CONSULTANCY REPORT
ON
HOSPITAL DURATION ANALYSIS FOR BACK SURGERY PATIENTS WITHIN NHS TRUST

PREPARED FOR:

CHIEF EXECUTIVE OFFICER
NHS TRUST

PREPARED BY:

SUSHANT THAPA MAGAR
ADVANTAGE ANALYTICS GROUP
27TH MARCH 2024

1. INTRODUCTION

The National Health Service (NHS), as one of the largest healthcare providers globally, encounters significant operational hurdles due to budget constraints, resource shortages, and the ever-increasing demand for quality patient care. Considering these challenges, predictive analytics and forecasting play a pivotal role in optimizing resource allocation, reducing costs, and enhancing patient outcomes within the NHS.

Accurate prediction of hospital stay durations is imperative for efficient resource management, cost reduction, and enhancing patient care and satisfaction. Recognizing the critical nature of these challenges, the NHS Trust's Chief Executive Officer has commissioned the Advantage Analytics Group for a consultancy assignment to leverage data analysis and predictive modeling techniques to address these issues effectively.

1.1 Context and Objectives

The Advantage Analytics Group has been tasked by the NHS Trust's Chief Executive to conduct a pilot project involving data analysis and predictive modeling to analyse the hospital stay duration of back surgery patients to achieve the following objectives:

- Optimal allocation of NHS resources, reducing waiting times and improving patient flow amidst budget constraints and resource shortages.
- Enabling targeted interventions to reduce unnecessary hospitalization days and manage expenditures effectively, thereby controlling costs.
- Facilitating improvements in patient care by revising treatments or practices to reduce recovery times and enhance care quality, leading to better patient outcomes and satisfaction rates.
- Accurately aiding in strategic planning and policy-making, informing decisions about service expansion, staffing, and infrastructure development to meet future demands.
- Ensuring better preparedness and response without compromising care quality during pandemics or mass casualty events.

2. DATA PROCESSING DECISIONS

The dataset for analysis comprises Patient Information, Surgical Information, and ICD codes. The decisions essential for ensuring the data is structured, standardized, and ready for in-depth analysis and modeling are as follows:

- **Calculation of Time Intervals:** Hours till Ambulation, Hours till release from Ambulation, Total Recovery Time and Surgery Preparation Time were computed to identify correlations and patterns that could strengthen predictive models.

- **Patient Age Calculation and Grouping:** Patient age was derived from the admission date and date of birth to ensure consistency and relevance. Also, to enhance data analysis patient age data was organized into categories for better interpretation.

- **Surgery Type:** ICD code file parsing extracted surgery-type data. Other relevant fields like Surgery Area/Process and Approach were also parsed to enrich the dataset for predictive modeling.

- **Temporal Analysis:** Data preparation for Temporal analysis was conducted to monitor admission and discharge patterns, aiding in hospital performance evaluation, and identifying areas for improvement. Inconsistent admission and discharge rates could indicate inefficiencies necessitating optimization efforts.

- **Data Blending**: Patient Information and Surgical Information were merged initially, followed by merging with ICD-10 codes data after necessary preparation, aligning with forecast sheet requirements.

3. EXPLORATORY DATA ANALYSIS

Exploratory Data Analysis (EDA) is a critical preliminary step in understanding the characteristics and insights inherent in a dataset. The following analyses have been conducted to serve the objectives of the project:

A. Patient Demography Analysis

a. Gender based analysis:

In analyzing the distribution of hospital stay lengths based on gender demographics, non-standard gender categories (ASKU and OTH) were filtered out to ensure a focus on more representative data.

Inference:

- Males and Females constitute the majority of patients, accounting for 49.80% and 47.15%, respectively.
- Significant variations in the average length of hospital stays across different gender categories. Notably, Bigender and Other (OTH) exhibit longer average stays (129.78 hours and 94.31 hours, respectively).

Recommendations:

- Tailor patient care plans and resource allocation strategies to accommodate variations in hospital stay lengths across different gender demographics, promoting equitable healthcare.
- Adopt inclusive data collection to accurately capture gender diversity, ensuring tailored services and equitable care.

b. Age based analysis:

Performed an analysis on hospital stay lengths, total recovery time, hours till ambulation, and hours till release from ambulation across various age groups. Patients were categorized into four age brackets (0-20, 21-40, 41-60, and 61-80) for clarity and ease of interpretation.

Inference:

- Older patients, particularly those aged 41-80, tend to experience slightly longer recovery times and hospital stays compared to younger patients (0-40 age groups). While individual differences in recovery times may appear minor, their cumulative impact can be substantial.
- However, the association analysis demonstrates a correlation coefficient of 0.02 between age and length of hospital stay suggesting a very weak linear relationship between the two variables.

B. Surgery Type/Medical Condition Analysis

The Plot of Means has been used to find the length of stay at the hospital based on the medical condition data available.

Inference:

- The type of surgery and approach utilized significantly impact the length of hospital stay and total recovery time. Surgeries involving replacements and extirpations, tend to result in longer hospital stays compared to procedures like insertions and removals.
- Percutaneous Endoscopic Approach leads to the highest length of stays, while Percutaneous Approach results in the lowest.

Recommendations:

- NHS could prioritize the adoption of less invasive techniques, such as removal or insertion procedures, whenever clinically appropriate. Ultimately, the recommendation to prioritize less invasive procedures should be applied judiciously, considering the individual needs and circumstances of each patient

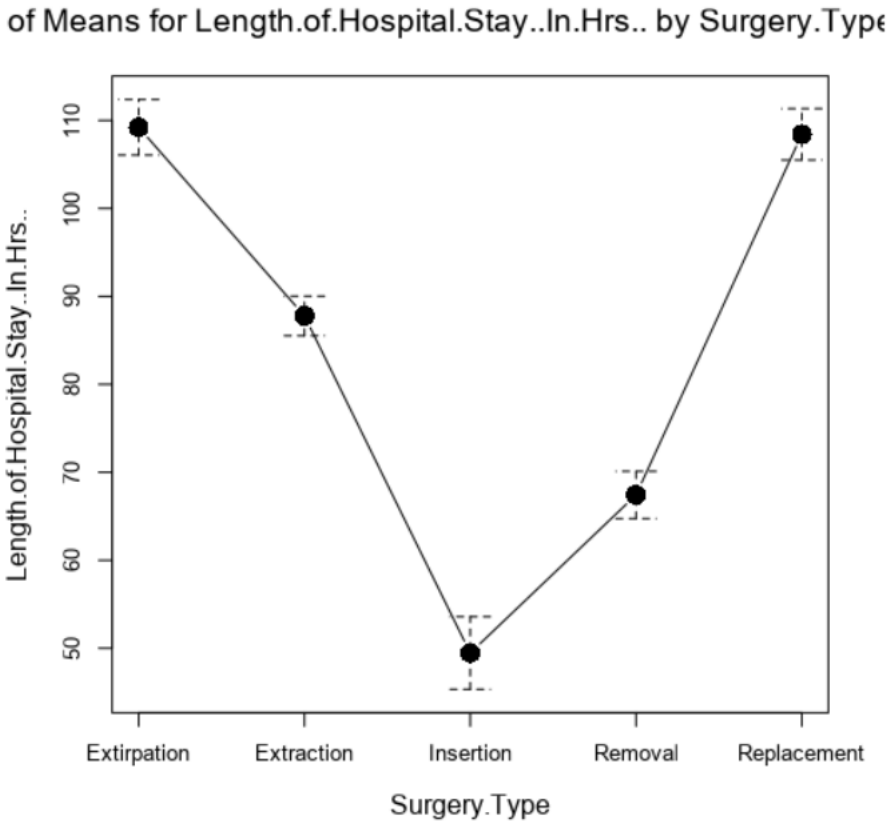

Figure 1 Hospital Stay based on Surgery Type

C. Temporal Pattern Analysis

By understanding hospital stay patterns, NHS can anticipate patient volume fluctuations, align staffing and resources accordingly, and ensure optimal service delivery.

Inference:

- Patients admitted on Saturdays and Thursdays tend to have longer stays, while those admitted on Wednesdays have shorter stays.
- Patients discharged on Sundays, Fridays, and Saturdays have prolonged stays.
- The counts indicate August as a month with consistently high patient volumes for both admission (194) and discharge (193), suggesting significant healthcare activity during this period within the NHS.

Recommendations:

- Increase staffing levels on days with historically higher admission rates, such as Saturdays and Thursdays, to ensure timely and efficient patient care.
- Implement strategies to expedite discharges, particularly on weekends (Sundays, Fridays, and Saturdays), when patients tend to have prolonged stays.
- Prepare for Seasonal Peaks

  While the current recommendations offer valuable insights based on the available one-year dataset, it's essential to view them as preliminary guidelines that may evolve with continued data collection and analysis.

D. Ambulation Timing Impact Analysis

Categorized ambulation within 24 hours as early and beyond as late. The categorization can be altered at NHS's discretion.

Inference:

- Patients with late ambulation exhibit notably longer average lengths of hospital stay and recovery times compared to those with early ambulation.

Recommendations:

- Promote early ambulation post-surgery.

E. Co-relation Analysis

Conducted correlation analysis crucial to identify potential relationships between different variables and the length of hospital stays.

Inference:

- Strong positive correlations were found between the length of hospital stay and both the time until ambulation (0.79) and the time until release from ambulation (0.88), indicating that delayed ambulation can prolong hospital stays.
- A moderate positive correlation (0.41) was observed between pre-surgery preparation time and hospital stay length, suggesting that longer preparation times may contribute to extended stays.

Recommendations:

- To reduce hospital stays, optimize ambulation protocols and post-operative care, focusing on minimizing delays in ambulation and streamlining pre-surgery preparation processes.

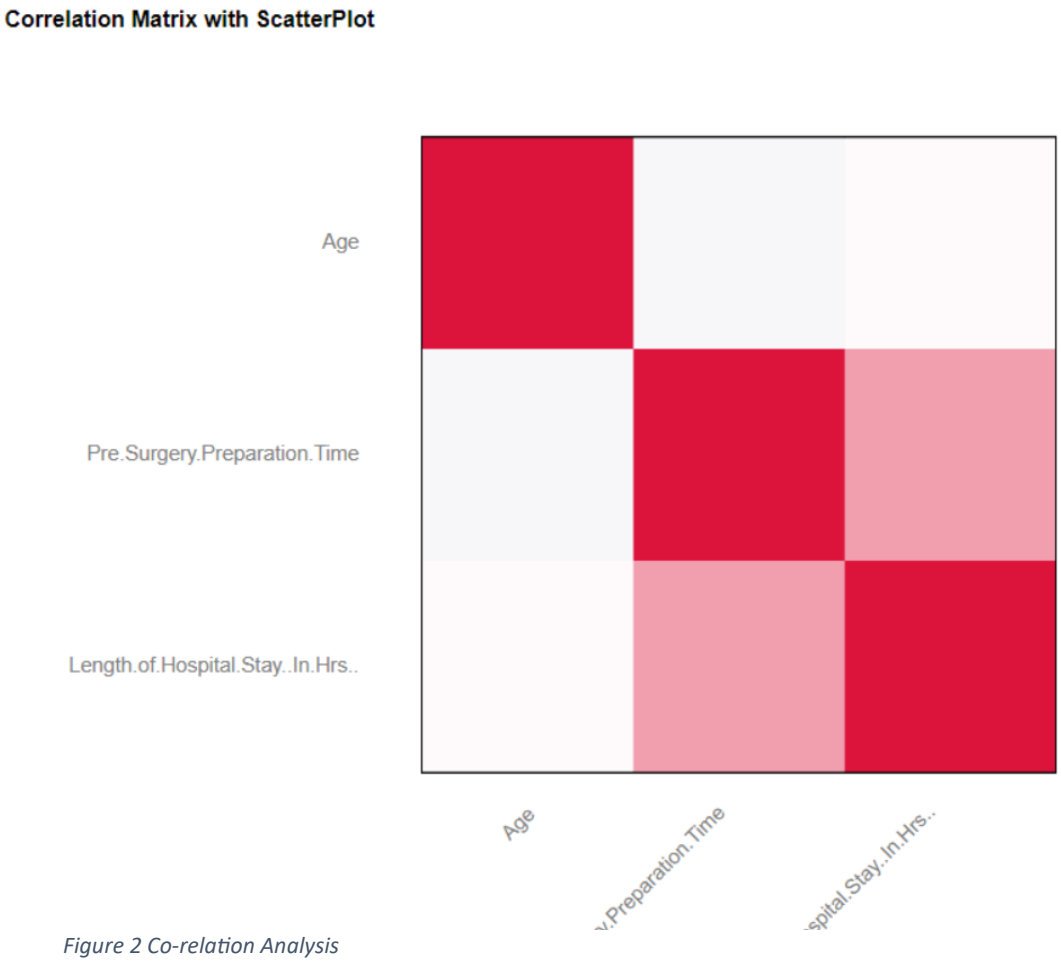

Figure 2 Co-relation Analysis

4. METHODOLOGICAL DECISIONS AND ASSUMPTIONS

A. Use of decimals for calculation of hours

To enhance accuracy, remaining time was represented with decimals instead of rounding off for time calculations aiding data granularity, crucial for precise analysis and decision-making.

B. Calculation of Patients Age

Patient's age was calculated using admission date and not the date of analysis to ensures relevance to the patient's healthcare situation at admission. It also promotes uniformity in data collection and analysis.

C. Choice of Predictive Models

The choice of Linear Regression, Decision Tree, and Random Forest models for predictive analytics is justified based on their distinct strengths and suitability for the given dataset. Linear Regression offers simplicity and interpretability, making it suitable for scenarios where the relationship between variables is linear or approximately linear. Decision Trees capture non-linear relationships and provide insights into feature importance, while Random Forest combines multiple trees to improve accuracy and handle complex datasets effectively.

D. Selection of a Predictive Model

Linear Regression demonstrated the highest correlation coefficient of 0.8903, indicating a strong linear relationship between the predictor variables and the target variable. It also achieved the lowest Root Mean Squared Error (RMSE) of 13.7452, suggesting that it has the smallest average deviation from the actual values. Additionally, Linear Regression exhibited the lowest Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), further highlighting its accuracy and precision in prediction. The Forest Model and Decision Tree, while also performing well, showed slightly lower correlation coefficients and higher error metrics compared to Linear Regression.

Model Comparison Report

Fit and error measures

Model	Correlation	RMSE	MAE	MPE	MAPE
Linear_Regression	0.8903	13.7452	10.8521	-2.3544	13.2354
Forest_Model	0.8840	14.5130	11.5379	-3.9158	14.5564
Decision_Tree	0.6855	21.8896	16.8705	-5.9988	21.3801

Model: model names in the current comparison.
Correlation: correlation between the predicted values and the actual values.
RMSE: root mean square error.
MAE: mean absolute error.
MPE: mean percentage error. Note: based on its definition, MPE may be positive or negative infinity if the target variable has 0 values. In this case, we return a weighted percentage error (WPE).
MAPE: mean absolute percentage error. Note: based on its definition, MAPE may be positive or negative infinity if the target variable has 0 values. In this case, we return a weighted absolute percentage error (WAPE).

Figure 3 Model Comparison

5. PREDICTED OUTCOMES

After conducting thorough data analysis and applying predictive modeling techniques, the predictions for hospital stay durations has been generated. The predicted outcomes are based on factors such as gender, age, surgery type, and hours till ambulation. These predictions offer valuable insights into anticipated hospital stay durations for back surgery patients, providing guidance for resource allocation, cost management, and patient care improvement initiatives within the NHS Trust.

Gender	Age	Surgery Type	Hours till Ambulation	Length of Hospital Stay (Hrs.)	Length of Hospital Stay (Days)
Female	53	Extirpation	46.00	137.14	5.71
Female	42	Extirpation	30.00	112.57	4.69
Male	73	Extirpation	29.00	110.61	4.61
Female	48	Replacement	30.00	110.71	4.61
Female	67	Replacement	28.00	107.14	4.46
Female	43	Extirpation	26.00	106.33	4.43
Female	71	Removal	33.00	90.35	3.76
Male	59	Removal	31.00	87.89	3.66
Male	48	Removal	27.00	81.94	3.41
Male	36	Extraction	13.00	75.89	3.16
Male	23	Removal	21.00	73.24	3.05
Female	25	Removal	17.00	66.62	2.78
Male	47	Insertion	24.00	62.34	2.60
Male	39	Removal	10.00	55.77	2.32
Female	57	Insertion	15.00	47.76	1.99

Figure 4 Predicted Outcome

6. SUMMARY OF INSIGHTS AND ACTIONABLE RECOMMENDATIONS

A. Data quality, Completeness and Precision

- For accuracy and precision, the use of decimals is recommended for representing remaining time instead of rounding off. However, if rounding off is necessary for presentation or stakeholders' preferences, it should be documented and communicated to ensure transparency.

- The absence of 10 patient IDs in the surgical information dataset suggests potential data quality issues. Asking why could raise questions about data quality, data completeness, reasons for patients not opting for surgery or missing surgery records, and so on.

- Further investigation into missing surgical procedure (Drainage) is necessary as it could simply mean there were no such surgical cases or incomplete records. Resource constraints within NHS facilities may also contribute to missing surgical procedures, highlighting the impact of resource availability on healthcare service provision.

- The presence of categories like OTH and ASKU raises concerns about data integrity. Implement standardized data collection protocols and coding systems that capture diverse gender identities accurately and inclusively.

B. Healthcare Policies and Practices

- Develop and implement gender-sensitive healthcare policies and practices that recognize and address the unique healthcare needs of gender-diverse individuals. Gender identity may have clinical relevance in healthcare, influencing factors such as treatment response, healthcare utilization patterns, and health outcomes.

- Prioritize less invasive procedures when clinically appropriate to potentially reduce patient recovery times, optimize resource utilization, and enhance patient outcomes. However, this approach should be applied judiciously, considering each patient's individual needs to ensure optimal care while minimizing hospitalization and recovery burdens.

C. Resource Allocation and Preparedness

- Adjust staffing and resources based on patient admission fluctuations, prioritize discharge planning to expedite patient release, and collect multi-year data for accurate forecasting.

- Prioritizing early ambulation enhances patient recovery, optimizes resource use, and potentially reduces costs. Implementing standardized protocols, staff education, incentives for compliance, patient monitoring, technology integration, and quality improvement initiatives can promote early ambulation and improve recovery outcomes.

The recommendations offer a comprehensive approach to enhancing healthcare delivery within NHS facilities, ensuring critical aspects of healthcare management are effectively covered to improve patient outcomes, operational efficiency, and cost management.

7. EVALUATING RECOMMENDATIONS

To evaluate the effectiveness of the recommendations, a multifaceted evaluation framework is required. It involves assessing the alignment of the recommendations with predefined objectives, analysing measurable outcomes such as changes in hospital stay duration and cost savings, gathering feedback from diverse stakeholders including healthcare professionals and administrative staff. It is also important to conduct thorough comparative analyses between pre- and post-implementation data, and evaluate the sustainability of the implemented recommendations over the long term to ensure lasting benefits for healthcare services within the NHS Trust.

Additional data points that would enhance the evaluation of the recommendations include patient outcome metrics like post-surgery complications and readmission rates, resource utilization data detailing staff hours and equipment usage, longitudinal data for trend analysis, emergency response data for assessing preparedness, and continuous feedback from stakeholders for ongoing insights. Incorporating these data points would provide a comprehensive understanding of the recommendations' impact on patient care, resource allocation, emergency response, and long-term sustainability within the NHS Trust.

8. DATA COLLECTION STRATEGY

- **Identify Data Sources:** Gather data from electronic health records (EHRs), administrative databases, patient surveys, and clinical systems.
- **Standardize Data Collection:** Implement uniform protocols and coding systems across departments for consistent and accurate data capture.
- **Post-Discharge Parameters**: Collect systematic data on post-discharge metrics such as re-admittance and recovery to evaluate treatment efficacy.
- **Ensure Security and Privacy:** Employ robust measures to protect patient privacy and comply with regulations like GDPR.
- **Monitor Data Quality:** Regularly audit data, validate information, and solicit feedback for continuous improvement.
- **Collect Longitudinal Data:** Gather data over time to track trends, outcomes, and intervention effectiveness.

9. CONCLUSION

The analysis conducted offers valuable insights into hospital stay durations for back surgery patients within the NHS Trust. Through meticulous data analysis and predictive modeling, critical factors impacting hospital stays, including gender, age, surgery type, temporal patterns, and ambulation timing, have been identified. The recommendations provided focus on optimizing resource allocation, cost management, patient care, strategic planning, and emergency preparedness. Implementation of these strategies can lead to more efficient healthcare delivery and improved patient outcomes. Continuous evaluation and data collection are vital for sustaining these initiatives. By monitoring outcomes and gathering feedback, NHS facilities can refine their strategies and ensure long-term benefits.