Hotel Booking Demand & Cancellations Analysis Business Analyst Report

1. Problem Statement

The goal of this analysis is to understand hotel booking demand patterns and identify key factors influencing booking cancellations. This helps hotels optimize revenue, reduce cancellations, and plan resources effectively.

2. Dataset Overview

The dataset includes booking details from City and Resort hotels. Key fields include lead time, stay duration, guests, booking changes, cancellation status, price (ADR), month/year of arrival, and customer type.

3. Key Business Questions & Solutions

Q1: What is the overall cancellation rate?

A: Cancellations account for approximately 37–40% of total bookings.

Q2: Which hotel type has more cancellations?

A: City hotels show higher cancellation rates compared to Resort hotels.

Q3: What factors contribute most to cancellations?

A: High lead time, higher ADR (price), previous cancellations, longer waiting lists, and OTA (online travel agency) bookings show higher cancellation likelihood.

Q4: Do prices influence cancellations?

A: Yes. Higher ADR values correlate strongly with increased cancellations.

Q5: Which months experience the highest cancellation activity?

A: June–August show peak cancellations due to increased seasonal demand.

Q6: Do repeat guests cancel often?

A: No. Repeat guests show extremely low cancellation behaviour.

4. Exploratory Data Analysis Insights

• Lead time was the strongest predictor of cancellation likelihood.

• ADR (average daily rate) is significantly higher for cancelled bookings.

• City hotels show higher booking volume but also higher cancellations.

• Online booking channels (OTA) contribute the highest cancellations.

• Direct bookings and corporate bookings show lowest cancellations.

• Peak months see higher prices and higher cancellations.

5. Machine Learning Model Summary

A Random Forest classification model was trained to predict cancellations.

Model Inputs: - lead_time - total_stay - total_guests - arrival_month - previous_cancellations - booking_changes - adr

Key Findings:

• Lead time was the most important feature.

• ADR and previous cancellations also had strong predictive power.

• The model performed well in identifying high-risk bookings.

6. Final Recommendations

✓ Reduce lead times for OTA bookings by offering incentives for early confirmation.

✓ Implement dynamic pricing strategies during high-risk months (June–August).

✓ Reward repeat guests to reduce churn.

✓ Monitor customers with long waitlist durations.

✓ Encourage direct bookings through loyalty programs and discounts.

✓ Improve forecasting using the trained ML model to flag high-risk bookings.

7. Conclusion

The analysis provides clear evidence of cancellation patterns and their drivers. Insights and predictions from this study enable hotels to improve revenue management, reduce cancellations, and enhance customer satisfaction.