

Capstone Project-1

Hotel Booking Analysis

By

Neeraja C

Data Science Trainee|AlmaBetter

OVERVIEW

INTRODUCTION

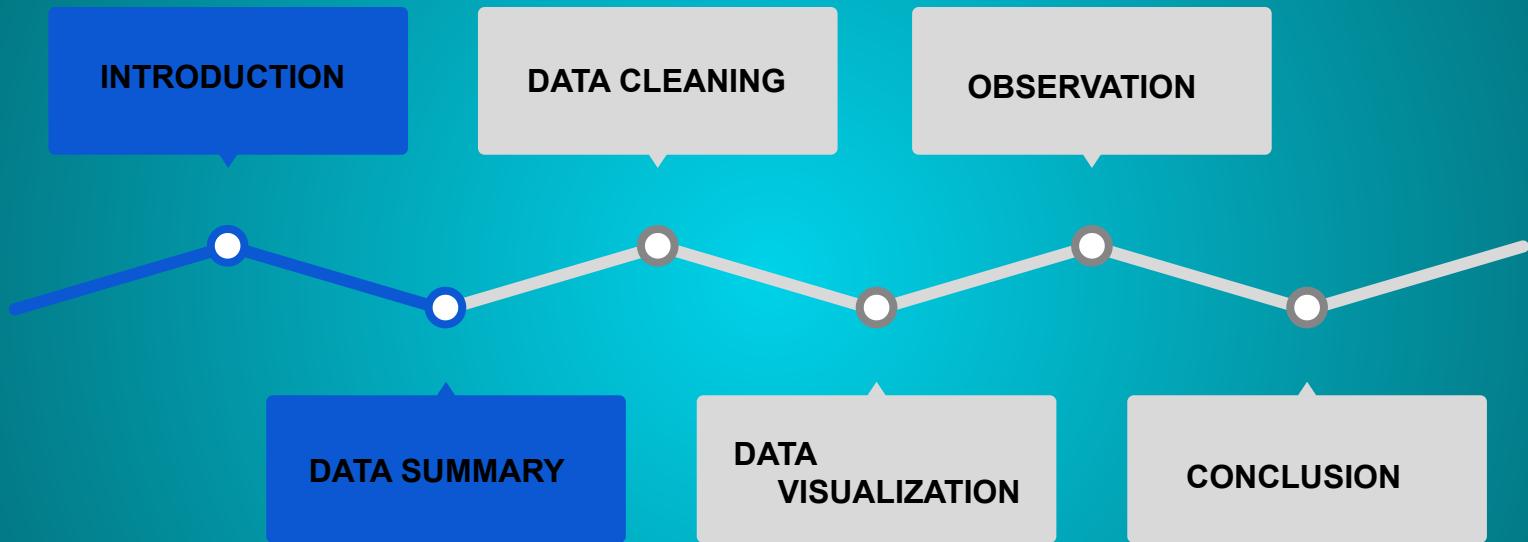
DATA CLEANING

OBSERVATION

DATA SUMMARY

DATA
VISUALIZATION

CONCLUSION







INTRODUCTION

- ❖ Hotel industry is a very volatile industry and the booking depends on many factors.
- ❖ The main objective behind this project is to explore and analyse data to discover important factors that govern the bookings and give insights to hotel management, which can perform various campaigns to boost the business and performance.
- ❖ We are here to explore a hotel booking dataset to discover important factors that govern the bookings.

PROBLEM STATEMENT

- ❖ Have you ever wondered when the best time of year to book a hotel room is? Or the optimal length of stay in order to get the best daily rate?
- ❖ What if you wanted to predict whether or not a hotel was likely to receive a disproportionately high number of special requests?
- ❖ This hotel booking data set can help you explore those questions!

DATA SUMMARY

	NUMBER OF ROWS	119390
	NUMBER OF COLUMNS	32
	MISSING VALUES	129425
	MISSING VALUES(%)	3.4%

COLUMN DETAILS

hotel: The names of the hotel are City Hotel and Resort Hotel

- **is_canceled:** Cancellation type, if the booking was cancelled or not.

Which takes 2 values 0 and 1.

0 indicates not cancelled.

- **lead_time:** Time between reservation and actual arrival.
- **arrival_date_year:** Year of arrival date
- **arrival_date_month:** Month name of arrival date.
- **arrival_date_week_number:** Week number of arrival date

COLUMN DETAILS

arrival_date _day_of_month: Day of the month of arrival date

stays_in weekend_nights: Number of weekend nights the guest stayed or booked to stay at the hotel

stays_in_week_nights: Number of week nights the guest stayed or booked to stay at the hotel

adults: Number of adults

children: Number of children

babies: Number of babies

meal: Type of meal booked

country: Country of origin of customer

market_segment: Market segment designation.

COLUMN DETAILS

market_segment: Market segment designation. In categories, the term “TA” means “Travel Agents” and “TO” means “Tour Operators”.

distribution_channel: The medium through booking was made

previous_cancellations: (0 or 1) Indicates whether or not the guest has previous cancellations

is_repeated_guests: Value indicating if the booking name was from a repeated guest (1) or not (0)

previous_cancellations: (0 or 1) Indicates whether or not the guest has previous cancellations

previous_bookings_not_canceled: (0 or 1) Indicates whether or not the guest has previous bookings which are not cancelled.

booking_changes: Number of changes made to the booking from the moment the booking was entered on the PMS until the moment of check-in or cancellation

reserved_room_type: Code of room type reserved.

COLUMN DETAILS

- assigned_room_type:** Code of room type assigned.
- deposit_type:** Whether refundable/non-refundable/no-deposit made
- agent:** ID of the travel agency that made the booking
- days_in_waiting_list:** number of days the booking was in the waiting list before it was confirmed
- customer_type:** Type of customers (Transient, group, etc.)
- adr:** Average daily rate is the average revenue that a hotel receives for each occupied guest room per day
- required_car_parking_spaces:** Number of car parking spaces required
- total_of_special_requests:** Number of special requests made
- reservation_status:** Reservation last status
- reservation_status_date:** Date of last reservation status

DATA VISUALIZATION

Here we are using Matplotlib and Seaborn library and the following graph and plots had been used.

Bar Plot

Histogram

Pie Chart

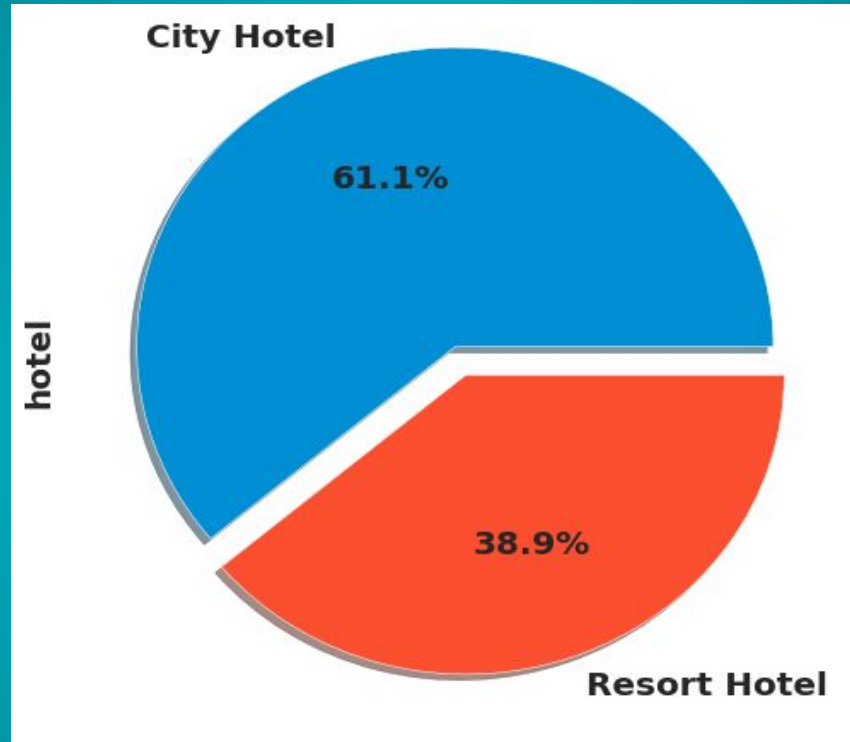
Line Plot

Heatmap

Box Plot

PERCENTAGE OF GUESTS IN EACH HOTEL

AI



More than 60% of the bookings are of City Hotel. Resort hotel tend to be on the expensive side and most people will just stick with City hotel.

CANCELLATION OF BOOKING



More than 70% of the people did not cancel the booking. Majority of bookings were not canceled, still some half of the bookings were canceled.

CANCELLATION OF BOOKING



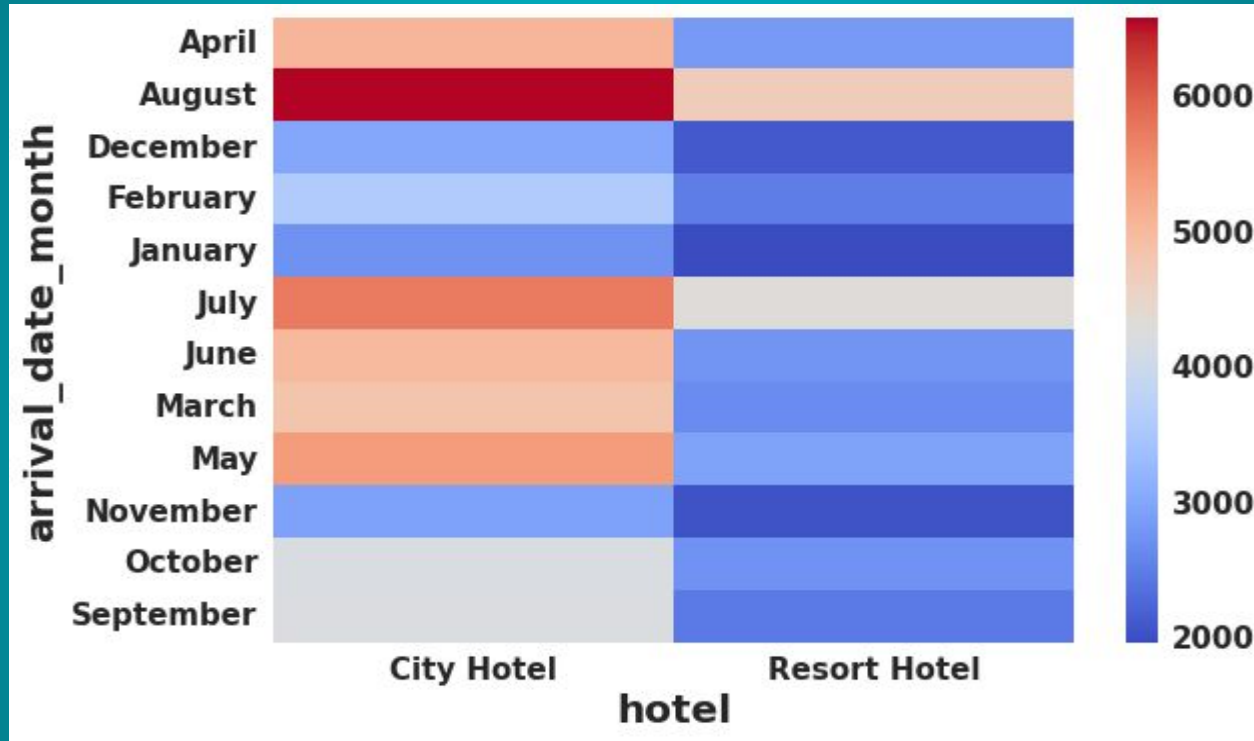
When lead time increases, the chances for cancellation increases. Bookings with a lead time less than 100 days have fewer chances of getting canceled.

CANCELLATION OF BOOKING



City hotel has more number of cancellation compared to Resort hotel. Most of the people prefer city hotel & most of the cancellations occur in city hotel.

BUSIEST MONTH



August is the most busiest month followed by July and May..Also the months after september, are the least demanded months.

NUMBER OF BOOKINGS OVER YEAR



2016 has the highest number of bookings.

MOST BUSY MONTHS



Both the hotels have more guests in the months of June, July and August. After August, we can see that the number of guests are decreasing.

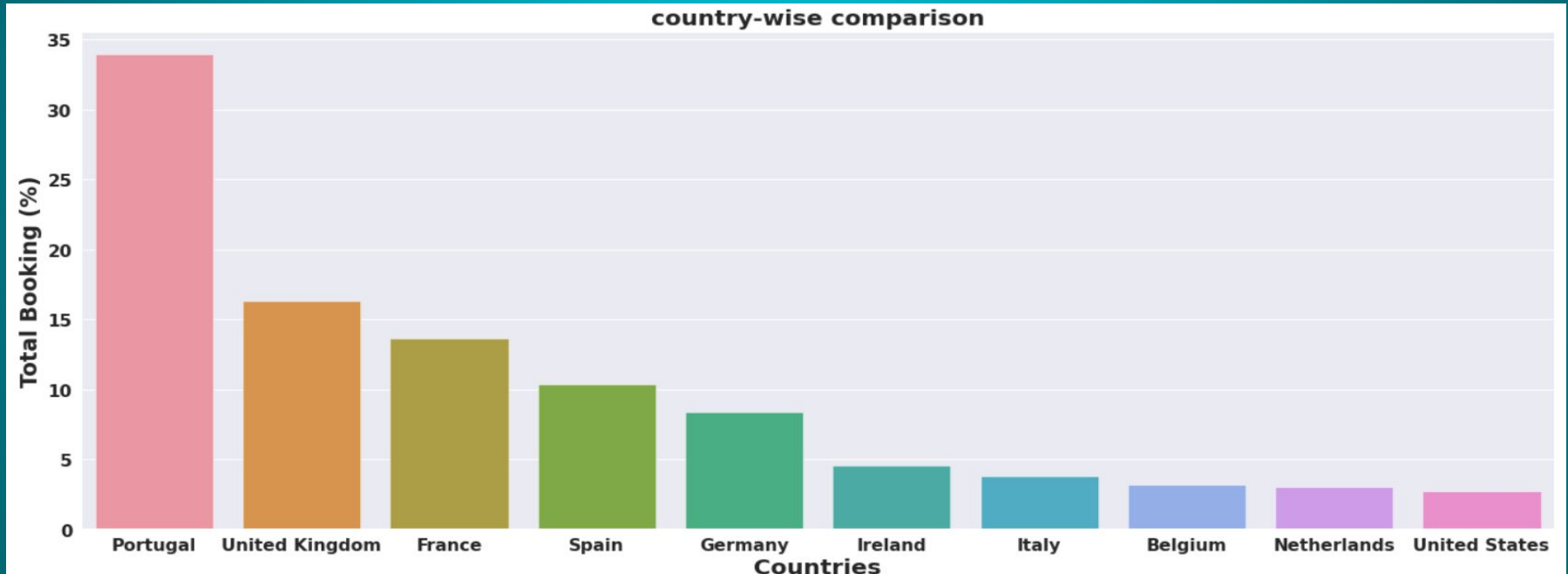
PRICE VARIATION

Room price per night over the Months



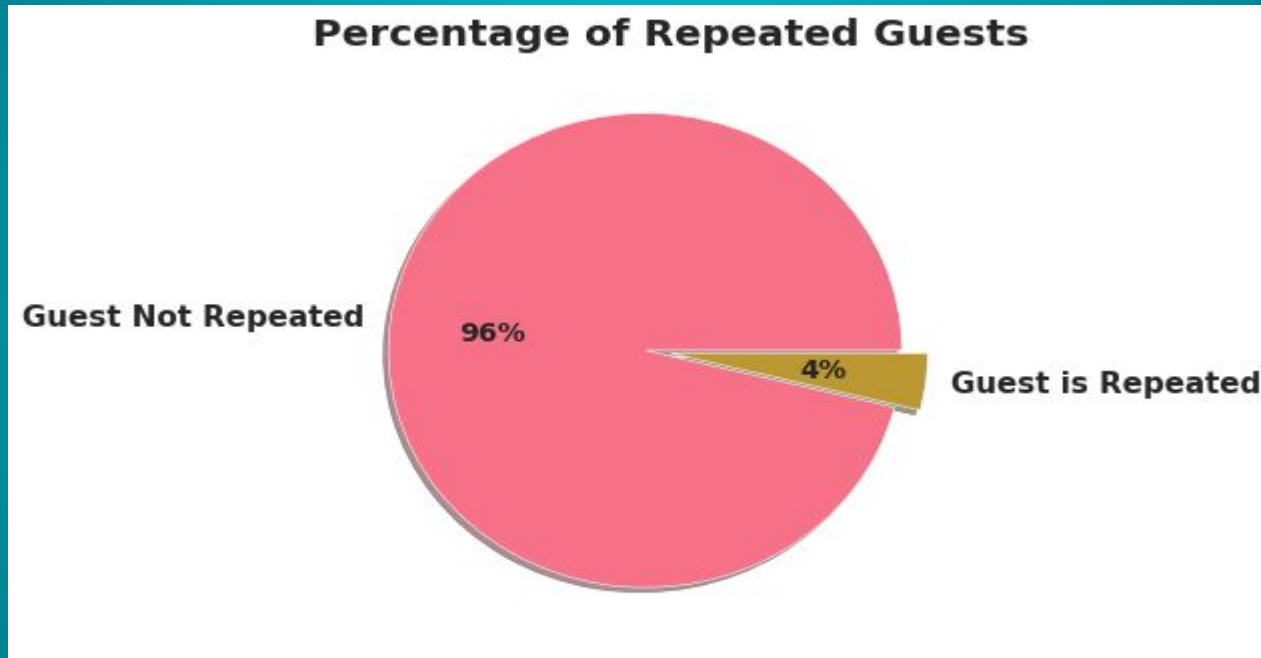
The plot clearly shows that prices in the resort hotel are much higher during the summer and the prices of city hotel varies less and is most expensive during summer and spring.

TOP 10 COUNTRIES WITH MAX NUMBER OF PASSENGERS



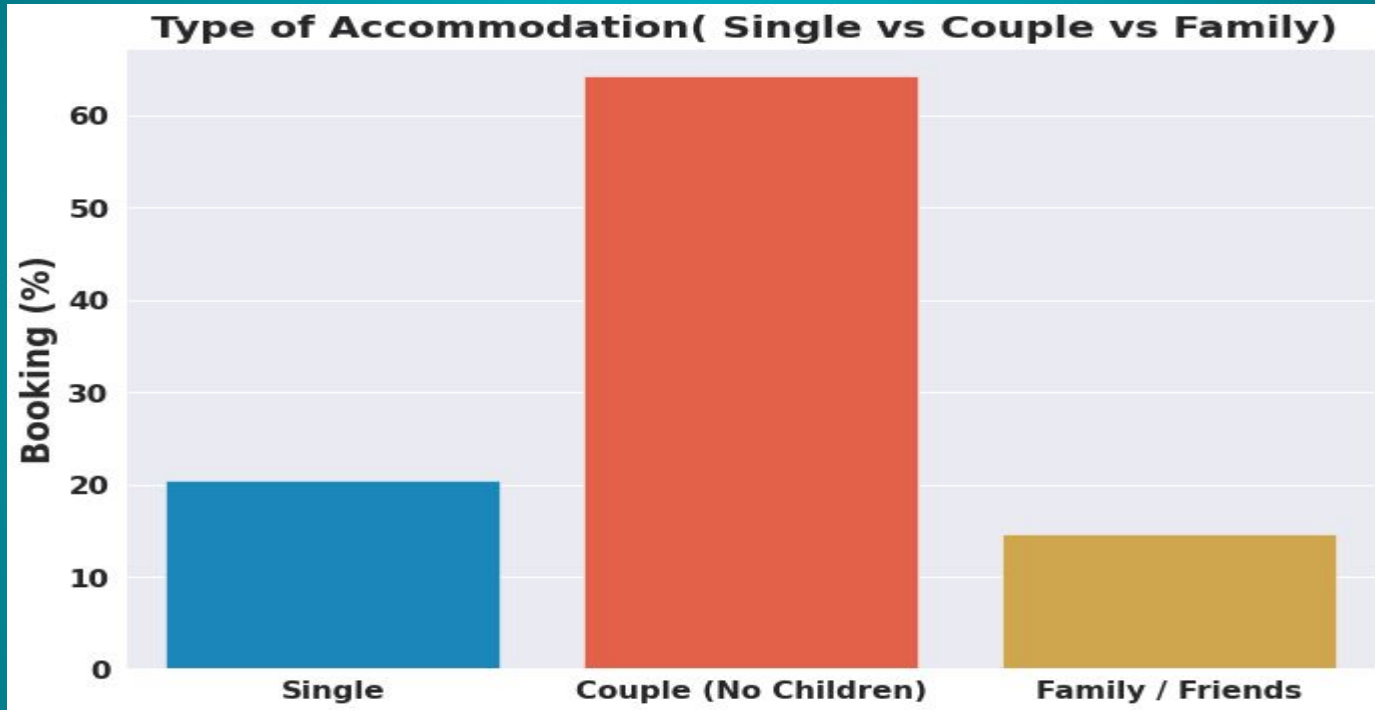
Most guests are from Portugal and other countries in Europe.

REPEATED GUESTS



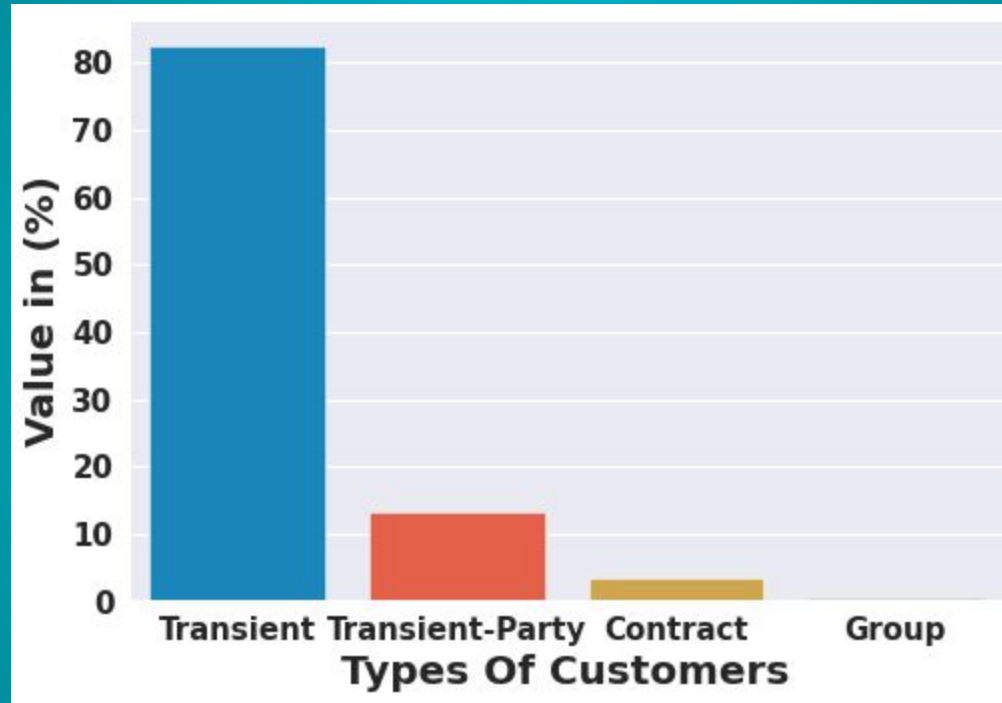
The chart shows only 4% guest is repeated in hotel. It means 3358 guests out of 83981 is repeated. A need to target repeated guests since they have booked before.

GUEST TYPE



Around 65% of the people visit with their family. While most of the visitors are couples. Booking of singles and people with family/friends are less as compared to couples.

CUSTOMER TYPE



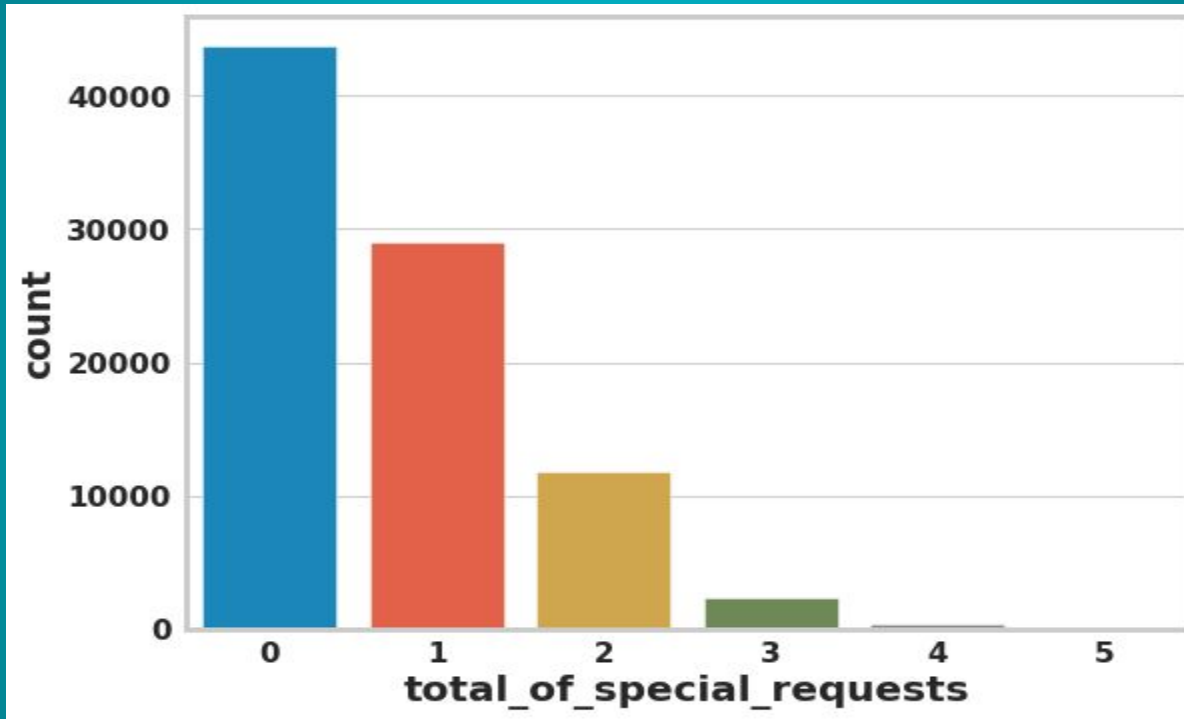
Most customers are of type Transient. They are simply individual guests requiring a short stay at the hotel. Such guests are usually not looking for anything more than a basic service.

NIGHT STAY DURATION



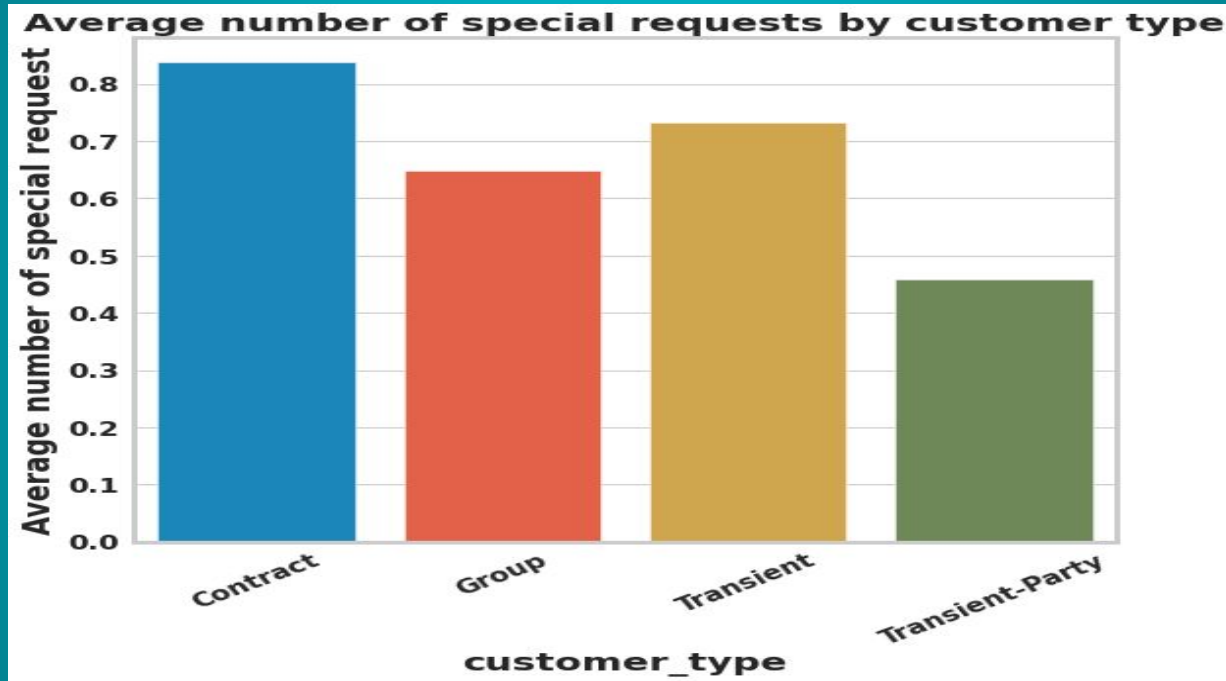
In Resort hotel people like to stay 1 day and in City hotel people like to stay 2-3 days. As number of days increases, people prefer Resort hotel the most. As number of days increases, the number of booking in both City hotel and Resort hotel decreases.

SPECIAL REQUESTS



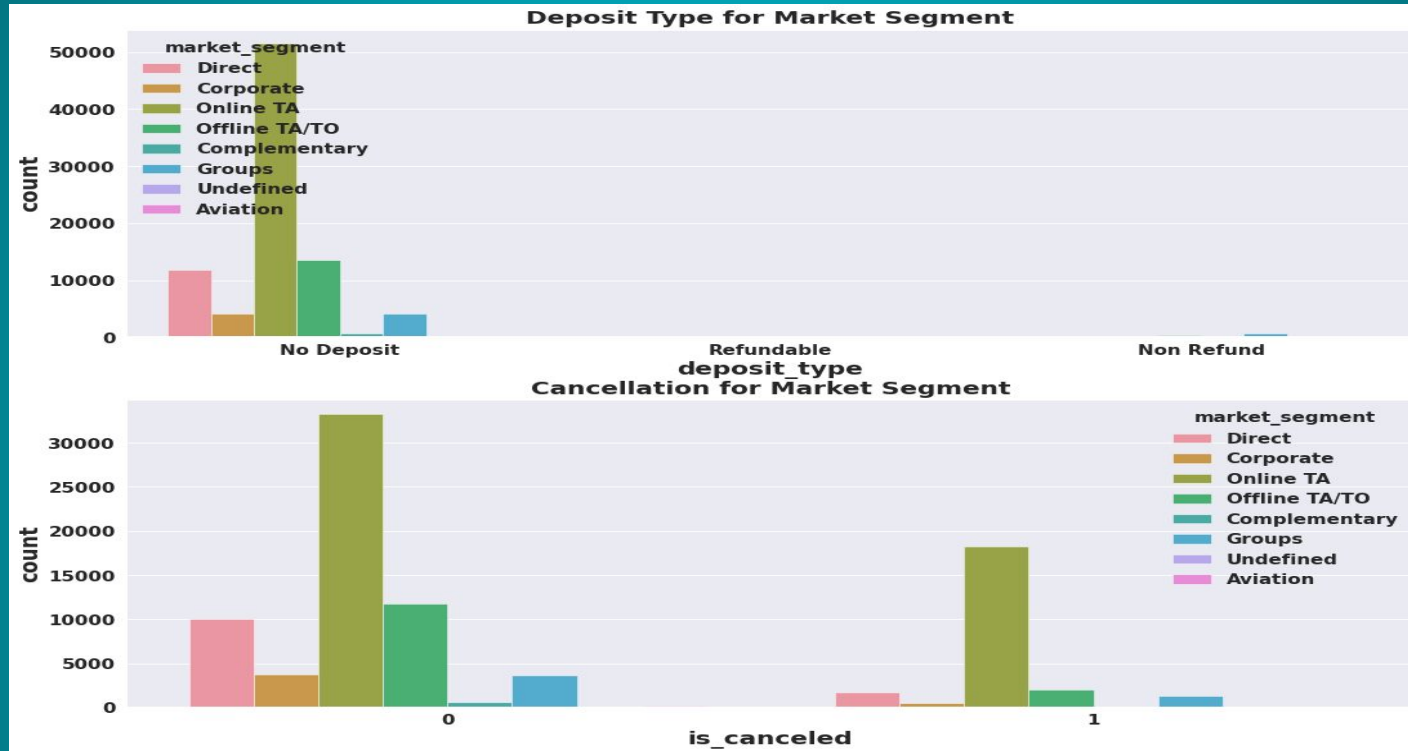
Majority of people does not require any special requests. The maximum number of special requests is 4.

SPECIAL REQUESTS



They are not entertaining much special requests. Contract people have more special requests.

BOOKING AND MARKET SEGMENT



Most of the bookings are done through online TA segment and most cancelation is done through online TA segment only. People don't want to pre-deposit the money. No deposit may lead to cancel the bookings.

DISTRIBUTION CHANNEL



The most preferred distribution channel is TA/TO.

CONCLUSION

- Majority of hotels booked are City hotel. Definitely need to spend the most targeting fund on those hotel.
- We found out there is huge demand for City Hotel and Resort Hotel in months of June, July and August. In other side, we found out that the prices in the above months are higher. So the best time to book the hotel will be between Nov-Mar, where the demand is low and prices too. The user can get better price and room easily.
- In a business perspective, we should focus more on summer season due to high demand.
- If people stay for less than 5 days, they will get the best price.
- They are not entertaining much special requests.

CONCLUSION

- When users book the hotel much earlier, they are more likely to cancel the booking due to various reasons. So we have to reduce lead time to decrease cancellation rate.
- We also realise that the high rate of cancellations can be due to high no deposit policies.
- Majority of guests are from Western Europe. we should spend a significant amount of our budget on those area.
- Most of the bookings are coming through online.
- Most of the bookings are of 2 adults.
- Given that we do not have repeated guests, we should target our advertisement on guests to increase returning guests.

thank
you