

ECE 143 Final Presentation

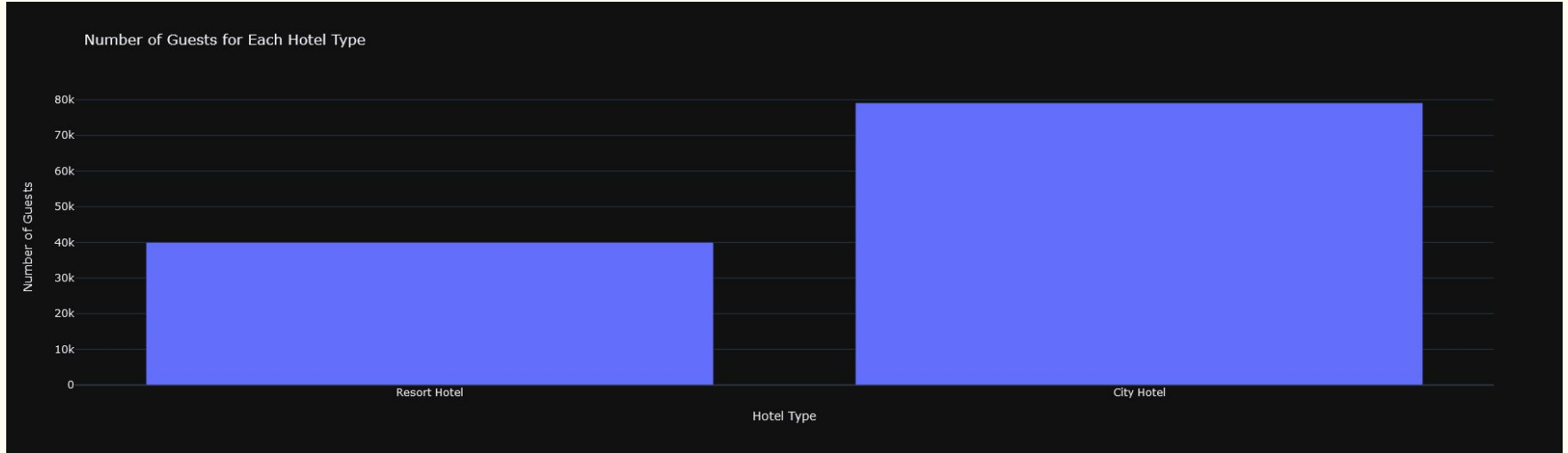
Hotel Booking Data Analysis

Group 6 : Jiajie Yu , Yidan Wang, Xinyi Liang, Varun Surapaneni

Dataset Overview.

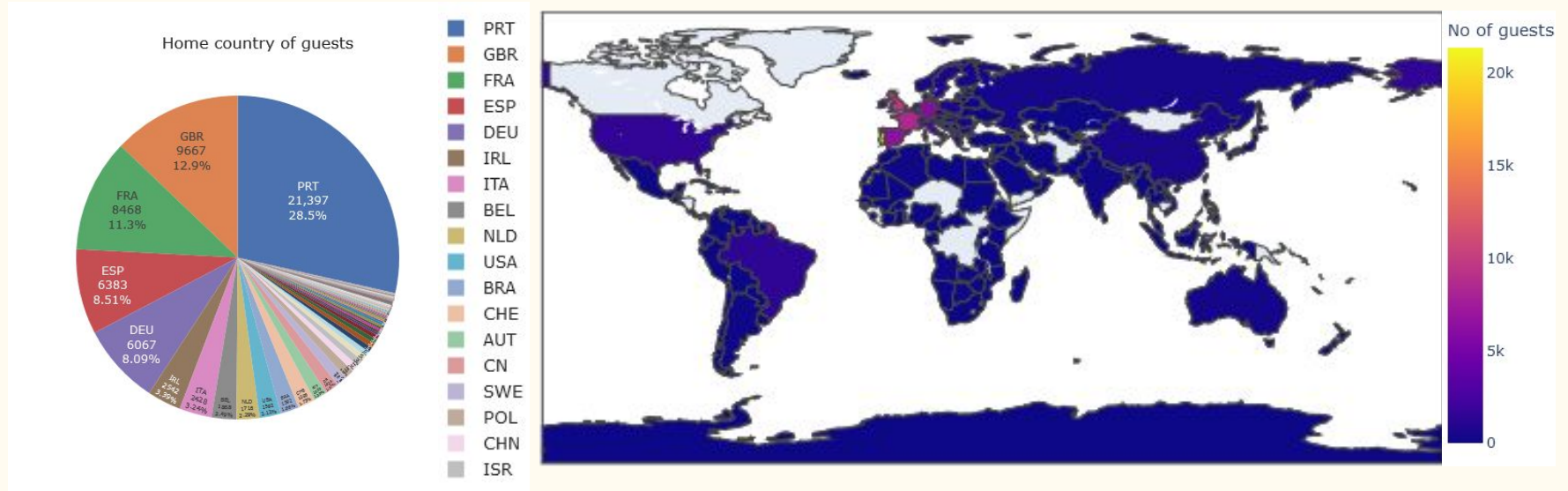
- 2 hotels considered a city hotel and a resort hotel.
- Duration of booking, lead time before booking, cancellations, number of people, any special accommodations.
- Market segment of guest, guest's previous statistics, deposit type, reservation, agent.
- ADR: Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights

Some Exploratory Data Analysis.



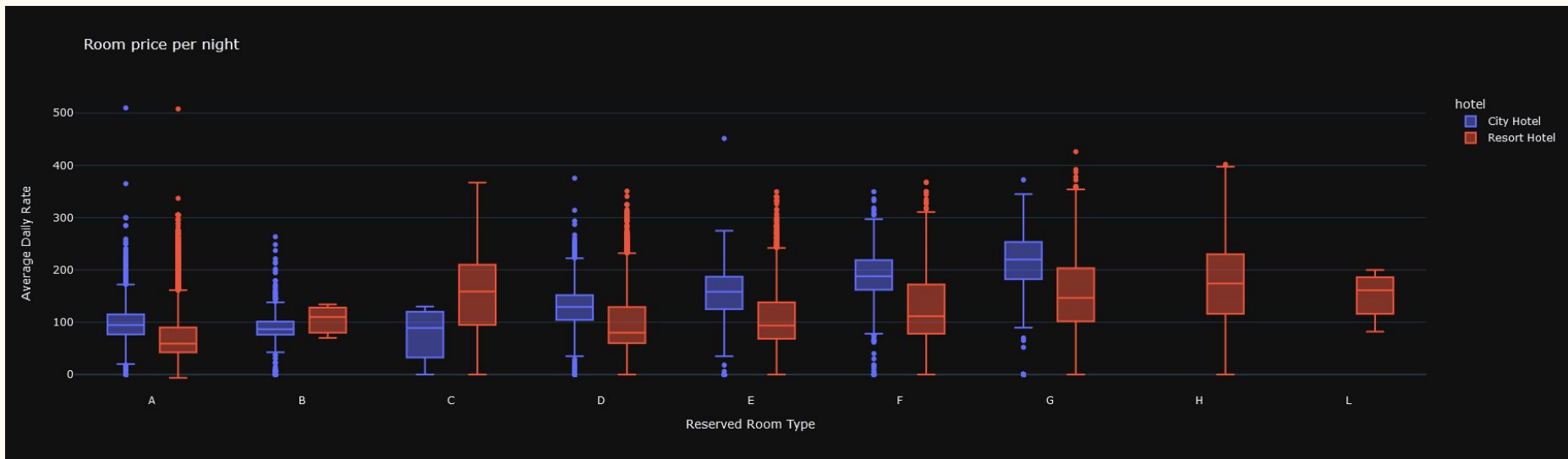
- Resort Hotel sees much less traffic in terms of numbers as compared to the city Hotel. Consistent with real life trends.

Guest Source Analysis



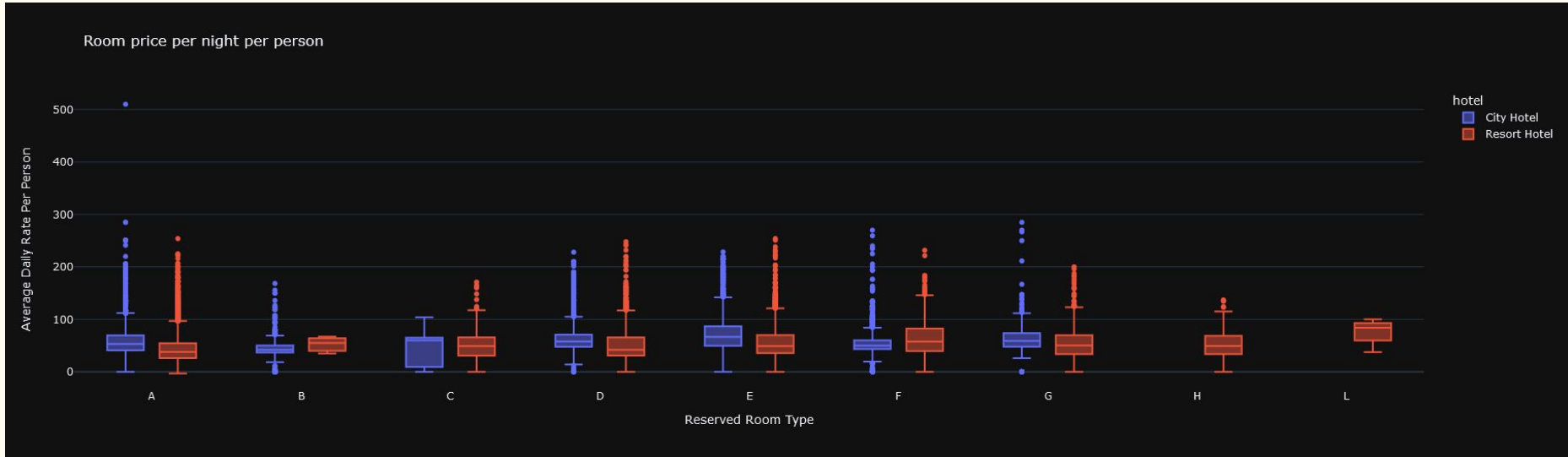
Destination is most popular among European Guests. Other countries provide a small number of guests. Strong travel between location and countries near Portugal suggesting a lot of ties between the places.

Room price per night



- Resort Hotel rooms have more consistent room pricing as they have a narrower interquartile range.
- City Hotels generally exhibit a higher median ADR across most room types, with the exception of room type B.

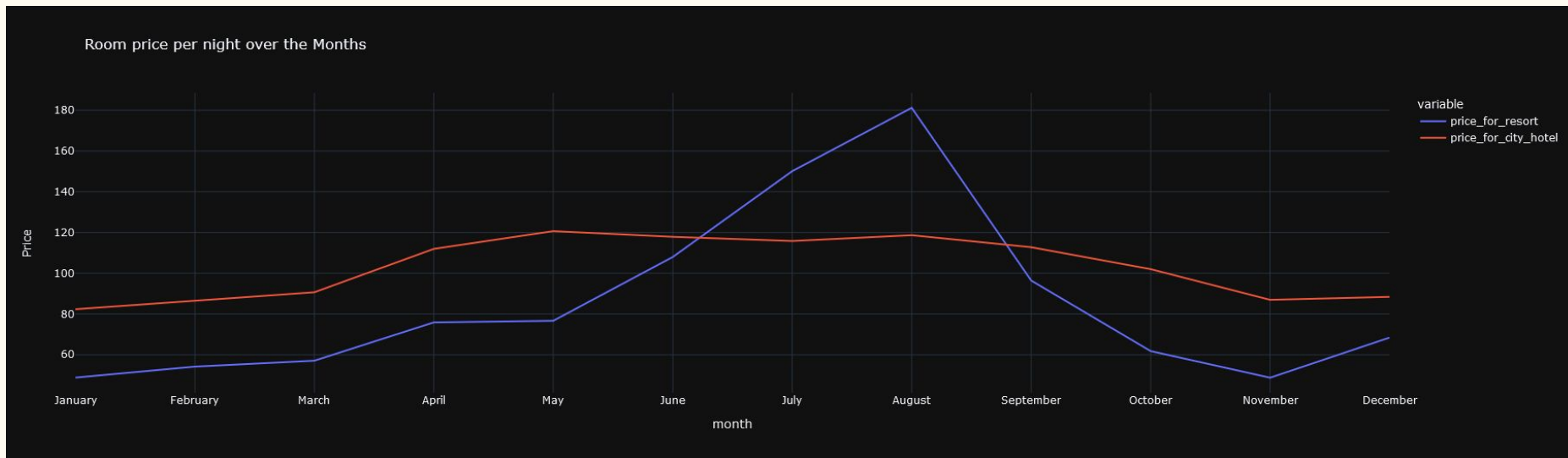
Average Daily Rate per Person



Comparison:

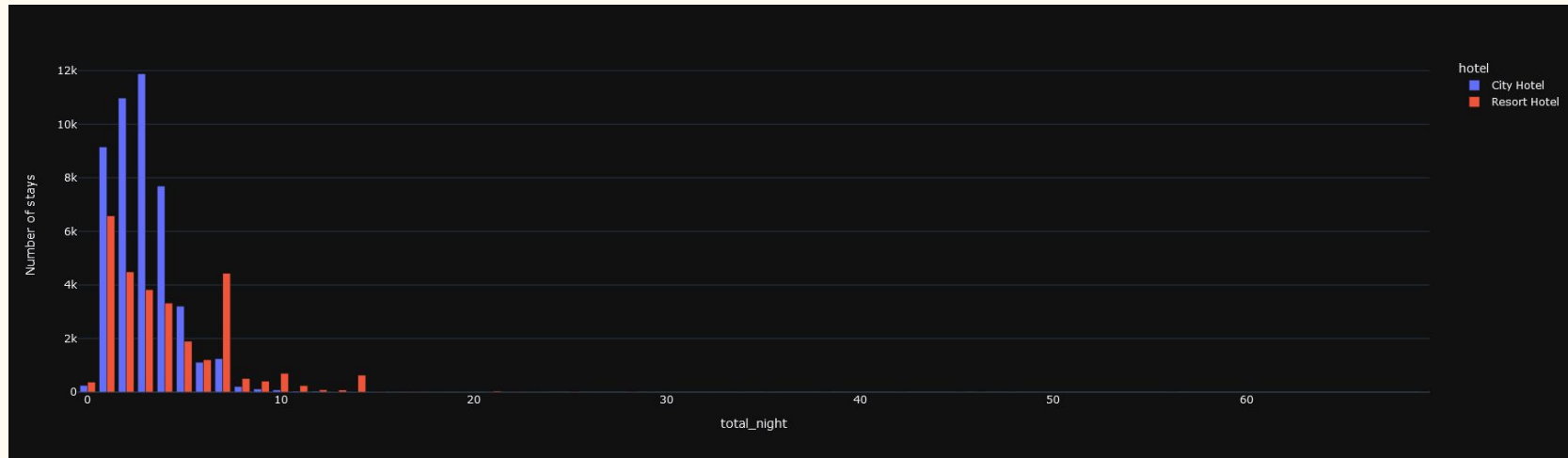
- City Hotel: Typically charges more per room type, with greater price variability, potentially catering to a wider range of customers or offering more varied room qualities. Seems more catered to business guests.
- Resort Hotel: Shows less variability and fewer extremes in pricing, which may appeal to guests looking for predictable pricing. Seems catered more towards leisure guests.

Room price per night throughout the year.



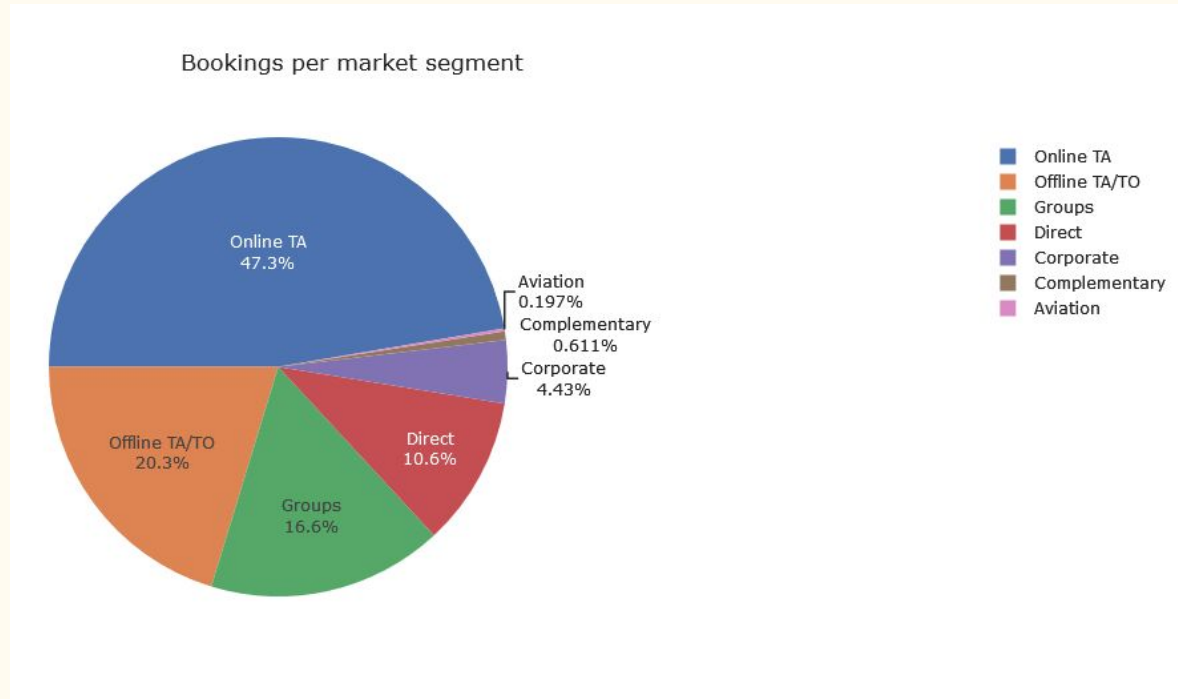
- **Seasonal Trends:** Prices for resort hotels peak sharply in August, which is likely due to the summer vacation period when demand is high. City hotel prices also rise during the summer months but do not exhibit as sharp a peak, suggesting a more consistent demand throughout the year or possibly a different pricing strategy that doesn't capitalize on seasonal spikes as aggressively.
- **Off-Peak Pricing:** Both types of hotels show a decrease in prices starting from September, with the lowest prices in November and December. This off-peak period may present an opportunity for promotions and special offers to attract guests when demand is typically lower. This graph also gives metrics for seasonal staff hires and shift workloads.

Duration of Stay



- Shorter Stays at City Hotels
- Longer Stays at Resort Hotels

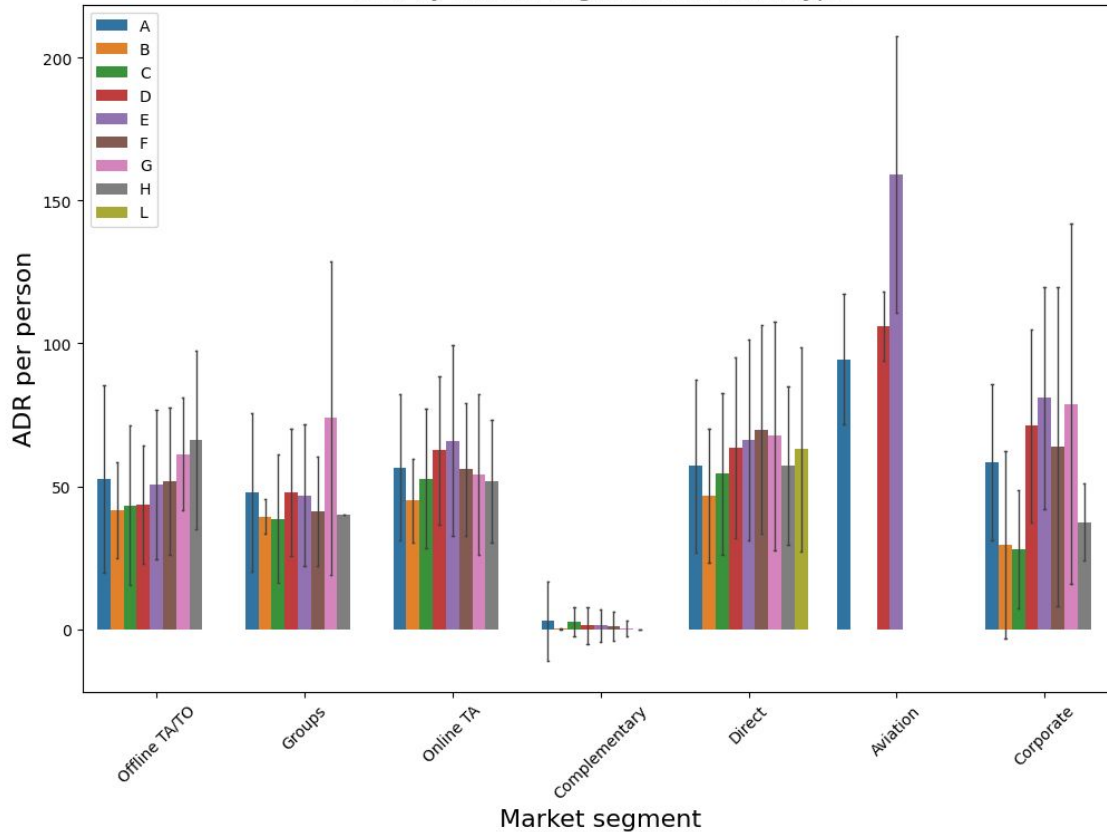
Booking per market segment



TA - Travel Agent

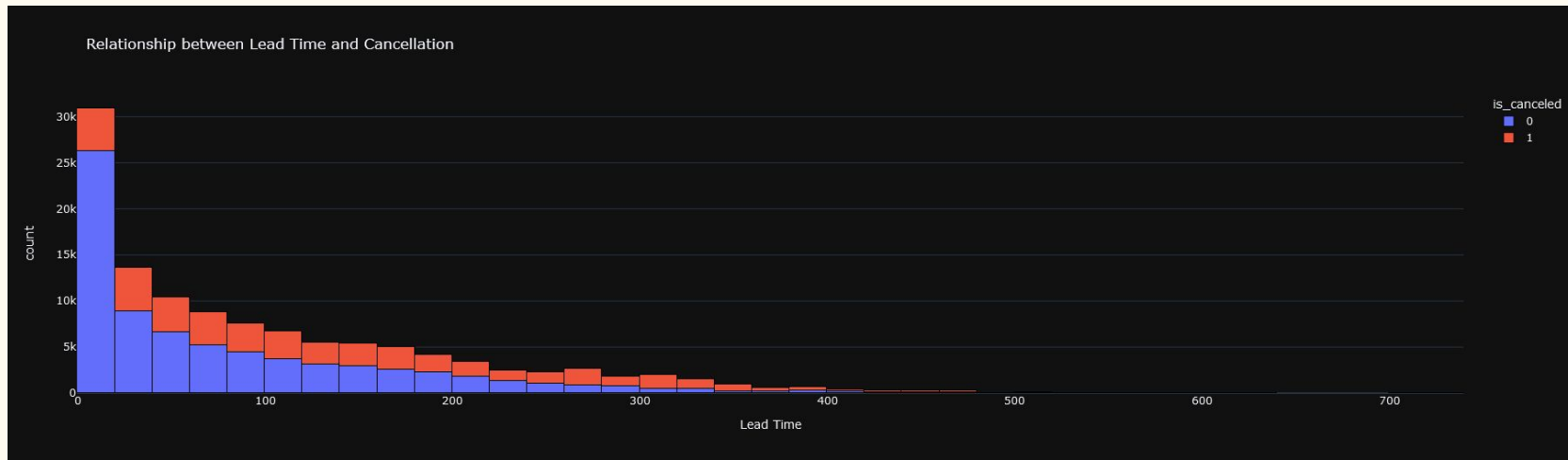
Booking per market segment

ADR by market segment and room type



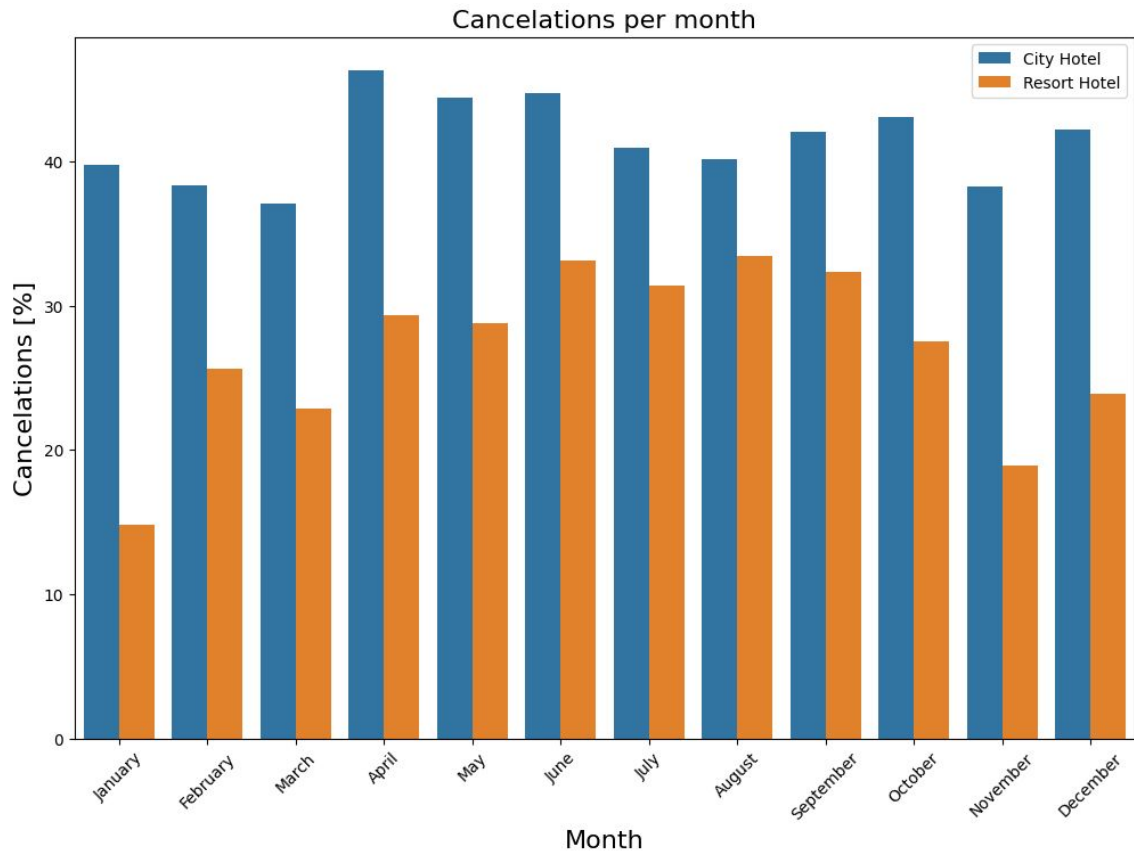
- Market Segment Variation: The 'Aviation' and 'Corporate' market segments have some of the highest ADRs, which could indicate a willingness to pay more among these travelers, possibly due to the costs being covered by businesses or the airlines themselves.

Cancellation Insights



- **Cancellation Trends:** Bookings with a shorter lead time (close to zero) have a lower likelihood of being canceled, as indicated by the red bars (canceled bookings) being shorter in that range. As lead time increases, the proportion of canceled bookings (red) also increases, suggesting a trend where bookings made well in advance are more likely to be canceled.
- **Booking Behaviour:** There are fewer bookings with long lead times, but these bookings show a higher relative proportion of cancellations.

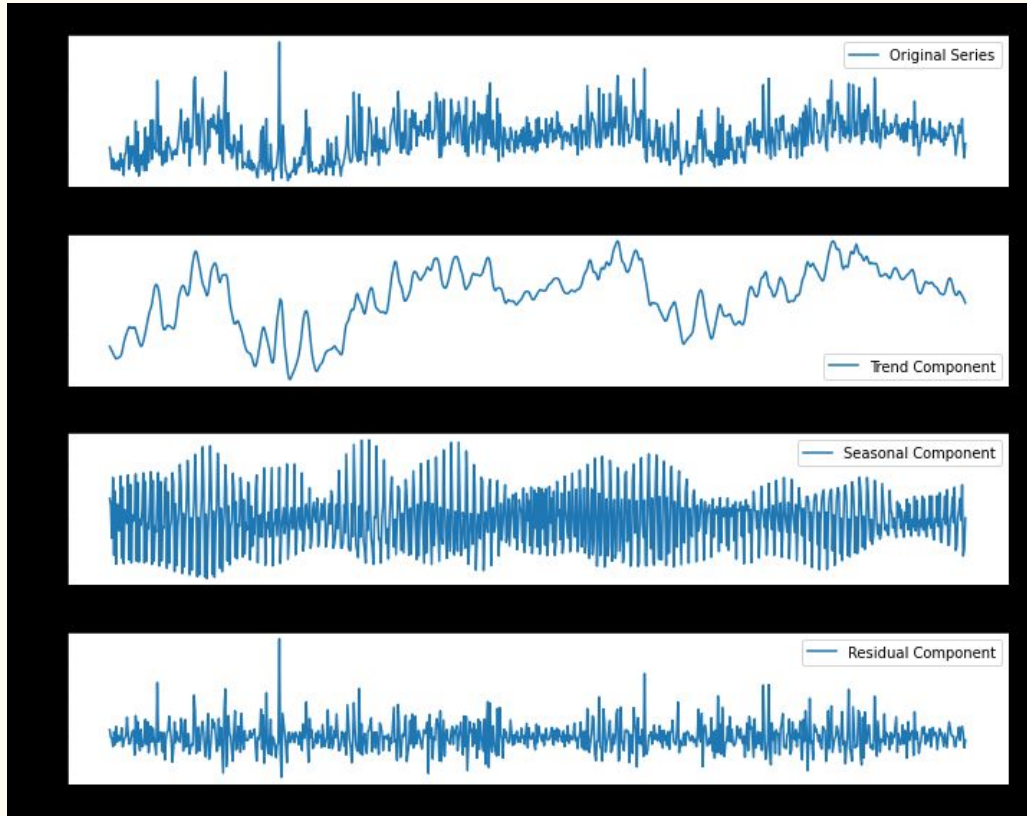
Cancellations per month



Seasonal Variation: Cancellations for the City Hotel are highest in the middle of the year, around June and July, as well as in November. For the Resort Hotel, cancellations peak during the summer months, with the highest in August.

City Hotel Cancellations: The City Hotel has a higher percentage of cancellations compared to the Resort Hotel for almost every month. This could be due to the nature of city hotel bookings, which might be more prone to business-related changes or short-term travel adjustments.

STL Decomposition(Seasonal and Trends Decomposition)



Y-axis : Total number of hotel booking per day.

X-axis: Date of the year from 1 Jan to 31 Dec

- Isolates Seasonal,Trend and Residual Component
- Seasonal Shows peaks in March,Sept,Nov and December.

Random Forest Classifier

Training Accuracy Score of Random Forest is : 1.0

Testing Accuracy Score of Random Forest is : 0.9547666185698467

Confusion Matrix :

```
[[18608   126]
 [ 1222  9845]]
```

Classification Report :

	precision	recall	f1-score	support
0	0.94	0.99	0.97	18734
1	0.99	0.89	0.94	11067
accuracy			0.95	29801
macro avg	0.96	0.94	0.95	29801
weighted avg	0.96	0.95	0.95	29801

Key Takeaways

As a Hotel Owner:

- Knowing the peak cancellation months give you more room for overbooking, better data for hiring temporary staff.
- Off-peak season gives opportunity for renovations, reduced staff, steeper discounts and promotions.
- Knowing previous history of guest can predict cancellation percentage and also be used to reward loyalty as desired.

As a guest :

- Knowing the peak booking prices and off season variations is very useful to decide when to book and visit the place.
- We believe that despite high prices through aviation people book here, because for everything booked through airlines, a good return is offered on miles, customers like to maintain status with airlines.
- Knowing ADRs per person per room is useful in booking for large parties of guests to save costs.