Healthcare

Case Study: Advancing Healthcare Analysis through Data Insights

Background

You are a data analyst at HealthStat Solutions, a company specializing in healthcare analytics. You have been given two datasets: 'Patient Medical Records' and 'Hospital Treatment Details'. The 'Patient Medical Records' dataset contains detailed information on patients, including their age, gender, blood type, diagnosis, treatments, admission and discharge dates, and total bills. The 'Hospital Treatment Details' dataset provides insights into the hospitals treating these patients, including the treating doctor, room number, daily costs, treatment types, and recovery ratings.

In a healthcare industry that relies heavily on data to make informed decisions for patient care and hospital management, your role is crucial. Your task is to analyze these datasets, uncovering trends and insights that could improve patient outcomes and optimize hospital operations.

Objective

As an analyst at HealthStat Solutions, your objective is to leverage Power BI for a deep dive into the provided healthcare datasets. This task encompasses meticulous data cleaning and sophisticated data modeling, utilizing DAX for advanced analytics. Your goal is to create a comprehensive, interactive dashboard in Power BI that presents a cohesive narrative of the healthcare data. This dashboard should serve as a tool to uncover and visualize important trends, such as the interplay between patient demographics and treatment outcomes, cost implications of various medical procedures, and overall hospital performance metrics. Your analysis will provide invaluable insights, aiding healthcare providers in enhancing patient care and operational efficiency, and positioning HealthStat Solutions at the forefront of healthcare analytics.

Data Source:

HealthcareDataset1.xlsx

The "HealthcareDataset1.xlsx" file contains the following columns:

1. **PatientID**: A unique identifier for each patient. (Primary Key)
2. **PatientName**: Name of the patient.
3. **Age**: Age of the patient.
4. **Gender**: Gender of the patient.

5. **BloodType**: Blood type of the patient.
6. **Diagnosis**: The diagnosis given to the patient.
7. **Treatment**: The treatment provided to the patient.
8. **AdmissionDate**: Date when the patient was admitted.
9. **DischargeDate**: Date when the patient was discharged.
10. **TotalBill**: The total bill amount for the patient's treatment.
11. **Full Prescription Details**: Detailed prescription information including medication names, dosages, frequency, and duration

HealthcareDataset2.xlsx

The "HealthcareDataset2.xlsx" file contains the following columns:

1. **PatientID**: A unique identifier for each patient, corresponding to 'PatientID' in "HealthcareDataset1.xlsx". (Foreign Key)
2. **Hospital**: The name of the hospital where the patient was treated.
3. **DoctorName**: Name of the doctor who treated the patient.
4. **RoomNumber**: The room number assigned to the patient.
5. **DailyCost**: The daily cost of the patient's treatment.
6. **TreatmentType**: Type of treatment provided.
7. **RecoveryRating**: A rating of the patient's recovery (out of 10).

Part 1: Data Cleaning, Modeling, and DAX in Power BI

1. **Data Importing and Initial Examination**
   ○ Import both datasets into Power BI. Perform a preliminary examination of the data. Identify any data quality issues or inconsistencies.
2. **Merging and Relating Datasets**
   ○ Merge the datasets using a suitable column as a key. Ensure that the merge is accurate and retains all necessary information.
3. **Cleaning: Handling Missing and Irrelevant Data**
   ○ Identify and address missing data in both datasets. Address duplicate entries and irrelevant data points, ensuring data quality.
4. **Data Type Conversion**
   ○ Convert 'AdmissionDate' and 'DischargeDate' to the appropriate date formats. Calculate the length of stay for each patient.
5. **Categorizing Age Groups**
   ○ Create a new column categorizing patients into age groups (e.g., Child, Adult, Senior). What are the predominant age groups?
6. **Analysis of Treatment Costs**
   ○ Calculate the total cost of treatment for each patient (consider 'TotalBill' and 'DailyCost'). Which treatments are the most expensive on average?
7. **Gender Distribution in Diagnosis**
   ○ Analyze the distribution of genders across different diagnoses. Are there any noticeable patterns or trends?

8. **Blood Type Analysis**
   - Investigate the distribution of blood types among the patients. Is there a blood type that is more prevalent?
9. **Recovery Rating Analysis**
   - Using DAX, analyze the average recovery rating by treatment type. Which treatment type has the highest average recovery rating?
10. **Hospital Utilization Analysis**
    - Calculate the average number of patients per room in each hospital. Which hospital has the highest utilization?
11. **Doctor's Patient Load**
    - Analyze which doctor has treated the most patients. Does a higher patient load correlate with lower recovery ratings?
12. **Treatment Effectiveness**
    - Compare the length of stay against the recovery rating for different treatment types. Which treatment shows the most effectiveness?
13. **Cost Analysis by Hospital**
    - Investigate the average treatment cost per patient in each hospital. Which hospital has the highest and lowest costs?
14. **Patient Admission Trends Over Time**
    - Analyze the trends in patient admissions over time. Are there any seasonal patterns?
15. **Correlation Between Age and Recovery**
    - Explore if there's a correlation between patient age and recovery rating.
16. **Analyzing Room Efficiency**
    - Create a measure to analyze the efficiency of room usage (number of patients per room versus average length of stay).
17. **Impact of Doctor on Recovery**
    - Analyze if the treating doctor has a significant impact on the patient's recovery rating.
18. **Advanced DAX: Length of Stay and Cost Correlation**
    - Using DAX, explore the correlation between the length of stay and total treatment cost.
19. **Recovery Trends by Gender and Age Group**
    - Analyze recovery trends across different genders and age groups.
20. **Hospital Performance Analysis**
    - Create a comprehensive performance analysis for each hospital, considering factors like patient load, recovery ratings, and treatment costs.
21. **Extracting Key Information**
    - Using the 'Full Prescription Details' column in the updated Healthcare Dataset 1, create a new column that lists only the names of the medications and their dosages for each patient. Exclude other details like frequency and duration from this column."
22. **Complex DAX: Predictive Modeling for Recovery Rating**

- ○ Use DAX to create a predictive model estimating the recovery rating based on factors like age, diagnosis, and treatment type.
23. **Data Modeling: Cohort Analysis Based on Admission Date**
    - ○ Perform a cohort analysis of patients based on their admission date. Analyze the recovery ratings and treatment costs for different cohorts.
24. **Advanced Data Transformation: Predicting Future Hospital Capacity Needs**
    - ○ Using historical admission trends, create a model in Power BI to predict future hospital capacity needs. How does this vary by hospital?

Part 2: Dashboard Building

1. **Comprehensive Healthcare Dashboard**
   - ○ Create a dashboard in Power BI showcasing key metrics like patient demographics, diagnosis distribution, treatment costs, and recovery ratings. Include filters for hospitals, diagnosis, and treatment types.
2. **Design and Usability**
   - ○ Focus on the dashboard's design and usability. Ensure it is intuitive, informative, and visually engaging.
3. **Time-Based Analysis in Dashboard**
   - ○ Incorporate time-based analysis in your dashboard, such as trends in patient admissions and length of stay over time.
4. **Interactive Hospital Performance Comparison**
   - ○ Create an interactive section comparing different hospitals' performance based on patient load, treatment costs, and recovery ratings.
5. **Treatment Effectiveness Visualization**
   - ○ Implement visualizations that show the effectiveness of different treatments based on length of stay and recovery ratings.
6. **Key Insights and Data Storytelling**
   - ○ Provide a section summarizing key insights or stories from the data. Highlight any significant findings or trends that emerged from your analysis.