Capstone Project - 1

Hotel Booking Analysis

Team Members

Shyam Shubham Mishra

Nisha Kumari

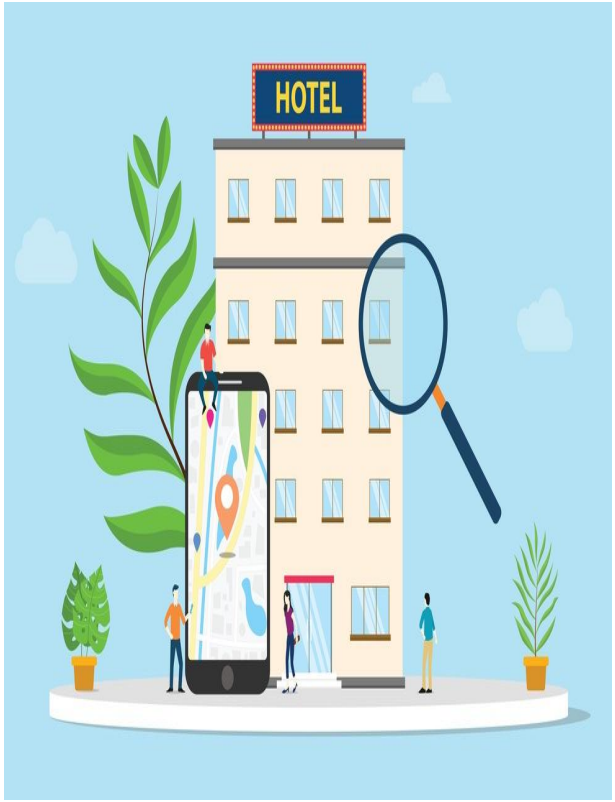
Mohd. Firoz Chaudhri

Ishan Singh

Hiten Mewada

LET’S DIG INTO THE HOTEL DATA

1. Defining the problem statement
2. Observing the dataset
3. Processing the data
4. Visualising the dataframe
5. Conclusion



Hotel industry is a very volatile industry and the booking depends on many factors. The main objective of hotel analysis is to explore and analyze data to discover important factors that govern the bookings and give insights to hotel management, which can performs various campaigns to boost the business and performance.

Exploring different problem parameters:

1. Examining the guest distribution
2. Exploring the bookings
3. Inspecting the hotel traffic
4. Scanning special requests & amenities
5. Study on average daily rates

Data Summary:-

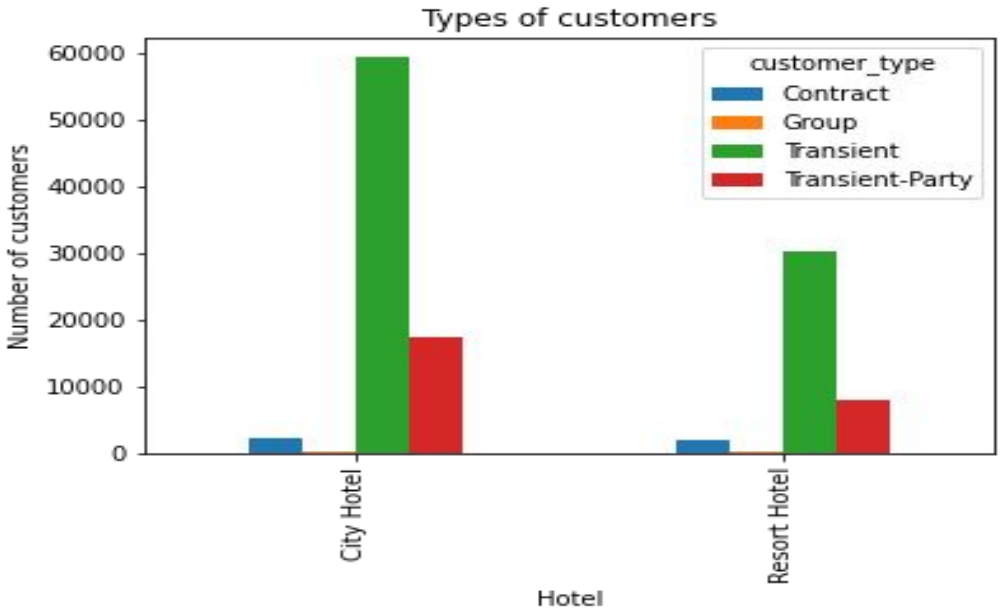
In the hotel booking dataset we have total 1193990 rows and 32 columns.

- **Hotel: Name of two hotels - City Hotel & Resort Hotel.**
- **Is_cancelled: (0,1) Indicates if booking cancelled or not.**
- **Lead_time: Time b/w reservation and actual arrival.**
- **Arrival_date_year: Year of arrival date.**
- **Arrival_date_month: Month of arrival date**
- **Arrival_date_week_no: Week no. of arrival date**
- **Arrival_date_day_of_month: Date of month of arrival date**
- **Stays_in_weekends_nights: no. of weekend nights guests stayed at the hotel**
- **Stays_in_week_nights: no. of week nights guests stayed at the hotel**
- **Adults,children,babies: no. of adults, children & babies arriving**
- **Meal: Type of meal preferred by guests**
- **Country: Countries from which guests are arriving**

- Market_segment: reservation sources
- Distribution_channel: Booking mediums
- Is_repeated_guest: (0,1) represents whether guests were repeated
- Previous_cancellations: (0,1) represents whether guests have previous cancellation
- Reserved_room_type: Type of room booked
- Assigned_room_type: room assigned by hotel to guests
- Booking_changes: number of changes made in bookings
- Deposite_type: whether payment was made at the time of booking or not
- Agent: charges that hotel provide as booking commission charges
- Company: company Id that made the booking
- Days_in_waiting_list: no. of waiting days for booking confirmation
- Customer_type: Types of customers
- Adr: Average Daily Rates
- Required_car_parking_spaces: Required parking spaces
- Total_of_special_requests: Special request demanded by guests
- Reservation_status: Current status of booking
- Reservation_status_date: Date of Reservation

Examining the guest distribution

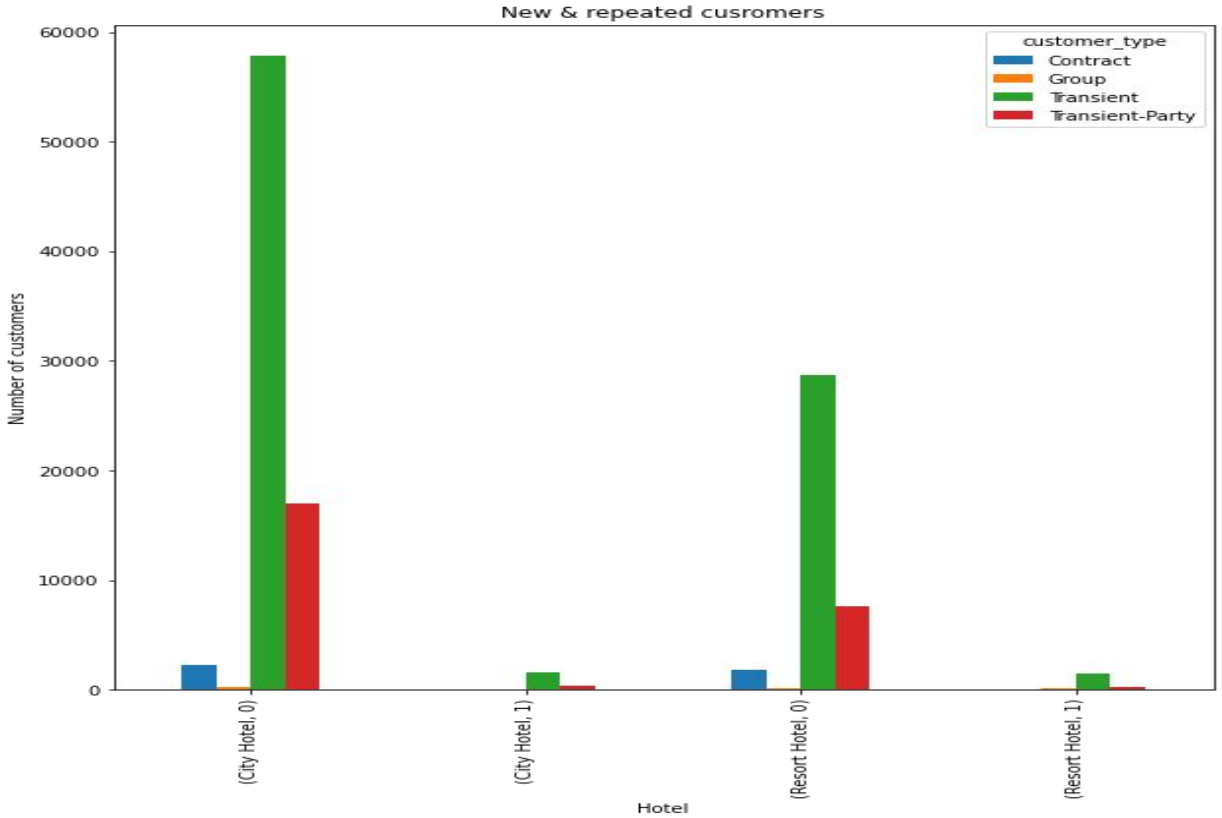
- **Types of customers w.r.t the hotels**



From the above bar plot we find that both the City Hotel & Resort Hotel are attracting 'Transient' type of customers the most followed by 'Transient-Party' type. Whereas, 'Group' type of customers are the least attended by both the hotels.

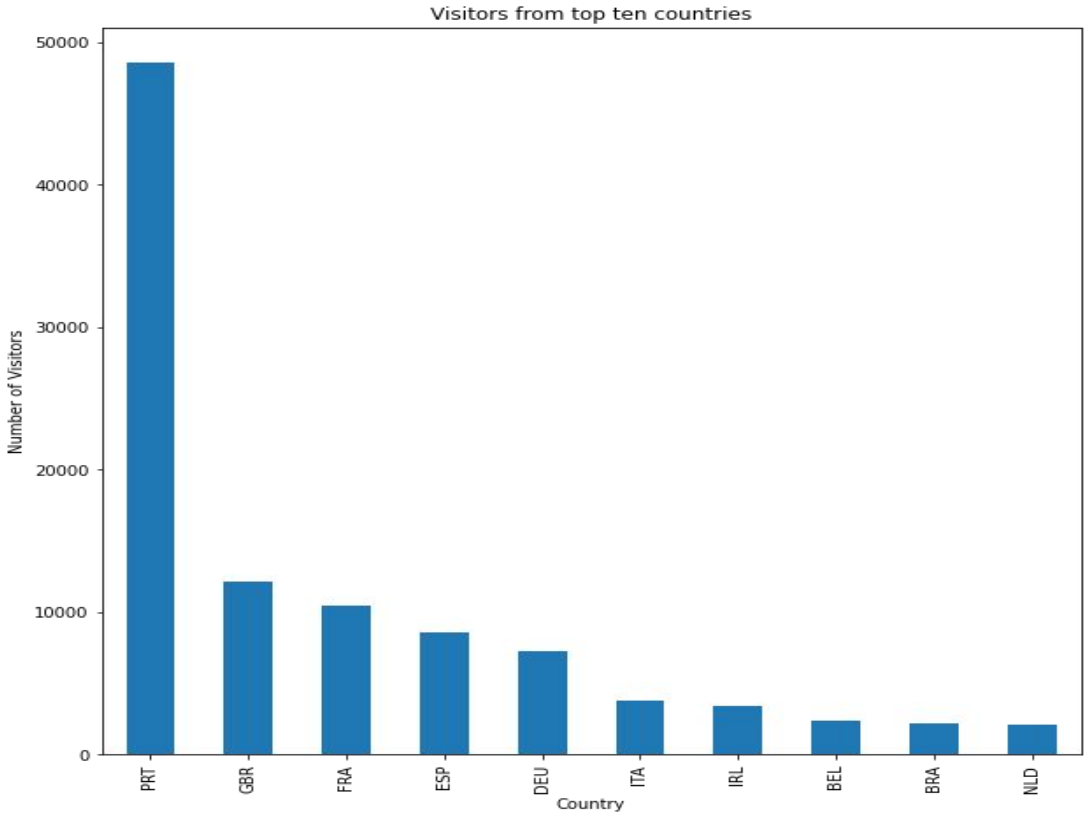
- **Ratio of new customers & repeated customers**

From the above data frame which represents the total number of repeated guests in both the hotels, we conclude that 'Transient' type of customers are most repeated in both the hotels, 'Contract' type customers being the least repeated.



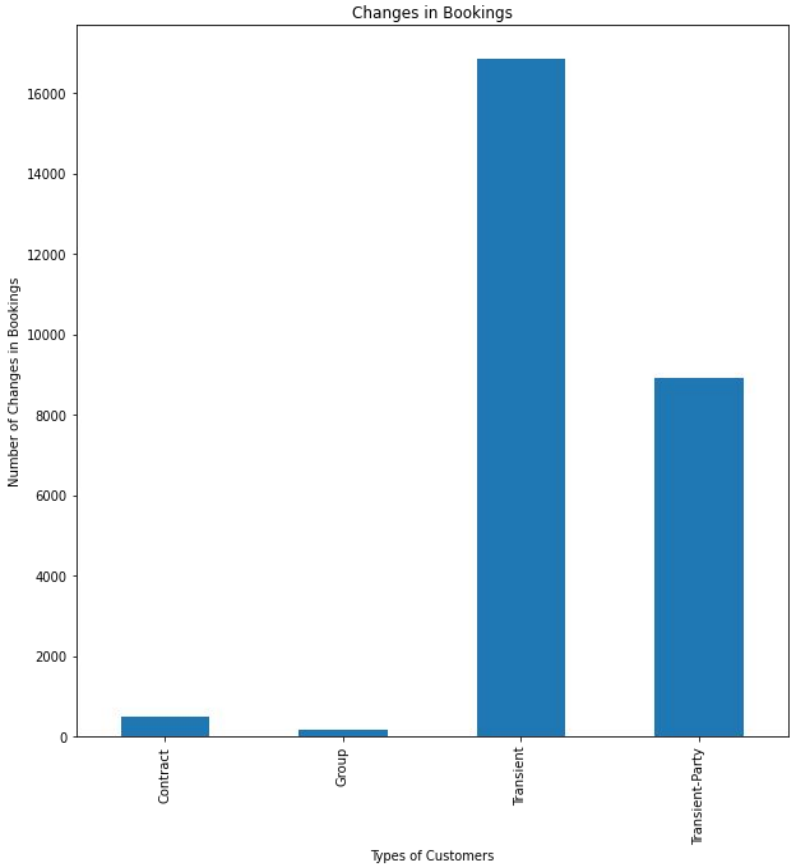
- **Customers from different countries**

The hotels are attending most guests from PTR followed by GBR & FRA**.**



- **Change in bookings**

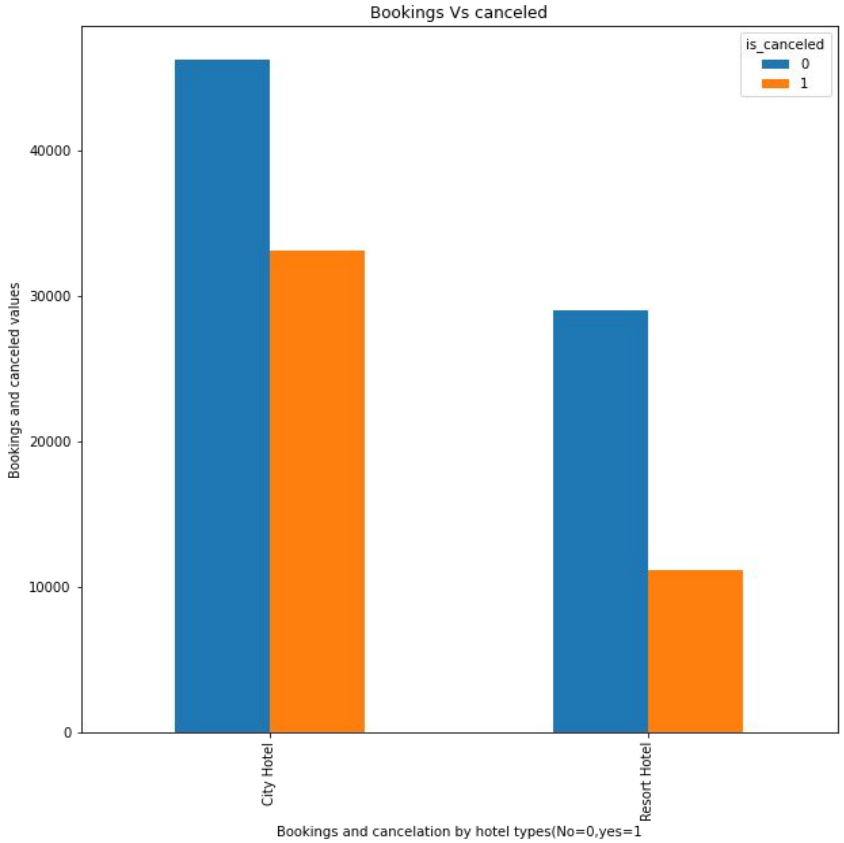
Here we conclude that 'Transient' type customers are making the most changes in bookings, whereas 'Group' type customers are making the least changes in bookings.



Exploring the bookings

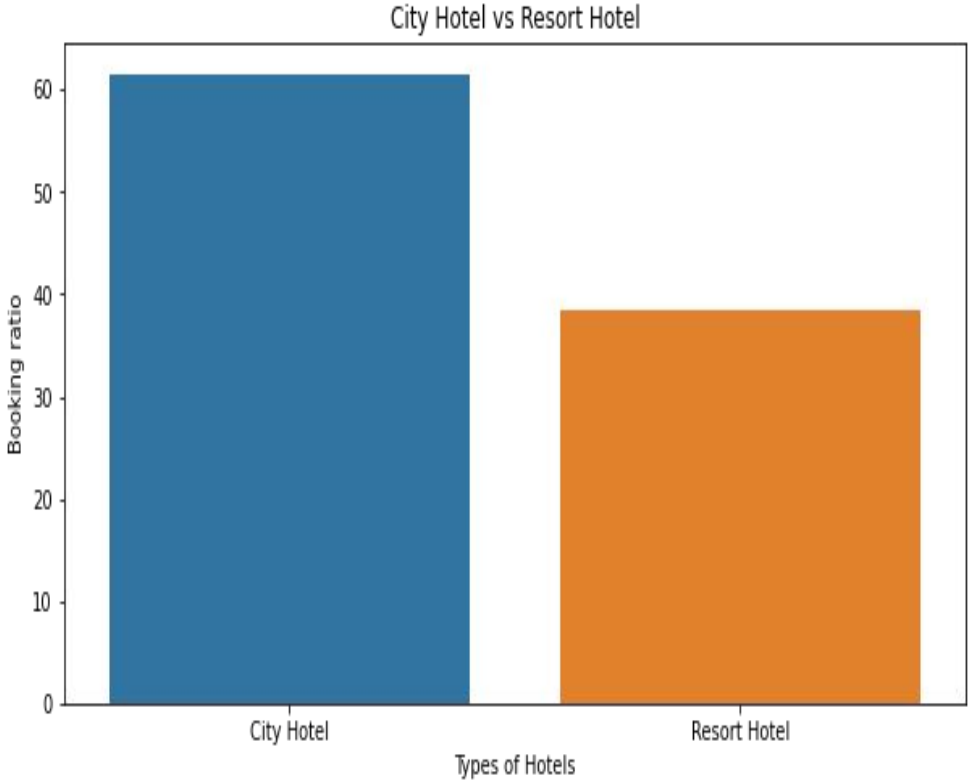
- **Total number of cancelled Bookings by hotel type**

From, the result We can conclude here number of Booking and cancellation are more in city hotel than resort hotel



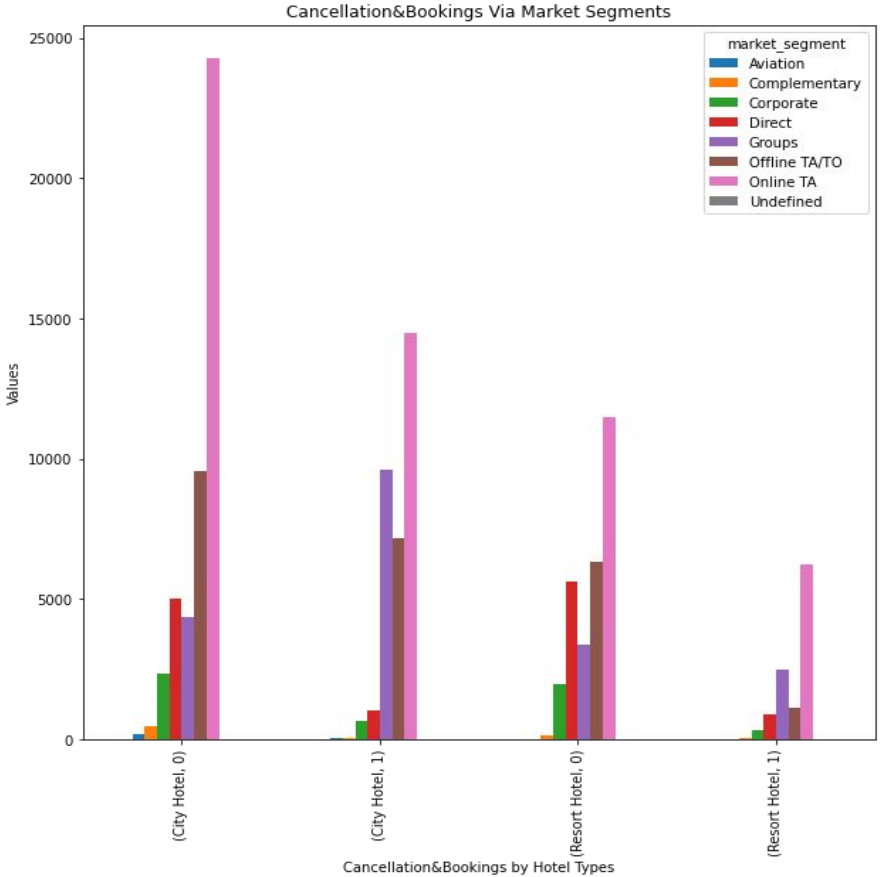
- **Booking ratio between Resort Hotel and City Hotel**

As a result we can see more than 60% of the people booked from the City Hotel compared to Resort Hotel



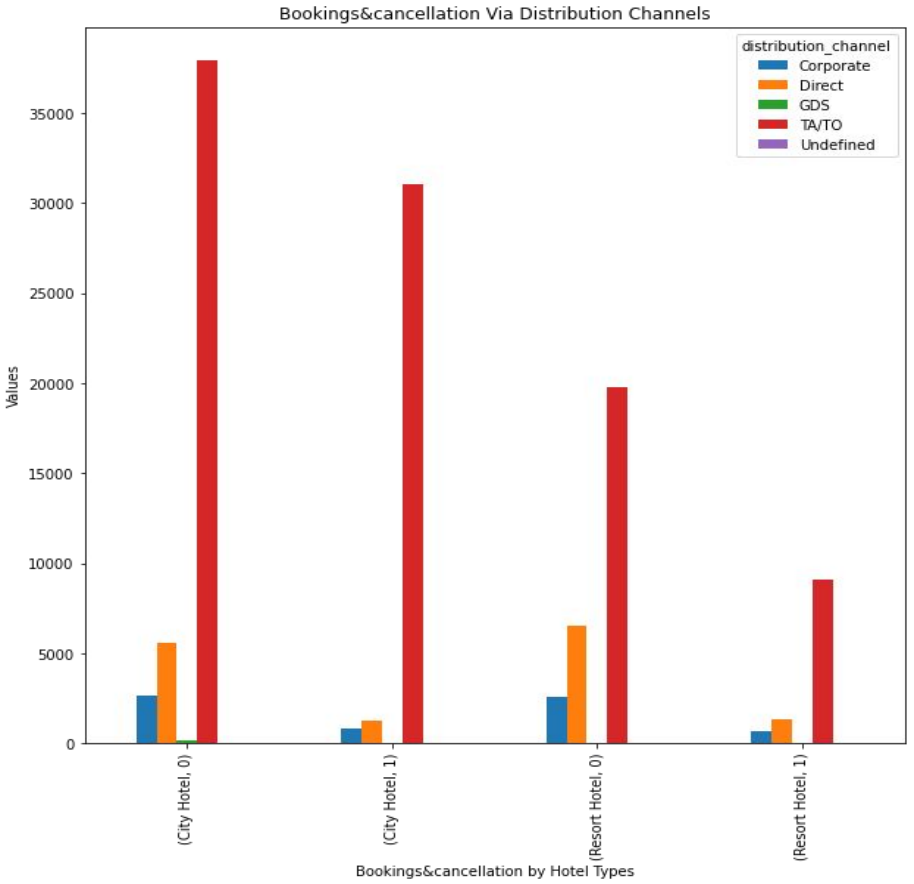
- **Total number of Cancellations through market segments**

We can find the conclusion from the diagram the highest segment from where the booking and cancellation done are Online TA



- **Total number of Bookings through different Distribution Channels**

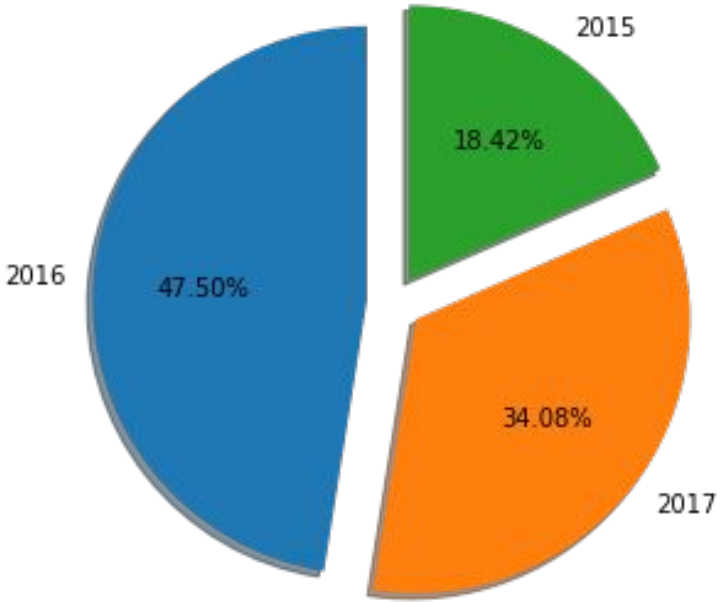
From,the result we can extract the higher channels from where the bookings and cancellation done is TA/TO from City Hotel.



Inspecting the hotel traffic

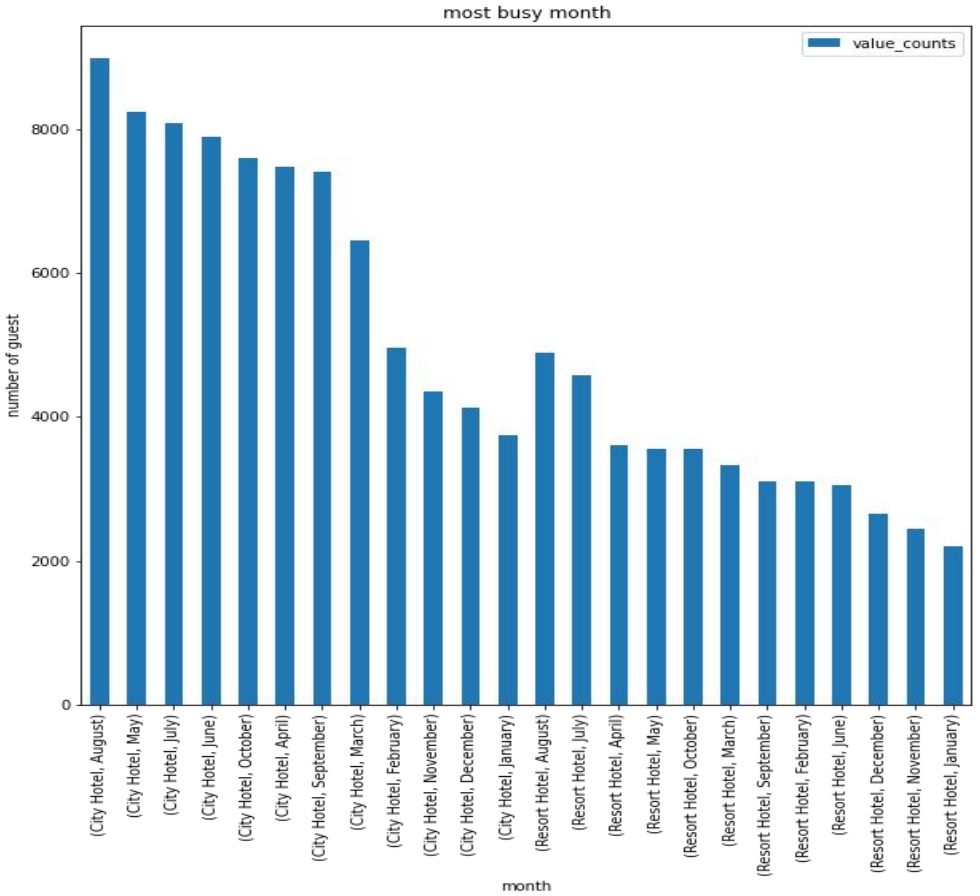
- **Most busy year of hotel booking**

Most busy year is 2016 followed by 2017 and 2015



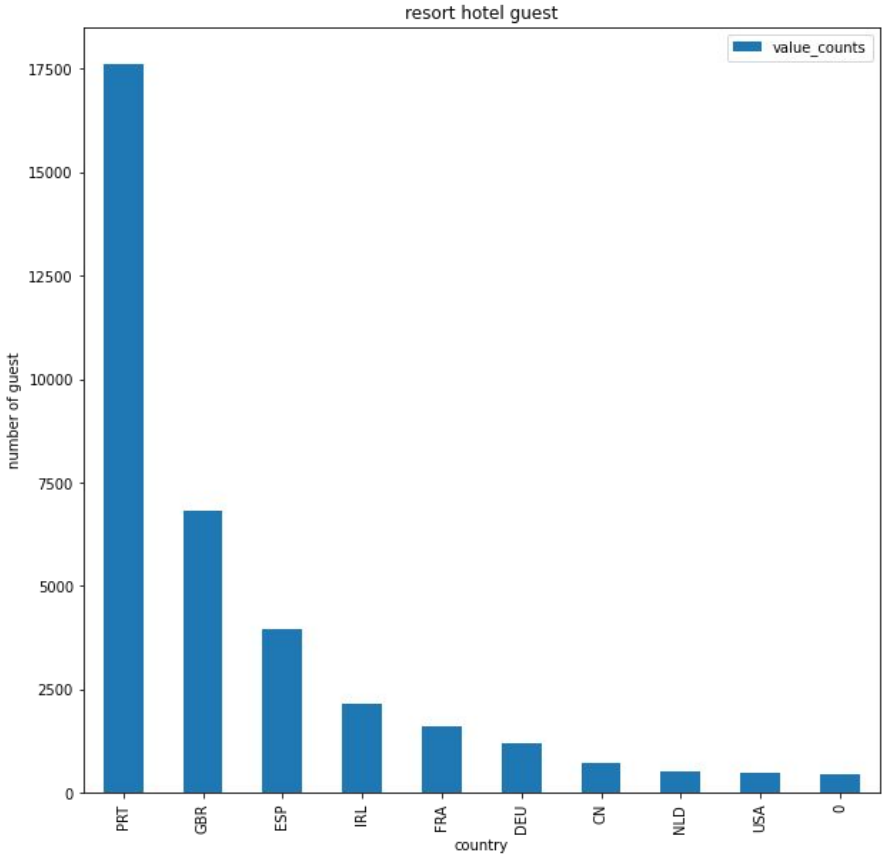
- **Most busiest month of hotel booking**

Most busy month of city hotel is August and Most busy month of resort hotel is also August



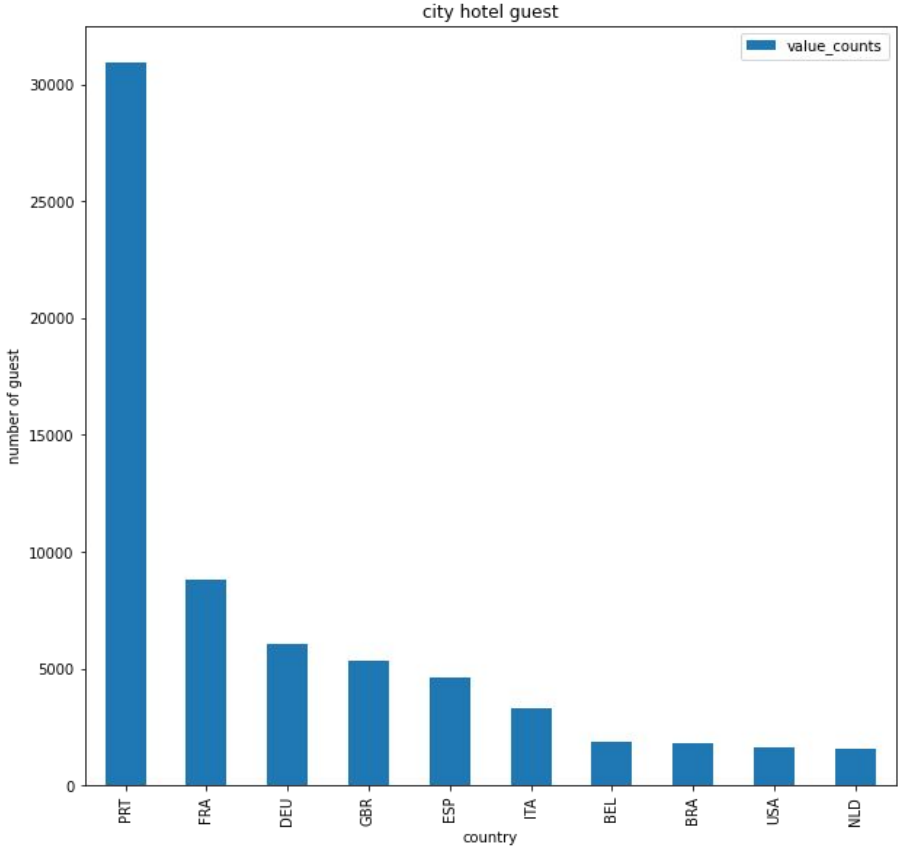
- **Guests coming from mentioned countries in Resort Hotel**

Most guest are coming from PRT country in Resort hotel



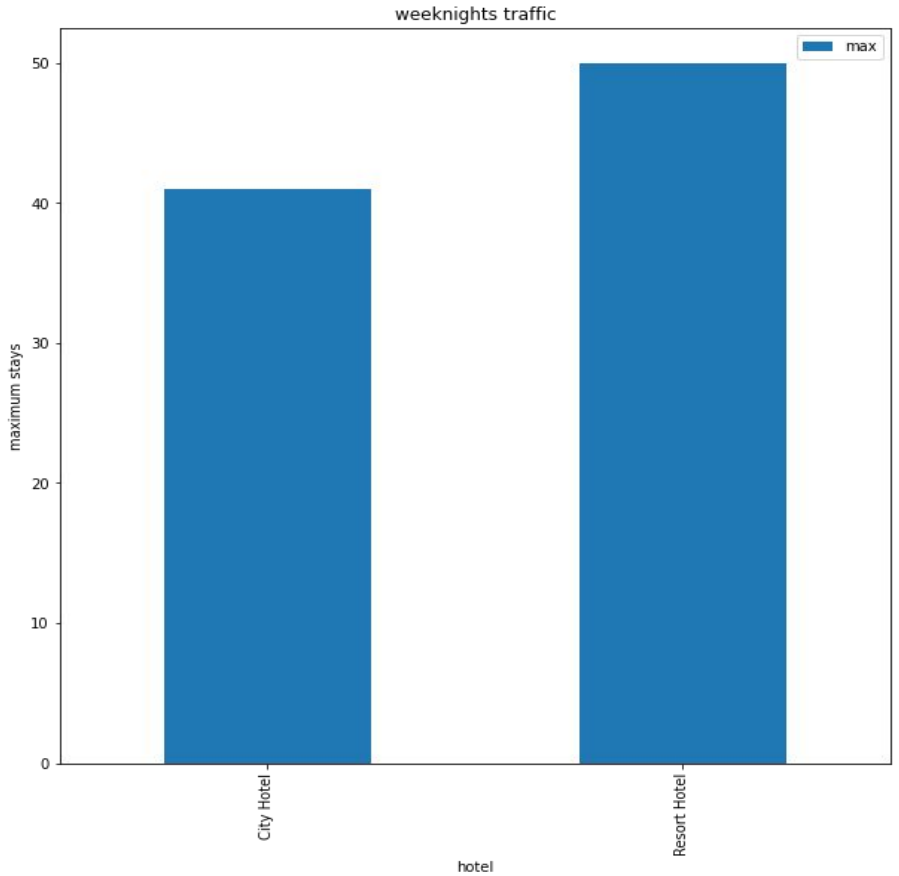
- **Guests coming from mentioned countries in City Hotel**

Most guest are coming from PRT country in city hotel



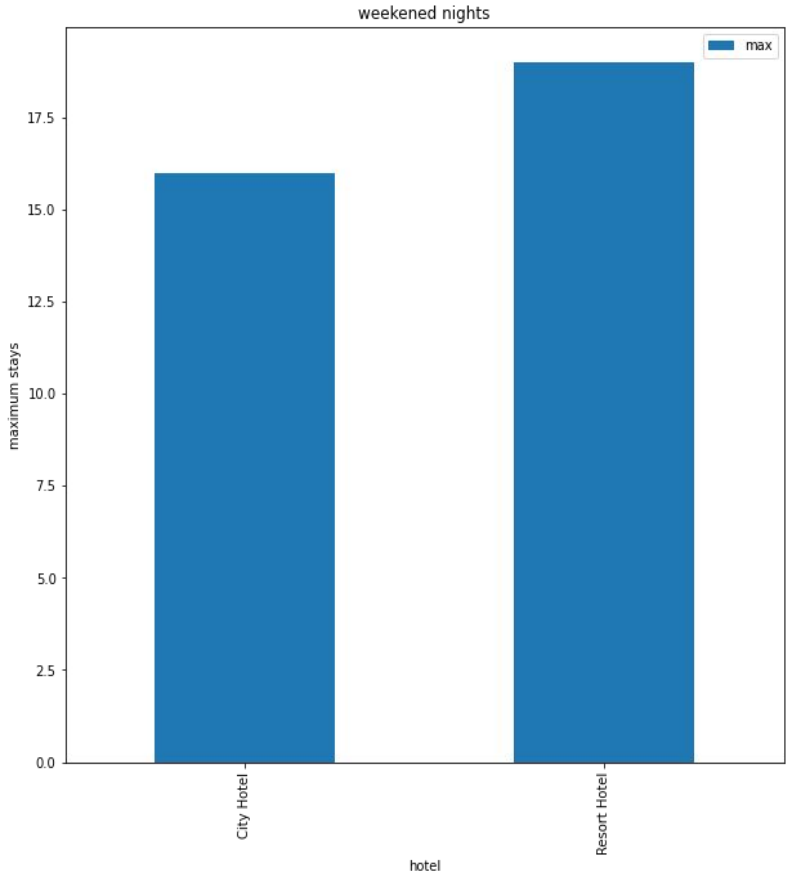
- **Number of stays in week nights**

Maximum number of stays on weeknights in Resort hotel followed by city hotel.



- **Number of stays in weekend nights**

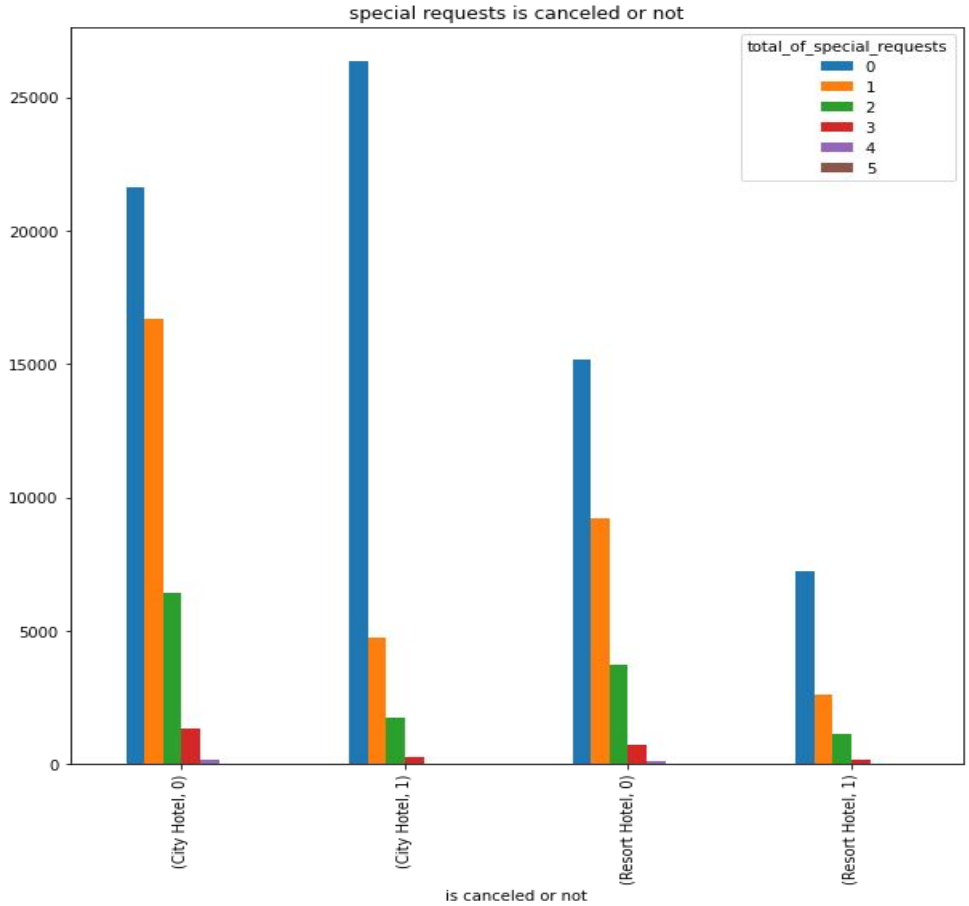
Maximum number of stays in weekend nights in Resort hotel followed by city hotel.



Examining special requests and amenities

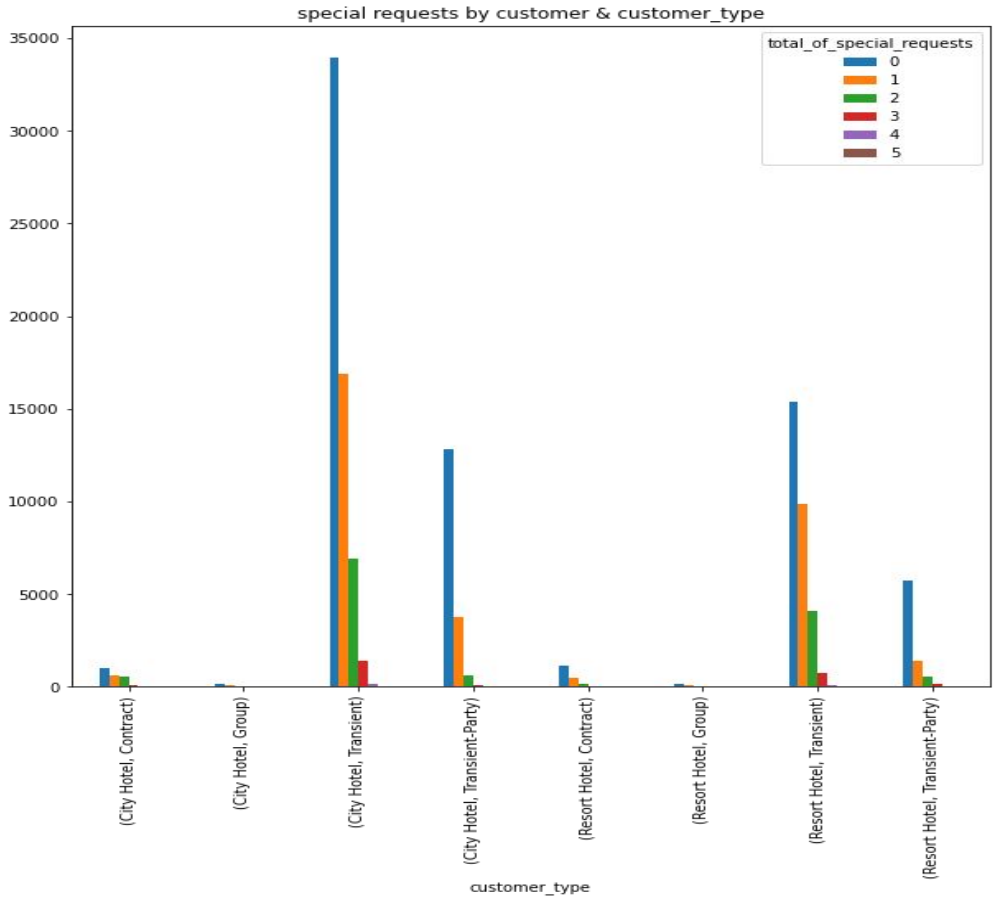
- **Relations between 'total of the special requests' , 'is canceled' or 'not'**

In both City Hotel and Resort hotel type 1 special requests are the most cancelled, whereas type 5 are the least cancelled.



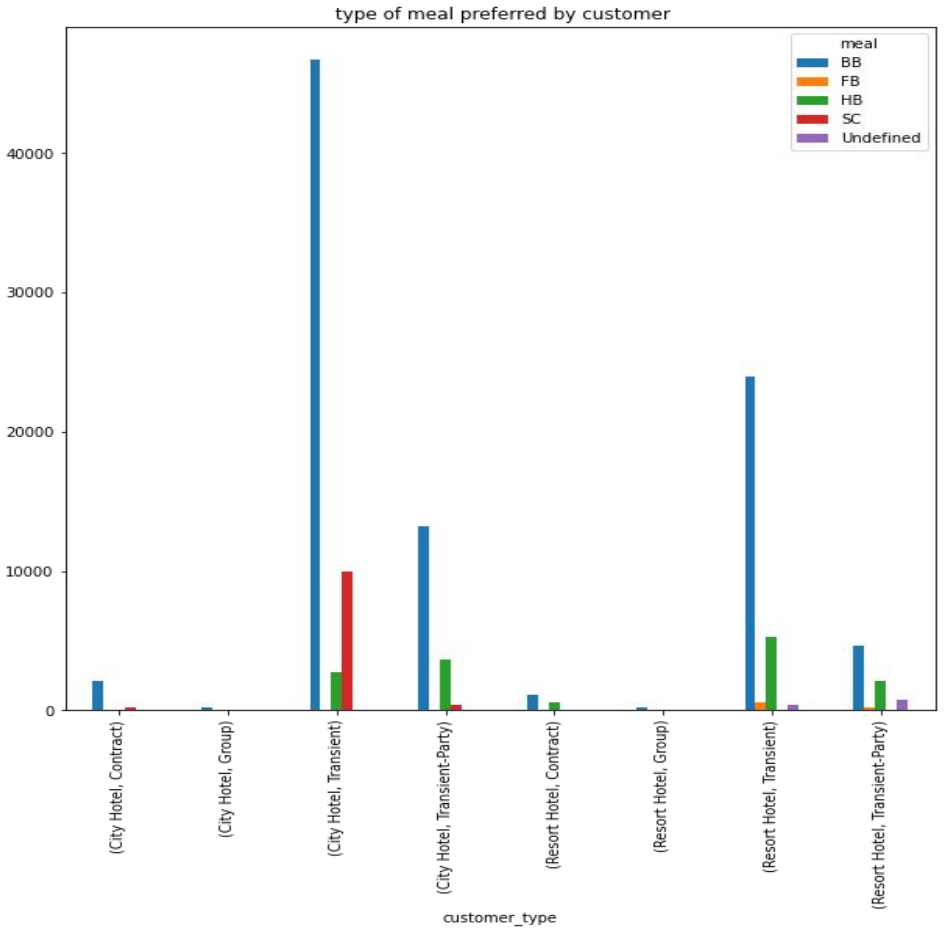
- **Relations between 'total of Special requests' with respect to 'customer**

In city hotel, most special requests are coming from transient and transient party type of customers



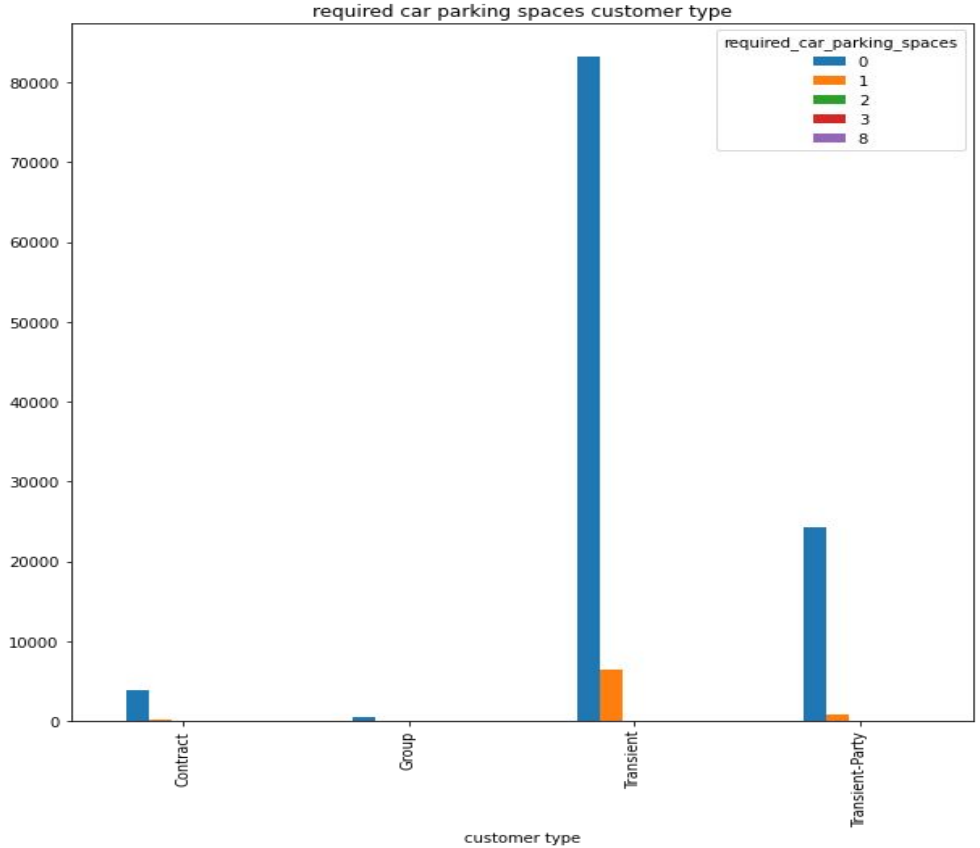
- **Type of meals are preferred by the customers**

In both the hotels BB meal is preferred most by the guests followed by SC & HB.

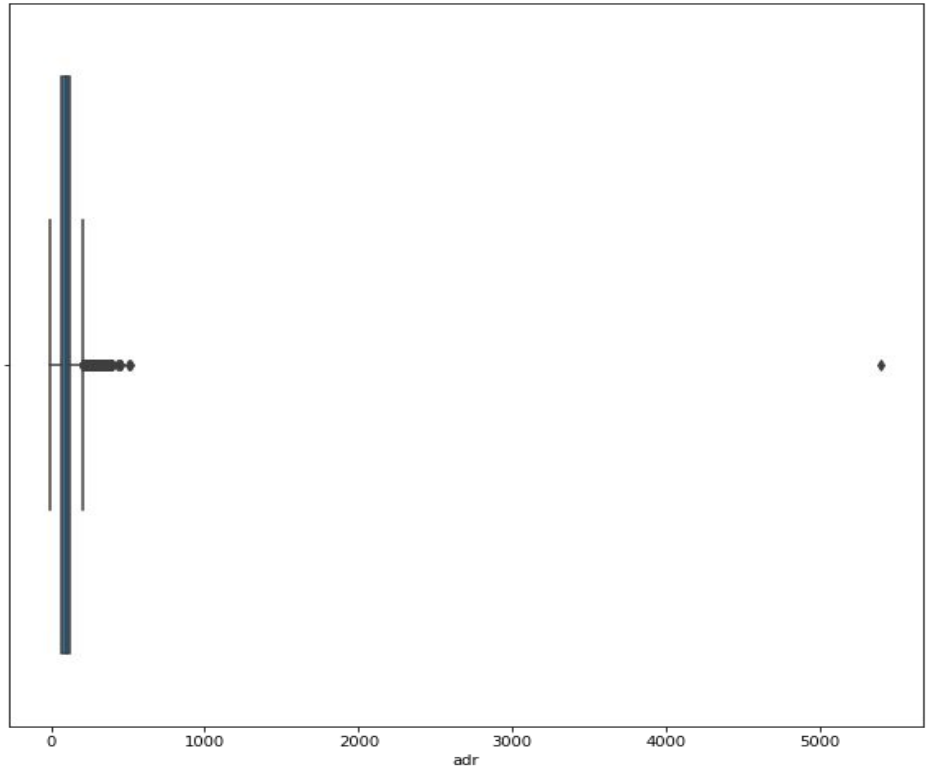


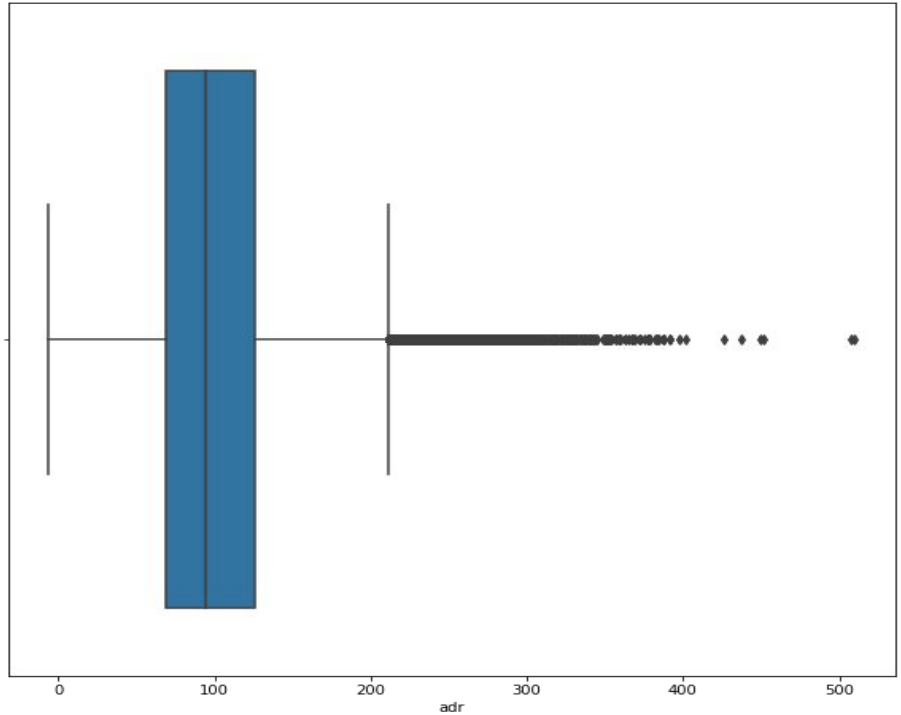
- **'Required car parking spaces' with respect to 'customer**

Transient type of customers are requiring car parking the most followed by Transient Party type of customers



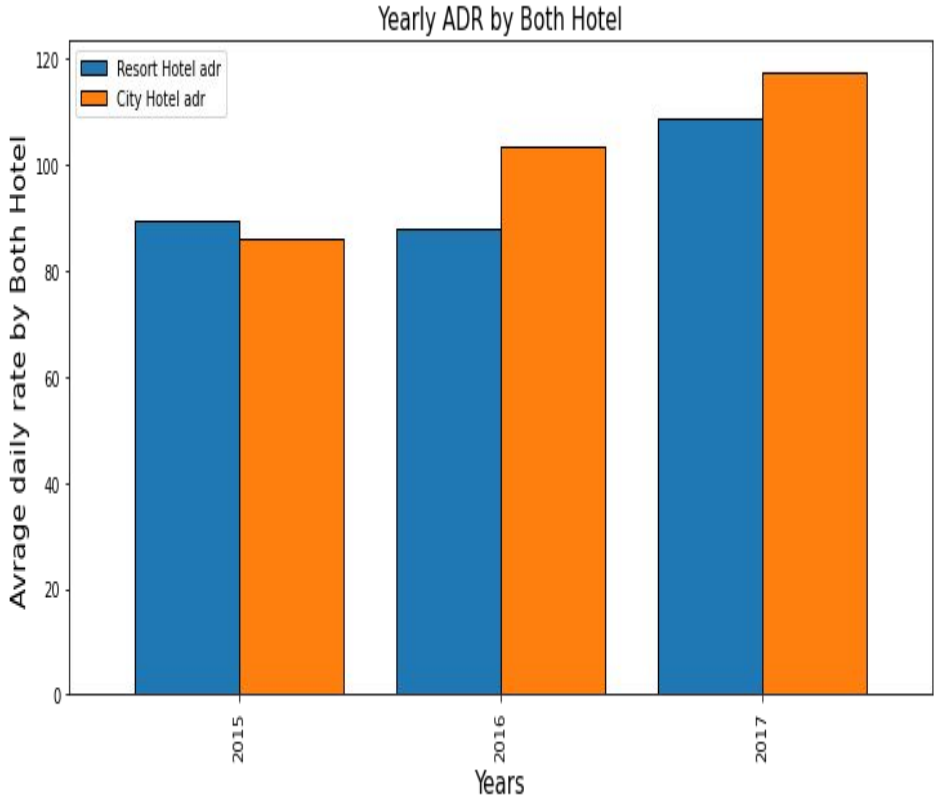
- **Study on Average Daily Rates (ADR)**





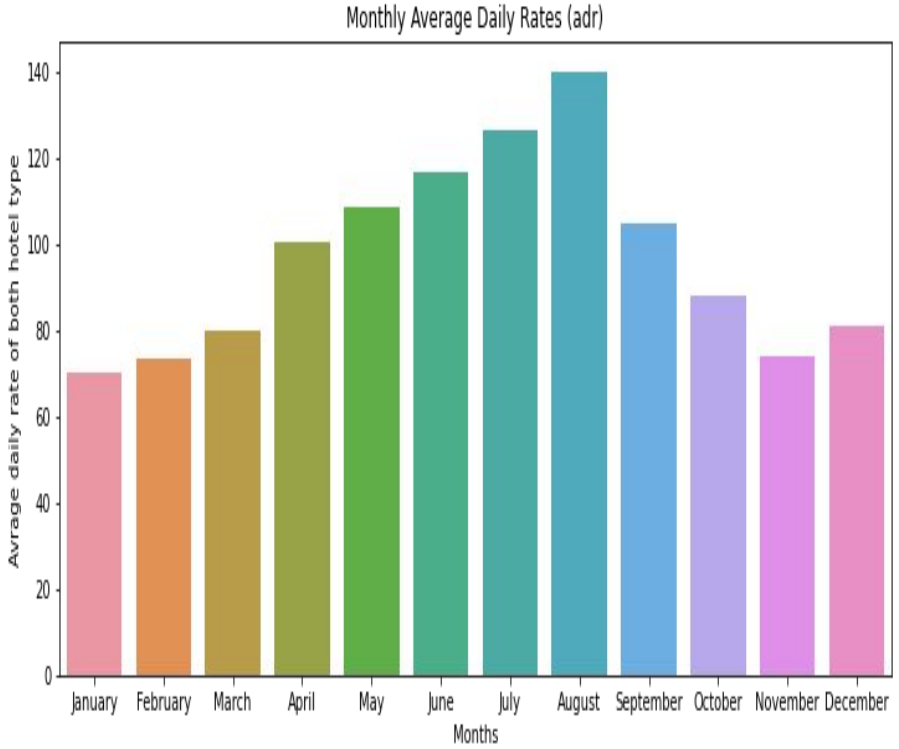
- **Yearly average of ADR with respect to hotel type**

As we can see over the years, the average daily rate has increased in both hotels. The city hotel is generating more adr than the resort hotel, but in 2015 the resort hotel is slightly ahead compared to the city hotel.

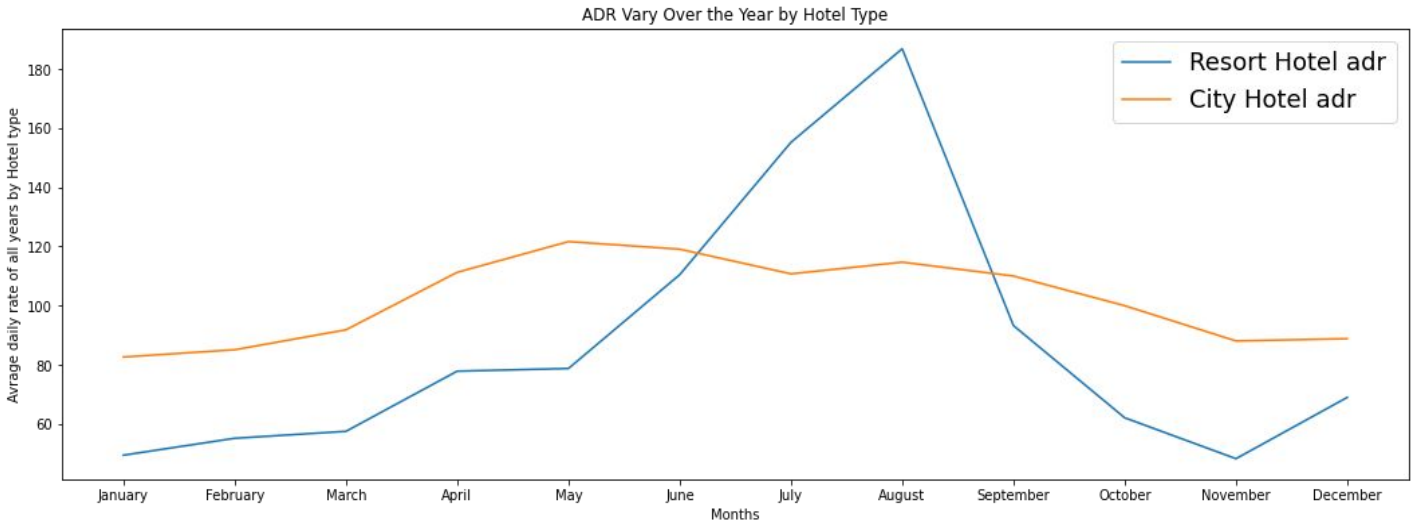


- **Monthly Average daily rates (adr)**

As we can see, the highest average daily rate (ADR) has occurred in August. It means both hotels are generating more adr in August and in January, rooms are more available due to the lowest average daily rate.

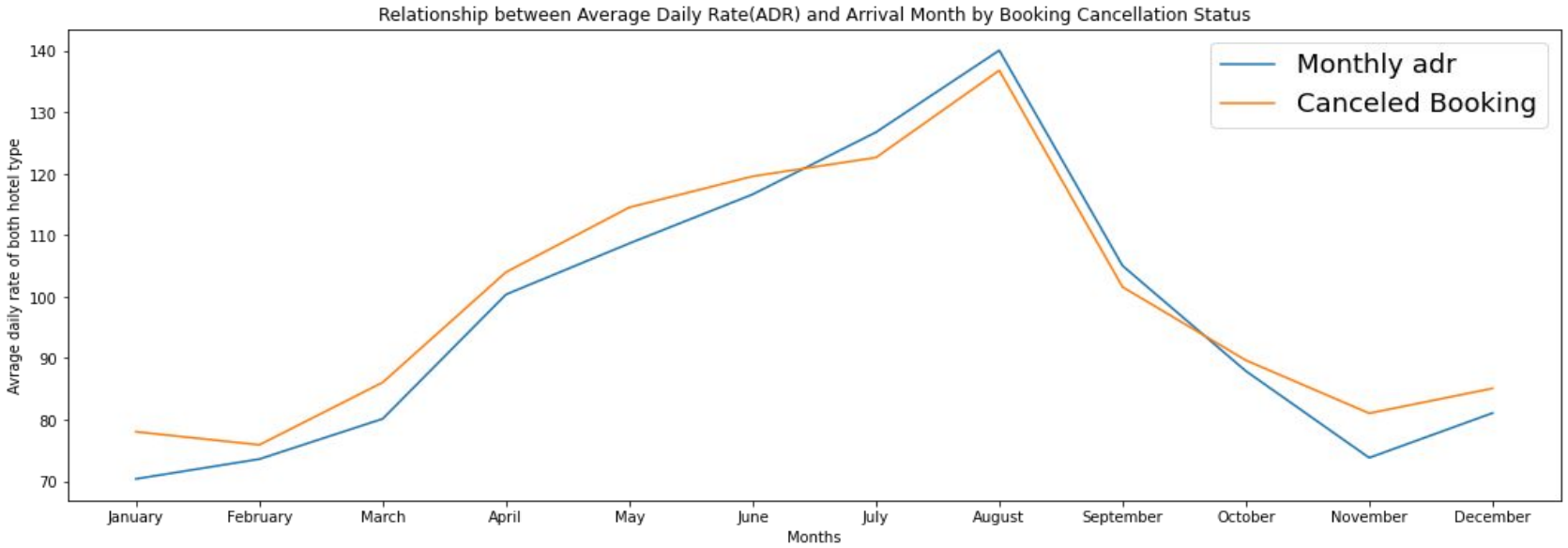


ADR Vary Over the Year by Hotel Type



At the Resort Hotel, rooms are more demanding during July & August, and for the City Hotel, rooms are more demanding during April,May & June. Also, over the years, city hotels have been more consistent than resort hotel in generating adr.

- **Relationship between Average Daily Rate(ADR) and Arrival Month by Booking Cancellation Status**



The highest Average Daily Rate (ADR) has occurred in August and, due to the highest ADR in August, rooms are more demanding and unavailable is also one of the reasons for more cancellations in August.

Conclusion :-

From the analysis of five problem parameters we come to a conclusion that ‘Transient’ type of customers are the most frequent, repeating, making most number of changes in the bookings, availing most special requests and requiring car parking spaces the most. The City Hotel is generating more average daily rate as it is getting more number of bookings (around 60% of the bookings) as well as cancellations than the Resort Hotel. August is seen to be the busiest month due to which the hotels are generating more average daily rate than rest of the months over the years and in may be the also the reason of most number of cancellations in the month of August. The hotels are attending most of their guests from Portugal followed by Great Britain and France. The bookings as well as cancellations to the hotels are mostly being done by the online TA segment through TA/TO channel. ‘Type 1’ special request is mostly being done by the customers and ‘BB’ type meal is the most preferred by the customers.

THANK YOU