

# Predictive Modeling of Hotel Booking Cancellations

Group 6:

Khurem Qazi

Priyadarshan Parida

Ria Choksi

Soniya Rajappan

Sunayana Jana



# Executive Summary

## Problem Statement

High hotel booking cancellation rates over **35%**, and notably higher for city hotel, resulting in an estimated annual revenue loss of over **\$5.5 M.**

## Objective

Prepare a comprehensive marketing plan to proactively target guests at risk of churn and prevent revenue loss from cancellations.

## Research Questions

What are the primary drivers of cancelation?  
How effectively can we predict booking cancellation behavior?

## Hypothesis

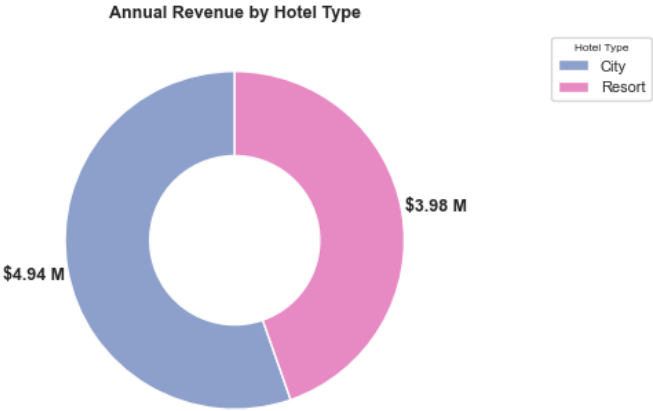
Cancellation likelihood is influenced by factors including **booking behavior** (e.g., lead time), **price sensitivity**, **deposit policies** (e.g., non-refundable), **customer type**.

## Data Insights - A Sneak Peek

- **Booking Horizon** (lead time)
- **Price Sensitivity**
- **Guest Geography** (domestic booking)
- **Guest Engagement & Experience** (room allocation accuracy, special request handling etc.)

Predictive Model with  
~90% Accuracy in  
identifying bookings at  
high risk of cancellation.

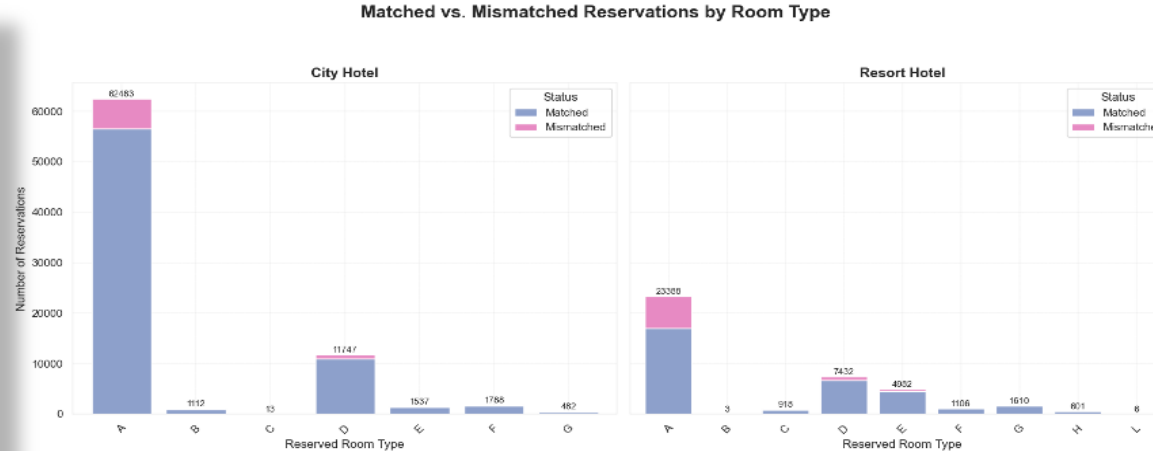
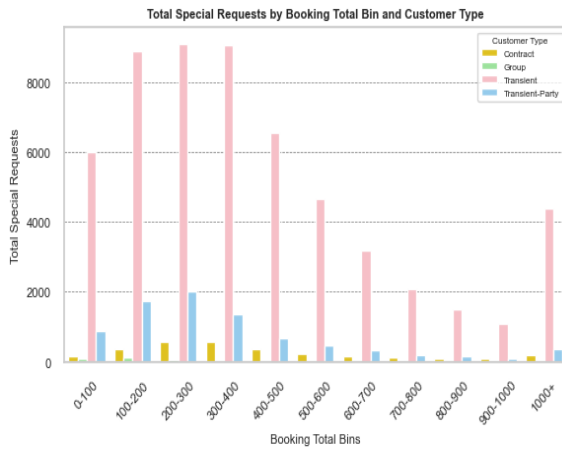
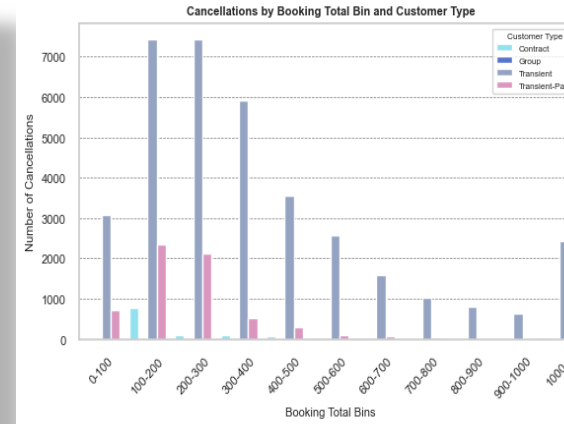
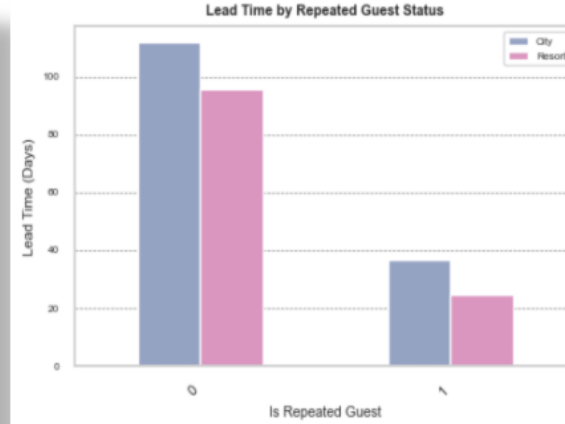
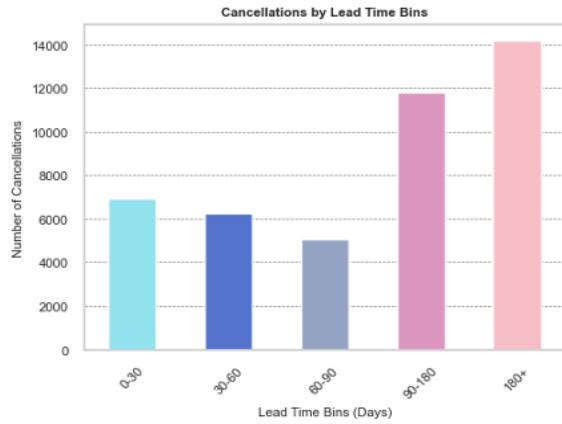
Hotel Type	Booked	Ratio of Total Reservations	Cancelled	Cancellation Rate
City	79162	65%	33078	42%
Resort	40046	35%	11120	28%
Total	119208	-	44198	37%





# Data Trends and Early Insights

## Key Insights



Higher cancellations with > 90 days lead time.

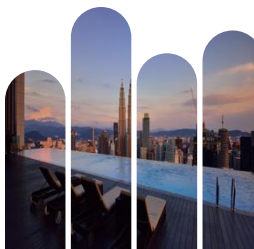
Repeat guests have lower lead time and cancel less.

More cancellations in \$100-500 booking total bins – price sensitivity.

Fewer special requests at higher booking total could indicate unmet service expectations.

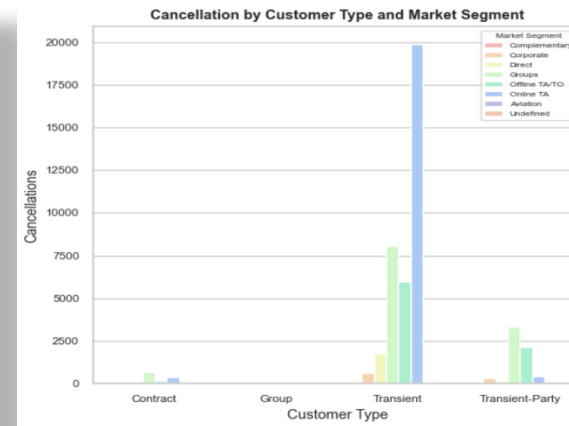
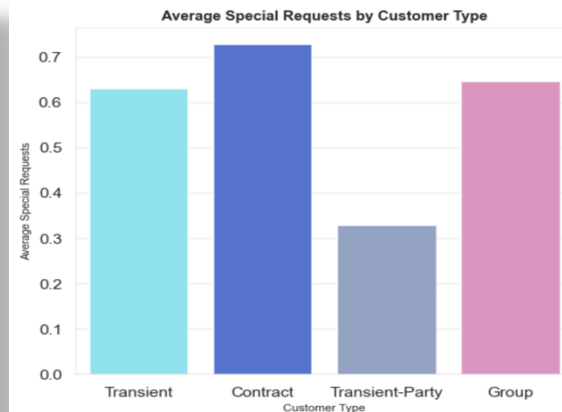
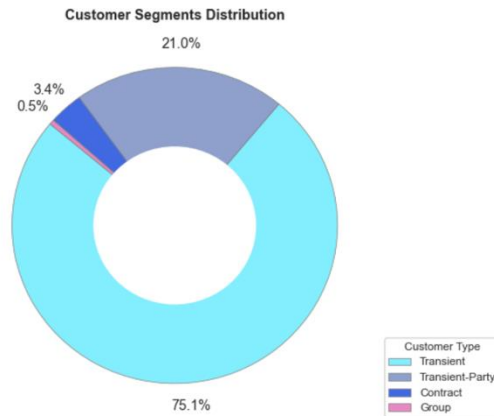
Special Requests positively [correlated](#) with booking total (ADR).

Room mismatch rate 12.4% overall. Maximum mismatches in Room Type A which is also the most booked.



# Data Trends and Early Insights Contd.

## Key Insights



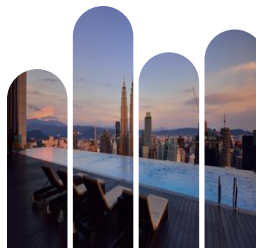
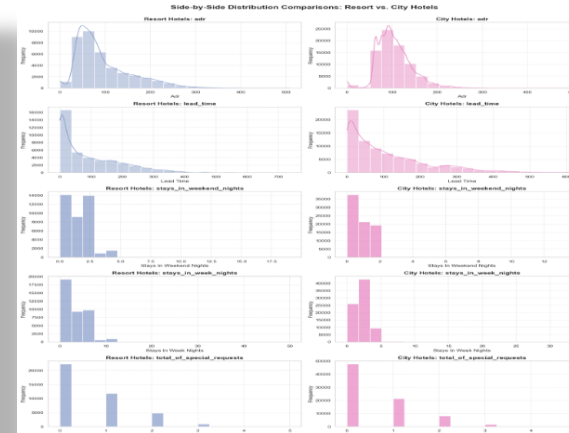
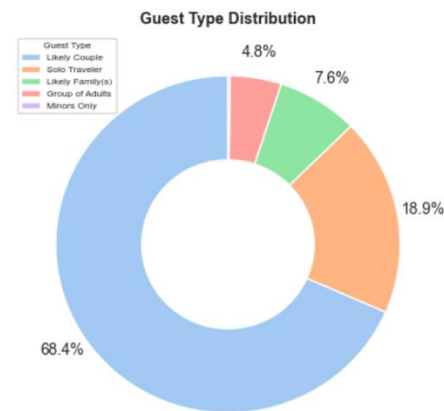
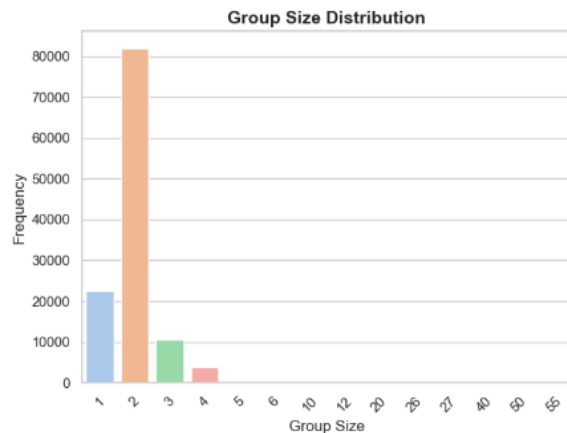
Majority are Transient customers (75%) followed by Transient-Party.

Transient-Party customers make fewer special requests indicating lower engagement/commitment.

Dominant customer type Transient exhibits most cancellations through Online TA channel.

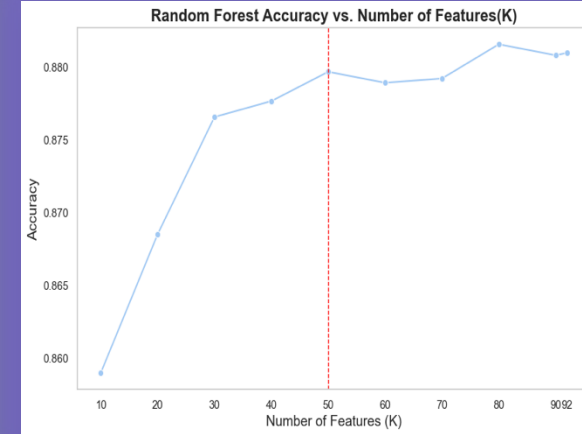
Group size of 2 (likely couple) represents ~70% of bookings.

Similar data trend between City Hotels and Resorts indicating amenability toward uniform marketing treatment.



# Interpretation of Model Outcome

	Interpretation / Insights	Directionality	Department to Intervene
Lead time	Longer lead times associated with higher cancellations likely due to tentative bookings	+	Marketing & Revenue Management
Booking Total (& Total Stay)	Higher booking values may lead to more cancellations due to greater financial commitment or unmet service expectations.	+	Revenue Management and Finance along with Marketing
Country of Origin	Domestic guests cancel more often likely due to higher flexibility to change plans.	+	Marketing
Special Requests	Special requests are associated with lower cancellation probably as it shows stronger guest engagement and commitment	-	Front Office and Guest Relations
Cancellation History	Past cancellation behavior predicts future cancellations	+	Revenue Management
Room Mismatches	Mismatches may cause dissatisfaction, increasing cancellation risk	+	Operations and Front Office
Customer Type & Channel	Transient Customers booking through Online TA are likely to cancel.	+	Marketing

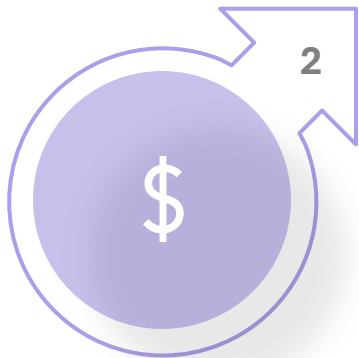


- **Accuracy 88%**
- 50 features selected based on importance scores
- Reduces model complexity, training time, and overfitting risk

# Data Driven Marketing Plan

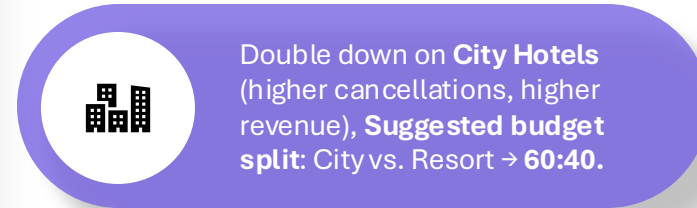


Early Birds

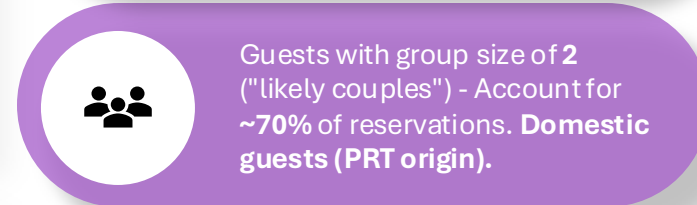


Value Seekers

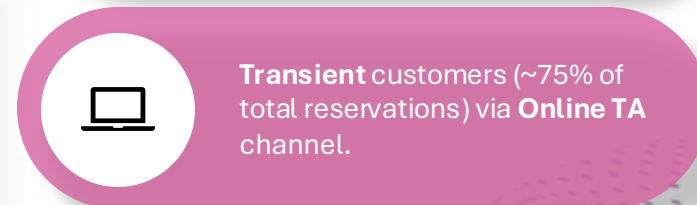
Problem	Supporting Insights	Action / Incentive
Long lead-times are associated with higher cancellations (tentative bookings)	<ul style="list-style-type: none"><li>Higher cancellations in lead time &gt;90</li><li>Repeat guests have lower lead times and cancel less</li></ul>	<ul style="list-style-type: none"><li><b>Early-bird discounts</b> and extra loyalty points for non-refundable bookings &gt;60 days out</li><li>Personalized <b>early-engagement</b> emails with attractions, toll tags, etc.</li><li><b>Post-stay engagement</b> with exclusive offers for high value guests</li></ul>
High booking totals may lead to more cancellations, likely due to <b>financial commitment</b> or <b>unmet expectations</b>	<ul style="list-style-type: none"><li>\$100–\$500 booking value - higher cancellations</li><li>Fewer special requests indicate lower engagement, and perceived value</li><li>Room mismatch (12.4%) - unmet expectations</li></ul>	<ul style="list-style-type: none"><li><b>Pre-stay engagement</b> through value-added offers: perks, early check-in, free breakfast</li><li><b>Loyalty</b> points for high value guests for completion of stays</li><li><b>Priority upselling</b> to reinforce commitment: e.g., <b>free spa access</b></li></ul>



Double down on **City Hotels** (higher cancellations, higher revenue), **Suggested budget split**: City vs. Resort → **60:40**.



Guests with group size of **2** ("likely couples") - Account for **~70%** of reservations. **Domestic guests** (PRT origin).



**Transient** customers (~75% of total reservations) via **Online TA** channel.



A wide-angle photograph of a rooftop infinity pool at dusk. The pool's edge is perfectly aligned with the horizon, creating a seamless transition between the water and the city skyline. In the foreground, four dark-colored lounge chairs with light-colored cushions are arranged on a dark wooden deck. The city skyline in the background features numerous skyscrapers, with the Petronas Twin Towers being the most prominent. The sky is a mix of deep blue and soft orange from the setting sun. The text "Thank you!" is centered over the image in a large, white, sans-serif font.

Thank you!



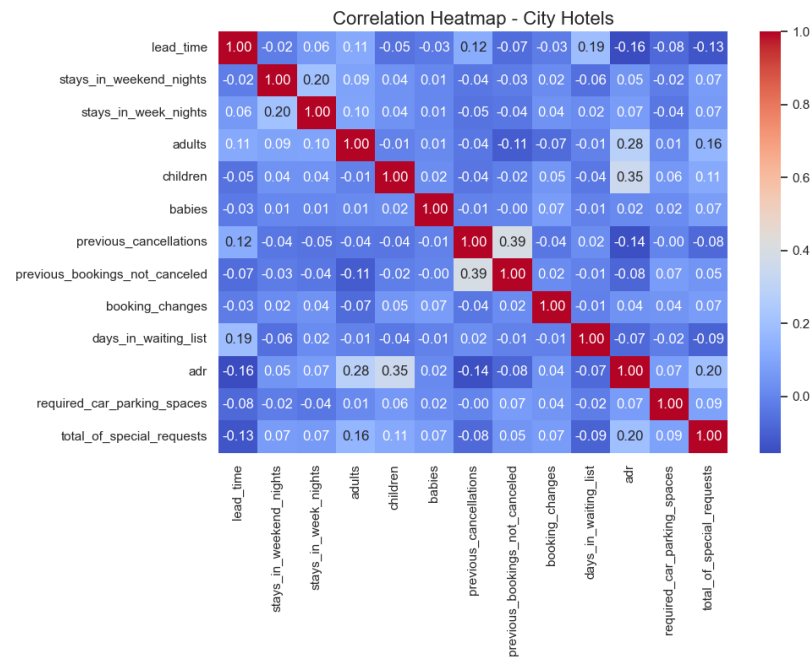
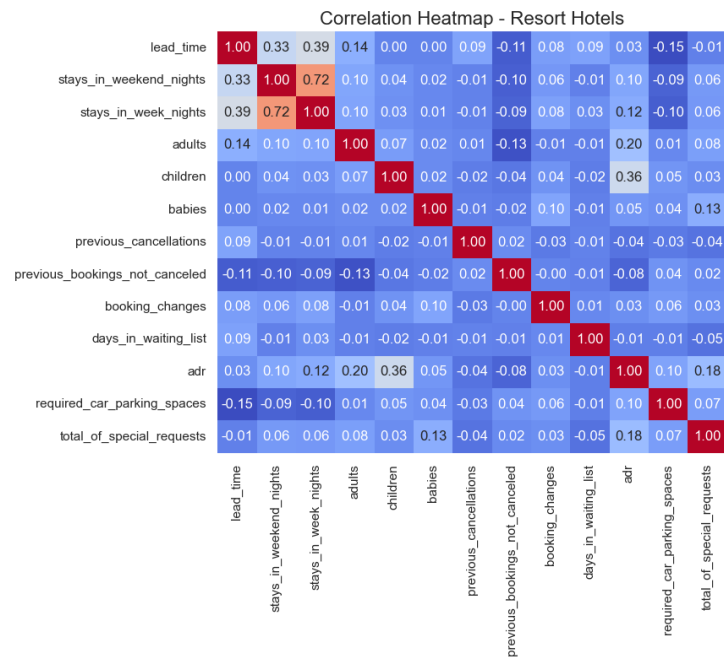
# Appendix

A wide-angle photograph of a rooftop infinity pool at dusk. The pool's edge is perfectly aligned with the horizon, creating a seamless transition between the water and the city skyline. In the foreground, four dark-colored lounge chairs with light-colored cushions are arranged on a dark wooden deck. The city skyline in the background features numerous skyscrapers, with the Petronas Twin Towers being the most prominent. The sky is a mix of deep blue and soft orange from the setting sun. The overall mood is serene and luxurious.



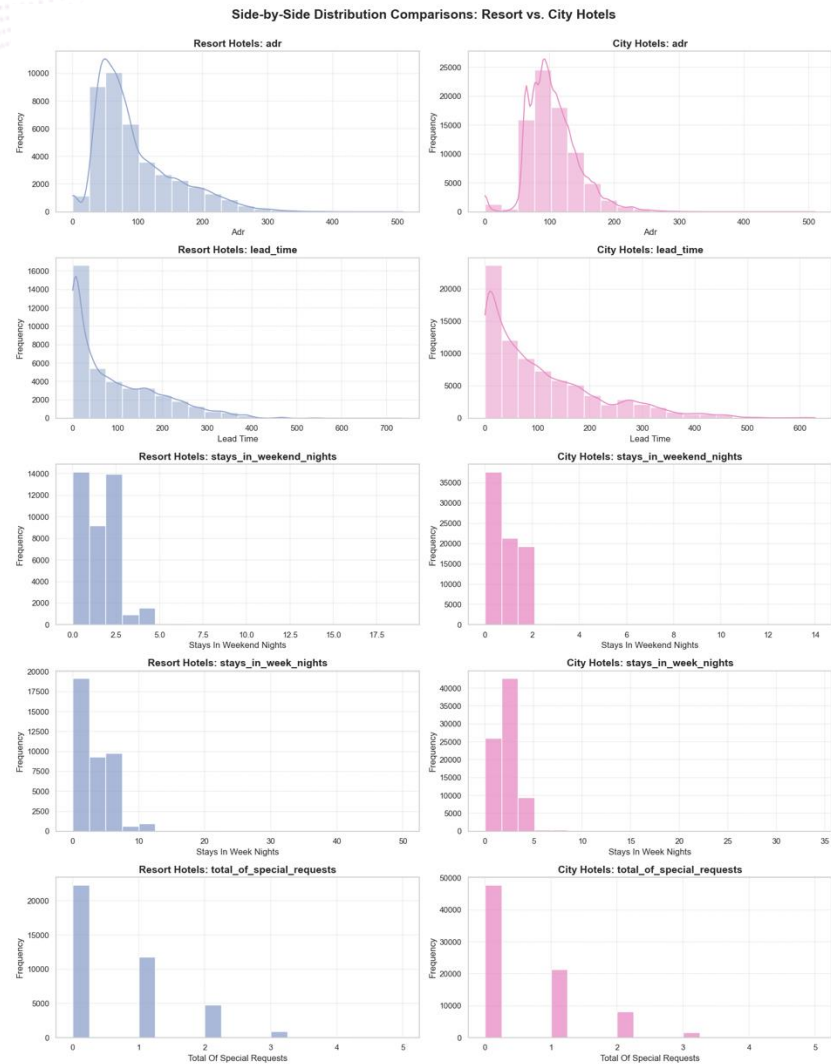
# Correlation Analysis

Feature Correlation



Back

# Side-by-Side Distribution Comparisons: Resort vs. City Hotels



[Back](#)