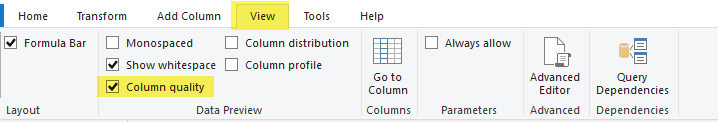
**Objective Questions:**

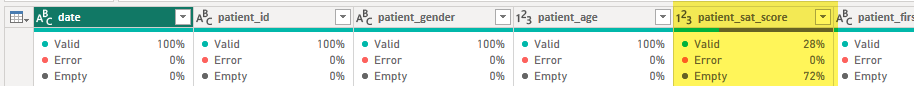
1. In analyzing the hospital dataset with Power BI, ensure data cleaning to address inconsistencies and missing values before further analysis.

In order to ensure the data cleaning to address the inconsistencies and missing values , we have load the data in the power BI and selected the option “Transform Data” to open it in the power query editor .

In Power Query editor we have routed to view option and add the column quality.



In both the sheets , the column quality was 100% valid except one column “*Patient\_sat\_score”* in sheet *“Hospital ER”.*



**We have removed the null values from here.**

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

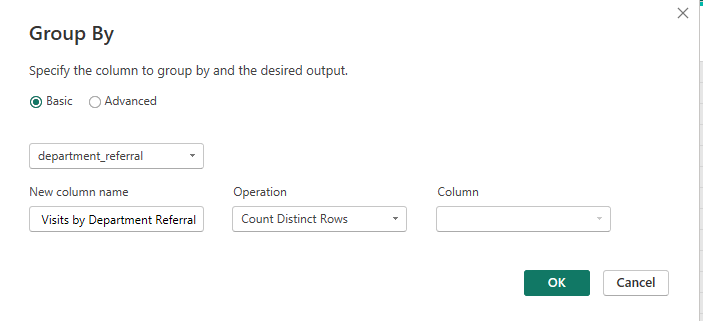
We have assessed the average waiting time by creating a new measure in the sheet “*Hospital ER*”.

Patient\_Average\_Wait\_time = AVERAGE('Hospital ER'[patient\_waittime])

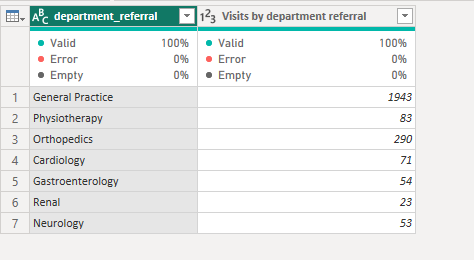


1. **Visits by Department Referral:** Calculate the total number of visits to each department based on referrals to understand which departments are most frequently visited.

In order to find out the visits by Department Referrals , we have used the Group by feature of Power BI in Power Query Editor and selected the operation as “Count Distinct Rows” and got the desired results.

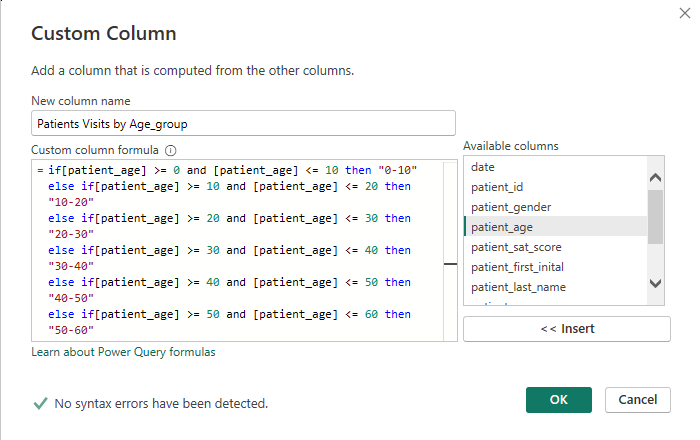
****

Results without filtering the empty satisfaction scores

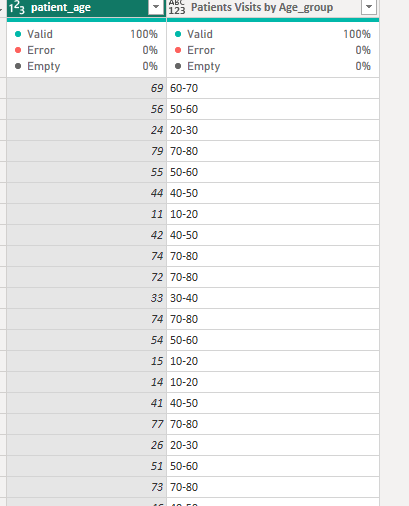
****

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

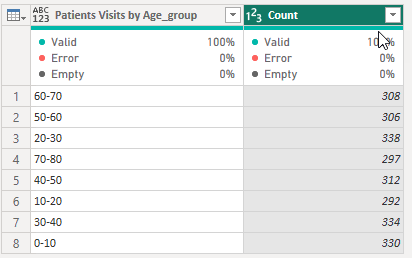
**In order to segregate the patients by their Age group we have created a custom column**

****

By doing this , it have created a new column which states if the patient age is 33 , he will be falling under Age group 30-40.



Then we have selected the “**Group by”** for this new Custom column “*Patients Visits by Age\_group*” which has provided us the desired results.

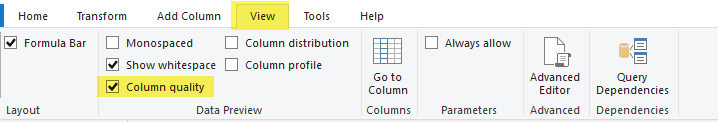


1. Were there any Null values in the data? What would be the best way to handle these Null values and which approach have you opted for?

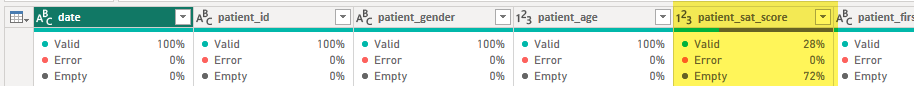
Yes , there were null values in the data as mentioned in the Objective Question 1 .

We have loaded the data in the power BI and selected the option “Transform Data” to open it in the power query editor .

In Power Query editor we have routed to view option and add the column quality.



In both the sheets , the column quality was 100% valid except one column “*Patient\_sat\_score”* in sheet *“Hospital ER”.*



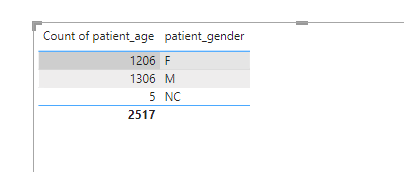
**We have removed the null values from here.**

There are numerous approach to remove the null values but I we’ve opted for the above.

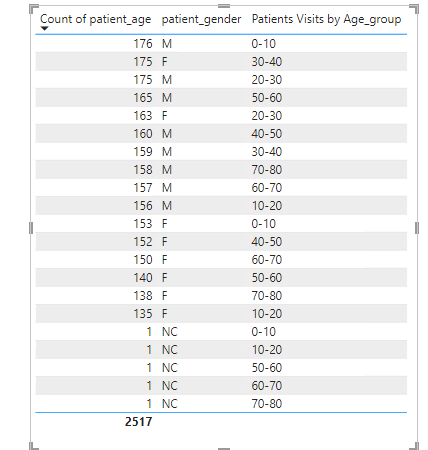
1. Is there any relation between the number of visits and the Gender of the patients?

As per the data , there were more male patients than Female. We have used the table visual here.

Below is the data.

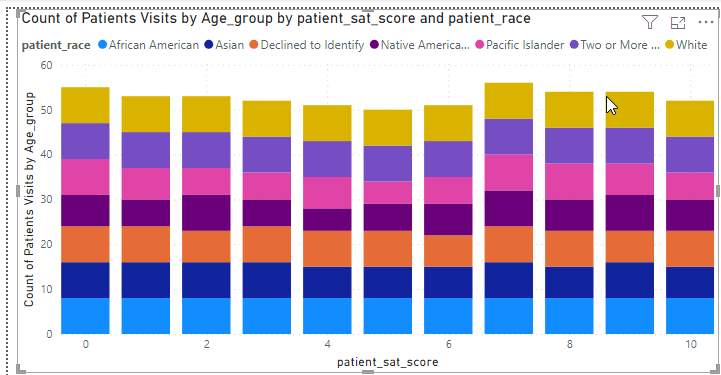
****

Furthermore , we can categorize it by age group as well as we already have the age group column.



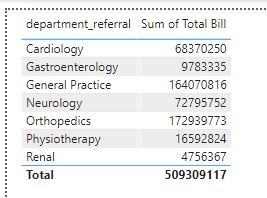
1. Average Satisfaction by Demographics: Determine the relationship between patient satisfaction scores, their age groups, and racial backgrounds to pinpoint areas for improvement in patient experience.

We have used the Stacked column chart here to determine the relationship between patient satisfaction scores, their age groups, and racial backgrounds.

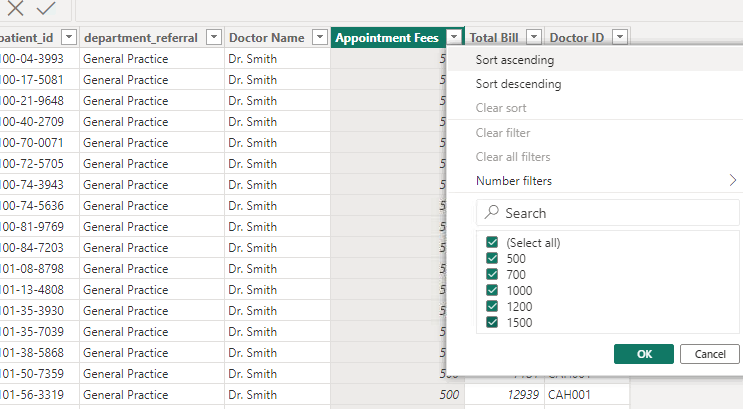


1. The hospital's managing director seeks to evaluate the revenue of each department to understand how much revenue is generated by each.

We have used the table visualization here and selected the columns “department\_referral” and “Total Bill” which provided us the total revenue generated by each department.



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

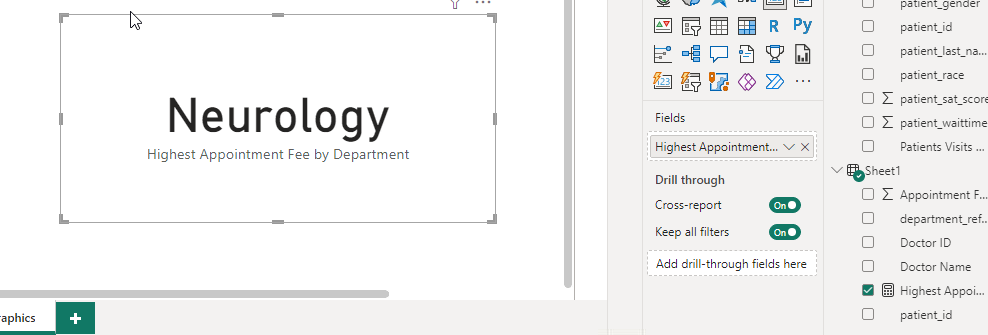
In order to find the department who is charging the highest appointment fees in general ,we can simply sort the *“appointment fee*” column in Descending order 

We can see that “Neurology” department is charging the most. 

But as instructed, we have to use the DAX function so we have created a new measure which is below

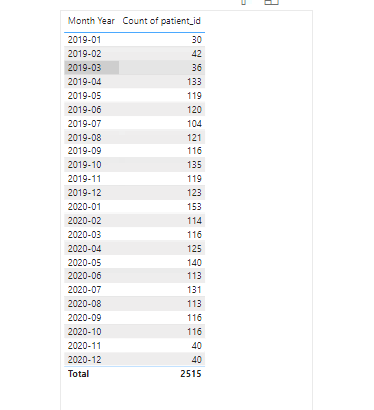
**Highest Appointment Fee by Department = SELECTCOLUMNS(TOPN(1,Sheet1,Sheet1[Appointment Fees],DESC),"Department",Sheet1[department\_referral])**

It has provided us the desired result **.**



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.

We have created a new column for the Month and year using FORMAT(‘Date’,”YYYY-MM”) and then created a tabular visualization for the same . The below table is indicating the count of patients in the particular month.

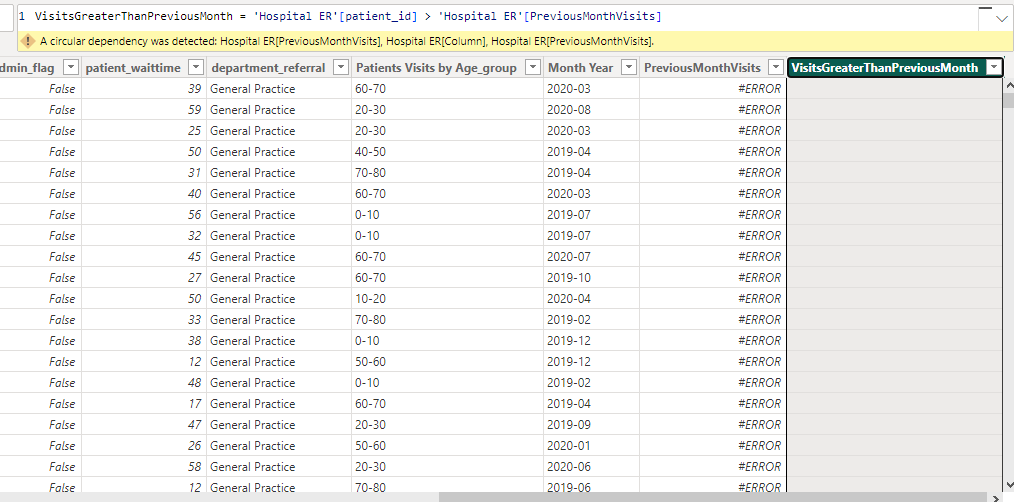


For PreviousMonth we have created a new column using formula *PreviousMonthVisits = CALCULATE(COUNTROWS('Hospital ER'), PREVIOUSMONTH('Hospital ER'[date]))*

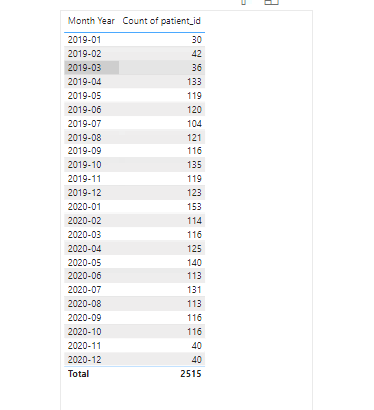


We have removed the duplicates from Model view as well as from PowerQueryEditor but it’s giving us the error. I’ve tried it for few days but unable to solve it.

Also to compare the current month and previous month ,we could have used the below column and add it in the tabular visualization but unfortunately we couldn’t find the previous months visits.



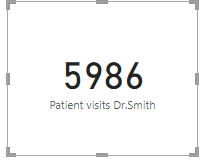
So this is the final output.



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

We have used the below DAX function here to solve the same

*Patient visits Dr.Smith = CALCULATE(COUNTROWS(Sheet1),FILTER(Sheet1,Sheet1[Doctor Name]="Dr. Smith"))*

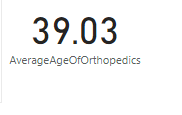
**

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?

In order to determine the average age of patients who visits Orthopedics dept. we can use the below DAX function

*AverageAgeOfOrthopedics = CALCULATE(AVERAGE('Hospital ER'[patient\_age]),FILTER('Hospital ER','Hospital ER'[department\_referral] = "Orthopedics"))*

And it’s giving us the result 39.03



If we were to determine the average age of all the department ,we can just take the average of patient’s age . Here in orthopedics we have filtered the department.

*Average Age All dept = AVERAGE('Hospital ER'[patient\_age]) = 39.95*

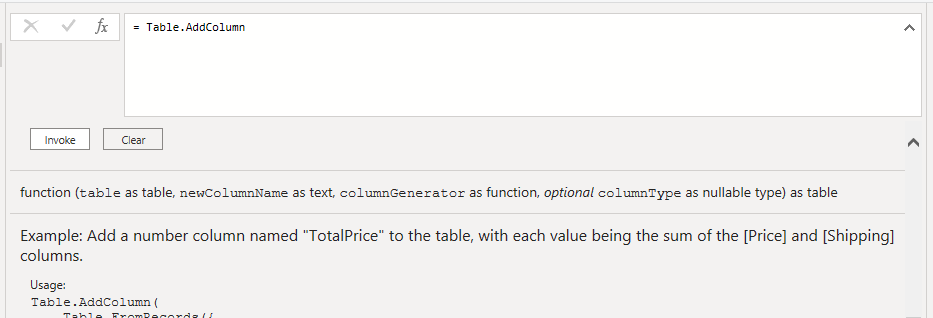
1. Were there any data format issues in the data, and if there were/are how you handle them?

Yes , there is an issue in the date format . I can’t use the PreviousMonth function ,error stated there is duplicate values , removed those

Still unable to do so .

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

= Table.AddColumn(#"PreviousStep", "NewColumn", each [Column1] & [Column2])



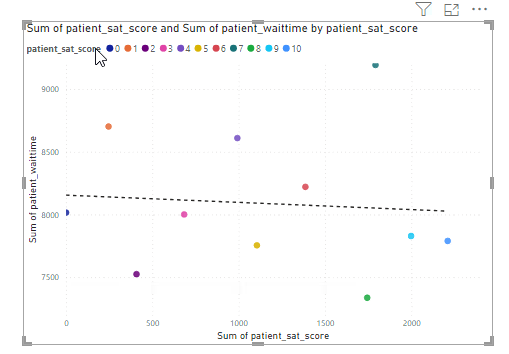
M-Query represents the transformation applied in that specific step.

Power Query steps are cumulative ,so each step builds upon the previous.We can see it in the formula bar.

**Subjective Questions**

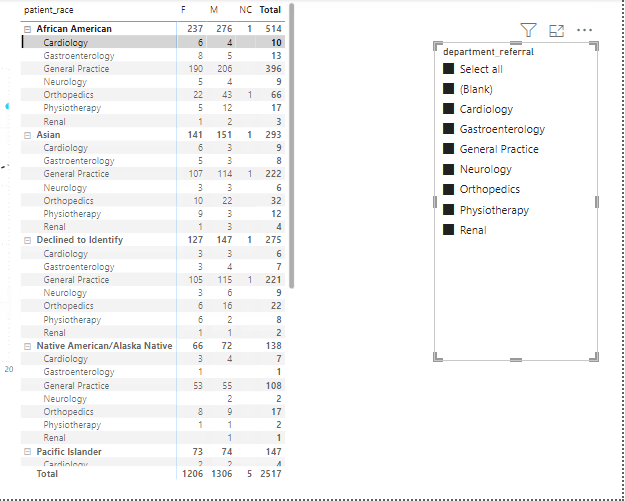
1. What is the relation between patient wait time and satisfaction scores?

In order to determine the relation between Patient’s Wait-time and satisfaction score, we have used the scatter plot. However we have aggregated it on the basis of sum as it’s not displaying the raw data.



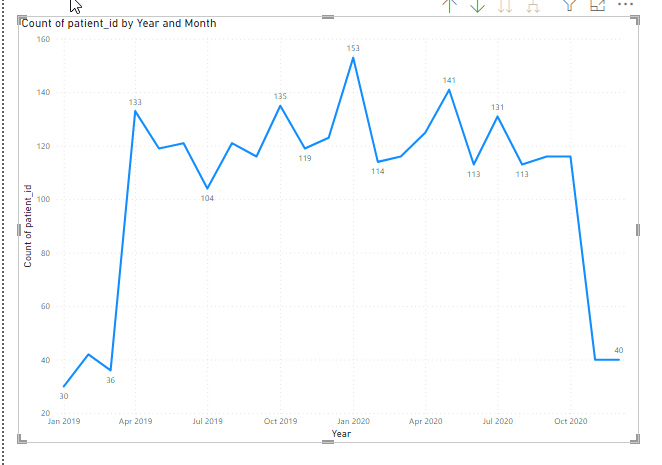
1. How do patient demographics affect the frequency of visits to different departments?

Here we have used the matrix visualization and added different fields such as Patient’s age , Race, department referral and gender . Also we have added slicer as well to categorize.



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

We have used line chart here to check the trend , on one axis we have taken year and month and on another axis we took count of Patient ID and as we can see in the below chart , in Jan 2020 there are comparatively more patients.

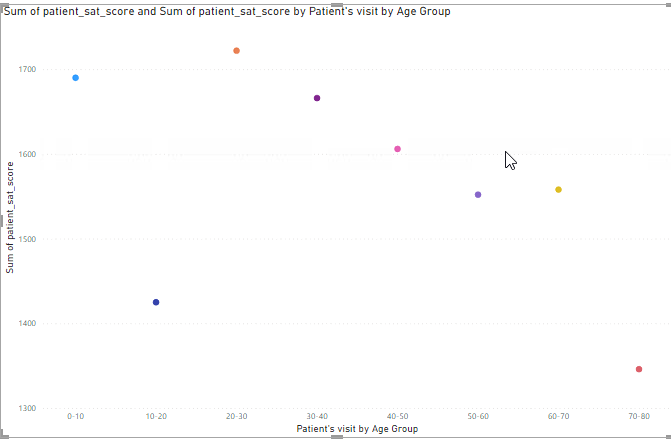


1. Which age groups report the highest and lowest satisfaction scores?

Age group from 20-30 are having the highest satisfaction score.

We have used scattered plot here and on x axis , we have used the age group which we created earlier and on Y axis we have took the sum of the satisfaction score.

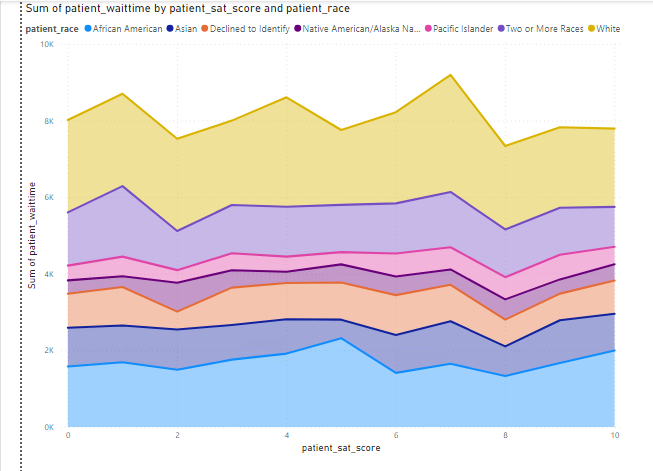
As we can see it in the below chart.



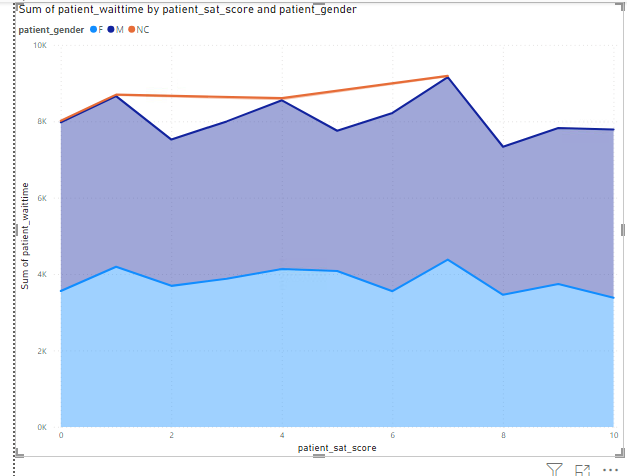
1. Say someone outside of the hospital claims that there is racial or gender-based discrimination in the hospital, how will you identify whether the claim was right or not?

That can only be observed when we enter in the hospital. We shouldn’t pay much heed to such claims but if we hear such claims , even if we’re not paying much attention to it they will always be there in the back of the mind. For such claims we can use the data we have to verify if there is any truth or not .

We have used the Area chart and verified that the People with white race are having better satisfaction scores compared to African American which are at bottom . We can conclude here that , race based discrimination is definitely there.

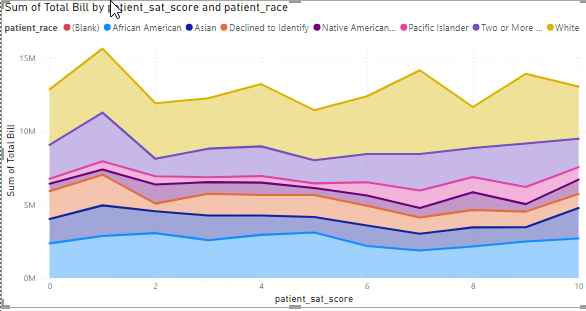


Also we have used the Gender on one axis and verified the patient’s satisfaction score. As we can see male are on top tier however females sum of satisfaction is comparatively below.



If we check the data of Total sum collected by patients and distribute them on the basis of race , it’s same as the satisfaction score.

People with White race are having high satisfaction score also they are charged comparatively higher , this concludes that number of white race are higher which resulted in their higher satisfaction score as well as high total bill. Below is the data on the basis of total bill.



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

There are number of factors which hospital management have to look to offer the discounts/offers.

It can be on the basic of Financial need , suppose if a patient is poor and having critical condition but we don’t have the data here for the financial condition of the patients.

It can be based on the medical condition or the severity of the medical condition.

It can be based on the stay , if a patient is admitted in the hospital and his condition is somewhat where he/she has to stay for longer , that can also be counted as a factor.

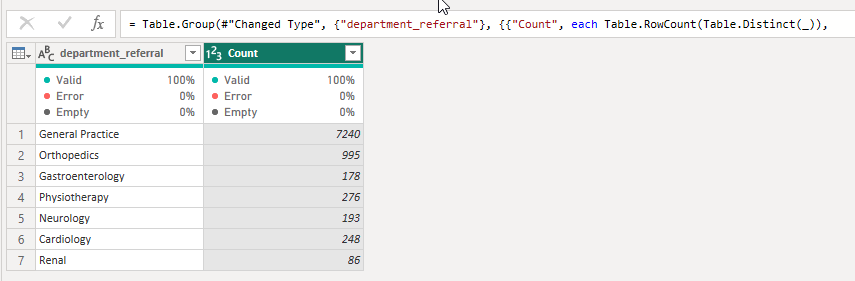
It can be on the basic of Age, race or gender .

Frequent visits can also be considered as a factor here.

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

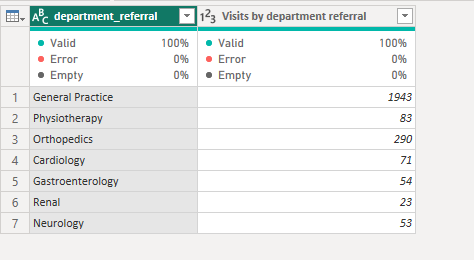
As per our data , we have seen in above objective question #3 that there are many patients for the general practice (with null as well as not null values)

#With NULL Values



#Without NULL values

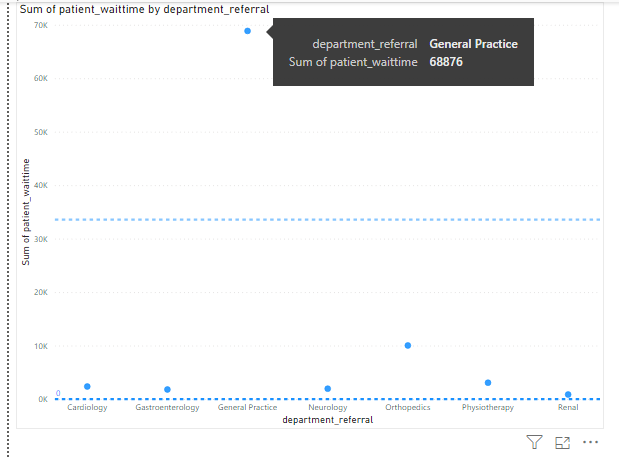
#Without NULL values

****

Also we have used the scattered chart over here to verify if we really need additional doctors for the general practice.

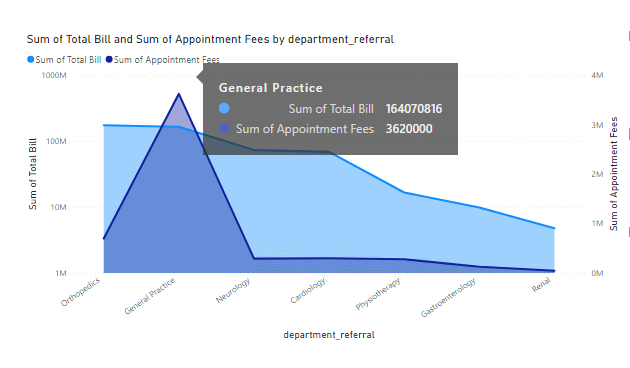
So we have looked at the patients wait-time and as we can see the patient visiting hospital for the general practice have to wait longer for the doctor comparative to other department. So we can hire additional doctor for the general practice if we are having a budget for the same .

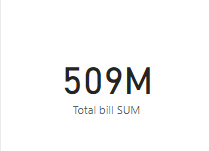
Below is the chart



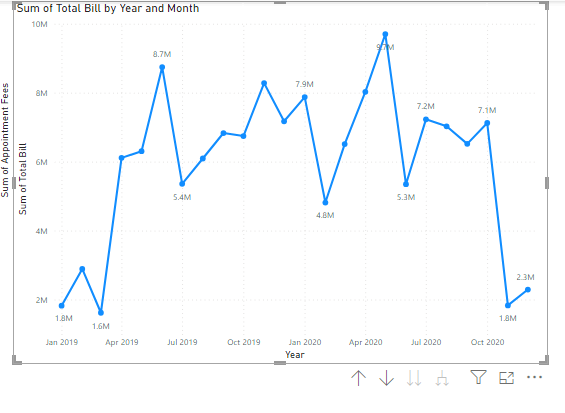
1. Is the hospital profitable? How will you determine the profitability?

We don’t have data where hospital is spending . We have only the data available for the total bill, appointment fee by department.

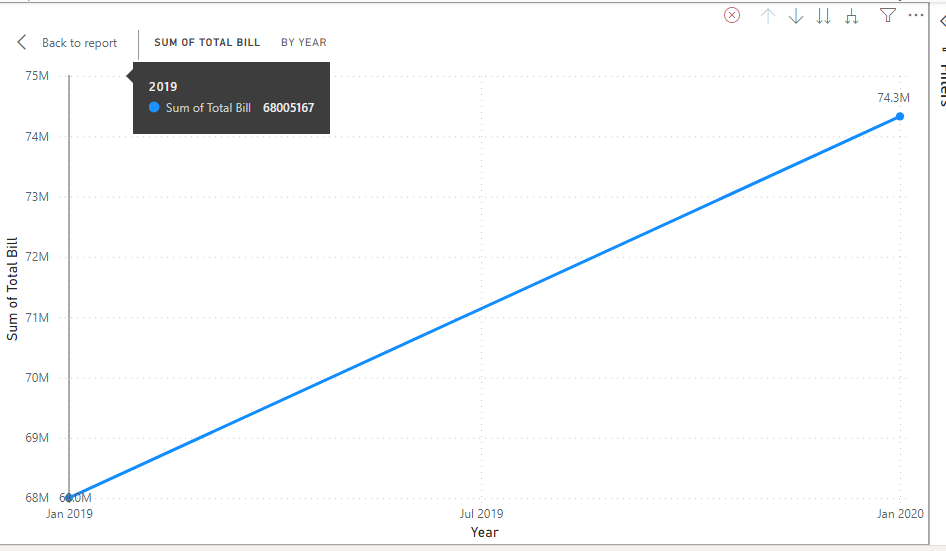




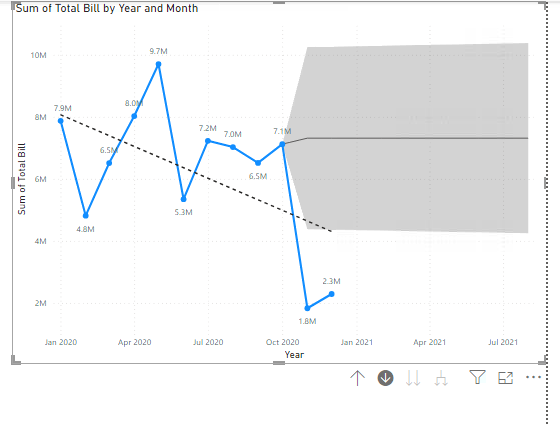
Below is the line chart for the revenue of the hospital by Year and month and as we can see in May 2020 the revenue peaked.



On the basis of the below chart we can say the hospital is profitable .



If we forecast it on the basis of the data provided , below are the expected results.



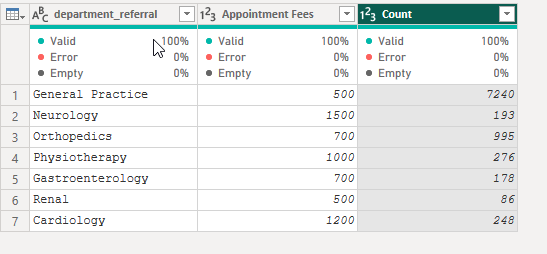
1. Any Department for which the waiting time is oddly large?

Yes, it’s for the General practice ,we have seen the data/chart in the subjective Question #7.

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

Designing effective strategies for providing discounts to the patients involves various factors.We have already discussed it in the subjective question #6.

Income based discount: Offer higher discounts to patients with lower income source to support their wellbeing also Neurology department is charging higher appointment fees , Discount can be given on the appointment fees as well for those patients.



Discount can be offered on the basis of the stay period in the hospital if the patient is admitted. We don’t have data for the stay period of the patient that’s why we couldn’t support this statement with the appropriate graph.

If a patient is availing multiple services then we can be offered some discount.

Frequent visit based : If a patient is visiting the hospital frequently then also he can be offered discount.

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?

As we have seen in the above questions , the patients visiting the hospital for the General Practice are comparatively very higher than other department.

The timings can be decided on the basis of the rush. During the day , the number of patients will be more so we can assign multiple doctors during the day and comparatively less number of doctors during the night.

The availability of the doctors will always be beneficial for the patients. They don’t have to wait for longer to see the doctor.

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

The Power BI Gateway is an application that helps developers access data located within an on-premises network. The developers can choose from both cloud-based and on-premises locations when requesting their data.

According to [**Statista**](https://www.statista.com/statistics/590031/worldwide-business-analytics-software-vendor-market-share/) Power BI has a 7.1 per cent market share amongst all Business Intelligence and Data Analytics platforms. This share is constantly on the up and owing to the wide suite of features it offers.  Power BI Gateway is a powerful tool for integrating on-premises data with Power BI. Learn how to use Power BI Gateway to securely connect to data sources.

The Power BI Gateway tool is a software application that helps users access the required data residing in an on-premises network. The tool acts as a gatekeeper for the source of data, and any requests made by users to access the data from a cloud or web-based application go through this gateway. Users are granted access depending on their authentication and data needs.

It's important to note that the tool does not transfer any data from the on-location source to the user's platforms. Instead, the tool connects the data source to the platform directly, and the user or client can then access their data for creating reports and analysing data. The Power BI Gateway tool is also designed to establish connections between the data source(s) and the source on-location

Use cases :

* Data Visualization
* Business Reporting
* Data Exploration
* Real-time analytics
* Self-Service BI
* Data Consolidation
* Predictive , Financial, Operational , Human Resources, Health Case and various other Analytics
* Data Collaboration