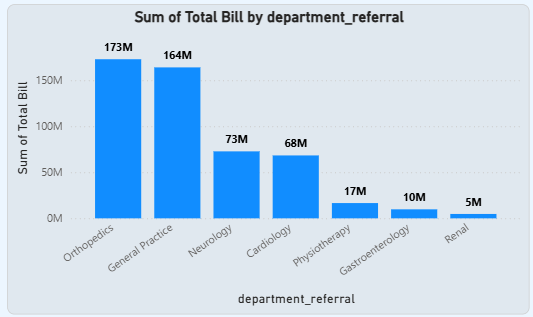
**Insights**

* **Middle-aged** patients tend to have scores falling between4.93 and 5.11.
* **Children** report the highest satisfaction, particularly in **Neurology and Gastroenterology**, both scoring 5.5.
* **Senior Citizens also** reflect strong satisfaction, with Gastroenterology peaking at 5.5 for this group too.
* The **Renal department** shows the lowest ratings across all age brackets, with the lowest at 4.5 for Middle-aged patients and 4.7 for Teenagers.
* **Teenagers generally report lower satisfaction**, especially in Renal and Orthopedics, both at 4.7.
* **Cardiology and Gastroenterology** consistently achieve solid satisfaction results, mostly ranging from 5.1 to 5.5 across age groups.
* Scores for the **Middle-aged segment** are relatively stable, with Neurology at the top (5.3) and Renal at the bottom (4.5).
* **Young Adults** show moderate and less varied scores, staying within 4.6 to 5.2.

**Conclusion:** Children and senior citizens have the highest patient satisfaction, especially in Neurology and Gastroenterology. The Renal department shows the lowest satisfaction across all age groups, with teenagers generally less satisfied overall. Cardiology and Gastroenterology maintain consistently high satisfaction, indicating strong patient trust.

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

**Answer:**

****

**Insights:**

* **Orthopedics** has the highest total billing at **173M,** closely followed by **General Practice** at **164M.**
* **Neurology** and **Cardiology** show moderate billing amounts of **73M** and **68M**, respectively.
* **Physiotherapy**, **Gastroenterology**, and **Renal** have significantly lower total bills, with Renal at only **5M**.
* There is a clear gap between high-billing departments (Orthopedics, General Practice) and low-billing ones (Renal, Gastroenterology).

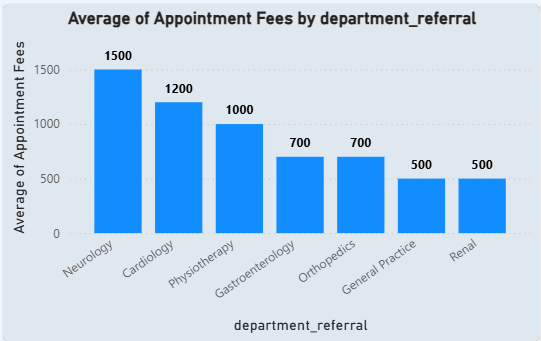
**Conclusion:** Billing is heavily concentrated in Orthopedics and General Practice, indicating higher patient volume or cost per visit, while departments like Renal and Gastroenterology contribute minimally to total revenue.

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

**Answer:**

****

* **DAX Formula:** Highest Appointment Fees by Department = CALCULATE(MAX('Doctor Patients'[Appointment Fees]), ALLEXCEPT('Doctor Patients','Doctor Patients'[department\_referral]))
* **Explanation:** It calculates the highest appointment fee within each department, ignoring any other filters except the department itself. In simple terms: For each department, it finds the maximum appointment fee, no matter what other filters are applied to the data. Below graph shows without DAX formula

****

1. **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.**

**Answer:**

* **DAX Formula:**

1. Month. = FORMAT('Hospital ER'[Date1].[Date], "MMMM")
2. Previous\_Month\_Visits. =

VAR LatestDate = MAX('Hospital ER'[Date1])

VAR PreviousMonthDate = EOMONTH(LatestDate, -1)

RETURN

CALCULATE(

    COUNT('Hospital ER'[patient\_id]),

    FILTER(

        ALL('Hospital ER'),

        YEAR('Hospital ER'[Date1]) = YEAR(PreviousMonthDate) &&

        MONTH('Hospital ER'[Date1]) = MONTH(PreviousMonthDate)

    )

)

1. Visits\_Flag =

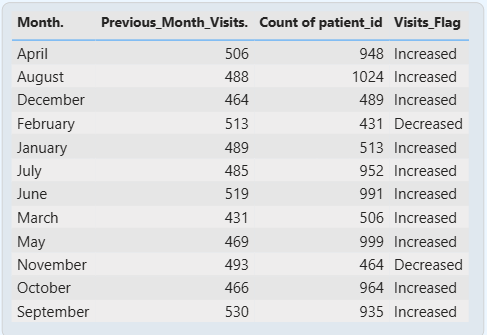
IF(

    [Count of patients] > [Previous\_Month\_Visits.],

    "Increased",

    "Decreased"

)

****

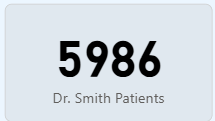
**Insights**

* Pronounced Seasonal Trend: Hospital visits demonstrate clear seasonality, with a decline in May, a strong summer/early autumn surge (June-October), and a year-end decrease.
* Peak Demand Period: The highest visit volumes consistently occur in September and October.
* Dynamic Monthly Shifts: Visit counts fluctuate significantly month-over-month, marked by both notable increases and decreases.
* Mid-Year and Year-End Declines: Both mid-spring (April-May) and year-end (November-December) periods show reduced patient traffic.

**Conclusion:** The hospital experiences predictable seasonal variations in patient visits. Recognizing these patterns is vital for optimizing resource management, staffing, and operational efficiency to effectively handle fluctuating demand throughout the year.

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

**Answer:**

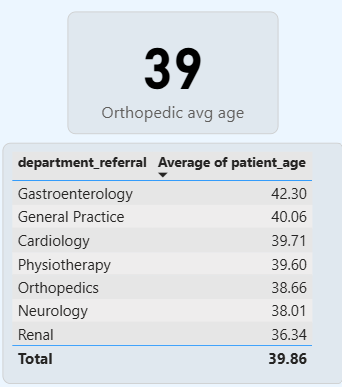
****

* **DAX Formula:** Dr. Smith Patients = CALCULATE(COUNTA('Hospital ER'[patient\_id]),'Doctor Patients'[Doctor Name]="Dr. Smith")
* **Explanation:** It counts the total number of patients treated by Dr. Smith. In simple terms it counts how many patient IDs exist in the Hospital ER table, but only for rows where the doctor’s name is Dr. Smith.

1. **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?**

**Answer:**

* **DAX Formula:** Orthopedic avg age = CALCULATE(AVERAGE('Hospital ER'[patient\_age]),'Hospital ER'[department\_referral]="Orthopedics")
* **Explanation:** It calculates the average age of patients who visited the Orthopedics department. In simple terms it takes the average of all patient ages in the Hospital ER table, but only for those whose department is Orthopedics.

****

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

**Answer:** There were indeed data formatting problems in the *Hospital Info* table, specifically in the **patient satisfaction scores** and **date** columns.The **date** field was stored as text, so I reformatted it properly into a date type to enable correct time-based analysis.

**Addressing blanks in patient satisfaction scores:**

* The **patient satisfaction scores** column contained empty spaces that could affect calculations.
* Originally stored as text, I first converted this column to a numeric type.
* Then, I replaced any blanks (nulls) with the average score for that column to maintain data quality and prevent calculation errors.

These steps ensured the dataset was clean, consistent, and ready for accurate reporting and meaningful insights.

1. **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?**

**Answer:**

When we add a column in Power Query, the M language code you see in the formula bar looks like this:

#"Inserted Merged Column" = Table.AddColumn(#"Changed Type", "Full\_name", each Text.Combine({[patient\_first\_inital], ". ", [patient\_last\_name]}), type text)

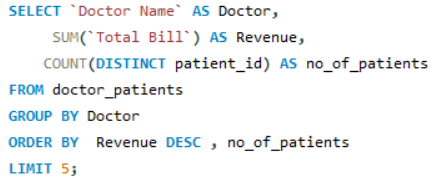
This line adds a new column named **"Full\_name"** by combining two existing columns with custom logic.

**About M Query:**

* **M** is the language used in Power Query to shape, clean, and transform data.
* It’s case-sensitive, functional, and step-based each step depends on the previous one.
* It’s powerful for building repeatable, automated data transformations without manual editing.

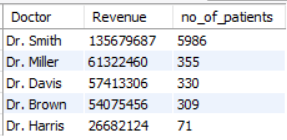
1. **Identify the top 5 doctors who generated the most revenue but had the fewest patients. (SQL)**

**Answer:**

****

**Explanation:**

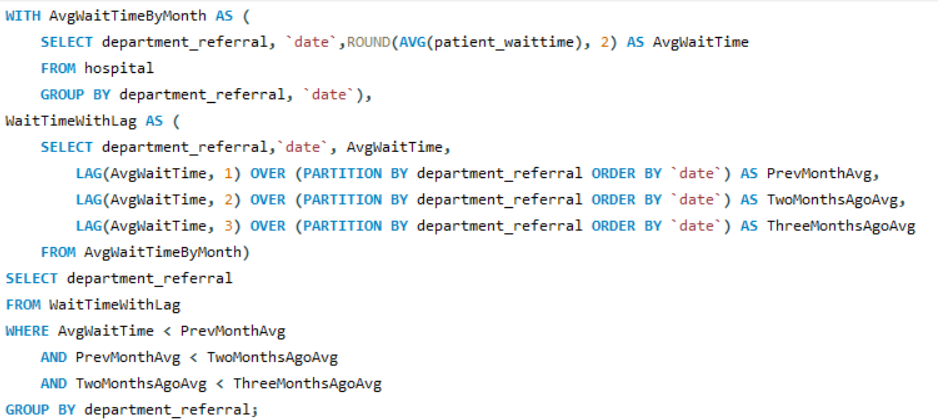
* + Selects: Doctor's name, total revenue (SUM(Total Bill)), and count of unique patients (COUNT(DISTINCT patient\_id)).
  + From: The doctor\_patients table.
  + Groups By: Each Doctor Name to aggregate data per doctor.
  + Orders By: Revenue in descending order, then by no\_of\_patients (ascending by default).
  + Limits: The output to the top 5 results.

****

* + **Insight:** Dr. Smith significantly outperforms other doctors in both revenue generated and the number of unique patients served. The remaining top doctors show a considerable drop-off in both metrics compared to Dr. Smith.

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

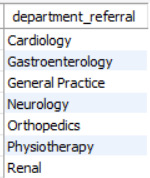
**Answer:**

****

* + AvgWaitTimeByMonth CTE: This Common Table Expression (CTE) calculates the average patient wait time, rounded to two decimal places, for each department\_referral and specific date (presumably representing a month).
  + WaitTimeWithLag CTE: This second CTE builds upon the first. For each department and month, it retrieves the AvgWaitTime and uses the LAG window function to get the average wait times from the previous month (PrevMonthAvg), two months ago (TwoMonthsAgoAvg), and three months ago (ThreeMonthsAgoAvg). This is partitioned by department\_referral and ordered by date to ensure correct month-over-month comparison within each department.
  + Final SELECT: The outer query then selects department\_referral from the WaitTimeWithLag CTE.

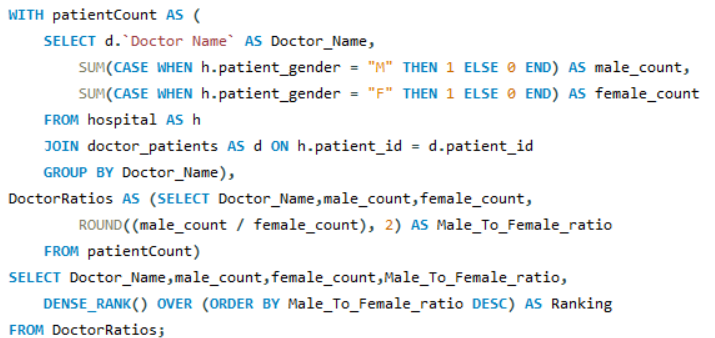
WHERE Clause: It filters these results to include only those departments where:

* + Current AvgWaitTime is less than PrevMonthAvg (current month is better than last).
  + PrevMonthAvg is less than TwoMonthsAgoAvg (last month was better than two months ago).
  + TwoMonthsAgoAvg is less than ThreeMonthsAgoAvg (two months ago was better than three months ago).
  + GROUP BY: Finally, it groups the results by department\_referral to show unique departments meeting these criteria.

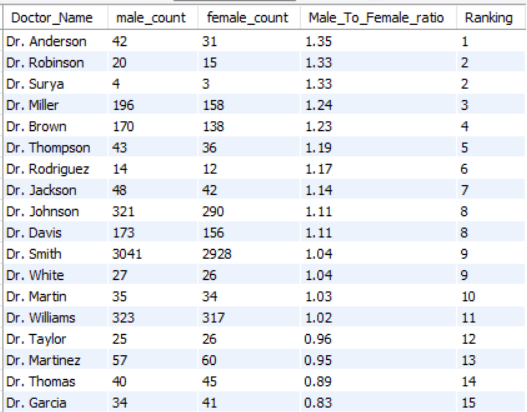
****

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

**Answer:**

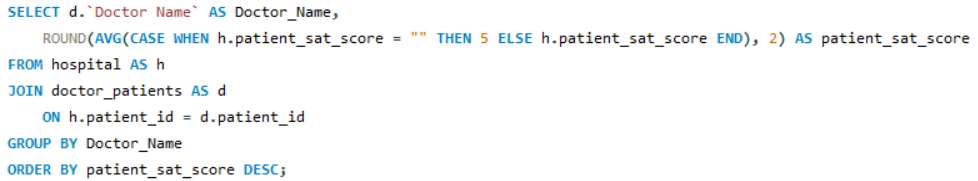
****

* + patientCount CTE: Calculates the total male and female patients per doctor by joining hospital and doctor\_patients tables.
  + DoctorRatios CTE: Computes the Male\_To\_Female\_ratio for each doctor from the patientCount data.
  + Final Select: Ranks doctors using DENSE\_RANK() based on their Male\_To\_Female\_ratio in descending order, showing which doctors have a higher proportion of male patients.

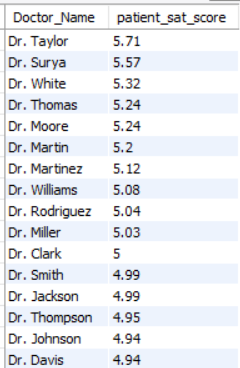
****

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

**Answer:**

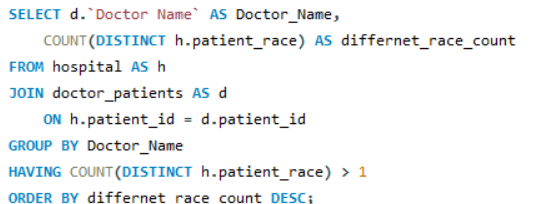
****

This SQL query calculates each doctor's average patient satisfaction score. It handles missing scores by treating empty values as 5. The query joins patient satisfaction data from the hospital table with doctor information from doctor\_patients, groups the results by doctor, and then orders them by the calculated average satisfaction score in descending order (highest satisfaction first).

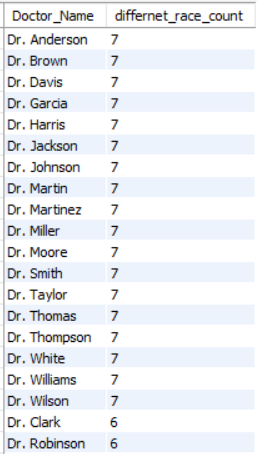
****

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

**Answer:**

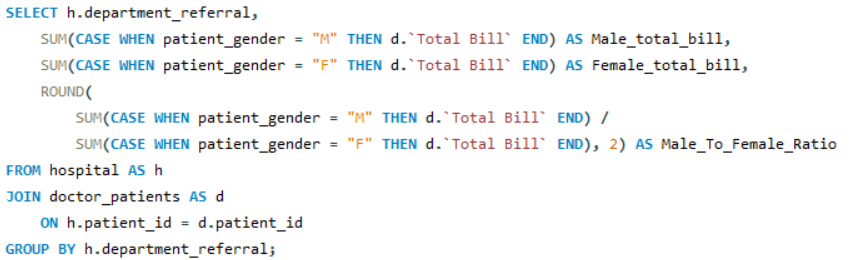
****

SQL query aims to find doctors who treat a diverse patient population. It joins patient data (including race) with doctor assignments. For each doctor, it counts the number of *distinct* patient races encountered. It then filters to include only doctors who have treated patients from more than one race. Finally, the doctors are ranked by the number of distinct races, showing the most diverse practitioners first.

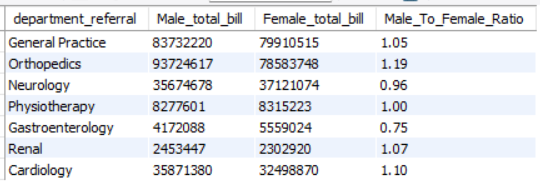
****

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

**Answer:**

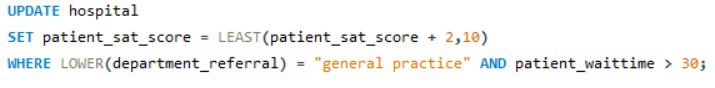
****

* + This SQL query calculates the total bill amount generated from male and female patients, and the male-to-female bill ratio, for each hospital department.
  + Data Source & Join: It retrieves data from the hospital table (aliased as h), which likely contains patient gender and department information, and joins it with the doctor\_patients table (aliased as d) on patient\_id to access Total Bill information.
  + Male/Female Bill Calculation: It uses SUM(CASE WHEN patient\_gender = "M" THEN d.'Total Bill' END) to calculate Male\_total\_bill and similarly for Female\_total\_bill. CASE statements sum the Total Bill only if the patient matches the specified gender.
  + Ratio Calculation: It then calculates Male\_To\_Female\_Ratio by dividing Male\_total\_bill by Female\_total\_bill and rounding the result to two decimal places.
  + Grouping: The results are GROUP BY h.department\_referral to ensure these calculations (sums and ratio) are performed for each distinct department.

****

1. **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)**

**Answer:**

****

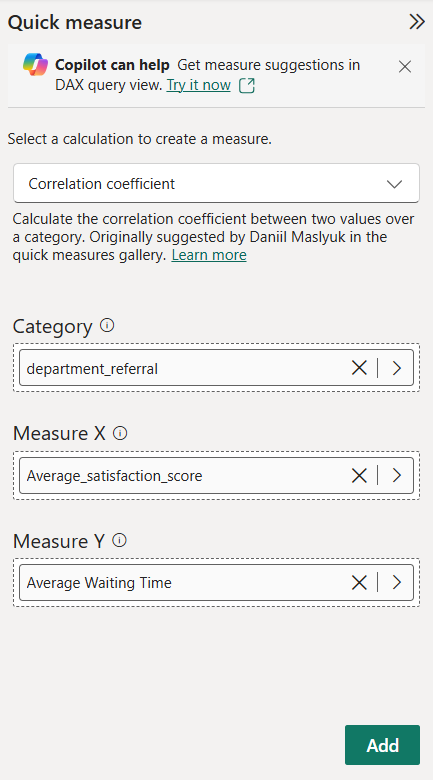
Query UPDATE statement increases patient satisfaction scores by 2, up to a maximum of 10, for patients in the "General Practice" department who had a wait time exceeding 30 minutes. This aims to simulate a post-visit intervention to improve perceived satisfaction for specific patients.

**Subjective Questions**

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

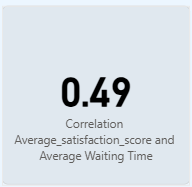
**Answer:**

* In this analysis, I leveraged Quick Measure, a Power BI capability that allows you to build standard calculations (such as correlations and averages) without manually writing DAX code.
* I applied the Correlation Coefficient option to check how patient wait time (Y-axis) relates to the satisfaction score (X-axis), categorized by department referral.
* Power BI used a statistical method to compute the coefficient, which came out to 0.49 — indicating a moderately positive link between wait time and satisfaction.
* This feature simplifies advanced calculations, reduces manual effort, and enables quick insights into data relationships even for users who aren’t familiar with DAX.

****

**Understanding the Correlation Coefficient:**

* A correlation coefficient can vary between -1 and +1:
* +1 means there’s a strong positive link — when one variable goes up, so does the other.
* 0 means there’s no link — the variables don’t affect each other.
* -1 means there’s a strong negative link — when one variable increases, the other goes down.
* A result of 0.49 suggests there’s a moderate positive connection between the variables.

****

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

**Answer:**

* **DAX Formula:**

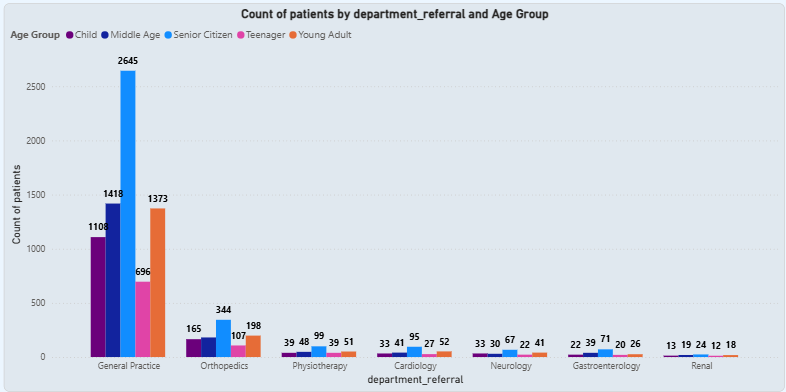
Age group-Gender/Race = {

    ("Age Group", NAMEOF('Hospital ER'[Age Group]), 0),

    ("patient\_gender", NAMEOF('Hospital ER'[patient\_gender]), 1),

    ("patient\_race", NAMEOF('Hospital ER'[patient\_race]), 2)

}

****

**Observations and Insights:**

* **General Practice dominates:** All age groups visit General Practice the most, with **Senior Citizens (2,645)** and Young Adults (1,373) being the largest groups.
* **Middle Age and Children**: Also show significant visits to General Practice (**1,418** and **1,103**, respectively).
* **Teenagers**: Have fewer visits overall but still contribute **696** to General Practice.
* **Specialty departments get fewer visits**:
  + **Orthopedics** has noticeable visits among Middle Age (165), Senior Citizens (344), and Children (107).
  + **Physiotherapy** is mainly used by Senior Citizens (99) and Middle Age (48).
  + **Cardiology** has a clear concentration among Senior Citizens (95) and Young Adults (52).
  + **Neurology** shows higher visits for Senior Citizens (67) and Children (33).
  + **Gastroenterology** and **Renal** departments have low counts across all age groups, with Senior Citizens showing the highest in both.

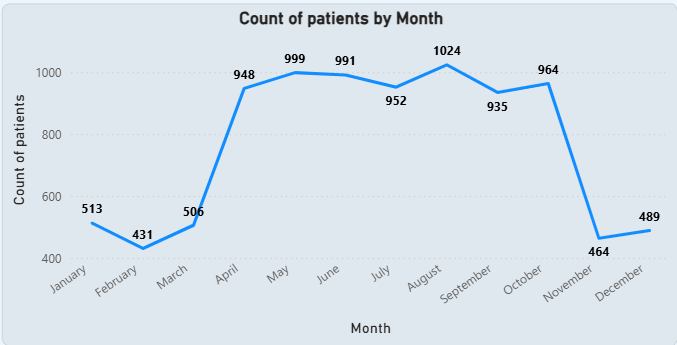
**Conclusion:**

* **Senior Citizens have the highest demand** for multiple departments, especially General Practice and Cardiology-related services.
* **Children rely mostly on General Practice and Neurology**, indicating common pediatric care and neurological concerns.
* **Young Adults and Middle Age groups use General Practice most**, with moderate use of Orthopedics and Physiotherapy.
* **Teenagers have the lowest visit counts overall**, mostly using General Practice.
* **Specialized departments** (like Renal and Gastroenterology) receive minimal visits across all demographics.

**Final Takeaway:** **Patient age groups strongly influence the frequency and type of department visits.** General Practice is the universal entry point, while older patients use more specialty services compared to younger groups.

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

**Answer:**

****

**Insights and Observations:**

* **Low Start to the Year:** Patient counts are relatively low in **January (513)** and reach the lowest point in **February (431)**.
* **Gradual Increase:** There is a steady rise from **March (506)** to **April (948)**, with visits peaking in the summer and early fall months.
* **Peak Period:** The highest patient count is in **August (1,024)**, with consistently high numbers from **April to October** (ranging from ~950 to 1,024).
* **Slight Dip:** A minor dip appears in **July (952)** and **September (935)** compared to surrounding months, but overall volume stays high.
* **End-of-Year Drop:** There’s a significant drop after **October (964)** — visits fall sharply in **November (464)** and stay low in **December (489)**.

**Conclusion:**

* There is **a clear seasonal trend**: patient volume is lowest at the start and end of the year and highest in the **middle months** (spring through early fall).
* The **busy season** runs from **April to October**, suggesting possible seasonal factors, weather, or planned checkups driving higher demand.
* The drop towards **November–December** could be due to holidays, weather conditions, or patients postponing non-urgent visits.

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

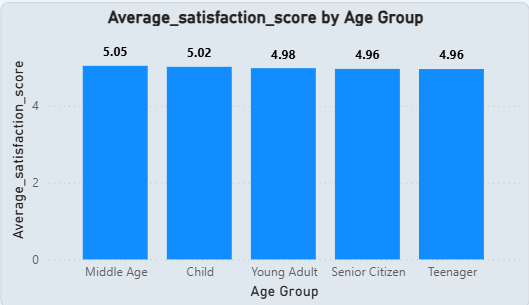
**Answer:**

**Insights and Observations:**

* **Highest Satisfaction:** The **Middle Age group** reports the highest average satisfaction score at **5.05**, slightly ahead of other groups.
* **Close Second:** **Children** also show strong satisfaction, averaging **5.02**, which is just slightly below the Middle Age group.
* **Lowest Scores:** **Senior Citizens and Teenagers** report the lowest average satisfaction, both at **4.96**, indicating slightly less satisfaction compared to other groups.

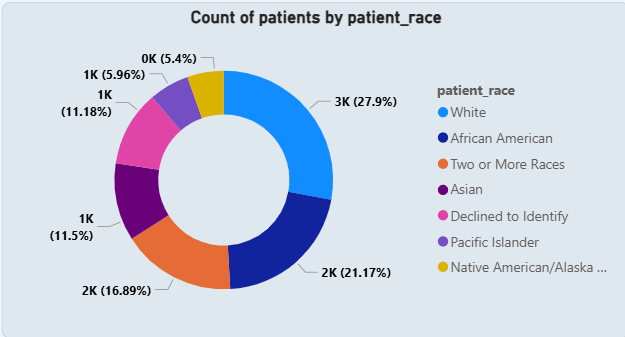
**Conclusion:**

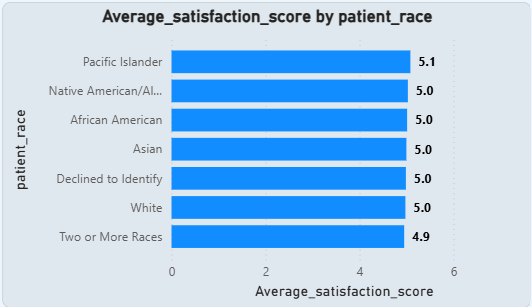
* **Middle Age patients are the most satisfied overall.**
* **Senior Citizens and Teenagers are the least satisfied groups.**
* Overall, **satisfaction scores are quite close across all age groups**, with only a small gap between highest and lowest.

****

1. **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?**

**Answer:**

****

****

**Insights and Observations:**

* **Patient Distribution by Race:**
  + The largest racial group visiting the hospital is **White patients (27.9%)**, followed by **African American (21.17%)**, **Two or More Races (16.89%)**, **Asian (11.5%)**, and others with smaller proportions.
  + There is also a portion of patients who **Declined to Identify** their race (~11%).
* **Satisfaction Scores by Race:**
  + The **average satisfaction scores** are very close for all racial groups:
    - **Pacific Islander patients** have the highest average at **5.1**, but they represent the smallest share.
    - Most other racial groups — **White, African American, Asian, Native American/Alaska Native, Declined to Identify** — all have the same average score of **5.0**.
    - **Patients of Two or More Races** have a slightly lower score (**4.9**), but the difference is minimal.
* **Key Observation:**
  + There is **no significant gap** in satisfaction scores between racial groups.
  + All groups have nearly equal levels of reported satisfaction, indicating that services and experiences appear similar across demographics.

**Conclusion (How to verify claims):**

* **Check Data Trends:** The charts suggest **no evidence of racial bias** based on patient count and satisfaction. Most groups are equally satisfied.
* **Investigate Gender Too:** If gender-based discrimination is a concern, run similar analyses for **gender breakdown** — compare visit counts and satisfaction scores by gender.
* **Combine with Other Metrics:** Look at complaints, grievances, treatment times, and outcomes for different races/genders to see if there are patterns of unequal treatment.
* **Take Action if Needed:** If any group consistently reports lower satisfaction or worse outcomes, it could point to areas needing improvement.

**Final Answer:** **Based on this data, there’s no clear indication of racial discrimination in how patients are treated or how satisfied they are.** However, a complete analysis should **also include gender-specific data** and look beyond just satisfaction scores to confirm fairness in treatment, wait times, and outcomes.

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

**Answer:**

**Insights and Observations:**

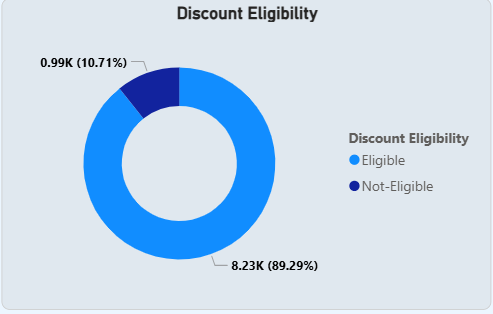
* **Dax Formula:**

Discount Eligibility = if(calculate(SUM('Doctor Patients'[Total Bill]),ALLEXCEPT('Doctor Patients','Doctor Patients'[patient\_id])) > 10000 ,"Eligible","Not-Eligible")

* **Patients whose total bill amount exceeds ₹10,000** are marked **Eligible** for a discount otherwise, they are **Not-Eligible**.
* According to the chart:
  + **89.29% (8.23K patients)** are eligible for discounts.
  + **10.71% (0.99K patients)** are not eligible.
* This shows that **most patients meet the billing threshold**, indicating high-value patients are being rewarded.

**Conclusion**

* **Basis for Discount:** The discount is based on the **total bill amount per patient** — patients spending more than ₹10,000 qualify.
* **Why:**
  + Encourages **loyalty** and **repeat visits** among high-value patients.
  + Rewards patients who generate significant revenue.
  + Makes discounts meaningful without over-extending hospital finances.
* **Recommendation:** This approach is fair and data-driven — **high spenders get benefits**, ensuring discounts target patients who contribute most to revenue.

****

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

**Answer:**

**Insights:**

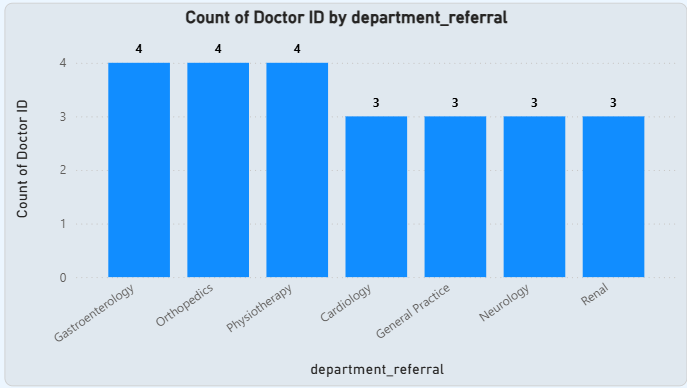
* The *Doctor Count by Department* chart shows Gastroenterology, Orthopedics, and Physiotherapy have 4 doctors each, while General Practice, Cardiology, Neurology, and Renal have only 3 doctors each.
* The *Patients by Department* chart shows General Practice has by far the highest patient count (7,240 patients) but only 3 doctors, indicating an unbalanced workload.
* Orthopedics has the second highest patient volume (995) but already has 4 doctors.
* Physiotherapy, Cardiology, and Neurology see moderate patient numbers (200–300 patients) with similar or fewer doctors.
* Renal and Gastroenterology have the lowest patient numbers and do not urgently require extra staffing.

**Recommendation:**

* Prioritize hiring for *General Practice*, which clearly needs at least 1–2 more doctors to manage its large patient volume.
* Consider adding 1 doctor to *Cardiology* or *Neurology* to help balance workload and maintain quality care.

**Conclusion:**

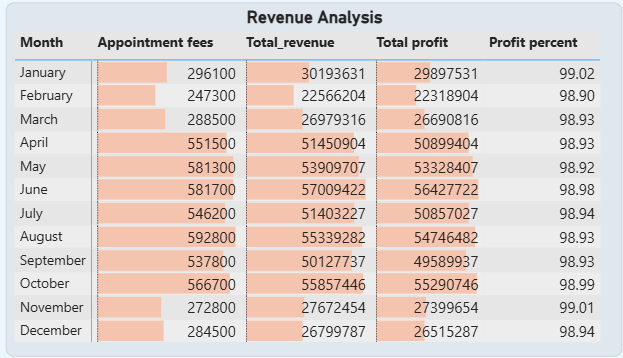
* General Practice should receive the highest priority for new hires.
* Cardiology or Neurology could also benefit from an additional doctor.
* No immediate need to hire more doctors for departments with lower patient volumes and balanced doctor counts.

****

****

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

**Answer:**

****

**How profitability is determined:**

* *Appointment Fees* represent revenue directly from consultations.
* *Total Revenue* includes all billing amounts collected.
* *Total Profit* is the difference between total revenue and appointment costs.
* *Profit Percent* shows profit as a percentage of total revenue.

**Insights:**

* The profit percentage stays extremely high throughout the year, consistently above 98.9%.
* Monthly profits closely follow total revenue trends, showing that costs beyond appointment fees are minimal.
* The steady profit margins indicate that operational costs are well-controlled.

**Conclusion:**

* The hospital is clearly operating with strong profitability.
* The consistently high profit percentage shows that the business model is financially healthy and sustainable.

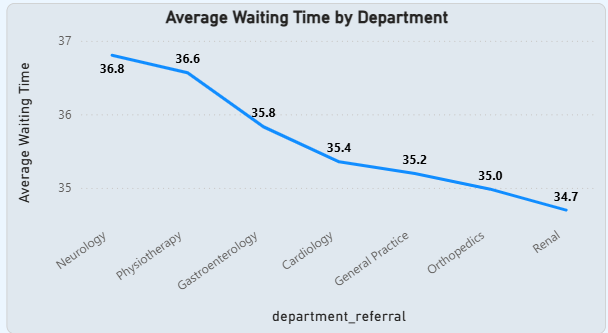
1. **Any Department for which the waiting time is oddly large?**

**Answer:**

**Insights:**

* Neurology has the Longest Wait Time: Neurology consistently shows the highest average waiting time at 36.8, indicating it's the department with the most significant patient wait.
* Significant Gradient in Wait Times: There's a clear decreasing trend in average waiting times across departments, from Neurology down to Renal, suggesting varying levels of demand or efficiency.
* Renal Department is Most Efficient (Wait Time Wise): The Renal department boasts the lowest average waiting time at 34.7, suggesting it is the most efficient or least burdened in terms of patient waiting.
* Specialized Departments Often Have Longer Waits: The top three departments with the longest waiting times (Neurology, Physiotherapy, Gastroenterology) are all specialized fields, potentially indicating higher demand for their specific expertise.

**Conclusion:** Neurology appears to have an unusually long waiting time, indicating a potential bottleneck or high demand in that specific department that requires further investigation.

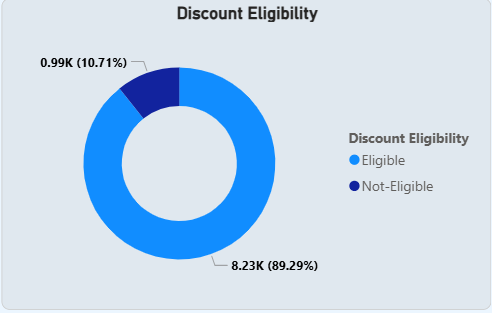
****

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

**Answer:**

* **Dax Formula:**

Discount Eligibility = if(calculate(SUM('Doctor Patients'[Total Bill]),ALLEXCEPT('Doctor Patients','Doctor Patients'[patient\_id])) > 10000 ,"Eligible","Not-Eligible")

****

**Insights**

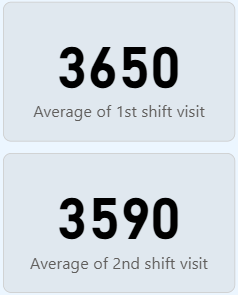
* **Patients whose total bill amount exceeds ₹10,000** are marked **Eligible** for a discount otherwise, they are **Not-Eligible**.
* According to the chart:
  + **89.29% (8.23K patients)** are eligible for discounts.
  + **10.71% (0.99K patients)** are not eligible.
* This shows that **most patients meet the billing threshold**, indicating high-value patients are being rewarded.

**Conclusion**

* **Basis for Discount:** The discount is based on the **total bill amount per patient** — patients spending more than ₹10,000 qualify.
* **Why:**
  + Encourages **loyalty** and **repeat visits** among high-value patients.
  + Rewards patients who generate significant revenue.
  + Makes discounts meaningful without over-extending hospital finances.
* **Recommendation:** This approach is fair and data-driven — **high spenders get benefits**, ensuring discounts target patients who contribute most to revenue.

1. **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?**

**Answer:**

****

* **Dax Formula:**

1. 1st shift visit = CALCULATE([Count of patients],'Hospital ER'[Hours] >=0 && 'Hospital ER'[Hours]<12 && 'Hospital ER'[department\_referral] ="General Practice")
2. 2nd shift visit = CALCULATE([Count of patients],'Hospital ER'[Hours] >=12 && 'Hospital ER'[Hours]<=23 && 'Hospital ER'[department\_referral] ="General Practice")

* 1st shift visit aggregates patients in "General Practice" from midnight to noon (0 <= Hours < 12).
* 2nd shift visit sums patients in the same department from noon to midnight (12 <= Hours <= 23). These formulas help identify patient load distribution across shifts.

**Insights:**

* Higher Demand in 1st Shift: The "Average of 1st shift visit" is 3650, which is notably higher than the "Average of 2nd shift visit" at 3590. This indicates a greater patient load during the first shift.
* Shift Timings: The DAX formulas define the shifts as 1st Shift (00:00-11:59) and 2nd Shift (12:00-23:59), providing clear timing for doctor alignment.
* Uneven Patient Distribution: While both shifts have significant patient visits, the slightly higher average in the first shift suggests an uneven distribution of patient traffic, which should be considered for doctor allocation.

**Final Conclusion:**

Given the higher average patient visits in the 1st shift (3650 vs. 3590), the hospital should align more doctors or allocate more resources to the 1st shift (00:00-11:59) for the General Practice department. A 2-shift policy based on these timings is beneficial as it aligns doctor availability with periods of higher patient demand, thereby optimizing patient care and resource utilization.

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

**Answer:**  
**Power BI Gateway** is a bridge that securely connects on-premises data sources with the Power BI cloud service. It enables users to keep their reports and dashboards up to date by allowing scheduled or live data refreshes from local servers without manually uploading data to the cloud.

**Key Use Cases:**

* **Automated Data Refresh:** Keeps reports current by automatically syncing on-premises data with the Power BI service.
* **Secure Connectivity:** Transfers data through encrypted channels, ensuring that sensitive on-premises information stays protected.
* **Real-Time Access:** Supports DirectQuery, allowing users to run live queries on on-premises data without storing it in the cloud.
* **Hybrid Reporting:** Combines data from on-premises systems (like SQL Server or file shares) with cloud data sources for comprehensive reporting.
* **Personal vs. Enterprise Use:** There are two main types — the **Personal Gateway**, which is suitable for individual users managing their own datasets, and the **Standard/Enterprise Gateway**, which serves multiple users and various data sources within an organization.

**Benefits:**

* Ensures data security and compliance by avoiding unnecessary data movement.
* Reduces costs by eliminating the need to duplicate large data volumes in the cloud.
* Enables up-to-date, reliable dashboards and reports for better decision-making.

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

**Answer:**

* **Clarify the Goal:** Understand the overall purpose of the analysis — what decisions need to be supported and what outcomes are expected.
* **Bring in the Data:** Import all relevant datasets into Power BI to get started with exploration.
* **Prepare the Data:** Use Power Query to transform and shape the data — filter out unnecessary rows, merge related tables, and create the needed structure.
* **Clean and Validate:** Fix any errors, handle missing or duplicate records, correct data types, and ensure consistent formatting.
* **Define Key Metrics:** Identify important KPIs that align with the business context — such as workforce requirements, patient volumes, revenue drivers, or departmental performance.
* **Analyze Broadly:** Explore trends and patterns in these KPIs, even without specific questions, to uncover meaningful insights.
* **Workforce Planning:** Examine not just doctors but also nursing staff, admin, and support teams to get a complete view of staffing needs.
* **Patient Strategies:** Look at revenue contribution by age group or segment to design targeted discount strategies that benefit both patients and the organization.
* **Connect Insights to Action:** Consider how insights could lead to broader improvements, like developing specialized services for high-revenue segments.
* **Visualize Findings:** Build clear dashboards and reports with visuals like charts and tables to present the results.
* **Add Interactivity:** Include slicers, filters, and drill-downs so stakeholders can explore different aspects of the data as needed.
* **Validate and Deliver:** Review the final analysis for accuracy and provide clear, actionable recommendations to guide decisions.

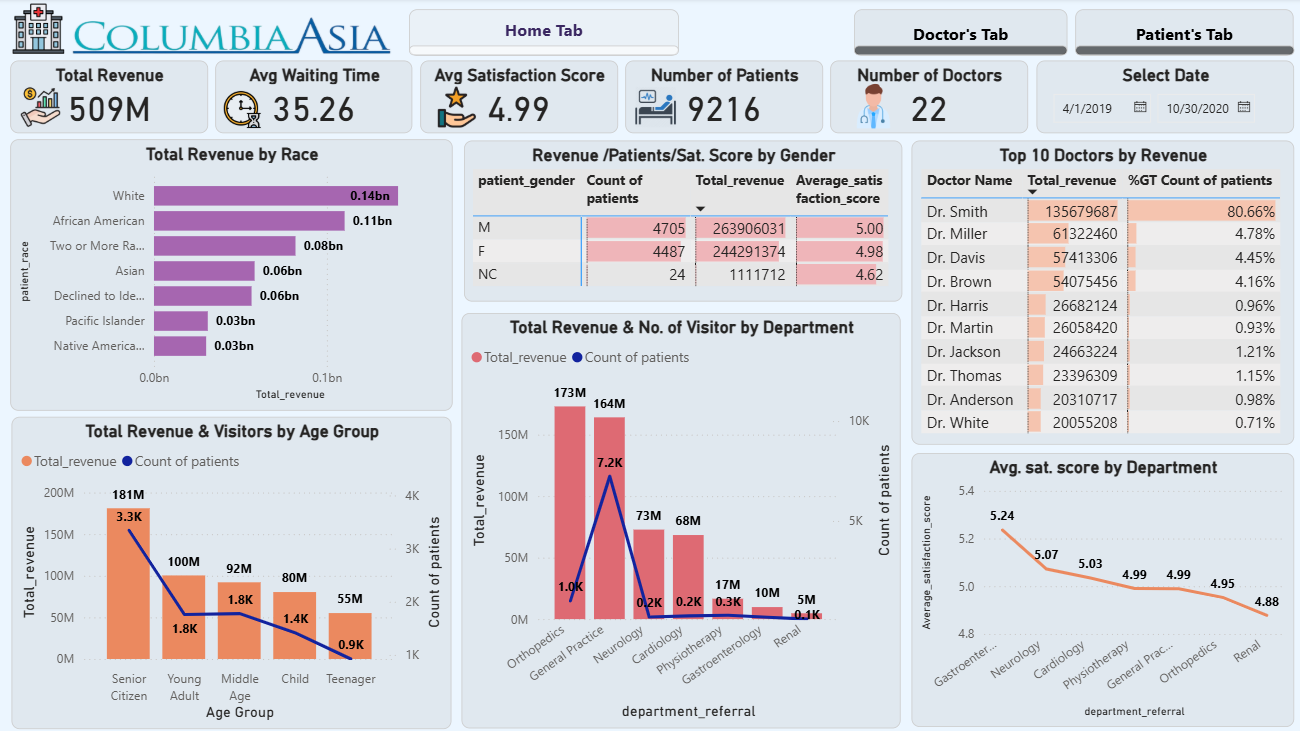
1. **Can you analyze and write the type of relationship between the doctor id and department, is it one-to-one?**

**Answer:**

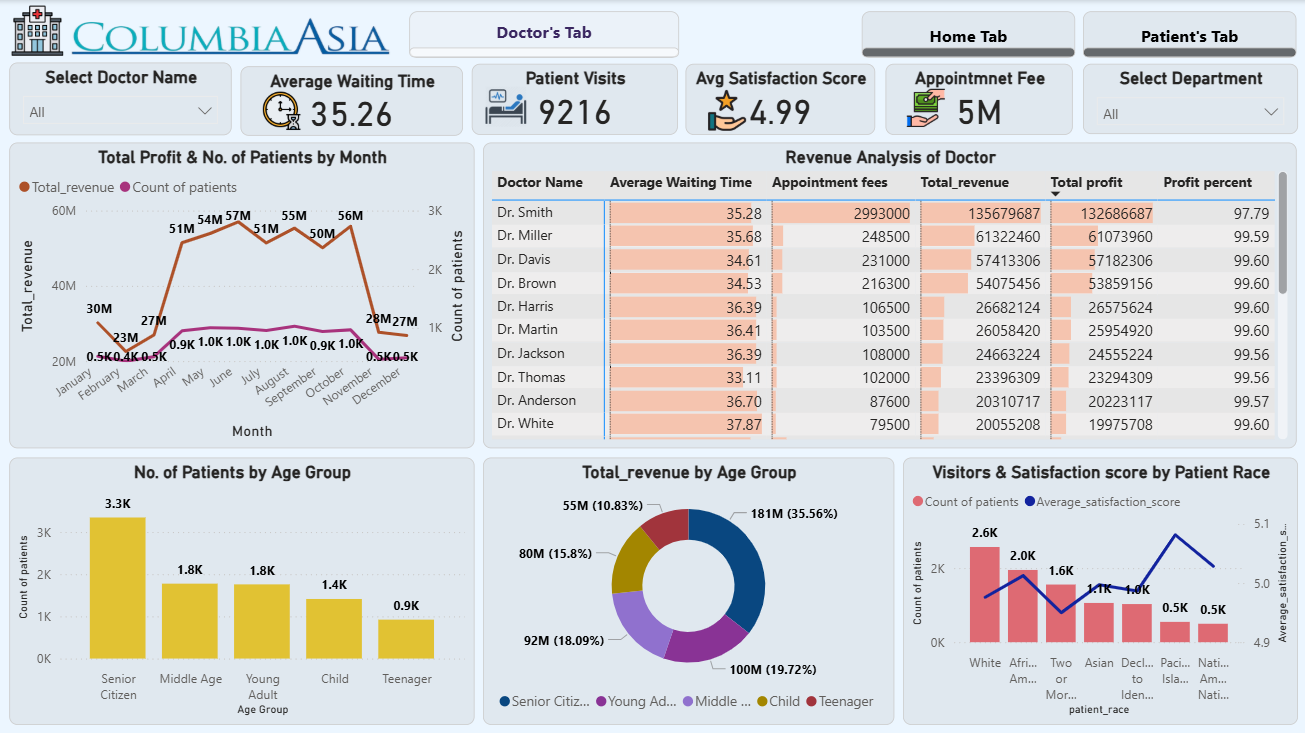
* Each **doctor ID** is linked to exactly one department.
* A single department can have **multiple doctors** assigned to it.
* This means the relationship is **one-to-many** from the doctor’s perspective to the department one department, many doctors.
* Alternatively, you could describe it as **many-to-one** from the department’s view — many doctors report to one department.
* So, it is **not** a one-to-one relationship — it is **one-to-many** (doctor to department).

Please find the below Dashboard Images

**Home Tab**



**Doctor’s Tab**



**Patient’s Tab**

