EffiCare Healthcare Efficiency & Cost Optimization

By Maithily Manohar

Final Project Report

Use the pages below to navigate the report by section.

Project overview

The EffiCare Healthcare Analytics project focuses on improving hospital efficiency and financial performance through data-driven analysis. The objective of this project is to study inpatient discharge data, identify operational inefficiencies, understand cost behavior, and support hospital decision-making using analytical dashboards and predictive models.

Over the course of the internship, the project progressed through multiple stages including data preprocessing, exploratory data analysis, feature engineering, interactive dashboard development, and machine learning model implementation. The final outcome provides a consolidated analytical framework that enables hospitals to monitor performance, control costs, and enhance patient care outcomes.

Dataset description

The dataset consists of hospital inpatient discharge records containing demographic, clinical, and financial information. Key variables include age group, gender, severity of illness, length of stay, total charges, total costs, risk of mortality, patient disposition, and hospital identifiers.

This dataset enables a combined analysis of operational metrics such as length of stay and clinical severity, as well as financial indicators such as costs, charges, and reimbursement behavior. These attributes together provide a comprehensive view of hospital performance and patient outcomes.

Data preprocessing and feature engineering

Data preprocessing was performed to ensure analytical reliability and model readiness. Duplicate records were removed and missing values were handled appropriately. Outliers in key numerical variables such as total charges, total costs, and length of stay were treated using the Interquartile Range (IQR) method to reduce skewness and improve stability.

Categorical variables were encoded to make them suitable for analysis and modeling. In addition, new analytical features were engineered to enhance insight generation. These included a charge-to-cost ratio to understand billing efficiency, an efficiency index to represent operational performance, and a profit margin to evaluate financial outcomes. These engineered features strengthened both the dashboard insights and the predictive modeling phase.

Dashboard insights

Key findings from the dashboard workstream, split by dashboard.

- [Dashboard 1: Hospital efficiency and cost analysis](#)
- [Dashboard 2: Patient demographics and outcome trends](#)
- [Dashboard 3: Payment type and cost optimization](#)
- [Dashboard 4: Insurance and billing optimization](#)
- [Dashboard 5: Cross-dashboard financial and operational KPIs](#)

Dashboard 1: Hospital efficiency and cost analysis

This dashboard examined hospital-level performance using key operational and financial metrics such as average length of stay, average total cost, average charges, discharge volume, and efficiency indicators.

The analysis revealed that hospitals with shorter average length of stay demonstrated higher operational efficiency. Cost variations across hospitals were strongly influenced by patient age and illness severity. Facilities with high discharge volumes showed better resource utilization, while hospitals with prolonged stays experienced increased operational burden and cost accumulation.

Attachements 👍

	Hip_Replacement_EffiCare_DataAnalysis.ipynb	Download	Open
383KB			

|	Week one Task.pdf	Download	Open
652KB	PDF		

Dashboard 2: Patient demographics and outcome trends

This dashboard focused on understanding patient characteristics and their impact on outcomes. It analyzed mortality risk by age group, length of stay by age, cost distribution by gender and race, and patient disposition categories.

The results indicated that mortality risk increases significantly with age and severity of illness. Severe cases required longer hospitalization and higher treatment costs. The majority of patients were discharged home, suggesting effective treatment pathways, while transfers and mortality cases were concentrated in higher severity categories.
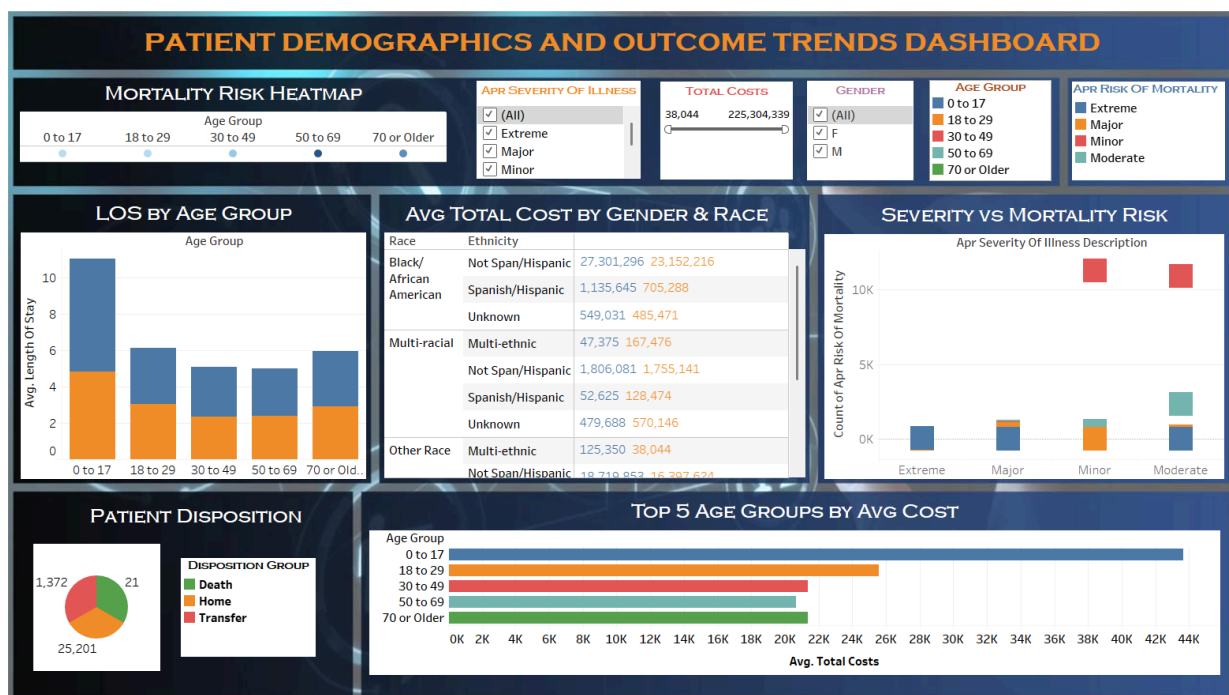
Dashboard 2
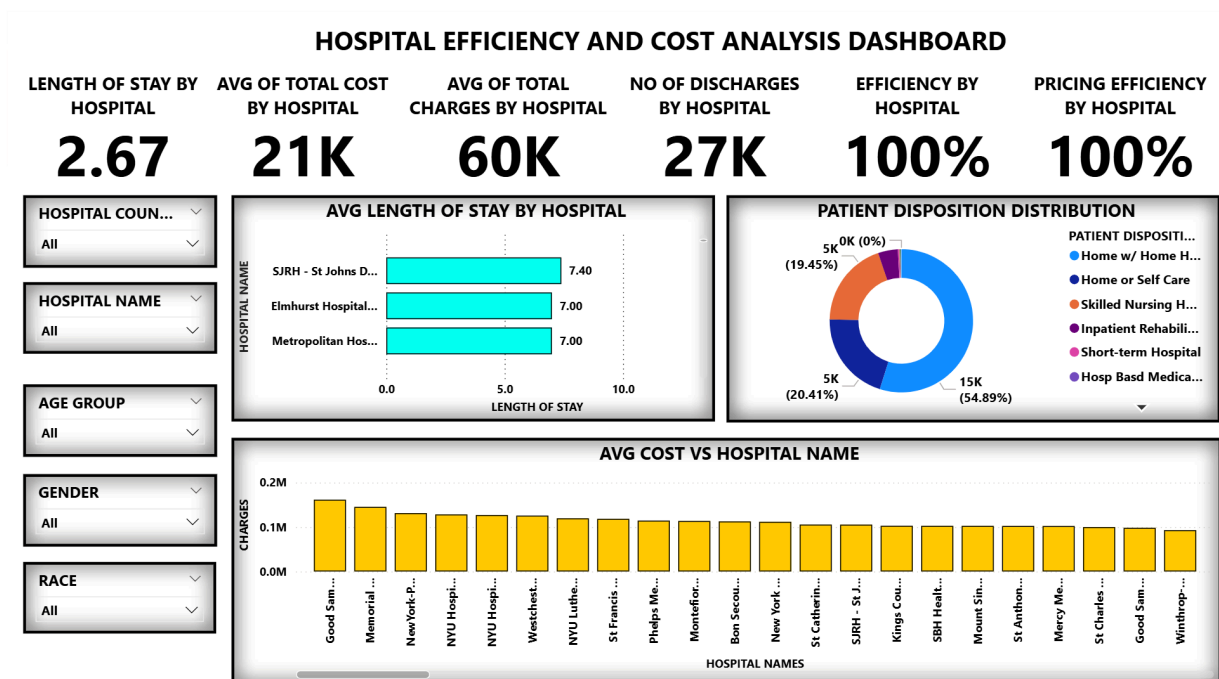
Dashboard 3: Payment type and cost optimization

This dashboard explored the relationship between payment type and hospital cost behavior. It analyzed average cost by payment type, length of stay by payment category, payment distribution, and severity versus cost patterns.

The findings showed that certain payment types were associated with higher treatment costs and longer length of stay. Elective admissions were observed to be more cost-efficient compared to emergency cases. Severity of illness had a direct impact on cost irrespective of payment method, highlighting the importance of case-mix management in financial planning.
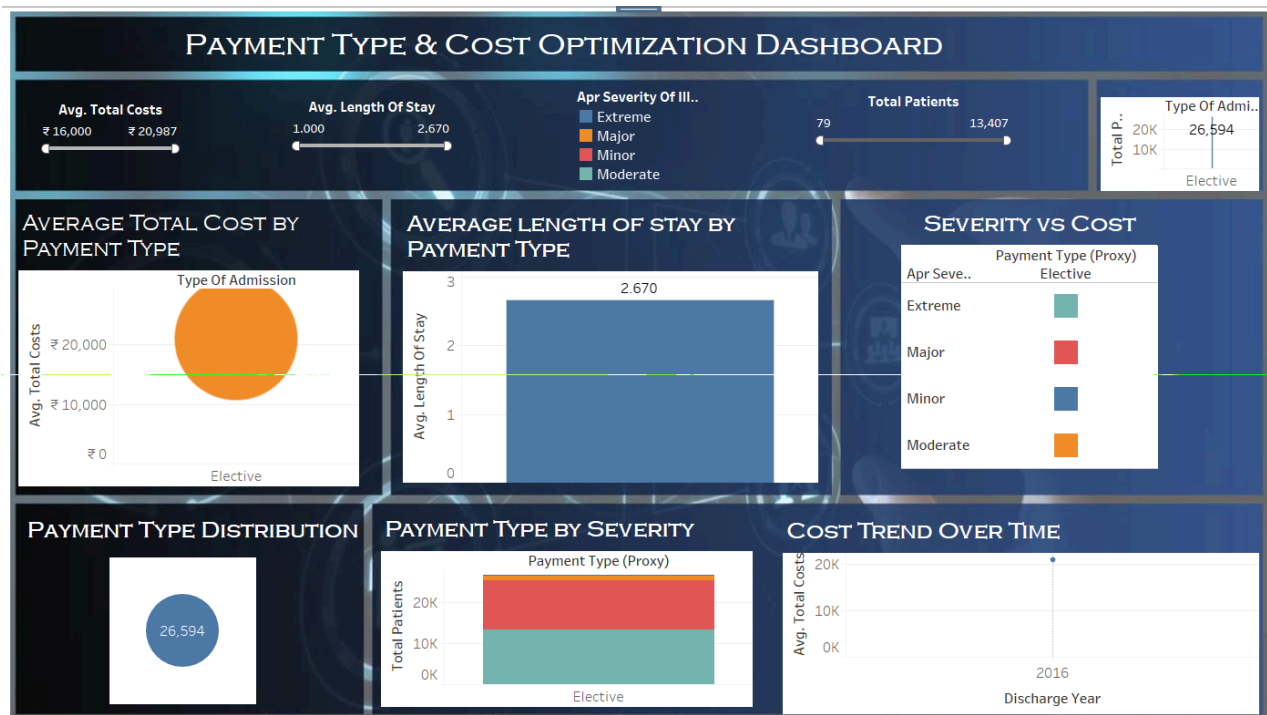
Dashboard 3

Dashboard 4: Insurance and billing optimization

This dashboard evaluated financial efficiency by examining revenue leakage, insurance reimbursement rates, profit, and revenue distribution across payment categories.

The analysis identified significant revenue leakage, particularly in moderate and major severity cases. Reimbursement rates directly influenced hospital profitability. Inefficient billing and delayed reimbursements were key contributors to financial loss, emphasizing the need for optimized insurance claim processing and billing workflows.

Dashboard 4

Dashboard 5: Cross-dashboard financial and operational KPIs

This dashboard consolidated insights from all dashboards to present key performance indicators such as total revenue, profit, reimbursement trends, and cost-to-charge relationships.

Hospitals demonstrating strong reimbursement performance and cost control achieved higher profit margins. Operational efficiency, particularly reduced length of stay, played a critical role in improving financial outcomes and sustaining hospital performance.

Dashboard 5
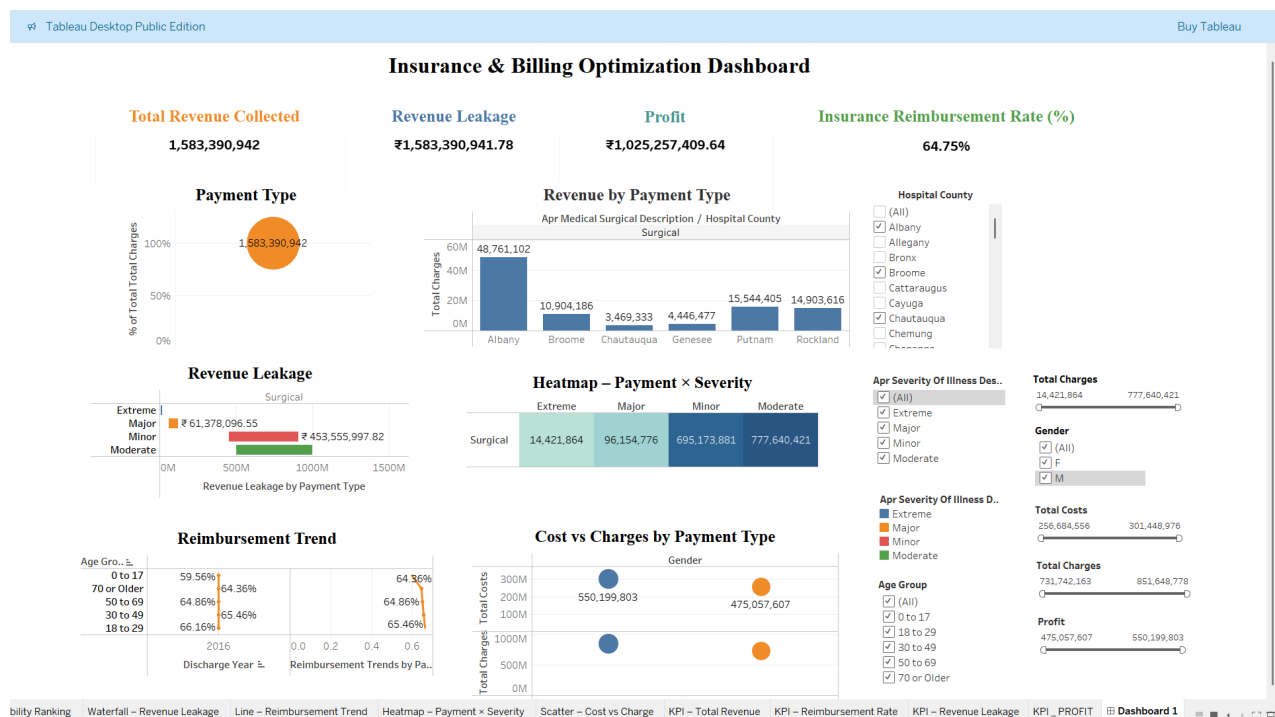
Machine learning models

Three major predictive objectives were addressed using machine learning.

For length-of-stay prediction, Linear Regression and Random Forest Regressor models were implemented to estimate expected hospitalization duration. These predictions support bed allocation and workforce planning.

For cost prediction, Linear Regression and Random Forest Regressor models were used to forecast treatment expenses based on patient and clinical attributes. This enables hospitals to plan budgets and manage financial risk more effectively.

For readmission-risk prediction, Logistic Regression and Random Forest Classifier models were applied to identify patients with a high probability of readmission. These predictions help hospitals implement targeted post-discharge interventions and improve long-term patient outcomes.

Business impact

This project provides hospitals with actionable intelligence to support operational and financial decision-making. It enables identification of high-cost patient segments, optimization of length of stay, improvement of insurance reimbursement efficiency, and better allocation of clinical resources.

By combining dashboards with predictive modeling, hospital administrators can proactively manage patient flow, reduce unnecessary costs, improve billing performance, and enhance overall quality of care.

Limitations and future scope

Limitations

The analysis is limited to the scope of the available dataset and does not include real-time operational data. External socio-economic and lifestyle factors were not incorporated. Additionally, regional constraints of the dataset may limit generalization across different healthcare systems.

Future scope

Future work can integrate real-time hospital data streams and expand the model using deep learning techniques. Additional clinical attributes such as lab results and treatment protocols can further improve prediction accuracy. The dashboards can be deployed as live monitoring systems for hospital management teams.

Conclusion

The EffiCare Healthcare Analytics project demonstrates how structured data analysis and machine learning can improve hospital efficiency and financial sustainability. The integrated approach of dashboards and predictive models offers a comprehensive decision-support system for healthcare administrators. This project successfully bridges analytical insights with practical business value for hospital operations.