Typhoid Treatment Insight: Data Analysis

KLE TECHNOLOGICAL UNIVERSITY
Creating Value, Leveraging Knowledge
Belagavi Campus
Dr.M.S.Sheshgiri College of Engineering & Technology
Department of Computer Science & Engineering

BACKGROUND

This EDA project analyzes treatment outcomes in typhoid patients using clinical and demographic data to identify factors influencing success. It explores correlations between symptoms severity, medication efficacy, and treatment duration to optimize therapies. The study aims to improve diagnostic accuracy, personalize treatment plans, and mitigate typhoid-related liver complications.

PROBLEM STATEMENT

The dataset encompasses patient demographics, symptoms severity, laboratory results, bacterial cultures, medications, treatment duration, and outcomes (success/unsuccessful). This analysis aims to uncover factors influencing treatment success through correlation exploration, feature engineering, and predictive modeling, aiming to optimize treatment strategies and improve patient outcomes in related medical scenarios.

DATA SET:

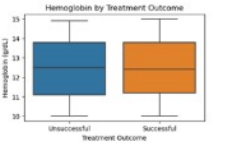
Source: Downloaded from Kaggle
Size: 554 KB
Rows: 5760
Attributes: 13

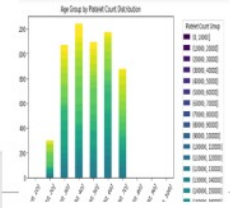
Features of the Dataset

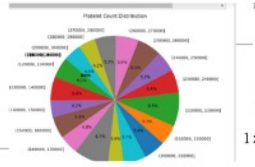
1.Patient ID: Categorical, Unique, Identifier
2.Age: Numerical, Continuous
3.Gender: Categorical (Male, Female)
4.Symptoms Severity: Categorical (Low, Moderate, High)
5.Hemoglobin (g/dL): Numerical, Continuous
6.Platelet Count: Numerical, Continuous
7.Blood Culture Bacteria: Categorical, Nominal
8.Urine Culture Bacteria: Categorical, Nominal
9.Calcium (mg/dL): Numerical, Continuous
10.Potassium (mmol/L): Numerical, Continuous
11.Current Medication: Categorical, Nominal
12.Treatment Duration: Numerical, Continuous
13.Treatment Outcome: Categorical (Successful, Unsuccessful)Feature

Data Pre-processing:

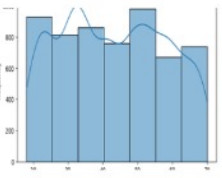
It involves cleaning, normalizing, and encoding data for analysis to ensure accuracy. Data Distribution Plots visualize patterns, outliers, and skewness. Missing Values Imputation fills missing data with estimated values. Outliers Handling identifies and addresses extreme values. Feature Engineering enhances model performance. Correlation Analysis assesses variable relationships. Feature Selection selects the most relevant variables for improved model accuracy and interpretability.





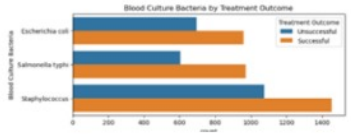


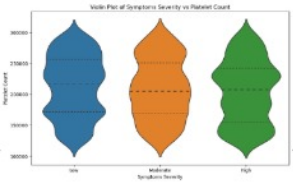
Uni-variate Hypothesis



Bi-variate Hypothesis

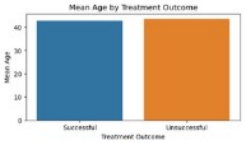
3: Bacterial infections and treatment success rates

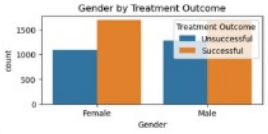




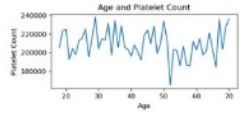
HYPOTHESIS

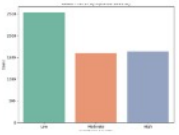
1: Demographic factors (age, gender) correlate with treatment outcomes

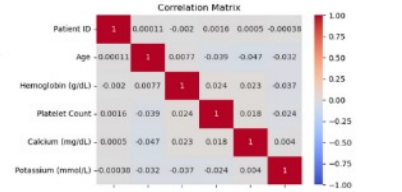




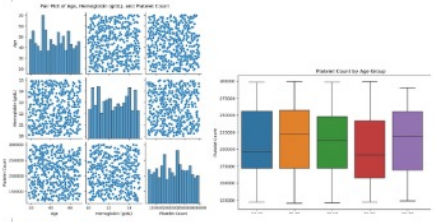
"2.Age affects platelet count."







Multi-variate Hypothesis



Conclusion

By meticulously processing and refining the dataset through data cleaning, visualization, and feature engineering, this project identified crucial factors influencing treatment outcomes in typhoid patients. Correlation analysis guided the selection of predictive features, optimizing model accuracy and relevance. These findings underscore the importance of early intervention and tailored treatment strategies in mitigating typhoid-related complications and improving overall patient care.

Guided By: **PROF. PRIYANKA .G.**

Team Members:
1.Gayatri 02FE22BCS037
2 Saqeeb 02FE23BCS417
3.Rachan 02FE23BCS415
4.Yashwanth 02FE23BCS432