**NAME: KIRAN H D  
PROJECT: COLUMBIA ASIA HOSPITAL PROJECT**

**OBJECTIVE QUESTIONS**

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

**Ans: Power BI Data Cleaning Steps:**

1. **Load Data:**  
   I started by loading both datasets *Doctor\_Patients\_Data* and *Hospital\_ER* into Power BI. To do this, I selected "Get Data" from the "Home" tab, chose the file format, and imported the datasets to begin my analysis.
2. **Open Power Query:**  
   I used Power Query Editor by selecting "Transform Data" under the "Home" tab in Power BI. Power Query is where I performed all the necessary data cleaning steps.
3. **Handle Missing Values:**  
   I carefully inspected each column for missing values. For example, in the *patient\_sat\_score* column, any missing values were filled with 5 (mean) to avoid skewing results with nulls. If any Doctor Name was missing, I replaced it with "Unknown" to keep records complete. I used the "Replace Values" option to make these substitutions easily.
4. **Standardize Text Fields:**  
   To ensure consistency, I standardized the text fields, such as *patient\_race* and *department\_referral*. In Power Query, I selected these columns and went to "Transform" > "Format" > "Capitalize Each Word." This formatting step helped maintain uniformity across the text fields, making the data more visually consistent in reports.
5. **Ensure Uniform Date Format:**  
   I checked the *date* column and set its data type to “Date/Time” to ensure consistency in all date entries. This step prevented issues that could arise from varied date formats.
6. **Apply and Load:**  
   After completing these cleaning steps, I clicked "Close & Apply" to save and load the cleaned data back into Power BI, making it ready for further analysis and visualizations on the report tabs.
7. **Assess the Average Waiting Time:** Analyse the patient wait times to identify the average duration a patient spends before receiving care.

**Ans:** Average Waiting Time Analysis:

* **Formula Used:** DAX

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

* **Result:** The average patient waiting time is 35.26 minutes.



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.

**Ans:**

* **Formula Used:** DAX

Count of patients = COUNT ('Hospital ER'[patient\_id])

* **Result:** Identified the total number of visitors per department based on referrals.
* **Visualization:** Created a chart with

**X-**axis: Count of Patients (measure).

**Y-**axis: Department Referral

A screenshot of a medical survey

AI-generated content may be incorrect.

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

**Ans:**

* **Approach:** To better understand patient demographics and how different age groups utilize healthcare services
* **Formula Used:** Created a new column using SWITCH to categorize patients by age group:

Patient\_age\_bucket = SWITCH (TRUE (),

'Hospital ER'[patient\_age] < 18, "0-18",

'Hospital ER'[patient\_age] >= 18 && 'Hospital ER'[patient\_age] < 35, "18-34",

'Hospital ER'[patient\_age] >= 35 && 'Hospital ER'[patient\_age] < 50, "35-49",

'Hospital ER'[patient\_age] >= 50 && 'Hospital ER'[patient\_age] < 65, "50-64",

'Hospital ER'[patient\_age] >= 65, "65+",

"Unknown")

* **Visualization:** Added the new Patient\_age\_bucket column to the chart's X-axis and used a measure to count the number of visitors for each age group.
* **Result:** You successfully created age buckets and visualized the number of visits for each demographic group.

A graph of patients by age group

AI-generated content may be incorrect.

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?

**Ans:** Null values were present in the patient satisfaction scores, which were subsequently filled with the average satisfaction score to ensure data completeness.

* **Identified Null Values:** Inspected the patient\_sat\_score column in Power Query to locate rows with missing values.
* **Calculated the Average:** Used Transform > Statistics > Average to calculate the column's average (4.9978298611111107)
* **Replaced Nulls with the Average:** Rounded the average to 5 and replaced all null values using Transform > Replace Values.

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

**Ans:** Analysis of Visits by Gender

* A breakdown of patient visits reveals the following gender distribution:
* Male: 51.05%
* Female: 48.69%
* Not Mentioned: 0.26%
* **Visualization:** This indicates that the number of visits is almost evenly distributed between males and females, with a slightly higher proportion of male patients. The Not Mentioned category is negligible at 0.26%.
* **Result:** The donut chart visually represents this distribution, highlighting the near parity in visits across genders.

A blue circle with white text

AI-generated content may be incorrect.

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.

**Ans:**

For Analysing the relationship between patient satisfaction scores, age groups, and racial backgrounds, I used a **Clustered Column Chart** Titled **Average satisfaction by Demography** on the **Main Tab** of my Power BI report. The chart visualizes the average satisfaction scores for different age groups and racial backgrounds, using a measure I created to calculate the average satisfaction score, **Avg\_Satisfaction**. Age group and patient race were placed in the axis field, while the average satisfaction score was shown in the values field. Data labels were enabled for easy interpretation of satisfaction scores for each demographic category.

**Insights**:

* **High and Low Satisfaction Groups:** The chart helps identify which age groups and racial backgrounds report higher or lower satisfaction scores. For instance, if older age groups (e.g., 60+) report lower satisfaction, this may indicate areas for improvement in patient care for these groups.
* **Age-Specific Trends:** Younger age groups (18-29) might report higher satisfaction scores compared to older patients, suggesting that younger patients may be more satisfied with the services provided, while older patients might require more personalized care or attention.
* **Racial Satisfaction Gaps:** The analysis may reveal if certain racial or ethnic groups have consistently lower satisfaction scores, highlighting potential disparities that could be addressed with targeted patient engagement strategies or improved care practices.
* **Overall Satisfaction Trends:** The chart provides an overall view of satisfaction trends, helping to spot patterns across demographics. Identifying low satisfaction scores in certain age or racial groups could inform strategies for improving the overall patient experience.

A graph of numbers and colors

AI-generated content may be incorrect.

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

**Ans:**  
To evaluate the revenue generated by each department, I used Power BI to display the financial contributions from all departments in a clear and effective manner.

The **Total Revenue by Departments** bar chart on the **Main Tab** was used to visualize the revenue generated by each department. This chart categorizes the revenue by department, clearly showing the financial contribution from each area. The departments with the highest revenue are represented with larger bars, offering a quick and comparative view of their financial performance.

For this analysis, the **department\_referral** field was used to group the data, and the total revenue for each department was calculated by aggregating the **Total Bill** values for each department. The **Revenue** metric is then presented in the bar chart, showcasing the contribution of each department, such as **General Practice**, **Orthopedics**, and **Neurology**, among others.

**Insights:**

* **Revenue Distribution:** The **Total Revenue by Departments** chart visually demonstrates how revenue is distributed across departments, allowing the identification of both high-revenue and low-revenue departments.
* **Identifying High-Performing Departments:** Departments like **General Practice** and **Orthopedics** generate the highest revenue, showing areas of strength within the hospital.
* **Resource Allocation:** The chart helps identify which departments are the primary revenue drivers, informing decisions regarding resource allocation and supporting the most profitable areas.
* **Strategic Growth Opportunities:** For departments with lower revenue, this visualization points out areas that could benefit from enhanced services, marketing, or additional support to improve their financial performance.

A graph of revenue

AI-generated content may be incorrect.

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

Ans:

* Highest Appointment Fee by Department
* To identify which department is charging the highest appointment fees in general.
* I used the LOOKUPVALUE DAX function to find the department corresponding to the maximum appointment fee. The measure created is:

Highest Appointment fee Department = LOOKUPVALUE(Doctor\_Patients\_data[department\_referral], Doctor\_Patients\_data [Appointment Fees], MAX (Doctor\_Patients\_data [Appointment Fees]))

Highest Appointment Fees = MAX (Doctor\_Patients\_data [Appointment Fees])

A close-up of a white square

AI-generated content may be incorrect.

* Using this measure, I visualized the data by displaying the department charging the highest appointment fees in a card.
* **Insights:**
* The analysis revealed that the Neurology department charges the highest appointment fee among all departments.
* This insight highlights Neurology’s premium positioning, possibly due to specialized expertise, advanced equipment, or demand for its services.
* **Result:** This analysis helps management better understand pricing strategies across departments and evaluate their alignment with the hospital's revenue goals and market positioning.

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.

**Ans:** Month-wise Total Visits with Previous Month Comparison Table

* To create a tabular visualization displaying Month-wise Total Visits, Previous Month's Visits, and a comparison column (Visit Increased?)

1. **Create a Calendar Table**

* You created a Calendar Table using the CALENDAR function to cover the date range in the Hospital ER Table:
* Mark as Date Table: Mark this table as the official Date Table under Modelling > Mark as Date Table.

Calendar = CALENDAR (MIN ('Hospital ER'[date]), MAX ('Hospital ER'[date]))

* **Purpose:** Ensures the date table spans the entire range of dates in the hospital data.

1. **Define Year, Month, and Month Name Columns**

* To organize data, you added calculated columns in the Calendar Table:
* Year Column: Year = YEAR('Calendar'[Date])
* Month Column: Month = MONTH('Calendar'[Date])
* Month Name Column: Month Name = FORMAT('Calendar'[Date], "MMMM")
* These columns help structure and sort data chronologically.

1. **Establish Relationships**

* Create a One-to-Many relationship between the Hospital ER Table (Date Column) and the Calendar Table (Date Column).
* Set the cross-filter direction to Single.

1. **Create Measures**

You defined the following measures:

* Current Month Visits:

Month wise Total visits = COUNT ('Hospital ER'[patient\_id])

* Previous Month Visits:

Pre\_Month\_Total\_Visits = CALCULATE (COUNT ('Hospital ER'[patient\_id]), PREVIOUSMONTH('Calendar'[Date]))

* Visit\_Increased:

Visit\_Increased = IF(ISBLANK([Month\_Wise\_Total\_Visits]) || ISBLANK([Pre\_Month\_Total\_Visits]), BLANK (), IF([Month\_Wise\_Total\_Visits] > [Pre\_Month\_Total\_Visits], "Yes”, “No"))

1. **A screenshot of a data

   AI-generated content may be incorrect.Tabular Visualization:**

* **Result:** Add the following fields to your Table Visual:
* Year (from the Calendar Table)
* Month Name (from the Calendar Table)
* Month\_Wise\_Total\_Visits (measure)
* Pre\_Month\_Total\_Visits (measure)
* Visit\_Increased (measure)

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

**Ans:**

To determine the total number of patients who have visited Dr. Smith, I used a calculated measure in Power BI. This measure allowed me to filter the dataset specifically for visits to Dr. Smith and then count the number of these visits.

I created the measure named **Dr. Smith patients** in the **All Measures**

Table. Using CALCULATE with a row iteration function; I filtered the data to include only those rows where **"Doctor Name"** equals **"Dr. Smith"** By counting the rows that meet this condition, I arrived at the total number of patients who visited Dr. Smith, which is **5,986**.

This calculation provides valuable insight into Dr. Smith’s patient load, which can be useful for workload analysis and understanding patient distribution across doctors.

Dr. Smith patients = CALCULATE (COUNT ('Hospital ER'[patient\_id]), Doctor\_Patients\_data [Doctor Name] = "Dr. Smith”)

A close-up of numbers

AI-generated content may be incorrect.

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?

**Ans:**

* **Approach:**
* Orthopedics: Create a measure with a filter to calculate the average age for Orthopedics and display it in a Card Visual.
* All Departments: Create a general measure without filters, add it to a Table Visual with the department column to display averages for all departments.
* **Explanation:**
* To calculate the average age for Orthopedics, a measure is created using the following formula

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

* The AVERAGE function then computes the average age, and the result is displayed in a Card Visual, as it represents a single value.
* For all departments, a general measure is created without any filter:

Average age of patient = AVERAGE ('Hospital ER'[patient\_age])

* This measure calculates the overall average age for patients. Adding this measure to a Table Visual, along with the department\_referral column, groups the data by department and shows the average age for each.
* Finally, the measures are formatted as whole numbers using the data format options in Power BI to ensure the results align with the realistic representation of age.
* **Visualization:**

A screenshot of a computer

AI-generated content may be incorrect.

* **Result:**
* A Card Visual shows the average age for Orthopedics.
* A Table Visual displays the average age for each department.
* These distinct approaches ensure accurate and meaningful insights tailored to each requirement.

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

**Ans:** Steps Taken to Resolve Data Format Issues:

* **Combining Columns for Patient Name:**
* The patient\_first\_initial and patient\_last\_name columns were stored separately, making it inconvenient for analysis.
* Using Power Query, I merged these columns into a single patient\_name column, ensuring proper formatting for better usability in queries and reports.
* **Extracting Dates and Creating a Calendar Table:**
* The date column included both date and time, but only the date component was needed for analysis.
* I created a custom column to extract the date and generated a calendar table to enable advanced date-based filtering and trend analysis.
* **Handling Blank Spaces in Scores:**
* The patient\_sat\_scores column had blank spaces, which could disrupt calculations and analysis.
* Initially, the data type of patient\_sat\_scores was in text format, so I converted it to a number.
* After conversion, I replaced the blank spaces (null values) with the column's mean value, ensuring data consistency and avoiding skewed results during computations.
* **Result:** The dataset was transformed into a clean and consistent format, making it ready for reliable analysis and accurate 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?

**Ans:**

When adding a new column in Power Query, M language generates code in the formula bar to define the transformation. For example, if I add a custom column for a calculation, the code might appear as:

= Table.AddColumn(PreviousStep, "New Column", each [Column1] + [Column2])

This line adds a new column titled "New Column," based on a calculation involving values from existing columns.

**Understanding M-Query:**

* M is the powerful language behind Power Query, designed to support efficient data transformation and manipulation. Key aspects of M include:
* **Data Transformation:** M allows a wide range of transformations, like filtering, grouping, and aggregating data, making it versatile for various data preparation tasks.
* **Step-by-Step Process:** Power Query records each transformation as a step, which can be viewed in the Applied Steps pane, making it easy to track and manage changes.
* **Readable Syntax:** M uses a formula-like syntax that is straightforward, enabling clear and precise adjustments in data processing.

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

**Ans:**

**Analysis:** Identifying Top 5 Revenue-Generating Doctors with Fewest Patients

To identify the doctors who generated the highest revenue while serving the fewest patients, we used an SQL query to aggregate and filter the data effectively. The key steps were to group by doctor, count the unique patients, and sum their total revenue, with the results ordered by patient count (ascending) and total revenue (descending). This approach reveals the doctors who maximize revenue despite a smaller patient base.

**Explanation of Key Elements:**

* **Grouping**: The query groups data by `Doctor Name` to analyse each doctor’s patient count and total revenue.
* **Aggregation**: COUNT (DISTINCT patient\_id) counts patients per doctor, while SUM (`Total Bill`) calculates total revenue.
* **Sorting**: ORDER BY no\_of\_patients ASC, Revenue DESC ensures doctors with fewer patients but higher revenue appear first.
* **Limitation**: LIMIT 5 restricts results to the top 5 doctors fitting these criteria.

A computer code with text

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

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

**Ans:**

A white background with blue text

AI-generated content may be incorrect.The departments with consistently decreasing average waiting times over three consecutive months were identified. This indicates improved efficiency and better patient flow management, contributing to enhanced service quality.

A computer screen shot of a computer code

AI-generated content may be incorrect.

A white background with black and white clouds

AI-generated content may be incorrect.

* **Output:**  
  A screenshot of a computer

  AI-generated content may be incorrect.

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

**Ans:**

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

A computer code with colorful text

AI-generated content may be incorrect.

A close-up of a white background

AI-generated content may be incorrect.

* **Output:**

A screenshot of a table

AI-generated content may be incorrect.

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

**Ans:**

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

A computer code with text

AI-generated content may be incorrect.

* **Output:**

**A screenshot of a medical report

AI-generated content may be incorrect.**

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

**Ans:**

Doctors who have treated patients from multiple races are identified, and the diversity of their patient base is measured by counting the distinct races they have treated. This helps highlight doctors with a more diverse patient demographic.

A computer code with text

AI-generated content may be incorrect.

* **Output:**

**A screenshot of a computer

AI-generated content may be incorrect.**

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

**Ans:**

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

A screenshot of a computer program

AI-generated content may be incorrect.

* **Output:**

**A screenshot of a computer

AI-generated content may be incorrect.**

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)

**Ans:**

The satisfaction scores for patients in the "General Practice" department who had a waiting time greater than 30 minutes are updated by increasing their score by 2 points. If the new score exceeds 10, it is capped at 10. This ensures that patient satisfaction scores are adjusted fairly while maintaining the maximum allowable score.

A screenshot of a computer program

AI-generated content may be incorrect.

**SUBJECTIVE QUESTIONS**

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

**Ans:**

* **Approach:**

A quick measure was created in Power BI by assigning department\_referral as the category, patient\_sat\_score as Measure1, and patient\_waittime as Measure2.

The DAX formula was automatically generated to calculate the correlation between these two measures.

To visualize the correlation value, a Gauge Chart was used with a range set from 0 to 2, where the final correlation value observed was 1.

* **Explanation:**

Correlation analysis in this context helps determine the strength and direction of the relationship between patient wait times and satisfaction scores.

A correlation value of 1 represents a moderate positive relationship, meaning that as patient wait times increase, satisfaction scores tend to decrease, but not in a strong or direct manner.

This suggests that while wait times influence patient satisfaction, the relationship is not perfectly linear and may be influenced by other factors such as staff behaviour, quality of care, or facility conditions.

The Gauge Chart effectively illustrates this result, making it easier to interpret the correlation score and its significance within the range.

* **Visualization:**

A blue and white circle with black text

AI-generated content may be incorrect.

* **Insight:**

The analysis revealed a moderate positive correlation of 1, implying that patient wait times do have an impact on satisfaction scores, though not overwhelmingly.

Departments with shorter wait times or better management of patient expectations might achieve higher satisfaction, but other factors such as service quality and communication likely play a crucial role in shaping patient perceptions.

* **DAX Formula:**

Corelation\_waittime\_satscore =

VAR \_\_CORRELATION\_TABLE = VALUES('Doctor\_Patients\_data'[department\_referral])

VAR \_\_COUNT =

COUNTX (

KEEPFILTERS(\_\_CORRELATION\_TABLE),

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

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

))

VAR \_\_SUM\_X =

SUMX(KEEPFILTERS(\_\_CORRELATION\_TABLE),

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

VAR \_\_SUM\_Y =

SUMX(KEEPFILTERS(\_\_CORRELATION\_TABLE),

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

VAR \_\_SUM\_XY =

SUMX(KEEPFILTERS(\_\_CORRELATION\_TABLE),

CALCULATE (

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

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

VAR \_\_SUM\_X2 =

SUMX(KEEPFILTERS(\_\_CORRELATION\_TABLE),

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

VAR \_\_SUM\_Y2 =

SUMX(KEEPFILTERS(\_\_CORRELATION\_TABLE),

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

RETURN

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

SQRT (

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

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

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

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

**Ans:**

* **Approach:**
* Prepare the data with columns for Department, Patient Count, and demographics (Age Group, Race, Gender).
* Create a Stacked Column Chart with Department on the x-axis and Patient Count on the y-axis.
* Use Field Parameters to switch between demographic variables (Age Group, Race, Gender) to dynamically filter the data.
* **Explanation:**
* The goal of this analysis is to explore how patient demographics influence the frequency of visits to various departments. By using a Stacked Column Chart, we can see how different demographic groups contribute to the overall patient count per department.
* Field Parameters allow us to toggle between demographic categories (such as Age Group, Race, and Gender) to analyse the impact of each on department visit frequency.
* This helps in understanding whether certain departments are preferred by specific demographic groups and whether there are patterns or imbalances in departmental utilization based on patient characteristics.
* **Visualizations:**

A graph of patients with numbers and text

AI-generated content may be incorrect.

* **Insight:**
* Demographics do not affect the number of visitors to other departments because the number of visits depends entirely on the patient’s health and the location of the hospital. The demographics of different departments are unaffected directly.
* **Result:**
* The stacked column chart reveals that demographics do not significantly affect the overall visit counts across departments.

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

**Ans:**

**Trend in Patient Wait Times Throughout the Year**

* To examine the trend in patient, wait times over the year, I created a line chart titled "Patients Visit by Month" (as shown in the image provided). This chart tracks monthly changes in the average patient wait times across the specified time period.
* **Key Insights:**
* **Consistent Decline Over Time:** The chart shows a clear, consistent decrease in patient wait times, starting from 530 minutes in August 2020 and declining to 466 minutes by September 2020. This downward trend may suggest improvements in hospital efficiency or resource allocation over time.
* **Mid-Year Stability:** There is a period of relative stability in wait times between January and April, where the times remained close to 480 minutes. This may indicate a period when patient load and resource availability were balanced, allowing the hospital to maintain consistent service levels.
* **End-of-Year Drop:** The wait times continued to decrease in the later months, indicating possible operational changes, seasonal shifts in patient volume, or efficiency improvements as the year progressed.
* **Implications:**
* **Operational Efficiency:** The steady decrease in wait times suggests that the hospital might be implementing effective measures to reduce patient waiting periods. Monitoring this trend can help identify specific actions that led to these improvements.
* **Resource Planning:** Insights from this trend can support strategic planning, as understanding periods of higher efficiency could help replicate similar conditions in future months.

A graph with numbers and a line

AI-generated content may be incorrect.

A graph showing a line

AI-generated content may be incorrect.

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

**Ans:**

* **Age Group Analysis**:
* Patient satisfaction scores across different age groups range between 4.9 and 5.0.
* The 36-50 age group reported the highest satisfaction, with an average score of 5.0.
* The 65 and above age group showed the lowest satisfaction, with an average score of 4.9.
* **Visualization:**

**A graph of blue rectangular bars

AI-generated content may be incorrect.**

* **Insight:**
* This suggests that younger and middle-aged patients are slightly more satisfied compared to older patients, indicating potential areas for improvement in services for senior patients.

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?

**Ans:**

* **Approach:**
* Use a Clustered Column Chart with a Field Parameter in the X-axis, which allows switching between race and gender.
* For the Y-axis, plot the average waiting time and average satisfaction score for each group (race and gender).
* This chart helps in comparing how different race and gender groups experience their visits in terms of waiting time and satisfaction scores.
* **Explanation:**
* To investigate whether there is any evidence of racial or gender-based discrimination in the hospital, a Clustered Column Chart was used.
* The Field Parameter in the X-axis allows for a dynamic comparison between race and gender, while the Y-axis represents the average waiting time and average satisfaction score.
* By comparing these metrics, we can analyse whether any demographic group experiences significantly longer wait times or lower satisfaction, which could indicate discriminatory practices.
* **Visualization:**

A graph of blue bars with white text

AI-generated content may be incorrect.

* **Insight:**
* By examining the average waiting times and patient satisfaction scores displayed in the chart, we conclude that there is no evidence of racial or gender discrimination occurring in the hospital.
* The data shows consistent and equitable treatment across different race and gender groups, with no significant disparities in either waiting times or satisfaction.
* **Result:**

The analysis reveals that waiting times and satisfaction scores are similar across race and gender groups, suggesting that the hospital does not exhibit any discriminatory practices.

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

**Ans:**

* **Approach:**
* To determine which patients are eligible for discounts, I used a DAX formula based on two conditions:

1. A Satisfaction Score of 6 or higher.

* Using these conditions, I created a Discount Eligibility column that classifies patients as either "Eligible" or "Not Eligible." I then visualized the results in a donut chart showing the count of eligible and non-eligible patients.
* **Explanation:**
* The hospital management intended to offer discounts to high-spending and satisfied patients. I used a DAX formula to categorize patients based on their total bill and satisfaction score, creating a field to reflect whether a patient qualifies for the discount.
* The formula used is:

Discount\_Eligibility = IF ('Hospital ER'[patient\_sat\_score] >= 6, "Eligible", "Not Eligible")

* **Visualization:**

A blue circle with text

AI-generated content may be incorrect.

* **Insight:**

Out of the total 9,216 patients, 12.46% were eligible for the discount, and 87.54% were not.

* **Result:**

The donut chart shows that a small proportion (12.46%) of patients are eligible for the discount, while the majority (87.54%) do not meet the eligibility criteria.

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

**Ans:**

* **Approach:**
* I analysed the hospital's departments using key metrics: Count of Patients, Sum of Appointment Fee, and Sum of Total Bill to recommend which departments should hire 2-3 new doctors. This was done using a table visualization.
* **Explanation:**
* Departments with high patient volumes, appointment fees, and total bills were prioritized.
* Orthopedics had the highest patient count (995), appointment fees (696,500), and total bill (172,939,773).
* General Practice had a large patient count (7,240) and significant revenue.
* Neurology had fewer patients (193) but substantial fees and bills.
* **Visualization:**

A table with numbers and a blue line

AI-generated content may be incorrect.

* **Insight:**
* Orthopedics and General Practice have the highest patient load and revenue, making them top choices for new hires.
* Neurology, despite lower patient count, still shows significant demand.
* **Result:**

I suggest hiring new doctors for Orthopedics, General Practice, and Neurology based on their high patient volume and revenue generation.

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

**Ans:**

* To evaluate the **status of patient admissions**, I created a **pie chart** displaying the proportion of patients with the **patient admin flag** set to "True" (admitted) or "False" (not admitted). This pie chart visually demonstrates that **50.04%** of patients (4,612) were admitted, while **49.96%** (4,604) were not admitted. This insight is useful for understanding the hospital’s patient management and resource allocation.
* For further profitability analysis, I calculated the **Total Revenue** with a DAX measure to sum the total bills:

Total Revenue = SUM (Doctor\_Patients\_data [Total Bill])

* Additionally, I analysed **revenue by department** to assess which departments contribute most to the hospital's earnings. Using Power BI's Group By feature, I grouped data by **'department\_referral'** and summarized the **Total Bill**. This data was visualized in a **stacked column chart**, highlighting revenue distribution across departments.
* **Insights:**
* **Admission Proportion**: The pie chart shows an almost equal split in admissions, with half of the patients admitted, potentially indicating balanced usage of hospital resources.
* **Revenue Base**: Overall revenue is strong, showing financial stability.
* **Departmental Revenue Contribution**: Certain departments generate higher revenue, revealing key profit centres.
* **Recommendation**:

To improve profitability, focus on managing costs in departments with higher patient admissions and revenue. Implementing cost controls, especially for admitted patients, can enhance operational efficiency and profitability.

A screenshot of a computer screen

AI-generated content may be incorrect.A graph of revenue

AI-generated content may be incorrect.

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

**Ans:**

* **Approach:**
* To identify departments with unusually large waiting times, I created a line chart with Department on the x-axis and Average Waiting Time on the y-axis. This visualization allowed for a clear comparison of waiting times across all departments.
* **Explanation:**
* The line chart highlighted variations in waiting times for different departments.
* By analysing the chart, I observed that the Neurology department had the highest average waiting time of 36.80 minutes, which stands out compared to other departments.
* **Visualization:**

**A graph of blue rectangular objects

AI-generated content may be incorrect.**

* **Insight:**
* The Neurology department's high waiting time could indicate issues such as a lack of sufficient staff, high patient demand, or scheduling inefficiencies.
* **Result:**
* The Neurology department was identified as having the largest average waiting time of 36.80 minutes, which is significantly higher than other departments

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

**Ans:**

* To devise effective discount strategies for patients, I analyzed **patient demographics** alongside **satisfaction scores** and **total bills**. This analysis aimed to identify patterns that could guide targeted discounts for specific patient groups. For this purpose, I used a **Line Chart** on the Patient’s Tab to illustrate the relationship between patient race, average satisfaction scores, and average total bills. This visualization helped me observe demographic trends that could inform discount policies.
* From the line chart, it emerged that the **Native American/Alaska Native demographic** shows the highest average total bill and a high satisfaction score. This suggests that they contribute significantly to revenue and have a positive hospital experience, making them ideal candidates for targeted discounts.
* **Insights:**
* **Revenue and Satisfaction Correlation**: Patients in the Native American/Alaska Native group have both high satisfaction and higher billing, indicating they are valuable contributors to the hospital.
* **Targeted Discounts**: Offering discounts to this group could strengthen their loyalty and potentially boost visit frequency.
* **Recommendations**:
* **Special Discount Programs**: Implement tailored discount programs for high-contributing demographic groups, focusing on those with high satisfaction and billing.
* **Regular Strategy Review**: Regularly assess discount effectiveness to ensure continued relevance and positive impact on patient engagement.
* **Flexible Adjustments**: Be open to refining discount strategies based on ongoing analysis and feedback to better serve the diverse patient population.

A graph with blue lines and white text

AI-generated content may be incorrect.

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?

**Ans:**

* **Approach:**
* Created a table visualization with Doctor Name and Department. Used DAX to assign doctors to two shifts: "8:00 AM - 2:00 PM" and "2:00 PM - 8:00 PM".
* Applied conditional formatting to highlight the General Practice department.
* **Explanation:**
* To assign doctors to shifts, the following DAX formula was used:
* This formula ranks doctors by name and alternates the shift assignment using the MOD function.
* Dr. Johnson and Dr. Williams were assigned to the morning shift (8:00 AM - 2:00 PM) due to higher patient traffic, while Dr. Smith was assigned to the evening shift (2:00 PM - 8:00 PM) to handle lower patient volume.
* **Visualization:**

**A screenshot of a medical schedule

AI-generated content may be incorrect.**

* **Result:**
* The two-shift policy allows for even distribution of workload, ensures better patient care, and reduces wait times, particularly in the General Practice department.

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

Ans:

* A Power BI Gateway is a crucial component for securely transferring data between on-premises data sources and the Power BI service in the cloud. It acts as a bridge, enabling cloud services like Power BI to access and refresh data that is stored locally in an organization’s internal network. There are two main types of Power BI gateways:
* On-premises data gateway (standard mode): Used for scenarios where you want to connect Power BI, Power Apps, and Power Automate to your on-premises data sources.
* On-premises data gateway (personal mode): Designed for individual use, it’s typically used by a single user to connect Power BI Desktop to on-premises data sources for personal reports.
* **Key Use Cases for Power BI Gateway:**
* **Data Refresh:**

One of the most common use cases is enabling scheduled or real-time data refresh from on-premises databases to Power BI. When a report is published to the Power BI service, the gateway ensures that the data stays up-to-date, by automatically refreshing from local sources.

* **Connecting to On-Premises Data Sources:**

Power BI can connect to various on-premises data sources such as SQL Server, Oracle, SAP, and other enterprise databases through the gateway, allowing businesses to use local data with Power BI's cloud features.

* **Real-Time Data Processing**:

For organizations that require real-time insights from their on-premises data, the gateway allows Power BI to provide a live connection to the data source, with the ability to automatically push data into reports as it updates in the local system.

* **Hybrid Data Models:**

Many businesses use hybrid environments where some data resides in the cloud, and other data is stored on-premises. The gateway enables Power BI to integrate cloud and on-premises data sources within the same report or dashboard, providing a unified view.

* **Security & Compliance:**

The gateway ensures that all data connections between on-premises sources and Power BI are secure. It uses encryption and the data never leaves the corporate network unprotected. This is critical for compliance with regulations such as GDPR or HIPAA.

* **Power Apps and Power Automate Integration**:

Besides Power BI, the on-premises data gateway is also used by Power Apps and Power Automate for real-time or scheduled access to on-premises data, making it a versatile tool in the Microsoft ecosystem.

* **Support for Custom Data Connectors:**

If businesses use custom or specialized on-premises data systems, the gateway can also support custom data connectors that help extend Power BI’s capabilities beyond the built-in connectors.

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

**Ans:**

* **Clarify the Goal:**
* First, I would try to understand the core objective or purpose of the task. What is the end result that’s desired? In this case, the goal seems to be to explain or analyze Power BI gateways, but without explicit questions, I would clarify if the focus should be on:

1. **Technical understanding** (how Power BI gateway works),
2. **Use cases** (when and why to use it), or
3. **Implementation** (how to set it up or troubleshoot it).

* **Break Down the Key Areas:**
* If the specific questions aren’t provided, I would break down the topic into key areas. For Power BI Gateway, these could be:

1. **Definition and Overview:** Explaining what a Power BI Gateway is.
2. **Types of Gateways:** Differentiating between the two modes of the gateway: standard and personal.
3. **Use Cases:** Detailing how it’s applied in various scenarios, like data refresh, connecting to on-premises data sources, and real-time processing.
4. **Configuration:** Basic steps to set it up.
5. **Security Considerations:** How it ensures data protection.
6. **Troubleshooting:** Common problems and solutions.

* **Start with a General Overview:**
* Without questions, I would begin by explaining the concept at a high level. I'd start with the basics — what Power BI Gateway is, why it’s needed, and how it fits into the broader Power BI ecosystem.
* **Provide Examples and Scenarios:**
* To make the information more practical and relatable, I would provide real-life examples and scenarios where Power BI gateways would be used. This helps readers (or listeners) understand the relevance of the gateway in different business contexts.
* **Address Potential Issues:**
* Next, I would consider the challenges that one might face with Power BI Gateways. These could include common issues like connectivity problems, gateway performance, data refresh failures, and security configurations.
* **Conclude with Practical Steps:**
* Finally, without specific questions, I would conclude by summarizing how an individual or organization can get started with setting up the gateway, along with best practices for maintaining and optimizing it.

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

**Ans:** I analyzed the relationship between **Doctor ID** and **Department Referral** using the **bar chart** visual. The data shows the count of distinct **Doctor IDs** for each department. By plotting this chart, I can easily observe how many unique doctors are assigned to different departments.

* **Insights:**
* The departments **Gastroenterology**, **Orthopedics**, and **Physiotherapy** each have the highest number of distinct doctors, with 4 doctors each.
* Other departments like **Cardiology**, **General Practice**, **Neurology**, and **Renal** have 3 distinct doctors each.
* **Conclusion:**
* The analysis indicates a **one-to-many** relationship between **Department** and **Doctor ID**, where one department has multiple doctors assigned to it.
* **Recommendations:**
* If the goal is to streamline patient management, I could consider increasing the number of doctors in departments with higher patient loads.
* I might also explore whether some doctors can be shared across departments to optimize staffing, especially in departments with fewer assigned doctors.
* For deeper insights, I can check if any doctor is working in multiple departments to verify if there is any **many-to-many** relationship present.

**A blue bar graph with black text

AI-generated content may be incorrect.**

**Report**

The hospital has asked for a report with three tabs:

* Main Tab

A screenshot of a medical information

AI-generated content may be incorrect.

* Doctors’ Tab

A screenshot of a medical report

AI-generated content may be incorrect.

* Patients’ Tab

A screenshot of a medical report

AI-generated content may be incorrect.