

Capstone Project

Team 1 : Hotel Booking Analysis

Team Members

Chandrakala
Shubham Deshmukh
Pragnesh Hingu

OUTLINE

- Introduction
- Business Challenges
- System Overview
 - A) Data Preparation and analysis
 - B) Methodology
- Results
- Conclusion

INTRODUCTION

- Hotel industry is a very volatile industry and the bookings depend on variety of factors such as type of hotels, seasonality, days of week and many more.
- This makes analyzing the patterns available in the past data more important to help the hotels plan better.
- Exploratory data analysis will help us to better understanding of customer behavior using the historical data, hotels can perform various campaigns to boost the business.

Business Challenges

- Very less or zero number of Hotel Bookings during some months of a year.
- Booking cancellations.
- To deal with unexpected disproportionality of high number of special request.
- To reduce these uncertainties and maximize their product and revenue.

Dataset

Numerical Features :-

- is_canceled
- lead_time
- arrival_date_year
- arrival_date_week_number
- arrival_date_day_of_month
- stays_in_weekend_nights
- stays_in_week_nights
- adults
- children
- babies
- is_repeated_guest
- previous_cancellations
- previous_bookings_not_canceled
- booking_changes
- days_in_waiting_list
- adr
- required_car_parking_spaces
- total_of_special_requests

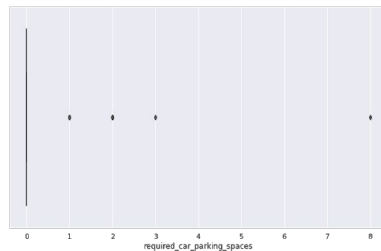
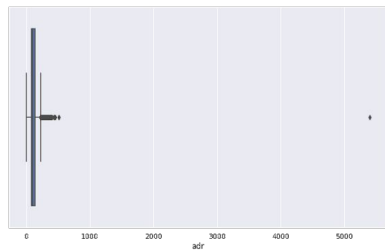
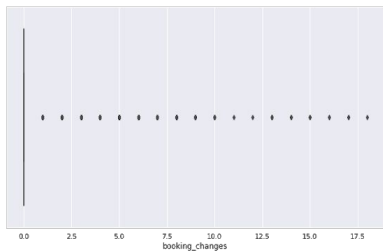
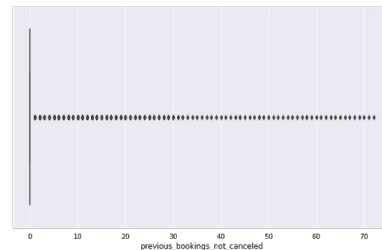
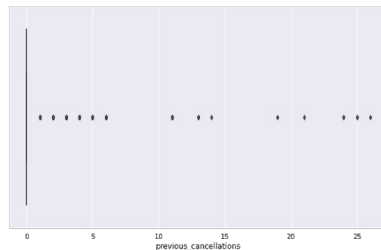
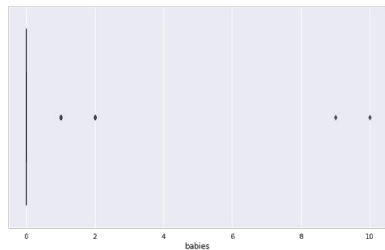
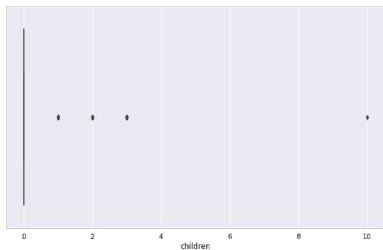
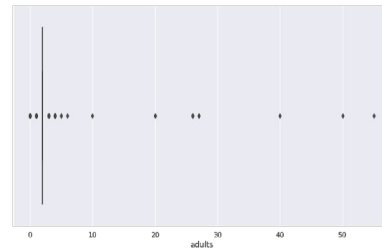
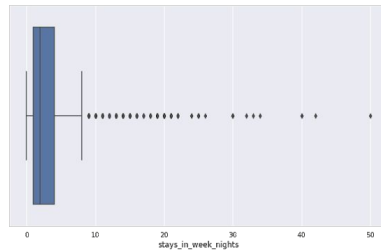
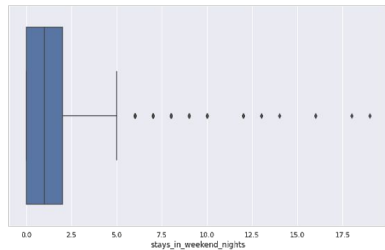
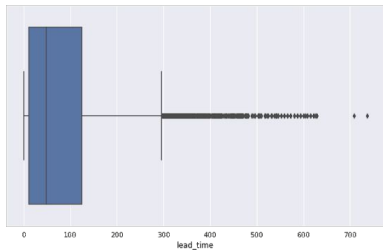
Categorical Features :-

- hotel
- arrival_date_month
- meal
- country
- market_segment
- distribution_channel
- reserved_room_type
- assigned_room_type
- deposit_type
- customer_type
- reservation_status.

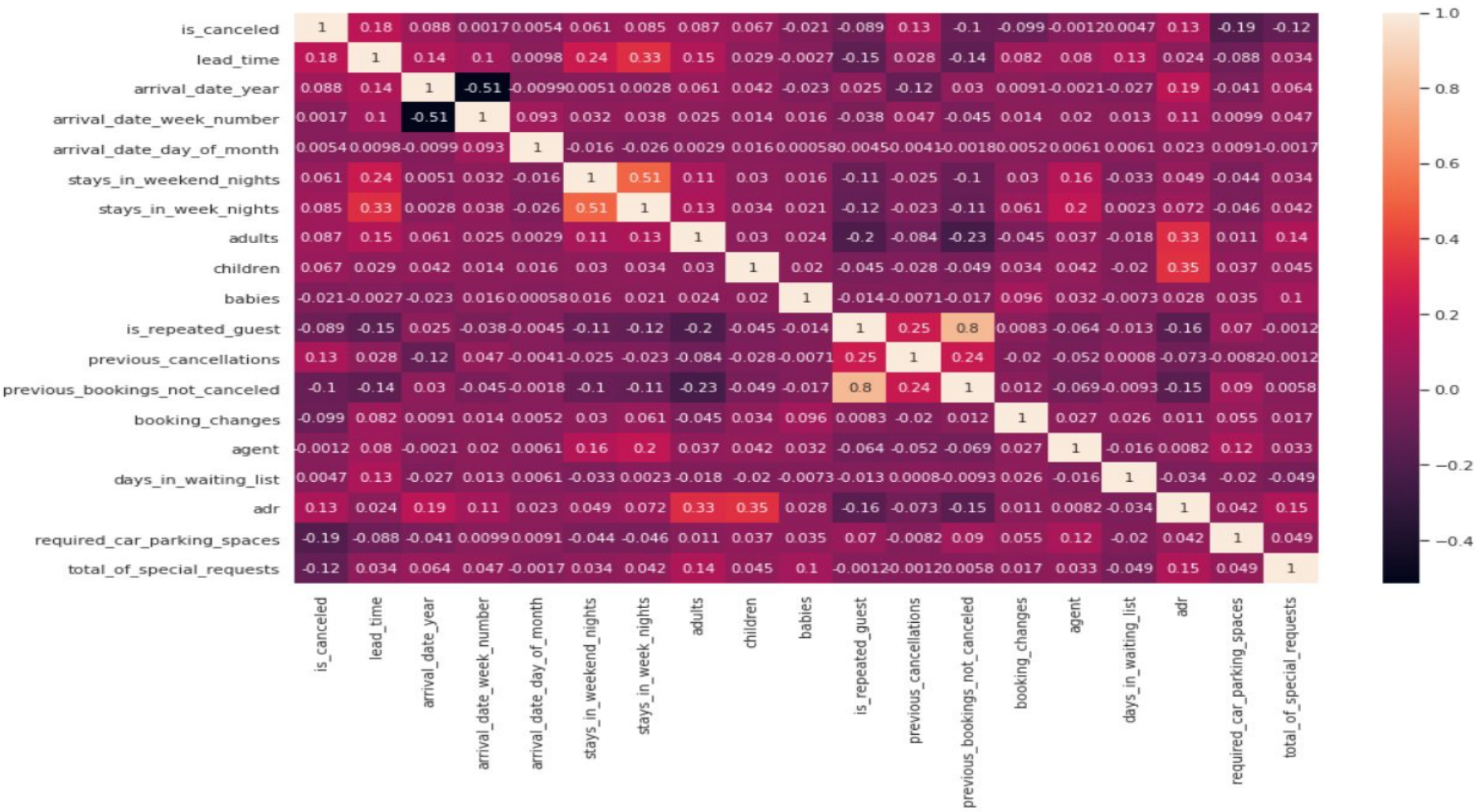
Data Pre-Processing

- Data Audit
Comprehensive first look at dataset.
- Missing Values
Dealing with Missing values
- Duplicate Rows
Removed 32,020 number of duplicate Rows to reduce the size of dataset.
- Converting Datatype
Converted reservation status date object type to datetime type.

Outliers



Correlation



For Simplicity purpose, we have divided the EDA to three parts

- EDA to understand the Customer's Preferences.
- EDA to Understand the booking Trend.
- EDA to get insights from Cancellation.

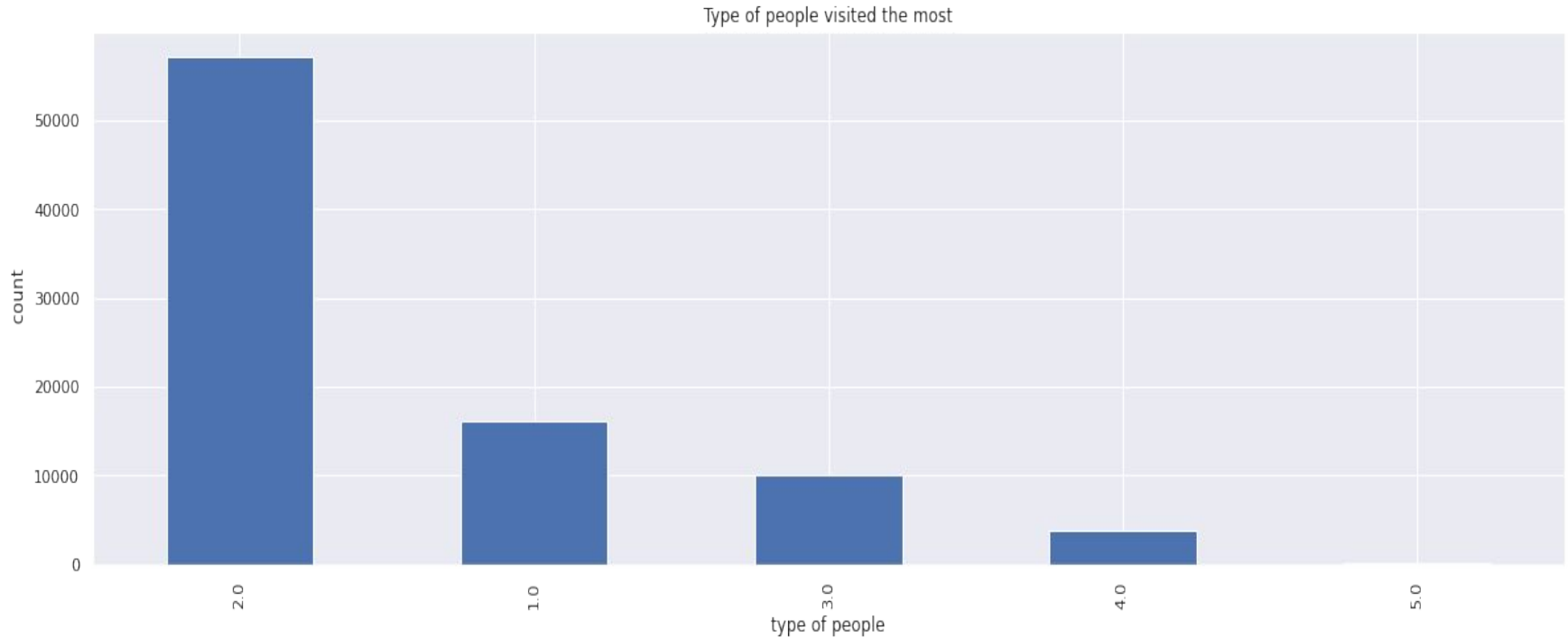
EDA on Customer's Preference's :

1.Which type of hotel people book in general ?



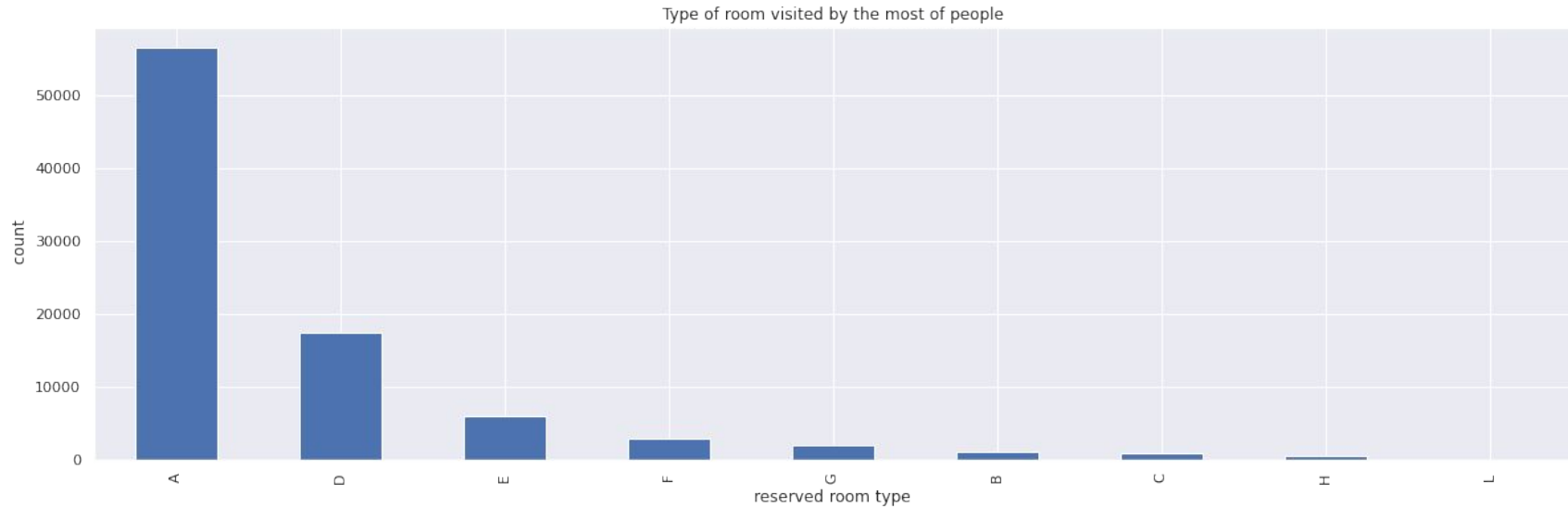
- The Above graph depicts information about which hotel type has occupied more guests.
- It's clear from the graph that city hotel was more occupied than the resort hotel.

2. which Type of people visited the most?



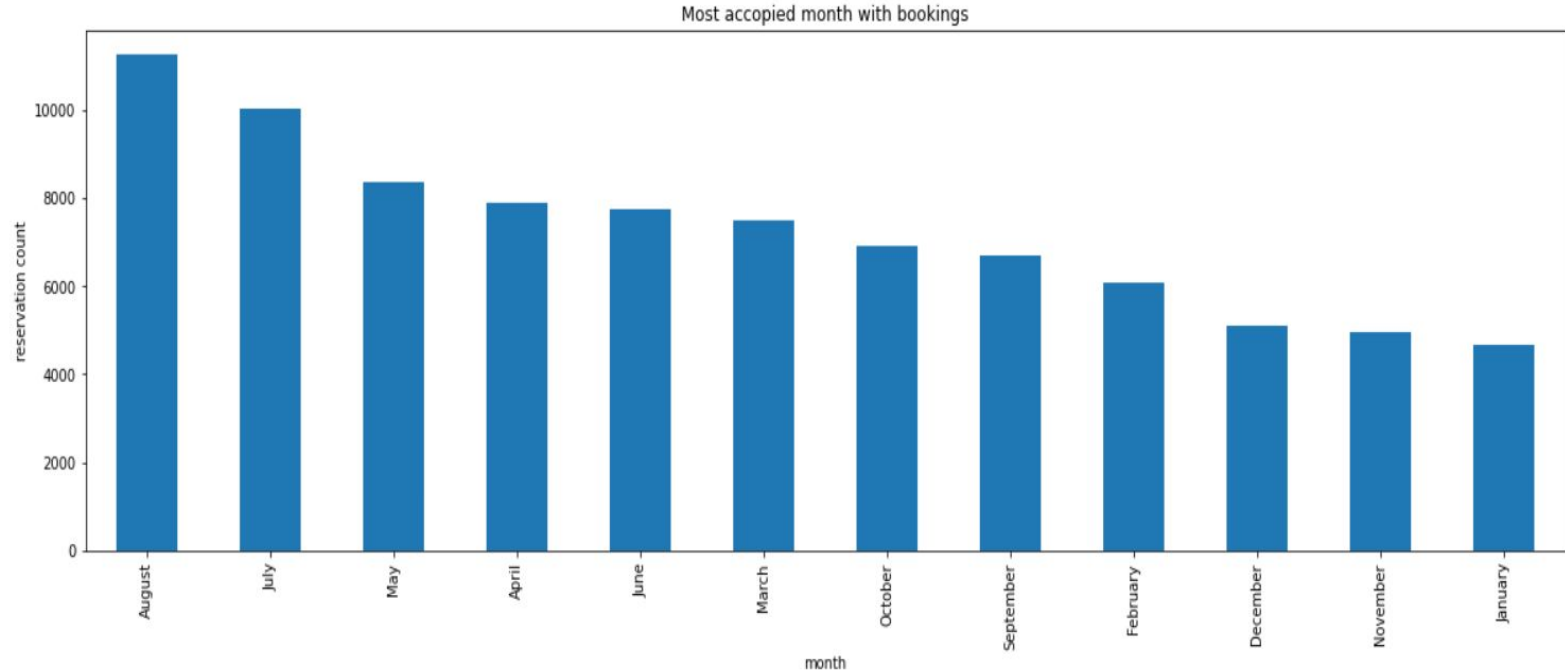
- From the above graph it has been observed that most visited people were couples

3. Room Category which are preferred by most people ?



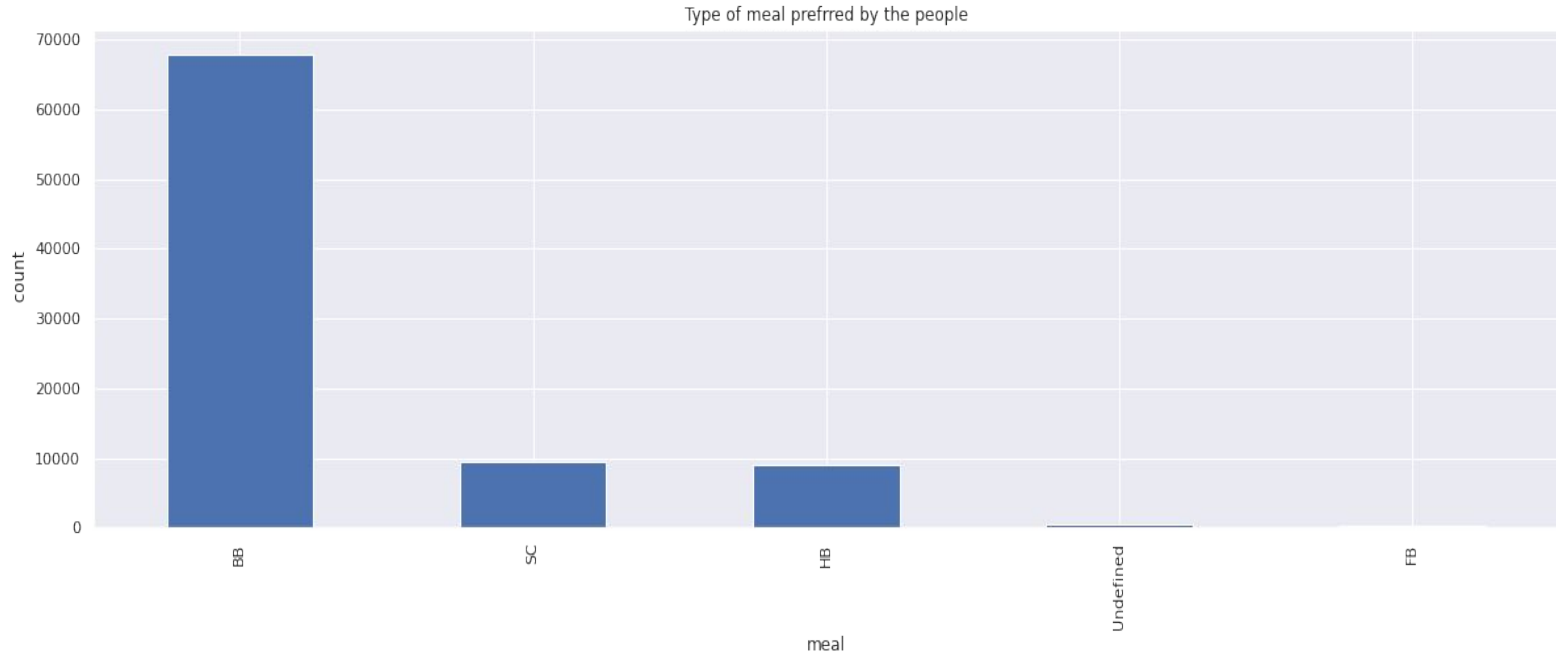
- From the above Graph it has been observed the most of the people have been preferred to Choose A_Type rooms followed by D_Type Rooms.

4.What is the Best time of a year to Book a Hotel Room?



- From the above graph it can be concluded that August and July were the most bussiest months in a Year.
- January and November months were the best time to book a hotel.

5. Type of meal preferred by customers ?

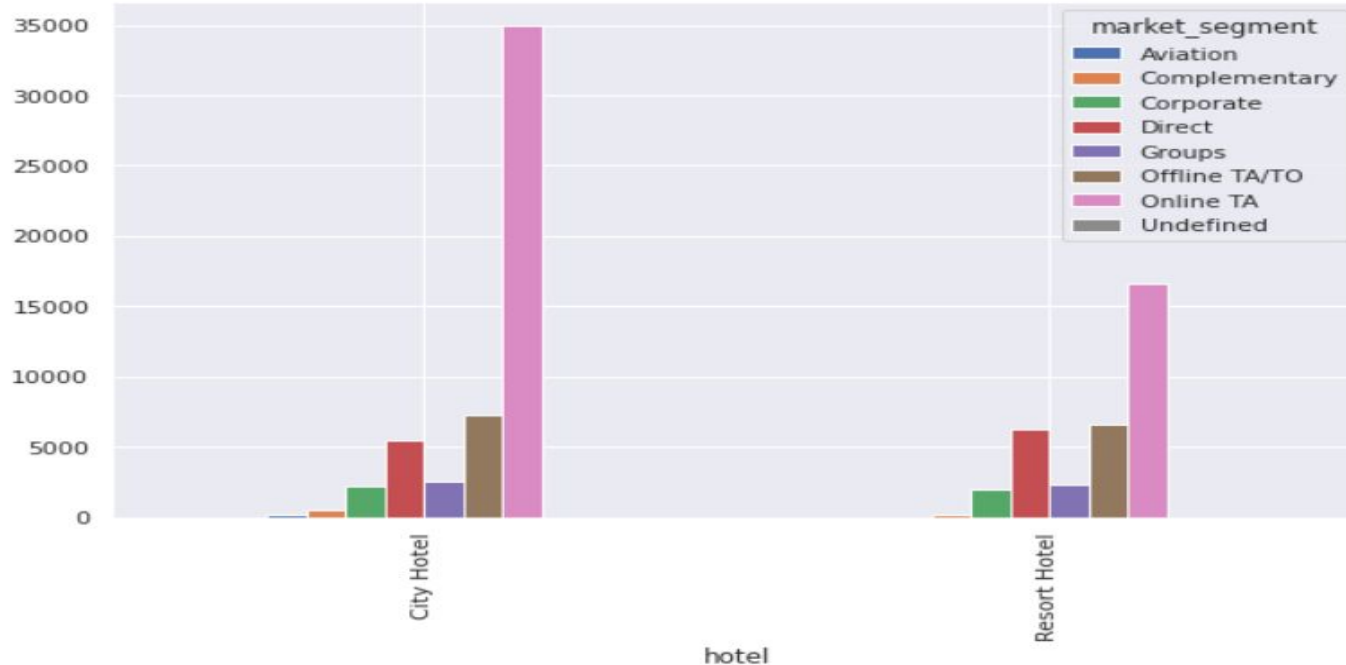


- So, it can be observed that most of the people were prefer to stay at Week end nights than Week day nights.

Booking Factors :

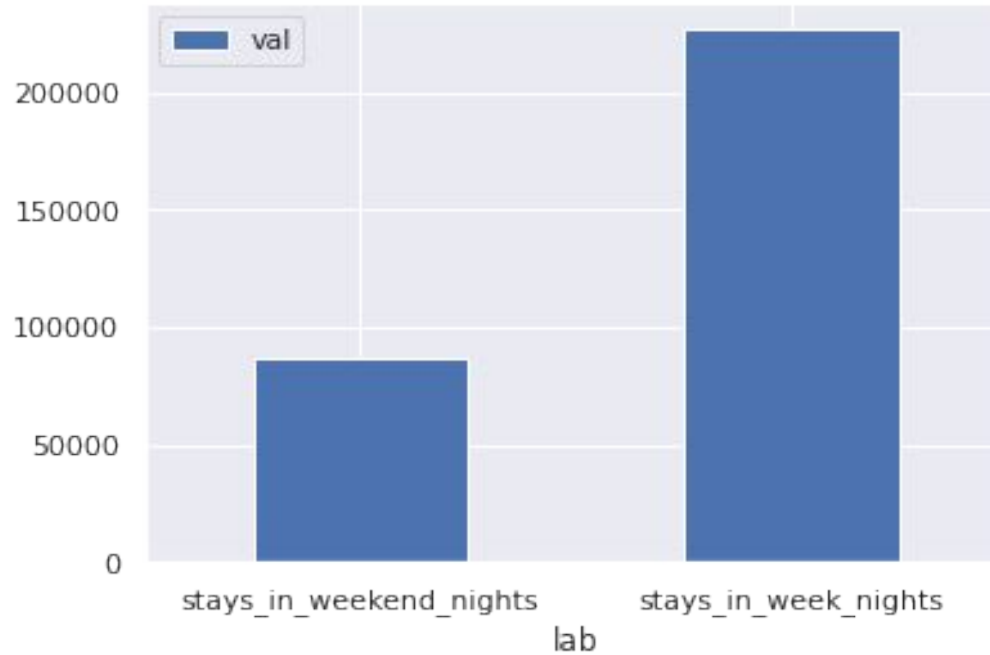
Here we are trying to analyze the data to get an idea of the booking trend.

Q6 - Which market segment makes highest number of booking?



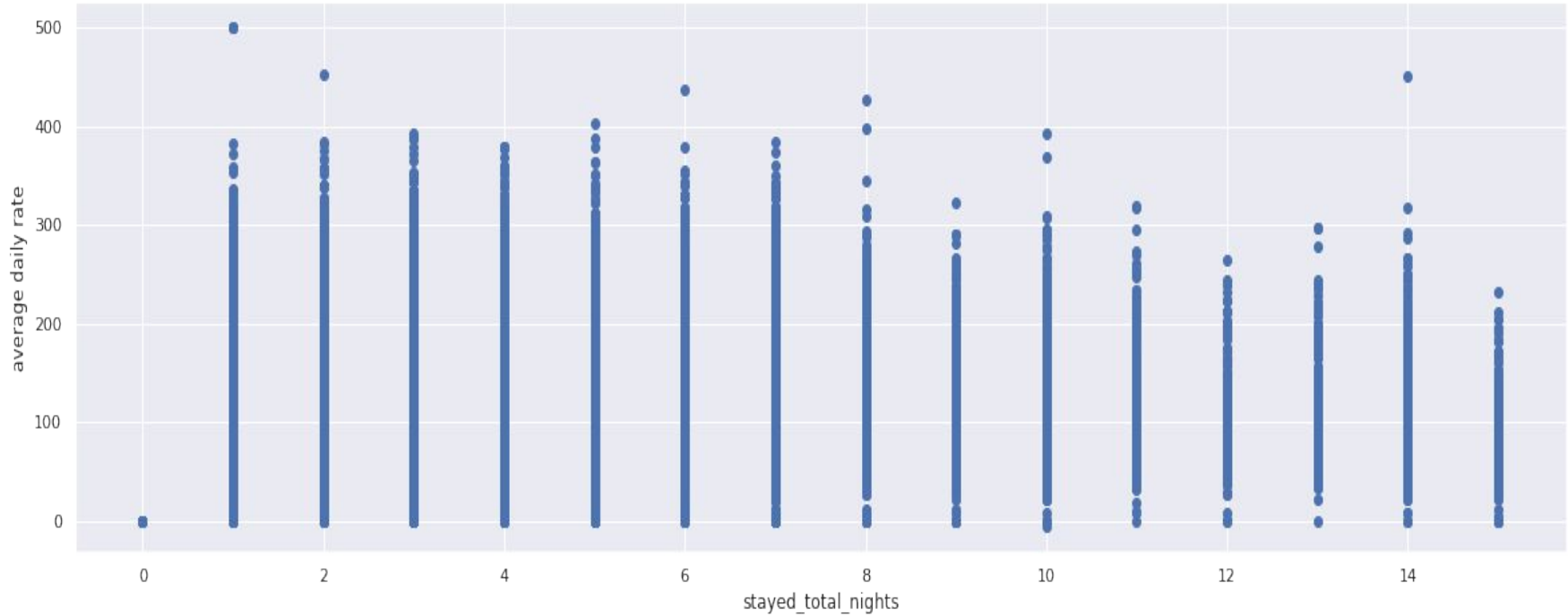
From above bar plot we can conclude that the online TA has highest number of bookings.

7.Comparision Between stay weekends and stay in week nights ?



- So, it can be observed that most of the people were prefer to stay at Week end nights than Week day nights.

8. Optimal length of stay in order to get best average daily rate?



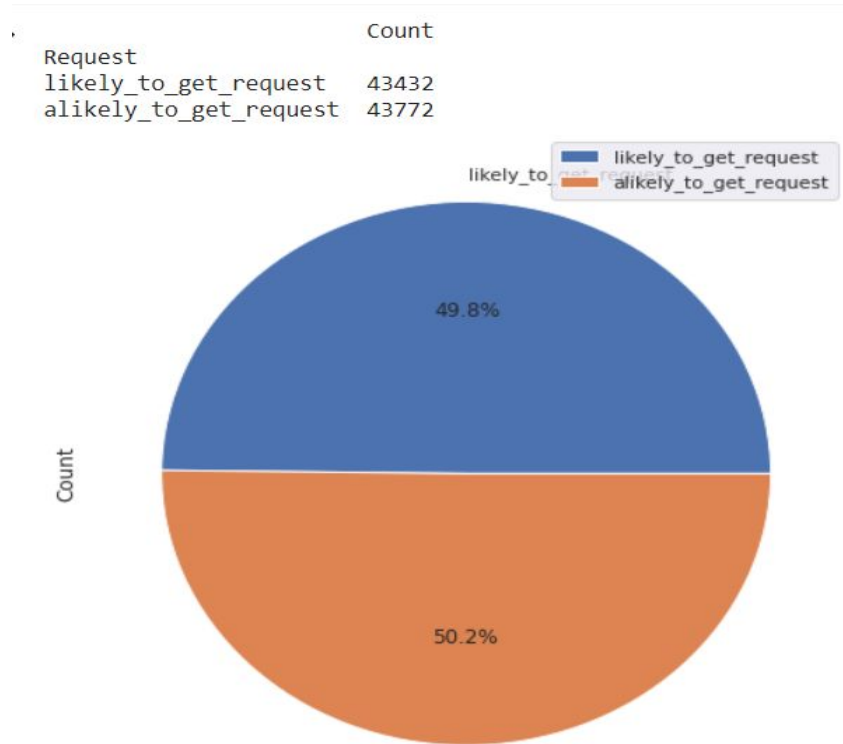
- Optimal length of stay in order to get best average daily rate is more than 10 days for both the hotels.

9.The country that has most number of guests?



- From above geographical map depicts the information about the nationality of the customers.
- From the above graph it has been observed that most of the people were coming from Portugal followed by France

10. To predict whether or not a hotel was likely to receive a disproportionately high number of special requests?

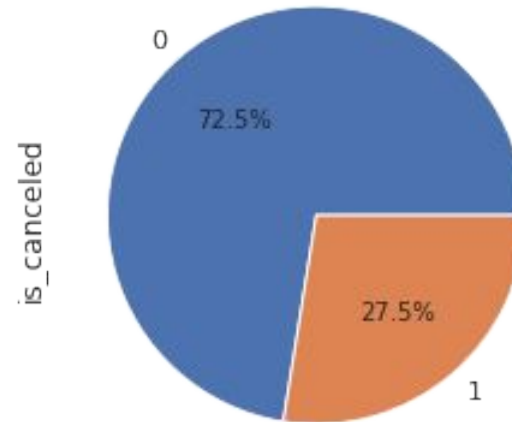


- The hotel was likely to receive a disproportionality high number off special request is 49.80%

EDA Performed to get an insight about the Cancellation:

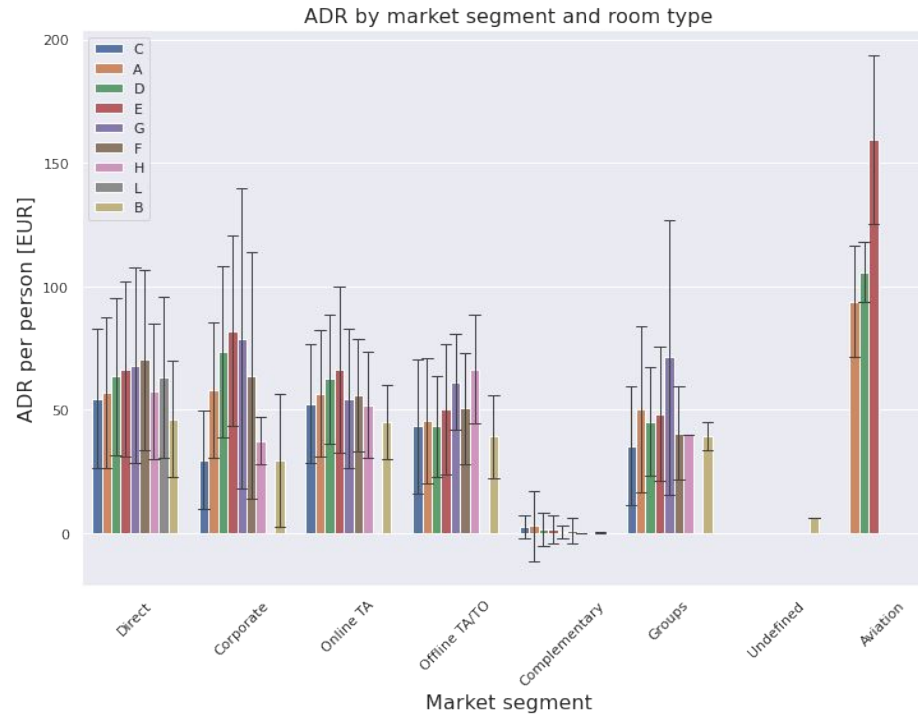
11. Percentage of Cancellation

proportion of canceled and Non canceled Bookings



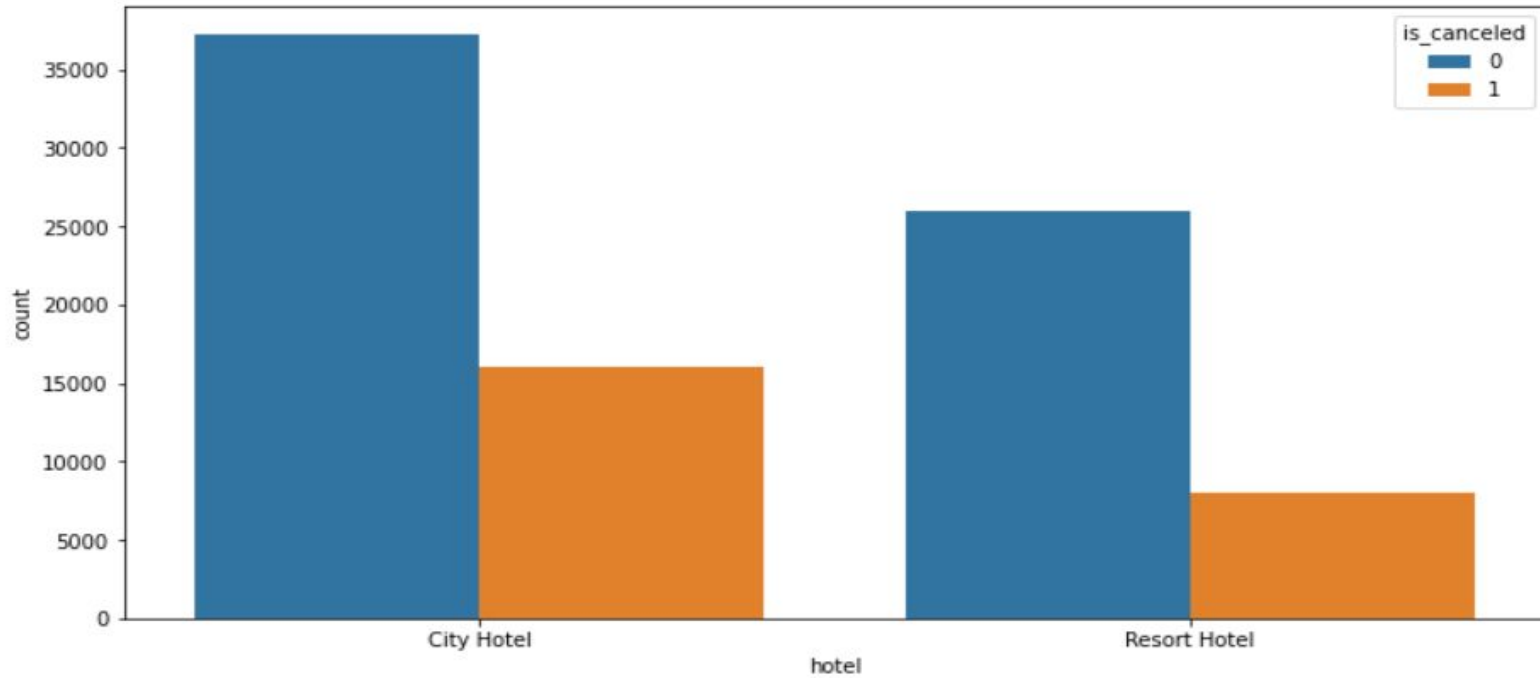
- we have noticed that around 27.5% of Bookings were canceled

12.Change in ADR with respect to market segment and room type ?

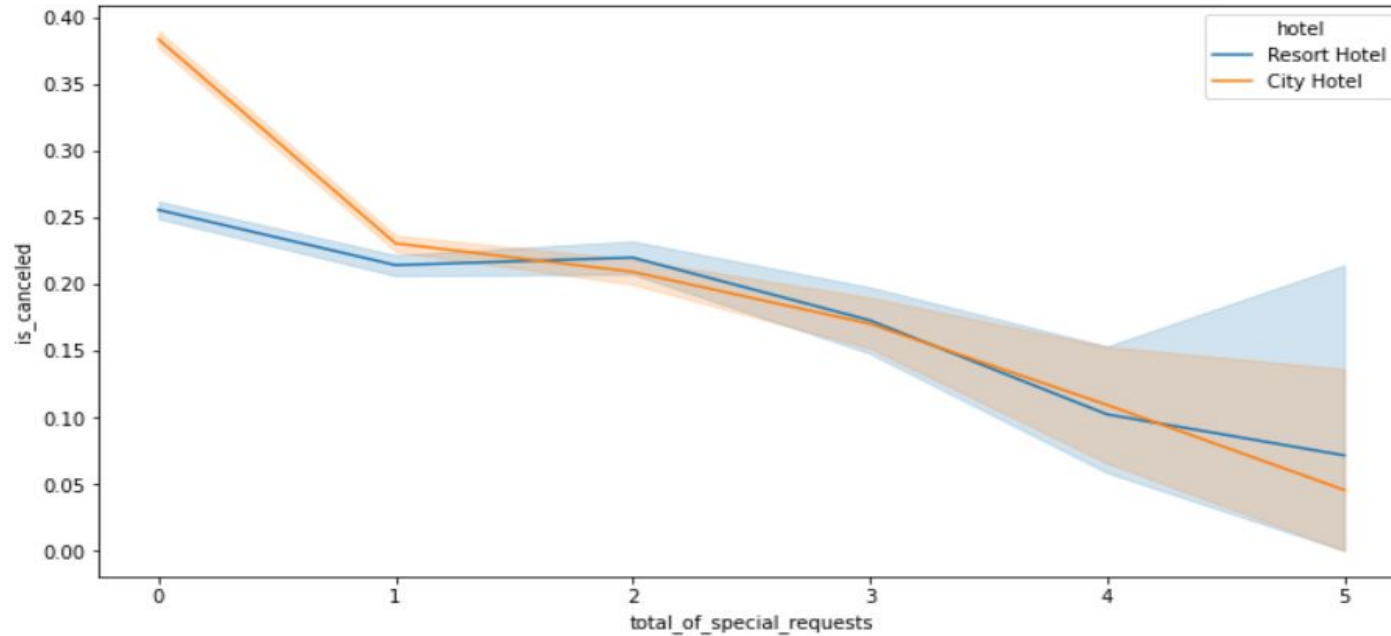


- On average, groups get the best prices and Airlines pay approximately twice as much
- What are possible reasons for the high prices paid by Airlines?
 - Airlines NEED a place for their crews to stay.
 - Airline personnel usually gets one room per person - more total rooms required compared to families.

13. Cancellation on both type of Hotels.

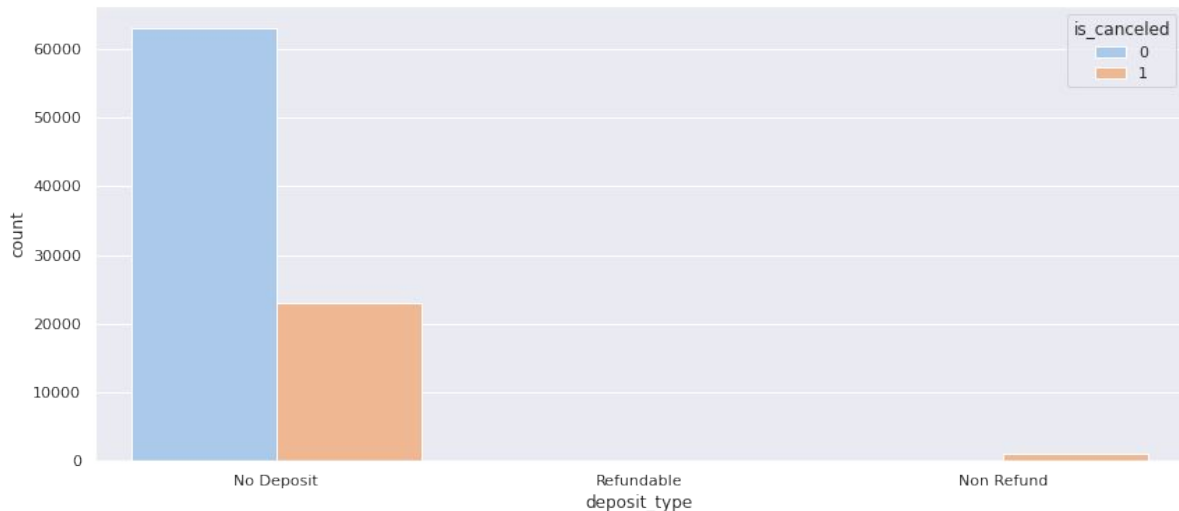


14. Does special requests decrease the cancellation rate?



- We notice that the cancellation rates are much lower if there are special requests.

15. Deposit type vs cancellation Rate

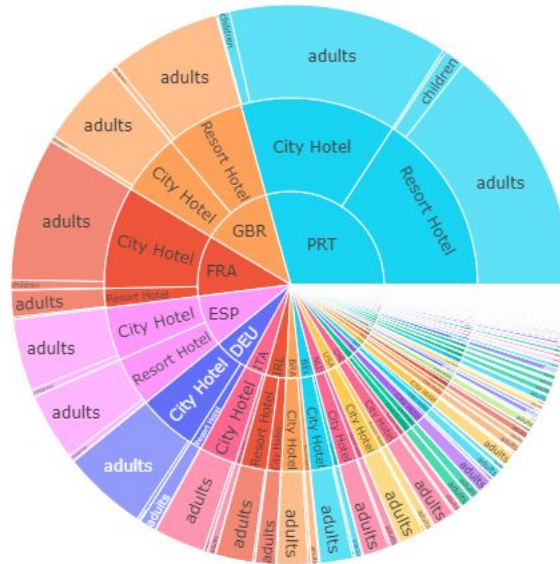


About 30k bookings of deposit type 'No Deposit' were cancelled. These numbers are huge if the hotels were not able to replace the cancelled bookings in time. It's a significant loss for the hotel. But in the next section, we will look at date of cancellation vs date of arrival to understand the impact of cancellation and how much time the hotel had to prepare for cancellations.

Also it is interesting to note that non-refundable deposits had more cancellation than refundable deposits. Logically one would have assumed that refundable deposits have more cancellation as hotel rates are usually higher for refundable deposit type rooms and customers pay more in anticipation of cancellation.

16.Total Hotel Booking Summary by Country, hotel and Accommodation type.

Country Wise Summary



Conclusion

After Performing the EDA, Below are some of the conclusions we came up with :

- Peak Seasons are August and July whereas Off Seasons are January and December, Hence January and December were the best months to book a Hotel.
- Most of the people were coming from Portugal followed by France
- The Hotel was likely to receive a disproportionately high number of special requests is 50%.
- Most of the people prefer City Hotel compared to Resort hotels, since Resort Hotels were costlier than City Hotels.
- The cancellation rates are much lower if there are special requests.
- Cancellation Rate in city hotel is very high when compared to the Resort Hotel.
- Most of the Bookings were from Online TA followed by Offline TA/To
- Most of the people were preferred stay at week end nights than Week day nights.
- Most of the people preferred to stay only one day followed by 7 days.
- The most visited type people were Couple followed by single person.

Thank You!