

Prescribing Growth: A Data-Driven Blueprint for Columbia Asia's Future

*-By Mohammed Hussain Mansuri*

**[OBJECTIVE QUESTIONS]**

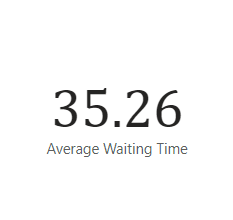
**Q1. In analyzing the hospital dataset with Power BI, ensure data cleaning to address inconsistencies and missing values before further analysis.**

* Before loading the dataset in PowerBI we can use **Power Query Editor** to transform the dataset.
* In the **Hospital ER** dataset, we could see there were around 72% null values in the “**patient\_sat\_score**” column.
* The best way to accurately clean data of null and blank values is to go back to the source of the data to get more information or get a better dataset, additionally we can improve data collection practice also.
* Since, we cannot do that here, the ideal way to fill the null values of numerical dataset such as “**patient\_sat\_score**” is to replace nulls with average value.
* In our case, the null values of the “**patient\_sat\_score**” column is replaced with the 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.**

* We can use a DAX Measure to find the average waiting time for patients

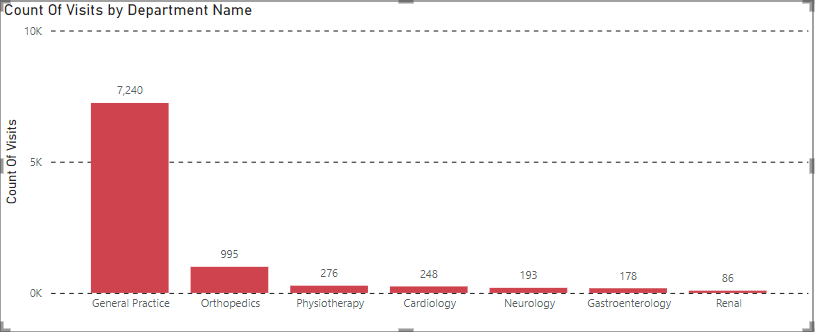
Average Waiting Time = AVERAGE('Hospital ER'[patient\_waittime])



* We can see from the above visual 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.**

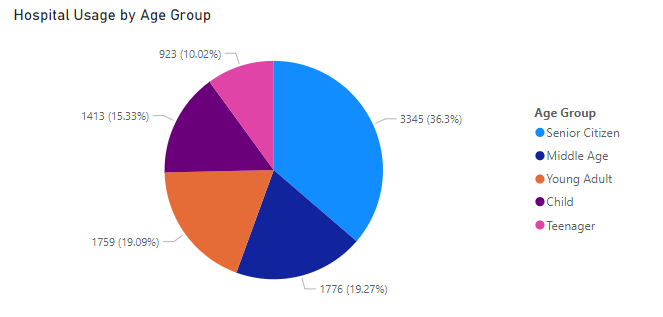
* We can use a bar chart in Power BI to calculate the number of visits per department like below:



* From the above chart, we can figure out that General Practice gets the most visits with around 7,240 visits.
* Followed by Orthopedics, Physiotherapy and Cardiology with 995, 276 and 248 visits.
* The lowest visited department in our hospital is Renal with 86 visits after Gastroenterology and Neurology with 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.**

* First we will create column “Age Group” which divides the ages of the patients into 5 categories.
* “Child” (Age<12), “Teenager”(12<Age<=20), “Young Adult”(20<Age<=35), “Middle Age”(36<Age<=50) and “Senior Citizen”(Age>50).
* We used the following DAX Formula to add a new column using the “Add new column” functionality.
* Age Group = SWITCH(TRUE(),'Hospital ER'[patient\_age]<=12,"Child", AND('Hospital ER'[patient\_age] >12,'Hospital ER'[patient\_age]<=20),"Teenager",AND('Hospital ER'[patient\_age] > 20, 'Hospital ER'[patient\_age]<=35),"Young Adult",AND('Hospital ER'[patient\_age]>35,'Hospital ER'[patient\_age]<=50),"Middle Age",'Hospital ER'[patient\_age]>50,"Senior Citizen")

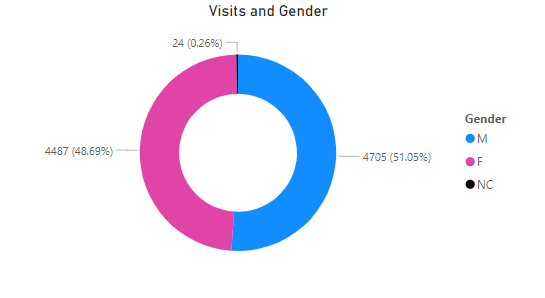


* We can observe in the above pie diagram that Senior Citizens utilize this hospital the most with around 3,345 patients or 36.3% of the total patient in the hospital.
* This is followed closely by Middle Age and Young Adult group with 1,776 people or 19.27% and 1,759 or 19.09% of peoples.
* The least 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?**

* Yes, the “patient\_sat\_score” column had 72% blank values which would significantly affect our data analysis.
* The best way to solve null values is to go back to the source and get the data again.
* Also, once can improve the data collection process to minimize or eliminate null values problem all together.
* The approach I took was first converted blank column to “null”.
* After making all values as null I replaced those values with column’s average value.
* So I replaced missing “patient\_sat\_score” with the average of “patient\_sat\_score”

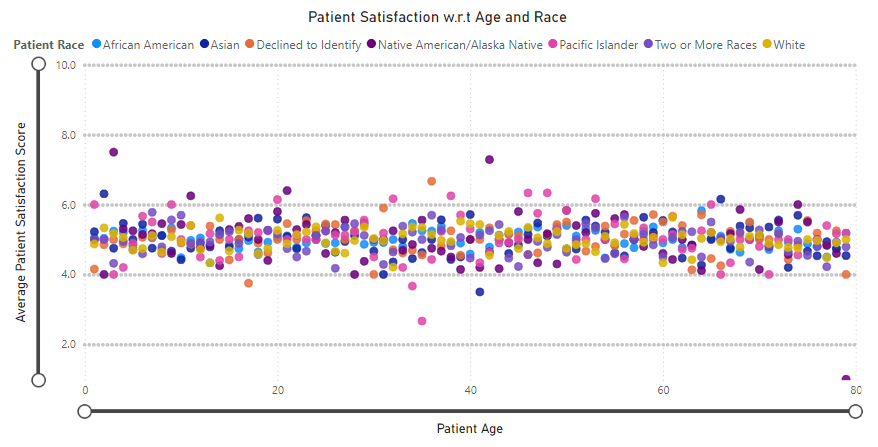
**Q6. Is there any relation between the number of visits and the Gender of the patients?**



* We can see from the above donut chart that Male and Female patients have equal distribution in the number of visits.
* Male patients visit are at 4705 or 51.05%, meanwhile female patient visits amounts to 4487 visits which is 48.69% of total visits.
* There are a few patients who either refuse to disclose their gender or do not confirm to any of the 2 gender categories which are marked as “Non-Confirmed (NC)” and have a share of 0.26% or 26 visits.

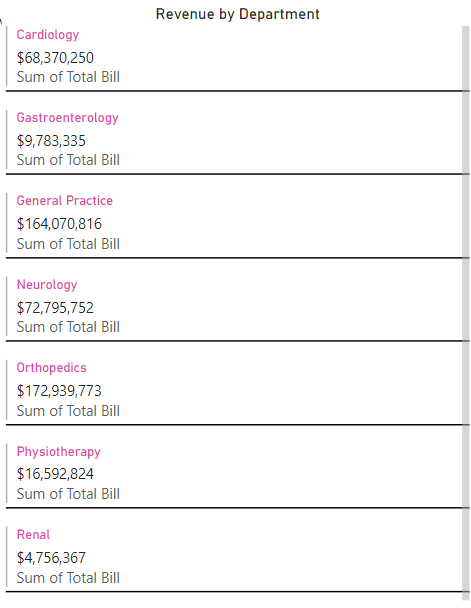
**Q7. Average Satisfaction by Demographics: Determine the relationship between patient satisfaction scores, their age groups, and racial backgrounds to pinpoint areas for improvement in patient experience.**

* We can use a scatter plot to observe a relation between patient satisfaction score, their age and racial background.

****

* We can observe that patients satisfaction score remains almost constantly in the range of 4.0 to 6.0 through out all the age groups
* We can also see that patients with different racial backgrounds all fall in this range irrespective of their age.
* We can conclude that there is no change in satisfaction score with the change of age and racial background.

**Q8. The hospital's managing director seeks to evaluate the revenue of each department to understand how much revenue is generated by each.**



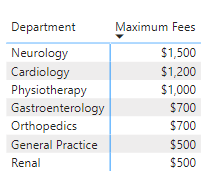
* As we can observe from the above multi-row cards visual, the contribution of different departments in hospital’s revenue.
* We can see that Cardiology nets us revenue with a sum of $68,370,250.
* Gastroenterology and General Practice comes in at with $9,783,335 and $164,070,816 respectively.
* Renal and Physiotherapy has generated $4,756,367 and $16,592,824 respectively, meanwhile Orthopedics and Neurology have a revenue of $172,939,773 and $72,795,752.

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

* We can use the following aggregation DAX function to calculate the highest appointment fees.

Max\_Department\_Fees = MAX('Doctor Patient Data'[Appointment Fees])

* We can further use a matrix visual and add departments to see maximum appointment fees for each department.



* We can observe from the above matrix that the maximum appointment fees charged is $1,500 by the Neurology Department.
* Cardiology comes in at 2nd place with $1,200 followed by Physiotherapy at $1,000.
* Renal and General Practice departments have the lowest fees $500 just behind Gastroenterology and Orthopedics with $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.**

* 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

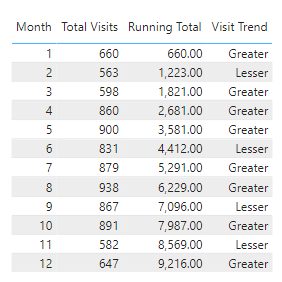
Running Total = RUNNINGSUM([Total Visits])

* 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 in to create our Visit Trends column in the below DAX query.

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

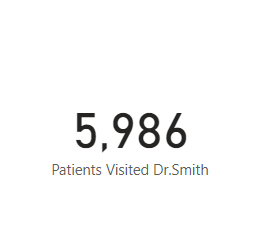


**Q11. Using ‘Calculate’ and a row iteration DAX function calculate the total number of patients who have visited Dr. Smith**

* We can use the COUNTDISTINCT() DAX function inside CALCULATE() and give filter as *“Doctor Name = “Dr. Smith”’*

Patients Visited Dr.Smith = CALCULATE(DISTINCTCOUNT('Doctor Patient Data'[patient\_id]),'Doctor Patient Data'[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.

**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?**

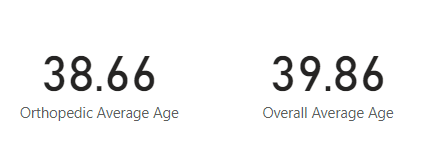
* For calculating average age of patients who visit Orthopedics Department we make use of the CALCULATE() DAX function along with AVERAGE() and filter on Department.

Orthopedic Average Age = CALCULATE(AVERAGE('Hospital ER'[patient\_age]),'Hospital ER'[department\_referral]="Orthopedics")

* 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.

Overall Average Age = AVERAGE('Hospital ER'[patient\_age])

* Below we 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?**

* Yes, there was data format issues for the date column in the “Hospital ER” table as well in the “Appointment Fees” and the “Total Bill” Column in the Doctor Patient Data table.
* For the data column we just clicked on the column and change its “Data Type” from the “Structures tab” in “Table view”.
* Same was done for “Appointment Fess” and “Total Bill” only this time we changed them to Currency data type from whole numbers.
* We are assuming since Columbia is in the US, the dataset has monetary amount mentioned in USD, so we converted both the columns in 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?**

* The M-query language is used to manipulate the tables and perform transformation on our dataset.
* Whenever we add/remove/manipulate columns in our dataset/table we can see the corresponding M-query in the formula bar.
* Power Query UI generates M-query whenever we perform any actions on our data.
* The following M-query code appears on the formula bar when you add a column in Power Query UI.

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

* In the above M-query we are merging the “patient\_first\_initial” and the “patients\_last\_name” column to create a new column called “Full Name” using the “Column from Examples” feature of Power Query UI.

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

* We can use the following SQL query to answer the above.

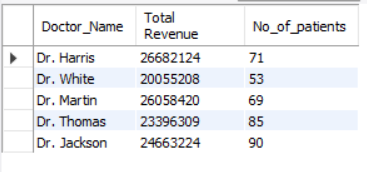
**SELECT Doctor\_Name, SUM(Total\_Bill) AS 'Total Revenue',COUNT(Patient\_ID) AS 'No\_of\_patients'**

**FROM doctor\_patients\_data**

**GROUP BY Doctor\_Name**

**ORDER BY Total\_Bill DESC, No\_of\_patients ASC**

**LIMIT 5;**



* We can see that Dr. Harris had the most revenue with the least patients.
* Followed by Dr. White, Dr. Martin and Dr. Thomas.
* Dr. Jackson has the lease amount of revenue with the most amount of person

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

* We can use the below query to identify the same

**WITH AvgWaitTimeByMonth AS (**

**SELECT `department\_referral`, CONCAT(Year, '-', Month) AS YearMonth,**

**ROUND(AVG(patient\_waittime),2) AS AvgWaitTime**

**FROM**

**hospital\_er\_data**

**GROUP BY `department\_referral`, Year, Month**

**),**

**WaitTimeWithLag AS (**

**SELECT `department\_referral`,YearMonth, AvgWaitTime,**

**LAG(AvgWaitTime, 1) OVER (PARTITION BY department\_referral ORDER BY YearMonth) AS PrevMonthAvg,**

**LAG(AvgWaitTime, 2) OVER (PARTITION BY department\_referral ORDER BY YearMonth) AS TwoMonthsAgoAvg**

**FROM**

**AvgWaitTimeByMonth**

**)**

**SELECT `department\_referral`,YearMonth,AvgWaitTime,PrevMonthAvg,TwoMonthsAgoAvg**

**FROM**

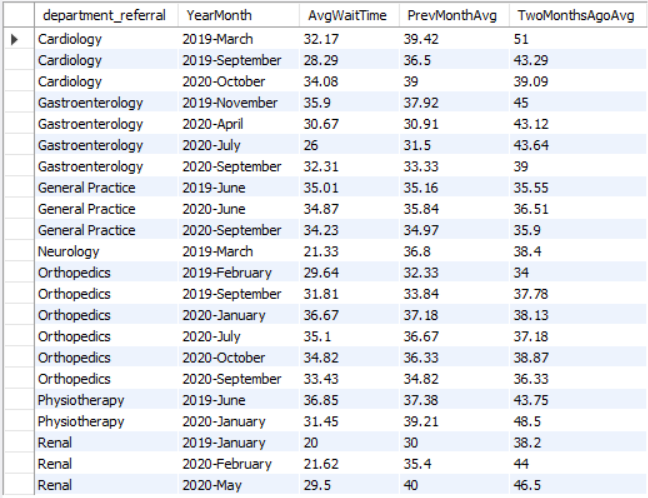
**WaitTimeWithLag**

**WHERE**

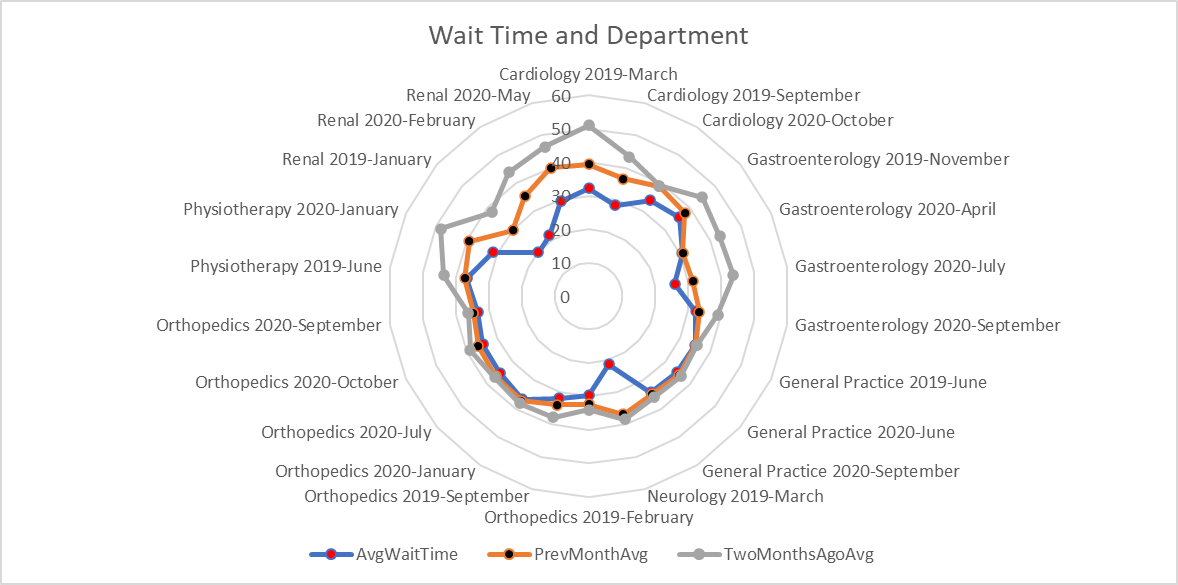
**AvgWaitTime < PrevMonthAvg**

**AND PrevMonthAvg < TwoMonthsAgoAvg**

**ORDER BY department\_referral, YearMonth ASC;**



* We can further visualize this using bar chart as below:



* We can observer that in all the departments the average wait time has been reducing significantly.
* Renal department has improved the most with 2 months ago the average being 38.2 minutes but now reduced to just 20 minutes, which equates to ~18 minutes of reduction in average wait time.
* Neurology follows closely with about 38.4 hours of average wait time 2 months go to 21.33 hours of average wait time, a drastic reduction of ~16 minutes of average wait time.
* Orthopedics department performs the most poor when it comes to reduction in average wait times with wait times being almost the same throughout.

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

**WITH male\_female\_count AS (**

**SELECT dp.Doctor\_Name, SUM(CASE WHEN patient\_gender = "F" THEN 1 ELSE 0 END) AS Female\_count,**

**SUM(CASE WHEN patient\_gender = "M" THEN 1 ELSE 0 END) AS Male\_count**

**FROM doctor\_patients\_data dp**

**JOIN hospital\_er\_data he**

**ON dp.patient\_id = he.patient\_id**

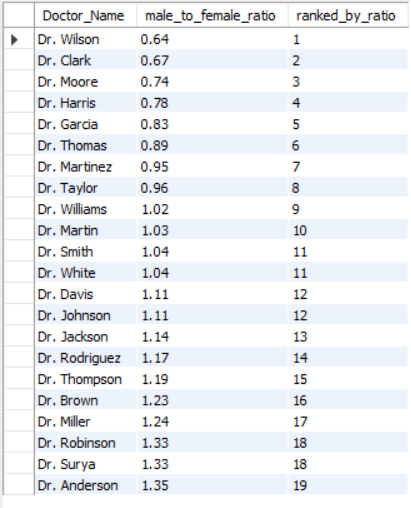
**GROUP BY dp.Doctor\_Name),**

**ratioed\_table AS**

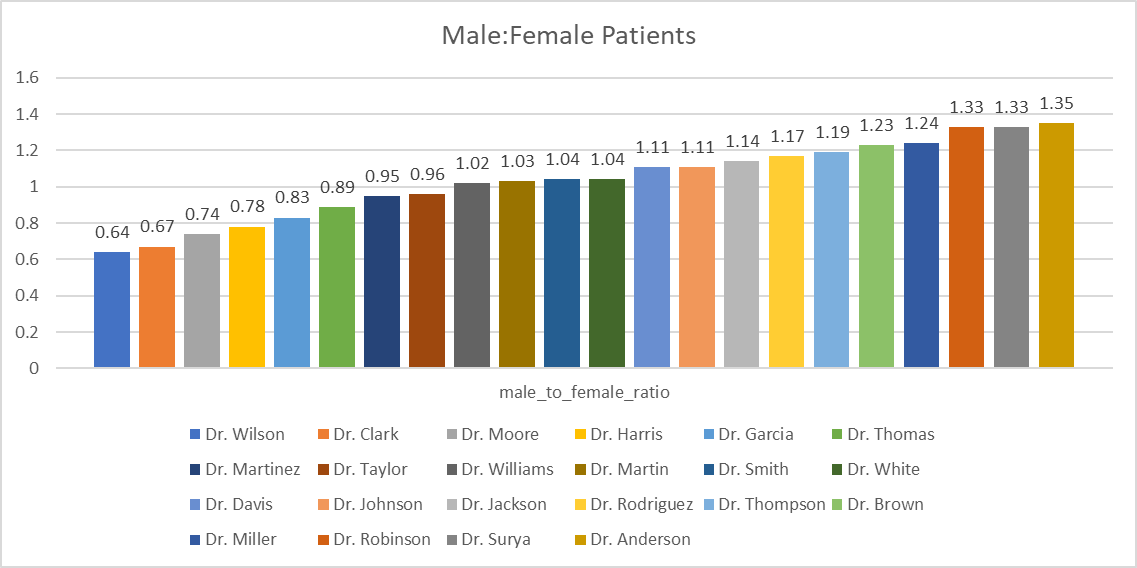
**( SELECT Doctor\_Name, ROUND((Male\_count/Female\_count),2) AS male\_to\_female\_ratio**

**FROM male\_female\_count)**

**SELECT \*,DENSE\_RANK() OVER(ORDER BY male\_to\_female\_ratio) AS ranked\_by\_ratio FROM ratioed\_table;**



* We can further visualize this as below:



* We can observe that Dr. Wilson has the lowest male: female patients ratio with 0.64 male patient for each female patient. Followed by Dr. Clark with 0.67 and Dr. Moore by 0.74.
* Interestingly, Dr. Williams, Dr. Martin, Dr. Smith and Dr. White all have almost the same male: female ratio of 1.02,1.03,1.04 and 1.04 respectively.
* They also seem to have the most balanced ratio of about 1 male patient to each female patient.
* Dr. Robinson, Dr. Surya and Dr. Anderson all have the same ratio of 1.33, 1.33 and 1.35 respectively.
* They are on the other extreme end with ~1.5 male patients to each female patient.
* Overall, the distribution looks quite balanced.

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

* We can use the following query for the above need:

**SELECT dp.Doctor\_Name, ROUND(AVG(he.patient\_sat\_score),2) AS `Average Satisfaction Score`**

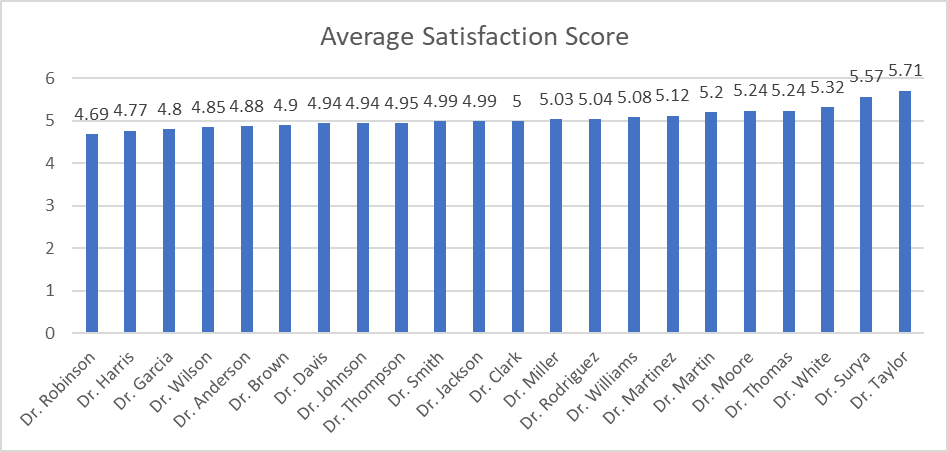
**FROM doctor\_patients\_data dp**

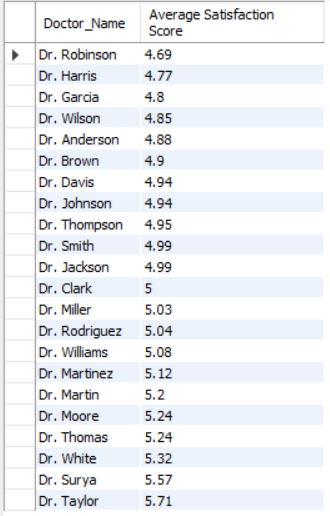
**JOIN hospital\_er\_data he**

**ON dp.patient\_id = he.patient\_id**

**GROUP BY 1**

**ORDER BY `Average Satisfaction Score**



****

* We can observe from the bar chart that Dr. Robinson has the lowest average satisfaction score with 4.69 out of 10 followed by Dr. Harris and Dr. Garcia with a score of 4.77 and 4.88 respectively.
* Dr. Smith, Dr. Jackson and Dr. Clark all have a very similar average of 4.99, 4.99 and 5 respectively. This puts them right in the middle of the scale.
* Dr. Taylor is our best rated doctor with a score of 5.71 after Dr. Surya and Dr. White with ratings of 5.57 and 5.32 respectively.

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

**SELECT dp.Doctor\_Name, ROUND((SUM(CASE WHEN LOWER(he.patient\_race) = "white" THEN 1 ELSE 0 END ))/COUNT(he.patient\_ID),2)\*100 AS percent\_Of\_white\_patients,**

**ROUND((SUM(CASE WHEN LOWER(he.patient\_race) = "african american" THEN 1 ELSE 0 END))/COUNT(he.patient\_ID),2)\*100 AS Percent\_of\_african\_american\_patients,**

**ROUND((SUM(CASE WHEN LOWER(he.patient\_race) = "asian" THEN 1 ELSE 0 END))/COUNT(he.patient\_ID),2)\*100 AS Percent\_of\_asian\_patients,**

**ROUND((SUM(CASE WHEN LOWER(he.patient\_race) = "native american/alaska native" THEN 1 ELSE 0 END))/COUNT(he.patient\_ID),2)\*100 AS Percent\_of\_native\_patients,**

**ROUND((SUM(CASE WHEN LOWER(he.patient\_race) = "pacific islander" THEN 1 ELSE 0 END))/COUNT(he.patient\_ID),2)\*100 AS Percent\_of\_pacific\_islandar\_patients,**

**ROUND((SUM(CASE WHEN LOWER(he.patient\_race) = "two or more races" THEN 1 ELSE 0 END))/COUNT(he.patient\_ID),2)\*100 AS Percent\_of\_mixed\_race,**

**ROUND((SUM(CASE WHEN LOWER(he.patient\_race) = "declined to identify" THEN 1 ELSE 0 END))/COUNT(he.patient\_ID),2)\*100 AS Percent\_of\_declined\_to\_identify,**

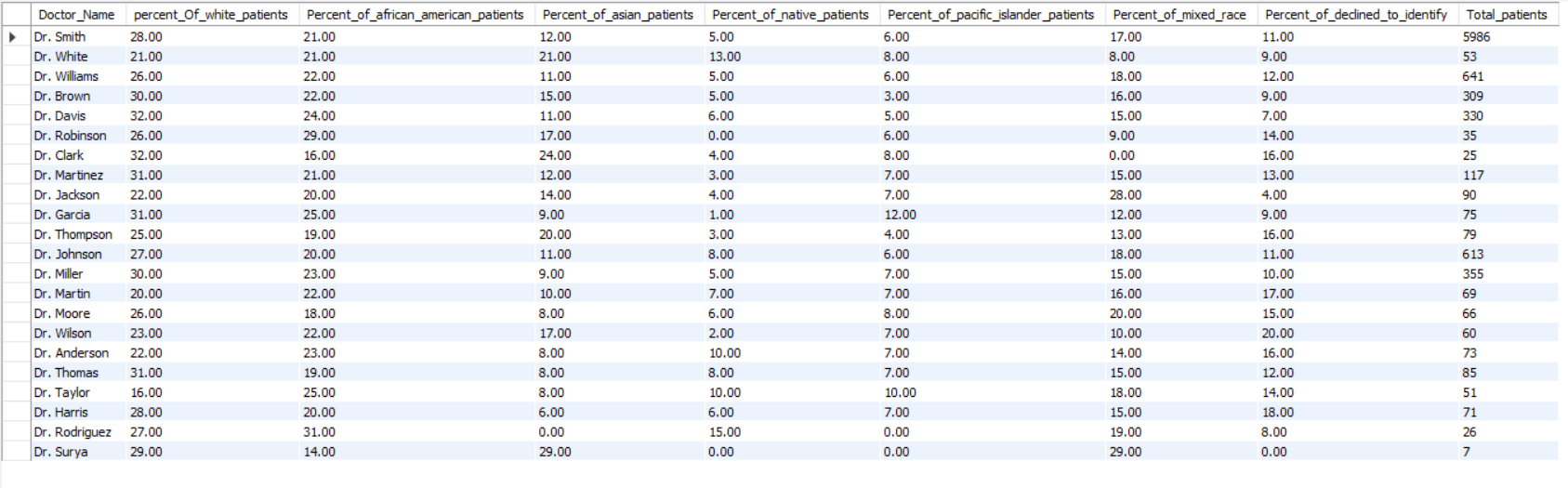
**COUNT(he.patient\_ID) AS Total\_patients**

**FROM doctor\_patients\_data dp**

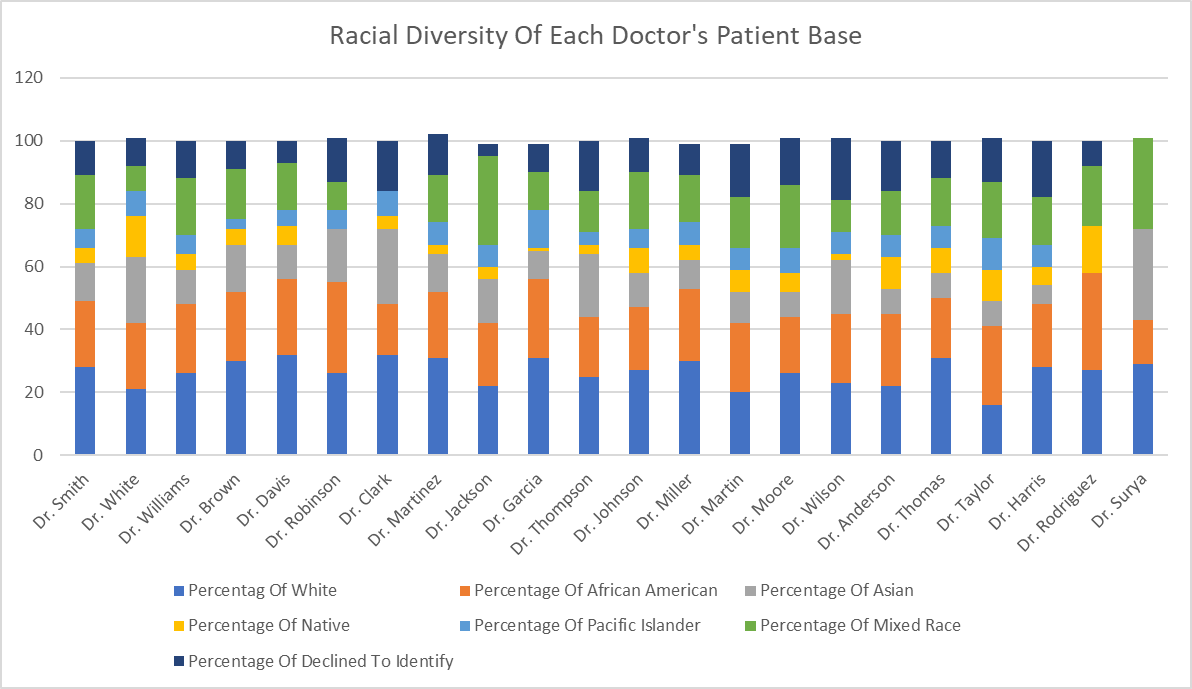
**JOIN hospital\_er\_data he**

**ON dp.patient\_id = he.patient\_id**

**GROUP BY dp.Doctor\_ID;**



* We can visualize the above output as follows:



* We can observe from the above visual that Dr. Surya has the least racially diverse base of patients out of all doctors
* Dr. Surya does not have any Native American/Alaskan Native and Pacific Islander patients, but total patient count is also low for them with just 7 patients.
* Overall, we can observer that all the doctors have good racial diversity in their patient base.

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

* We can use the below query to find the information requested above.

**WITH male\_female\_count AS (**

**SELECT dp.department\_referral AS Department\_Name, SUM(CASE WHEN patient\_gender = "F" THEN dp.Total\_Bill ELSE 0 END) AS Female\_Bill\_Total,**

**SUM(CASE WHEN patient\_gender = "M" THEN dp.Total\_Bill ELSE 0 END) AS Male\_Bill\_Total**

**FROM doctor\_patients\_data dp**

**JOIN hospital\_er\_data he**

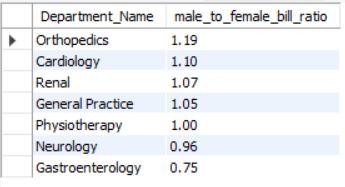
**ON dp.patient\_id = he.patient\_id**

**GROUP BY dp.department\_referral)**

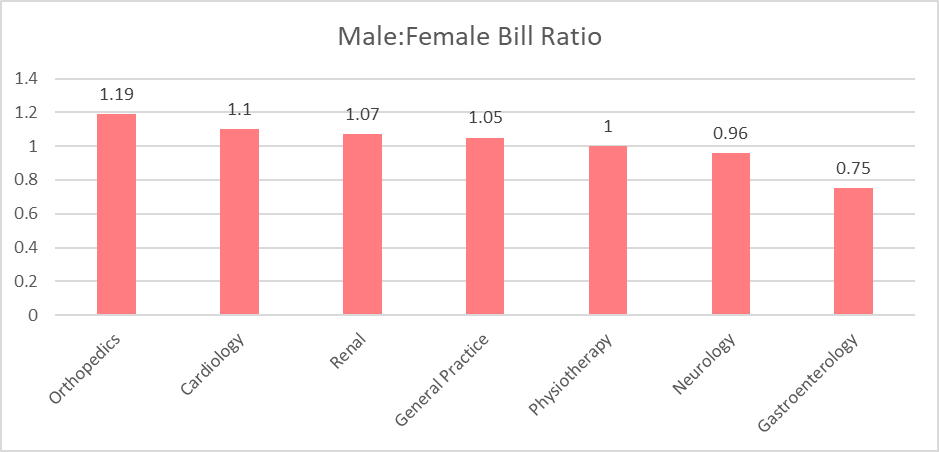
**SELECT Department\_Name, ROUND((Male\_Bill\_Total/Female\_Bill\_Total),2) AS male\_to\_female\_bill\_ratio**

**FROM male\_female\_count**

**ORDER BY male\_to\_female\_bill\_ratio DESC;**

****

* Lets visualize the above table.



* We can observe that Orthopedics Department has the most number of male bills to female bills with 1.19 bill generated by male patients for each bill generated by female patient.
* Followed by Cardiology with 1.1 bills generated by male patients for each female generated bill
* Physiotherapy department has the most balanced ratio of 1:1 for each female bill there is exactly 1 male bill.
* Neurology and Gastroenterology is slightly dominated by bill generated by females with 0.96 and 0.75 bills generated by males for each bill generated by female respectively.

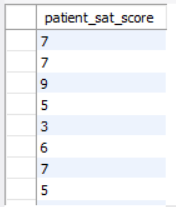
**Q21. Update the patient satisfaction score for all patients who visited the "General Practice" department and had a waiting time of more than 30 minutes. Increase their satisfaction score by 2 points, but ensure that the satisfaction score does not exceed 10. (SQL)**

* We can use the below query to update the values in the “patient\_sat\_score” column in the “hospital\_er\_data” table.

**UPDATE hospital\_er\_data**

**SET patient\_sat\_score = LEAST(patient\_sat\_score + 2,10)**

**WHERE LOWER(department\_referral) = "general practice" AND patient\_waittime > 30;**

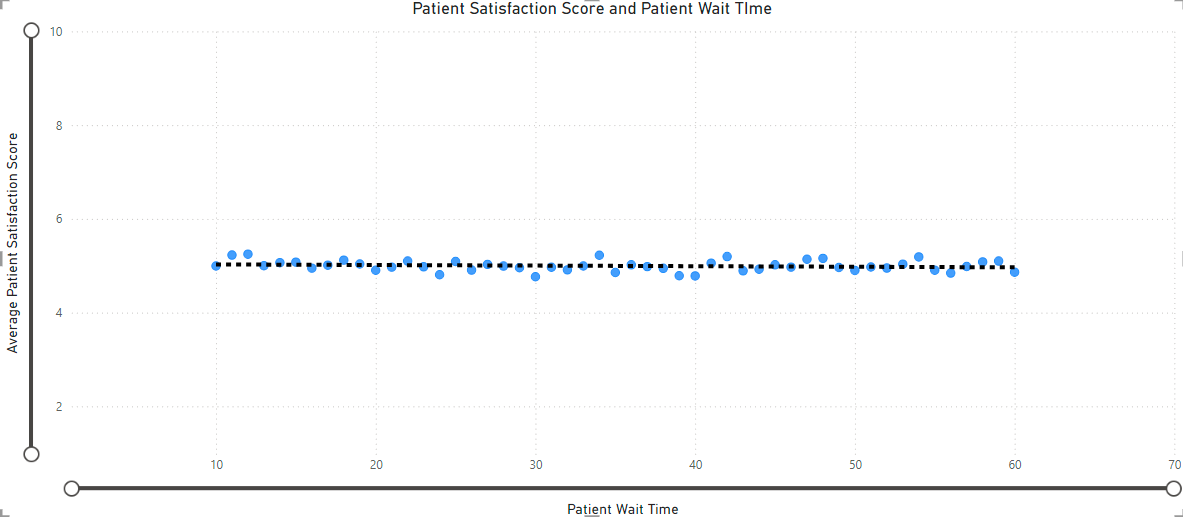


* This is just a snippet of data updated in the hospital\_er\_data table.

**[SUBJECTIVE QUESTIONS]**

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

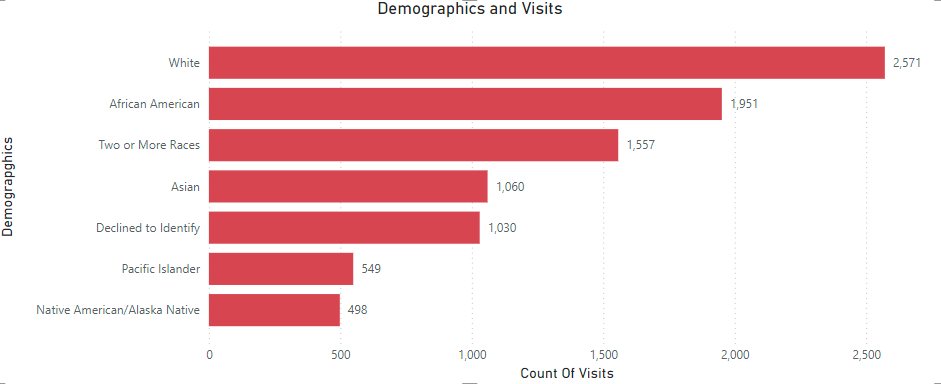
* We can use a scatter plot in Power BI to observer if there is any relation between Wait Time and Satisfaction Score.



* We can observe from the above scatter plot that there is a negative correlation between patient satisfaction score and patient wait time.
* The more the wait time is increase the lower the satisfaction score gets.
* Even though there are a few outliers in the dataset, but we can still conclusively say that as the wait time increases, satisfaction score decreases.

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

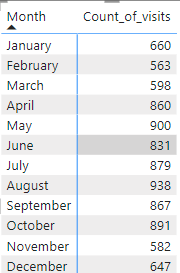
* Let’s visualize this using Horizontal bar Chart.



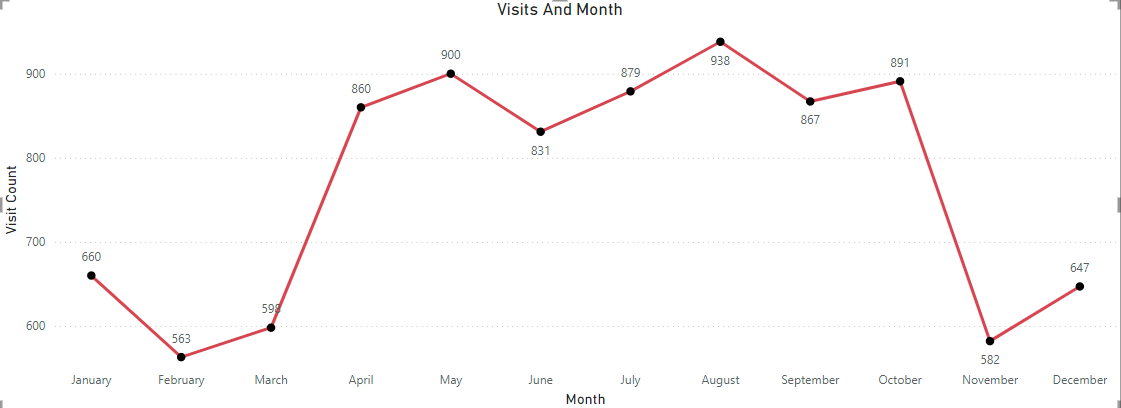
* We can see that the people identifying themselves as White visits the hospital most with 2,571 visits
* Followed by African American patients with 1,951 visits
* Interestingly, 1,030 patients refused to identify their race meanwhile 1,557 were of mixed heritage.
* Native American/Alaskan Native have the lowest visit count of 498 visits just after Pacific Islander having 549 visits.
* This could be due to relatively low population of people identifying as Native America/Alaskan Native or Pacific Islander demographic.

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

* Let us first get the monthly visit count.

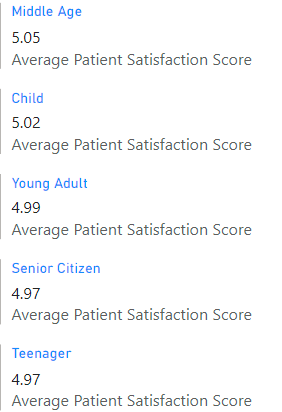
****

* Now, we can see that August was the month with the most visit and December was the month with the least.
* Lets, use a line chart to better visualize this for more clarity.

****

* We can observer that there is a high volume of patients all through April to October and then a sharp decline in November.
* Start of the year (January – March) and end of year (November – December) has low volume of patients.

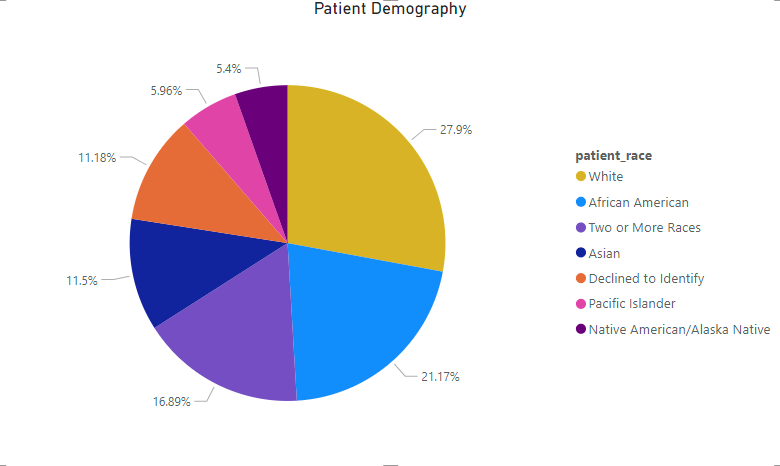
**Q4.Which age groups report the highest and lowest satisfaction scores?**



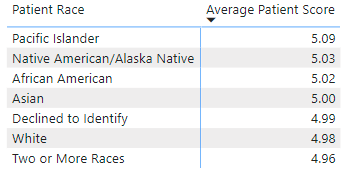
* We can see with the help of the multi-card visual that Middle age is the age group with the best average satisfaction score of 5.05.
* Senior Citizens and Teenagers are in tie for last position with average score of 4.97.
* Child age group was very close to getting 1st place with a satisfaction score of 5.02.
* Followed by Young Adult with a satisfaction score of 4.97.

**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?**

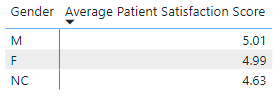
* We can deny this claim multiple ways through the visualization of our patients data.
* First, lets check the patient visits based on their demographics.



* We can clearly see that white patients only account for around 27.9% of our total patient base.
* The other ~70% of our patient base is non-white.
* We have patients from various walks of life such as African American (21.7%), Mixed Race (16.89%), Asian (11.5%), Pacific Islander (5.6%) and Native American/ Alaska Native with 5.4%.
* The remaining patients who declined to identify their race are around 11.18% which still if we assume all are White, still leave a significant gap between white patient base and others.
* Apart from that, we can also check the average patient satisfaction score given different demographics to again falsify these claims.



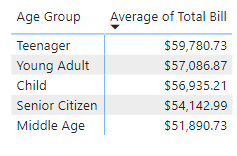
* As we can observer from the above matrix that Pacific Islander highly rate the hospital with an average satisfaction score of 5.09.
* Followed by Native America/Alaskan Native with a score of 5.03 and African American with the score of 5.02.
* Asians, Whites and Mixed race people rate the hospital in a very similar way with 5.0, 4.98 and 4.96 average satisfaction scores respectively.
* And, people who refused to disclose their race have an average patient satisfaction score of 4.99.
* All of the above metrics can put any claims of the hospital being racially discriminating false.
* We can even check for gender based metrics too.



* We can clearly see that there is not much disparity between each gender’s score.

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

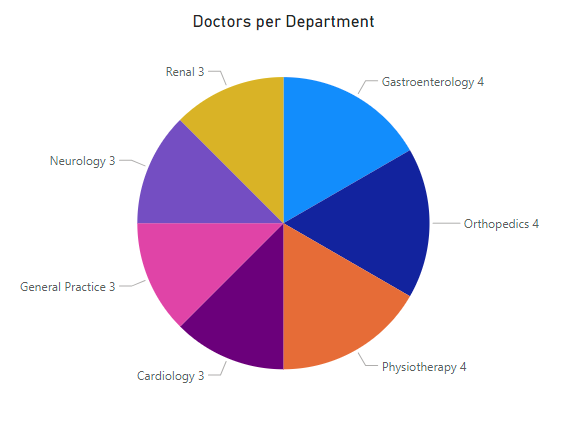
* The approach I would take when it comes to providing discounts is age group.
* We can provide discounts to age groups who spent the most at the hospital.



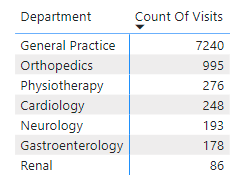
* We can see that Teenagers (Aged between 12 to 20) are the ones who spent the highest on average in the hospital.
* Amounting to over $59,780.73 on average they are hospital’s biggest customer.
* Secondly, if we attract more patients from the teenager age group, we might be able to create relationships with them and make a life long customers out of them.
* For all the above reasons I think the hospital should give patients discounts on the basis of their age and more specifically to the teenagers.

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

* We can look at the current departments and check which departments have more or less doctors.



* Most of the hospital’s department have 3 or 4 doctors, we can further analyze the number of visits each department get.

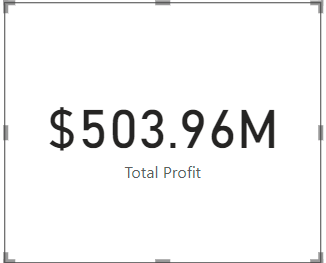


* We can see that General Practices gets the highest amount of visits with 7,240.
* Even Orthopedics which comes in at 2nd place with 995 visits is no where near that.
* My suggestion would be to hire all 3 new doctors to the General Practice department as it has only 3 doctors currently.

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

* We can use the following DAX function to determine profitability.

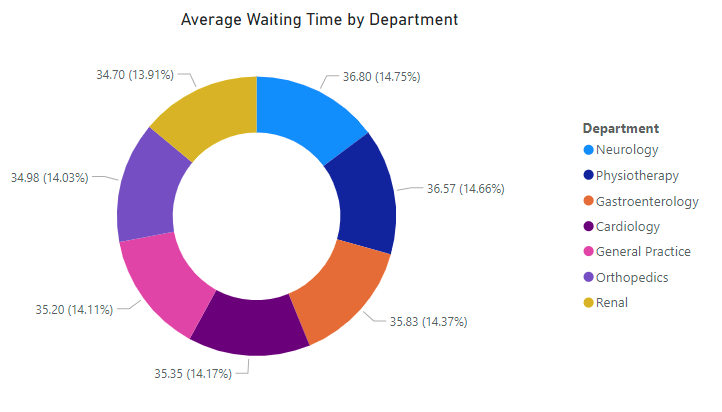
Profit Calculation = SUM('Doctor Patient Data'[Total Bill]) - SUM('Doctor Patient Data'[Appointment Fees])



* After our calculations we can see the hospital is in profit of about $503.96 Million.
* We calculated profit by subtracting the Doctor’s consultation charge from the total bill paid by the patients.

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

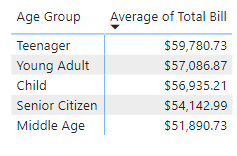
* Let us visualize the average waiting time per department by using a Donut chart.



* We can see that all departments almost have average waiting time in the same ball park.
* All the departments have the average wait time between 34.70 minutes to 35.83 minutes which is approximately a ~1.5% of difference between the maximum and minimum.
* This difference is very insignificant and hence we can conclude that there is no particular department with relatively very high wait time.

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

* The approach I would take when it comes to providing discounts is age group.
* We can provide discounts to age groups who spent the most at the hospital.



* We can see that Teenagers (Aged between 12 to 20) are the ones who spent the highest on average in the hospital.
* Amounting to over $59,780.73 on average they are hospital’s biggest customer.
* Secondly, if attract more patients from the teenager age group, we might be able to create relationships with them and make a life long customers out of them.
* For all the above reasons I think the hospital should give patients discounts on the basis of their age and more specifically to the teenagers.

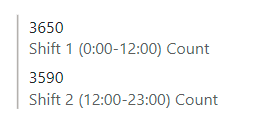
**Q11. Say you need to align the doctors of the “General Practice” department to work in one of the two shifts, how will you identify what will these two shifts' timings be, and how will you divide the doctors in these two shifts? And also will this 2 shift policy be helpful for the hospital?**

* Let us divide the first day in 2 shifts and calculate the count of visits using the below 2 DAX functions.

Shift 1 Visit Count = CALCULATE(COUNT('Doctor Patient Data'[patient\_id]),'Hospital ER'[Time Of Day]>=0 && 'Hospital ER'[Time Of Day]<12 && 'Hospital ER'[department\_referral] = "General Practice")

Shift 2 Visit Count = CALCULATE(COUNT('Doctor Patient Data'[patient\_id]),'Hospital ER'[Time Of Day]>=12 && 'Hospital ER'[Time Of Day]<=23 && 'Hospital ER'[department\_referral] = "General Practice")

* We can visualize this using the multi-row card for more clarity.



* We can see that Shift 2 has more patients then shift 1 therefore while assigning shifts to doctors we should assign more doctors in shift 2 then in shift 1.
* Since we have 3 doctors currently in General Practice Department, we will assign 2 to Shift 2 and 1 to Shift 1.
* The main benefit of dividing duty of the doctors in shifts is that the workload get distributed evenly.
* Which in turn leads to lesser wait time for patients and better satisfaction scores.

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

* Power BI gateway is a tool to access data present within an on-premises (Privat Cloud) network from anywhere
* When we need to access our data from on-premises databases, servers ,storage devices, etc without actually moving our data to the cloud we can use Power BI gateway.
* Power BI gateway acts like a bridge between the data source and the cloud service where the developer wants to access the data.
* Some benefits of Power BI gateway are:
  + The connection between the data source and cloud is very secure and reliable which help in protection of sensitive data.
  + Data can be accessed directly without moving it to expensive cloud storage solution hence saving money.
  + Dashboards can be refreshed using Power BI gateway with a connection to the on-premise data source.

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

* If the objective and subjective question were not given to me I would have approached the problem according to the goals.
* For maximizing revenue generation I would have asked for more data like hourly rate for a different categories of hospital rooms, the cost of running all the machines in those rooms and medicine cost along with doctor’s consultation fees.
* To gain insights for hiring new employees, I would not just focus on doctors but also the nursing staff and the support staff.
* Having a bigger picture for each department will help in sharing workload and hiring/firing employees according to the hospital’s requirements.
* For patients discounts, my strategy would remain the same, I would still give discounts to the age group which gives us the most business.
* Along with the immediate benefit of increased cash flow, catering to a specific age group could pave way for a specialty department which in turn would attract more of the same high revenue generating group.

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

* The relationship between “Doctor Patient Data” table and “Hospital ER” table is 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 “Hospital ER” table there is only entry in the “Doctor Patient Data” table.
* There is only one relationship established between these tables and there it is active by default.
* The cross filter between these tables is bi-directional.

