

Capstone Project

Hotel Bookings Data Analysis

By-Neeraj Sachan
Cohort-Seattle



Lets Analyse the hotel data

1. Defining Problem Statement.
2. Creating Questions for Solving the problem statement.
3. Cleaning the Data and getting the basic details about our Data.
4. Doing Exploratory Data Analysis for getting the Answers our questions and making final observations and conclusions.



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 helps in exploring those questions!.

Steps:-

- 1. Data Processing:-** In this step we had observed the data and for cleaning the data, we removed the unnecessary features and treated the null values.
- 2.EDA-** In this part, we done some exploratory data analysis(EDA) on the features selected in step 1 to get the answers of our questions.
- 3. Final Observations-** After doing EDA and making Observations, we had made the some final Conclusion.

Data Description

The given data set contains booking information for a city hotel and a resort hotel, and includes information such as when the booking was made, length of stay, the number of adults, children, and/or babies, and the number of available parking spaces, among other things. All personally identifying information has been removed from the data.

The given hotel dataset contains 32 features and 119390 observations. Each observation represents the complete detail about the booking.

Questions:-

We created following questions for our analysis:

- Q1. Which type of customers do more bookings?
- Q2. Let's have an overview which type of hotel generally people prefer to book?
- Q3. Let's have an overview which type of deposit is more preferred by the customers?
- Q4. Which kind of food is mostly preferred by the guests?
- Q5. Let's have an overview how much guests pay for a room per night?
- Q6. What is the most preferred room type?
- Q7. Looking into which countries mostly visitors are coming from?
- Q8. What is the percentage of repeating guests?
- Q9. Which hotel has high chance that its customer will return from another stay?
- Q10. Which hotels generating more ADR?
- Q11. What is the relationship between total number of guests and ADR?

Questions (Cont.)

Q12. What is the relationship between total stay and ADR?

Q13. Which year had the highest Bookings?

Q14. Is which month most of the bookings happened?

Q15. How does the price vary over the year?

Q16. What if you wanted to predict .Whether or not a hotel was likely to receive a disproportionately high number of special requests?

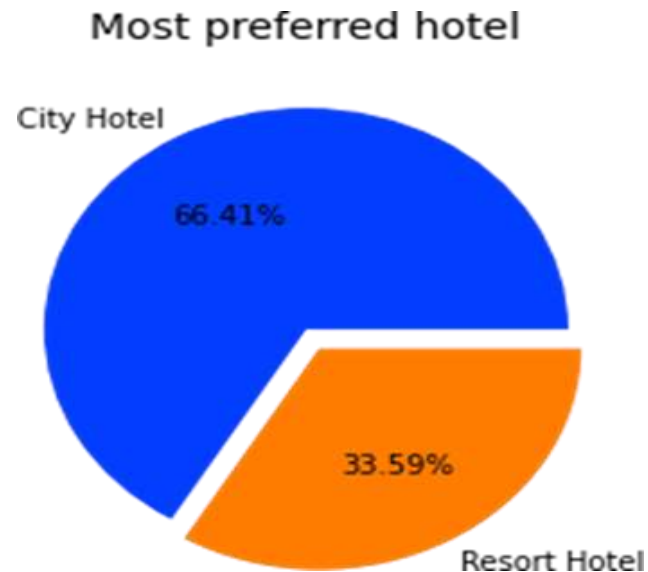
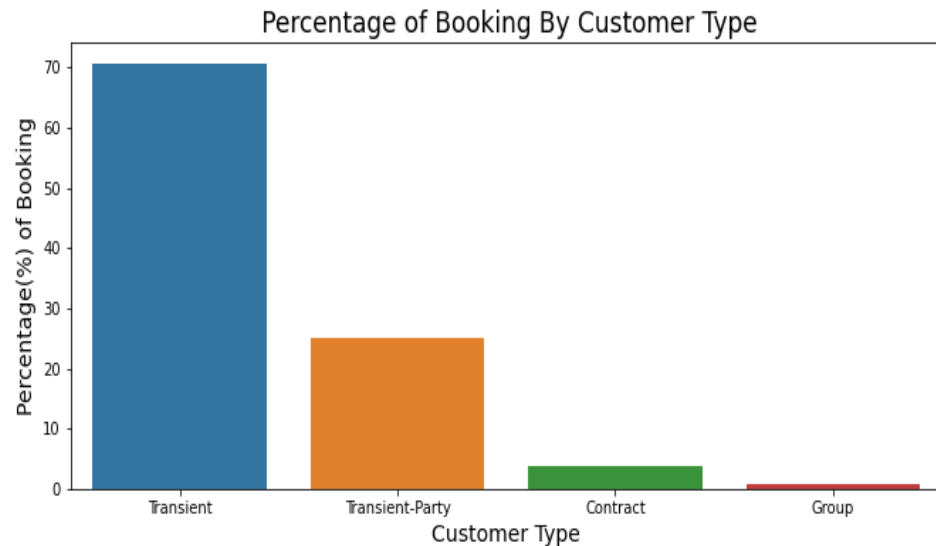
Q17. How long people stays in hotels?

Q18. Is customer canceled their bookings if they are not allotted with the same room type which was reserved by them?

Q19. Which distribution channel contributed more to generate high ADR?

Q20. Which distribution channel has the highest Cancellation rate?

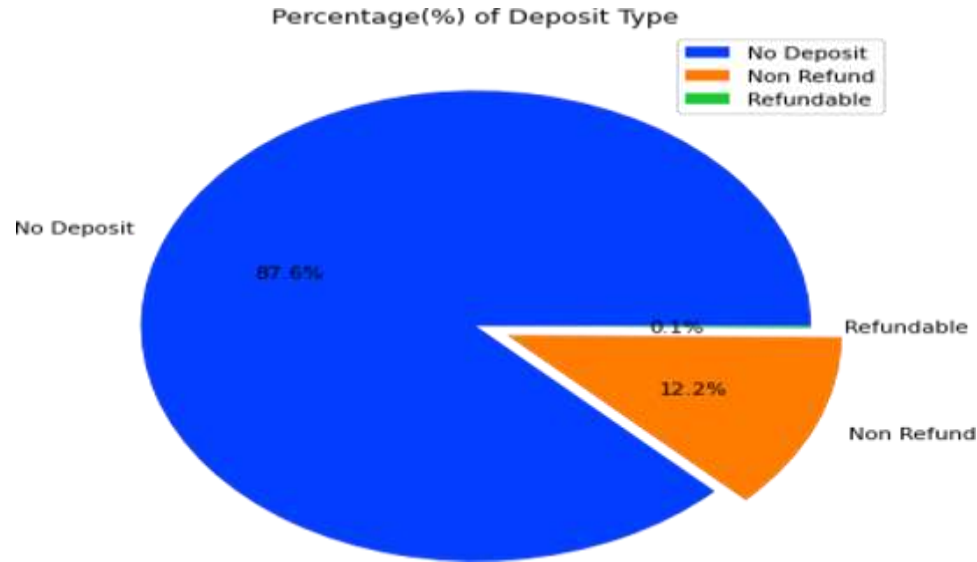
EDA & Observations



Observation:

- Bookings by Transient customer is more than other type of Customers (approx.. 70% of total bookings.).
- People prefer to book City Hotels more as compare to Resort Hotels.

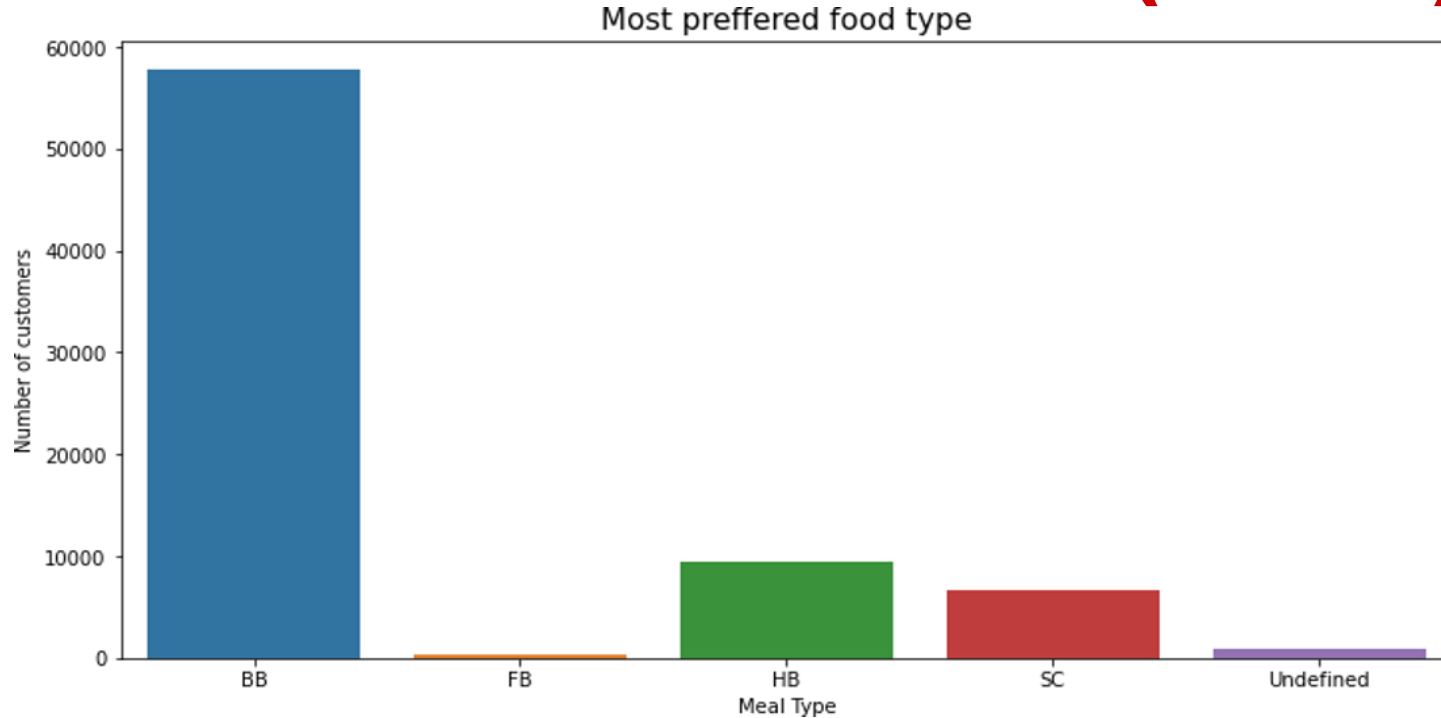
EDA & Observations (Cont.)



Observation:

- 87.6% of guests prefer “No Deposit ” type of deposit.

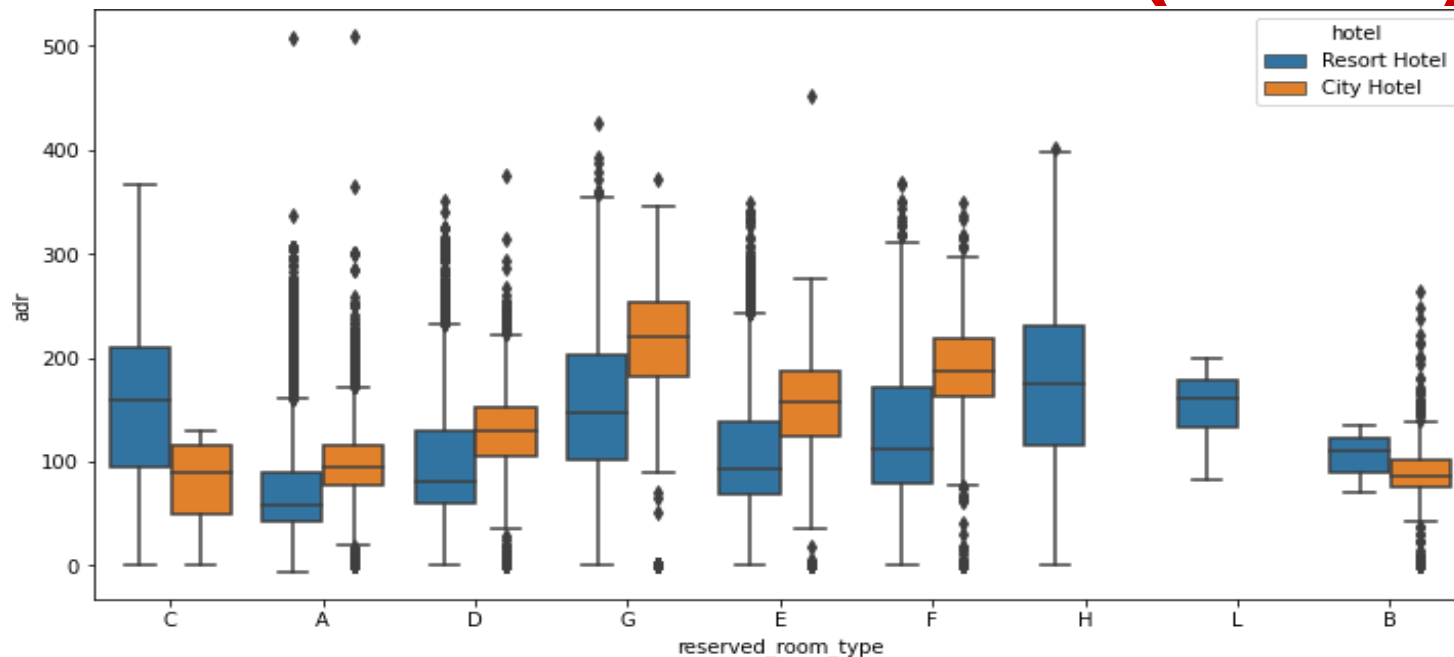
EDA & Observations (Cont.)



Observation:

- *Most preferred food by the customers is BB(Bed and Breakfast).*
- *HB(Half Board) and SC(Self catering) are approx. equally preferred.*

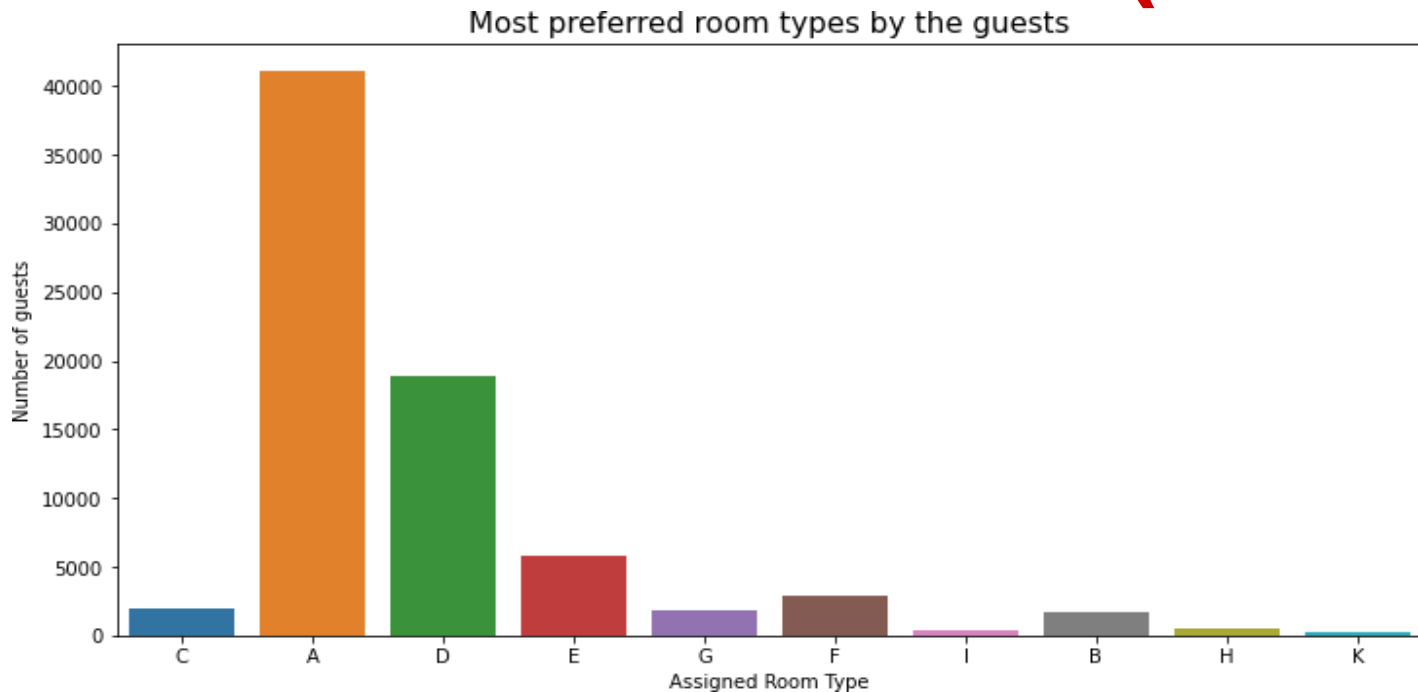
EDA & Observations (Cont.)



Observation :

- The figure shows that the average price per room depends on its type.
- For Resorts G,H and C type of Rooms are most costly.
- For City Hotels, G type rooms are most costly. They are more expensive than the resort hotels

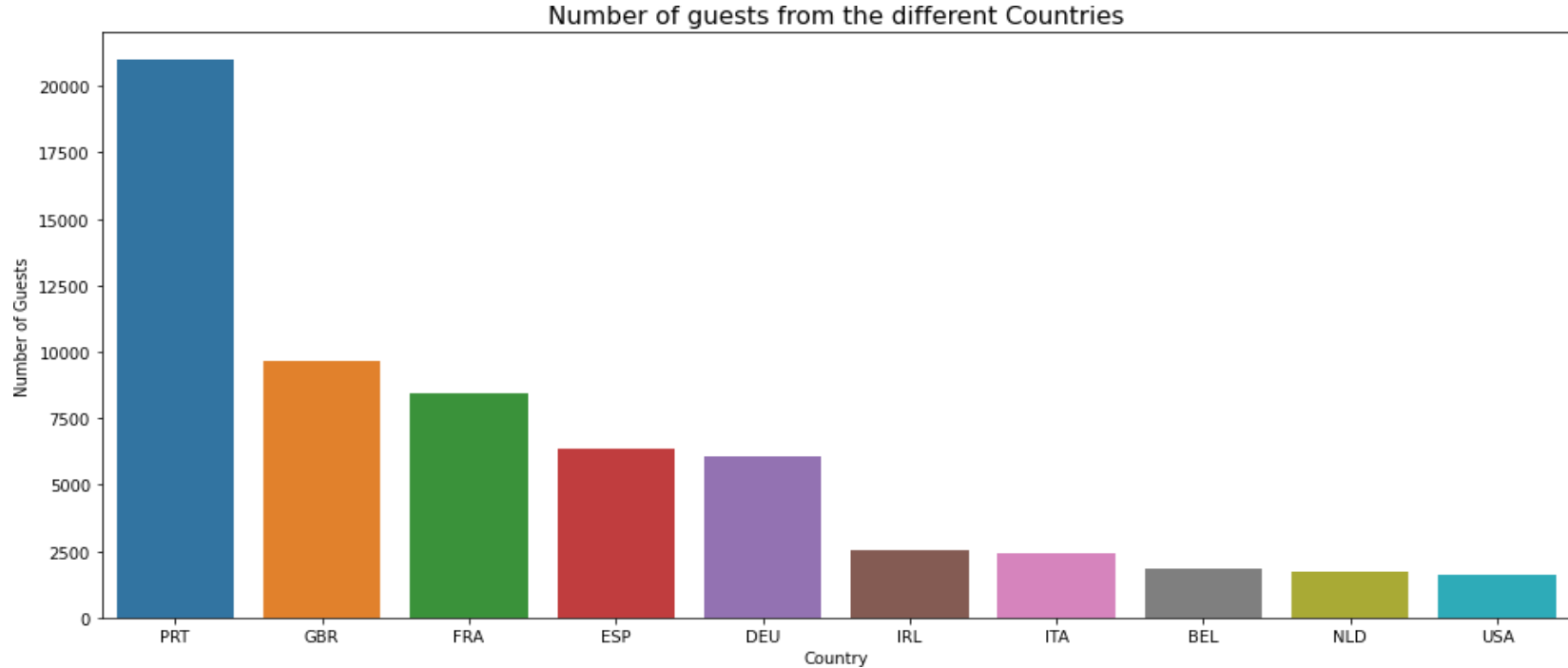
EDA & Observations (Cont.)



Observation:

- The most preferred room is A (Room Type).

EDA & Observations (Cont.)

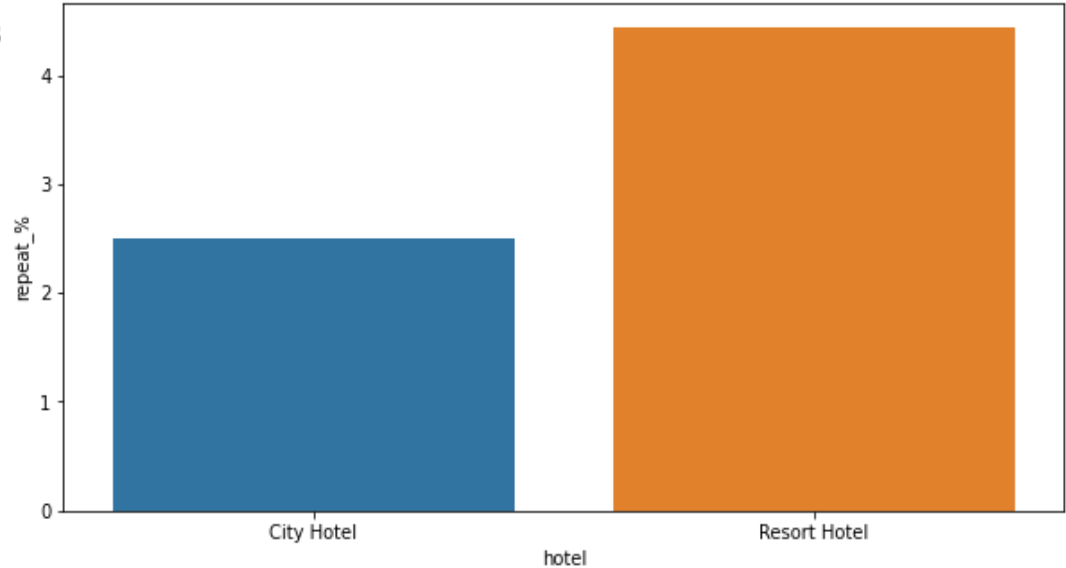
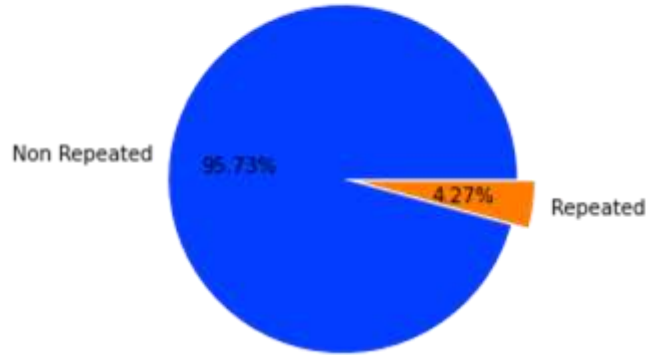


Observation:

- *Most of the guests are coming from Portugal. More than 25000 guests are from Portugal*

EDA & Observations (Cont.)

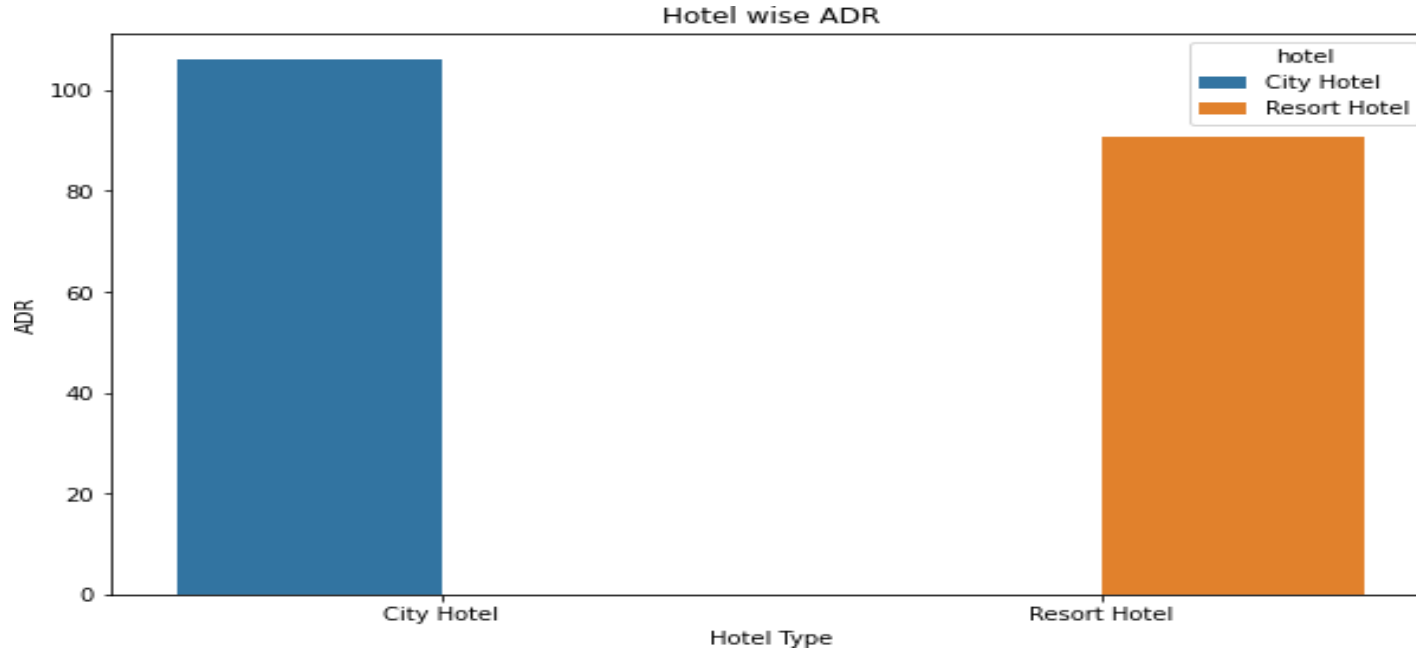
Percentage of Repeated and Non Repeated Guests



Observation:

- *Repeated guest are very few which only 4.27%.*
- *In order to retained the guest management should take feedbacks from guests and try to improve the services.*
- *Both hotels have very small percentage that customer will repeat ,but resort hotel has slightly high.*

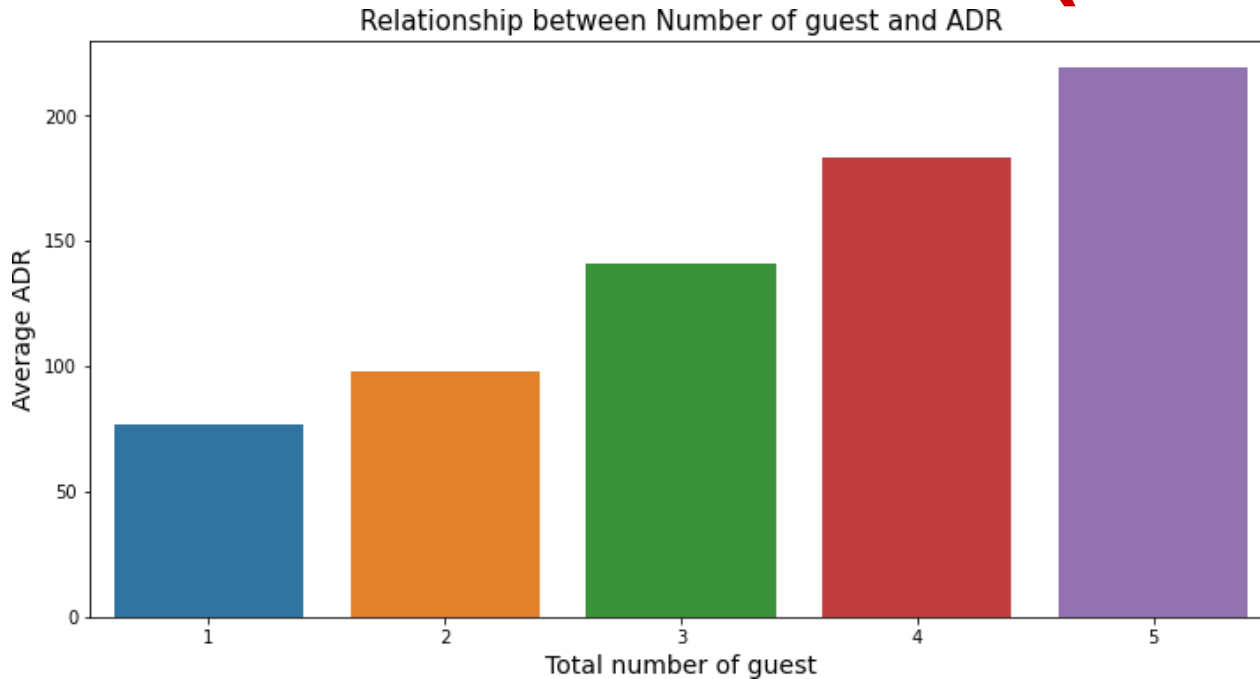
EDA & Observations (Cont.)



Observation:

- *Resort hotel has the lowest ADR. That means Resort hotels are generating less revenues than the city hotels.*
- *More the ADR more is the revenue.*

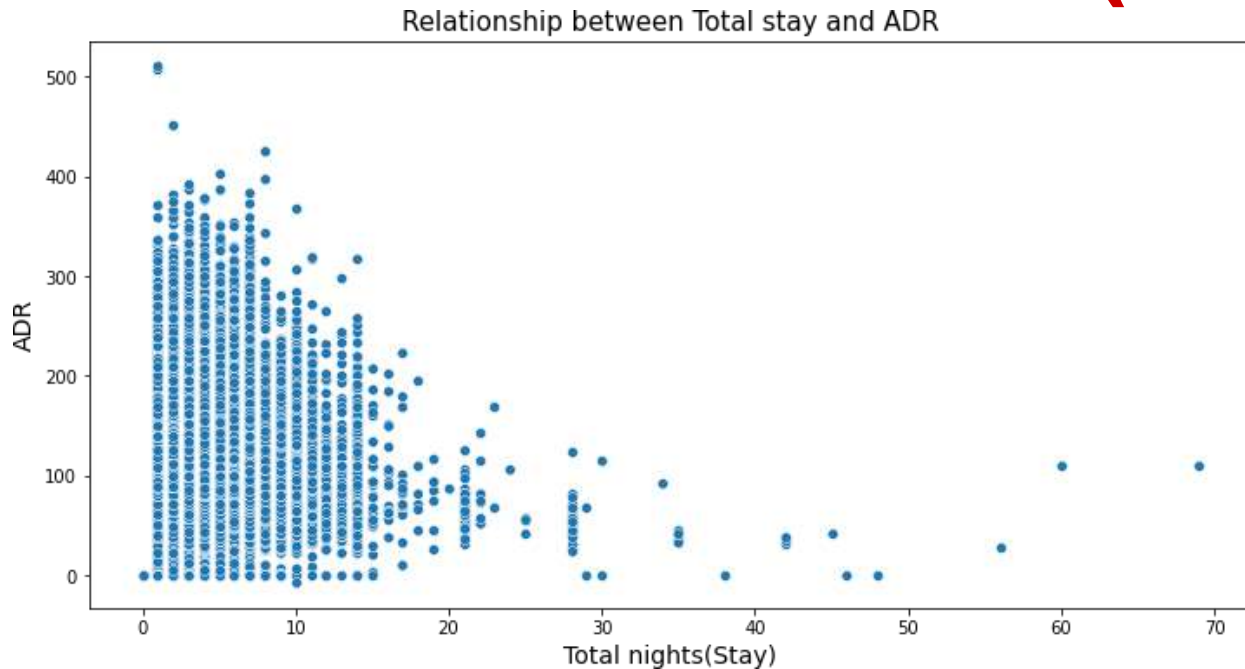
EDA & Observations (Cont.)



Observation:

- As the total number of people increase . Thus ADR and total people are directly proportional to each other .

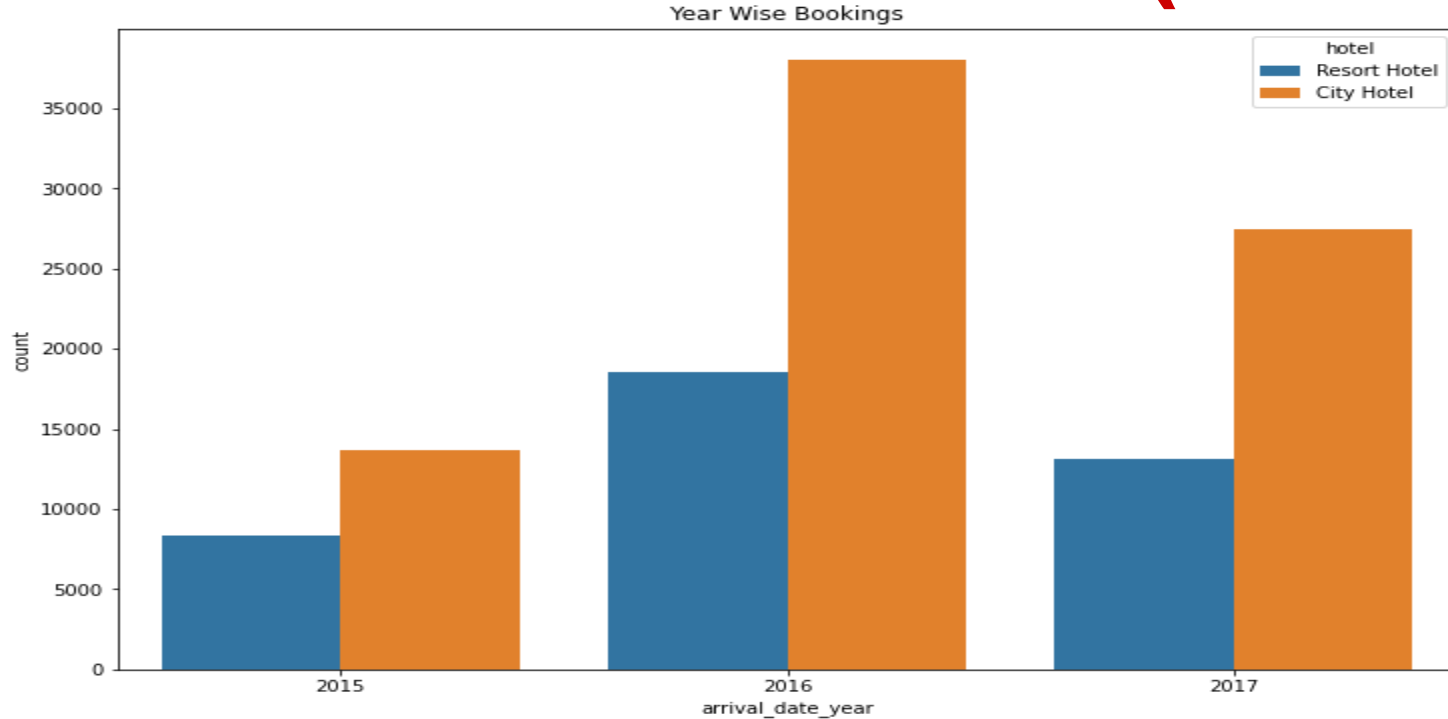
EDA & Observations (Cont.)



Observation:

- As the total stay increase ADR is decreasing . Thus , for longer stays customer can get good ADR(price).

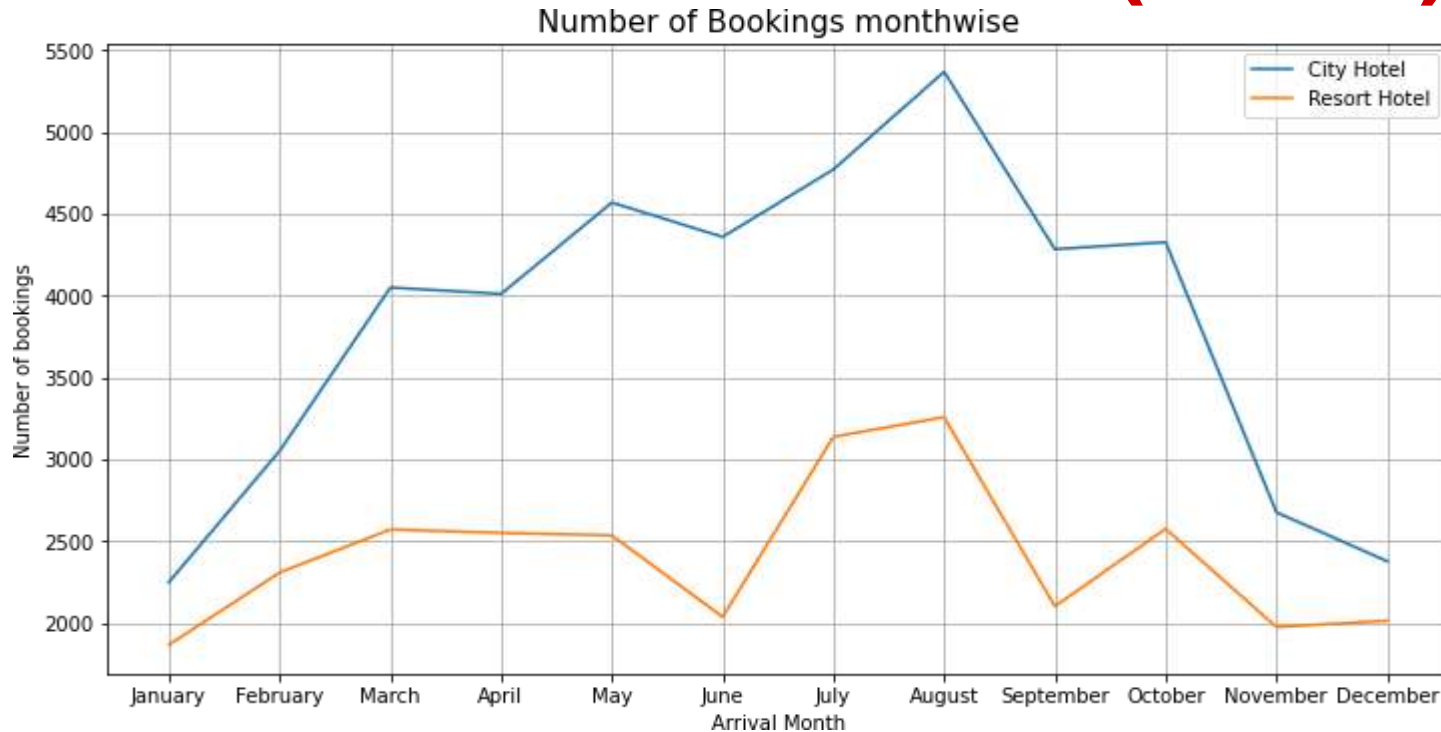
EDA & Observations (Cont.)



Observation:

- 2016 has the highest bookings . And 2015 had less Bookings.
- Overall City hotels had the most of the bookings.

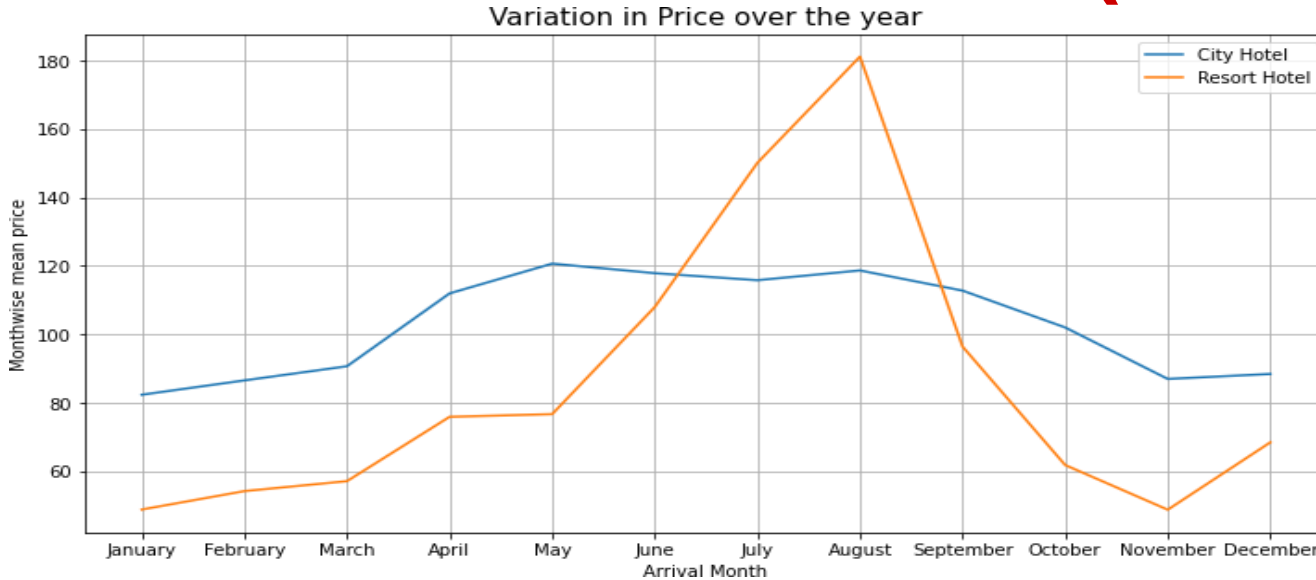
EDA & Observations (Cont.)



Observation:

- *July and august monthly had the most bookings .Sumer vacation can be reason for bookings.*

EDA & Observations (Cont.)



Observations:

- For Resort hotel ADR is high in the month June, July, August as compared to City Hotels. May be Customers/People wants to spend their Summer vacation in Resorts Hotels.
- From 9 observation, it is cleared that the, both type of hotels have the fewest guests during the winter. Therefore, the best time for guests to visit Resort or City hotels is January, February, March, April, October, November and December as the average daily rate in this month is very low.

EDA & Observations (Cont.)



Observation:

- ❑ *Maximum customer do not made any special requests along with their bookings.*
- ❑ *Number of special request is higher in case of city hotel, because the number of bookings is also higher.*
- ❑ *In the end, we can say that, very few customers made more than 2 special requests.*

EDA & Observations (Cont.)

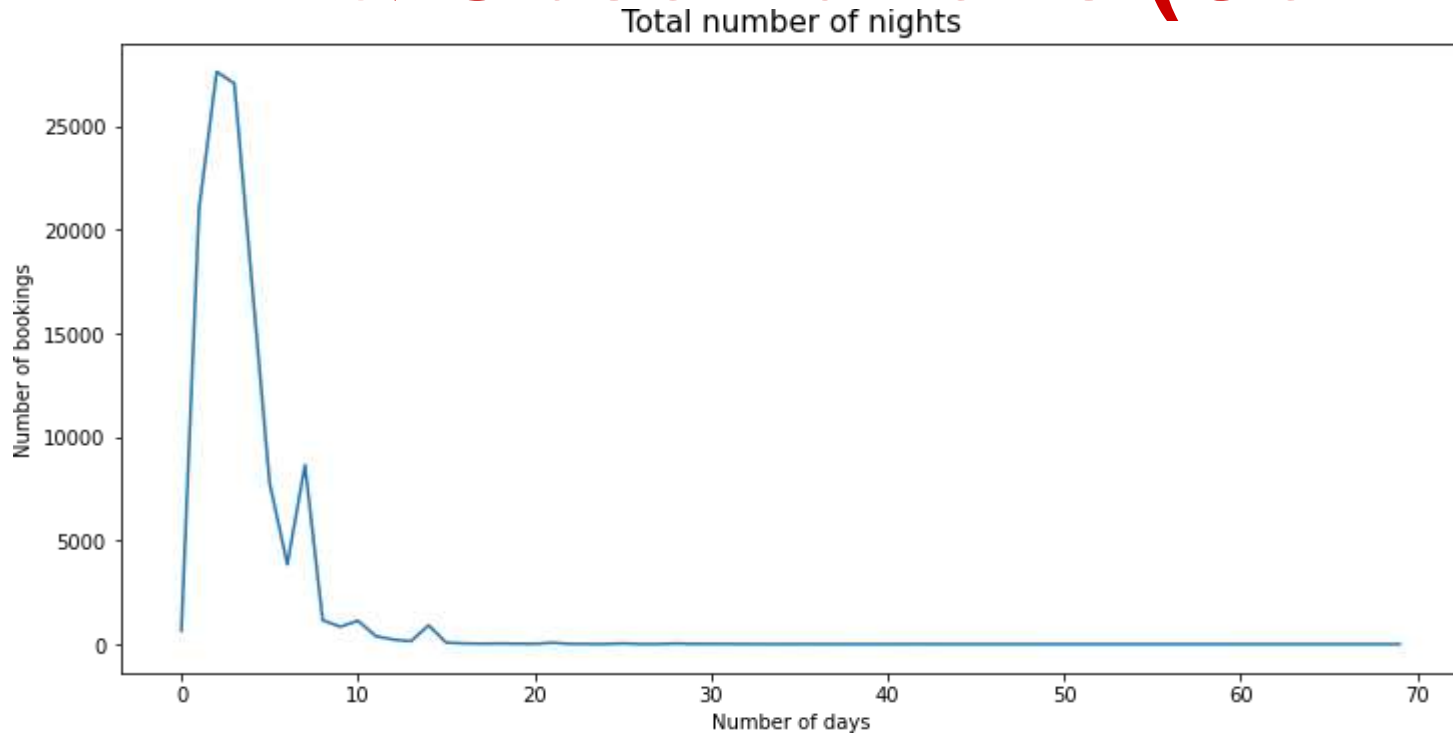


EDA & Observations (Cont.)

Observation:

- *is_canceled and same_room_alloted_or_not are positively correlated. That means customer is sometimes cancel his bookings if he don't get the same room as per reserved room.*
- *lead_time and total_stay is positively correlated. That means more is the stay of customer more will be the lead time.*
- *total number of guest is positively correlated to each other. That means more the people more will be adr.*
- *is_repeated guest and previous bookings not canceled has strong correlation. May be repeated guests are not more likely to cancel their bookings.*

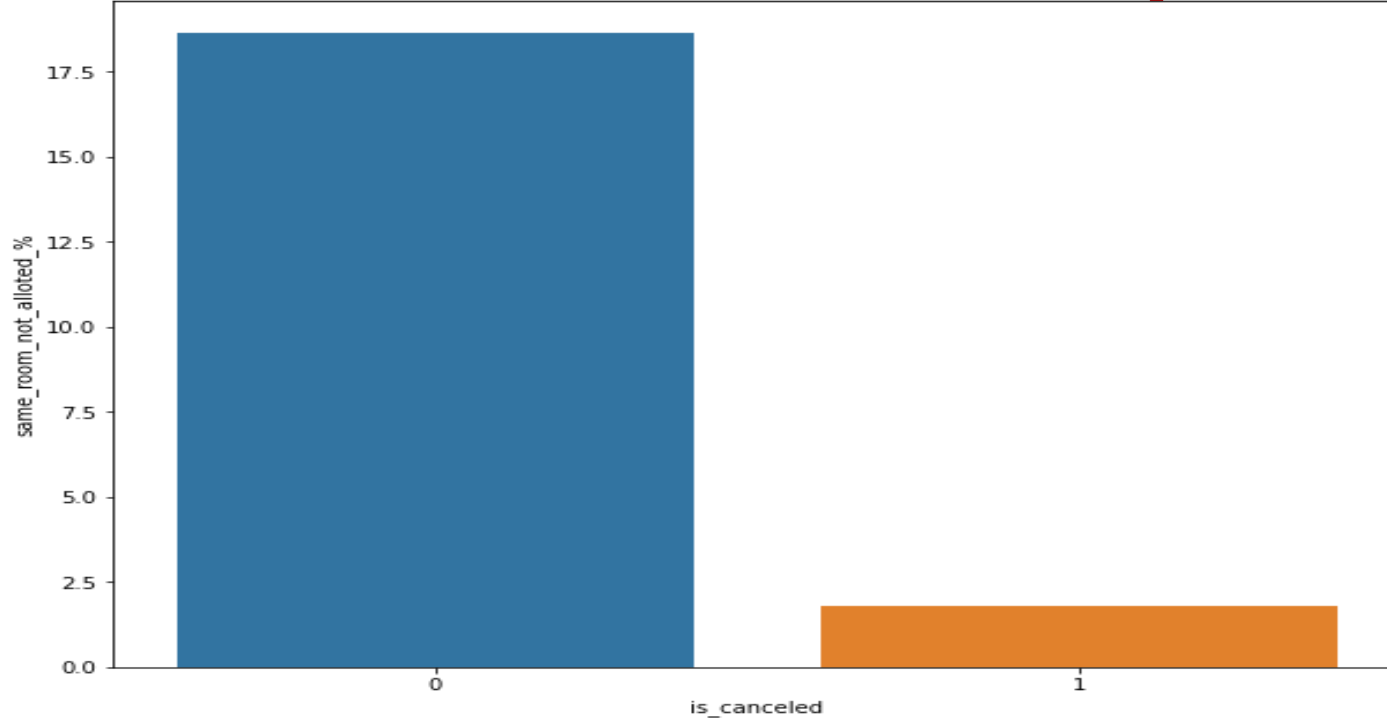
EDA & Observations (Cont.)



Observation:

- ❑ *Guests booked hotel mostly for 0 to 3 nights.*
- ❑ *Very few guests booked for more than 8 days.*

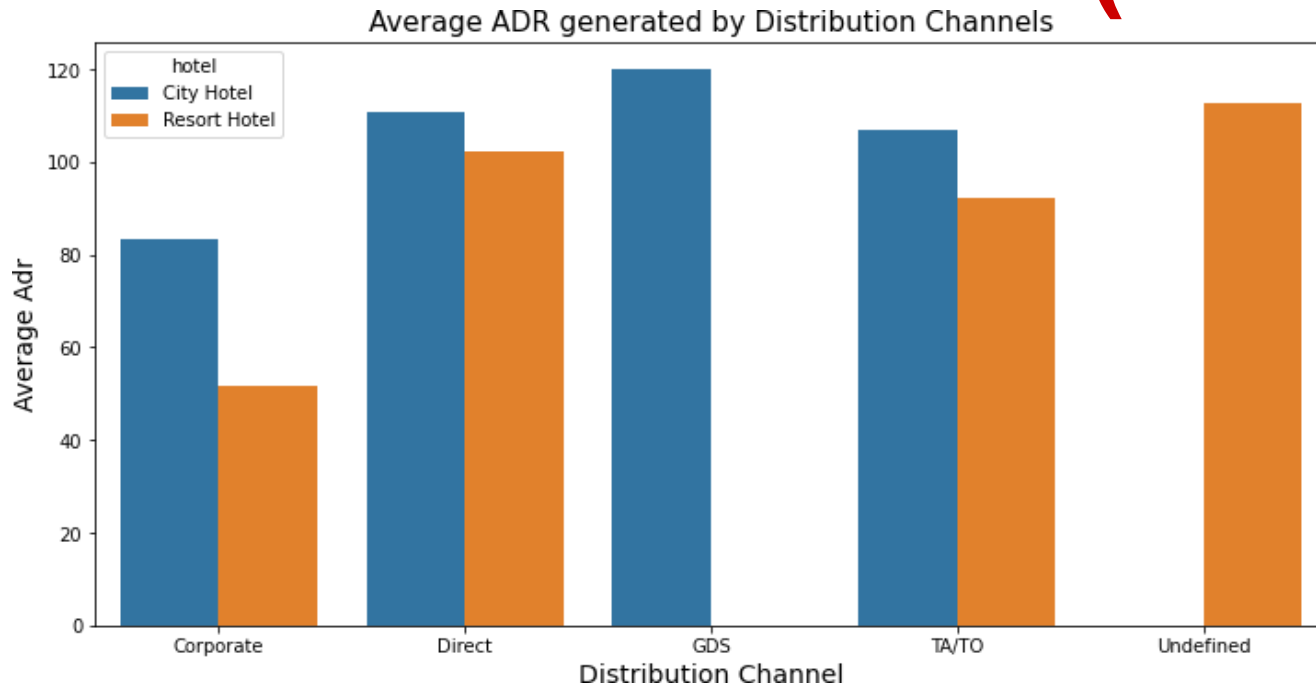
EDA & Observations(Cont.)



Observation:

- ❑ We see that not getting same room as demanded is not the case of cancellation of rooms. A significant percentage of bookings are not cancelled even after getting different room as demanded.

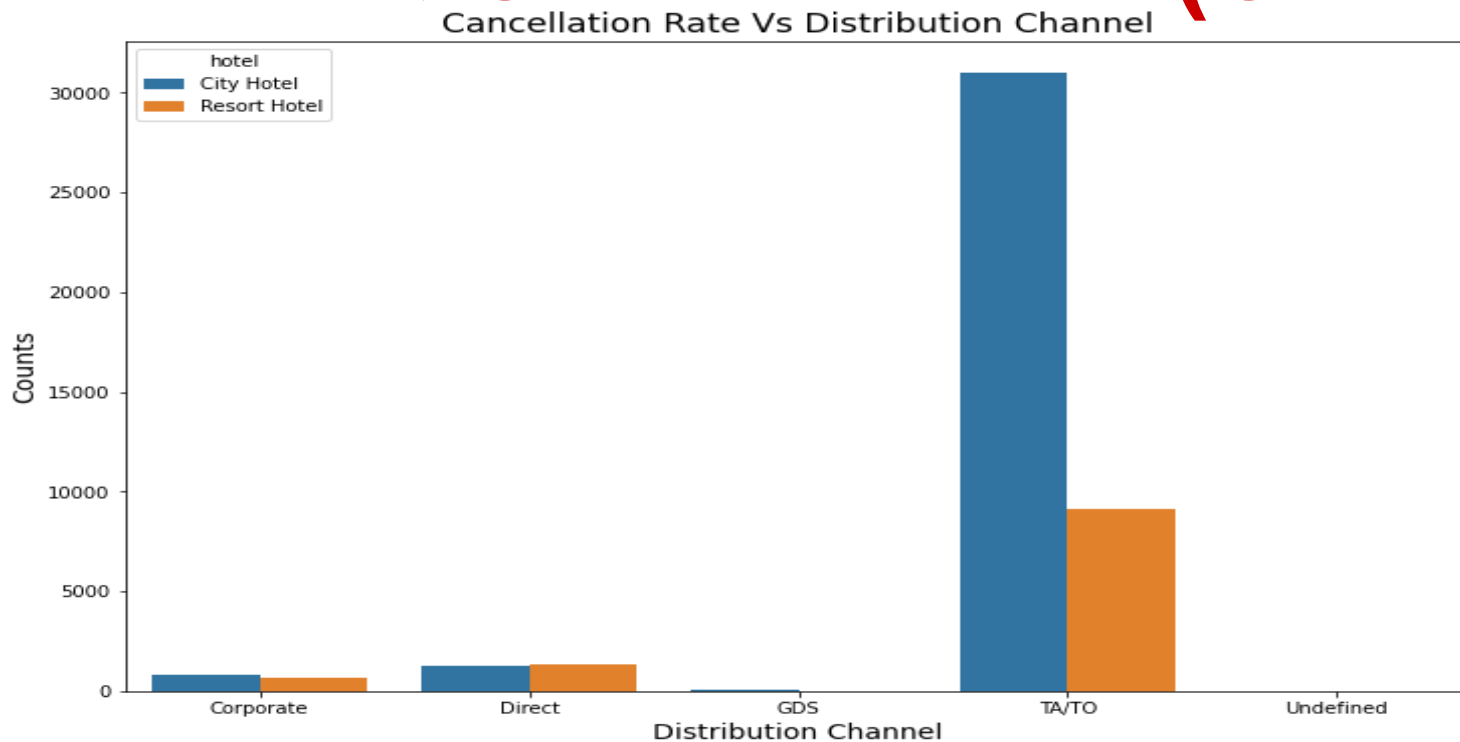
EDA & Observations (Cont.)



Observation:

- ❑ *Direct' and 'TA/TO' has almost equally contributed in ADR in both types of hotels.*
- ❑ *GDS has highly contributed in ADR in 'City Hotel' type.*
- ❑ *GDS need to increase Resort Hotel Bookings, for increasing its ADR.*
- ❑ *Resorts made high ADR by Undefined mode of booking also.*

EDA & Observations (Cont.)



Observation:

- ❑ In "TO/TA " City hotels has the high cancellation rate compared to resort hotels.
- ❑ In "Direct" both the hotels has almost same cancellation rate.

Final Conclusion

Above observations are very useful for Hotel team for the growth of their business. Alongwith above observations, I made the some important conclusion for the Hotels and Customer given below:

1. Best months for planning a trip are October to February, because prices for both the hotels are lessor as compare to other months due to less bookings.
2. Guest numbers for the Resort hotel go down slightly from June to September, which is also when the prices are highest. Thus, these months should be avoided for booking.
3. Very large number of customers are cancelling their bookings, so hotels need to make strict cancellation policy, like they can use non-refund options.

Final Conclusion(Cont.)

4.Cancellations are high when done through agents compared to direct booking. Hotels need to do marketing and give special incentives for direct booking as these may establish personal one to one relationship promoting customer loyalty.

5.The number of repeated guest is very low, in order to retained the guests management should take feedbacks from guests and try to improve the services.

Thank
You