**COLUMBIA ASIA HOSPITAL**

Columbia Asia started its operations in 1996. Currently it has 22 medical facilities across Southeast Asia: 14 in Malaysia, 3 in Vietnam and 5 in Indonesia. The company believes in setting up mid-size hospitals built in residential areas for accessibility and efficiency, and to better serve the respective communities. This also helps keep costs down for consumers with no compromise on healthcare quality, modern amenities and highly-trained teams of specialists and nurses.

Columbia Asia aims to be the preferred choice of healthcare services for families and businesses. With 100 to 200 beds per medical facility, two-storey buildings for smoother patient flow and new medical technologies to decrease the levels of invasive treatments, the efficiency of patient care is optimised resulting in a shorter length of stay. In turn, this results in transparency in rates with better affordability for patients. But as progressive as medical technology can be, the difference in healthcare relies much on its people – highly skilled medical consultants, caring nurses as well as teams of committed staff who see to every patient’s need, on the road to recovery.



Columbia Asia medical facilities provide a wide array of medical services such as General Surgery, Paediatrics, Obstetrics & Gynaecology, Orthopaedics and Internal Medicine. These are supported by a comprehensive list of ancillary services that include an Intensive Care Unit, Neonatal Care Unit, Physiotherapy, Laboratory, Pharmacy and Imaging.  
  
At Columbia Asia, comprehensive medical programmes demand ethics, excellence and strict clinical governance. All of its operations follow international quality assurance guidelines that meet the highest standard of patient care possible.

**OBJECTIVE QUESTIONS**

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

While analysing the hospital dataset with Power BI, there were some missing entries in the data set, like patient sat score had some missing values, which I removed from data as it didn’t have any strong basis to justify the missing value.

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

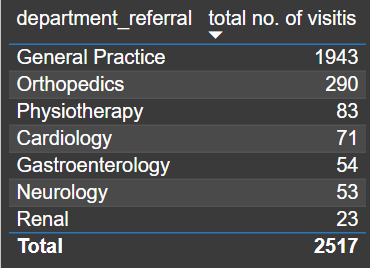
The average duration a patient spends before receiving care was about 35.35 minutes. The given image shows the dax used to calculate the average waiting time.



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

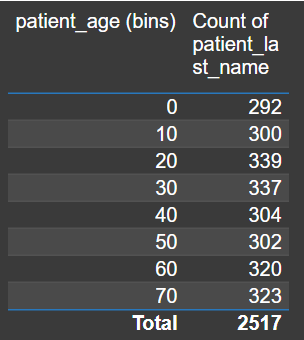
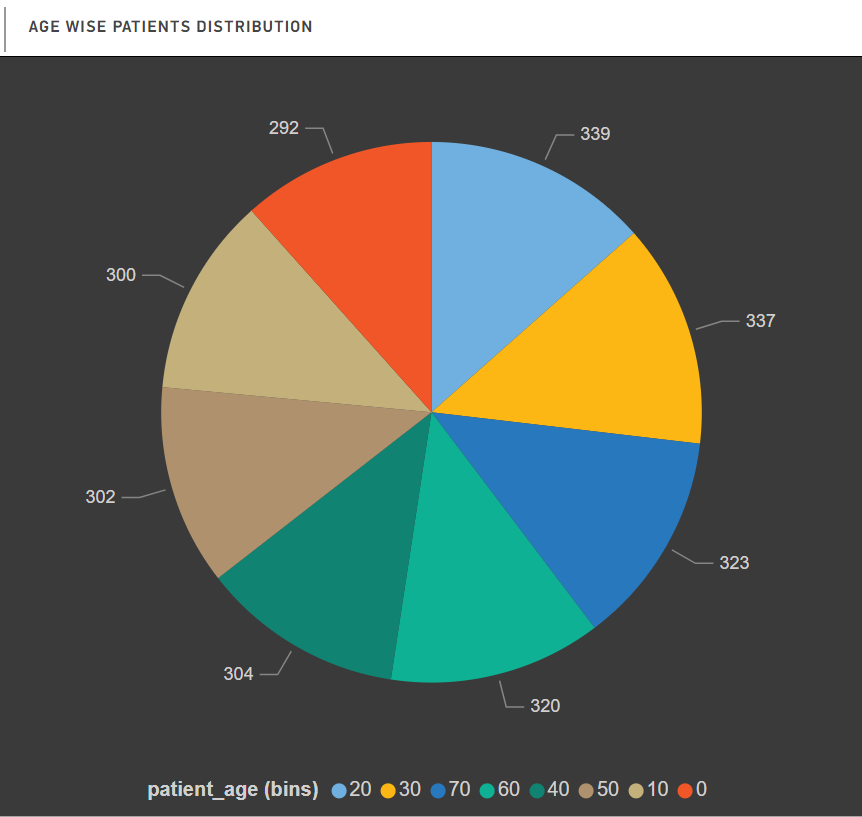
The total number of visits to each department based on referrals to understand which department are most frequently visited was found out first by using simple DAX function. And the further understanding was made that the General Practice (1943) was the most frequently visited department followed by Orthopedics (290).





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

For finding the above question I grouped the age column by using group measure and created a group bin of 10 years of interval.From the finding it was found out that the people from age group of 20 i.e. (20-30) utilizes the healthcare services the most followed by group of 30 i.3.(30-40). While the age group of 0 i.e. (0-10) utilizes the healthcare services the least.

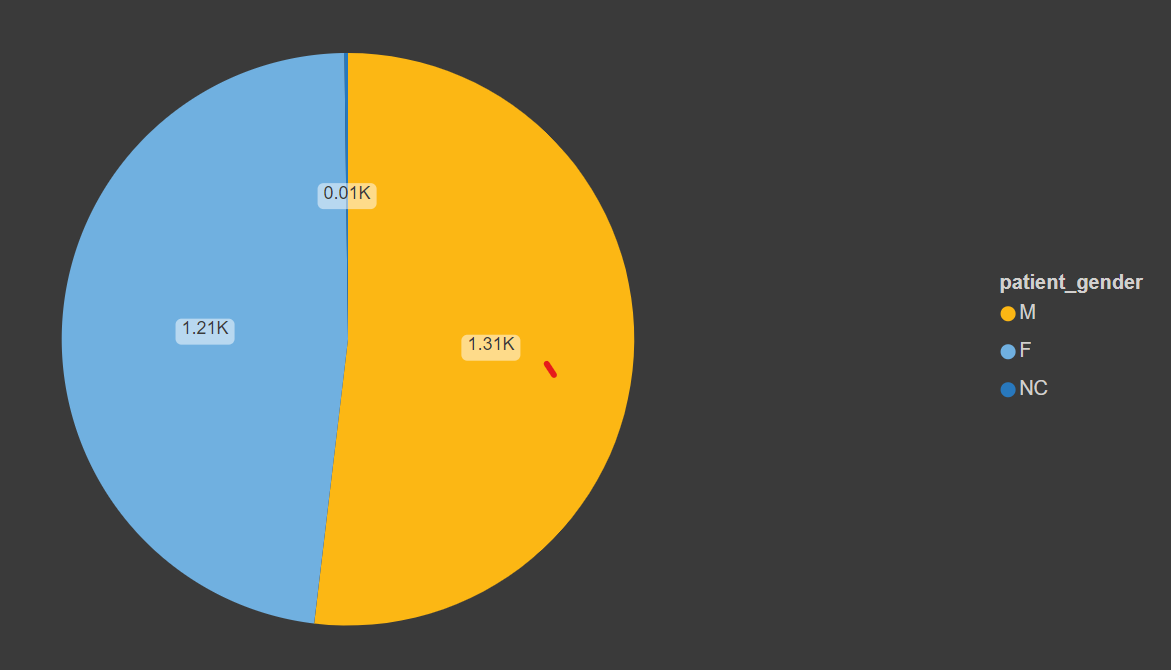
 

5. Were there any Null values in the data? What would be the best way to handle these Null values and which approach have you opted for?

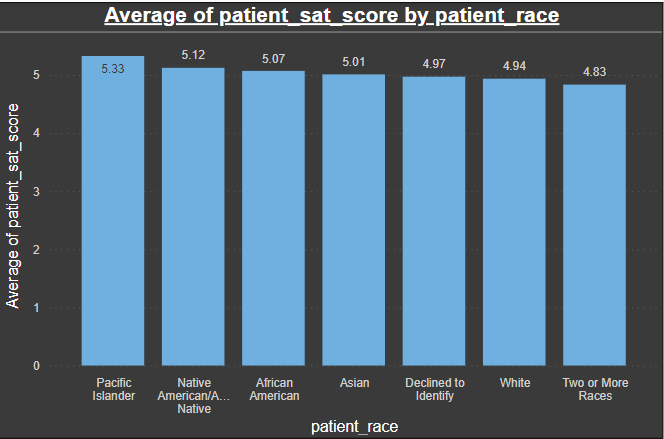
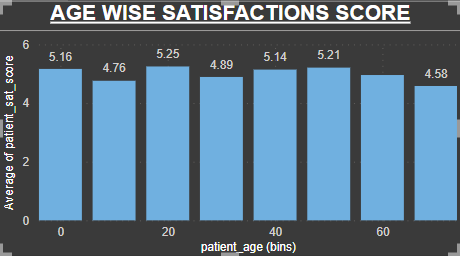
While analyzing the given hospital dataset it was seen that there were some null values present in the dataset named Hospital ER in the column named patient sat score I filtered out the null values as their were no specific reasoning provided by the team, so instead of replacing I replaced it as no true basis were made to fill the nulls.

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

From the provided dataset it can be seen that the maximum patient visited in the hospital are male (1306) while only 1206 female visited the hospital.



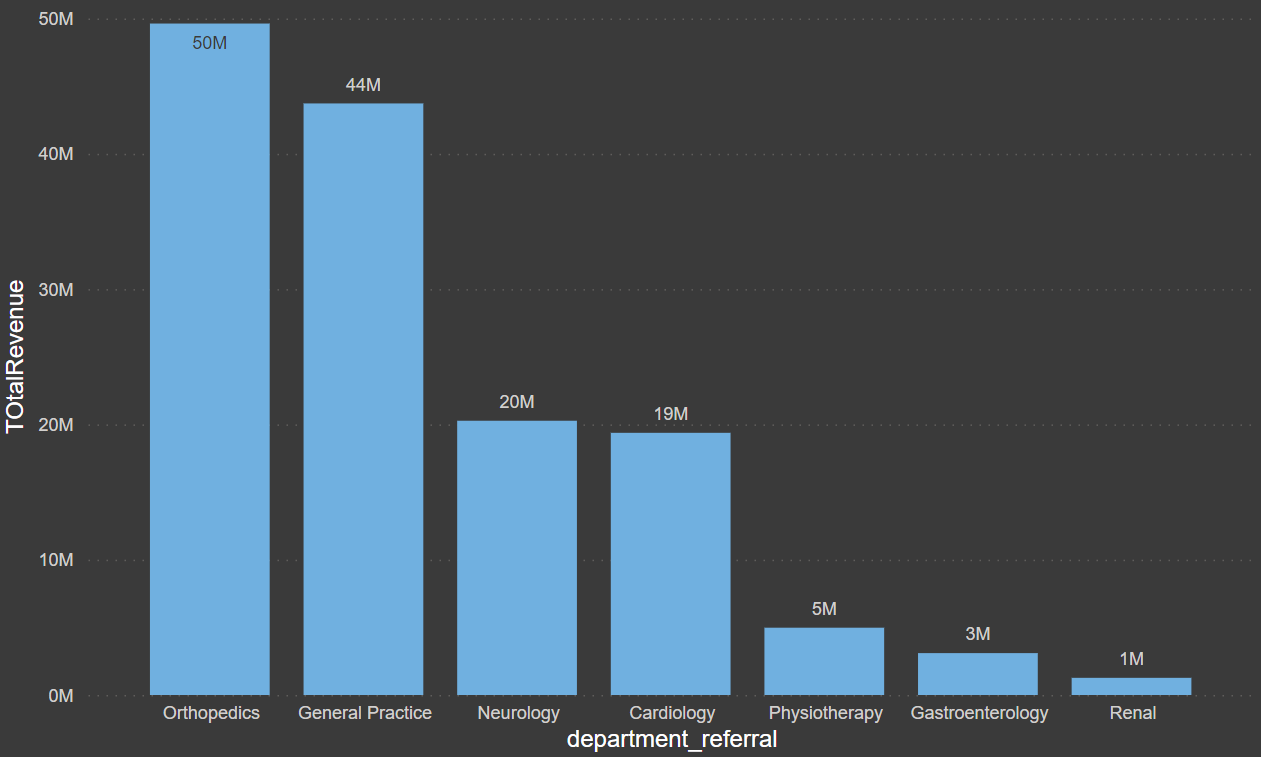
7. Average Satisfaction by Demographics: Determine the relationship between patient satisfaction scores, their age groups, and racial backgrounds to pinpoint areas for improvement in patient experience.

From the above graph shown it can be seen that the according to the age group there is declining trend in the satisfaction score of patients which can increased by giving prioritization to senior citizens in place of youths, which would help in increasing the patient’s satisfaction score. While in racial group it can be seen that the Pacific Islanders have highest satisfaction score that is 5.33. While the two or more races race have the least satisfaction score that is 4.83.

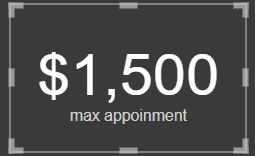
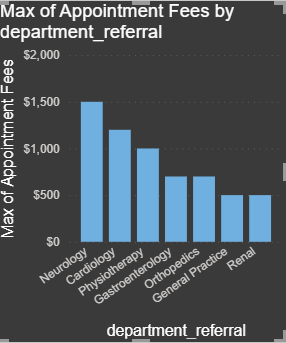
8.The hospital's managing director seeks to evaluate the revenue of each department to understand how much revenue is generated by each.

The overall revenue that was generated by the hospital was about 509M. And among all the department associated in the hospital, Orthopaedics has generated the highest revenue which is about 50M followed by General Practice department that is around 44M. the lowest revenue was generated by the Renal department that’s only 1M.



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

The Neurology department is charging the highest appointment fees in general that is around $1,500. The second highest appointment fees is charged by Cardiology that is $1,200. While the lowest appointment fees is of Renal and General practices department that is arounf $500.



(10. question is after 14th obj question)

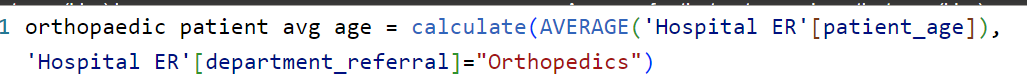
11. Using ‘Calculate’ and a row iteration DAX function calculate the total number of patients who have visited Dr. Smith.

The total number of patients who have visited Dr. Smith are 1592. The below pic shows the DAX function used to calculate the given statement.



12. 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?

The average age of the patients who visit the orthopedics is around 39.02. Whereas the average age of patients in all department is 39.94. the formulas used to determine the above statement is as follows:





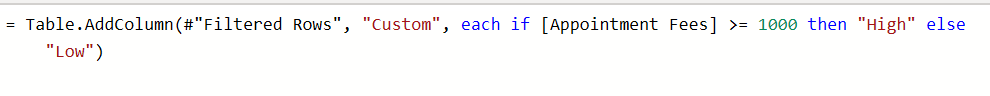
13. Were there any data format issues in the data, and if there were/are how you handle them?

There were some data format issues in the given dataset which was changed in different ways:

1. The datatype of the appointment fees and the total bills were in the text form which was changed into whole decimal number along with $ sign as it is depicting money.
2. The date column was given along with the time which I split it using split by column to get two different columns of date and time.

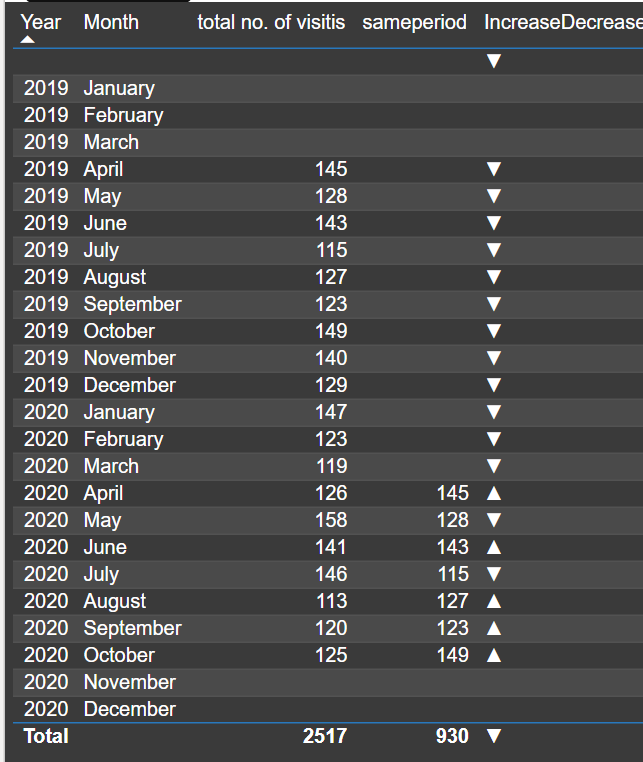
14. When we add a column in Power Query what’s the code that comes in M language in the formula bar? What do you know about M-query?

In order to answer the question above I added an additional column in power query editor based on appointment fees where appointment fees greater than or equal to 1000 are categorized to “High” or else “Low”. The below given image is the M-query of added column:



M-query is an acronym of word “mashup” query. It is used to query a large amount of data from data source. It is used during the data import which is the first step where data gets imported in data model of power bi desktop where the queries are using M in the background.

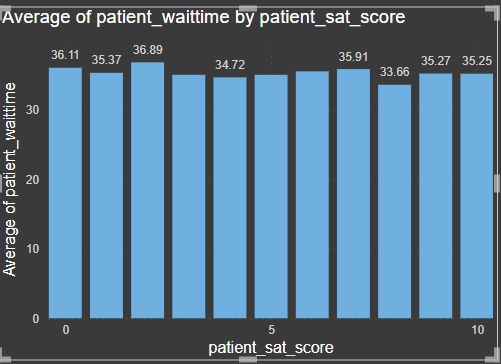
10.Create a tabular visualization in the Report view which consists of Month-wise total visits in the hospital.  Add a third column in the table that consists of the previous month’s total visits for each month’s row. Also, include a column that states whether the visits in a month are greater than that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

From the below given table it can be seen that there is increase in patients visit in 2020 as compared to 2019 in the month of April, June, August, September & October. There is a declining trend in the month of May & July ****

**SUBJECTIVE QUESTIONS**

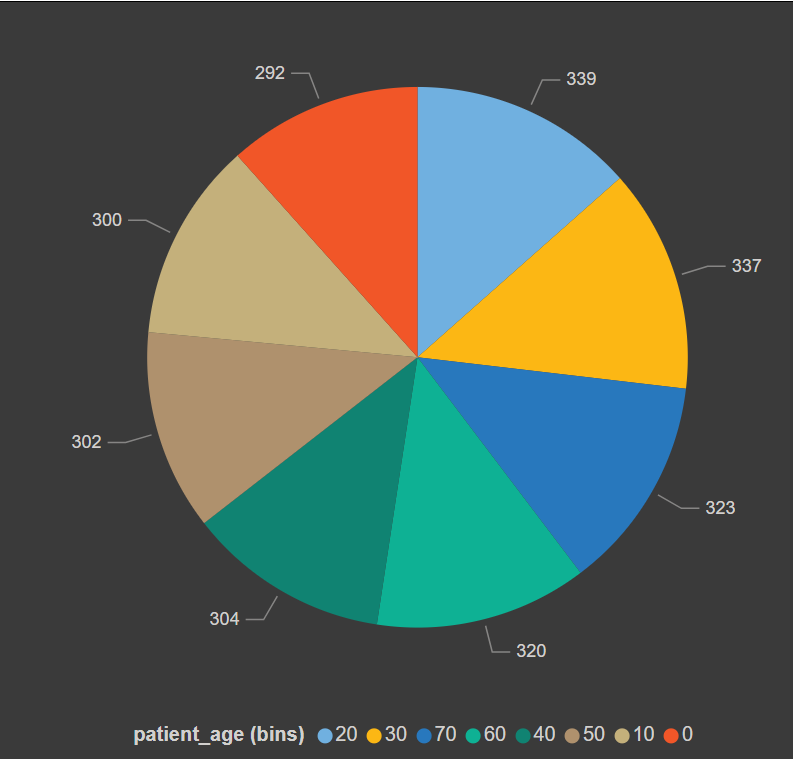
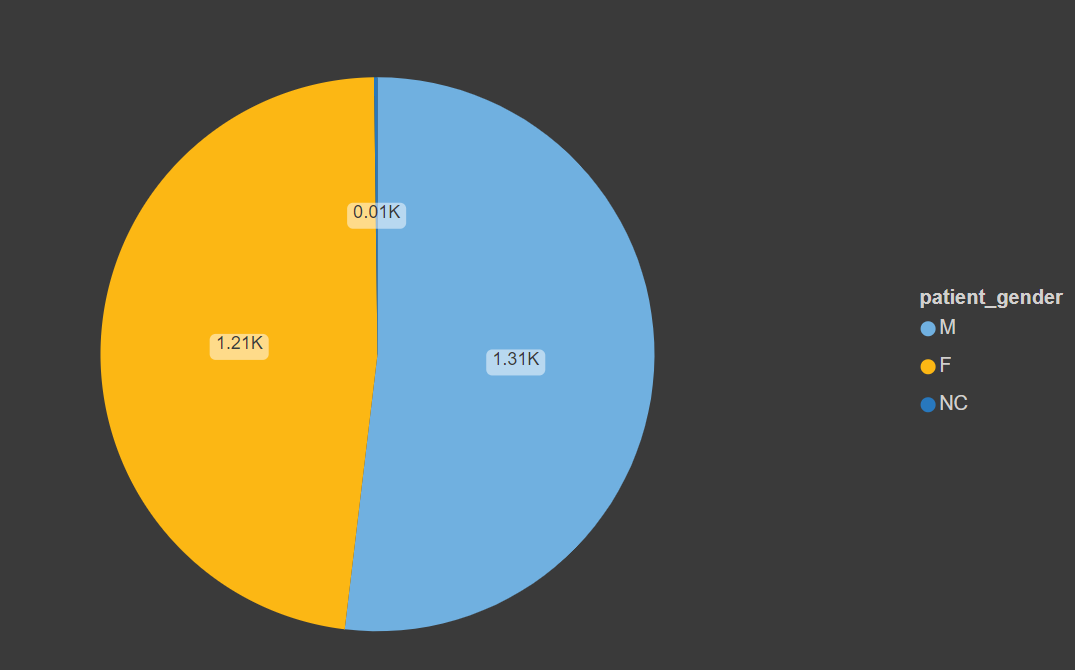
1.What is the relation between patient wait time and satisfaction scores?

We can say that the average patient wait time is inversely proportional to satisfaction score, in many cases its seen that the less average time wait has high satisfaction score as compared to the patient whose wait time is higher. It can be seen that low wait time has a score of 8-10 while high wait time has score rate of 0-3.

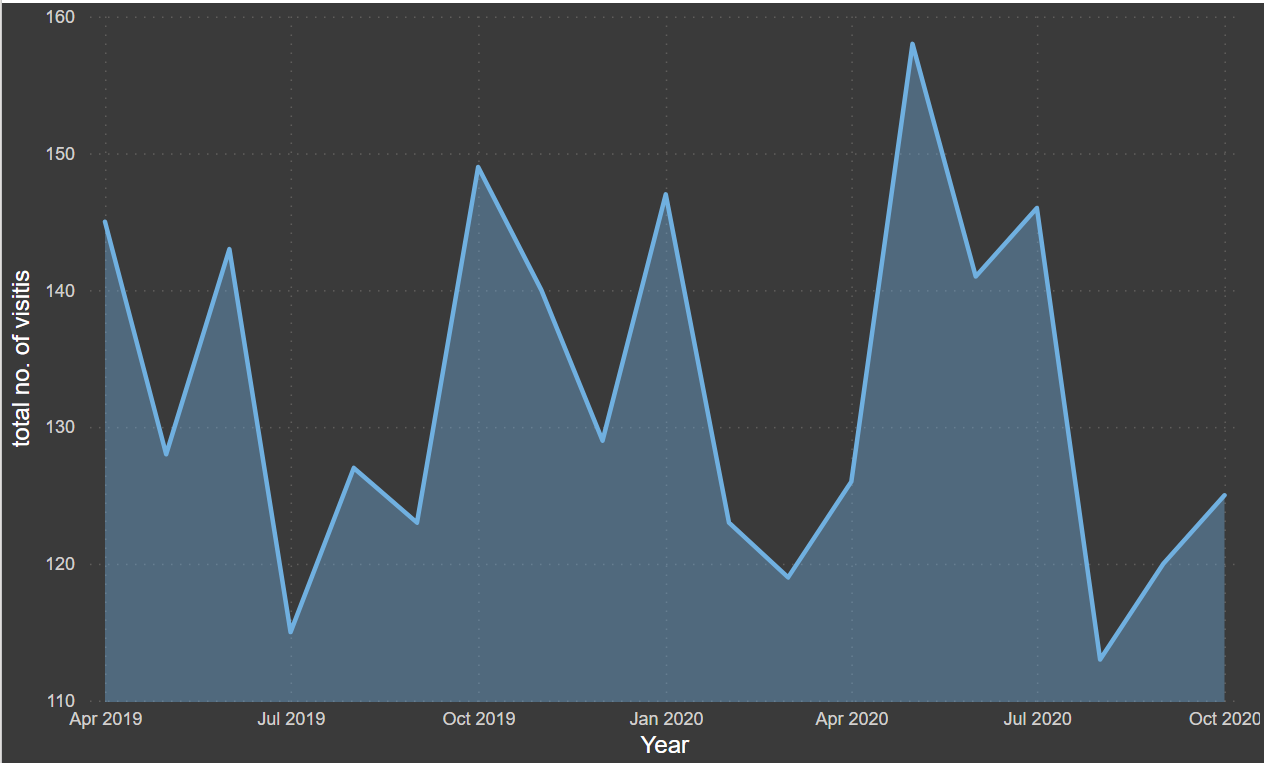


2. How do patient demographics affect the frequency of visits to different departments?

From the given data set it can be that the majority of the patient coming to the hospital are male that is around 1306 while only 1206 females visit the hospital. On the basis of the ages of patients the maximum number of patients that visits the hospital belongs to the age group of 20 (20-30) followed by the people of age group 30 (30-40). While the lowest number of patients belong to age group of 0 (0-10).

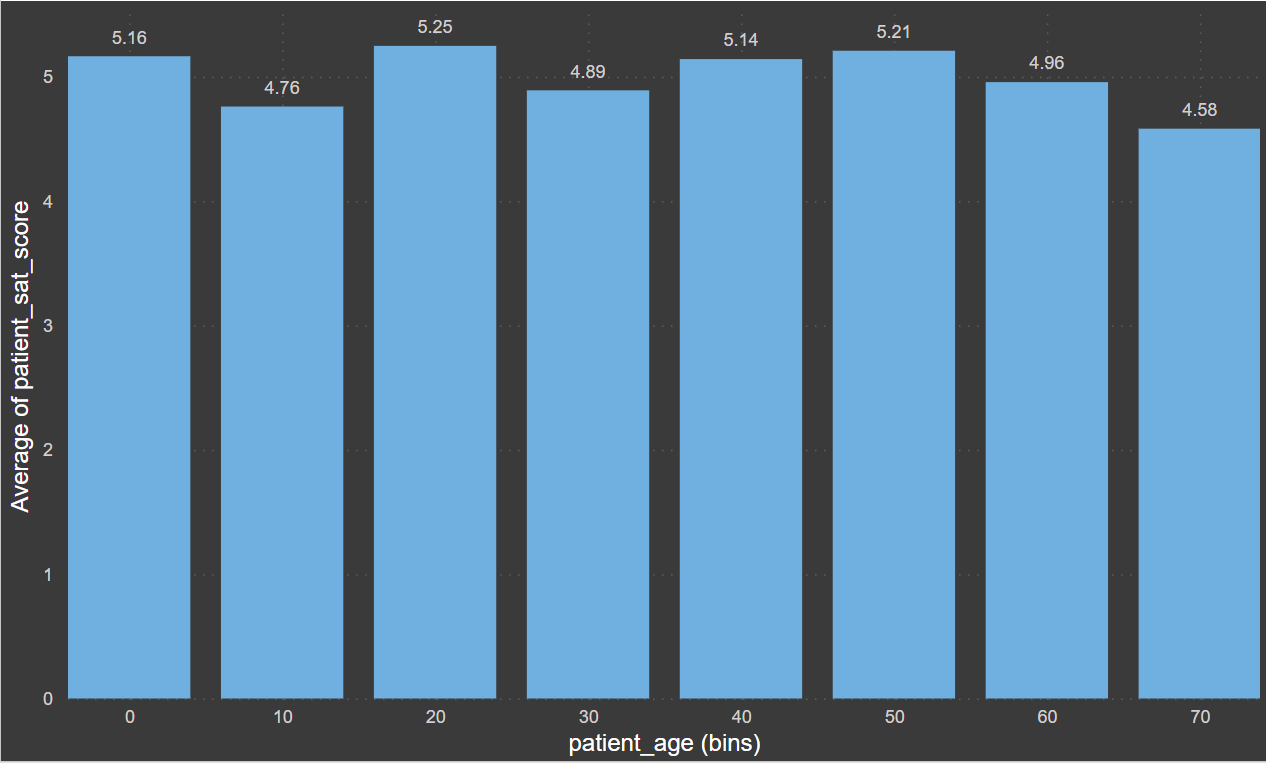
3. Is there a noticeable trend in the volume of patient visits throughout the year?



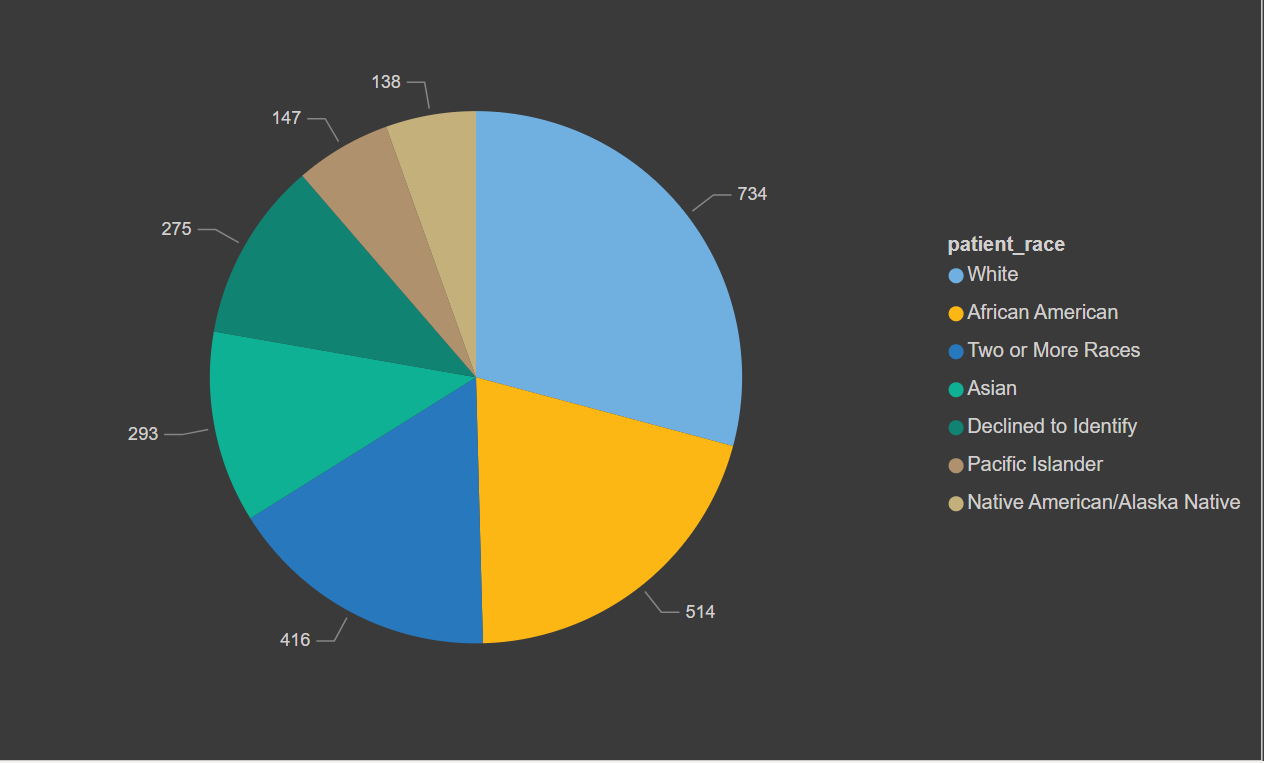
From the above graph it can be seen there is a decline in the number of visits from April 2019 to May 2019 which suddenly shoots up in June 2019 which later declined in July 2019. But as seen overall trend of number of visits to month it can be seen that there is fluctuation in every month. The highest number of visits can be seen in the month of May 2020 which is 158 patients which is followed by October 2019 which is 149. While the month of August 2020 shows the minimum patients.

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

The age group of 20 (20-30) shows the highest satisfaction score i.e. 5.25 which is followed by age group of 50 (50-60) i.e. 5.21. While the age group of 70 (70-90) shows the lowest satisfaction score i.e. 4.58. (used average satisfaction score vs age group)



5. 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?

****

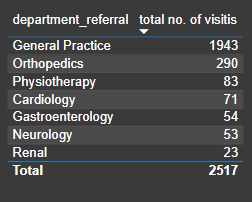
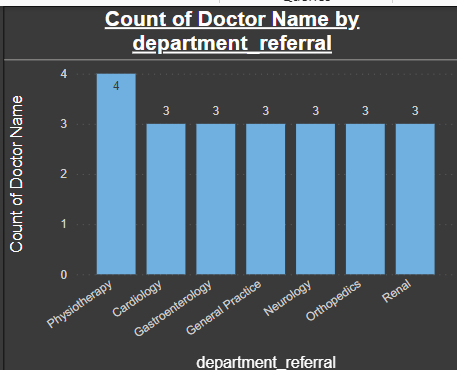
So, if someone outside of the hospital claims that there is racial or gender-based discrimination in the hospital, it can be said that the above claim was not true as it can be that second majority of the patient in the hospital visited are African American, not only that but more than 5 racial community comes here for the treatment. Not only by racial but along with that the female patients are in good number which is only 100 less than that of male patients. So the hospital is non-biased in terms of race and gender of patients.

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

The hospital management should offer discounts to the patients of the age group of 60 (60-70) & 70 (70-80). As they are seniors and need more assistance as compared to the youth generations, thus it will help in increasing the satisfaction score of patients and there can be increase in the patients of this age group also

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

If the hospital has a budget to hire 2-3 new doctors, they should be hired in the Orthopedics, General Practice and Physiotherapy, as they have highest numbers of visitors and a smaller number of doctors. Thus, it will help in increasing the revenue as these departments are busier as compared to others.

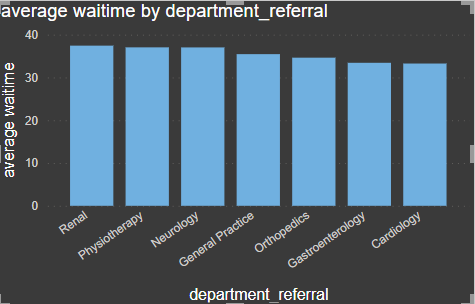
** **

8. Is the hospital profitable? How will you determine the profitability?

If as compared to the appointment fees taken by the doctors to the total revenue earned by hospital, it can be said that the hospital is profitable. As the total fees by the doctor is 5.35M. thus by excluding that from the total revenue determines the profitability.

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

The average waiting time of the renal department is oddly large as compared to other departments which is around 37.43 minutes. Which is followed by Physiotherapy department which is 37.04 minutes. While the cardiology has the least waiting time that is 33.25 minutes.

****

10. Come up with strategies to provide discounts to the patients.

The strategies we could use to provide discounts to the patients should be like giving 50% discount in the appointment fees of doctor to all the seniors category of 60 (60-70) & 70 (70-80). And we can also provide extra free facilities to the senior citizens like massage area or free regular tests. Thus, it will not increase the number of patients but also increase the satisfaction score of patients by attracting new patients.

11.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?

As the General Practices is the busiest department in all the department, we can align the doctors in two shifts based on demands. The shift timing can be divided into 12:00 AM to 12:00 PM and one is 12:00 PM to 12:00 AM. As of now there are total 3 doctors available in department so would recommend to hire 3 more doctors in the same department so that there will be 6 doctors overall, making which 3 doctors can work in 2 respectively given shifts.

12. What do you understand by Power BI gateway? What are its use cases?

Power Bi gateway is a tool provided by Microsoft that allows you to connect your on-premises data sources to Power Bi in the cloud. It acts as a bridge between your organization’s network and Power Bi, facilitating data transfer securely. This enables you to refresh datasets in Power Bi with data that resides on your local servers or databases, ensuring that your reports and dashboards are up to date.

It is primarily used in scenarios where organizations have on-premises data source that they want to connect to Power Bi in the cloud. Some common use cases include:

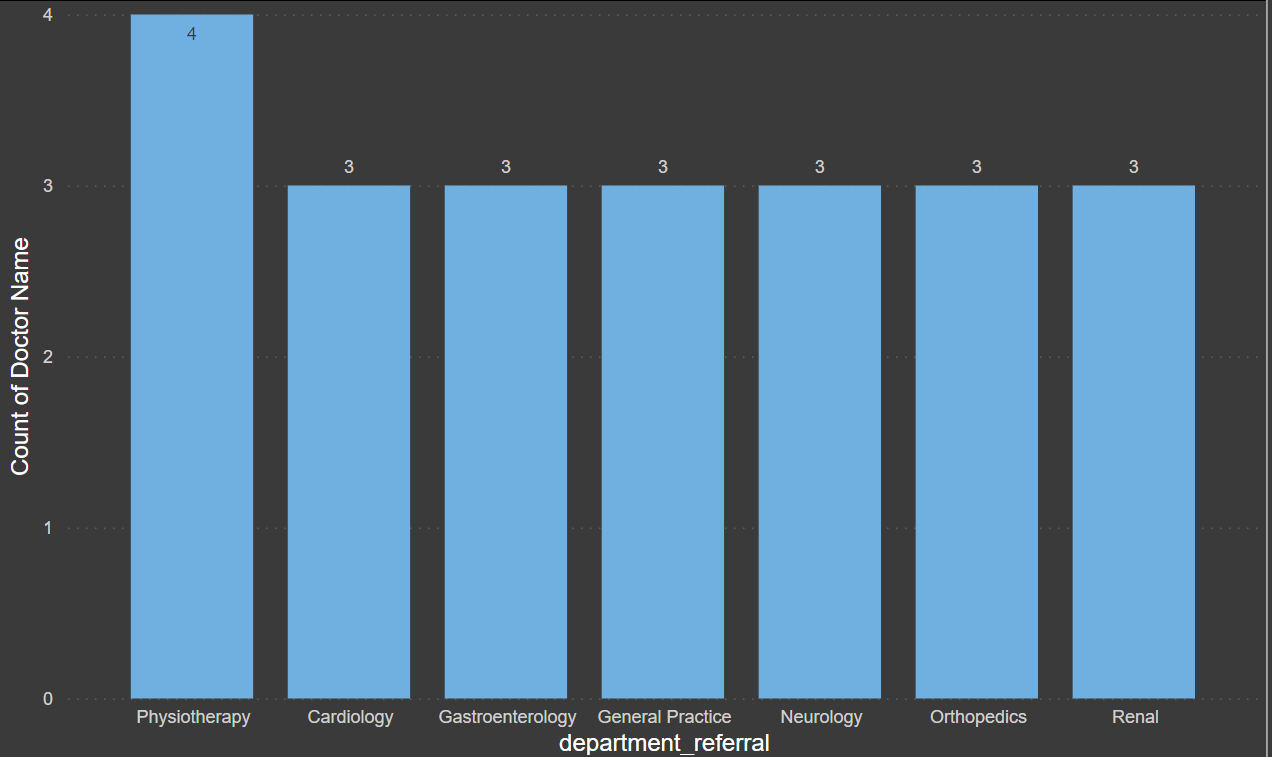
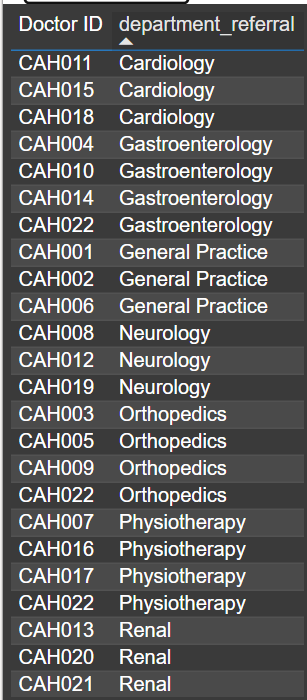
1. Accessing On-Premises Data.
2. Real-Time Data Refresh.
3. Hybrid Scenarios.
4. Scheduled Data Refresh.
5. Data Security and Compliance.

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

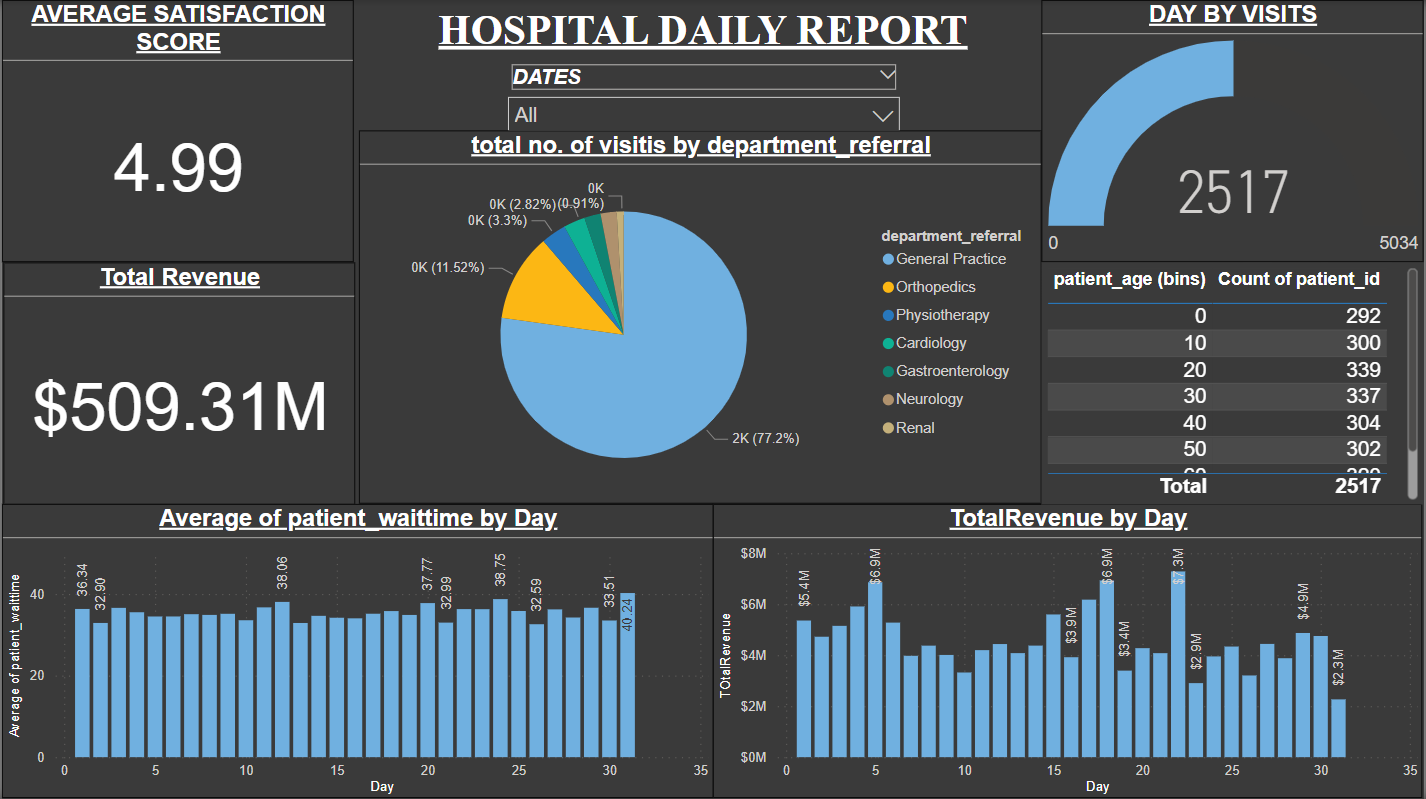
If the objective and subjective question weren’t given, I would have made a two-report tab. In one I would have made doctor’s tab and patient’s tab. Where I would have shown the same thing which is shown in this project. But I would have a different approach like how could the waiting time can be reduced, how we can improve the satisfaction score rating, etc.

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

There is a one-to-one relationship between the doctor id and department, as each doctor is assigned to single department. Like one doctor is associated to single department. There is no single doctor who are assigned to multiple departments.

**REPORTS**

****

