1. **PROJECT OVERVIEW**

* **Objective:** The goal of this project is to build an interactive analytics dashboard using Python, SQL, and Streamlit to assist healthcare organizations in making data-driven decisions. The dashboard will enable in-depth analysis of various key aspects, including treatment patterns, facility utilization, length of stay, billing versus insurance discrepancies, follow-up appointment trends, bed occupancy etc. By providing actionable insights into these areas, the project aims to optimize hospital operations, improve resource management, and enhance patient care outcomes. This will be accomplished through robust data preprocessing, effective database integration, and the development of a user-friendly platform that allows healthcare administrators to interact with and visualize the data in real time.
* **Scope:** This project focuses on developing an interactive analytics dashboard to provide insights into key healthcare metrics. The scope includes analyzing treatment patterns, bed occupancy, facility usage, billing discrepancies, follow-up appointment trends, and length of stay. Data will be preprocessed and stored in an SQL database for efficient querying and analysis. The dashboard, built using Python and Streamlit, will allow users to visualize trends, filter data, and generate actionable insights in real-time. The project aims to support healthcare administrators in making data-driven decisions for resource optimization, operational efficiency, and enhanced patient care.

1. **Data Analysis Methodology**
2. **Data Collection:** The data used in this project was provided in the form of an Excel file titled **"Healthcare-Dataset.xlsx."** The dataset contains detailed records of patient admissions, treatments, billing, and healthcare provider information.

It includes the following key columns:

* **Patient\_ID**: A unique identifier for each patient.
* **Admit\_Date:** The date the patient was admitted to the hospital.
* **Discharge\_Date:** The date the patient was discharged from the hospital.
* **Diagnosis:** The medical condition diagnosed for the patient.
* **Bed\_Occupancy:** The type of bed assigned to the patient (e.g., General, ICU, Private).
* **Test:** The diagnostic test performed on the patient (e.g., MRI, X-ray, CT scan).
* **Doctor:** The name of the doctor responsible for the patient’s treatment.
* **Followup Date:** The date scheduled for the patient’s follow-up appointment.
* **Feedback:** The rating provided by the patient for their care (on a scale of 1 to 5).
* **Billing Amount:** The total amount billed for the patient’s treatment.
* **Health Insurance Amount:** The amount covered by the patient's health insurance.

This dataset serves as the primary source for analysis, providing the foundation for understanding various hospital operations, patient trends, and treatment outcomes.

1. **Data Cleaning and Preprocessing:**

The dataset was preprocessed to ensure its quality and suitability for analysis. The following steps were performed:

* **Handling Missing Values**: The dataset contained some missing values in the “Followup Date” column. These missing values were addressed by dropping the rows with missing follow-up dates to ensure that only complete records were used in the analysis. The deleted rows accounted for only 1.7% of the total dataset, which is a very small portion. Therefore, the remaining data was sufficient for analysis and did not significantly affect the overall results.
* **Date Formatting**: The “Admit\_Date”, “Discharge\_Date”, and “Followup Date” columns were ensured to be in the correct date format to enable accurate calculations of length of stay and follow-up periods.
* **Checking and Modifying Data Types**: The datatypes for all the columns were checked to ensure consistency and accuracy. The datatype for the Billing Amount column was found to be an integer, so it was changed to a float data type. This modification was made to allow for more precise calculations, especially for handling fractional values, and to ensure it aligns with the proper numerical representation of currency values.
* **Handling Duplicates**: Duplicate rows were checked in the dataset, and it was confirmed that no duplicate rows were found. This ensures that the data is unique and reliable for analysis.

1. **AnalysisTechniques**:  
   For the analysis, the cleaned data was first inserted into an SQL database using a Python-MySQL connection. This enabled efficient querying and aggregation of the data to gain insights into key patterns such as treatment trends, billing discrepancies, and bed occupancy etc. SQL queries were used to extract relevant subsets of the data, which were then analyzed further.

Once the necessary data was extracted, **Plotly** was used to create interactive visualizations, including bar charts, line charts, scatter plots, and pie charts etc. Plotly’s capabilities provided dynamic, interactive plots that made it easier to identify trends, such as seasonal admission patterns and differences between billing and insurance amounts.

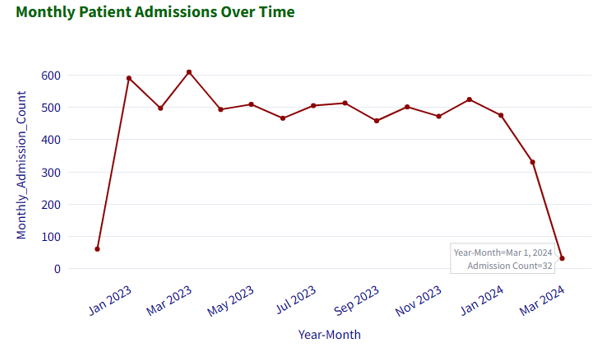
Additionally, a **Streamlit** application was developed to build an interactive dashboard, allowing users (e.g., healthcare administrators) to explore the data, filter results, and view visual insights in real time. The platform's interactive capabilities helped users interact with the data seamlessly and make informed decisions based on the visualized trends.

1. **Individual Use Cases, Visualization Insights and Recommendations:**

**USE CASE 1: Trends in Admission Over Time**

**Use Case Description**: Analyze monthly patient admissions to identify trends and seasonal patterns, helping hospitals manage resources and plan for peak periods effectively.

**Visusalization:** The graph depicts the trend of patient admissions over time, broken down by month and year. This visualization helps to uncover seasonal fluctuations and trends in admissions, enabling hospitals to anticipate demand and plan resources effectively. By analyzing monthly admission counts, hospitals can forecast peak periods, optimize staffing levels, and ensure sufficient resources are available to accommodate patient needs.



**Insights from the graph:**

The graph shows monthly patient admissions from January 2023 to March 2024. Here are some observations:

* **Peak Admission Season:** There is a significant peak in patient admissions during the first quarter of 2023, reaching its highest point in March 2023.
* **Decreasing Trend:** After the peak in March 2023, the number of admissions generally trended downwards until January 2024.
* **Seasonal Fluctuation:** The graph suggests a potential seasonal pattern, with admissions generally higher during the winter months and lower during the summer months.
* **Sharp Decline:** A sharp decline in admissions is seen between January 2024 and March 2024, with the number of admissions falling to a minimum in March 2024.

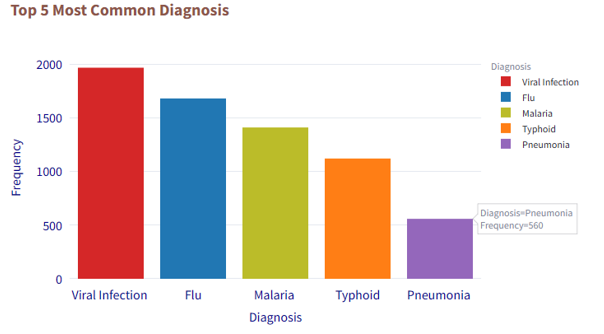
**Recommendations:**

1. **Increase Resources During Peak Admission Periods:** Given the significant peak in admissions during the first quarter of 2023, particularly in March, it is advisable to allocate additional resources—such as staffing, medical equipment, and bed capacity—during this period to ensure hospitals are prepared to handle the surge in patient volume.
2. **Monitor and Address Sharp Decline Post-January 2024:** The sharp decline in admissions from January to March 2024 suggests that hospitals should monitor the reasons behind this drop (e.g., changes in patient behavior or external factors). Understanding these trends can help adjust resource allocation and anticipate the needs of patients during periods of low admissions.
3. **Leverage Off-Peak Months for Operational Improvements:** During periods of lower admissions, such as spring months, hospitals can focus on optimizing processes, conducting staff training, and performing maintenance on equipment. This will ensure the hospital is better prepared for future peak periods and maintain efficient operations year-round.

**USE CASE 2: Diagnosis Frequency Analysis**

**Use Case Description:** Analyze the most common diagnosis to identify prevalent health conditions, enabling hospitals to prioritize care, allocate resources effectively, and plan for necessary treatments.

**Visualization:** This chart shows the top 5 most common diagnosis. It helps identify the most common health conditions, useful for healthcare planning and resource allocation.



**Insights from the Graph:**

* **Viral Infection:** The most frequent diagnosis with a count of 1971 cases.
* **Flu:** The second most frequent diagnosis with a count of 1684 cases.
* **Malaria:** The third most frequent diagnosis with a count of 1413 cases.
* **Typhoid:** The fourth most frequent diagnosis with a count of 1123 cases.
* **Pneumonia:** The fifth most frequent diagnosis with a count of 560 cases.

**Recommendations:**

**Allocate Resources for Viral Infections & Flu:** Increase staffing, medications, and equipment for managing viral infections and the flu, especially during peak seasons.

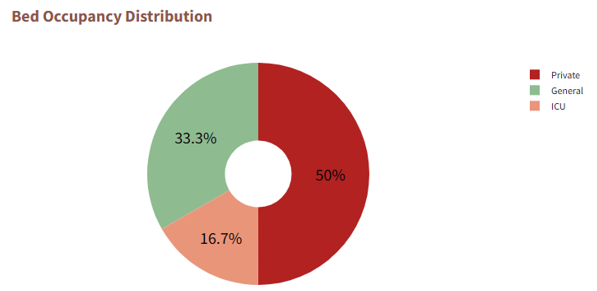
**Prioritize Malaria & Typhoid Treatment:** Focus on prevention and treatment of malaria and typhoid through awareness campaigns, vaccinations, and adequate medical resources.

**Enhance Pneumonia Care:** Ensure sufficient ICU capacity and respiratory support for pneumonia cases, despite their lower frequency.

**3.USE CASE 3: Bed Occupancy Analysis**

**Use Case Description:** Analyze bed occupancy patterns to optimize resource allocation and manage bed availability for improved operational efficiency.

**Visualization:** This chart shows the distribution of bed occupancy types, offering insights for effective resource management and planning.



**Insights from the Graph:**

* **Private beds:** Make up 50% of bed occupancy, indicating high demand for private care.
* **General beds:** Account for 33.3%, showing strong demand for general care services.
* **ICU beds:**  Represent 16.7% of occupancy, suggesting lower demand for critical care.

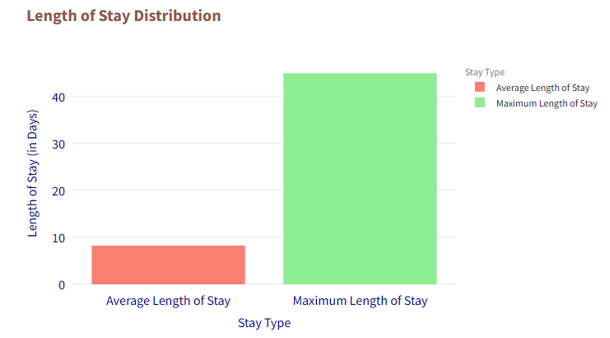
**Recommendations:**

* **Increase Private Bed Availability:** Given the high demand for private beds, consider expanding private bed capacity to meet patient preferences and needs.
* **Optimize General Bed Usage**: With 33.3% of beds occupied by general care patients, ensure efficient management and timely turnover to maximize bed availability.
* **Monitor ICU Utilization:** Track ICU bed usage and adjust capacity to ensure readiness for critical care.

**USE CASE 4: Length of stay distribution analysis**

**Use Case Description:** Analyze the average and maximum length of patient stays to identify trends in hospital resource usage, optimize bed management, and improve discharge planning for better operational efficiency.

**Visualization:** The graph shows the average and maximum length of stay for patients. It provides insights into patient stay durations, which can help optimize hospital resource allocation and improve care management.



**Insights from the Graph:**

* **Average Length of Stay:** The average length of stay is approximately 8 days.
* **Maximum Length of Stay:** The maximum length of stay is approximately 45 days.

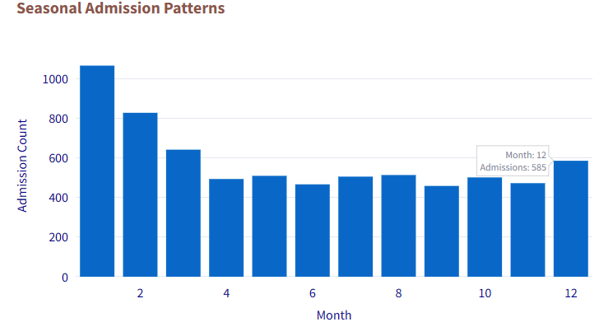
**Recommendations:**

* **Address Long Stays:** The maximum stay of 45 days suggests the need for a review of long-term cases. Hospitals should identify the causes (e.g., complex cases or delays in discharge) and implement strategies to reduce unnecessary extended stays.
* **Optimize Bed Management:** With an average stay of 8 days, improve bed turnover by streamlining discharge planning and patient flow to minimize delays in bed availability.
* **Implement Predictive Planning:** Use average and maximum length of stay data to forecast future bed usage and staffing needs, ensuring adequate resources are available for both short-term and long-term patients.

**USE CASE 5: Seasonal Admission Patterns**

**Use Case Description:** Analyze monthly patient admissions to identify seasonal trends, enabling hospitals to anticipate peak periods and optimize resource allocation, staffing, and capacity planning accordingly.

**Visualization:** The below graph reveals seasonal trends in patient admissions by month, helping to understand hospital usage patterns and plan resource allocation for peak periods.

****

**Insights from the graph:**

* **Peak Season:** Highest admissions in January, likely due to winter illnesses.
* **Decline After Peak:** Admissions drop from January to March as patients recover.
* **Lower Seasons**: Relatively stable admissions from April to October, with a slight rise in summer.
* **December Increase:** Admissions increase in December, possibly due to the holiday season.

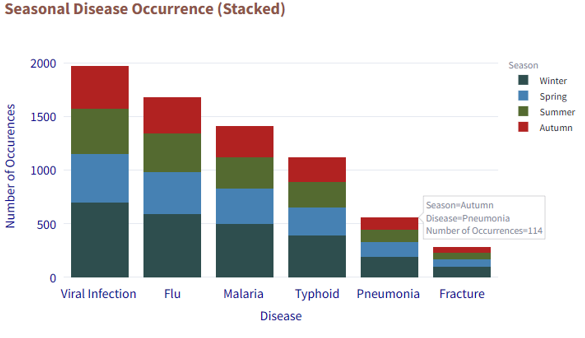
**Recommendations:**

* **Focus on Prevention:** Implement preventive healthcare programs during peak seasons to reduce admissions.
* **Continuous Monitoring:** Regularly monitor admission trends and analyze the data to identify patterns allocate resources like medical equipments and beds etc.

**USE CASE 6: Seasonal Trends in Disease Occurrence**

**Use Case Description:** Analyze seasonal patterns of disease occurrences to identify prevalent diagnoses in each season.

**Visualization:** The graph identifies which diseases or diagnosis are more prevalent during different seasons. It helps healthcare providers understand seasonal patterns of diseases, enabling proactive measures to prevent outbreaks and improve patient care during high-risk periods.



**Insights from the graph:**

* **Winter Peak:** All diagnoses are at their highest during the winter season.
* **Spring Trends:** Viral infections peak in spring, while fractures are least common during this season.
* **Gradual Decline:** Diagnosis counts gradually decrease from winter to spring, then to summer, and finally to autumn.

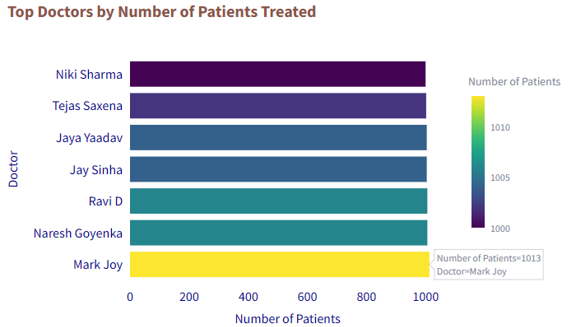
**Recommendations:**

* **Prepare for Winter Peaks:** Increase resources, such as staffing, medical supplies, and equipment, to handle the surge in diagnoses during the winter season.
* **Focus on Viral Infections in Spring:** Allocate more resources to manage viral infections during the spring, while ensuring that fracture care remains available but less prioritized.
* **Adjust Resource Allocation Throughout the Year:** As diagnosis decrease from winter to autumn, optimize staffing and medical resources accordingly to align with lower patient volumes, while maintaining readiness for seasonal shifts.
* **Launch Public Health Awareness Campaigns:** Educate the public about diseases that are more common in each season. Provide information on preventive measures, symptoms to watch for, and guidance on when to seek medical attention, helping to reduce seasonal outbreaks and improve early intervention.

**USE CASE 7: Top Doctors by Number of Patients Treated**

**Use Case Description:** Identify the most frequently consulted doctors to optimize staff allocation.

**Visualization**: This chart shows the doctors who have treated the most patients. Understanding which doctors have the highest patient load can assist in effective resource planning and ensure that staffing is aligned with demand.

****

**Insights from the graph:**

* **Top Doctor:** Mark Joy treated the most patients, with a total of 1,013.
* **High Patient Counts:** Naresh Goenka and Ravi D each treated 1,006 patients.
* **Declining Trend:** The number of patients treated by other doctors gradually decreases, with counts dropping to 1,000.
* **Top Performers:** The top 7 doctors have treated over 1,000 patients each.

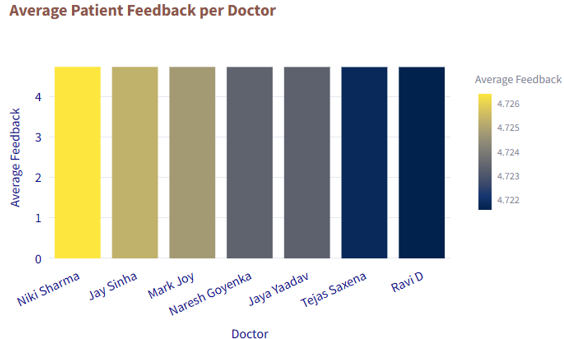
**Recommendations:**

* Allocate additional resources (staff, medical support) to Mark Joy, Naresh Goenka and Ravi D to manage his high patient load more effectively.
* Consider offering additional training or mentorship to them to ensure that he is able to provide high-quality care to all of their patients.
* Explore ways to improve the efficiency of the appointment scheduling system to reduce patient wait times.
* Distribute patient load more evenly across doctors by considering additional support or reassigning cases to help reduce the strain on the top performers.

**USE CASE 8: Patient Feedback Analysis per Doctor**

**Use Case Description**: Analyze patient feedback scores per doctor to identify areas for improvement in care and enhance patient satisfaction.

**Visualization:** This chart analyzes the average patient feedback for each doctor. By identifying areas with lower satisfaction scores, healthcare providers can focus on improving patient care, communication, and facilities to enhance overall patient experience.

****

**Insights from the graph:**

* **High Average Feedback:** Niki Sharma has the highest average patient feedback score, indicating a high level of patient satisfaction.
* **Consistent High Scores:** Jay Sinha, Mark Joy, Naresh Goyenka, and Jaya Yaadav all have similar average feedback scores that are above the average, suggesting consistent high patient satisfaction.
* **Lower Average Feedback:** Tejas Saxena and Ravi D have slightly lower average feedback scores compared to the other doctors.

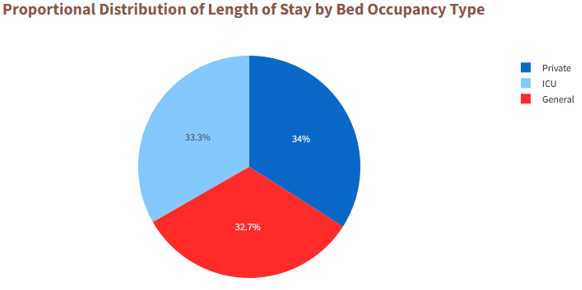
**Recommendations:**

* Recognize and share best practices from doctors like Niki Sharma, Jay Sinha, and Mark Joy to maintain high patient satisfaction levels across the team.
* For Tejas Saxena and Ravi D, review patient feedback to identify areas for improvement in care, communication, or facilities and implement targeted training or support.
* Regularly track patient feedback to ensure consistent quality care and promptly address any emerging issues to maintain overall satisfaction.

**USE CASE 9: Length of Stay and Bed Utilization Analysis**

**Use Case Description**: Analyze the average length of stay (LOS) for patients by bed occupancy type to identify which beds are occupied for longer periods.

**Visualization:** The below graph calculates the average length of stay for patients by bed occupancy type, helping to identify which bed types are occupied longer. This insight aids in optimizing bed management and resource allocation.

****

**Insights from the graph:**

The pie chart represents the proportional distribution of Length of Stay (LOS) by bed occupancy type. The following insights can be drawn from the data:

* **Private beds** have the highest proportion of LOS (34%).
* **ICU beds** have the second highest proportion of LOS (33.3%).
* **General beds** have the lowest proportion of LOS (32.7%).

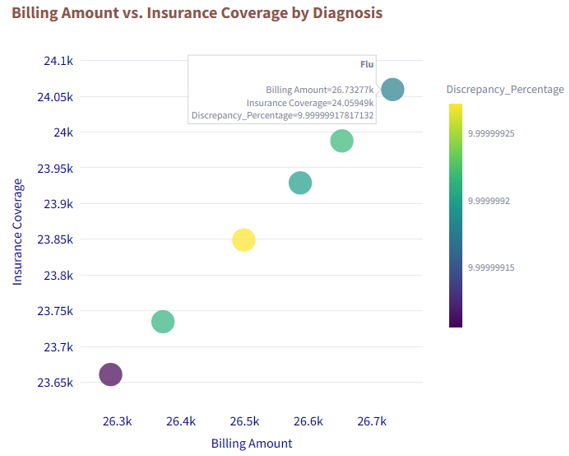
**Recommendations:**

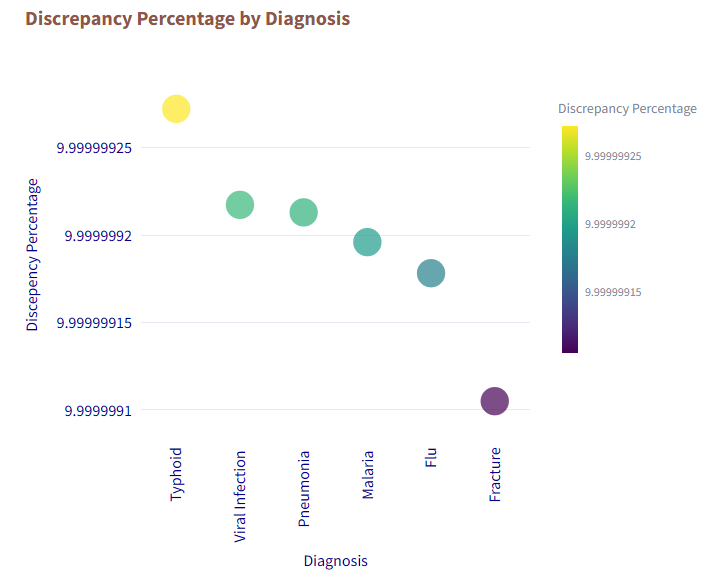
* Since private and ICU beds account for the highest proportion of Length of Stay (LOS), allocate more resources (staff, medical equipment) to these areas to manage prolonged patient stays efficiently.
* Track ICU and private bed usage regularly to avoid overcrowding and ensure patients are moved to the right care setting promptly.
* Since general beds have a lower LOS, focus on improving turnover to maximize capacity while maintaining quality care.

**USE CASE 10: Average Billing vs. Insurance Coverage Discrepency by Diagnosis**

**Use Case Description:** Analyze the discrepancy between average billing amounts and insurance coverage for different diagnosis to identify potential gaps in coverage and billing accuracy. This helps healthcare providers optimize billing processes and improve financial planning.

**Visualization:** The graph compares the average billing and insurance coverage by diagnosis, highlighting discrepancies. It helps in identifying areas where insurance coverage is insufficient compared to billing amounts, aiding in better resource management and negotiation strategies.

****

****

**Insights from the graph:**

* Typhoid shows the highest discrepancy between billing and insurance coverage.
* Viral Infection and Pneumonia have nearly identical discrepancies.
* Discrepancies decrease for Malaria and Flu, with the lowest being seen in Fractures

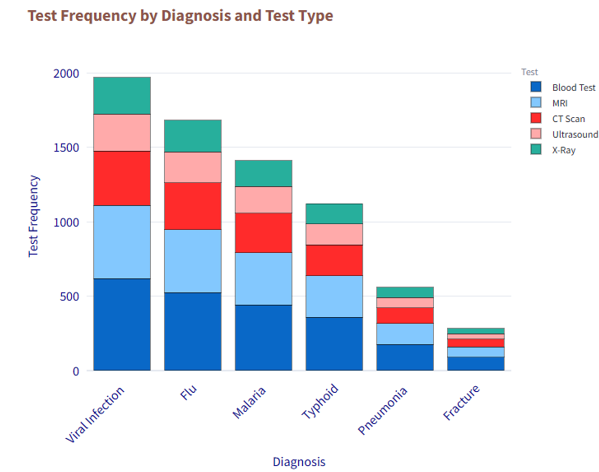
**Recommendations:**

Improve billing accuracy, hospitals should regularly review and update their billing and insurance policies to ensure they match correctly. Standardizing billing practices across different diagnoses can help reduce discrepancies. It’s also important to maintain clear communication with insurance providers to make sure patients are receiving the right coverage. Finally, hospitals should keep monitoring any discrepancies and make adjustments as needed to ensure all billing is accurate and fair for patients.

**USE CASE 11: Relation Between Test and Diagnosis**

**Use Case Description:** Analyze the relation between diagnosis and the tests most frequently ordered for each condition.

**Visualization:** The below graph identifies the most commonly ordered tests for each diagnosis, helping optimize test selection and improve patient care. It streamlines workflows and reduces unnecessary testing.

****

**Insights from the graph:**

* **Viral Infection:** Blood tests are the most common tests ordered for viral infections.
* **Flu:** Blood tests are the most frequent for flu, with MRI being the second most common.
* **Malaria:** Blood tests lead, followed by MRI and CT scans.
* **Typhoid:** Blood tests and MRIs are most frequently ordered for typhoid.
* **Pneumonia:** Blood tests and CT scans are the top tests for pneumonia.
* **Fracture:** Blood tests and CT scans are the most commonly ordered for fractures.
* **Ultrasound and X-rays:** These are the least commonly ordered tests across all diagnoses.

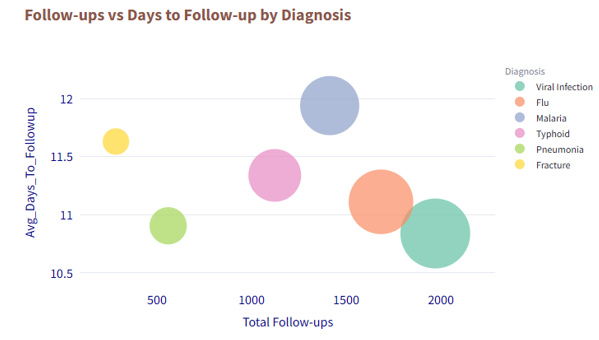
**Recommendations:**

* Since blood tests are the most common across several diagnosis, ensure they are ordered efficiently to avoid overuse and reduce costs, while ensuring they are necessary for accurate diagnosis.
* For conditions like Malaria, Pneumonia, and Typhoid, optimize the use of CT scans and X-rays to avoid excessive or inappropriate imaging while ensuring accurate diagnosis.
* Given the lower frequency of ultrasound and X-rays, evaluate whether these tests could be more useful for certain diagnosis. Consider incorporating them in case of fracture and avoid unnecessary tests.
* Establish clear guidelines for when each test should be ordered to improve consistency, reduce unnecessary testing, and ensure timely and accurate diagnosis.

**USE CASE 12: Track Follow-up Appointments and Patient Outcomes**

**Use Case Description:** Monitor patient follow-up appointments and the time taken for follow-ups after discharge, helping the hospital assess patient engagement and identify patterns that may impact treatment outcomes.

**Visualization:** This analysis tracks follow-up appointments and calculates the average time to follow-up for each diagnosis. It helps ensure continuity of care and improve treatment outcomes.

****

**Insights from the graph:**

The graph shows the relationship between the total number of follow-ups and the average number of days to follow-up, categorized by diagnosis.

Observations:

* Viral Infection has the highest number of follow-ups (2000) and the average days to follow-up is around 11.
* Flu has a slightly lower number of follow-ups (1684) with the average days to follow-up being around 12.
* Malaria has a considerably lower number of follow-ups (1413) with an average days to follow-up of around 12.
* Typhoid has 1123 follow-ups with an average days to follow-up of around 11.3.
* Pneumonia has 560 follow-ups with an average days to follow-up of around 11.
* Fracture has the lowest number of follow-ups (284) with an average days to follow-up of around 11.6.

Viral Infections and Flu require more frequent follow-ups, possibly due to the nature of these illnesses. Malaria and Typhoid have a similar average days to follow-up, indicating that the follow-up time may be related to the severity of the illness. Pneumonia and Fracture have a lower number of follow-ups, indicating a shorter duration of treatment and possible faster recovery.

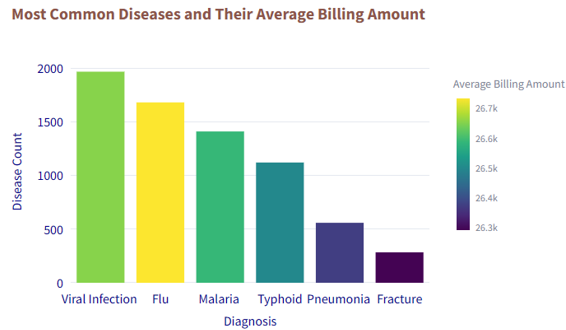
**Recommendations:**

* **Enhance Follow-Up for Viral Infections and Flu:** Given their higher follow-up frequency, ensure timely follow-ups for these conditions to monitor recovery and prevent complications.
* **Streamline Follow-Up for Malaria and Typhoid:** With similar follow-up timings, consider optimizing patient education and communication to ensure follow-ups are done on time, reflecting the severity of the illness.
* **Monitor Follow-Up Frequency for Pneumonia and Fractures:** Given their lower follow-up rates, ensure patients with these conditions are receiving adequate post-treatment care and check-ups to confirm complete recovery.

**USE CASE 13: Most Common Diseases and their Average Billing Amount**

**Use Case Description**: Identify the most common diseases and analyze the average billing amount for each.

**Visualization:** The below graph identifies the most common diseases and their average billing amounts, helping hospitals plan financial resources and allocate funds effectively.



**Insights from the graph:**

* Viral Infection is the most common disease, followed by Flu, Malaria, Typhoid, Pneumonia, and Fracture.
* The color gradient of the bars indicates the average billing amount, with yellow representing the highest billing amount (for Flu) and purple representing the lowest billing amount (for Fracture).

These insights can help the hospital focus on diseases with high treatment costs and manage resource allocation accordingly.

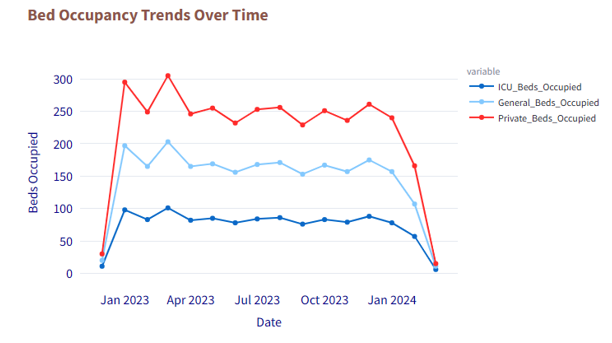
**Recommendations:**

* **Focus on High-Cost Diseases:** Given that Flu has the highest average billing amount, the hospital should prioritize cost management strategies, such as negotiating with suppliers or optimizing treatment protocols to reduce costs.
* **Enhance Resource Allocation:** For common diseases like Viral Infections and Malaria, ensure adequate resources (staff, medication, equipment) to handle higher patient volumes without compromising care.
* **Monitor Low-Cost Diseases:** While Fracture has lower average billing amounts, the hospital should still track these conditions for potential improvements in care delivery and efficiency.
* **Improve Financial Planning:** Use the analysis of disease frequency and billing amounts to refine financial planning, allocate budgets for high-cost diseases, and anticipate future healthcare needs.

**USE CASE 14: Track Bed Occupancy Trends and Identify Peak Admission Periods**

**Use Case Description:** Utilize bed occupancy data to analyze admission trends, pinpoint peak periods, and optimize bed management, staffing, and resource allocation to enhance preparedness for high-demand periods.

**Visualization:** The graph below shows the number of beds occupied in each category based on year and month.



**Insights from the graph:**

* **Peak Admissions:** Private bed occupancy generally peaked in April 2023 and experienced another rise in December 2023.
* **Trends in Private Bed Occupancy:** There is a notable decline in private bed occupancy from December 2023 to January 2024.
* **General Bed Occupancy:** General bed occupancy remained relatively stable with slight fluctuations throughout the year.
* **ICU Bed Occupancy:** ICU bed occupancy showed a steady trend, remaining consistently lower than both general and private bed occupancy.

**Recommendations:**

**Private Beds:**

* Prepare for higher demand in April and December by adjusting staffing and resources.
* Investigate the drop in private bed occupancy from December to January and consider strategies to fill beds, such as promotions or adjusting availability.

**General Beds:**

* Keep staffing flexible to handle small changes in occupancy throughout the year.

**ICU Beds:**

* Ensure ICU beds are efficiently used, and monitor for potential spikes, especially during flu season.

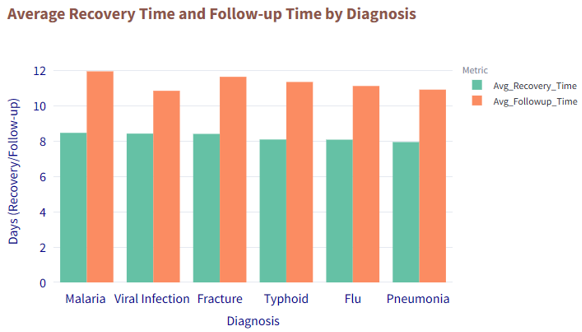
**Overall Resource Management:**

* Use past data to better predict peak times and plan staffing accordingly.
* Adjust staffing and resources ahead of peak periods to ensure smooth operations.

**USE CASE 15: Monitor Patient Recovery and Treatment Outcomes**

**Use Case Description:** Analyze patient recovery and follow-up times across different diseases to identify factors that contribute to faster recovery.

**Visualization:** The below graph evaluates recovery times and treatment outcomes for different diseases, providing insights into how treatment strategies impact recovery. It aids in refining treatment plans and enhancing patient care by optimizing recovery processes.

****

**Insights from the graph**

* Malaria and Viral Infection have the longest average recovery times.
* Pneumonia has the shortest average recovery time.
* Malaria and Fracture have the highest average followup times.
* Follow-up times are generally consistent across other different diagnosis.
* Average follow-up time is longer than average recovery time for all diagnosis.

**Recommendations:**

**Malaria and Viral Infections:**

* Review and improve treatment plans to reduce recovery times.
* Allocate more resources (staff, equipment) for better management.

**Pneumonia:**

* Study successful pneumonia treatments and apply effective methods to other conditions.

**Fractures:**

* Check if all follow-ups are necessary and explore ways to reduce follow-up times.

**Follow-up Times Consistency:**

* Streamline the follow-up process to make it more efficient across different conditions.

**Follow-up vs Recovery Times:**

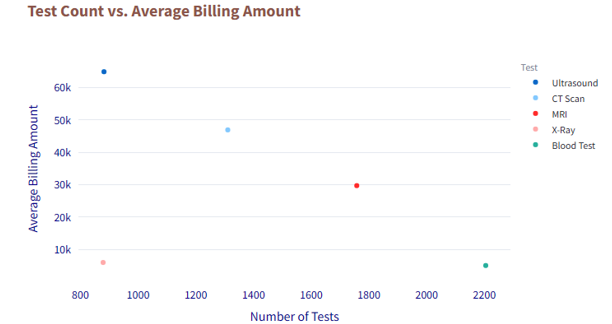
* Assess if follow-ups can be shortened or adjusted to avoid unnecessary visits.

Overall, optimizing treatment and follow-up processes can lead to faster recovery and better resource management across the hospital.

**USE CASE 16: Impact of Tests on Total Billing Amount**

**Use Case Description:** Analyze the medical tests that contribute the most to total billing amounts by evaluating their average billing and frequency.

**Visualization:** The below graph identifies the medical tests that contribute most to patient billing amounts. By understanding the tests that lead to higher costs, hospitals can optimize resource allocation and cost management strategies.

****

**Insights from the graph:**

* **Blood tests** have the highest test count at 2205, with an average billing amount slightly lower than **X-rays**, which have a test count of 879.
* **Blood tests** and **X-rays** have the lowest average billing amounts compared to other tests.
* **MRI** has the second-highest test count, with an average billing amount of approximately 29.67k.
* **CT scans** have the third-highest test count at 1311, surpassing **ultrasound** and **X-rays**, with an average billing amount around 46.89k.
* **Ultrasound** has the highest average billing amount compared to all other tests, with a test count similar to that of **X-rays**.

**Recommendations:**

1. **Blood Tests and X-rays:**

* Optimize Cost-Effectiveness: While they have high test counts, the lower average billing amount suggests that these tests are more affordable. Consider ensuring these are used efficiently, but also explore bundling or pricing strategies for higher volume to optimize revenue.

1. **MRI and CT Scans:**

* Review High-Cost Tests: With MRI and CT scans having high average billing amounts, the hospital should evaluate the necessity of these tests to avoid overuse. Ensure they are recommended based on clear medical need to control costs while maintaining quality care.

1. **Ultrasound:**

* Control Utilization: Since ultrasound has the highest average billing amount, the hospital should ensure that it is being used appropriately. Review cases where ultrasound is recommended and consider whether it is overused or if cheaper alternatives can be used in certain situations.

1. **Cost Management and Resource Allocation:**

* Strategize Resource Allocation: For high-volume and high-cost tests like CT scans and ultrasounds, ensure that resources are allocated efficiently (e.g., managing equipment, staff, and scheduling) to balance cost and patient care quality.
* Invest in Cost-Effective Tests: Focus on optimizing the use of low-cost tests like blood tests and X-rays, but consider cost management strategies to prevent underutilization of higher-cost, high-value tests.

1. **Conclusion:**

The project brings significant value to the hospital by providing data-driven insights into test utilization and billing. It helps the hospital make informed decisions on resource allocation, improve cost management strategies, and ensure that high-cost tests are used efficiently. These improvements can ultimately lead to better financial sustainability for the healthcare system while maintaining high-quality patient care.