Analyzing Hotel Booking Cancellations

Leveraging Data Mining Techniques for Business Insights

Group 21: Jalyn Hamilton, Tatyana Beeston, Joseph Maldonado & Collins Indiazi

AGENDA

1. Introduction & Business Problem
2. A Look at The Data Source
3. Target Variables & Predictors
4. Methodology and Analytics Part I
5. Methodology and Analytics Part II
6. Conclusion & Recommendations

Section 1

Introduction

Understanding the Data Set

Understanding The Business Problem

OBJECTIVE: Analyze the challenges presented by booking cancellations

- One of the significant challenges faced by hotels is predicting and managing cancellations effectively to minimize revenue loss and optimize resource utilization.
- This problem affects hotel management, revenue managers, and operational staff, as well as guests who may experience inconvenience due to cancellations. By leveraging analytics, hotels can analyze historical booking data, guest behavior patterns, and external factors to develop predictive models for identifying potential cancellations.
- These models can help hotels implement proactive strategies such as overbooking mitigation, targeted marketing campaigns, and flexible cancellation policies to reduce cancellation rates while maximizing revenue and guest satisfaction. Financially, reducing cancellation rates can lead to increased revenue by ensuring higher occupancy rates and minimizing revenue loss from empty rooms.
- Operationally, effective cancellation management allows hotels to optimize resource allocation and staffing levels, improving operational efficiency and cost-effectiveness.

A Look at The Data Source

Source, Content, Method & Technique

The Source

The data required to address this problem can be sourced from the hotel booking dataset available on Kaggle (https://www.kaggle.com/jessemostipak/hotel-booking-demand).

The Content

This dataset contains comprehensive booking information for city and resort hotels, including details such as booking dates, lead time, guest demographics, and reservation status (canceled or not canceled).

The Method

After sourcing the data, exploratory data analysis (EDA) will be conducted to gain insights into the distribution and patterns of cancellations, as well as relationships between cancellation rates and other variables such as lead time, market segment, and booking channel.

The Technique

Visualization techniques such as histograms, bar plots, and heatmaps will be used to explore the data visually. Additionally, preliminary predictive models will be built using classification algorithms such as logistic regression or decision trees to evaluate the feasibility of predicting cancellations based on available features.

Target Variable & Predictors

Key Components

Target Variable: The target variable in the dataset is typically referred to as the dependent variable or the response variable. In this analysis, the target variable is likely related to hotel booking cancellations. It could be a binary variable indicating whether a booking was canceled (1) or not canceled (0). Predicting this target variable is the main objective of the analysis.

Predictors (Features): The predictors, also known as independent variables or features, are the attributes or characteristics of hotel bookings that are used to predict the target variable. Some common predictors in hotel booking cancellation prediction models include:

Lead Time

The number of days between the booking date and the arrival date.

Duration of Stay

Number of nights the guest stays, including both weekday and weekend nights.

Booking History

Previous cancellations, previous bookings not canceled, etc.

Arrival Time

Month and week number of arrival.

Guest Demographic

Number of adults, children, and babies.

Room, Deposit & Customer Type

Type of room reserved.

Type of deposit made for the booking.

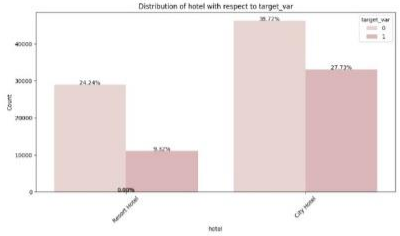
Type of customer (e.g., transient, contract, group).

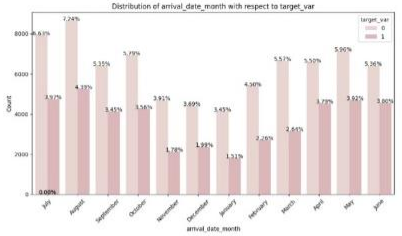
Average Daily Rate

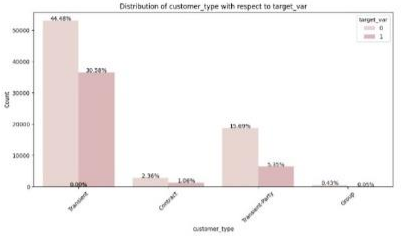
Average daily rate of the booking.

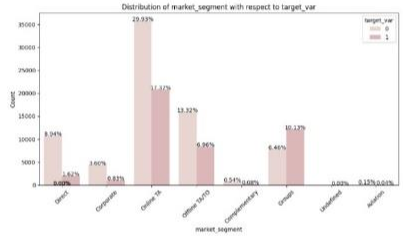
Booking Details

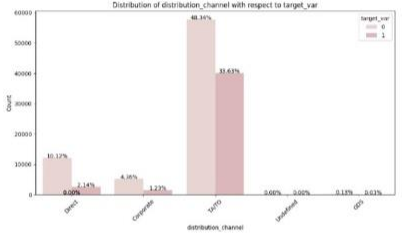
Meal type, market segment, distribution channel, etc.

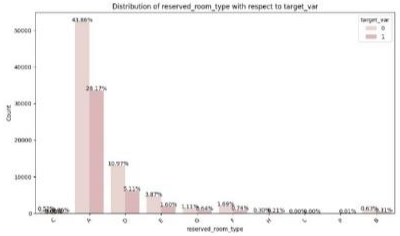


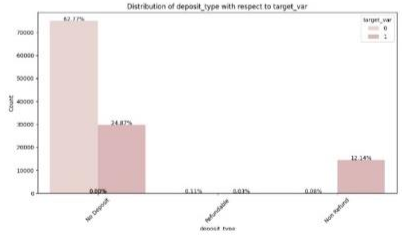


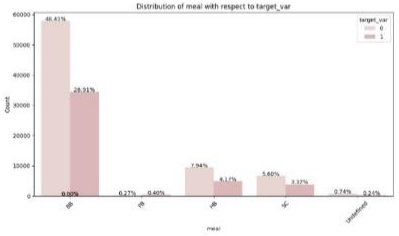












Distribution of Target Variable

Section 2

Model Approaches

Analysis & Testing

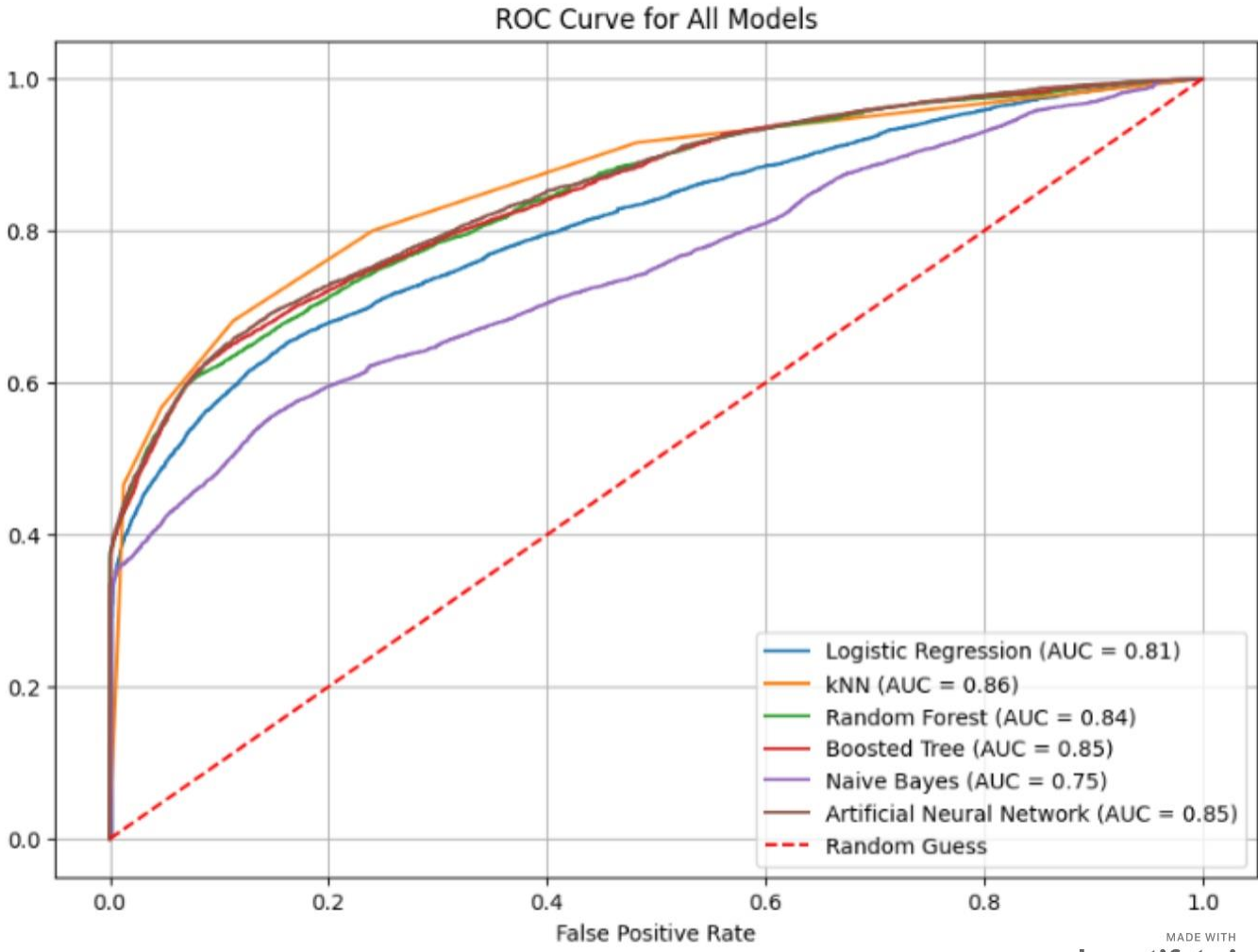
ROC Curve for All Models

Visual IV - Regression Analysis

Our Model Selection is kNN

- Classification Report:

	precision	recall	f1-score	support
0	0.83	0.89	0.86	22576
1	0.78	0.68	0.73	13240
accuracy			0.81	35816



Key Variables

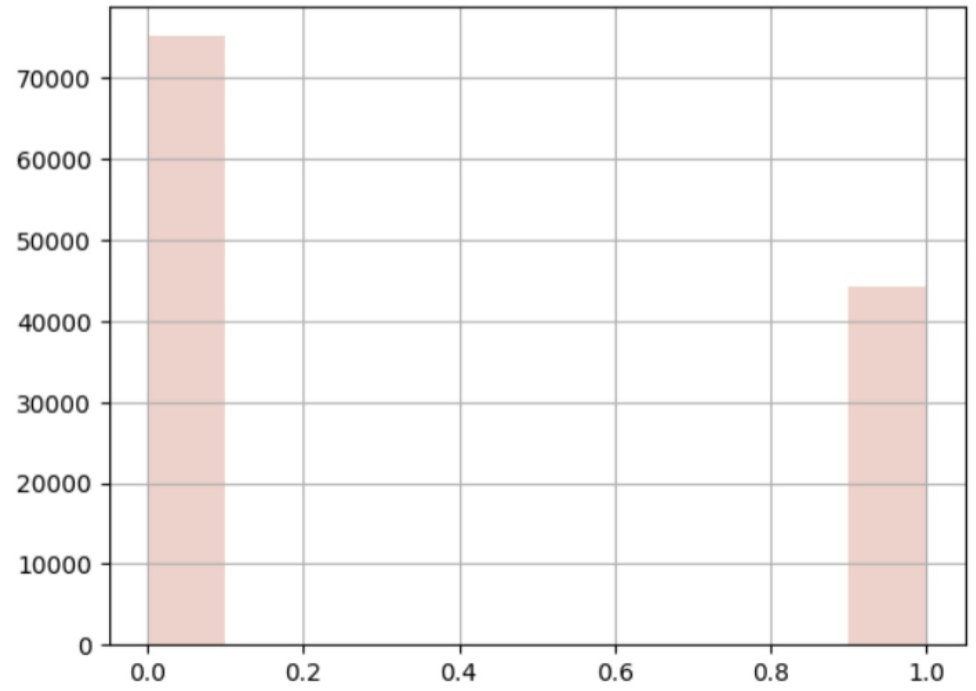
Visual II - Histogram

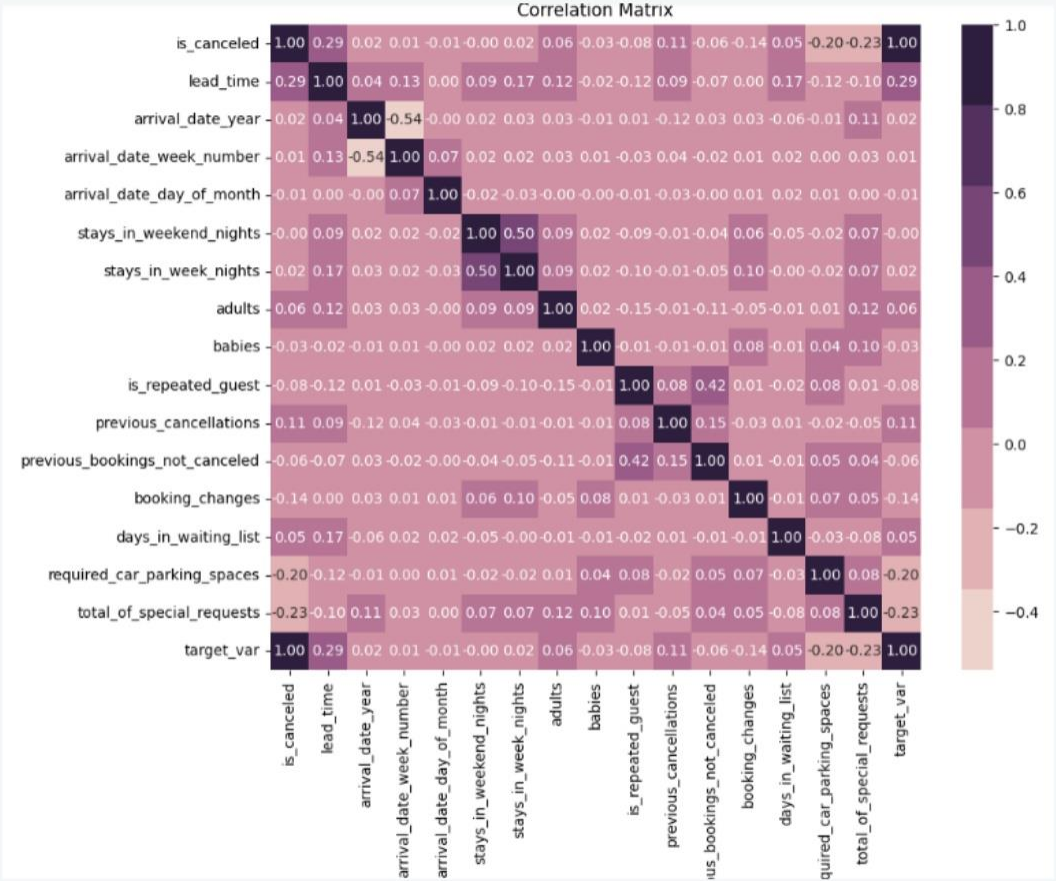
The Presented Histogram depicts the distribution of the target variables 0= Canceled; 1=Booked

```
# Generate Cubehelix color palette with 2 colors
cubehelix_colors = sns.cubehelix_palette(8)

# Plot histogram of 'is_canceled' variable
hotel_bookings['is_canceled'].hist(color=cubehelix_colors[0])
```

```
<Axes: >
```





Visual III - Heat Map

The correlation Matrix - The relation between numerical variables and target variables.

Section 3

Conclusion & Recommendations

Our Final Thoughts

Conclusion

Summary & Closing Statements

Pricing Strategy Optimization:

- Given the higher cancellation frequency observed in city hotels, it is crucial to reassess pricing strategies to remain competitive and adaptable.
- City-based hotels should consider revising their pricing tactics to better align with customer preferences and market demands, especially during peak periods characterized by heightened cancellation rates.

Enhanced Guest Retention Initiatives:

- City hotels should prioritize initiatives aimed at enhancing guest retention to mitigate the adverse effects of cancellations on revenue streams. Strategies may include personalized guest experiences, loyalty programs, and proactive communication to foster stronger relationships with guests and encourage repeat bookings.
- Resort hotels can leverage their unique offerings and amenities to create memorable guest experiences, thereby increasing guest satisfaction and loyalty. By delivering exceptional service and value, resort hotels can minimize the likelihood of cancellations and cultivate a loyal customer base.

Collaboration with Distribution Partners:

- Hotels should prioritize transparency and consistency in marketing materials and communication channels to build trust and credibility with guests. By providing accurate and reliable information, hotels can manage guest expectations effectively and minimize the likelihood of cancellations resulting from miscommunication or misunderstanding.

Continuous Monitoring and Adaptation:

- It is essential for hotels to continuously monitor booking trends, guest feedback, and market dynamics to adapt their strategies and offerings accordingly. By staying agile and responsive to changing consumer preferences and market conditions, hotels can maintain a competitive edge and drive sustainable growth in an increasingly dynamic and competitive landscape.