Capstone Project

Hotel Booking Analysis by EDA

OVERVIEW

- INTRODUCTION
- DATA SUMMARY
- DATA CLEANING
- DATA VISUALIZATION
- OBSERVATION
- CONCLUSION

INTRODUCTION

- This data collection includes hotel reservations for both resort hotel and city hotel.
- Includes details like the date the reservation was made, the number of adults, kids, and/or newborns staying, and the length of the stay.

DATA SUMMARY

DATA SET NAME

Hotel Booking Database including booking information for a city hotel and a resort hotel of various countries from 2015 – 2017.

SHAPE

- **The Data set contains the 119390 Rows**
- **And 32 columns**

COLUMN DETAILS

- Hotel - H1 city hotel, H2 Resort hotel
- is_canceled - 1:Canceled, 0:Not canceled
- Lead_time - Time between booking & arrival
- Arrival_date_year - year of arrival date
- Arrival_date_month - month of arrival date
- Arrival_date_week_number – Week number of year for arrival date
- Arrival_date_day_of_month – Day of arrival date

COLUMN DETAILS

- Stays_in_week_nights - stays in week nights
- adults – total number of adults in hotel
- Children – total number of children in hotel
- babies – total number of babies in hotel
- Meal – Bed & Breakfast (BB), Half Board (HB), Full Board (FB), Undefined contain no meal package
- Country – country of customers
- Market segment – A group of people who share one or more common characteristics used for business
- distribution_channel –Chain of business through which a

COLUMN DETAILS

- Previous_cancellations – previous cancellation by customer
- Previous_booking_not_canceled- confirmed booked by customer
- reserved_room_type – Pre-booked by customers
- assigned_room_type – single / couple / family
- booking_charges – booking charges we charged
- deposit_type – No Deposit, Non Refund, Refundable
- agent- ID of travel agency
- company- ID of the company
- days_in_waiting_list – Number of days in waiting

PROBLEM STATEMENT

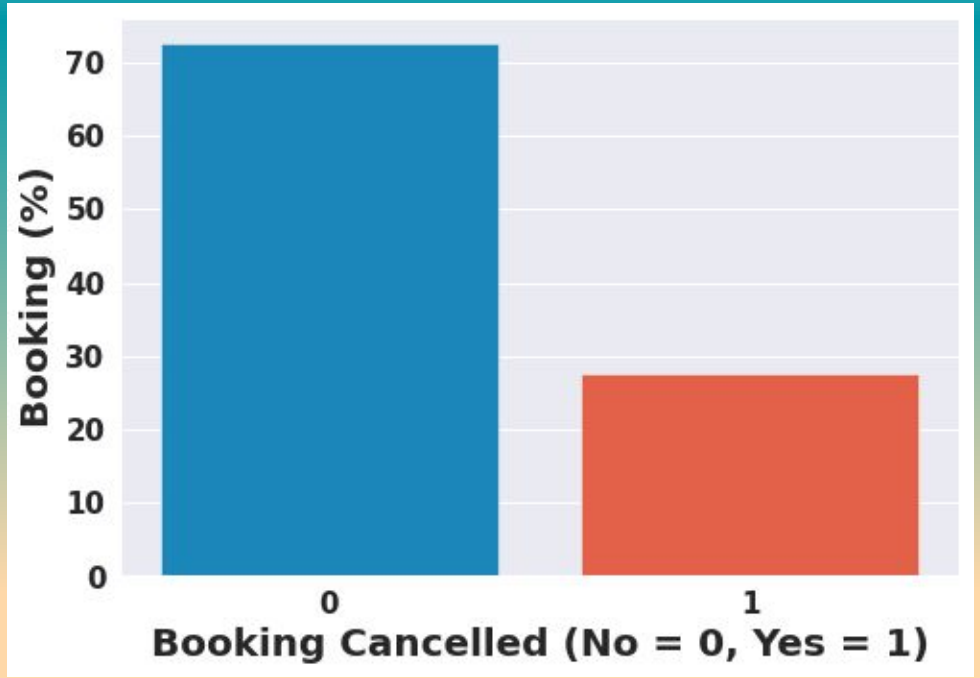
- Have you ever wondered when the best time of year to book a hotel room is? Or the optimal length of stay in order to get the best daily rate?
- What if you wanted to predict whether or not a hotel was likely to receive a disproportionately high number of special requests?
- This hotel booking dataset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

DATA VISUALISATION

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

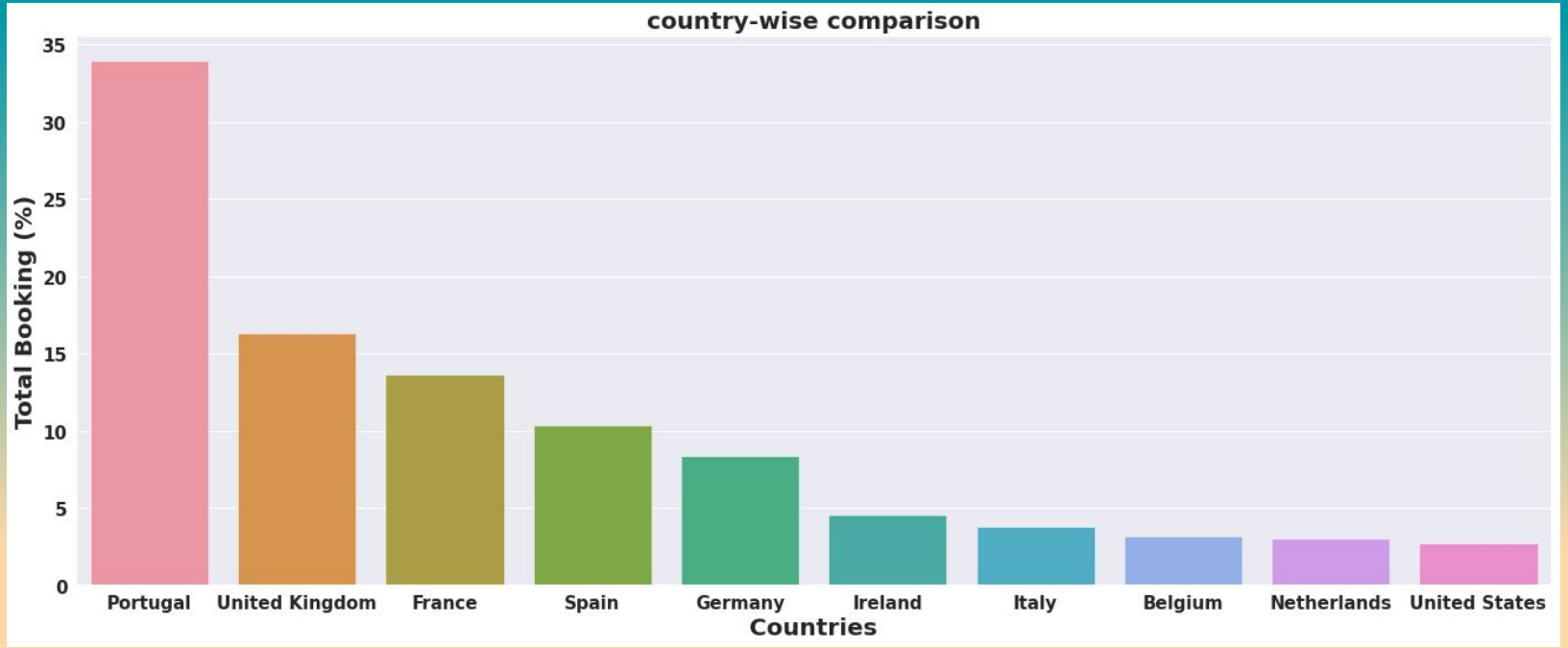
- Bar Plot
- Histogram
- Pie Chart
- Line plot
- Heatmap
- Box Plot

BOOKING CANCELLATION DETAILS



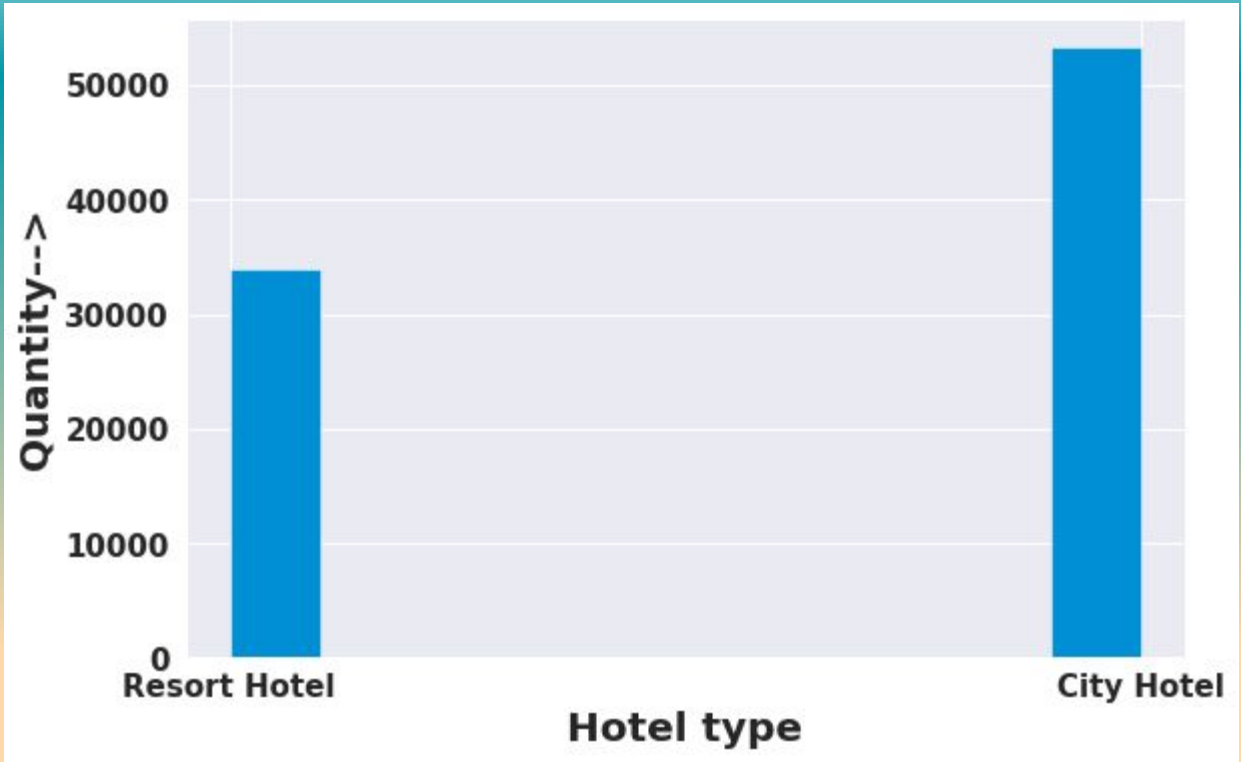
- More than 70% of people did not cancel bookings

TOP 10 COUNTRIES WITH MAX NUMBER OF PASSENGERS



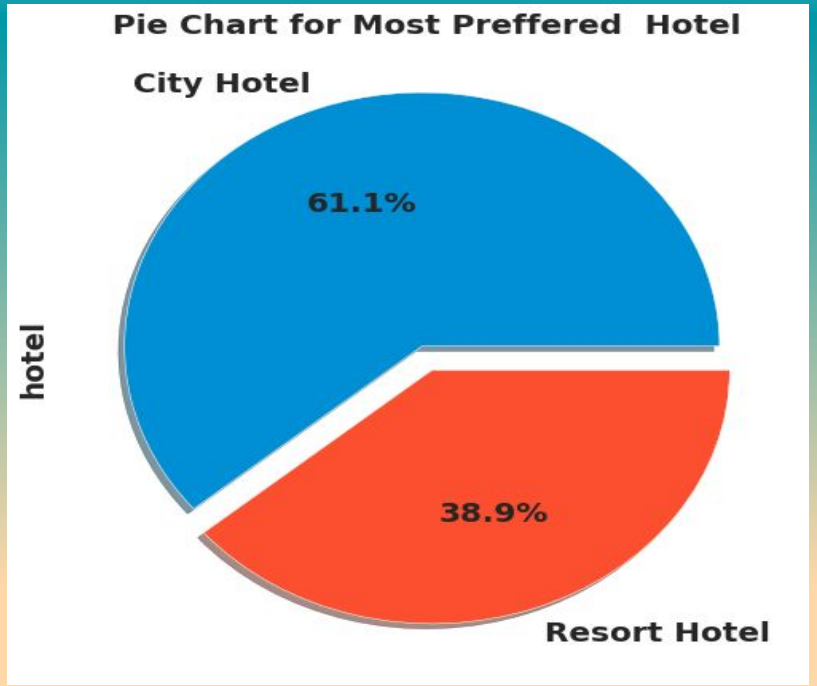
- **Most guests are from Portugal and other countries in Europe.**

TYPES OF HOTELS AND THEIR DETAILS



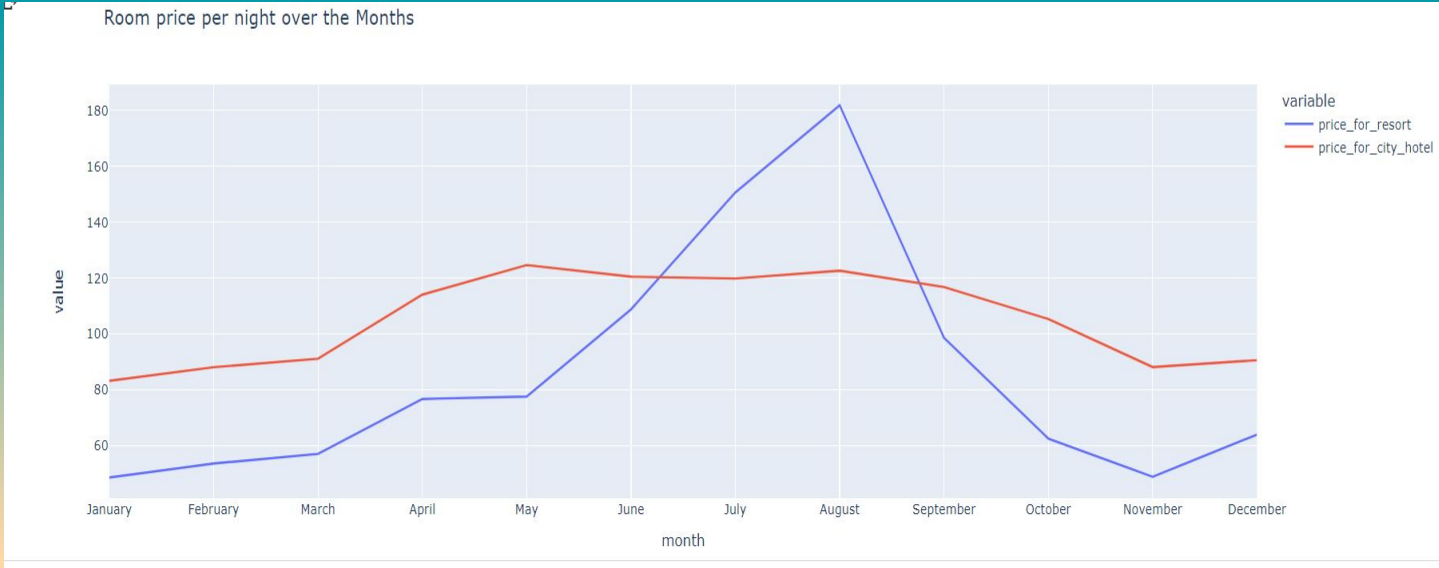
- **Most guests are from Portugal and other countries in Europe.**

PERCENTAGE OF GUESTS IN CITY AND RESORT HOTEL



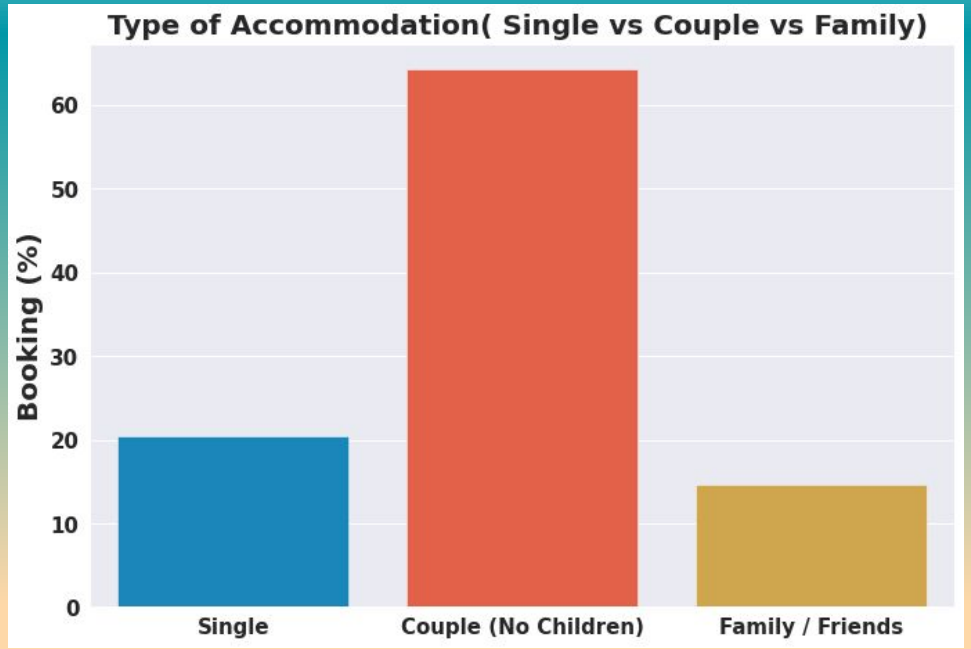
- **City Hotel contains maximum number of bookings of around 61% compared to Resort Hotel which is 38.9%**

COST OF ROOM PER NIGHT



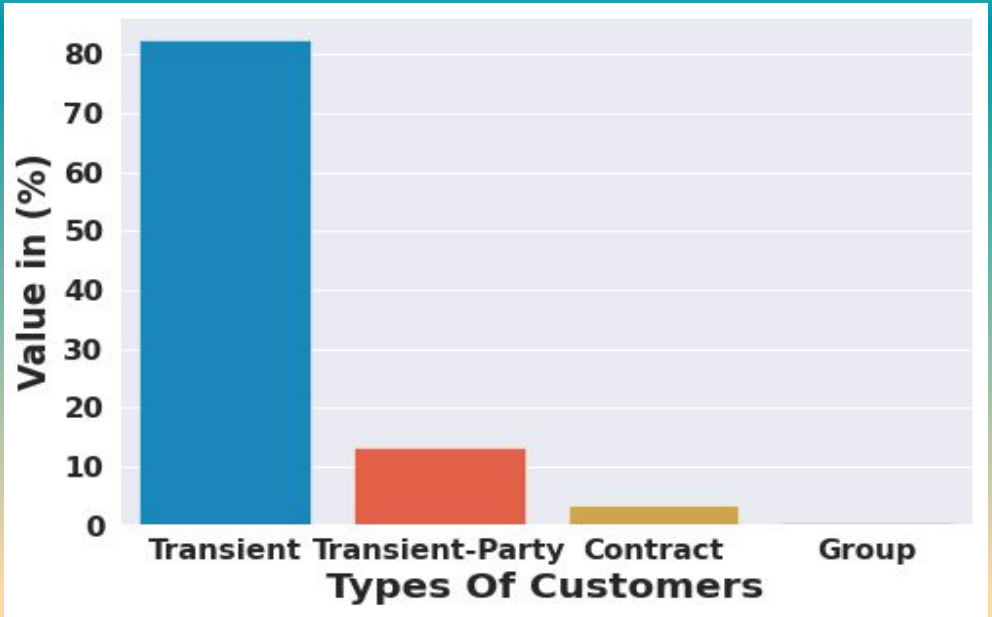
- **This plot clearly shows that prices in the Resort Hotel are much higher during the summer and prices of city hotel varies less and is most expensive during July and August**

MOST BOOKED ACCOMMODATION TYPE



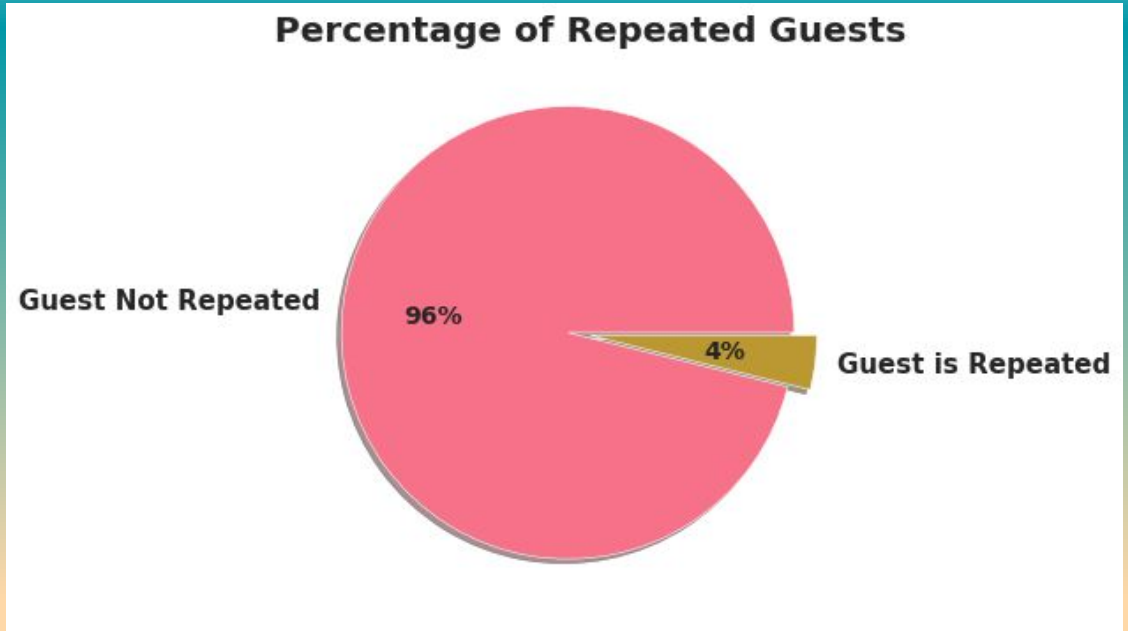
- **The most booked accommodation type is couple with no children**

TYPES OF CUSTOMERS



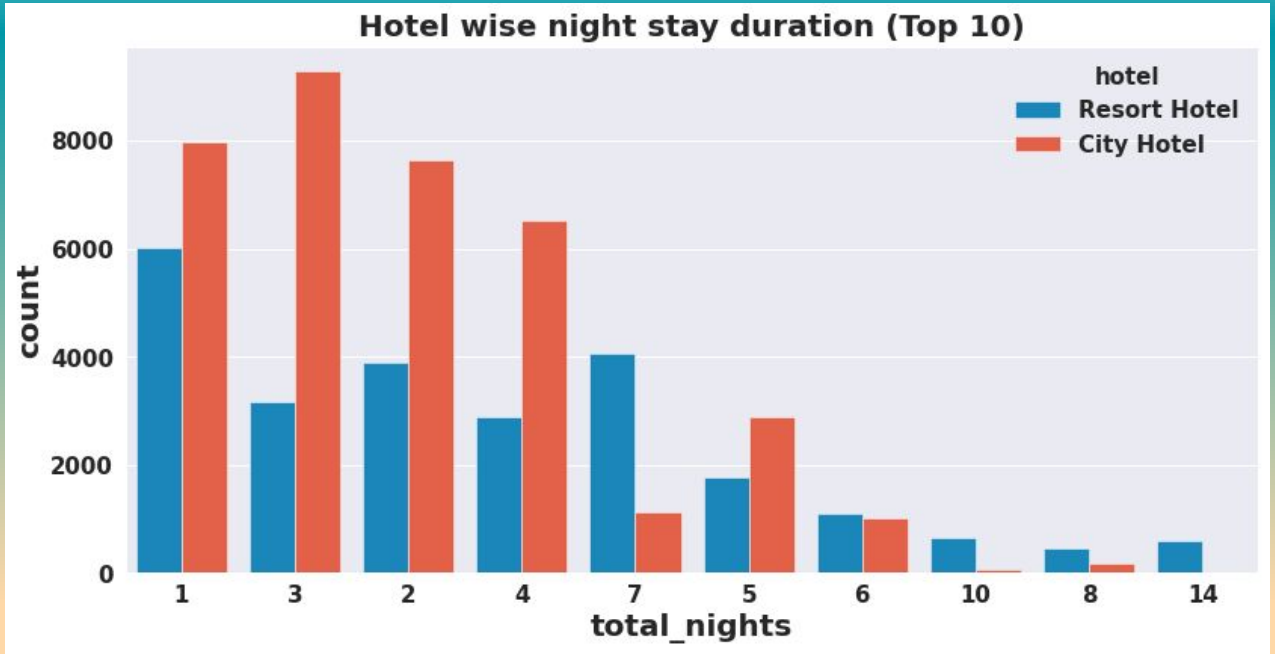
- **The percentage of different types of customers are Transient 82% , Transient-Party 13% , Contract 3% and Group 0%**

REPEATED GUESTS



- **The chart shows only 4% Guest is Repeated in Hotel. It means 3359 guest out of 83981 is repeated.**

DURATION OF STAY IN HOTEL



- **In resort hotel people like to stay 1 day and in city hotel people like to stay 2-3 days**

MOST BUSY MONTHS

- **The City hotel has more guests during spring and autumn and guest numbers for the Resort hotel go down slightly from June to September. Both hotels have the fewest guests during the winter.**

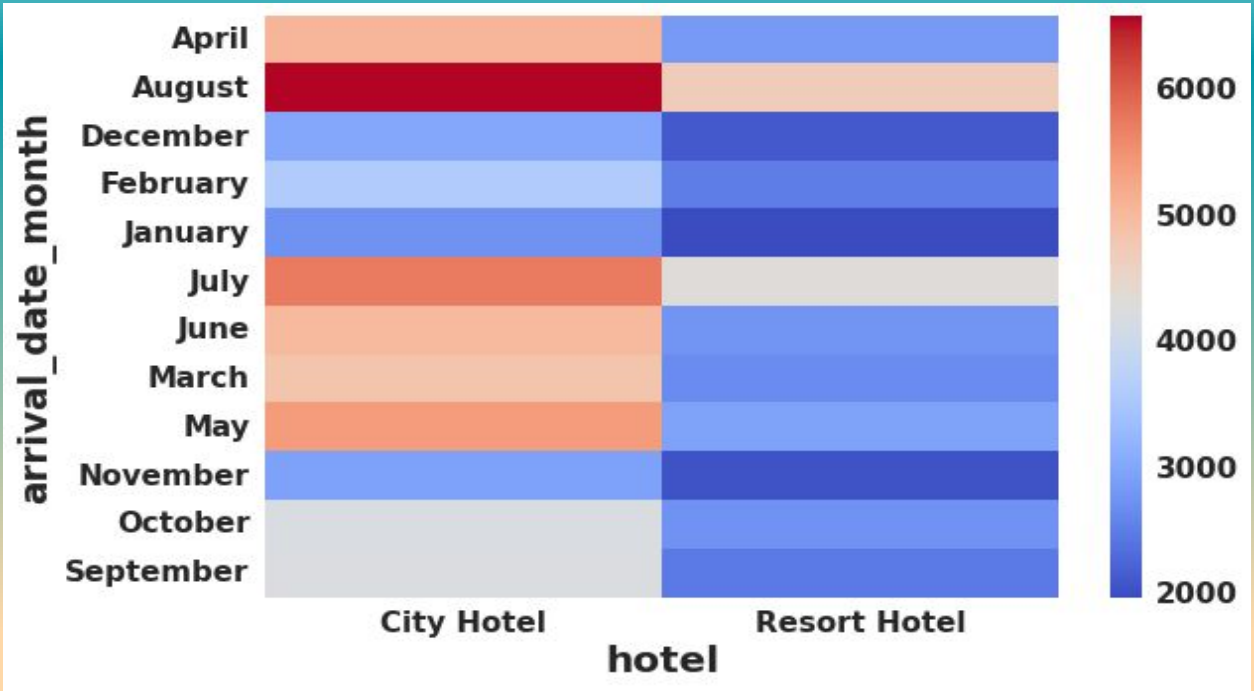
HOTEL WITH HIGHEST NUMBER OF CANCELLATION

- City hotel has more number of cancellation compared to Resort hotel

NUMBER OF BOOKINGS OVER TIME

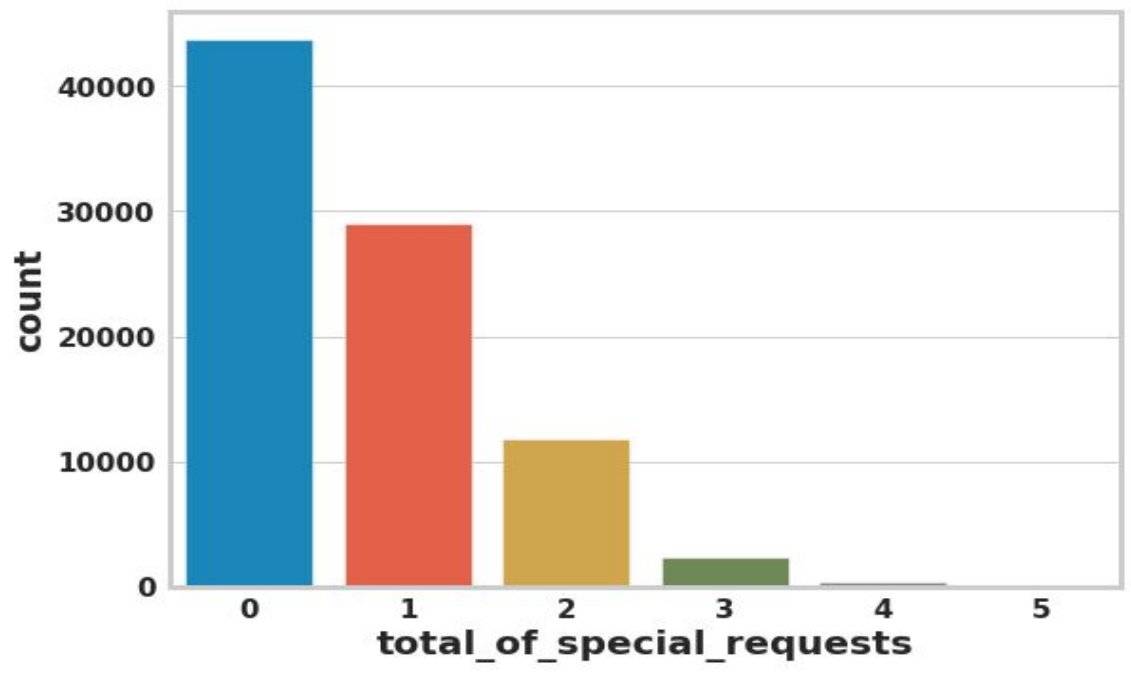
- **2016 has the highest number of bookings and there has been increase in the arrivals as years passes.**

NUMBER OF BOOKINGS OVER TIME



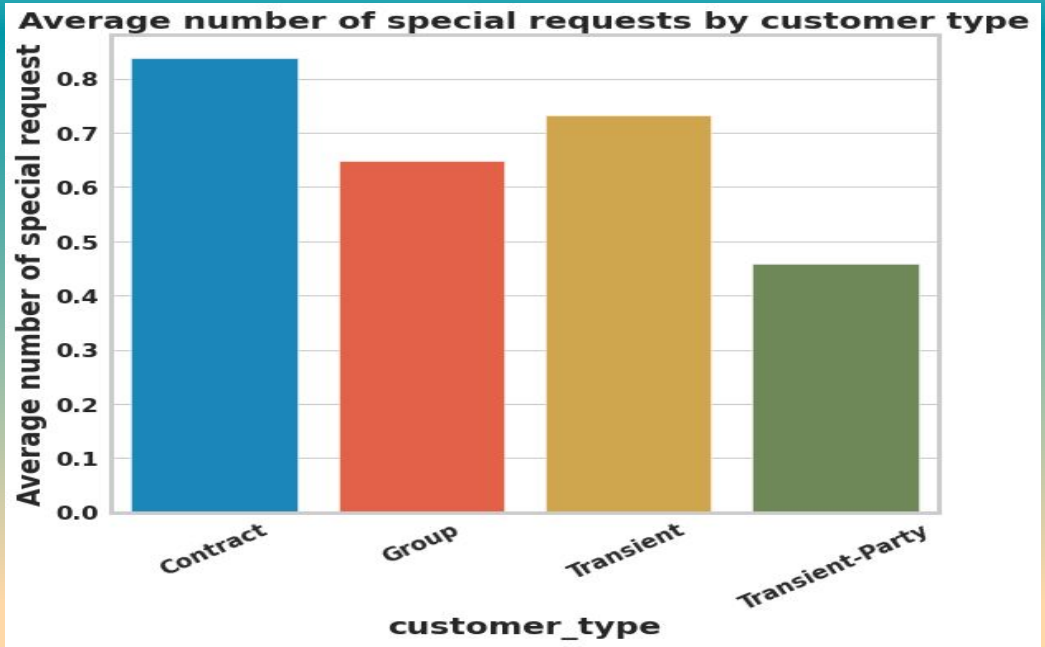
- **August is the most demanded month of the year. Interestingly, number of bookings decreased over year and June, the month before July, is the least demanded one. So, June is the best month of the year to go on holidays**

SPECIAL REQUESTS



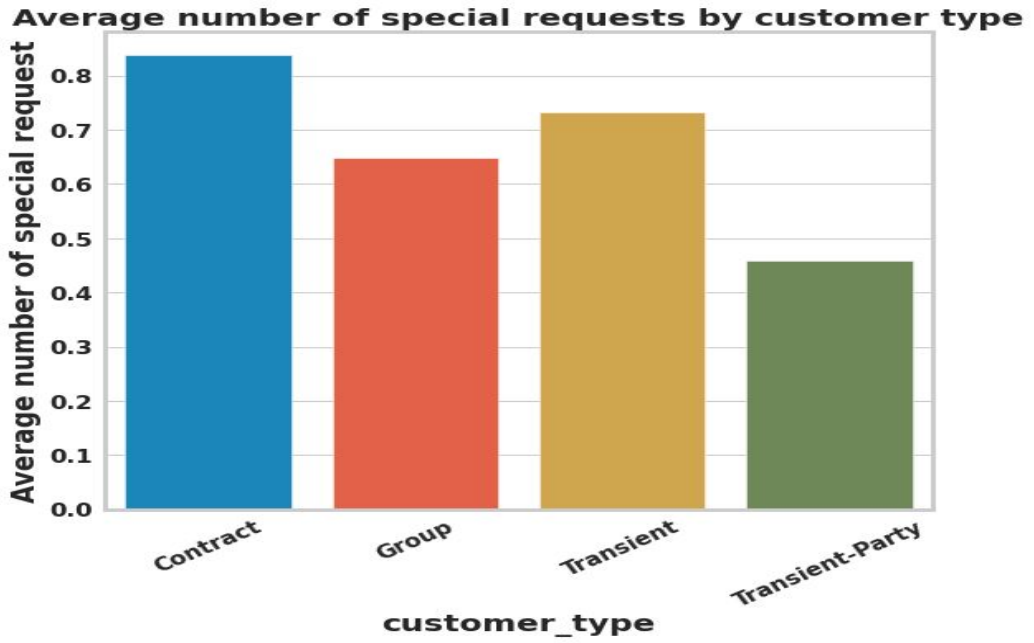
- **They are not entertaining much special requests.**

SPECIAL REQUESTS



- **Contract people have more special requests**

distribution channel is mostly prefered



- Contract people have more special requests