University of Austin Texas
PG-DSBA Certificate Program
Presenter: Uchenna Nwosu

InnHotels Presentation

Contents

Problem Overview and Objective

Business Description

INN Hotels Group has a chain of hotels in Portugal, they are facing problems with high number of booking cancellations. A significant number of bookings are called off as well as no-shows. The typical reasons include changed plans, scheduling conflicts, as well as customer specific behavior. This is made easier by the option to cancel for free or with little penalty. The cancellation of bookings negatively impact:

1. revenue when the hotel cannot resell the room.
2. marketing costs by increasing commissions paid for publicity to help sell these rooms.
3. profitability because last minute price reductions are necessary to sell the room in short notice.
4. personnel costs because extra hours are required to service guests.

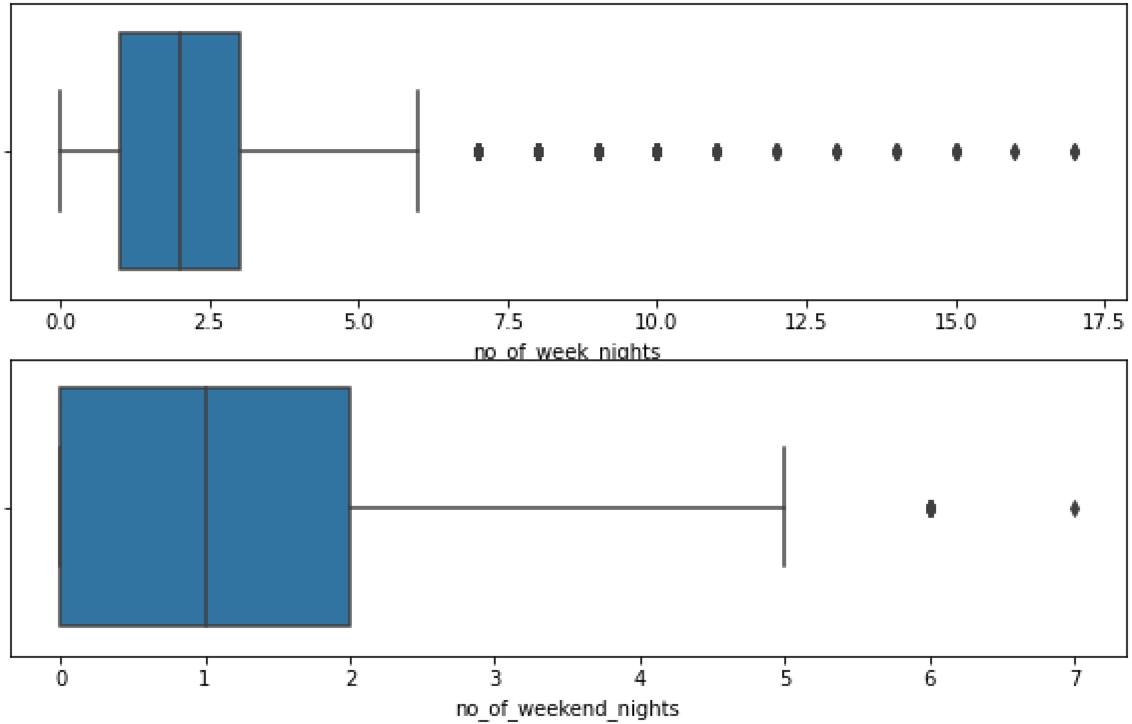
Objective

To develop a Machine Learning solution that can help predict cancellations. Identifying them in advance will help in implementing contingency and cancellation plans.

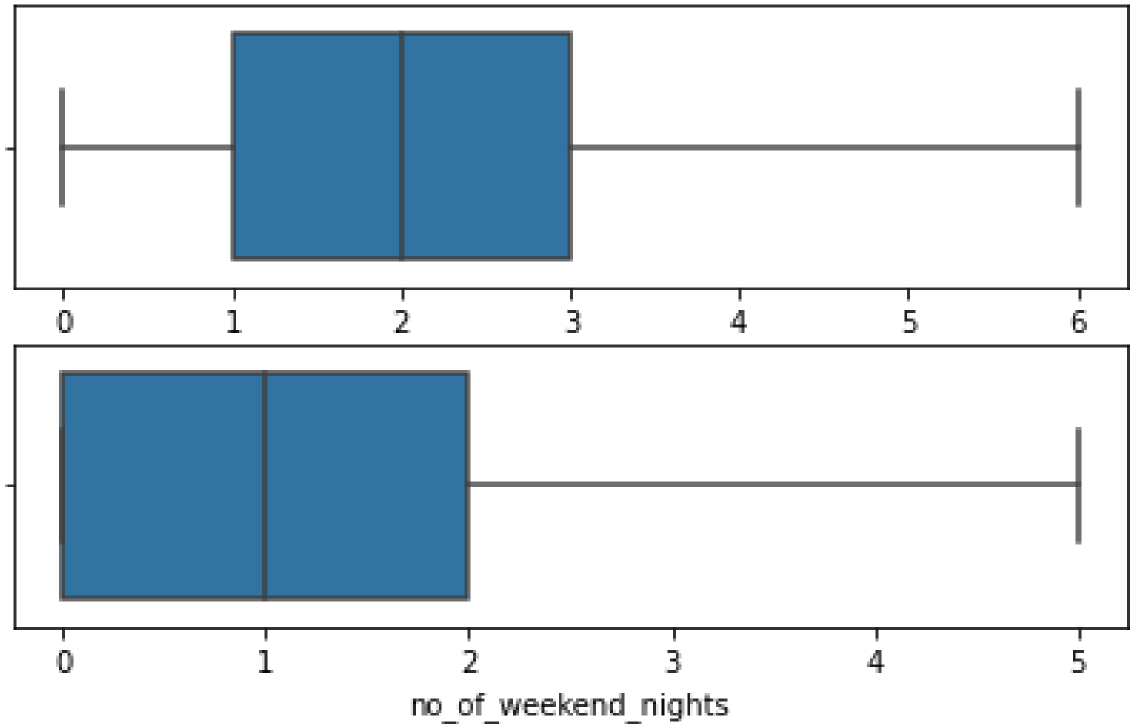
Data Overview

Booking_ID: the unique identifier of each booking
no_of_adults: Number of adults
no_of_children: Number of Children
no_of_weekend_nights: Number of weekend nights no_of_week_nights: Number of weeknights
type_of_meal_plan: Type of meal plan booked by the customer:
required_car_parking_space:
room_type_reserved:
lead_time: Number of days till arrival date
arrival_year: Year of arrival
arrival_month: Month of arrival
arrival_date: Day of arrival
market_segment_type: Market segment designation.
repeated_guest: Is the customer a repeated guest?
no_of_previous_cancellations: Number of previous cancellations by the customer
no_of_previous_bookings_not_canceled: Number of previous bookings not canceled
avg_price_per_room: Average price per day of the reservation
no_of_special_requests: Total number of special requests made by the customer
booking_status: Flag indicating if the booking was canceled or not.

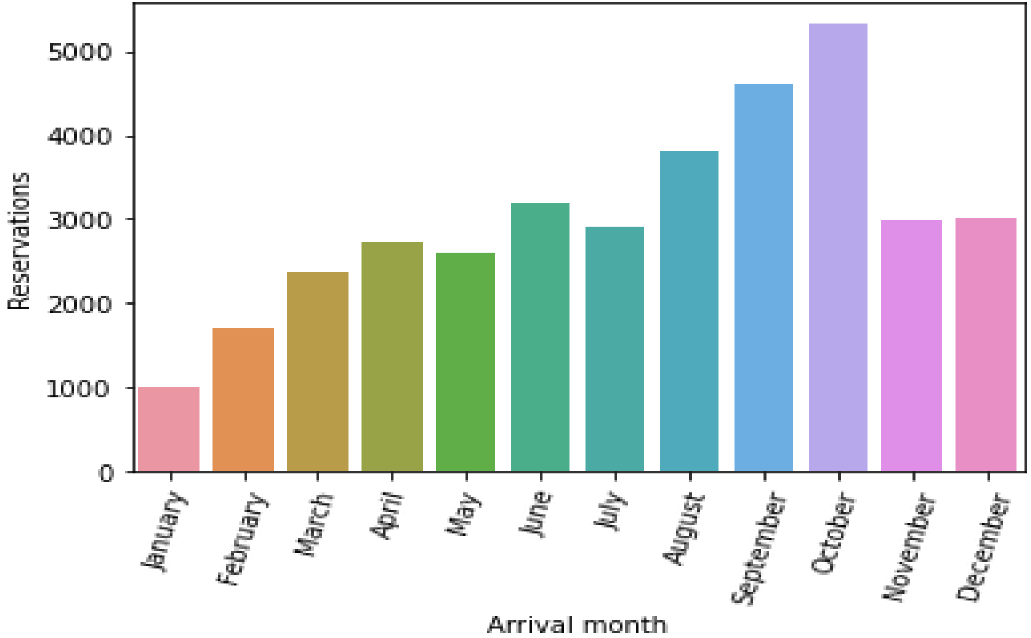
EDA: Before outlier treatment



EDA: After outlier treatment



EDA: Reservations by month

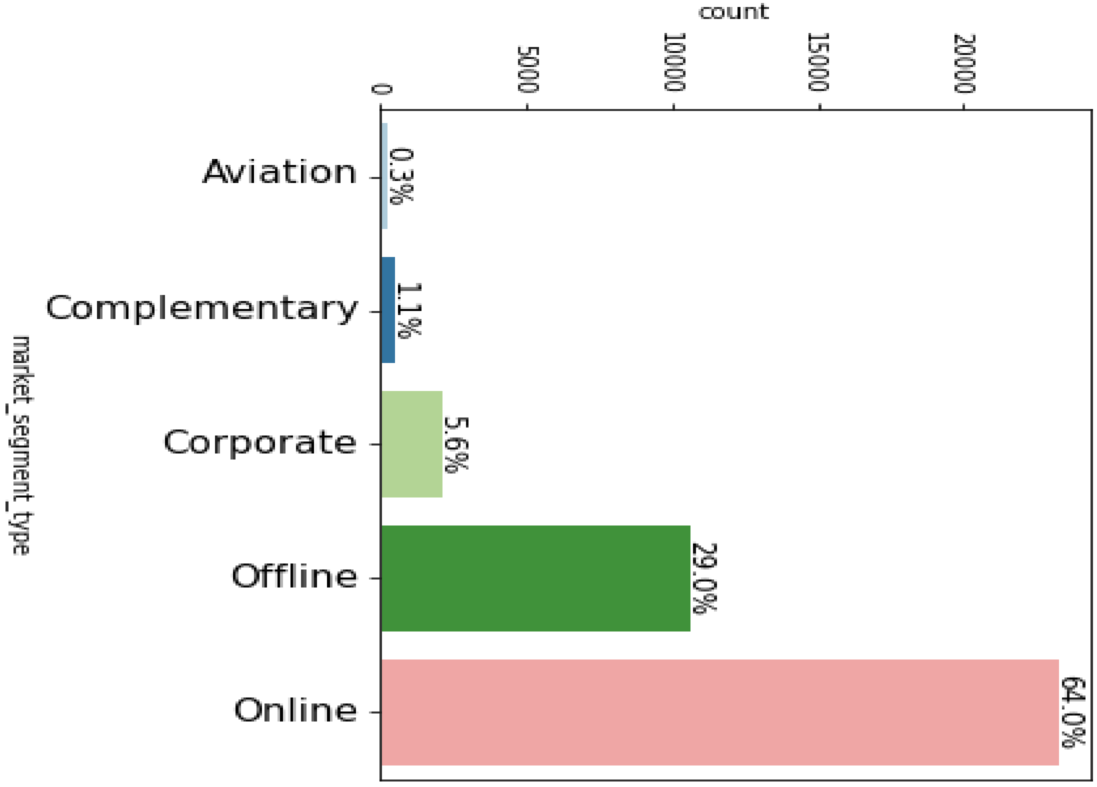


Observations:

October witnesses the most reservations

There is a gradual increase in reservations up till October then there is a gradual decline

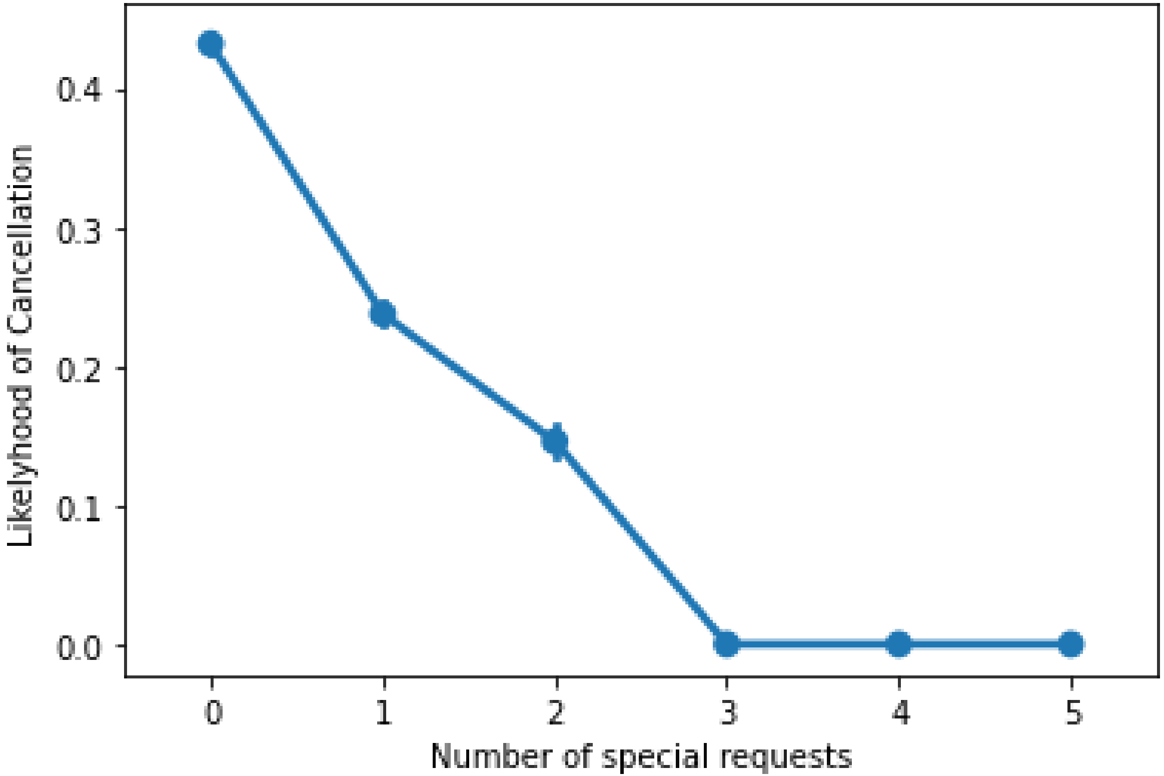
EDA – Market segment percentages



Observations:

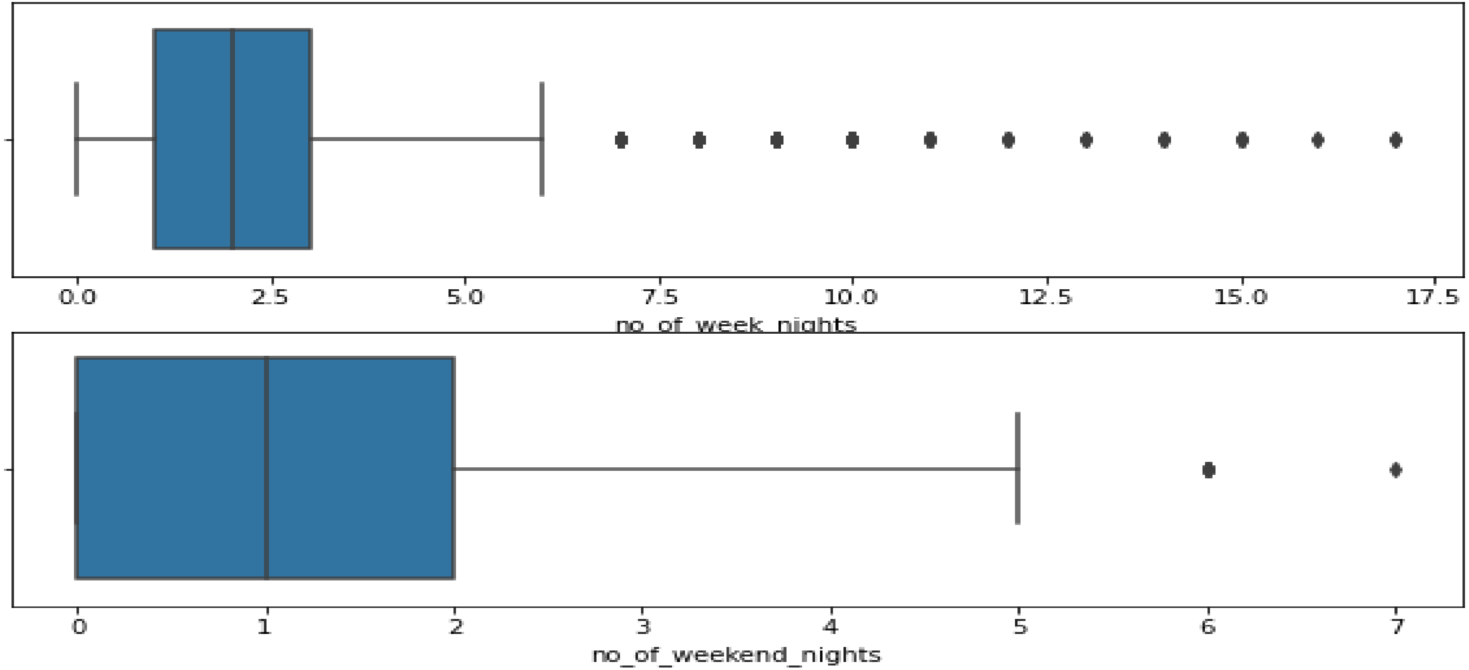
The online market accounts for more than half the reservations

EDA: Cancellations by Number of special requests



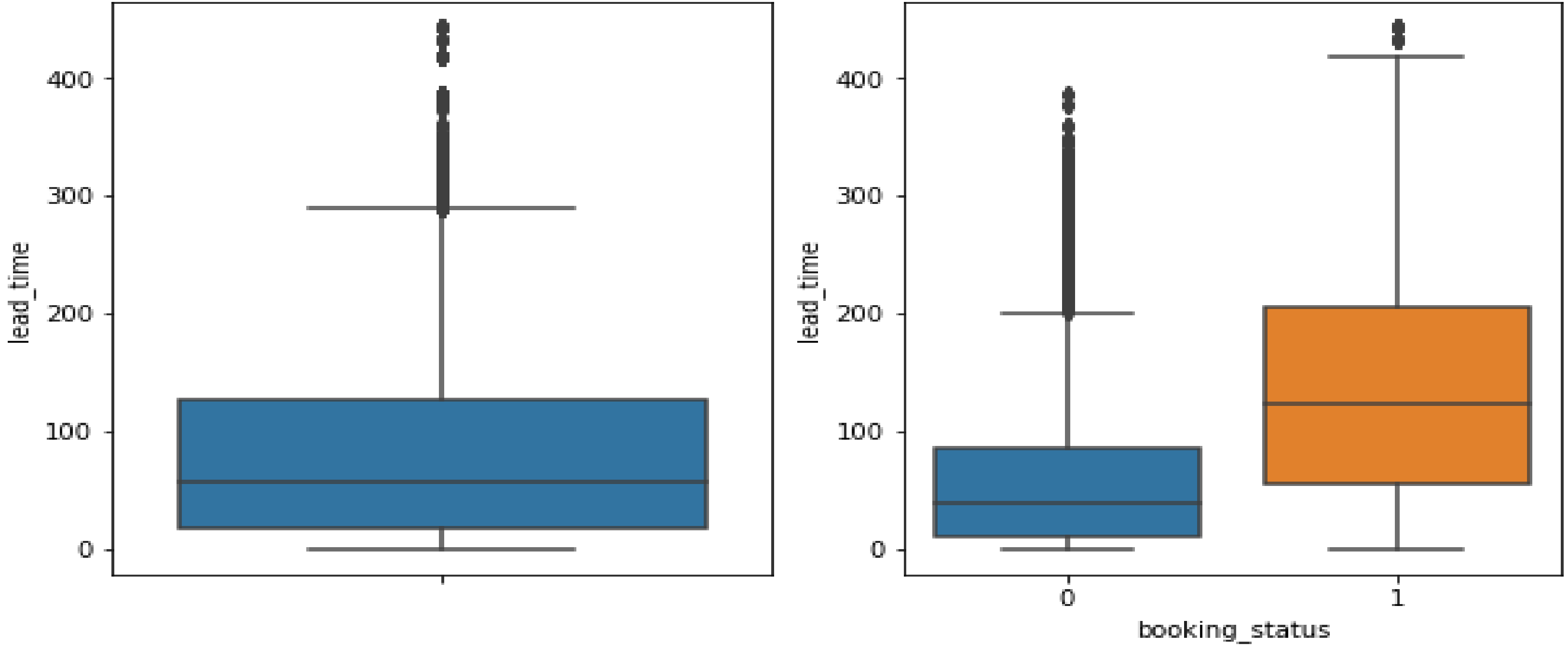
The likelihood of cancellations reduces with the number of special requests made by the customer. Customers that make 3 or more special requests tend not to cancel their reservations

EDA – Weekend vs Weekday patronage



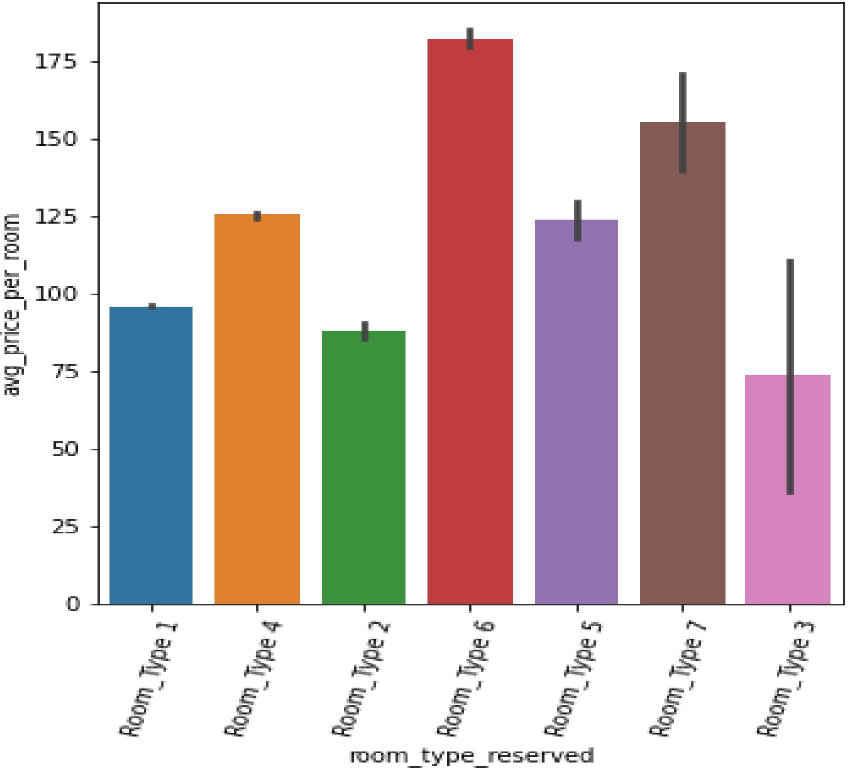
The average room stay is one day for weekends and about two days for weekdays.

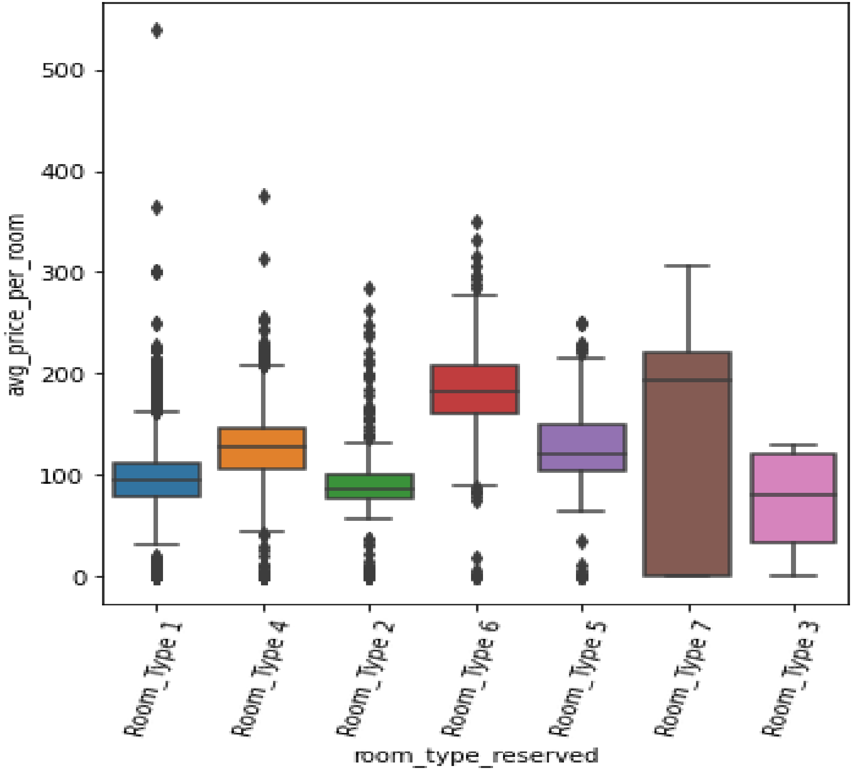
EDA – lead time vs booking status



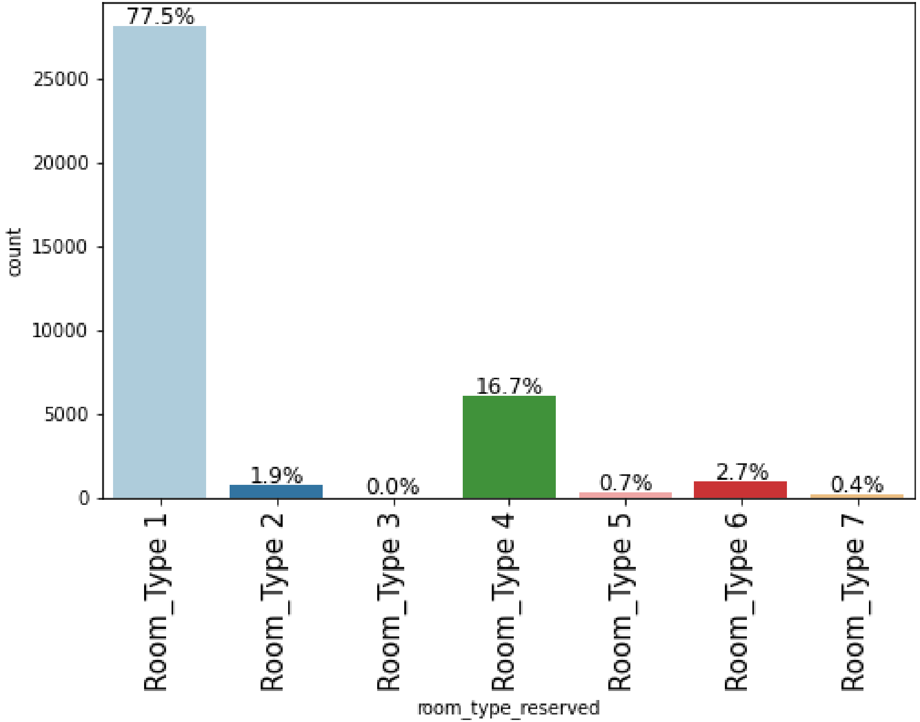
Cancellations are more likely as lead time increases.

EDA: Room type by Price

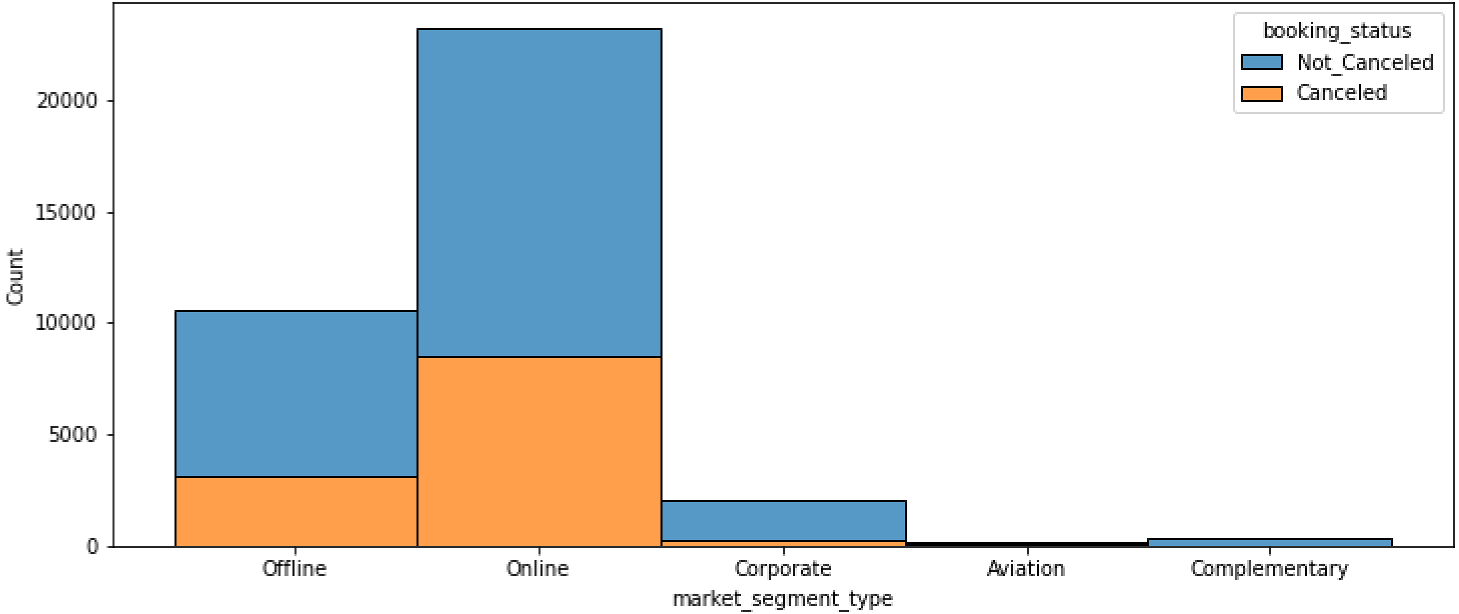




EDA – Room type by popularity

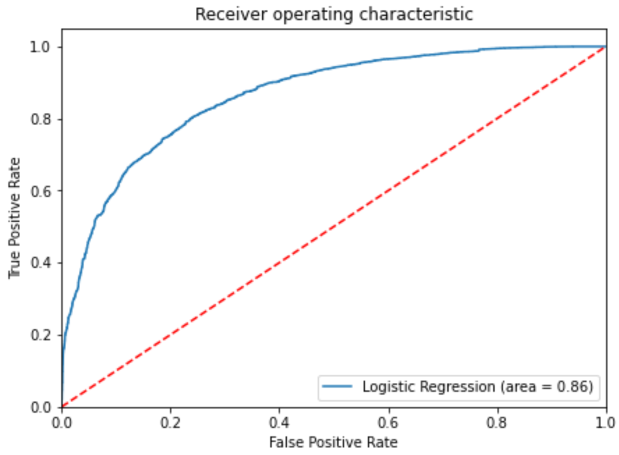


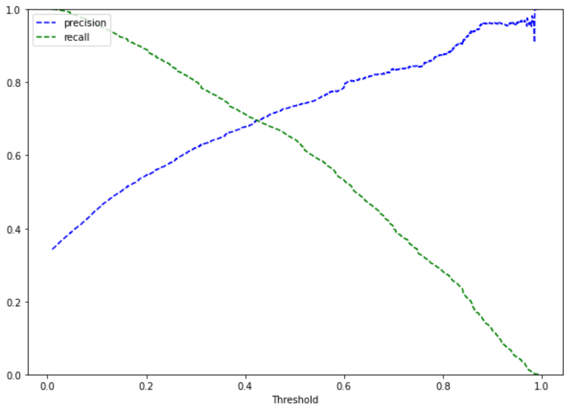
EDA: Booking status by market segment

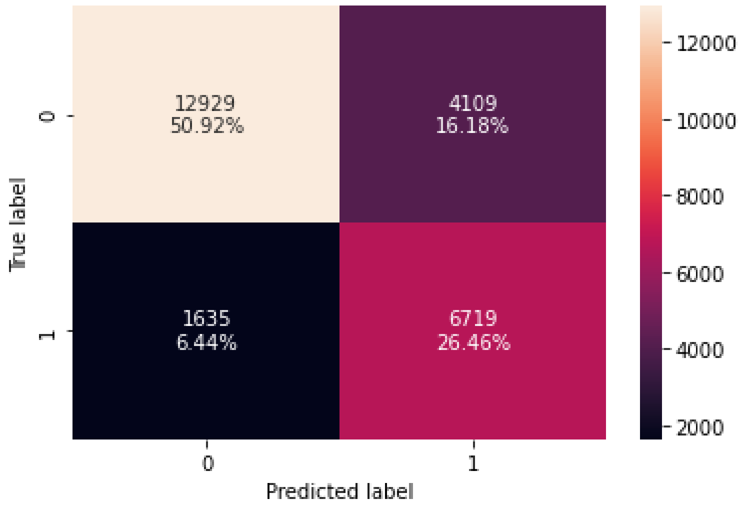


There is relatively little or no cancellation in the corporate market

Logistic Regression: Performance

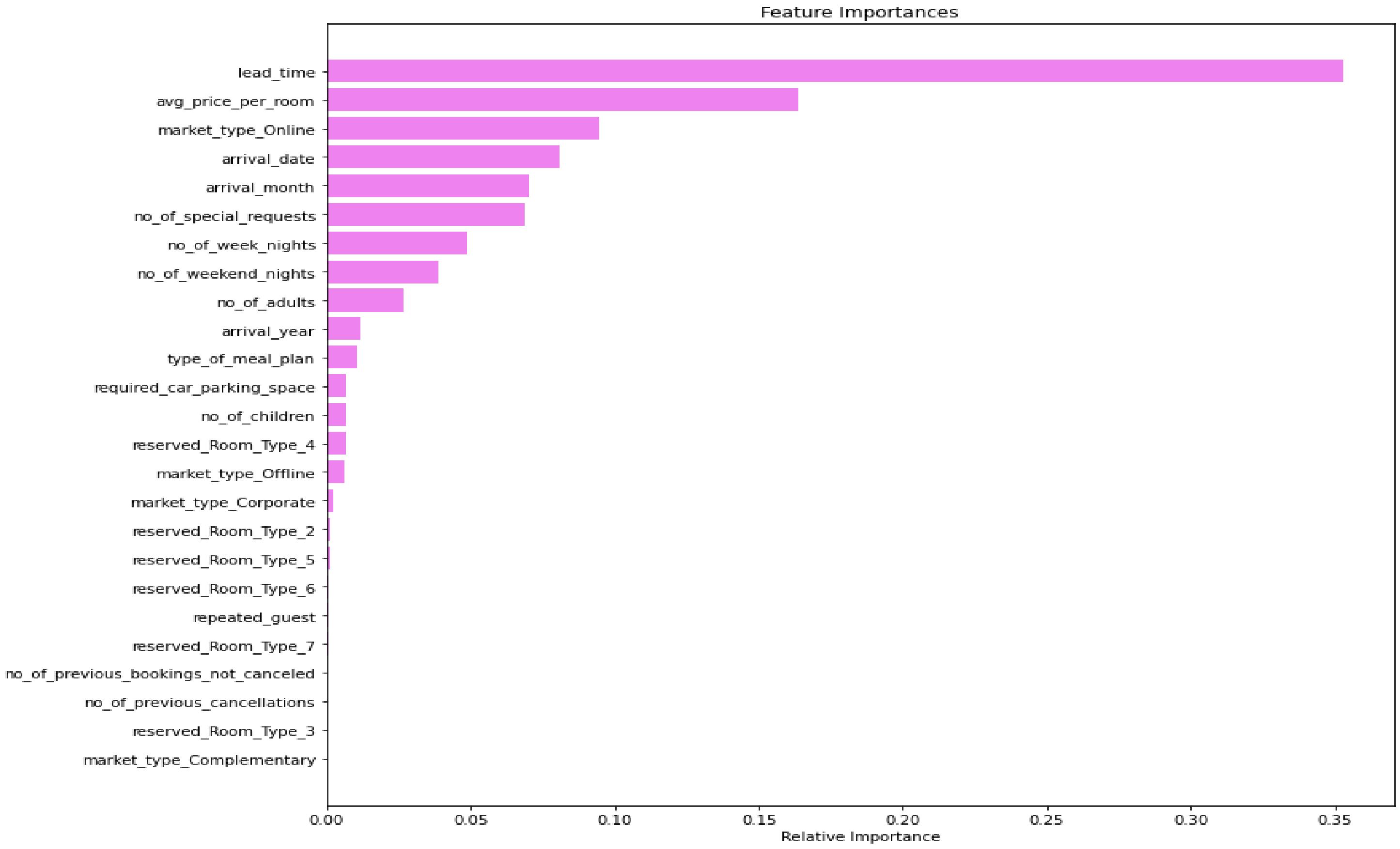


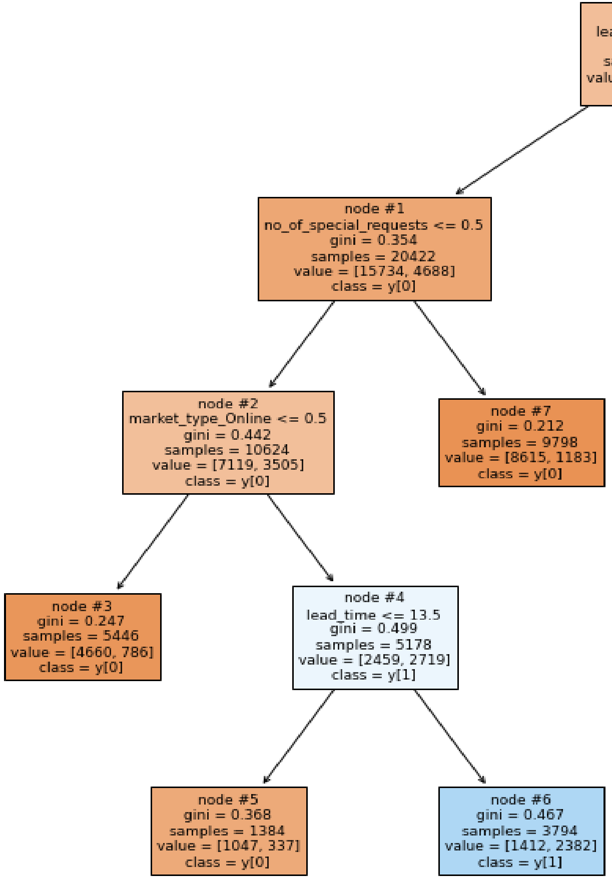




Accu.	Recall	Prec.	F1	Model
0.76	0.78	0.60	0.68	Logit
0.80	0.73	0.69	0.70	D. Tree

Decision Tree model feature importances





Generated decision tree showing the most important features selected in the model using GridSearchCV

Insights

- Overwhelmingly, the greatest factor that influences cancellations based on the adopted models is lead time.
- The greater the time between the booking and the checkin date, the greater the likelyhood of a cancellation.
- Our logistic regression model enables estimation of the likelyhood of a cancellation given lead time:
- A unit change of 0.229 in lead time will increase the odds of a cancellation by 1.017 times - or a 1.68% increase in odds. Therefore, a lead_time of 1 day will result in over 5% increase in odds. It would be instructive to encourage reservations as close as possible to the checkin date. The EDA of lead time revealed that most cancellations occurred above the median number of days - 57 days.
- The online market is the second most influential feature identified by our decision tree model. Reservations made online are more likely to get canceled than others.
- A hotel room reserved online is 5.53 times more likely to be canceled.
- A unit change in average room price will increase lead to a 2.21% increase in the odds of a cancellation.
- A unit increase in the number of special requests leads to a decrease in the odds of cancellation by 77%

Both EDA and machine learning models revealed that customers who make multiple special requests do not cancel. This is probably an indication of their level of seriousness. Corporate customers rarely cancel. Online customers are constantly shopping for better deals and will cancel whenever they find better options. These are the category of customers that must be handled with caution.

Final recommendations

Great Learning
POWER AHEAD

Penalize cancellations that are last minute with a non-refundable 20% deposit - exempt customers that maintain a track record of no cancellations

Long-term reservations should be either discouraged or allowed at a premium price. Free cancellation of long-term reservations available only within a company determined time window. For example, reservations whose lead times exceed 50 days should be canceled 2 weeks before check in. This gives enough window for the room to be resold.

Long-term bookings must be paid in full by a certain date before checkin and will be subject to hefty penalty before a refund is issued.

provide incentives for customers not to cancel. For instance, start a loyalty program. Offer complimentary reservations for customers in good standing.

Find out and log the specific reasons for cancellation: for example, if customer has found a better deal, verify and match it in order to prevent loss of the business.

Encourage non-corporate customers to book their rooms as close as possible to their check in dates

INNHotels must diversify their customer base so that they are not as heavily dependent on the online market. Market to families. The number of reservations being made for children is low. They also need to market more to the aviation market that has tourists, more apt to fulfill their bookings.

Room turnover rate is high between 1 and 2 days for weekends and weekdays, respectively, leaving a corresponding need to get new reservations. INNHotels must improve competitiveness by improved offerings both to get more bookings, and to motivate longer stay.

There is presently data for only repeated guests who are less than 3% of the reservations in the sample. For this reason, the importance ranking in the decision tree model for these related to cancellation behavior was insignificant.

greatlearning
Power Ahead

Happy Learning !