

Capstone Project

Hotel Booking Analysis

Suyash A. Patil

Points to Discuss

- Introduction
- Data Summary
- Univariate analysis
- Hotel wise analysis
- Distribution Channel wise analysis
- Booking Cancellation analysis
- Correlation Heatmap
- Conclusion

Introduction

We will discuss the analysis of given hotel bookings data set

Analysis Performed to derive key factors driving the hotel bookings trends:

- Univariate analysis
- Hotel wise analysis
- Distribution Channel wise analysis
- Booking cancellation analysis

Data Summary

Given data set has different columns of variables crucial for hotel bookings. Some of them are:

hotel: The category of hotels, which are two resort hotel and city hotel.

is_cancelled : The value of column show the cancellation type. If the booking was cancelled or not. Values[0,1], where 0 indicates not cancelled.

lead_time : The time between reservation and actual arrival.

stayed_in_weekend_nights: The number of weekend nights stay per reservation

stayed_in_weekday_nights: The number of weekday nights stay per reservation.

meal: Meal preferences per reservation.[BB,FB,HB,SC,Undefined]

Country: The origin country of guest

Is_repeated_guest: Tells us whether the customer has made any booking before(0 for No and 1 for Yes)

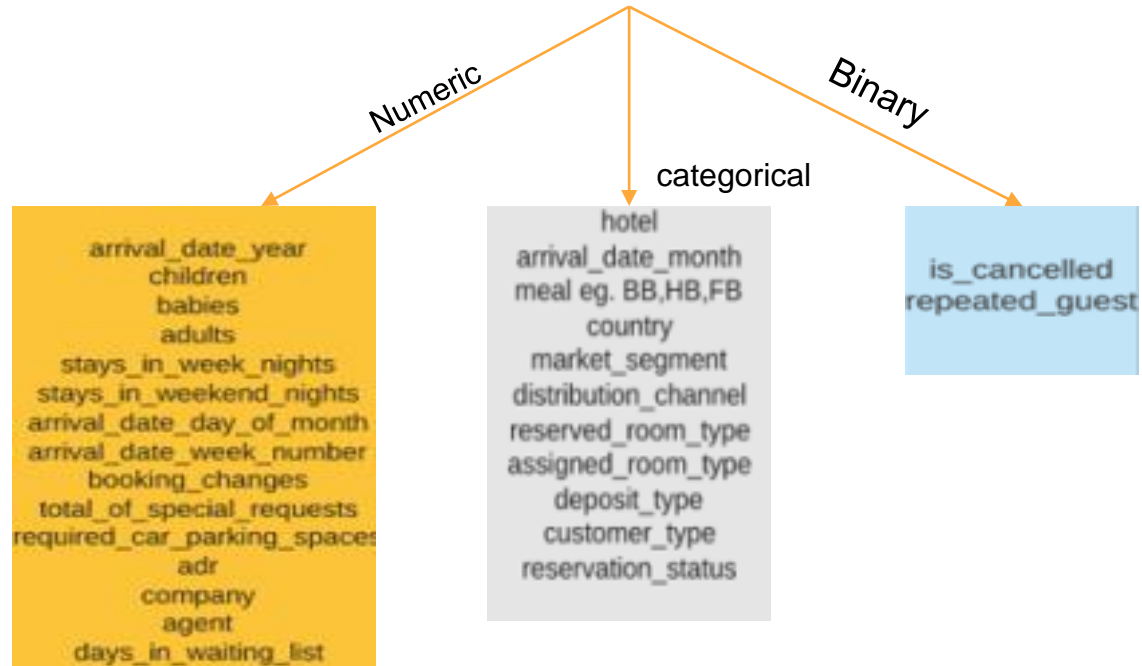
distribution_channel: The medium through booking was made.[Direct,Corporate,TA/TO,undefined,GDS.]

Is_repeated_guest: Shows if the guest is who has arrived earlier or not.Values[0,1]-->0 indicates no and 1 indicated yes person is repeated guest.

days_in_waiting_list: Number of days between actual booking and transact.

customer_type: Type of customers(Transient, group, etc.)

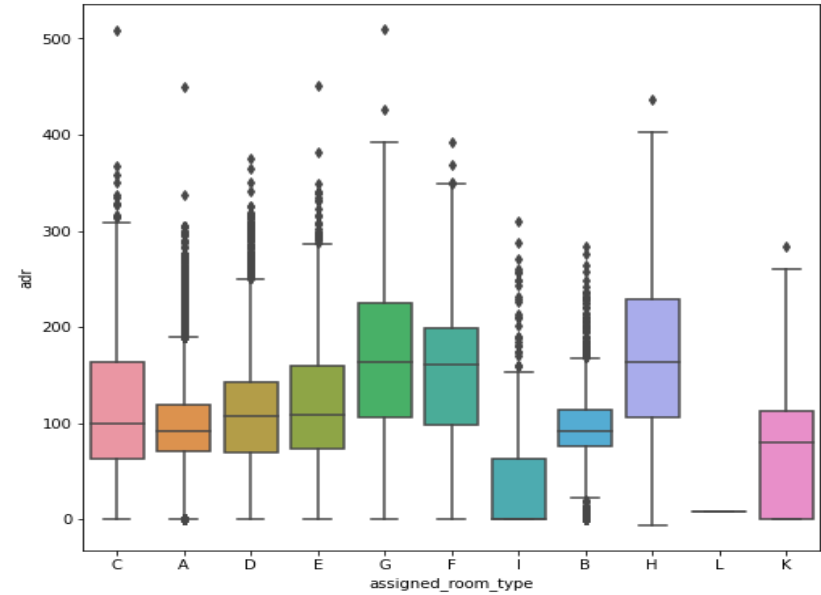
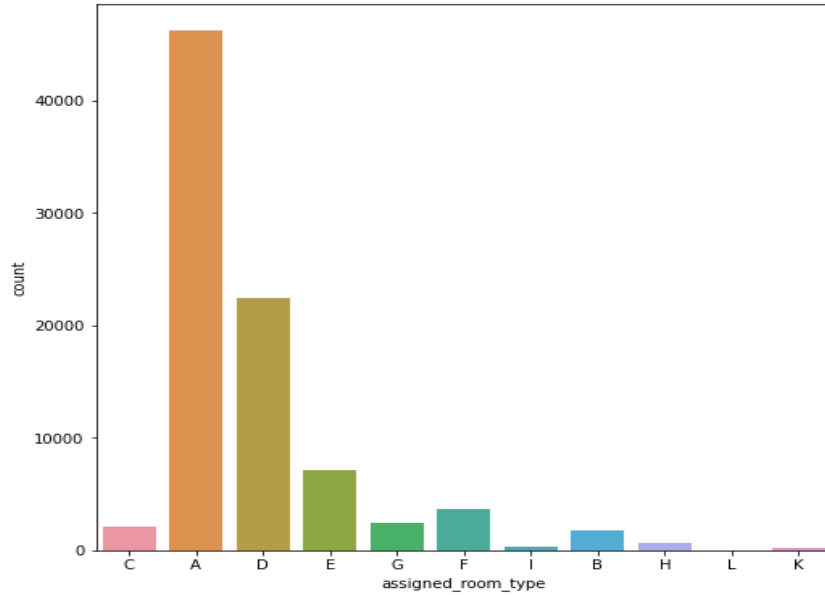
Dataset



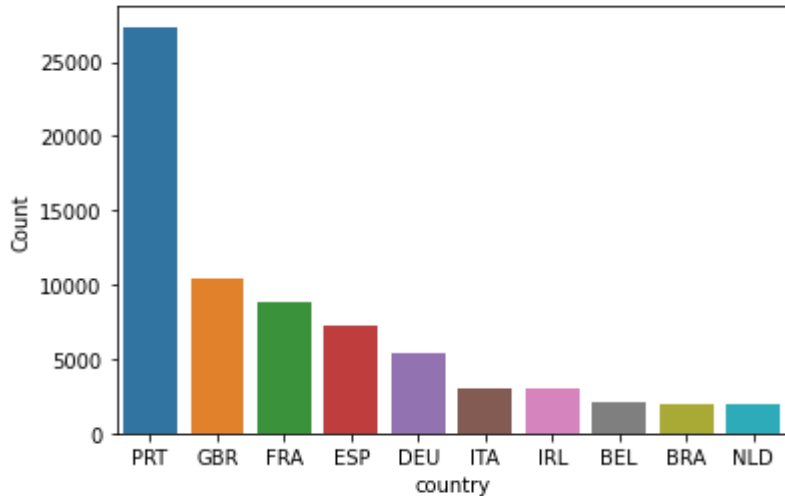
Univariate Analysis

While doing univariate analysis of given hotel booking dataset, we answered following questions:

- (1) Which room type is in most demand and which room type generates highest adr?
- (2) From which country most of the customers are coming?
- (3) What is the most preferred meal by customers

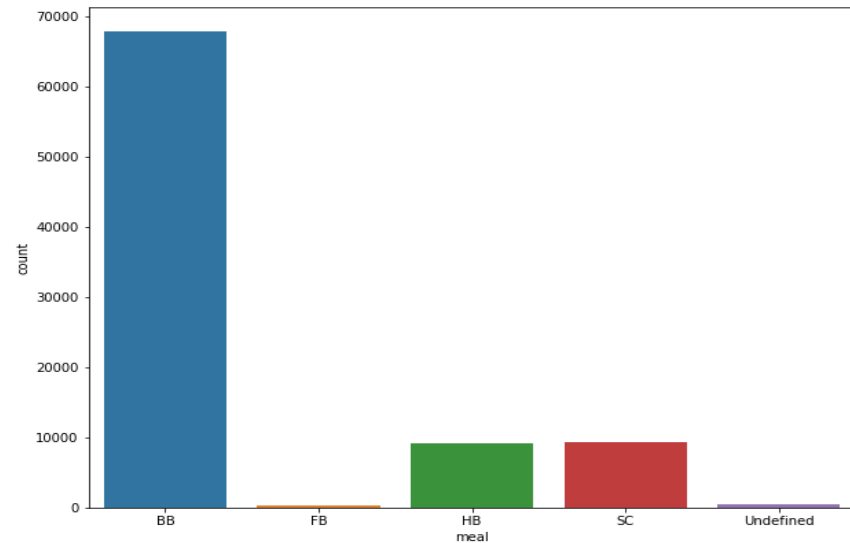


- Type A room is most demanded by customers.
- Room types C, G and H are some of the highest adr(average daily rate) generating rooms.



- Most preferred meal type is BB(Bed and breakfast).

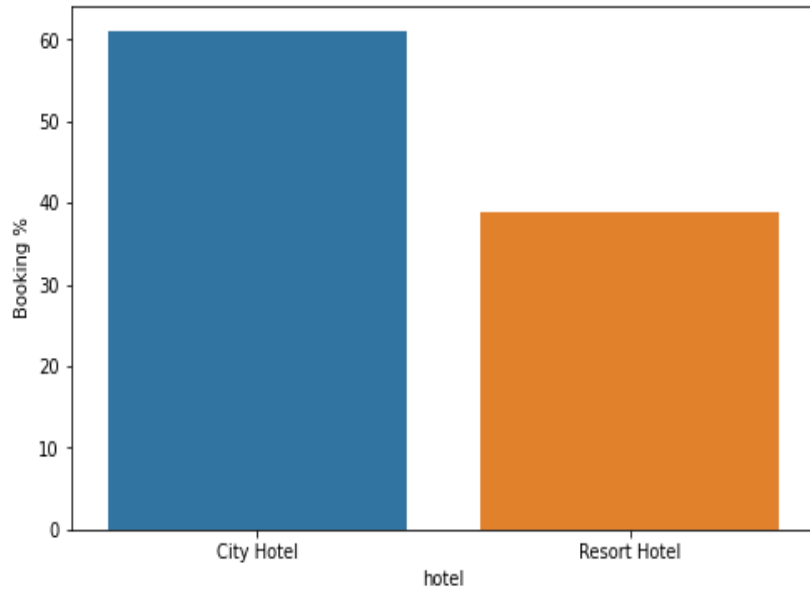
- Most of the customers from European countries like Portugal, Great Britain, France and Spain.



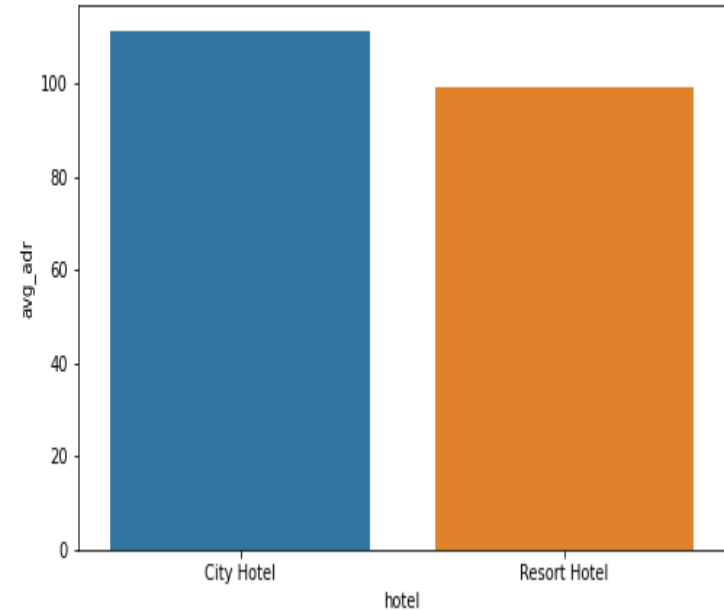
Hotel wise Analysis

While doing hotel-wise analysis of given hotel booking dataset, we answered following questions:

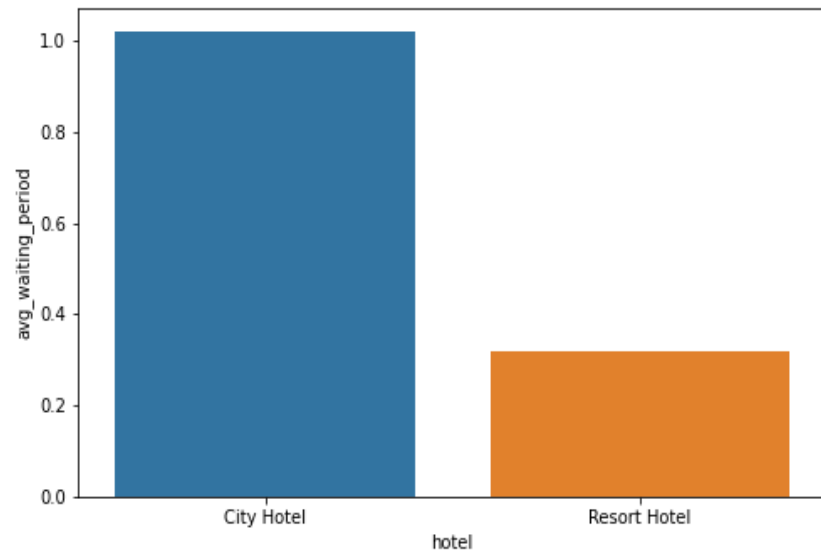
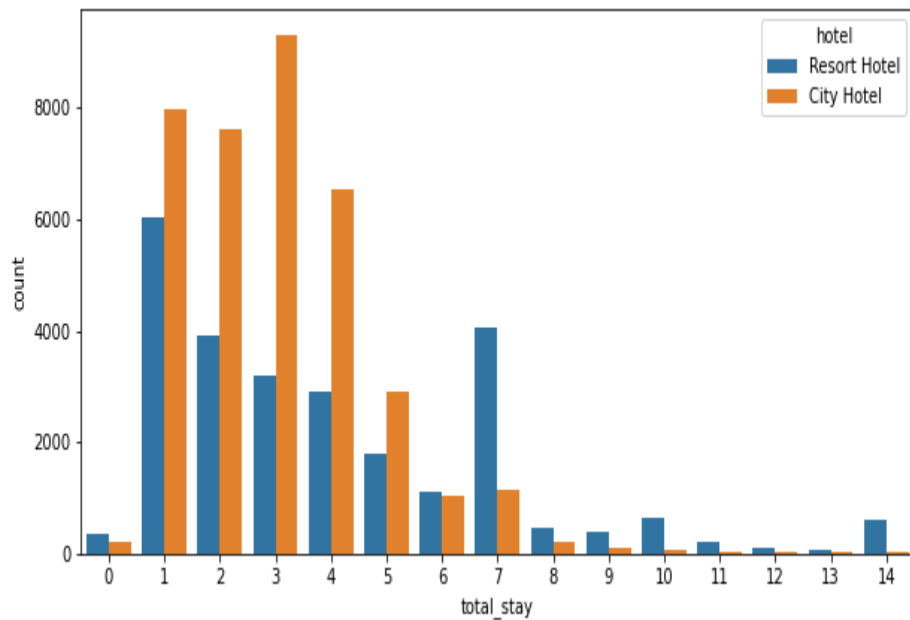
- (1) Percentage of bookings in each hotels?
- (2) Which hotel makes more revenue?
- (3) What is most preferred stay length in each hotel?
- (4) For which hotel, does people have to wait longer to get a booking confirmed?
- (5) Which hotel has higher booking cancellations rate?
- (6) Which hotel have higher customer returning rate?



- Around 60% bookings are for City hotel and 40% bookings are for Resort hotel.

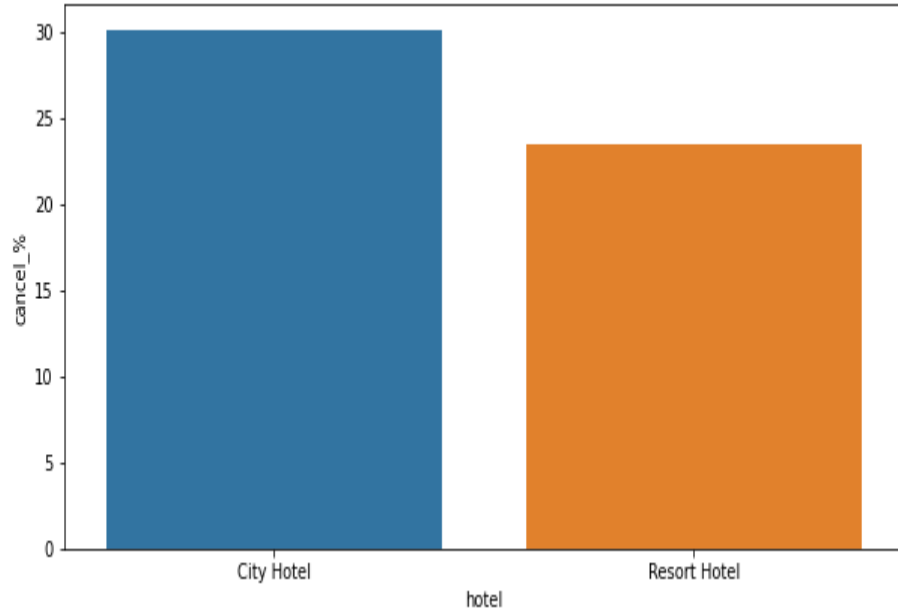


- Avg adr of Resort hotel is slightly lower than that of City hotel. Hence, City hotel seems to be making slightly more revenue.

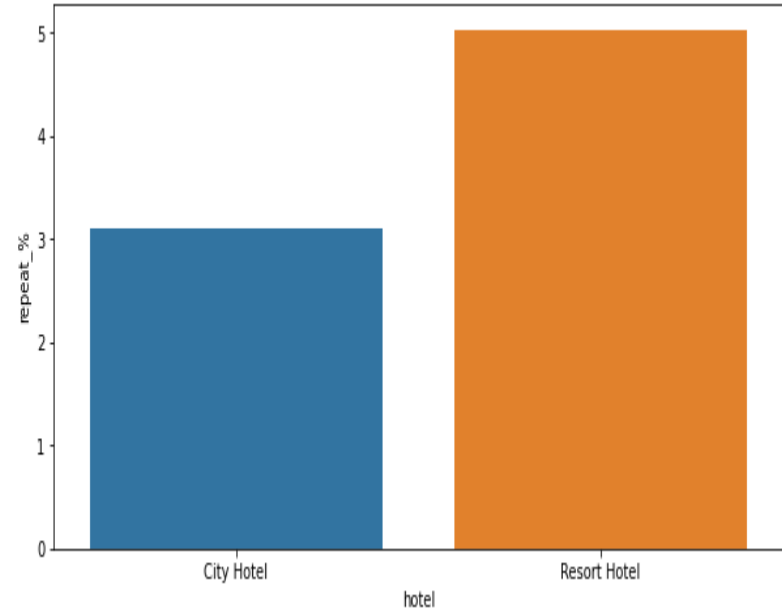


- Most of stays are less than 5 days. There are very few long stays at hotels but Resort Hotel is preferred for long stays.

- City hotel has significantly longer waiting time, hence City Hotel is much busier than Resort Hotel



- Almost 30 % of City Hotel bookings and 25 % of Resort hotel bookings got canceled

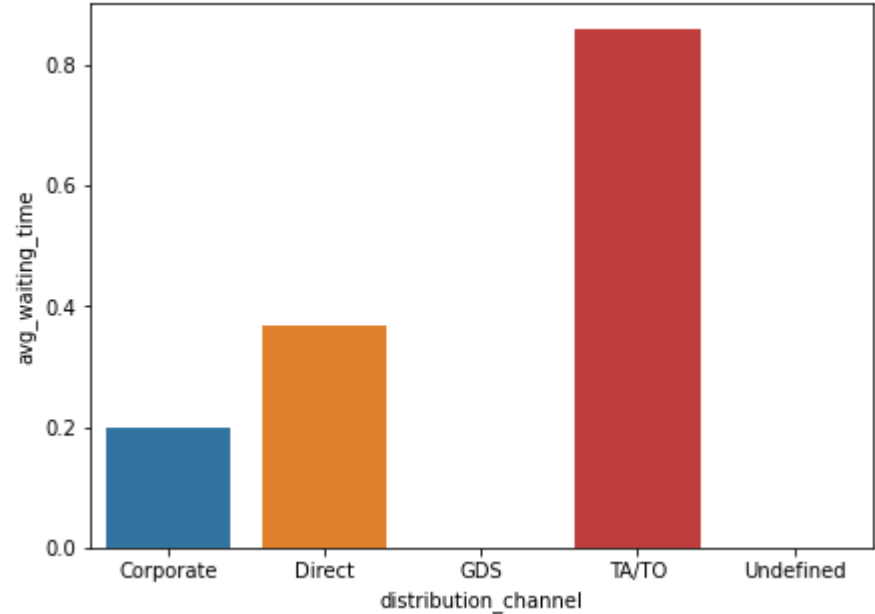
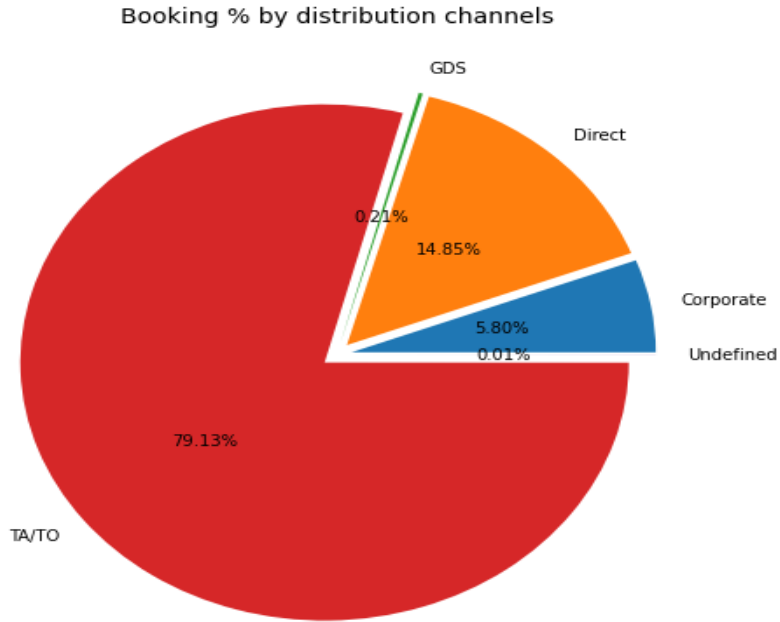


- Both hotels have very small percentage that customer will repeat, but Resort hotel has slightly higher repeat % than City Hotel

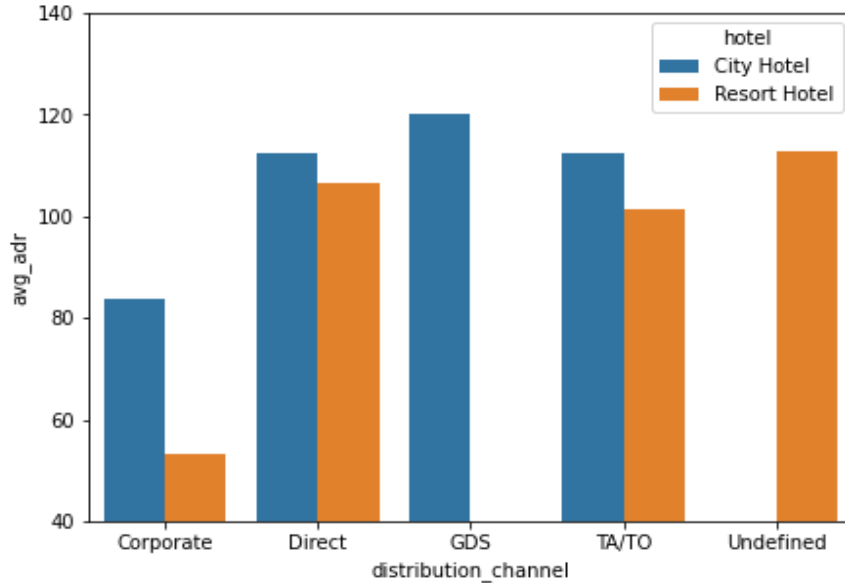
Distribution channel wise Analysis

While doing Distribution channel wise analysis of given hotel booking dataset, we answered following questions:

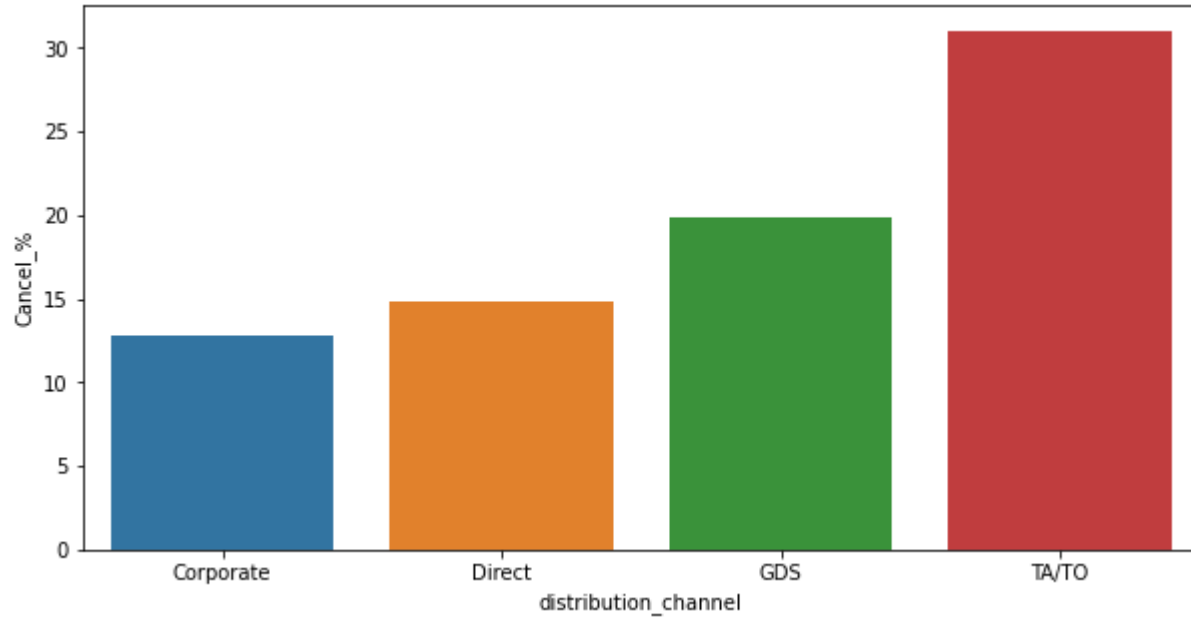
- (1) Which is the most common channel for booking hotels?
- (2) Which channel is mostly used for early booking of hotels?
- (3) Which distribution channel brings better revenue generating deals for hotels?



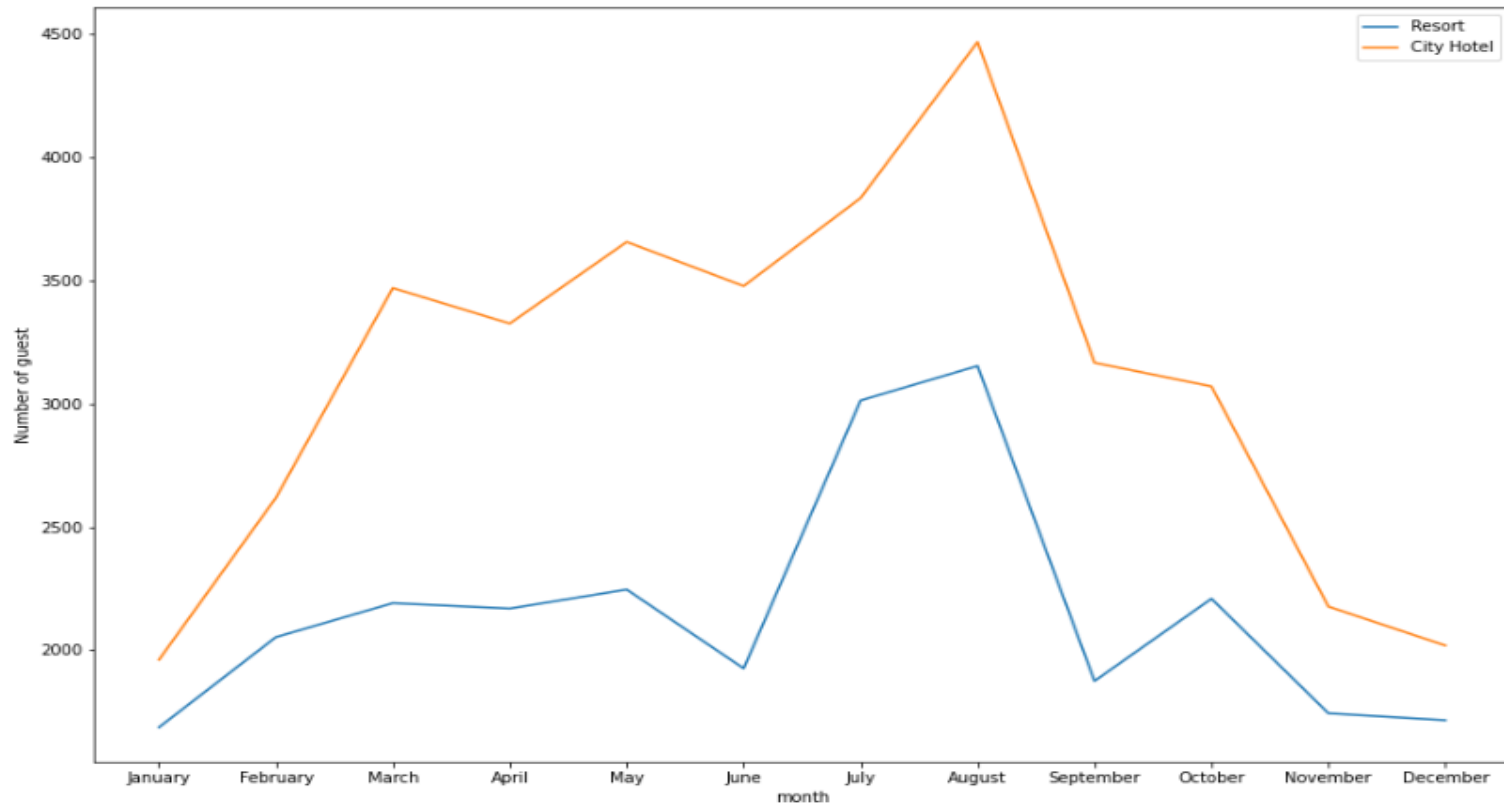
- Here we can see that the most of guest are making reservation through TA/TO channels which is travel agency and tour operator.
- The second most used channel is direct.
- Channel which is mostly used for early booking of hotels is also TA/TO



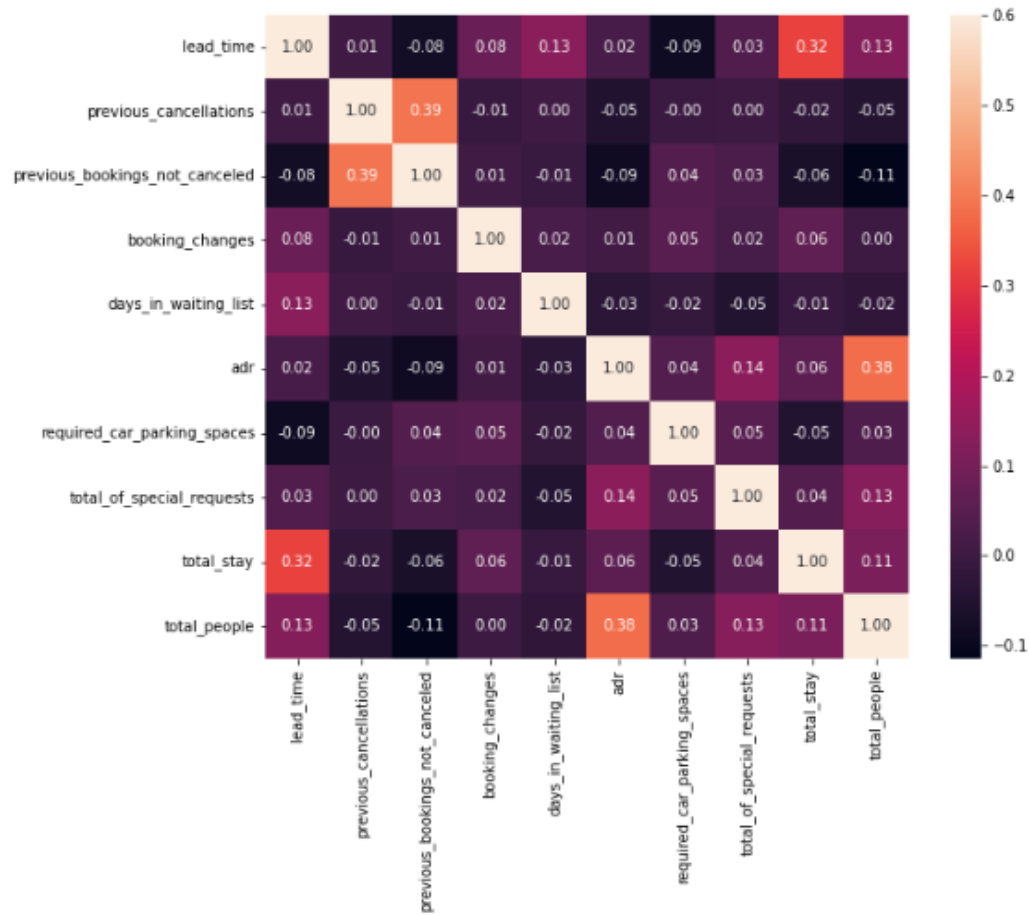
- GDS channel brings higher revenue generating deals for City hotel, in contrast to that most bookings come via TA/TO. City Hotel can work to increase outreach on GDS channels to get more higher revenue generating deals.
- Resort hotel has more revenue generating deals by direct and TA/TO channel. Resort Hotel need to increase outreach on GDS channel to increase revenue.



- TA/TO has highest booking cancellation %. Therefore, a booking via TA/TO is 30% likely to get cancelled.



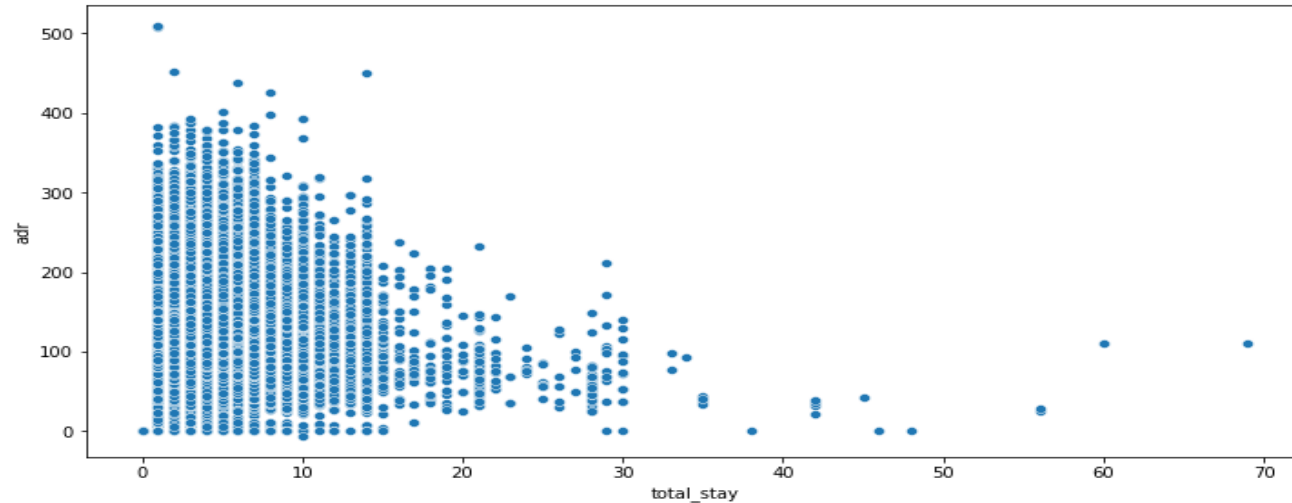
From the month of July to August the number of bookings increased and in August, City Hotel got most number of guests



- Total stay length and lead time are slightly correlated. This may mean that for longer hotel stays, people generally plan little before the actual arrival.

- adr is slightly correlated with total_people, which makes sense as more no. of people means more service to deliver, therefore more adr.

Optimal stay length for better deals in adr



For shorter stays the adr(average daily rate) varies greatly but for longer stays (> 15 days) adr is comparatively very less. Therefore, customers can get better deal for longer stays more than 15 days

Conclusion

- Most of the guests came from European countries, with most no. of guest coming from Portugal.
- Guests use different channels for making bookings out of which most preferred way is TA/TO.
- Mostly guests stay for less than 5 days in hotel and for longer stays Resort hotel is preferred.
- Around 60% bookings are for City hotel and 40% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel. Also the overall adr of City hotel is slightly higher than Resort hotel.
- Both hotels have significantly higher booking cancellation rates and very few guests less than 3 % return for another booking in City hotel. 5% guests return for stay in Resort hotel.
- July- August are the most busier and profitable months for both of hotels.
- For customers, generally the longer stays (more than 15 days) can result in better deals in terms of low adr.