Understanding and Reducing Booking Cancellations: A Data-Driven Approach

Alejandra Barrios

Capstone

Business Analytics & Insights

August 2024

Executive Summary

Building and analyzing a **predictive model** to identify which customers are more likely to cancel their booking.

Some **key factors** that influence hotel booking cancellations are:

- Deposit type method
- Region
- Booking modifications
- Customer history

Recommendations

- Enhance customer communication and service
- Improve targeted marketing efforts
- Optimize operational efficiencies

Agenda

High cancellations rates can significantly impact hotel revenue and operations

Average Cancellation Rates

Hotels experience an average cancellation rate of 20-25%

Revenue Impact

Cancellation can lead to revenue losses up to 15% anually

Operational Challenges

High cancellation rates result in overbooking, underbooking, and inneficient resource allocations

Effective predection of cancellation can help in optimizing bookings and improving customer satisfaction

Operational Efficiencies

Hotels using predictive analytics have seen a 10-15% increase in booking efficiency

Revenue Stability

Implementing dynamic pricing based on cancellation predictions has led to 5-10% increase in revenue

Customer Satisfaction

Enhancing customer communication and policies to reduce cancellations improves overall guest experience

Driving operational excellence through predictive analytics into Booking Cancellation
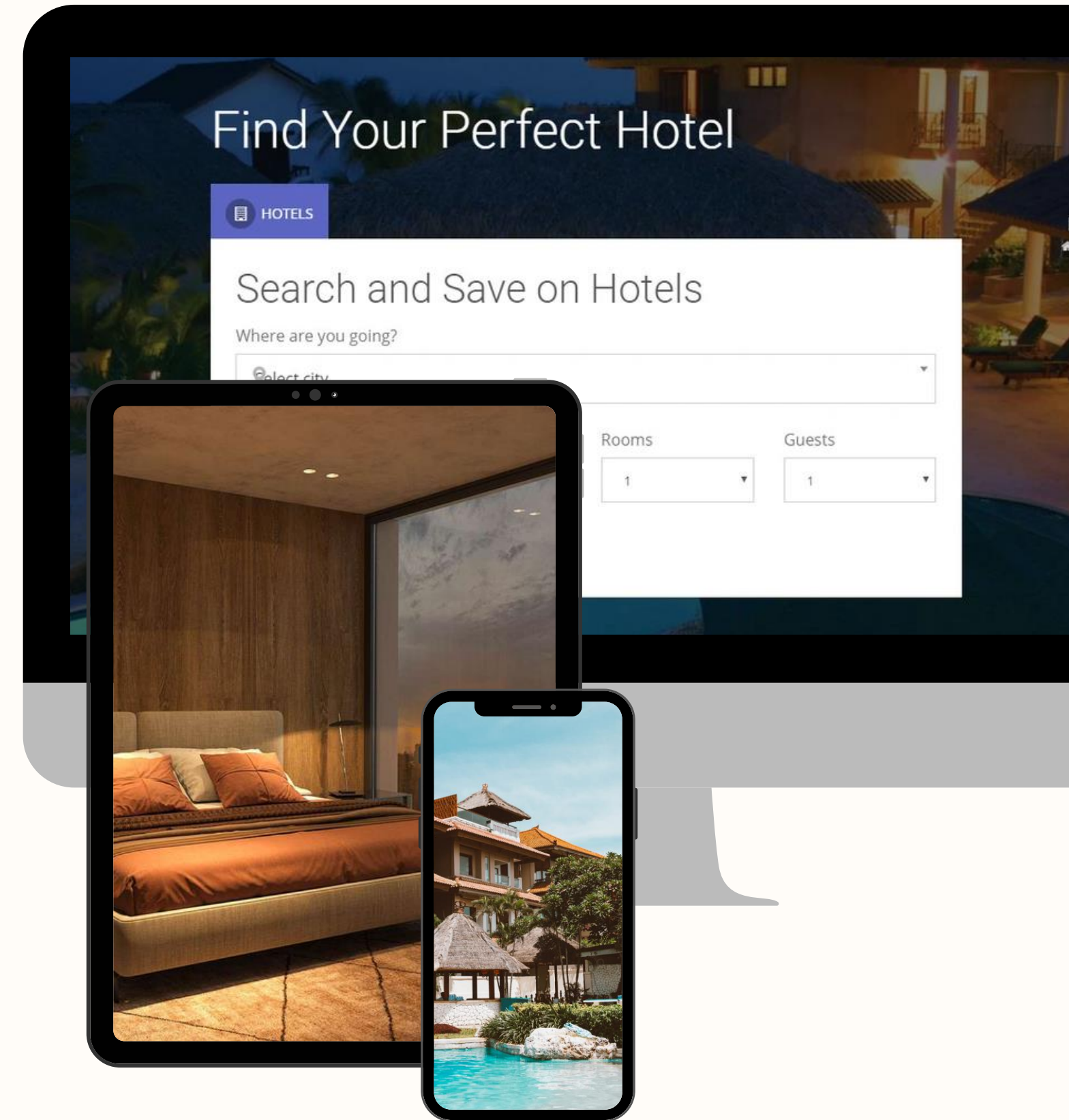
Problem Statement

High cancellation rates disrupt inventory management, staffing schedules, and financial forecasting, leading to revenue loss and operational inefficiencies.

Understanding and predicting cancellations is crucial for implementing proactive measures to mitigate their impact on hotel operations.

Objectives

- Identify the key drivers behind cancellations to improve operational inefficiency and reduce costs
- Develop targeted interventions and policies to reduce the likelihood of cancellation, thereby increasing guest satisfaction

Dataset encompases arrivals and cancellations for a Hotel Chain in Portugal

SOURCE

Datasets were obtained from the data article "Hotel Booking demand datasets" at the Journal of Science Direct.

City Hotel   Resort Hotel

TIMELINE

Hotel bookings with arrival scheduled from July 1, 2015 to August 31, 2017

STRUCTURE

119,390 records and 32 variables (i.e. lead time, number of guests, date of arrival, average daily price, etc.)

TARGET VARIABLE

The model will predict if the booking will be canceled or not.

0 - Not canceled
1 - Canceled

EXCLUSIONS/ASSUMPTIONS

- Date columns (year, month, week number, day of the month)
- Variable "reservation_satus" (Check Out and Canceled)
- Counter-intuitive variables

# The optimized decision tree balances high predictive accuracy with interpretability

METRICS

Accuracy - overall performance
**ROC AUC -** discrimination capability

WHY DECISION TREE OVER RANDOM FOREST?

- **Interpretability:** provides a clear and intuitive representation of decision rules
- **Transparency in Decision-Making: i**dentifying the most significant features that influence the likelihood of cancellation
- **Model Complexity and Maintenance:** simpler model structure, which can be easier to maintain and update
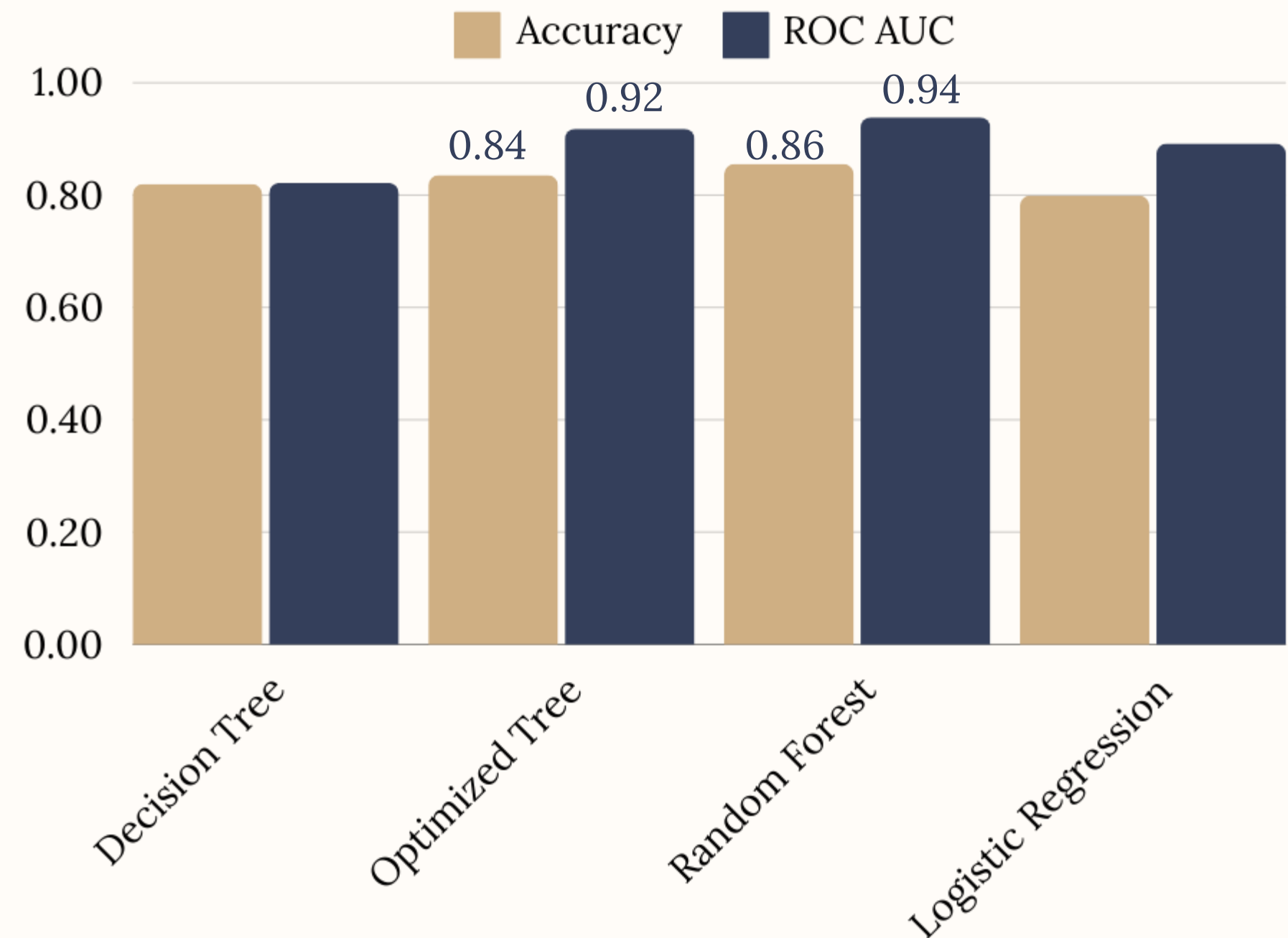
A decision tree is a powerful tool that helps us visualize and understand the factors influencing cancellations
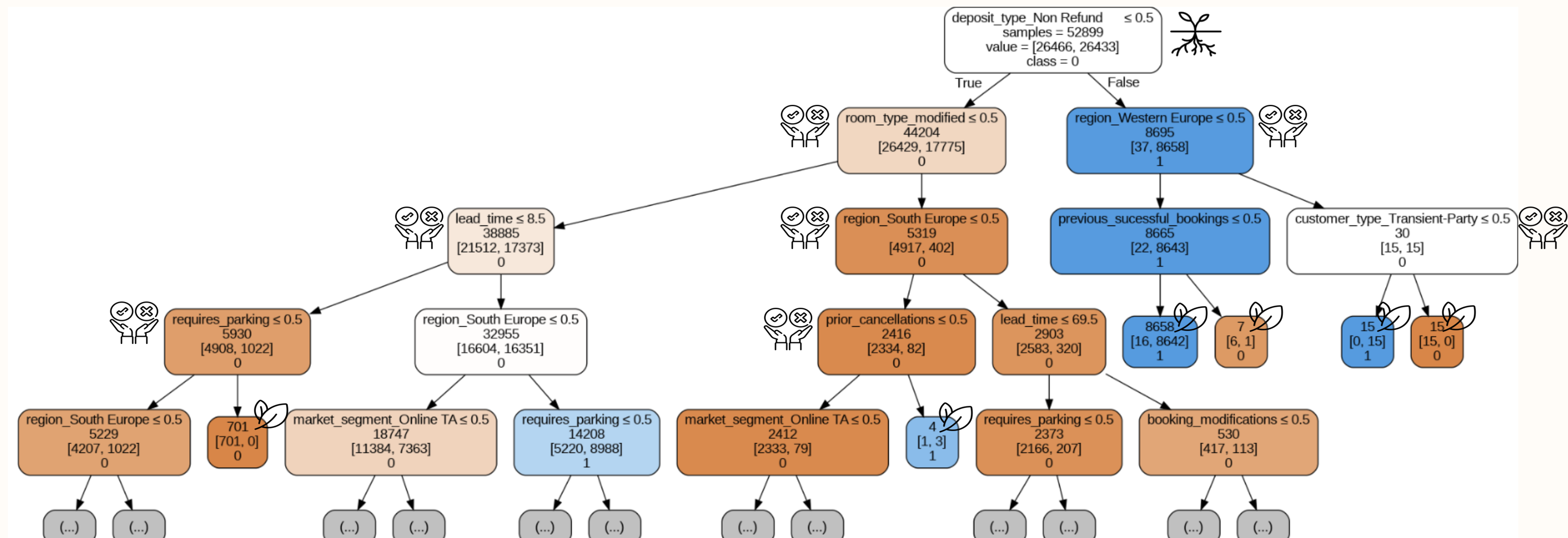
Root Node
Starting point of the tree

Decision Nodes
Specific feature or condition

Leaf Nodes
Outcome (Canceled and Not-Canceled)

The deposit type "Non Refundable" is the primary factor influencing whether a booking is canceled


IMPACT OF DEPOSIT TYPES

Customers who pay a **non-refundable** deposit have a much higher percentage of **canceled reservations (99.6%),** compared with the ones that do not pay a non-refund deposit (40.2%)

ROOM ASSIGNMENT EXPERIENCE

Exact room assignments experience a significantly higher **cancellation rate (44.6%)** compared to those with different room assignments (7.55%).

REGION INFLUENCE

Guests from **Western Europe** have **50% chance** to cancel their booking
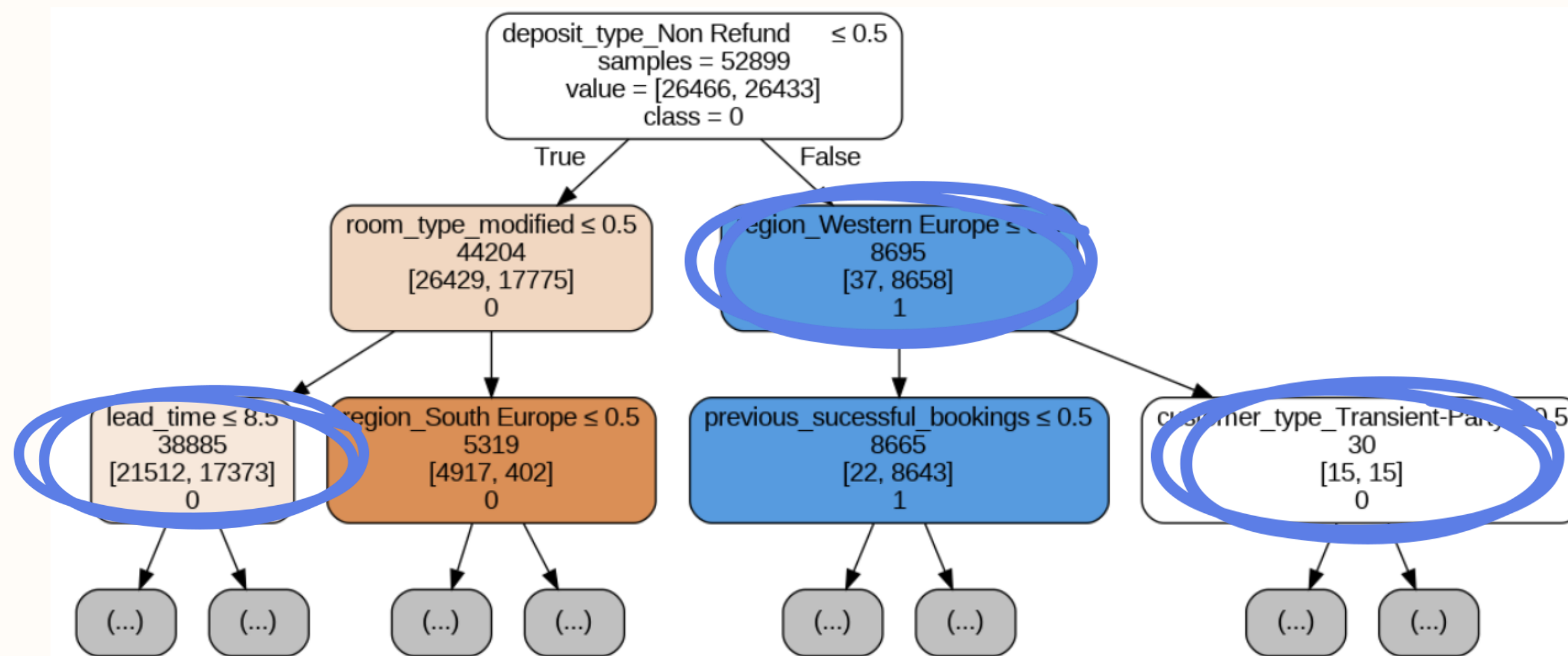
Customers from different regions exhibit distint booking behaviors

Western Europe
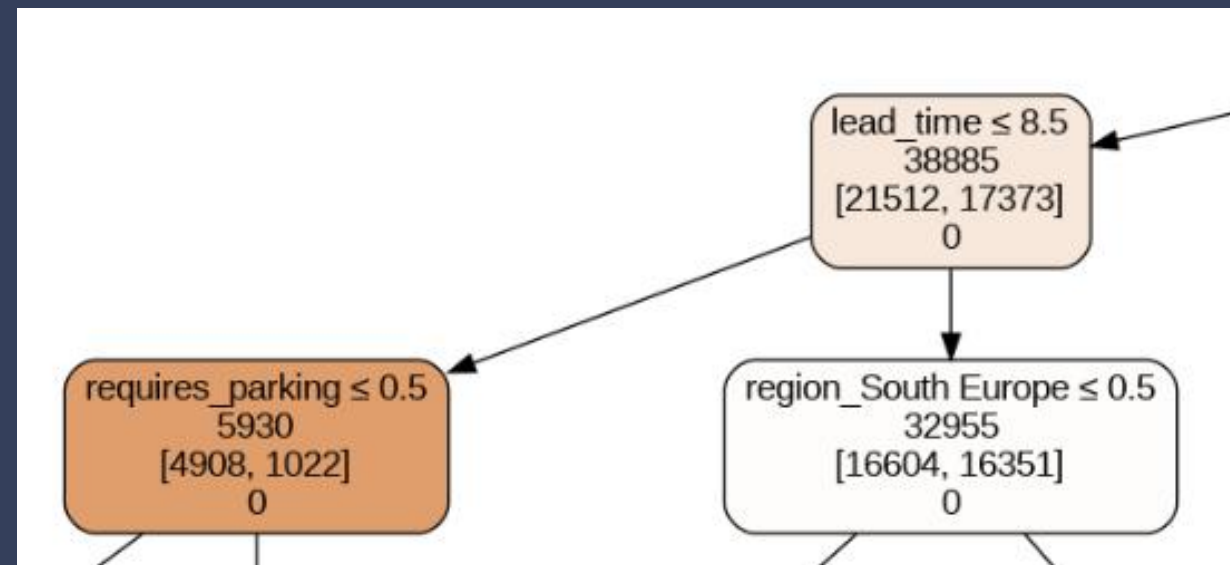
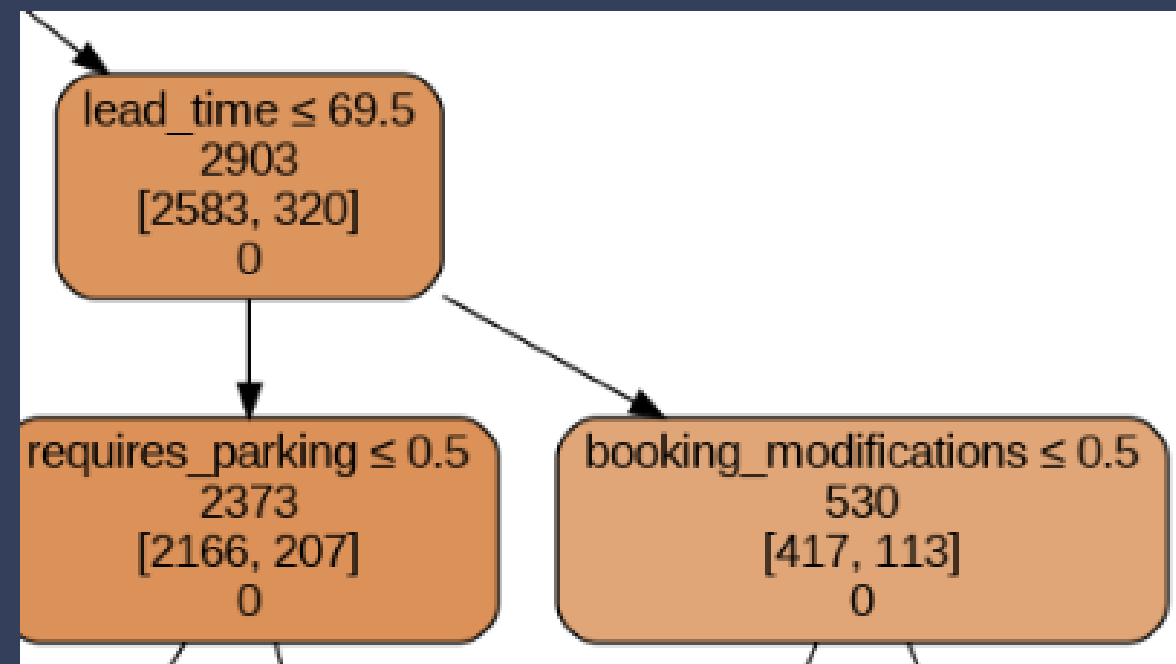Guests from Western Europe have a cancellation rate of 50% despite non-refundable deposit.

South Europe

Guests from South Europe have 11% chance of cancellation

Longer lead times present a higher cancellation rate on bookings

- Bookings with lead times of less than 8.5 days have a lower cancellation rate (17.23%) compared to those with **longer lead times (49.61%)**.
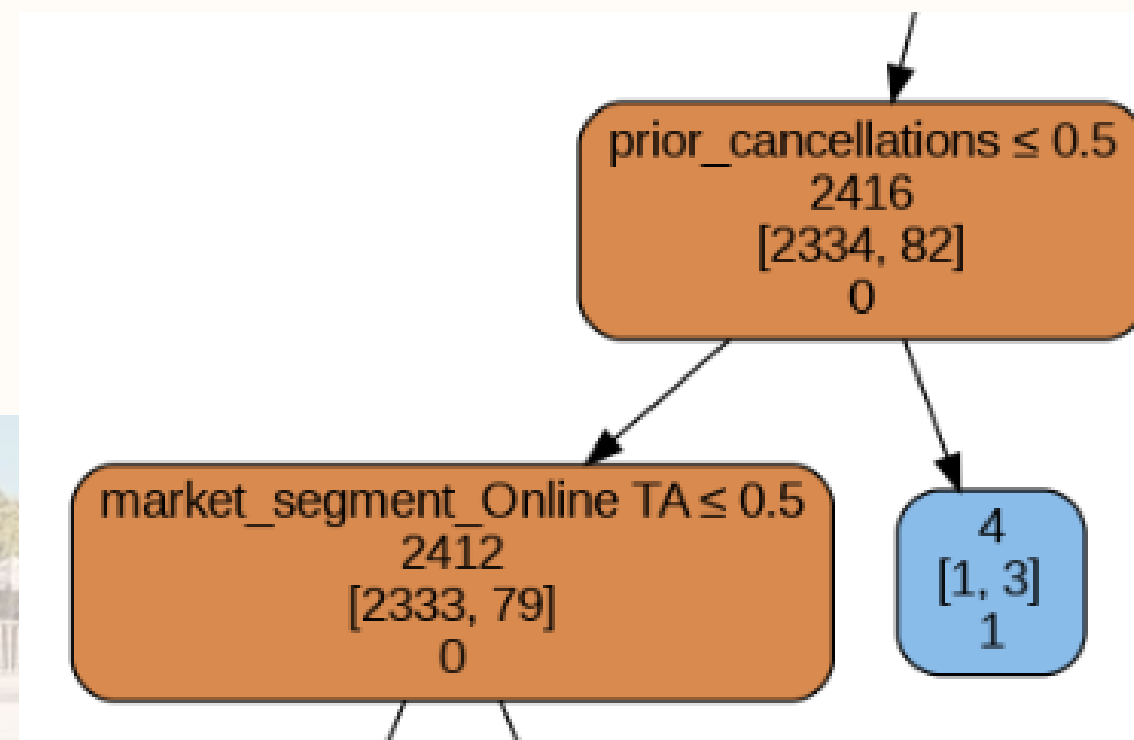
- Bookings with **lead times longer** than 69.5 days have a **higher cancellation rate (21.32%)** than shorter lead times (8.72%).
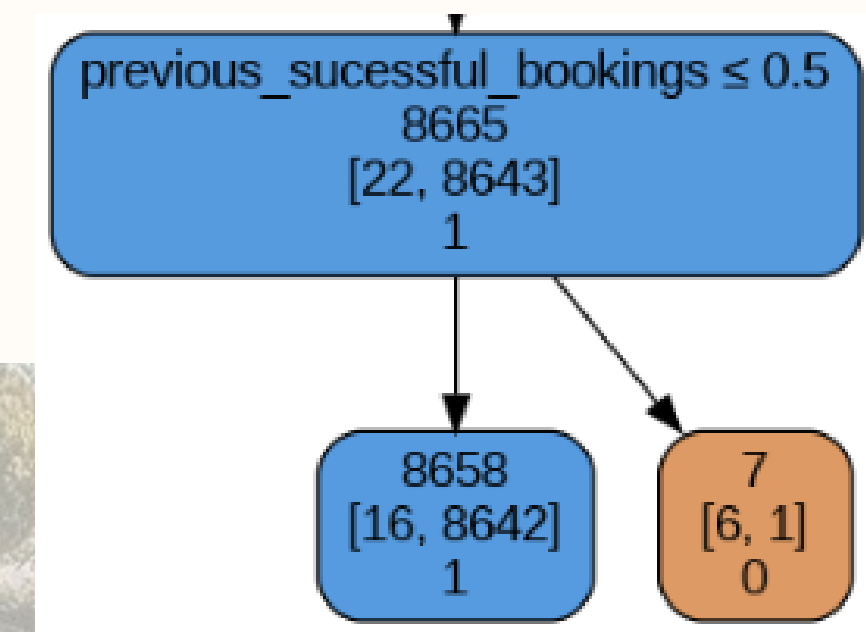
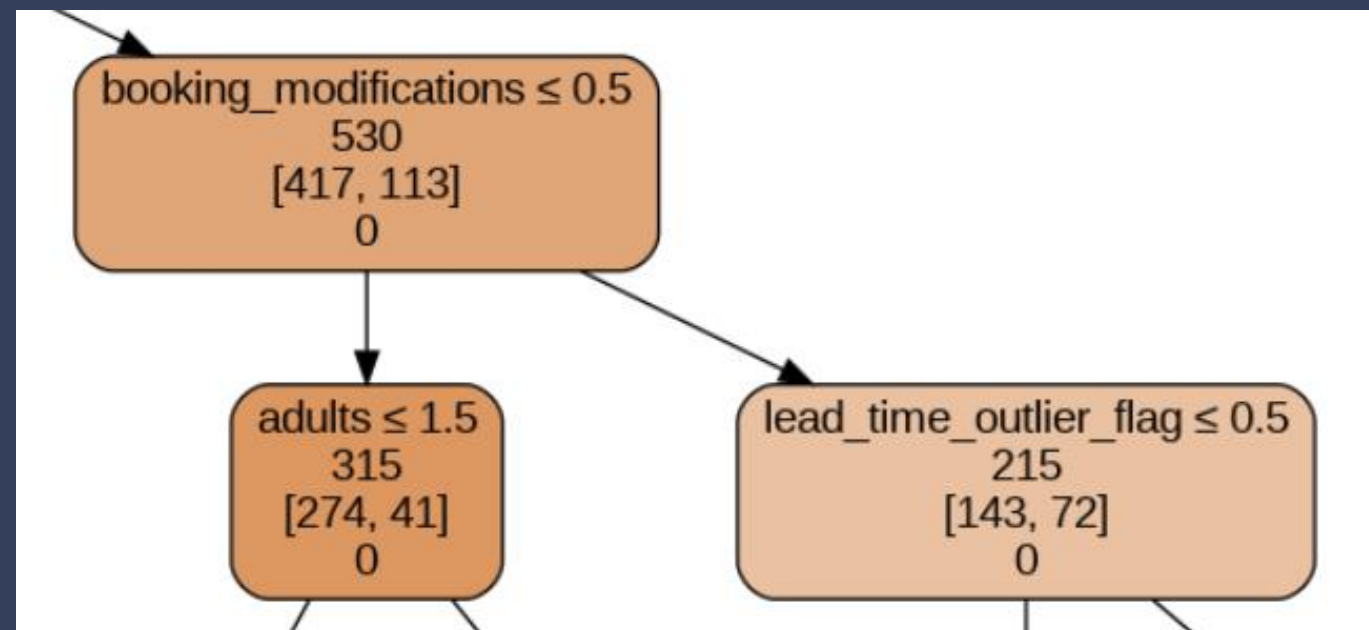# Customers with previous cancellations are at a greater risk of canceling

Customers with **previous cancellations** have a significantly **higher chance** of canceling again **(75%)** compared to those without any cancellation history (3.2%)

First-time guests with no previous successful bookings have **higher chances of canceling (99.8%),** compared to a much lower chance (14.3%) for guests with prior successful bookings.
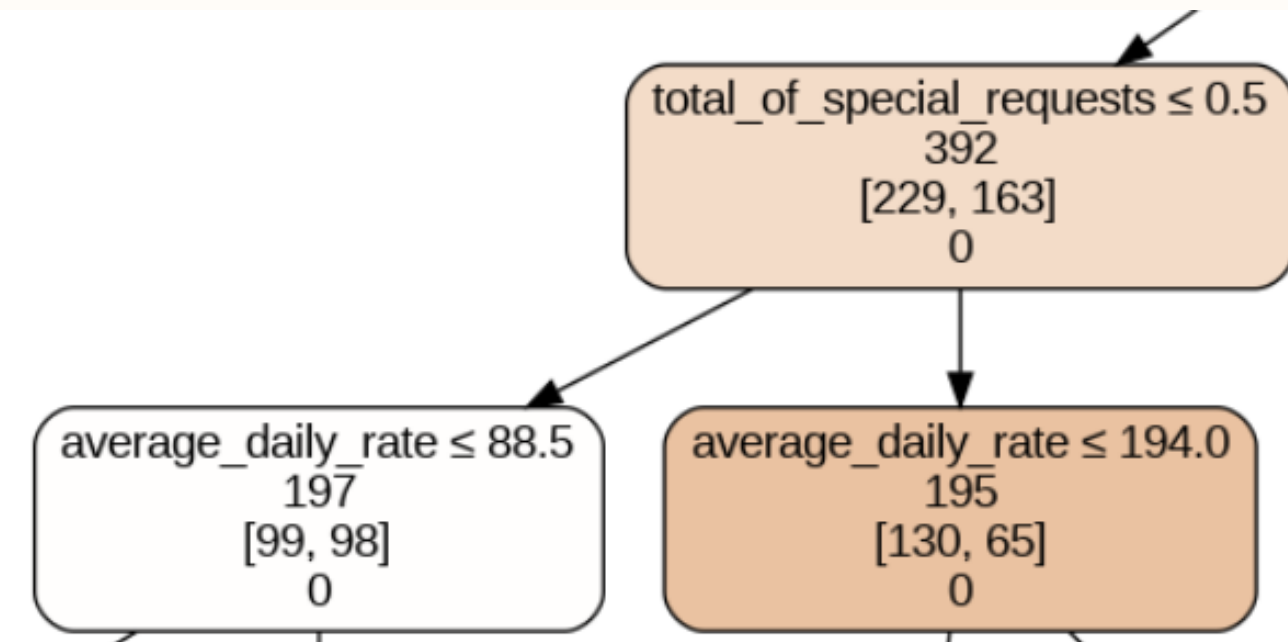
Customers that modify their bookings are more likely to cancel

booking_modifications ≤ 0.5
530
[417, 113]
0

adults ≤ 1.5
315
[274, 41]
0

lead_time_outlier_flag ≤ 0.5
215
[143, 72]
0

Customers that **modify their booking** are **more likely to cancel (33%)**, compared with 13% of chances to cancel if they do not modify their booking.

Bookings with **special requests have lower chances to cancel (33%)**, compared to 59% of chances to cancel when the booking does not have any special requests.

total_of_special_requests ≤ 0.5
392
[229, 163]
0

average_daily_rate ≤ 88.5
197
[99, 98]
0

average_daily_rate ≤ 194.0
195
[130, 65]
0

Conclusion

Decision Tree (Cross-Validation) is the model with the highest predictive power, classifies whether a booking will be canceled with 84% accuracy and 92% ROC AUC.

Deposit Method	• Non-Refundable bookings have a higher impact in the cancellation rate, highlighting a gap in customer expectations or communication.
Customer Region	• Customer from Western Europe are more likely to cancel than South European customers.
Lead Time	• Longer lead times have a higher cancellation rate.
Customer History	• Customers with cancellation history present a higher risk for cancellation. • Customers with previous sucessfull bookings have a lower chance to cancel.
Booking Modifications	• Modification in the room reserved present a lower chance to cancel. • Customers that do modify their booking are more likely to cancel. • Bookings with special requests help to reduce the probability of cancellation.

Recommendations: Model 2.0

Investigate the quality of data since there are some variables that are counter-intuitive.

Explore customer segmentation with techniques like K-means, or Hierarchical Clustering.

Business Recommendations

CUSTOMER COMMUNICATION

- Enhance customer service by providing clear booking terms of Non-Refundable Bookings.
- Engage with customers who book far in advance through regular updates or personalized offers.
- Offer additional support and guidance to customers making changes to manage expectations and reduce cancellation

PERSONALIZED INCENTIVES

- Offer incentives or rewards for repeat guests to encourage continued positive booking behavior.
- Provide discounts or special rates for customers who book well in advance.
- Provide additional incentives for bookings with special requests, such as complimentary upgrades or services.
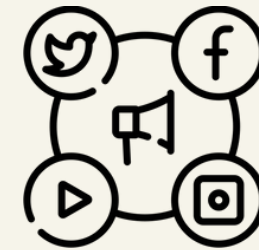
MARKETING CAMPAIGNS

- Implement targeted campaigns focusing on regions with high cancellation rates particularly Western Europe.
- Promote the benefits of maintaining the original reservation and offering flexible rebooking options.

SPECIAL REQUESTS

- Offer promotions or discounts to customers who include special requests.
- Improve the handling and fulfillment of special requests can reduce cancellations and secure more stable bookings.

References

- Nuno, A. Almeida, A. Nunes, L. February 2019. Hotel booking demand datasets. Science Direct. https://www.sciencedirect.com/science/article/pii/S2352340918315191#t0040
- Kramser, V. Octuber 17, 2023. Unlocking Predictive Analytics: A game changer for hotel operations. NeedNect Blog. https://blog.neednect.com/predictive-analytics/
- RateBoard. January 21, 2020. Hotel cancellations pose a great challenge. RateBoard. https://www.rateboard.io/en/blog/hotel-cancellations-pose-a-great-challenge
- Elyssa, A. February 22, 2023. Hotel Booking Trends. Hotel Business. https://hotelbusiness.com/siteminder-traveler-confidence-surges-post-pandemic/

Q&A Session

Thank you for your listening!