

Capstone Project

HOTEL BOOKING ANALYSIS

Team Members

Shubhashis Singh
Sagar Balu Kumbhar
Ashwini Verma
Abhijeet Roy

Points of Discussion:

- Define Problem statement.
- Data Pipeline.
- Data Summary.
- EDA (Exploratory Data Analysis).
- Observation.
- Conclusion.
- Strategies.

Problem Statement :

To explore and analyze the data to discover important factors that can impact growth of the hotel bookings.

Data Summary :

The provided data set has following different columns of variables necessary for hotel bookings:

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

is_cancelled : The value of column shows if the booking was cancelled or not. Values[0,1], where 0 indicates not cancelled and 1 indicates the cancellation.

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, which are:

Undefined/SC – no meal package

BB – Bed & Breakfast

HB – Half board (breakfast and one other meal – usually dinner)

FB – Full board (breakfast, lunch and dinner)

Data Summary :

Country: The origin country of guest

market_segment: This column shows how reservation was made and what is the purpose of reservation. Eg, corporate means corporate trip, TA for travel agency and TO for tour operators.

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

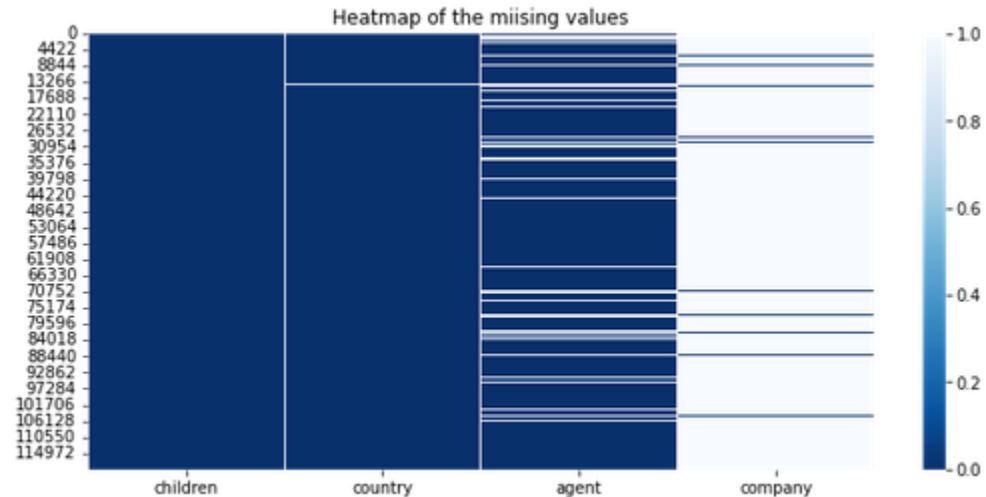
Is_repeated_guest: Shows if the guest have arrived the hotel before or not. Values[0,1]-->0 indicates no, and 1 indicates yes.

days_in_waiting_list: Number of days between actual booking and the transaction.

customer_type: Type of customers(Contract, group, transient, transient party, etc.)

Data Pipeline

- **Data Processing:** In Data Processing we have imported necessary libraries and data set. We used these libraries to understand the data and presenting the necessary information in the form of Charts.
- **Data Cleaning:** After understanding the data, we got to know about the null values and discrepancies in the data. We replaced the null values with 0 for numerical data and 'undefined' for categorical data. We removed unnecessary columns which have most of the null values.



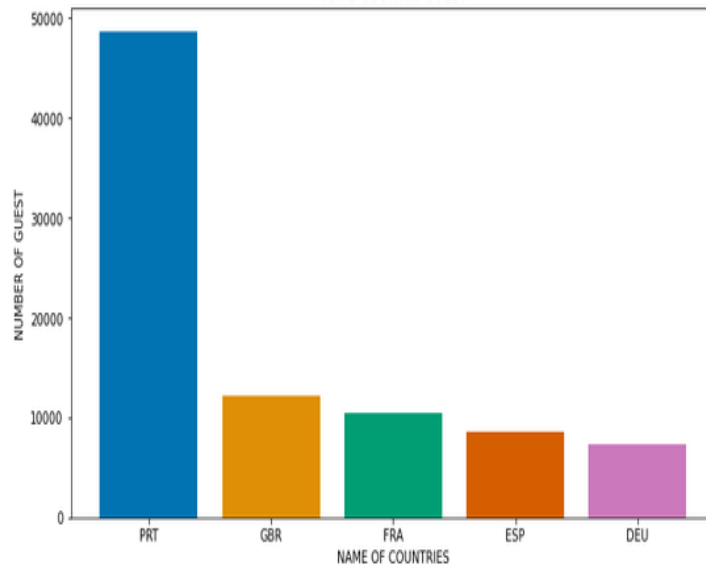
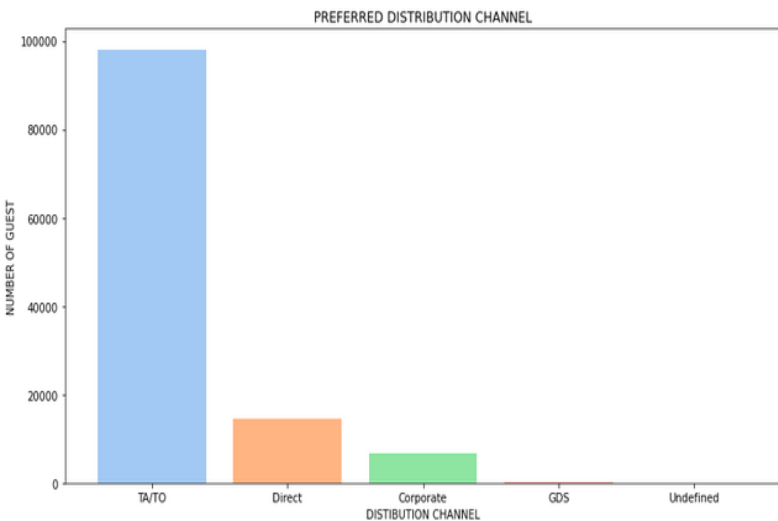
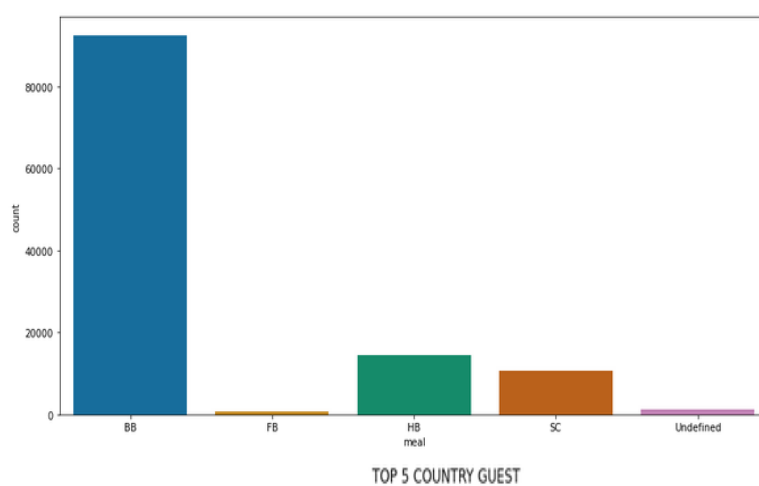
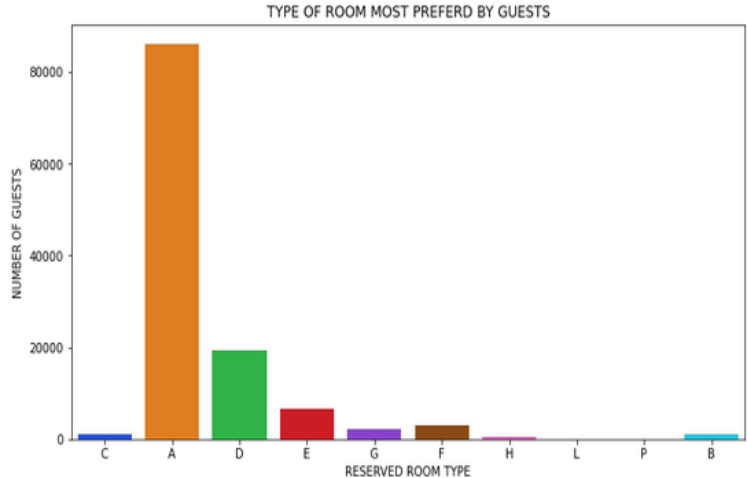
Data Pipeline

Exploratory Data Analysis

Univariate Analysis:

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

- Which distribution channel gave most of the bookings?
- Which room type is in most demand?
- From which country most of the customers are coming?
- What is the most preferred meal by customers?

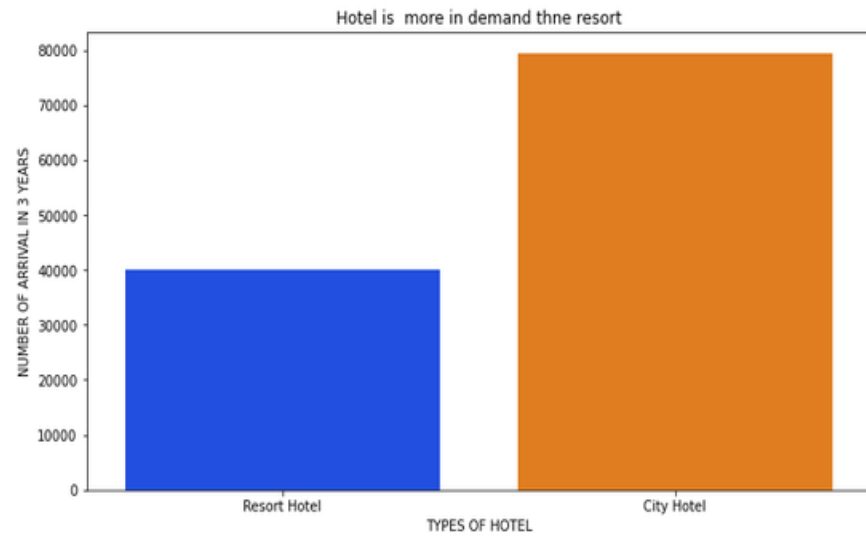
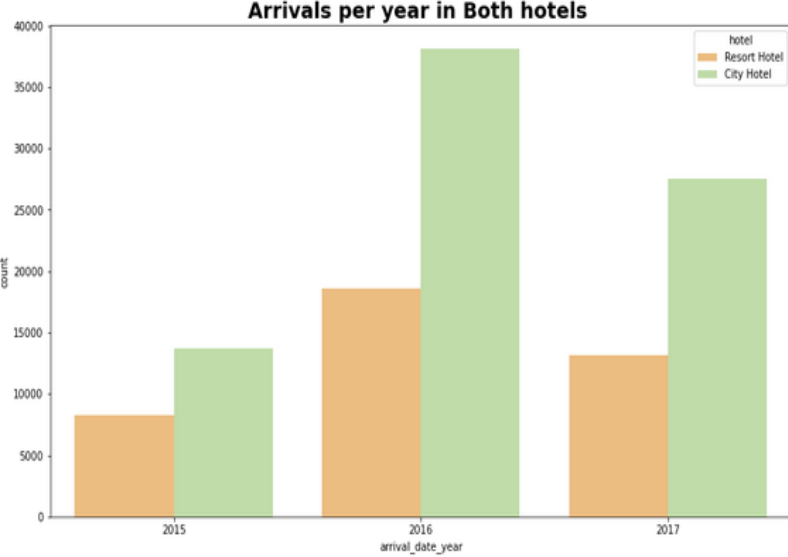


- Type A Room is preferred by customer.
- More than 70% of hotels provide bed and breakfast
- 40% guests are from Portugal
- Most of channel distribution is through Ta/To, which is more than 80%

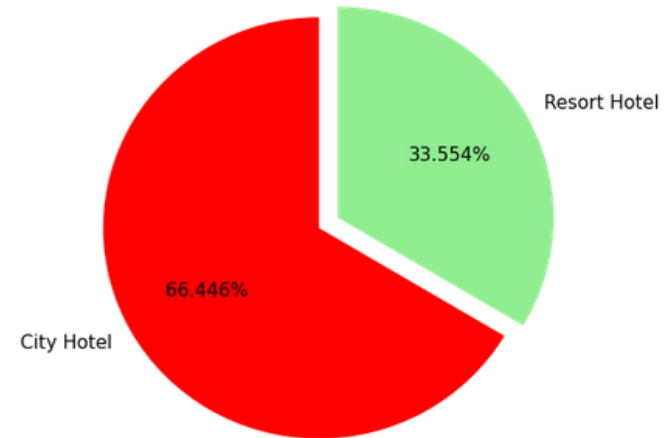
Hotel wise Analysis

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

- Which type of hotel was more engaging and in most demand?
- Which type of hotel receives more guests?
- Which type of hotel makes more revenue?
- What type of guests have most check-ins?
- What is most preferred stay length in each hotel?
- Which hotels have higher and how much customer returning rate?

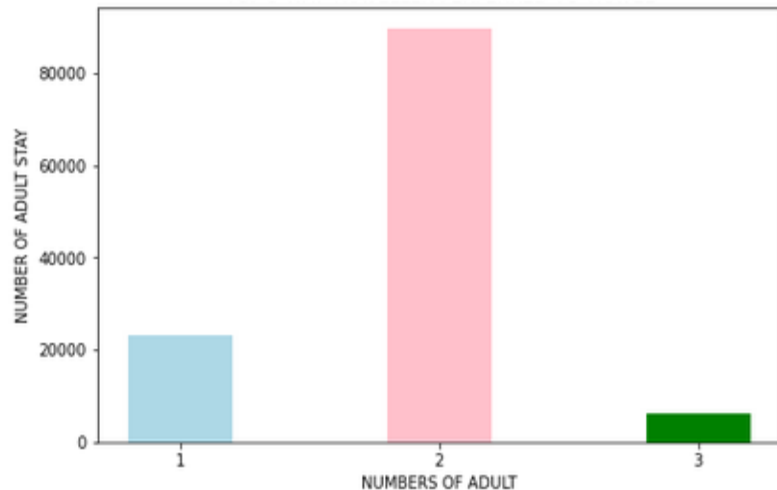


- With the above graph, it is evident that city hotels are more engaging and in demand. Hence, city hotels are more profitable.



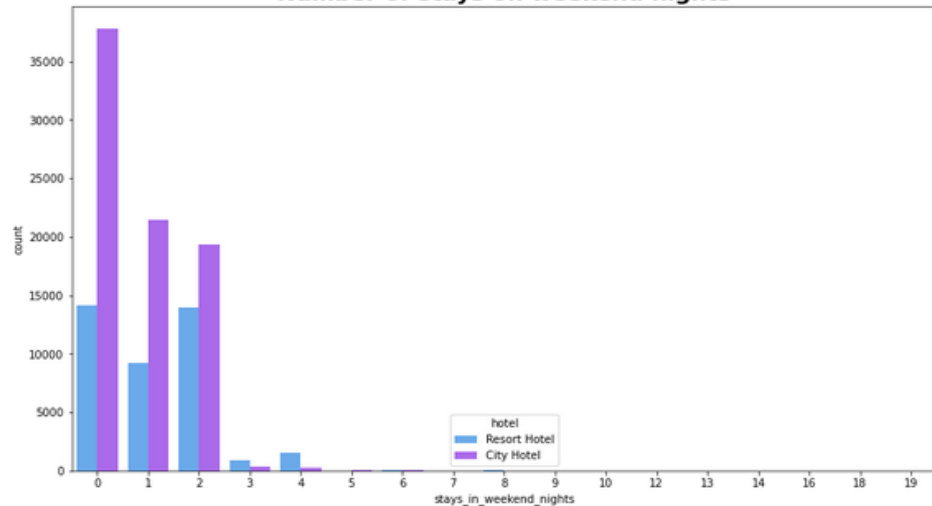
66% customer staying in city hotel.

TOP 3 WAY TRAVELLER PERFERRED TO TRAVEL

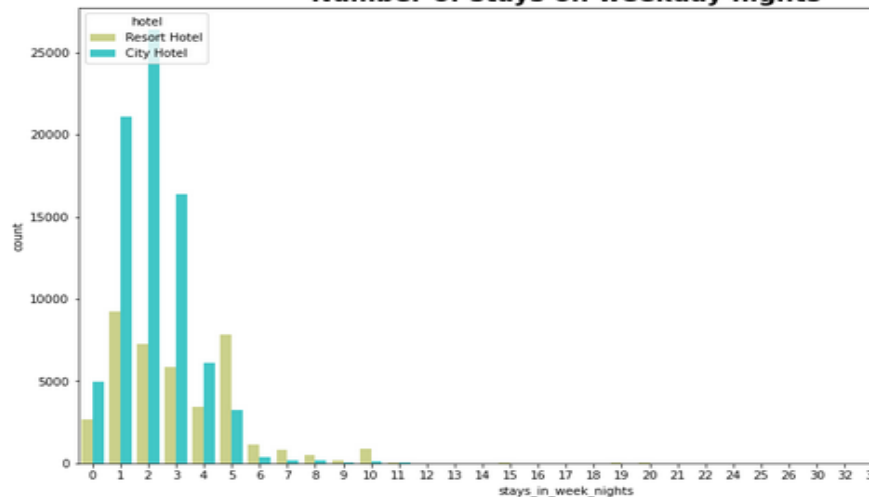


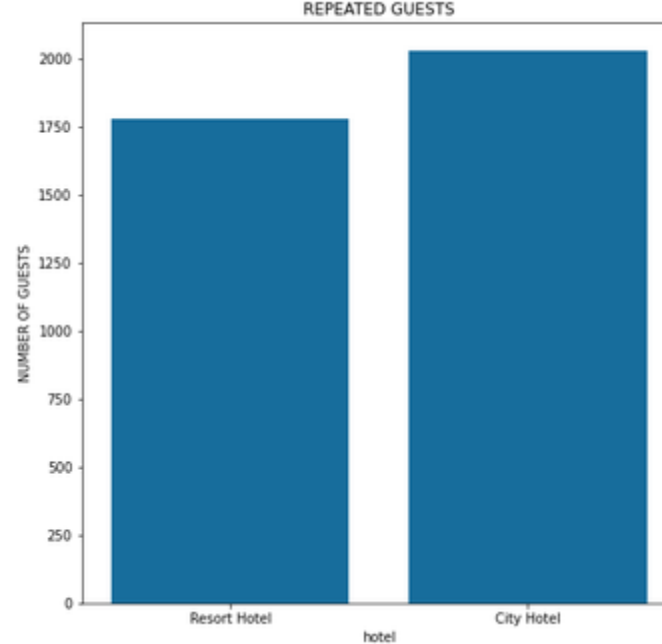
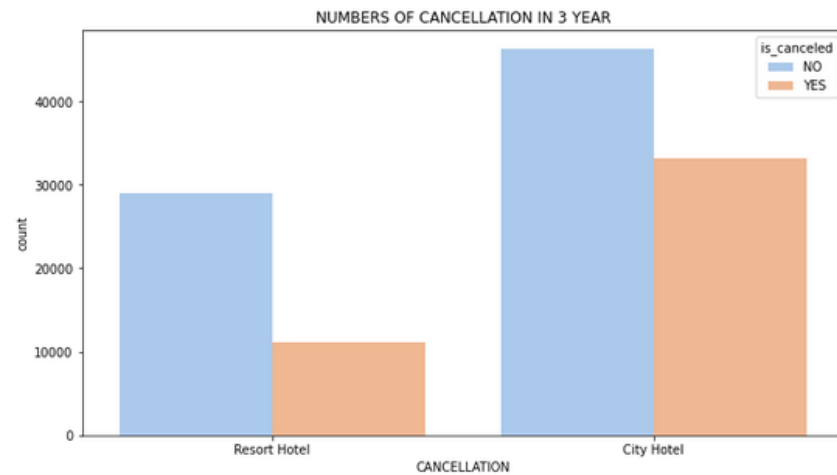
- Couples have more check-ins than any type of guests.
- Majority of the travellers who arrived at the hotel on weekend left the hotel on the same day.
- Majority of the travellers who arrived at the hotel on week day left the hotel next day

Number of stays on weekend nights



Number of stays on weekday nights





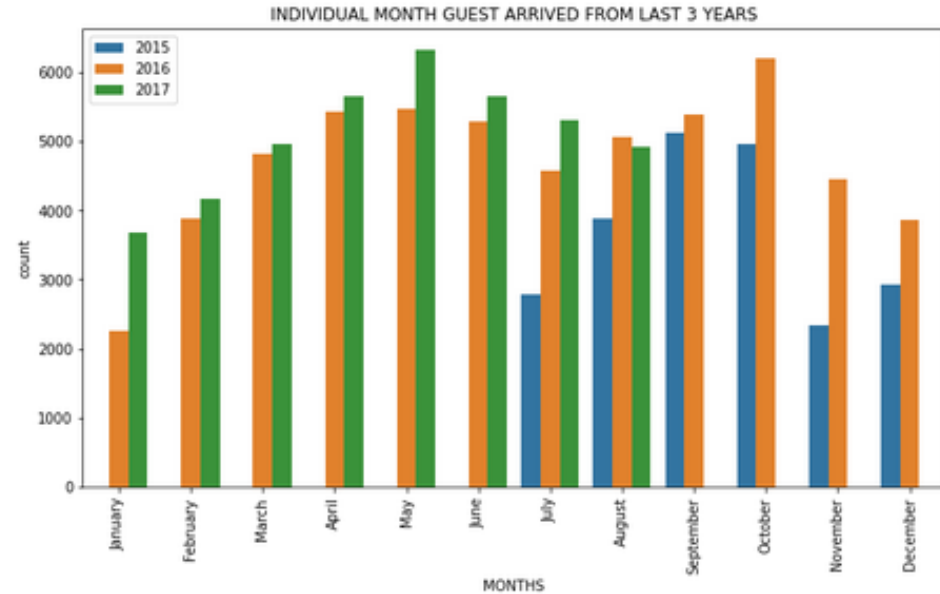
- City hotels receives a higher number of guests than the resort hotels. Cancellation rate is more for city hotels.
- City hotels have a higher number of repeat customers. However, the ratio of repeat customers is more for resort hotels as the city hotels receives a higher number of guests than the resort hotels.

Time wise Analysis

While doing time-wise analysis of given hotel booking dataset, we answered the following question:

What are the busiest months for hotels?

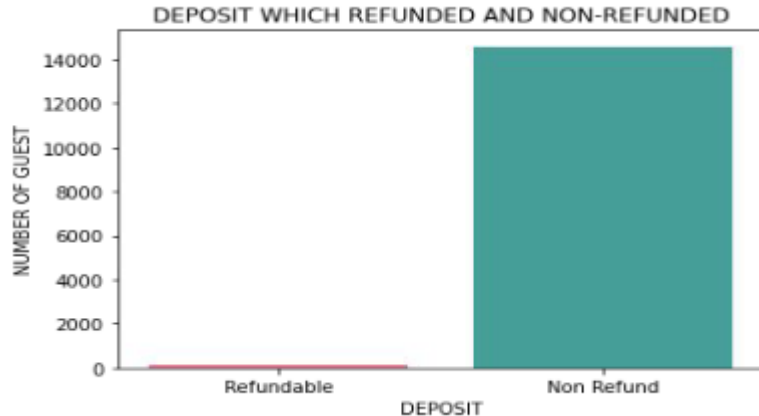
- May is the busiest and profitable month for the hotels in 2017.
- On an average, May-June are the most busier and profitable month for the hotels



Some other important questions

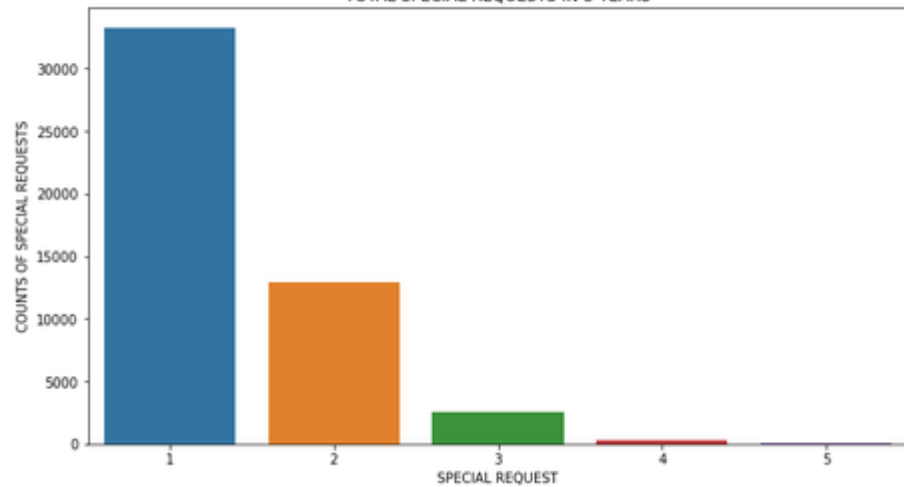
Some other analysis are also done, which are as follows:

- How is the deposit type affecting the profit?
- Which hotel type received more special requests?
- How many special requests were received?

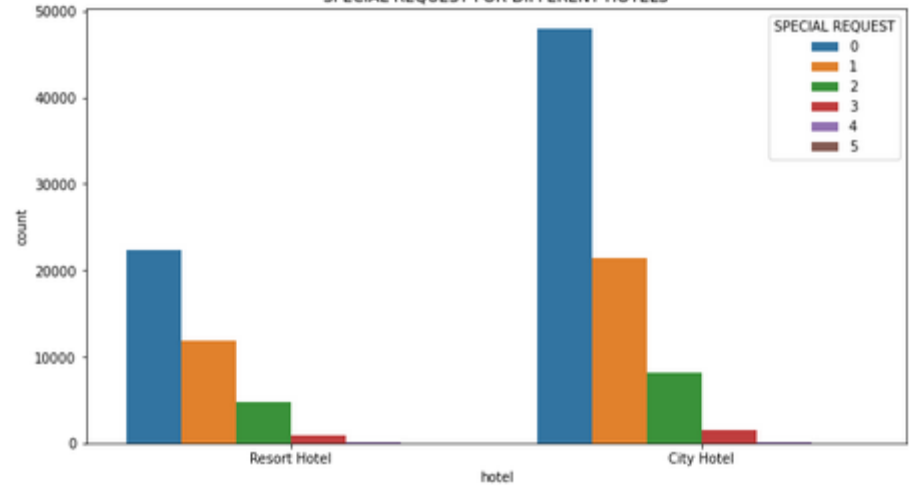


- Non-refundable deposits tends to convert to a successful transaction.

TOTAL SPECIAL REQUESTS IN 3 YEARS

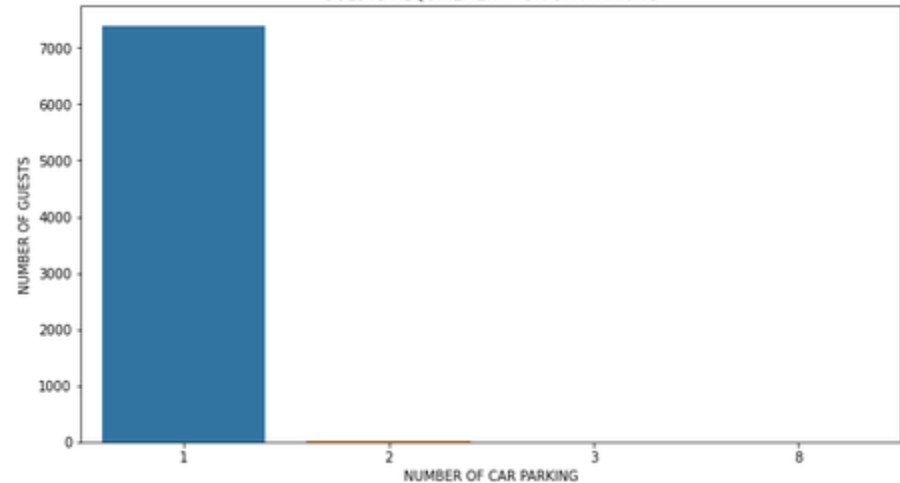


SPECIAL REQUEST FOR DIFFERENT HOTELS



- City hotels have received the most special requests.
- 1 number of special requests is the highest.
- 1 car parking was needed for most of the guests.

GUESTS REQUIREMENT FOR CAR PARKING



Observations Summary

- Type A room is in most demand.
- Most booking were made from online Travel Agents.
- Majority customers are from Portugal.
- Most customers prefer Bed and breakfast meal type.
- City hotels are more engaging and in demand. Hence, city hotels are more profitable.
- City hotels receives a higher number of guests than the resort hotels. Therefore, the cancellation rate is more for city hotels.
- City hotels have a higher number of repeat customers. However, the ratio of repeat customers is more for resort hotels as the city hotels receives a higher number of guests than the resort hotels.
- Non-refundable deposits tends to convert to a successful transaction.
- May is the most busiest and profitable month for the hotels in 2017.
- On an average, May-June are the most busiest and profitable month for the hotels.
- City hotels have received the most special requests.
- 1 car parking was needed for most of the guests.

Conclusion

- Around 60% bookings are for City hotel and 40% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel.
- Mostly guests stay for less than 5 days in hotel, and for longer stays Resort hotel is preferred. City hotels can offer bundled offers like 2 days and 3 nights, or 4 days and 5 nights, etc., in order to increase the customer stay.
- 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. Hotels can provide offers and discounts to the repeat customers in order to increase the number of repeat customers.
- Most of the guests came from European countries, with most no. of guest coming from Portugal. Hotels can work on attracting other countries guests with special offers and discounts along with organizing an event, etc.
- Guests use different channels for making bookings out of which most preferred way is TA/TO. Hotel can offer incentives for higher bookings.
- Non-refundable bookings tend to convert to a successful transaction.
- May is the busiest and profitable month for the hotels in 2017.
- On an average, May-June are the most busier and profitable month for the hotels.
- Couples are the most common guests for hotels. Hence, hotels can plan services according to couples needs to increase revenue.

Strategies to Counter High Cancellations at the Hotel

1. Set Non-refundable Rates, Collect deposits, and implement more rigid cancellation policies.
2. Encourage Direct bookings by offering special discounts
3. Monitor where the cancellations are coming from such as Market Segment, distribution channels, etc.

Thank you