

## Introduction

The project aims to analyze factors influencing hotel booking cancellations using detailed Kaggle data. This analysis supports the business justification focusing on enhancing revenue and operational efficiency in the hospitality industry.

## Business Justification

Precise forecasts aid in maximizing room rates and overbooking strategies, directly impacting hotel revenues. Revenue management techniques have shown to improve profitability, especially noted during economic downturns.

## Research Questions and Hypotheses

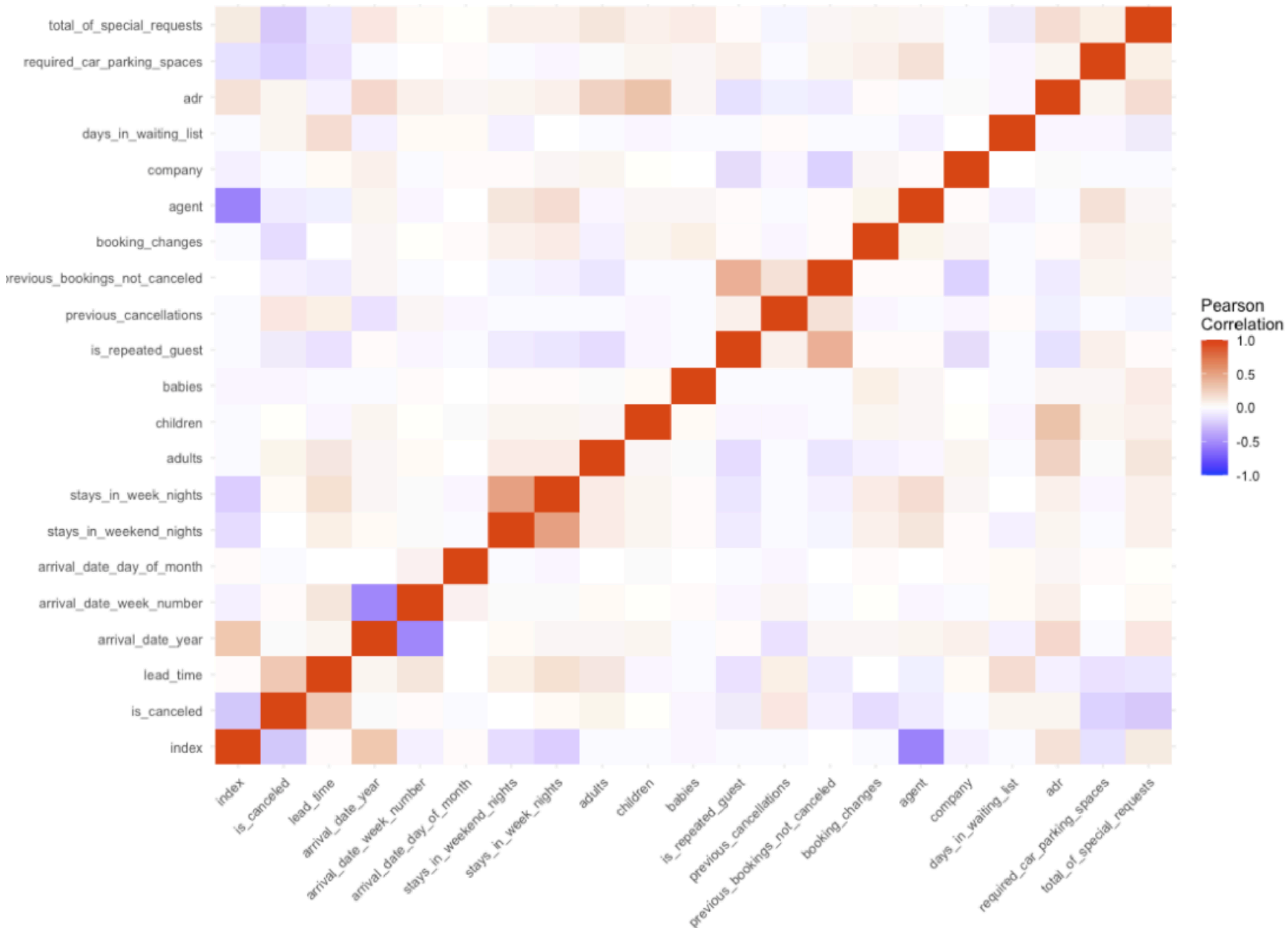
- 1. How do customer demographics and booking behaviors correlate with cancellation likelihood?
- 2. Impact of seasonal trends, room types, and advance booking periods on cancellation rates.
- 3. Influence of pricing strategies and payment conditions on cancellation probability.

Hypotheses: Specific variables such as customer demographics, time of booking, and chosen room types significantly influence the likelihood of cancellations.

## Methodology

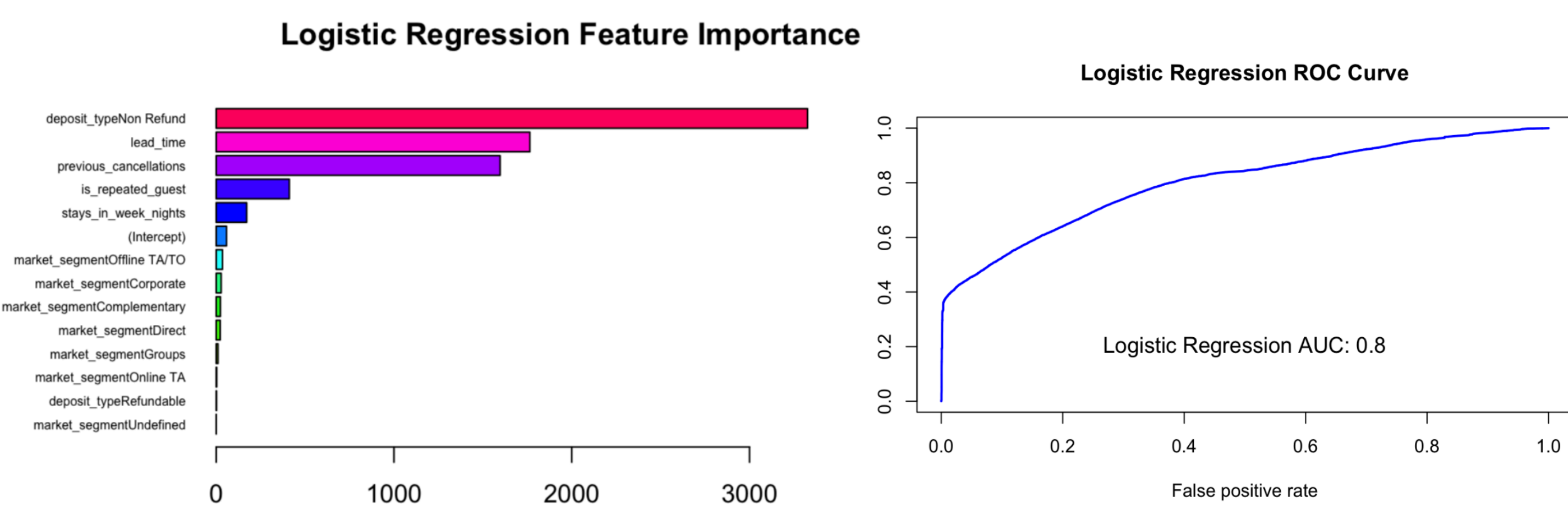
Data cleaning focused on removing columns with high missing values and replacing some missing data with mean or most frequent values.

The predictive model uses key variables like lead\_time and previous\_cancellations, which indicate higher cancellation risks, while total\_of\_special\_requests and required\_car\_parking\_spaces suggest lower risks. Booking\_changes are included to reflect how flexibility can reduce cancellations, effectively capturing critical guest booking behaviors.



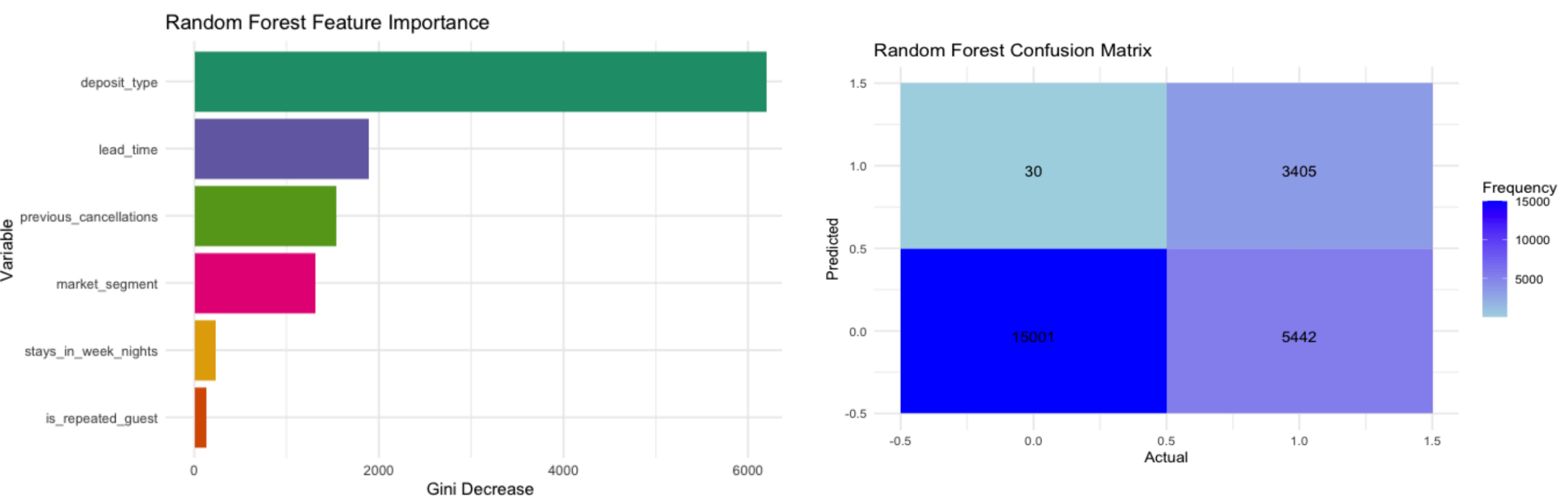
## Logistic Regression

Longer lead times and non-refundable deposits increase cancellation likelihood. Repeat guests have a lower cancellation rate, suggesting loyalty reduces cancellations.



## Random Forest

Deposit type, lead time, and previous cancellations are significant predictors of cancellations. High recall rate: effective at identifying actual cancellations but prone to false positives.

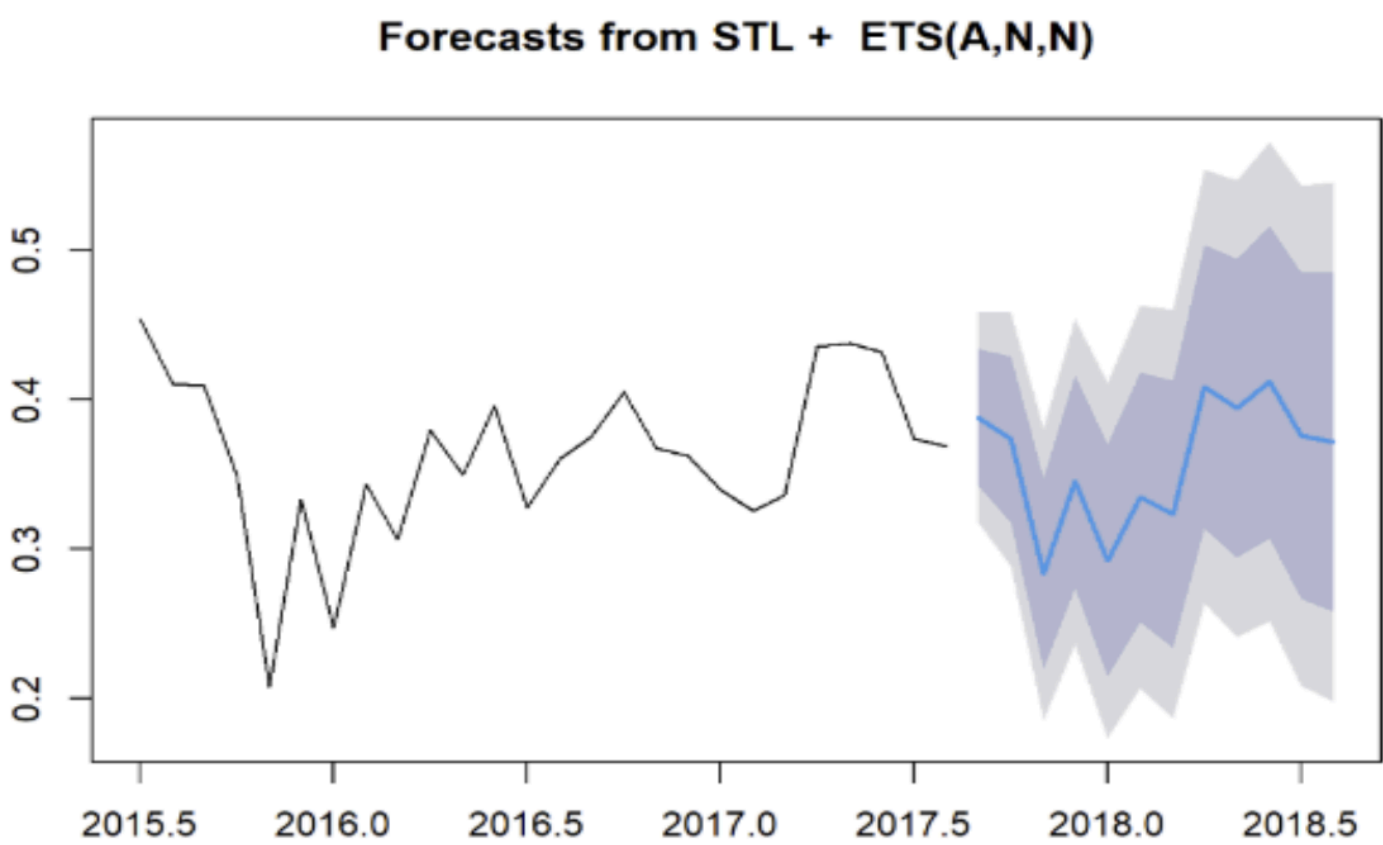


## Support Vector Machines (SVM)

High precision but moderate recall: reliable in predicting cancellations but misses some true cancellations. Needs further tuning to improve balance between detecting true positives and negatives.

## Holt-Winters Analysis

Seasonal trends show higher cancellations in summer. Upward trend in cancellations over the years, suggesting a need for strategic adjustments during peak seasons.



## Conclusions

The study significantly contributes to understanding and managing hotel booking cancellations, aiding strategic decisions in the hospitality industry. However, the study faces limitations such as geographical specificity and reliance on pre-pandemic data. Future research could include broader economic factors or external conditions to refine predictive models.