

Capstone Project

Hotel Booking Analysis

Capstone Project Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

NO	NAME	EMAIL ID	CONTRIBUTION
1	SONKAR ABHISHEK	sonkarabhi1406@gmail.com	PREPARED THE PYTHON PROJECT
2	GOURAV PATIL	patilgourav84@gmail.com	HELPED WITH ANALYSIS
3	AKSHAY PAWAR	akshaypawar459@gmail.com	HELPED WITH ANALYSIS AND DETAILS
4	RITESH NEULKAR	riteshneulkar@gmail.com	HELPED WITH ANALYSIS AND PPT
5	KEVIN VARSANI	varsanikevin@gmail.com	HELPED WITH ANALYSIS AND TECHNICAL DOC.

Github Link:- <https://github.com/abhi14062000/hotel-booking-analysis.git>

To discuss the analysis of given hotel bookings data set from 2015-2017.

POINTS TO DISCUSS :

- Introduction To EDA
- Abstract
- Agenda
- Data Summary
- Cancelled Booking Analysis
- Hotelwise Analysis
- Agent Analysis
- Reservation Analysis
- Timewise Analysis
- Conclusion

What is EDA ?

Exploratory Data Analysis (EDA)



approach to analyze the data using visual techniques. It is used to discover trends, patterns, or to check assumptions with the help of statistical summary and graphical representations.

ABSTRACT :

- A. 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?
- B. This hotel booking dataset can help you explore those questions!
- C. The analysis tries to answer three questions:
 - ❖ How strong is the seasonality in these hotels?
 - ❖ At what time in week the hotels are more busy ?
 - ❖ Can we predict a cancellation, just with the information available at the moment this reservation has been made?

AIMS TO THE EDA

```
graph TD; A[AIMS TO THE EDA] --> B[Maximize Insight Into A Dataset]; A --> C[Uncover Underlying Structure]; A --> D[Extract Important Variables]; A --> E[Detect Outliers And Anomalies]; A --> F[Test Underlying Assumptions]; A --> G[Develop Valid Models]; A --> H[Determine Optimal Factor Settings (Xs)]; B --> C; D --> E; F --> G;
```

Maximize
Insight Into
A Dataset

Uncover
Underlying
Structure

Extract
Important
Variables

Detect
Outliers And
Anomalies

Test
Underlying
Assumptions

Develop
Valid Models

Determine
Optimal
Factor
Settings (Xs)

AGENDA

To discuss the analysis of given hotel bookings data set from 2015-2017.

We'll be doing analysis of given data set in following ways :

1. Cancelled Booking Analysis
2. Hotelwise Analysis
3. Agent Analysis
4. Reservation Analysis
5. Timewise Analysis
6. Week Stay Analysis

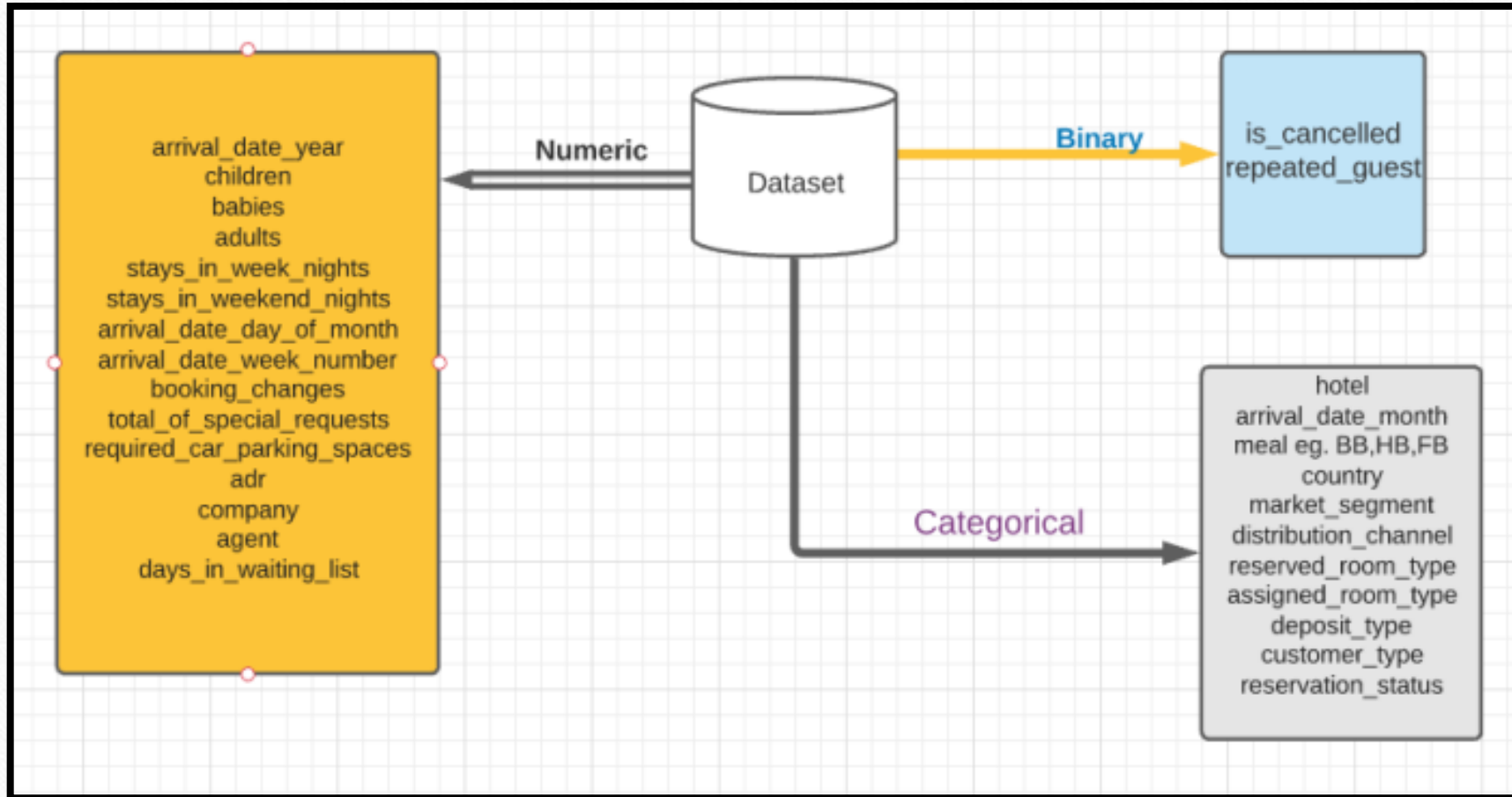
Data Summary :

- Given data set has different columns of variables crucial for hotel bookings. Some of them are:
- hotel: The category of hotels, which are two resort hotel and city hotel.
- is_cancelled : The value of column show the cancellation type. If the booking was cancelled or not. Values[0,1], where 0 indicates not cancelled.
- 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.[BB,FB,HB,SC,Undefined]
- Country: The origin country of guest.

Data Summary (contd..)

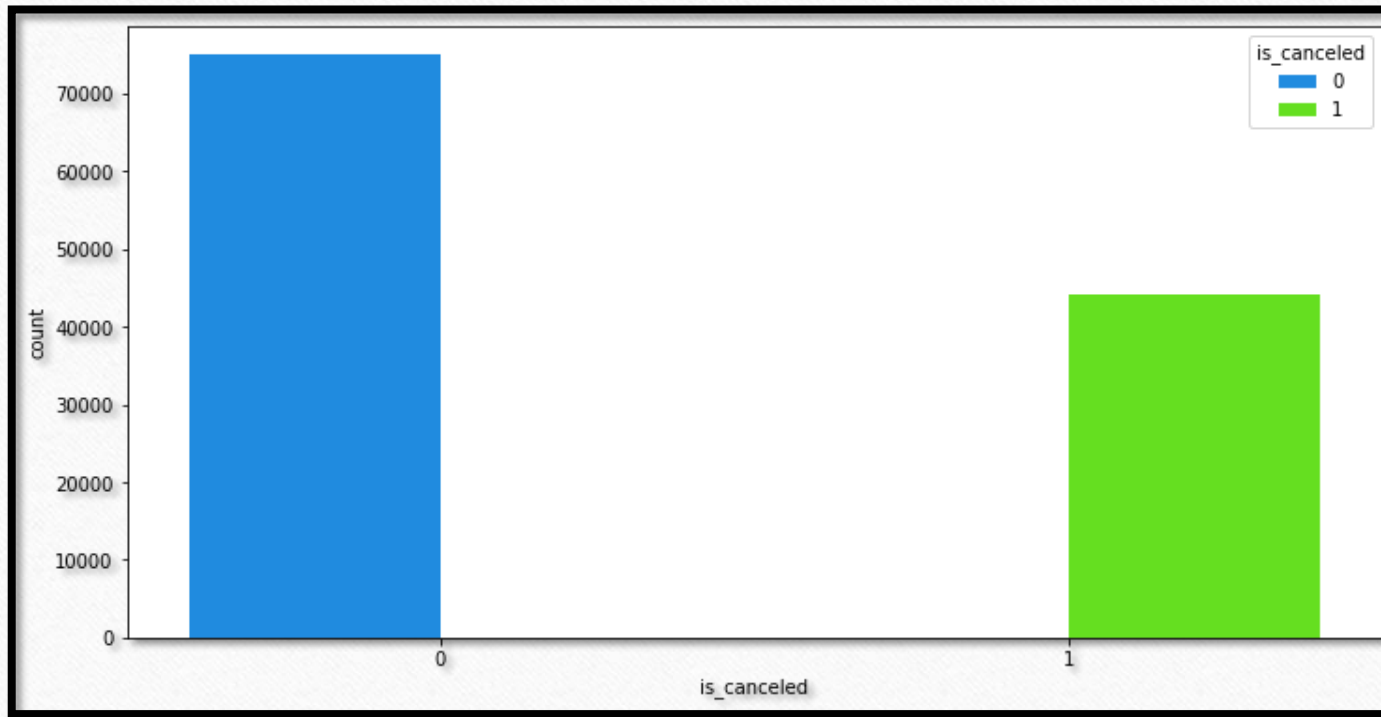
- market_segment: This column show how reservation was made and what is the purpose of reservation. Eg, corporate means corporate trip, TA for travel agency.
- distribution_channel: The medium through booking was made.[Direct,Corporate,TA/TO,undefined,GDS.]
- Is_repeated_guest: Shows if the guest is who has arrived earlier or not.Values[0,1]-->0 indicates no and 1 indicated yes person is repeated guest.
- days_in_waiting_list: Number of days between actual booking and transact.
- customer_type: Type of customers(Transient, group, etc.)

Data Summary :



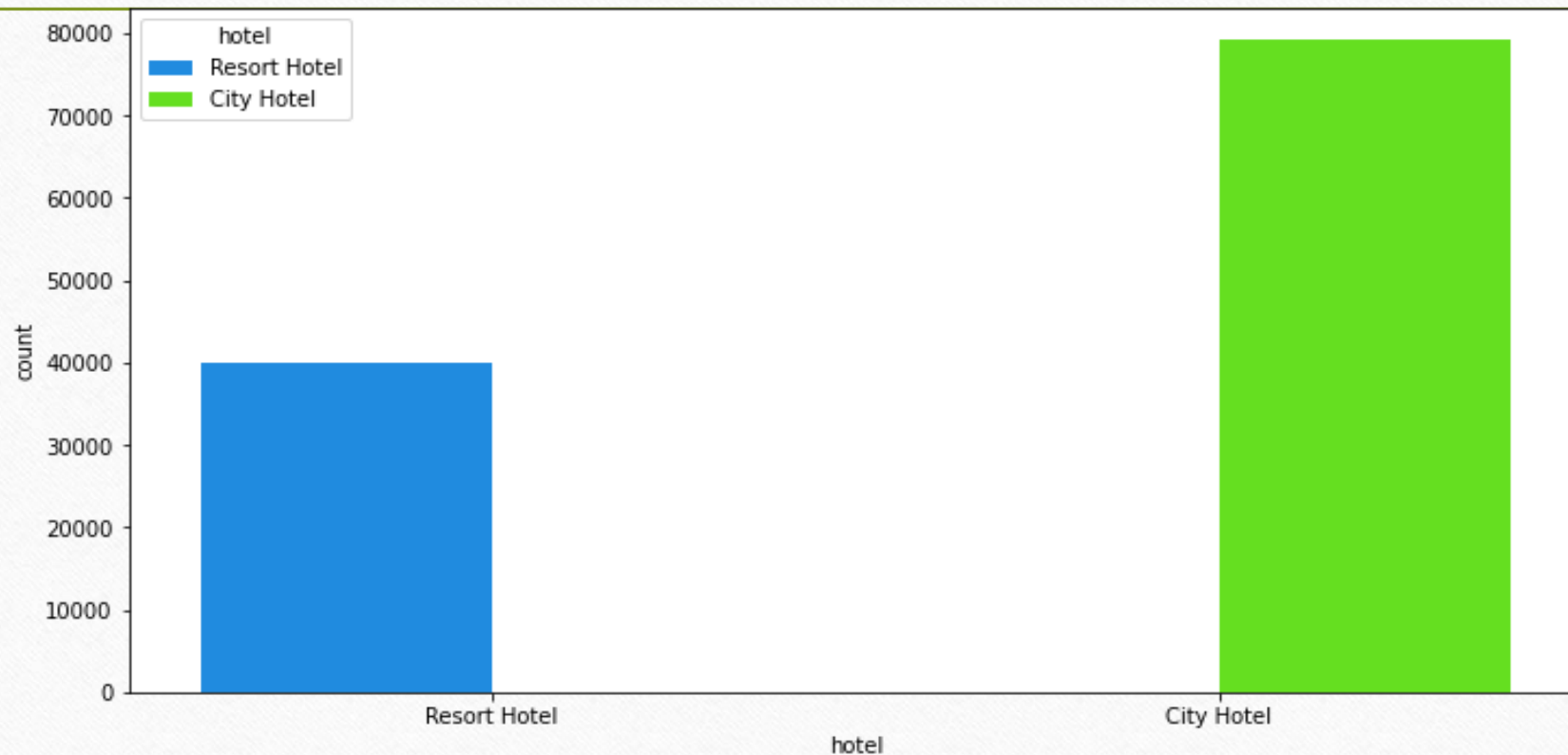
Cancelled Booking Analysis :

So here we can see that almost 37% times the bookings got cancelled



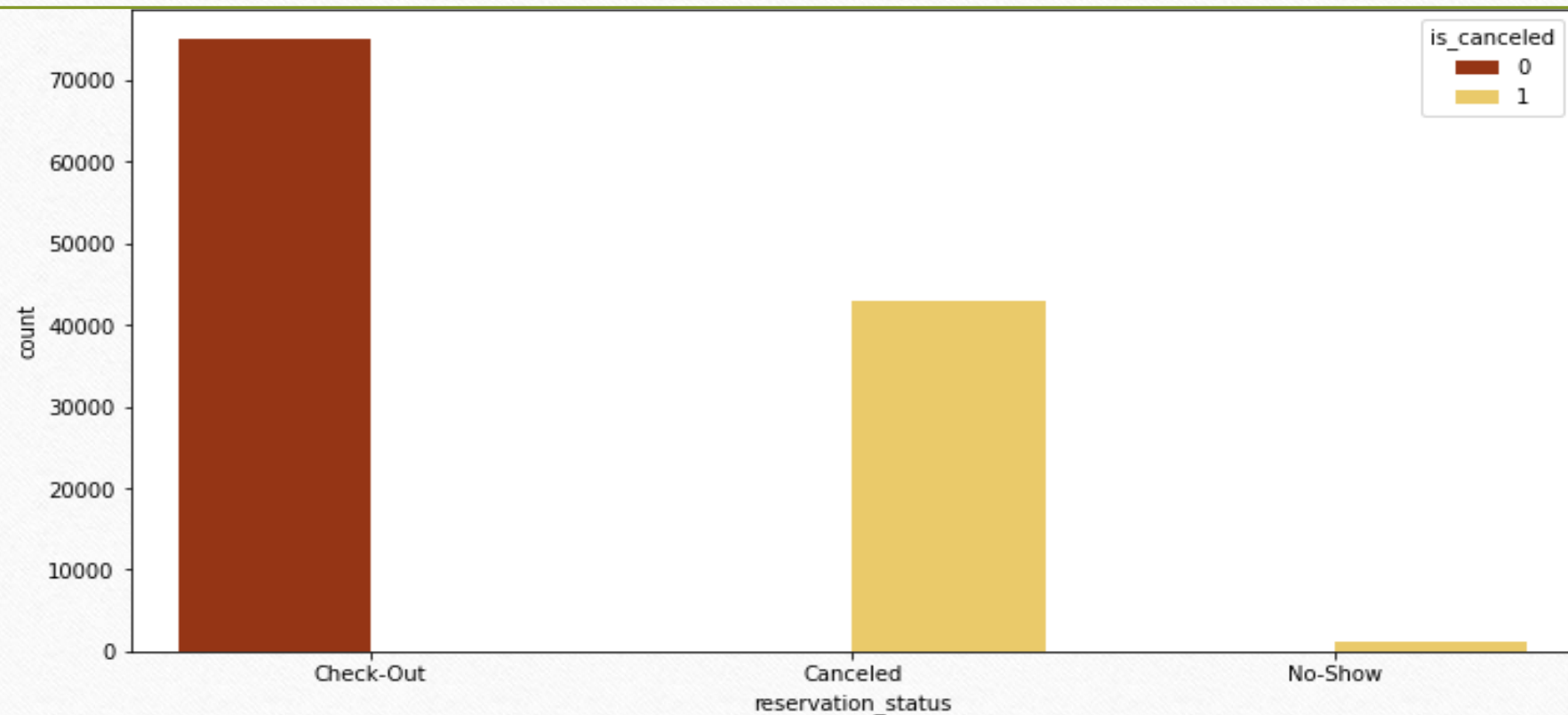
Hotelwise Analysis :

From the following graph we can see that more than 60% peoples prefer city hotels



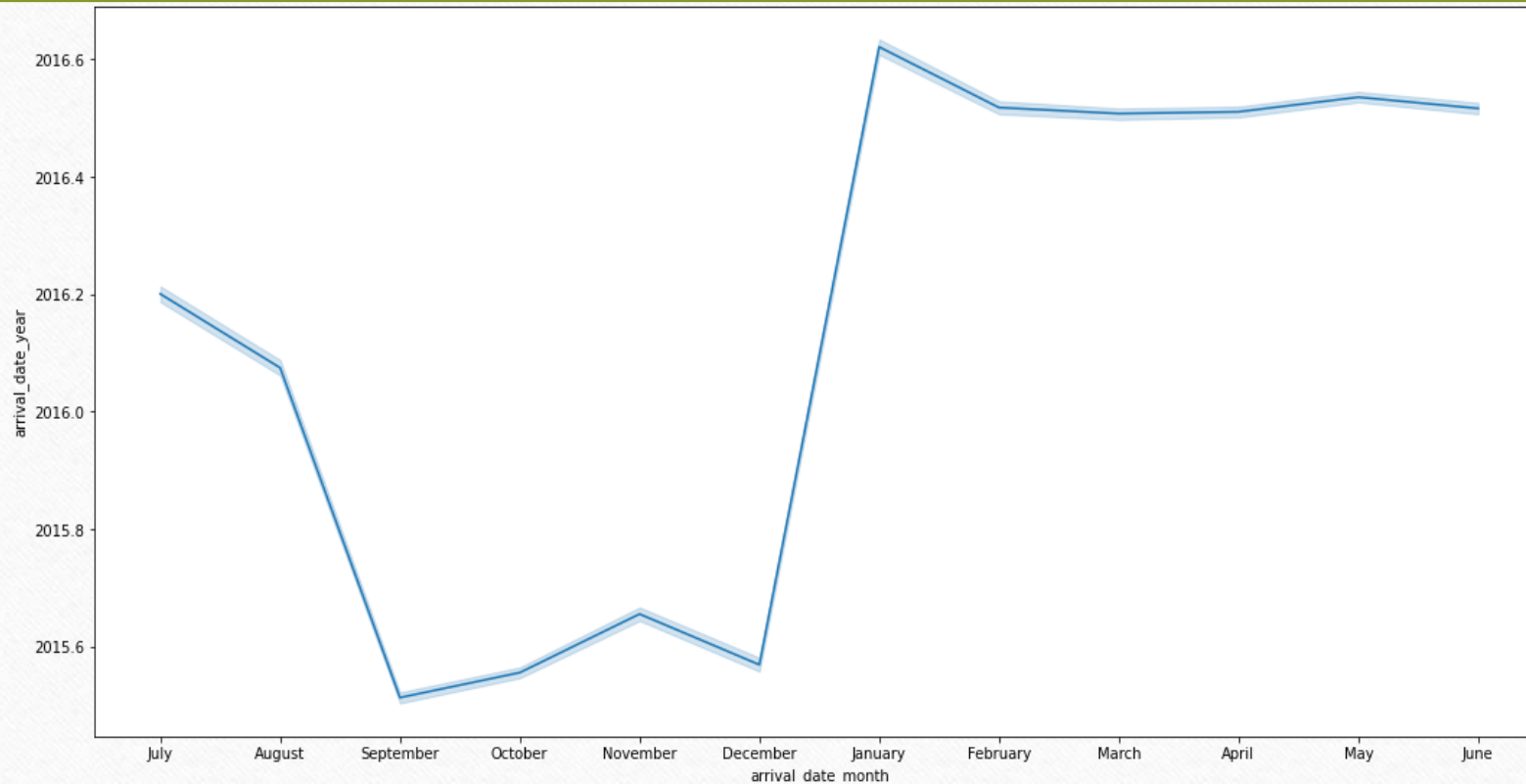
Reservation Analysis :

- In the below graph, we can see almost 40% is in cancel reservation



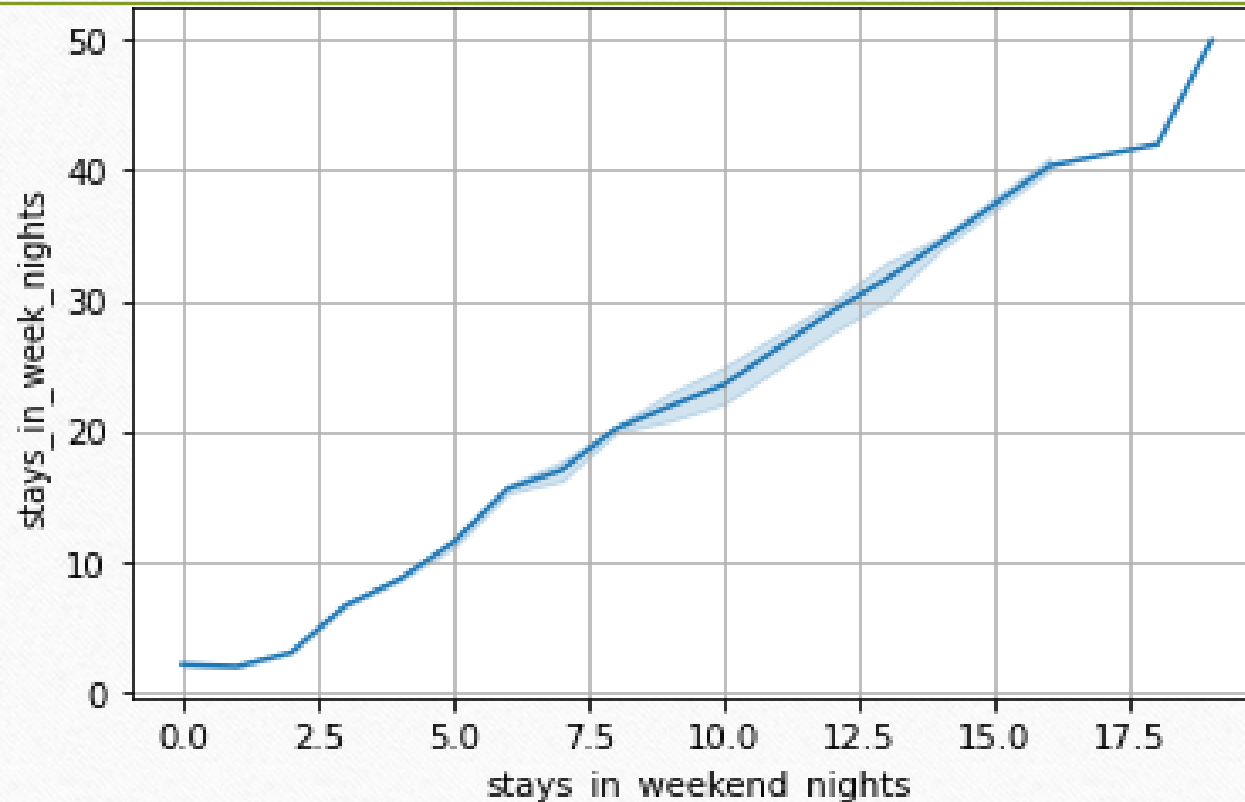
Timewise Analysis :

- From the month of July to August the number of bookings increased and in August, City Hotel got most number of guests.



Timewise Analysis(contd..)

- So here we can see that stays in week nights is more than weekend nights



Conclusion :

- Around 60% bookings are for City hotel and 40% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel.
- Both hotels have significantly higher booking cancellation rates
- July- August are the most busier and profitable months for both of hotels.
- Stays in week nights is more than weekend nights.
- Almost 40% is in cancel reservation.