The Second Chance: An Analysis on Transforming Patient Discharge and Recovery

(2024-2025)

Content

- **Business Case –** To build a predictive analytics model that proactively identifies patients at high risk for 30-day readmission. The goal is to use these predictions to target interventions at specific patient profiles and the critical discharge process failures revealed by the model, ultimately reducing costly readmissions.
- **Data Acquisition –** Patient discharge records, including demographics, diagnoses, and care plans, were obtained in CSV format.
- **Data Preparation –** The data was cleaned and segmented by patient complexity and discharge destination to uncover the root causes of readmission.
- **Data Visualization –** Tableau was used to visualize the patient journey, revealing a "Broken Handoff" as the key driver of readmissions.

Data Preparation

- **Total number of records = 12980**

- **No duplicate values found.**

- **Missing values treatment:** No critical predictor variables had significant missing data that required removal. The analysis was performed on the complete dataset.

- **Key Strategic Preparation:** Beyond standard cleaning, the analysis hinged on two key transformations:
  - **Building a 'Risk Profile':** Patient data was segmented using Chronic_Conditions and ER_Visits.
  - **Analyzing Transition Paths:** Care_Plan_Following_Discharge destinations were grouped to pinpoint process failures.

Exploratory Data Analysis

The Features (i.e., variables) are segregated into multiple Categories namely:

- **Dependent Variable (Y) –** This is the **Readmit30** variable (1 for Yes, 0 for No), which our model aims to predict.

- **Primary Predictor Variables –** Our analysis identified a core set of variables with the most significant, direct impact on the readmission rate. These will be detailed on the next slide.

- **Feature Prioritization & Secondary Variables –** The goal was to identify variables with the strongest direct impact for a clear, interpretable model. While routine vitals (Temperature) and demographic data (Marital_Status, Insurance_Provider) were thoroughly analyzed, they showed a weaker direct correlation with readmissions. They are retained as secondary variables for potential use in more complex feature interaction models.

Exploratory Data Analysis (contd..)

Features Reviewed for Relevance:

A comprehensive review of all variables was conducted to prioritize the most direct drivers for this model.

- While features like routine vitals (Temperature, Pulse) and general demographics (Marital_Status, Gender) showed low individual correlation, they were not discarded.
- These variables are retained for potential inclusion in future, more complex multivariate machine learning models, where their interactive effects may provide additional predictive value.

Primary Predictor Variables:
Our analysis revealed two core themes among the strongest direct predictors of readmission:

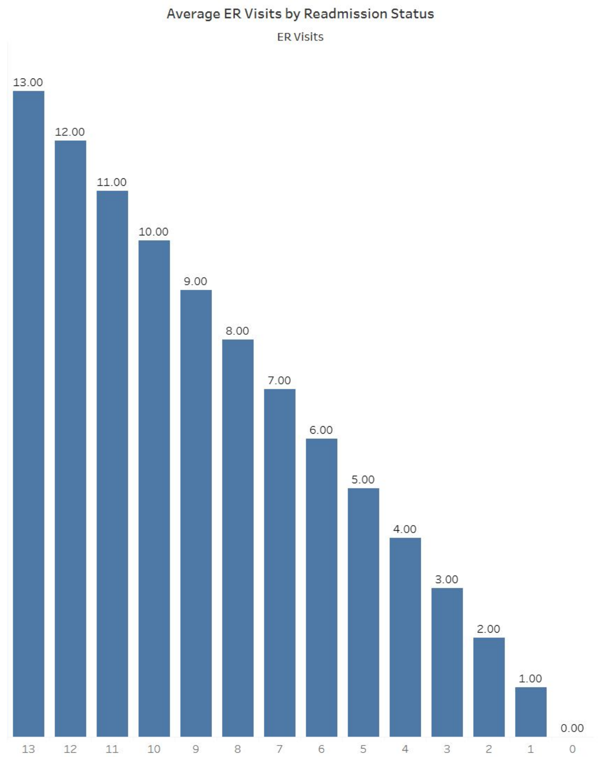
1. **The 'Chronic Crisis' Patient Profile:**
   - ER_Visits (Number of prior Emergency Room visits)
   - Chronic_Conditions (Total number of chronic conditions)
   - Cost_Of_Initial_Stay (Indicates initial medical complexity)
2. **The 'Broken Handoff' Process Failure:**
   - Care_Plan_Following_Discharge (The destination after discharge)
   - Condition (The primary diagnosis, e.g., Heart Failure)

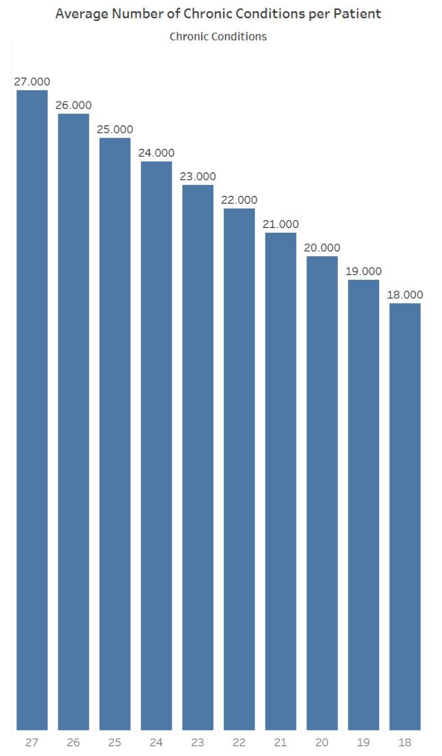
Data Visualization

The following visuals will tell the story of the "Second Chance." The charts will demonstrate:

1. **WHO** is at the highest risk of readmission (The 'Chronic Crisis' Patient).
2. **WHERE** in our process we are failing them (The 'Broken Handoff').
3. **HOW** we can begin to fix it with a focused approach.

Profile of a Second Chance: The Revolving Door Patient



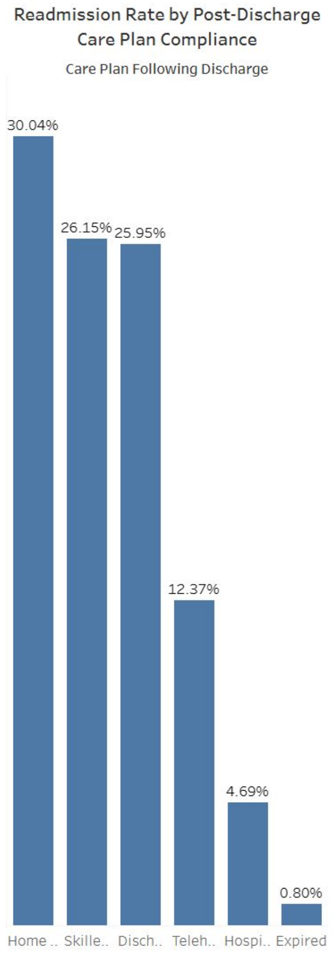


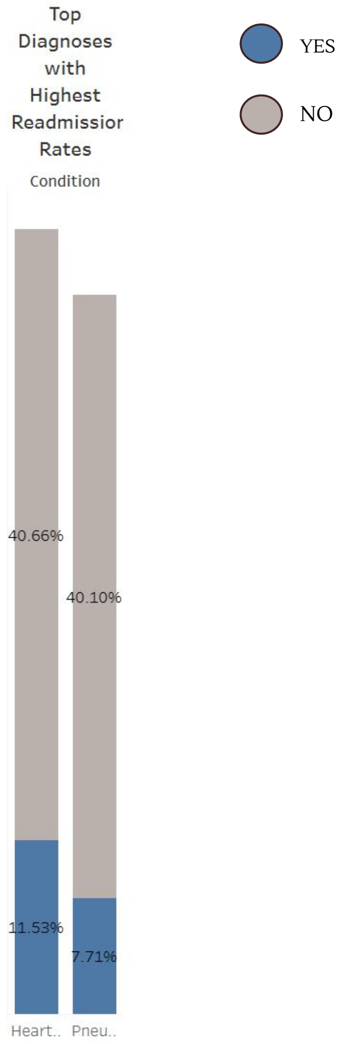
Key Highlights

High-utilizers drive risk. The readmitted cohort shows **more ER visits on average** and a **higher average count of chronic conditions**. As ER-visit frequency and multimorbidity rise, the likelihood of 30-day readmission increases.

Action: flag patients with multiple chronic conditions and frequent ER use for case management and early follow-ups.

Profile of a Second Chance: The Weight of Extreme Comorbidity



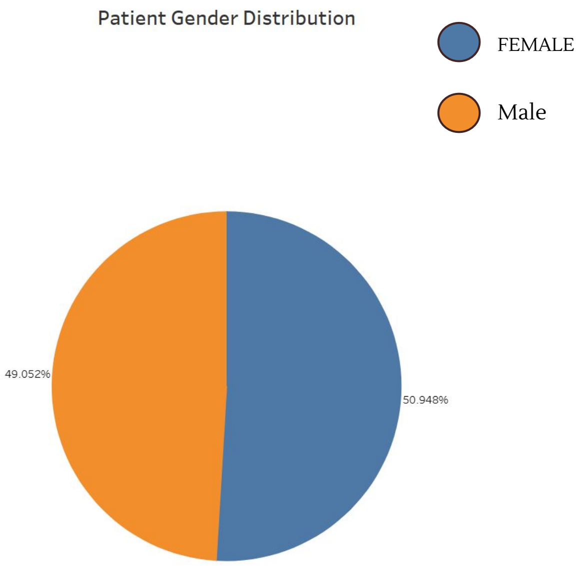


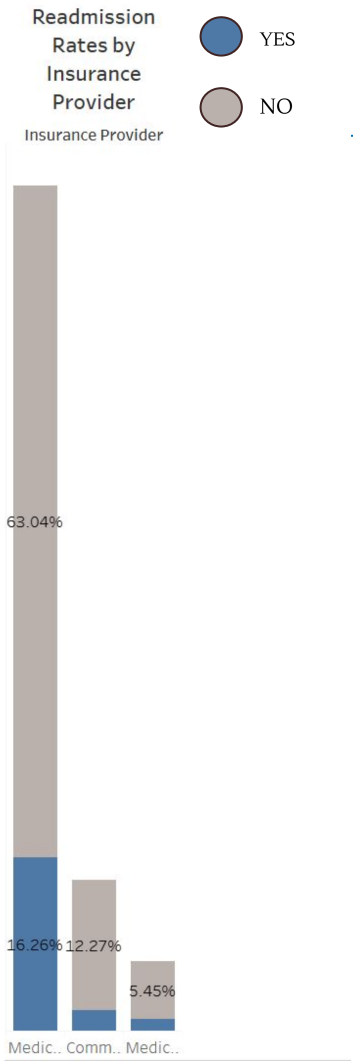
Key Highlights

Post-discharge pathway matters. Your chart shows **higher readmission percentages** for home/skilled-nursing/discharge-clinic (~26–30%) and **lower rates** when there's structured follow-up like **telehealth (~12%)** or **hospital follow-up (~5%)**. Clinically, **heart failure (~11%)** and **pulmonary/pneumonia (~7%)** emerge as the **top readmission diagnoses**.

Action: prioritize telehealth/hospital follow-up protocols for HF and pulmonary patients.

The Second Chance Moment: The Critical Handoff to Post-Acute Care



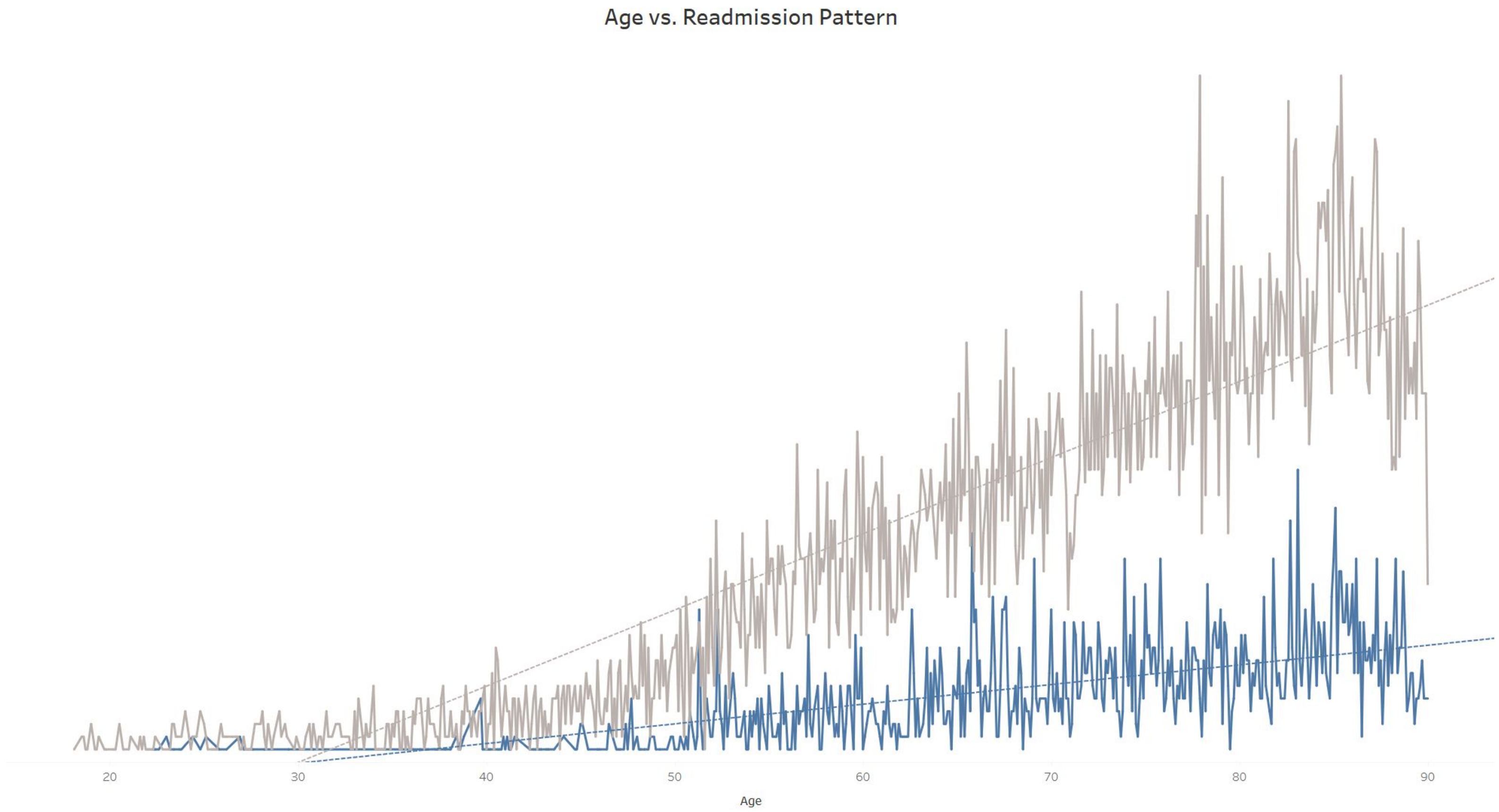


Key Highlights

Gender is essentially **balanced (~51/49)**, so it’s not driving the pattern. Readmissions are **concentrated in Medicare (~63%)**, with far fewer in Commercial/Medicaid (~12–16% / ~5%), consistent with an older, more multi-morbid population.

Action: tailor interventions to Medicare beneficiaries (transitional care, medication reconciliation, early PCP follow-up).

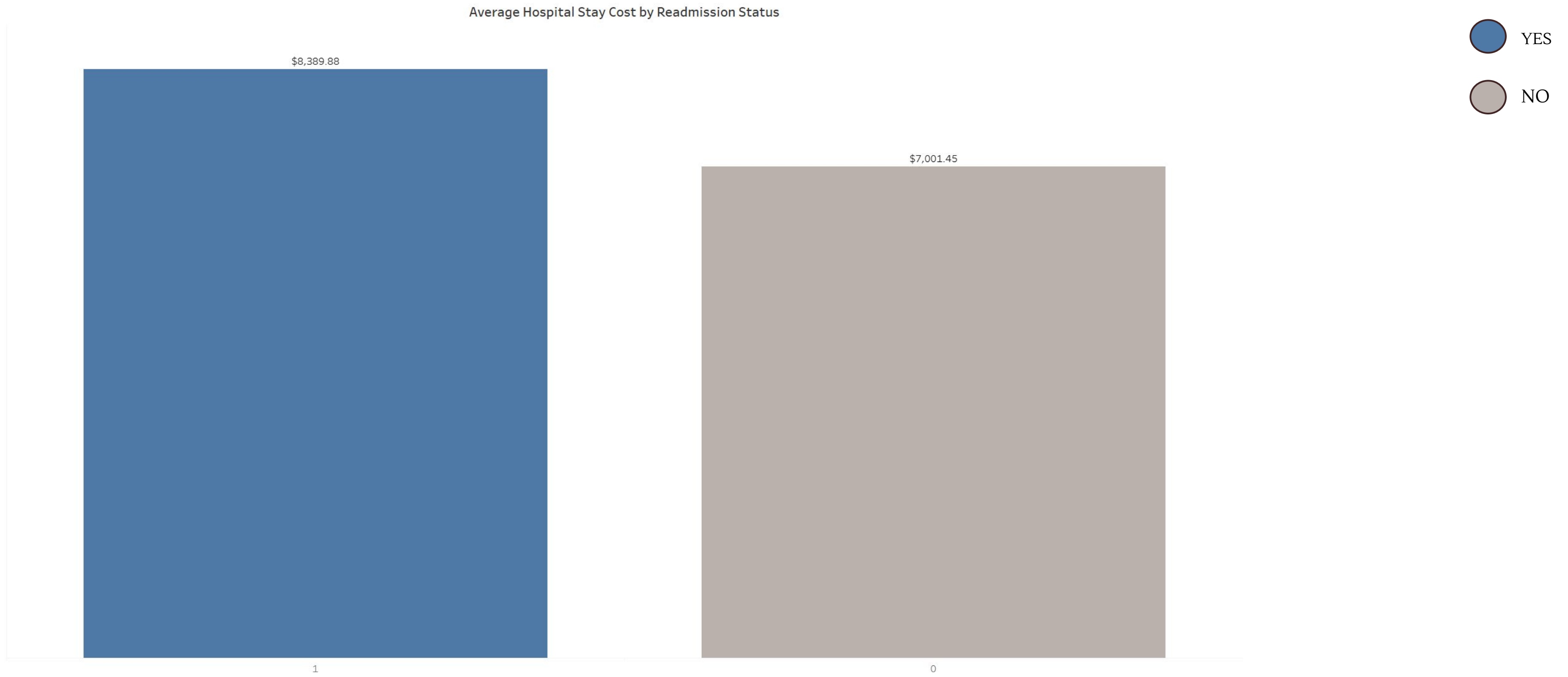
Our First Second Chance: Focusing on Heart Failure



Key Highlights

Readmission **increases with age**, with a visibly steeper climb after ~60–65. Older patients show both higher admission volume and higher readmission intensity.
Action: geriatric-focused discharge planning (teach-back, caregiver engagement, early phone check-ins).

The Value of Getting the Second Chance Right



Key Highlights

Readmitted patients had a **higher average initial-stay cost** (**$8,389.88** vs **$7,001.45**), a gap of **~$1.4K** at the first admission alone—pointing to greater initial complexity/severity.

Action: build a business case: prevention programs that reduce readmissions likely **pay for themselves** via avoided initial-stay intensity plus the avoided return stay.

Key Predictive Features & Model Performance

Significance and Correlation Analysis was performed to identify the features impacting the likelihood of being a Readmission Patient (`Readmit30.`).

Top 3 Categorical features impacting Readmission Rate:

1. ER_Visits	2. Chronic Conditions	3.Care_Plan_Following_Discharge

Top Numerical features impacting Donor Readmission Rate:

Independent Variable
Condition
Tabacco_User

Independent Variable
Cost_of_Initial_Stay
Insurance_Provider
Age

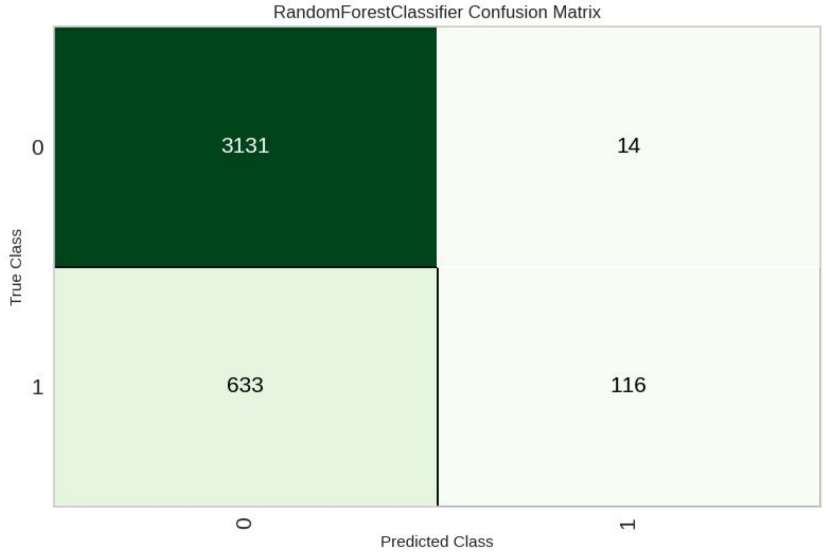
Actionable Conclusion:
The data strongly indicates that a new **"Second Chance Handoff Protocol"** should be piloted for Heart Failure patients being discharged to Skilled Nursing Facilities and Home Health to address the single largest driver of readmissions.

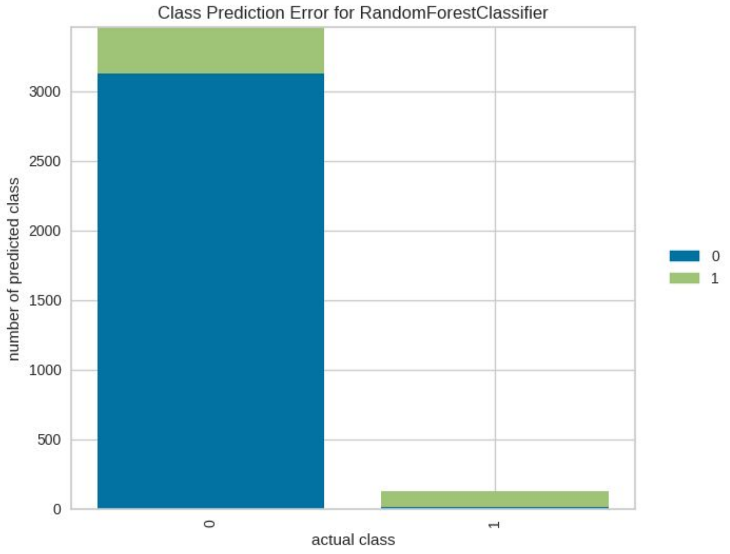
Random Forest Model Performance Overview

Performance Results

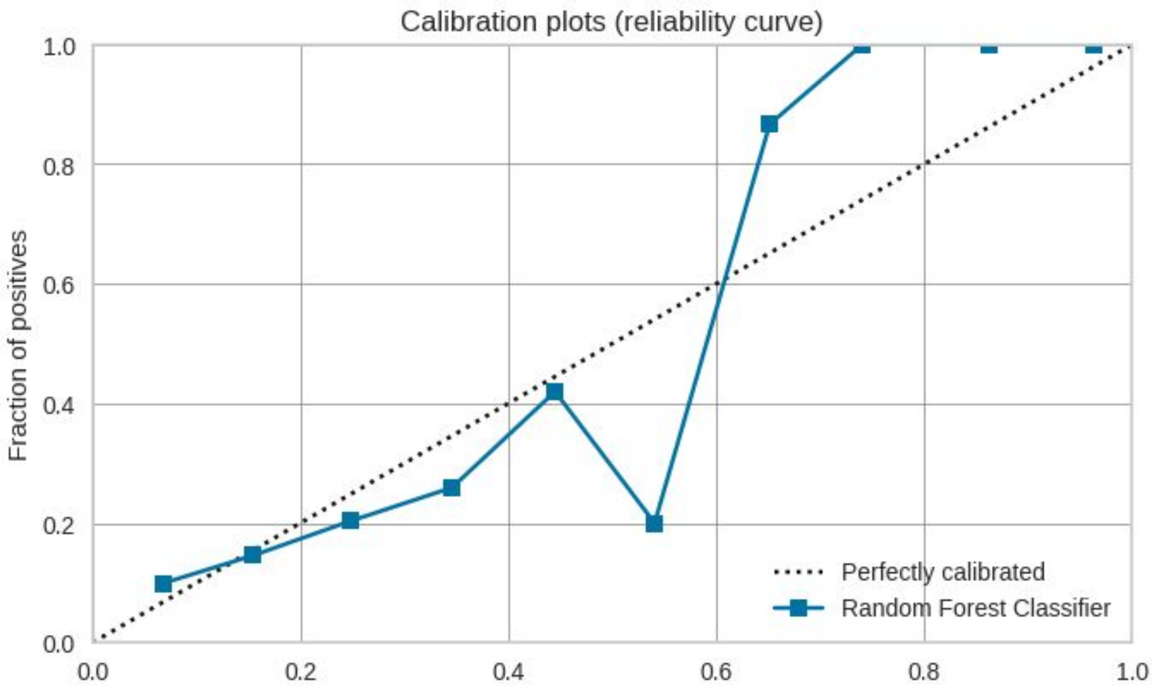
- **Accuracy:** 0.83
- **AUC:** 0.68
- **Precision:** 0.91
- **Recall:** 0.16

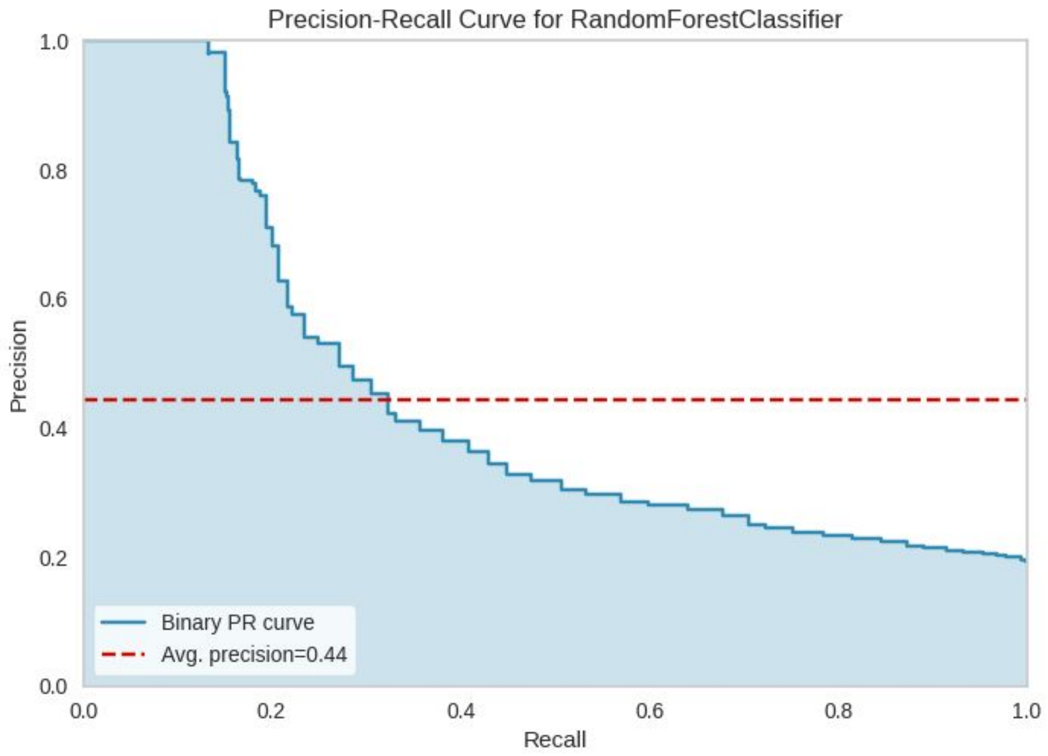
Insight:
The Random Forest model achieved **strong accuracy (83%)** and **very high precision (91%)**, meaning that when it predicts a readmission, it's almost always correct.
However, recall remains low (16%), showing that some true readmissions were missed.
This makes the model **conservative but dependable**—ideal for flagging the highest-risk patients.





Model Evaluation Visuals



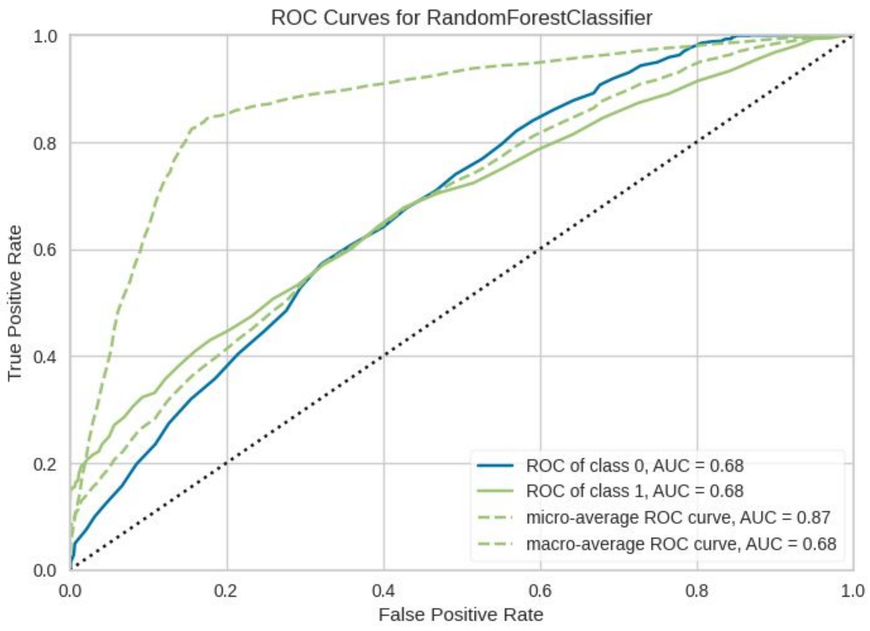


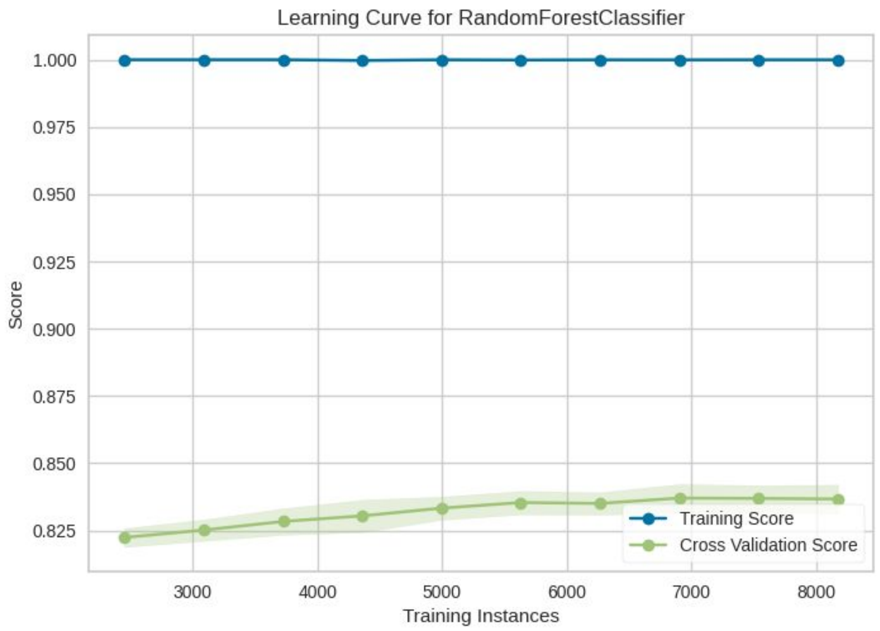
Key Highlights

The **precision-recall curve** shows a strong precision focus but lower recall—reflecting the model's cautious nature.

Finally, the **calibration plot** indicates good probability reliability at higher predicted scores.

Model Evaluation Visuals





Key Highlights

The **ROC curve (AUC = 0.68)** confirms moderate discriminative power between readmitted and non-readmitted patients.

The **learning curve** demonstrates **excellent training stability**, with cross-validation accuracy plateauing near 83%, proving no major overfitting.

Overall, the model is **balanced, consistent, and reliable** for hospital readmission prediction.

Conclusion

General Insights from the Dataset & Model

- About **20% of patients** were readmitted within 30 days.
- Readmitted patients show **more ER visits**, **higher average costs**, and **more chronic conditions**.
- **Heart failure** and **pulmonary diseases** lead readmission diagnoses.
- **Older adults** and **Medicare patients** represent the highest-risk population.
- Following **post-discharge care plans** and **telehealth follow-ups** drastically reduce readmissions.
- The **Random Forest Classifier** proved to be the **best-performing model**, offering accurate, actionable predictions.

Final Conclusion:
Targeting high-risk patients—those with frequent ER visits, multiple chronic conditions, and higher initial stay costs—can significantly reduce readmissions.
Implementing predictive analytics like this model enables hospitals to **allocate resources efficiently**, **improve patient outcomes**, and **lower overall costs**.

Questions?