

Analysis Report On Hospital Administration Data

Problems and background

Addressing high readmission rates is crucial for enhancing patient care quality, optimising hospital resource allocation, and complying with healthcare regulations. Identifying the underlying causes can lead to targeted interventions, improved patient outcomes, and potentially significant cost savings.

Scenario: You are a data analyst in a large hospital's administration department in Novartis, focusing on improving patient care while reducing unnecessary readmissions. Recently, the hospital has seen a spike in readmission rates, leading to increased healthcare costs and indicating potential quality of care issues. The administration is concerned about the impact on patient outcomes and the hospital's reputation. Your task is to analyse hospital admission data to identify patterns and factors associated with high readmission rates.

Project scope

This project aims to analyse hospital admission data to identify factors influencing readmission rates. The dataset includes patient demographics, clinical data, visit history, diagnosis codes, medication indicators, and readmission status. We will examine the distribution of readmission rates across age groups, the average length of hospital stays by medical specialty, and the correlation between emergency visits and readmission rates. Additionally, we will assess the proportion of readmissions in diabetic patients, the impact of changes in diabetic medication, the relationship between lab procedures and readmission rates, and variations in readmission rates by race and gender. We will also explore the distribution of patients by weight, the impact of medication quantity on hospital stay length, and the frequency of outpatient visits and their effect on readmission rates. The insights gained will help improve patient care, reduce readmissions, and optimise hospital operations.

Methodology

- **Data Sources:** The data used in this project was provided by our team and sourced from Kaggle.
- **Data Wrangling:** Data cleaning was performed in Python to handle missing values, normalise formats, and ensure consistency across all variables.
- **Data Analysis:** The analysis was conducted using both Python and SQL Server to calculate average revenue by industry, assess industry representation, identify geographic trends, and examine correlations between emergency visits and readmission rates.
- **Data Visualization:** Data was visualised using PowerBI and Python. Various charts and graphs were created to effectively communicate findings on revenue distribution, industry growth rates, geographic trends, and other key insights.

Goals and KPIs (3 - 4 Max)

KPIs

- The percentage of patients readmitted to the hospital within 30 days of discharge.
- The average number of days patients spend in the hospital during a single visit.
- The correlation between the number of emergency visits in the previous year and the likelihood of readmission.

Goals

- Reduce the readmission rate by 15% within the next 12 months.
- Decrease the ALOS by 10% within the next year.
- Implement targeted intervention programs for high-risk patients to reduce the correlation coefficient by 20% within 18 months.

Technical Processes

- Tools used: Python, PowerBi, SQL Server
- Clauses used in SQL: SELECT, GROUP BY, FROM, ORDER BY, SQRT, TOP, Covariance. LIMIT, AVG, COUNT, FETCH.
- Python library: Numpy, Pandas. Used Matplotlib, Plotly and Seaborn for visualisation in Python
- PowerBi: Used for visualisation and making reports.

Business Concepts Used

Market Understanding

Analyse patient demographics and behaviours to understand the hospital's target market. This involves examining patient age, race, gender, and geographic distribution to identify trends and needs within different population segments.

Customer Demographic

Study demographic information such as race, gender, age, and weight to understand the composition of the patient base. This helps tailor services to meet the specific needs of various groups and ensure inclusive healthcare.

Customer Behaviour

Evaluate patient behaviours, including the number of visits (outpatient, emergency, inpatient) and length of hospital stay, to identify patterns. Understanding these behaviours can help in optimising patient care and resource allocation.

Customer Retention

Assess factors influencing readmission rates, such as medical specialty, number of procedures, and medication changes. This analysis can guide strategies to improve patient retention through better follow-up care and targeted interventions.

New Customer Acquisition

Identify trends in patient admissions by specialty and geographic trends to target new patient groups effectively. Focusing on specialties with lower readmission rates and understanding the impact of demographic factors can aid in acquiring new customers.

Recommended Analysis:

1. What is the distribution of readmission rates across different age groups?

- The distribution of readmission rates across different age groups shows varied patterns:
- Readmissions are relatively low in younger age brackets ([0-10) to [30-40)).
- There is a steady increase from middle age ([40-50)) onwards, peaking notably in the 70-80 age group.
- The highest absolute count of readmissions occurs among individuals aged 70-80 years.
- Readmission rates then slightly decrease in the oldest age groups ([80-90) and [90-100)).

Understanding these age-related trends is crucial for optimising healthcare interventions, particularly focusing resources where the burden of readmissions is highest, such as in the elderly population.

2. Analyse the average length of hospital stays by medical specialty. Which specialties have the longest and shortest average stays?

→ Analysis of Average Length of Hospital Stays by Medical Specialty

Upon analysing the average length of hospital stays by medical specialty, it is evident that there is considerable variation across different fields. The specialties with the longest average stays include:

- Paediatrics-Pulmonology: 10.7 days
- Physical Medicine and Rehabilitation: 8.9 days
- Outreach Services: 8.2 days
- Pathology: 8.1 days
- Psychiatry-Addictive: 8.0 days

These specialties likely involve complex and prolonged treatments, necessitating extended hospital stays.

Conversely, the specialties with the shortest average stays are:

- Pediatrics-Emergency Medicine: 1.0 day
- Anesthesiology-Pediatric: 1.7 days
- Pediatrics-Hematology-Oncology: 2.0 days
- Speech: 2.0 days
- DCP TEAM: 2.0 days

These shorter stays are indicative of treatments or procedures that are either less intensive or designed for rapid resolution, allowing patients to be discharged quickly.

Q.3 How does the number of emergency visits in the previous year correlate with readmission rates?

- The data indicates a correlation between the number of emergency visits in the previous year and readmission rates. Patients with no emergency visits have a lower readmission rate, with 33,064 not readmitted (0) and 26,023 readmitted (1). As the number of emergency visits increases, the likelihood of readmission also rises. For instance, with one emergency visit, the number of readmissions increases to 3,004 compared to 2,113 not readmitted. This trend continues, with patients having multiple emergency visits showing progressively higher readmission rates. For example, at three emergency visits, 336 were readmitted compared to 136 not readmitted. This pattern suggests that frequent emergency visits are associated with a higher probability of hospital readmission.

Q.4 What proportion of patients with diabetes were readmitted within 30 days?

- Among patients with diabetes, a significant majority were readmitted within 30 days. Specifically, 79.8% of patients who were on diabetes medication experienced readmission. In contrast, only 20.2% of patients without diabetes medication were readmitted. This indicates a higher propensity for readmission among diabetes patients, highlighting the need for targeted interventions to reduce readmission rates in this group.

Q.5 Is there a significant difference in readmission rates between patients with changes in diabetic medication and those without?

- There is a notable difference in readmission rates between patients who had changes in their diabetic medication and those who did not. Among the patients, 53.9% who experienced changes in their diabetic medication were readmitted within 30 days. Conversely, 46.1% of patients without any changes in their diabetic medication were readmitted. This suggests that changes in diabetic medication are associated with a higher readmission rate, indicating the potential need for closer monitoring and management of patients undergoing such changes.

Q.6 Investigate the relationship between the number of lab procedures performed during the stay and readmission rates.

- The relationship between the number of lab procedures performed during a hospital stay and readmission rates shows a distinct pattern. Patients with a moderate number of lab procedures (31-50) have the highest readmission rates, with 12,380 readmissions for 31-40 procedures and 14,435 for 41-50 procedures.
- As the number of lab procedures increases beyond 50, the readmission rates start to decline. For instance, there are 11,409 readmissions for 51-60 procedures, dropping to 7,530 for 61-70 procedures, and further declining as the number of procedures increases.
- This trend suggests that while a higher number of lab procedures is initially associated with higher readmission rates, extremely high numbers of procedures correspond to a lower frequency of readmissions. This may indicate that patients undergoing more extensive testing either receive more comprehensive care, reducing the need for readmission, or that only the more severe cases, which require frequent monitoring, are being identified and managed more effectively.

Q.7 How do readmission rates vary by race and gender?

→ **Gender Differences:**

- ◆ Across all races, females generally have higher readmission rates compared to males, with Caucasian females showing the highest number of readmissions (26,198) followed by African American females (7,565).

Race Differences:

- ◆ **Caucasians** have the highest readmission rates overall for both genders.
- ◆ **African Americans** have the second highest readmission rates.
- ◆ **Asians** have the lowest readmission rates among the racial categories listed.
- ◆ **Hispanic** and **Other** categories have moderate readmission rates, with females slightly higher than males.

Unknown/Invalid Gender:

- ◆ Very few cases are recorded under the "Unknown/Invalid" gender, mostly within the Native American and Other categories, showing minimal impact on overall trends.

These variations indicate the need for targeted interventions to address the specific healthcare needs and challenges faced by different racial and gender groups to reduce readmission rates.

Q.8 What is the distribution of patients across different weight categories, and how does this relate to readmission rates?

- NA (Not Available): The majority of patients (64,454) have weight data unavailable, limiting analysis of readmission rates.
- [75-100) kg: Highest number of patients (900) with a notable readmission rate, indicating potential correlation between weight and readmission.
- [50-75) kg: 603 patients, showing significant readmissions, suggesting weight may impact healthcare outcomes.
- [100-125) kg: 416 patients, moderate readmission rate, indicative of weight-related health complexities.
- Others (0-50 kg and >125 kg): Lower patient counts and varied readmission rates, suggesting less frequent readmissions but potential for specialised care needs.

Insights:

- Limited data availability (NA category) highlights the need for improved weight data recording.
- Higher weight categories generally correlate with increased readmission rates, suggesting potential health impacts requiring closer monitoring and care management.

Q.9 Evaluate the impact of the number of medications prescribed on the length of hospital stay.

- Correlation between Number Of Medications and Time in Hospital: 0.94

Conclusion

Based on the comprehensive analysis of hospital administration data, several key insights have emerged. Significant correlations were found between higher numbers of emergency visits, changes in diabetic medication, and increased readmission rates, suggesting the need for targeted interventions in these areas. Additionally, the distribution of patients across weight categories revealed varying readmission rates, emphasising the potential impact of weight on healthcare outcomes. These findings underscore the importance of data-driven healthcare strategies to optimise patient care, reduce readmissions, and improve overall hospital management efficiency.

Project owner

Name: Ritika Pathak

Email id: ritupat707@gmail.com

Peerlist: https://peerlist.io/ritika_p

Kaggle: <https://www.kaggle.com/ritika01mp>

Discord: [r.k.0.7](#)