Capstone Project
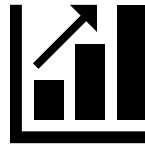## Hotel Bookings Data Analysis
By-Aakash Yadav
Cohort-Seattle

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 hotel generally people prefer to book?

Q2. What is the percentage of cancellation of Bookings?

Q3. Which type of customers do more bookings?

Q4. What is the percentage of repeated guest?

Q5. Which type of deposit is more preferred by the customers?

Q6. Which kind of food is mostly preferred by the guests?

Q7. From which country mostly guests are coming from?

Q8. What is the most preferred room type?

Q9. How much guests pay for a room per night?

Q10. Which months of the year are busiest for bookings?

Q11. How does the price vary over the year?

Questions (Cont.)

Q12.How long people stays in hotel?

Q.13 What is the most commonly used distribution channel for hotel bookings?

Q14. Which hotels generating more ADR?

Q15. Which type of hotel has longer waiting time?

Q16. Which distribution channel contributed more to generate high ADR?

Q17.Whether or not a hotel was likely to receive a disproportionately high number of special requests?

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

Q19. What is the relationship between total number of Guests and ADR?
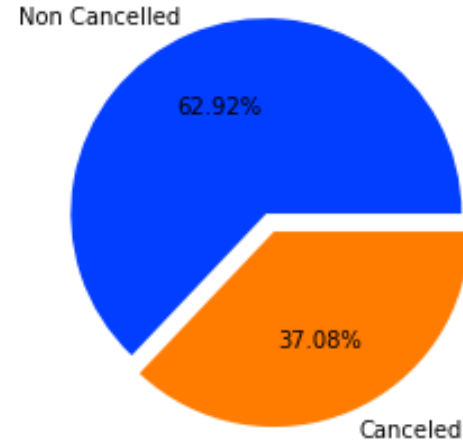
Q20. What is the relationship between total stay and ADR?

EDA & Observations

Most preferred hotel

Percentage of canceled/non canceled bookings
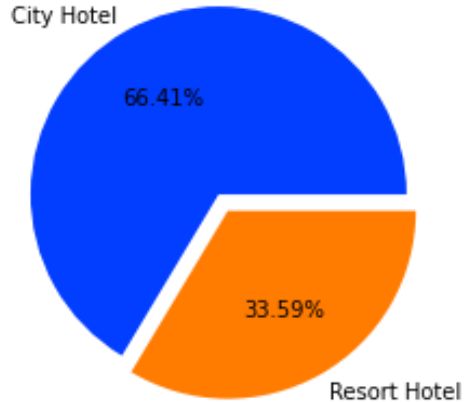
Observation: People prefer to book City Hotels more as compare to Resort Hotels.
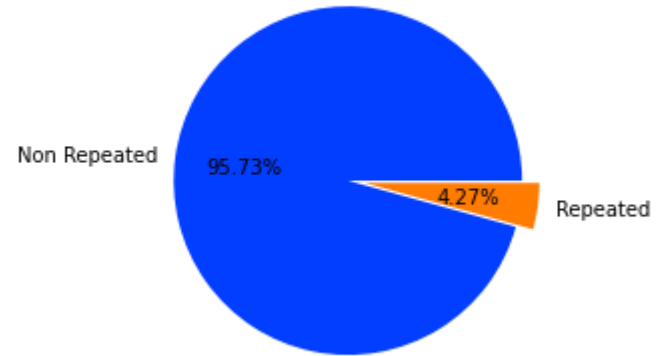
Observation: 37.08% bookings got canceled by the guests.

EDA & Observations (Cont.)

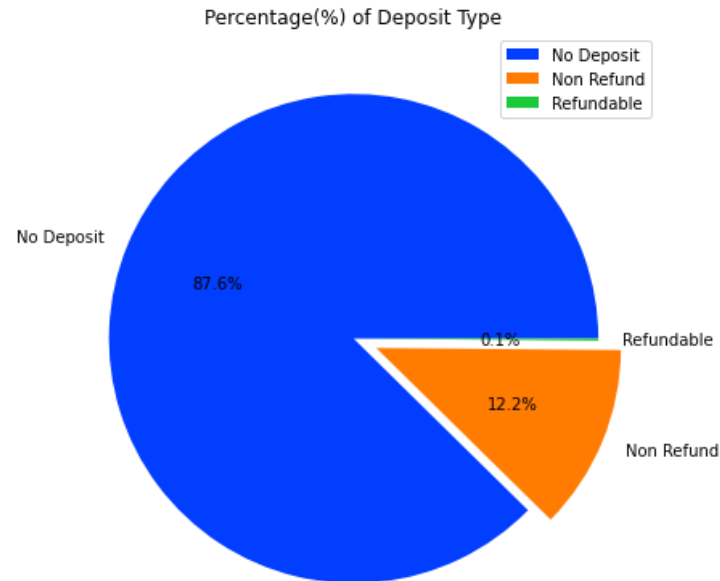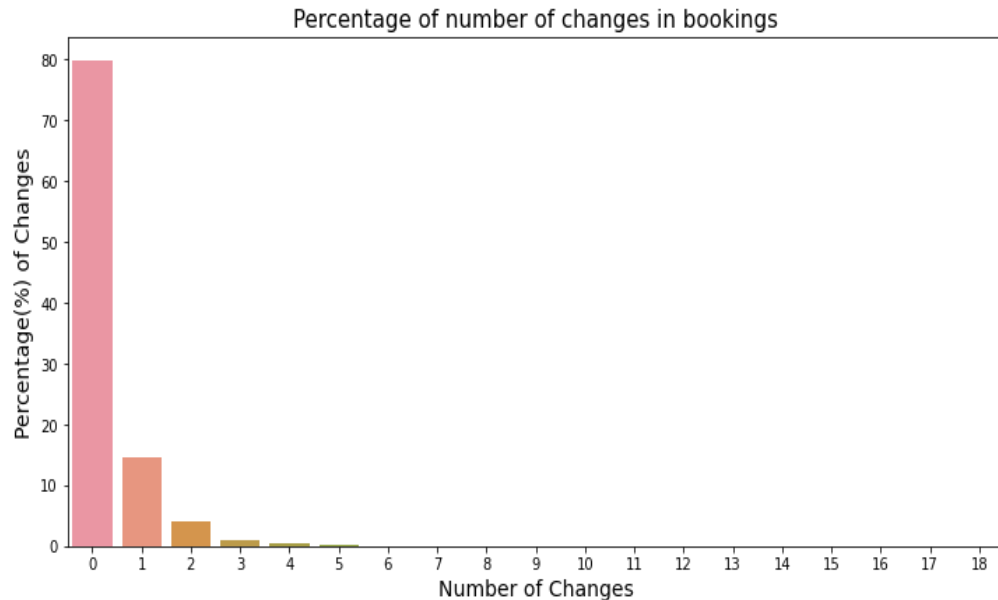Percentage of bookings by customer type
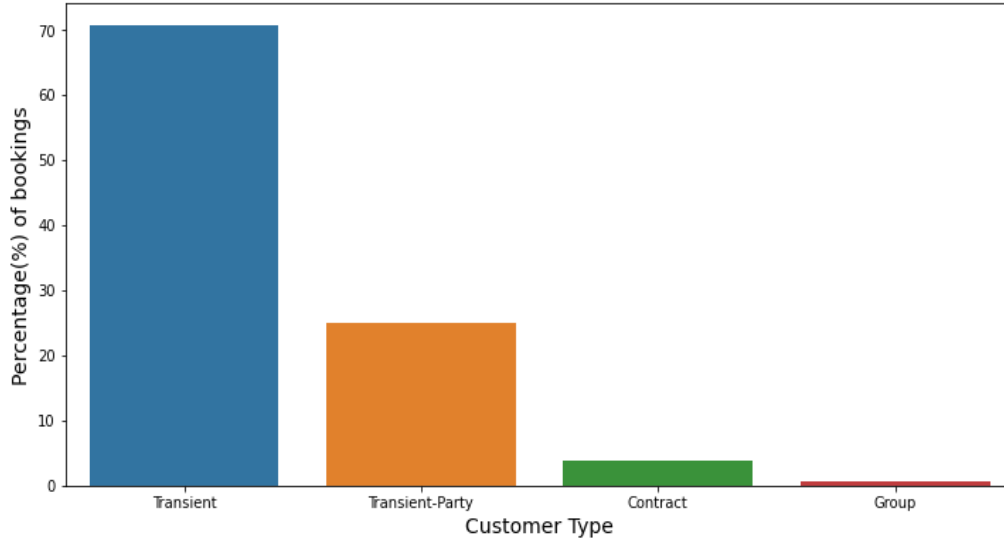
Percentage of Repeated and Non Repeated Guests

Observations:

•Booking by the Transient Customer is very high as compare to other type of customers(approx. 70% of all the bookings).
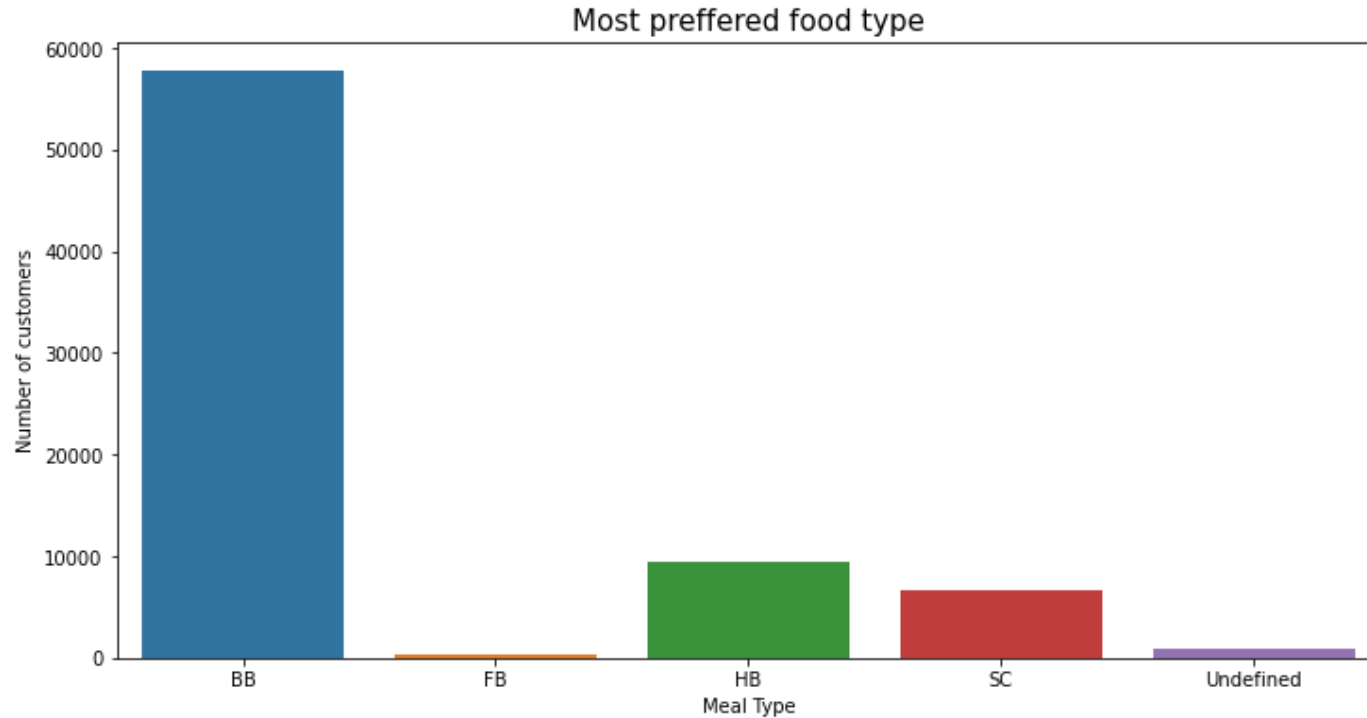
•Repeated guest are very few, which only 4.27%.

EDA & Observations (Cont.)

Observations:

- Almost 80% of the bookings were not changed by guests.
- Some time guests made 1 or 2 changes in their bookings. Otherwise, its very rare.
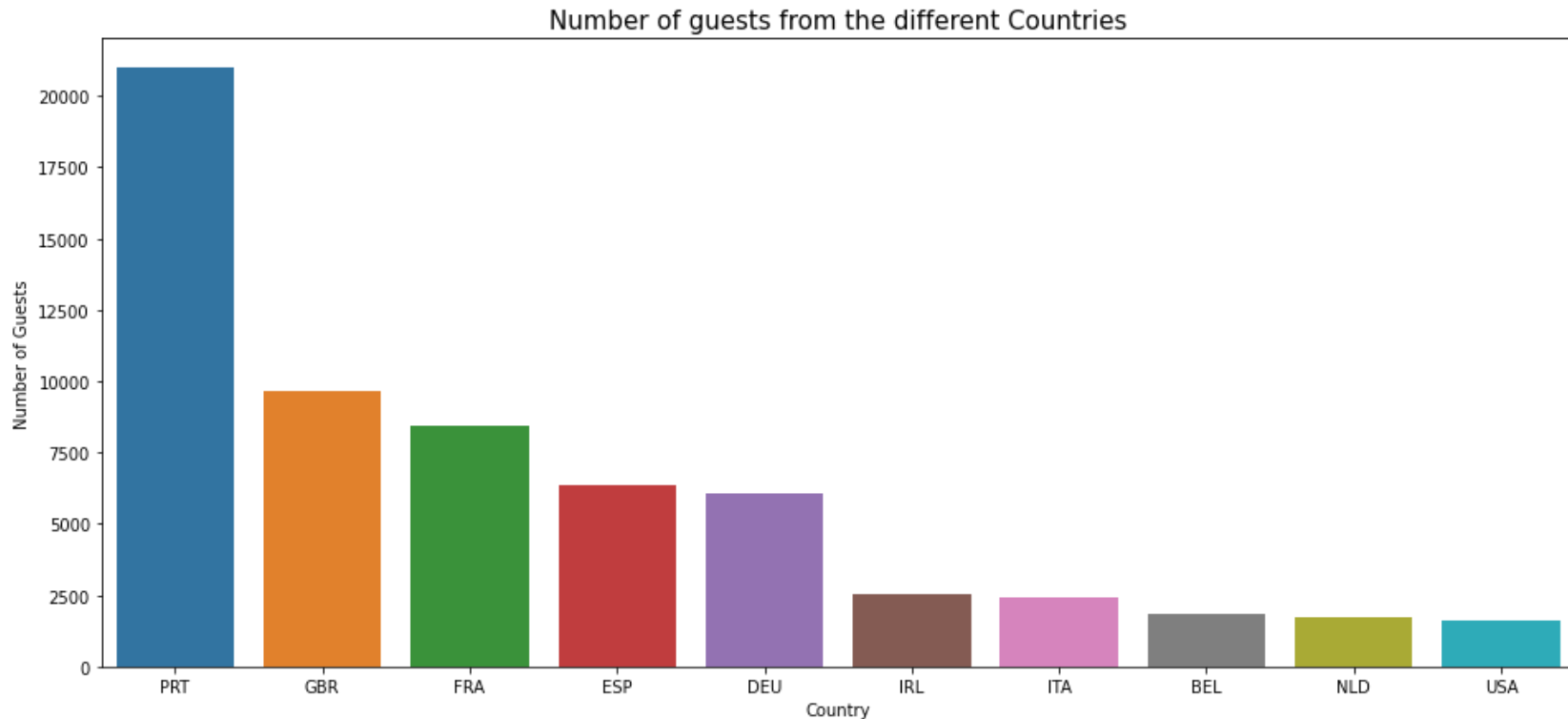- 87.6% of guests prefer "No deposit" type of deposit.

EDA & Observations (Cont.)

Most preffered food type

Observations:
- Most preferred food by the customers is BB(Bed and Breakfast).
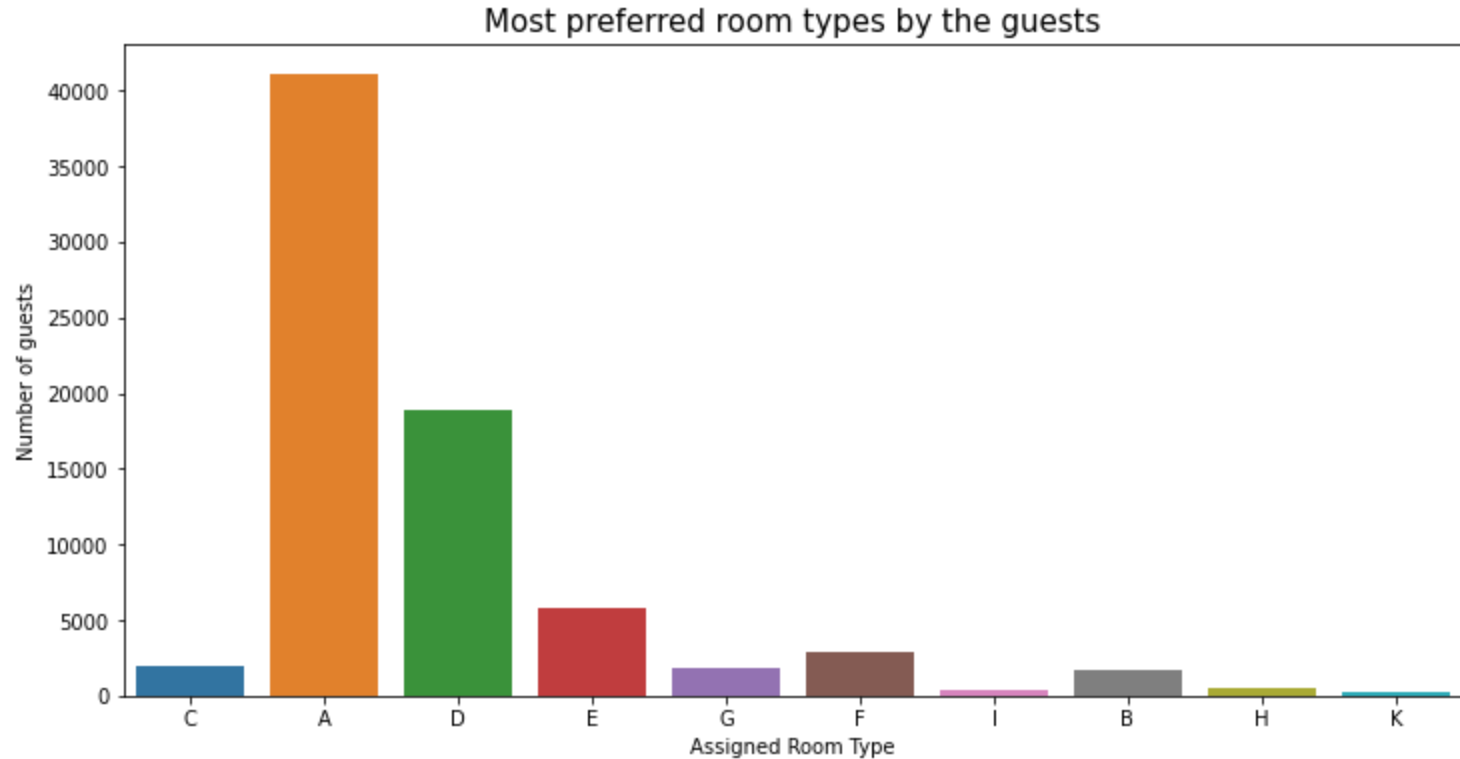- HB(Half Board) and SC(Self catering) are apprx. equally preferred.
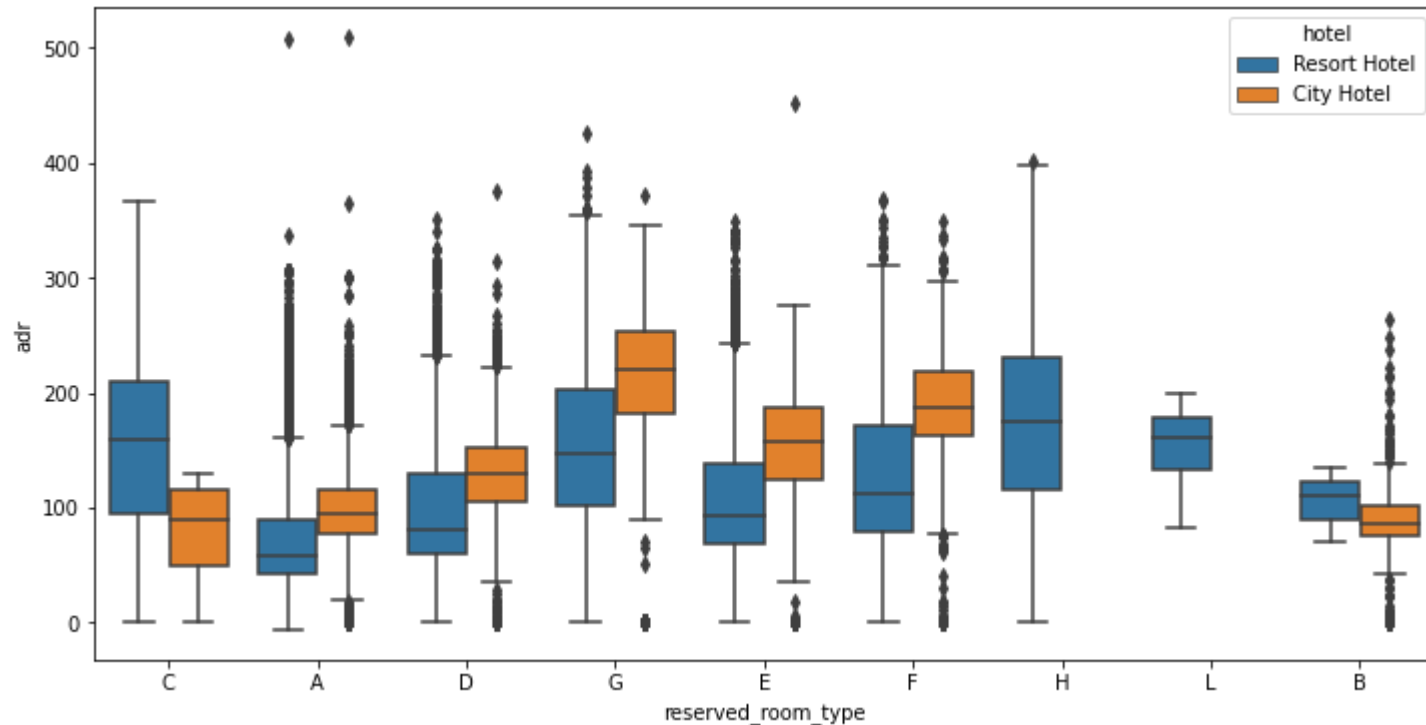
EDA & Observations (Cont.)

Number of guests from the different Countries

Observation:

•Most of the guests are coming from Portugal. More then 25000 guests are from Portugal.
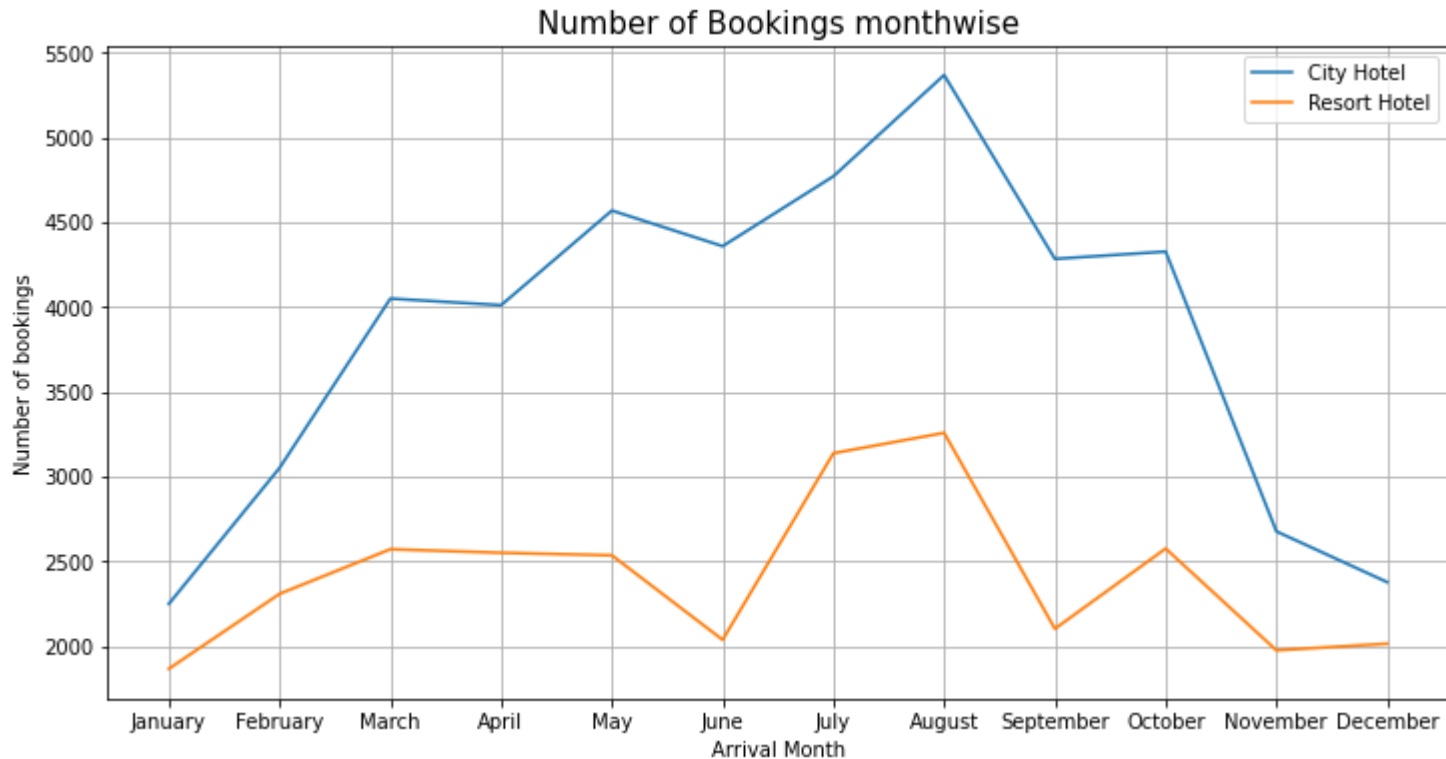
EDA & Observations (Cont.)

AI

Most preferred room types by the guests

Observation:
- The most preferred room is A.

EDA & Observations (Cont.)

Observations

- 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.
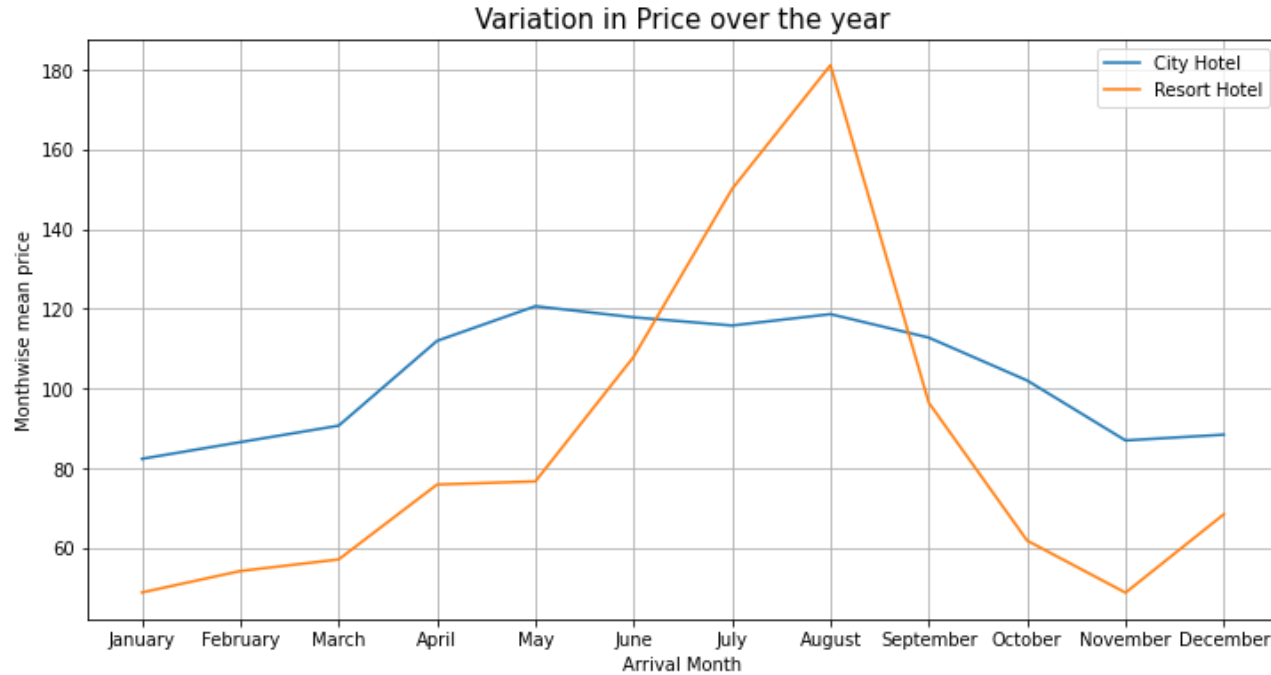- For City Hotels, G type rooms are most costly. They are more expensive then the resort hotels.

EDA & Observations (Cont.)

Number of Bookings monthwise

Observation:
•July and August months had the most Bookings. Summer vacation can be the 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.
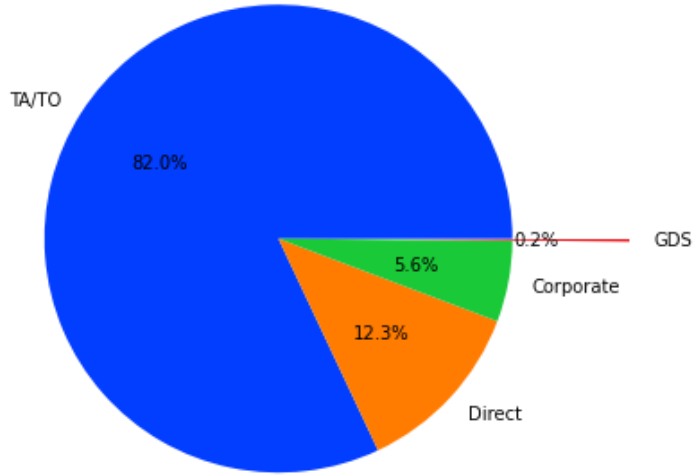
Total number of nights
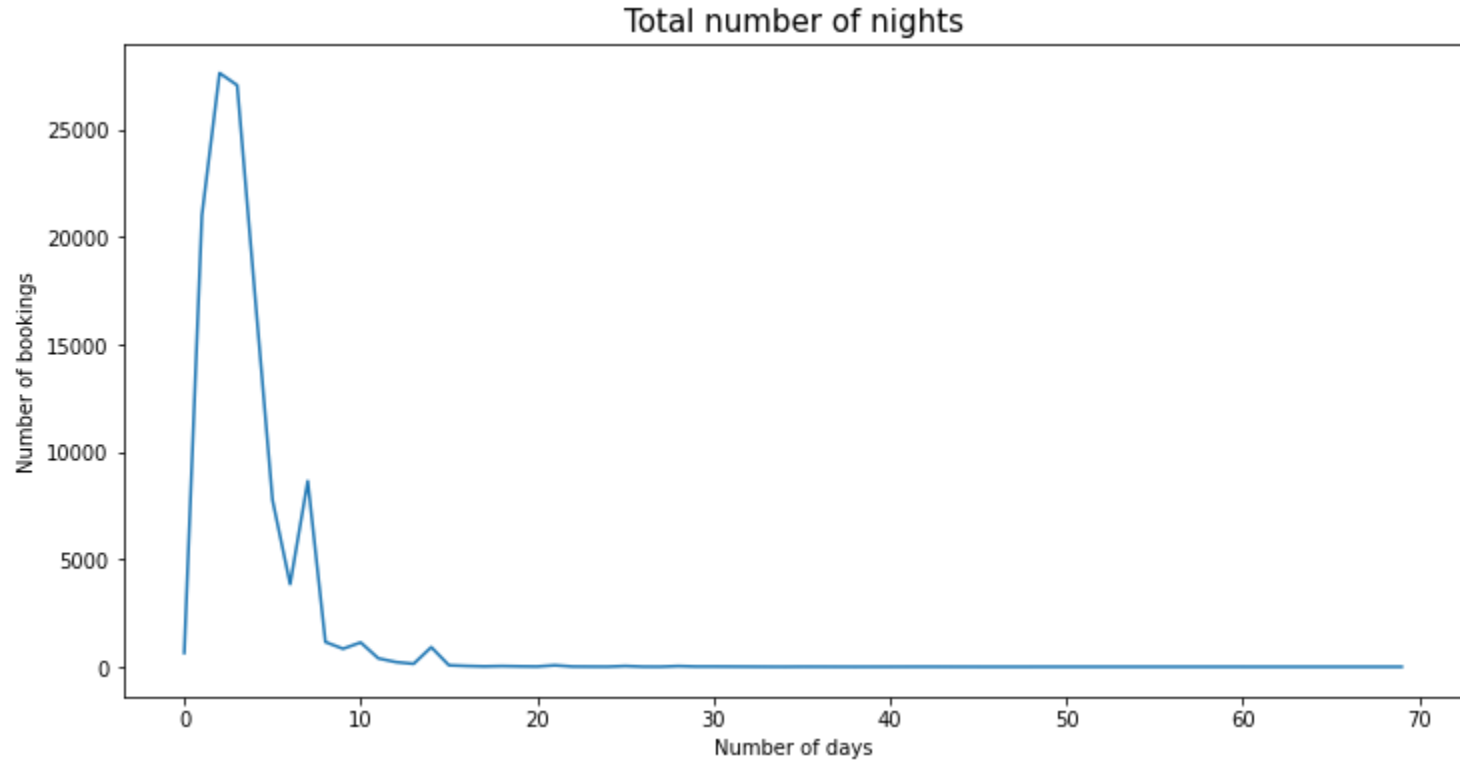
Observation:
- Guests booked hotels mostly for 0 to 3 nights.
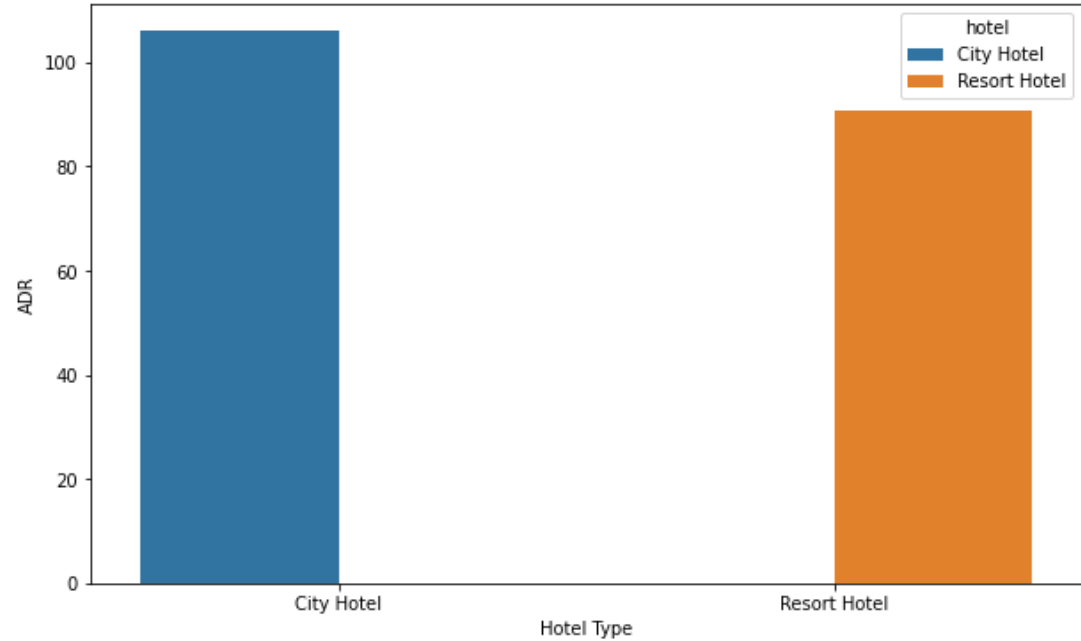- Very few guests booked hotels for more than 8 days.

EDA & Observations (Cont.)

Disribution channel used for bookings
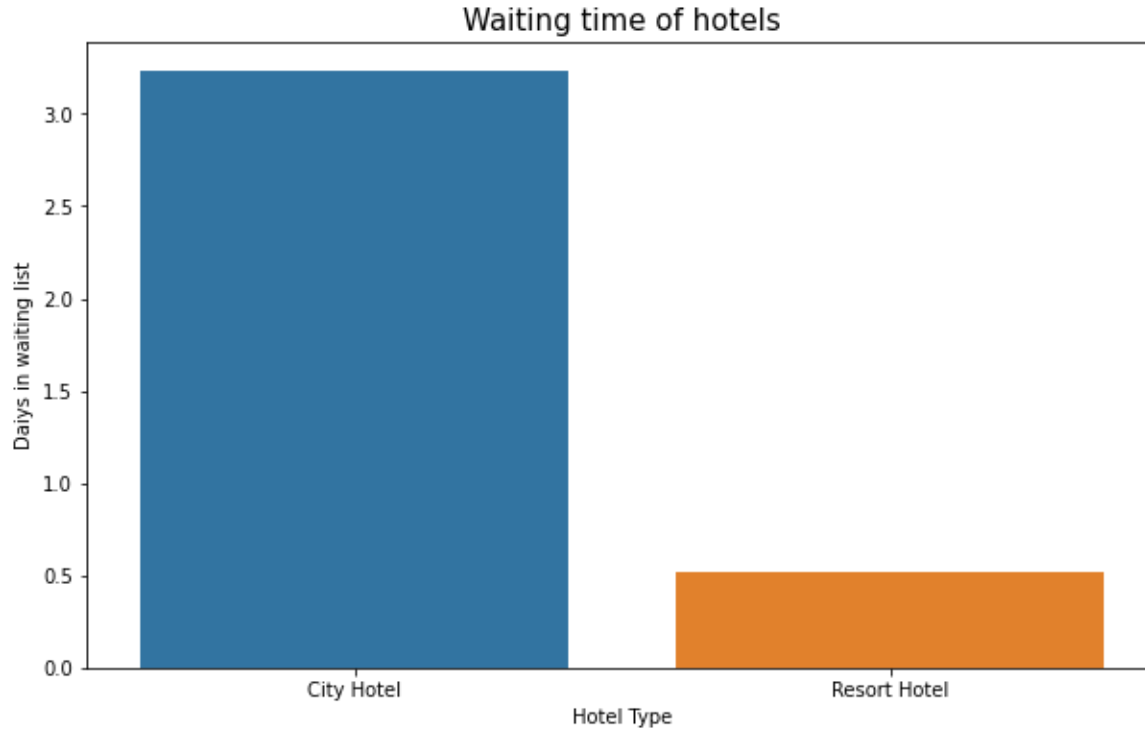

Hotel wise ADR

Observations:
- TA/TO is mostly (82%) used for booking hotels
- City hotel has the highest ADR. That means city hotels are generating more revenues than the resort hotels. More the ADR more is the revenue.
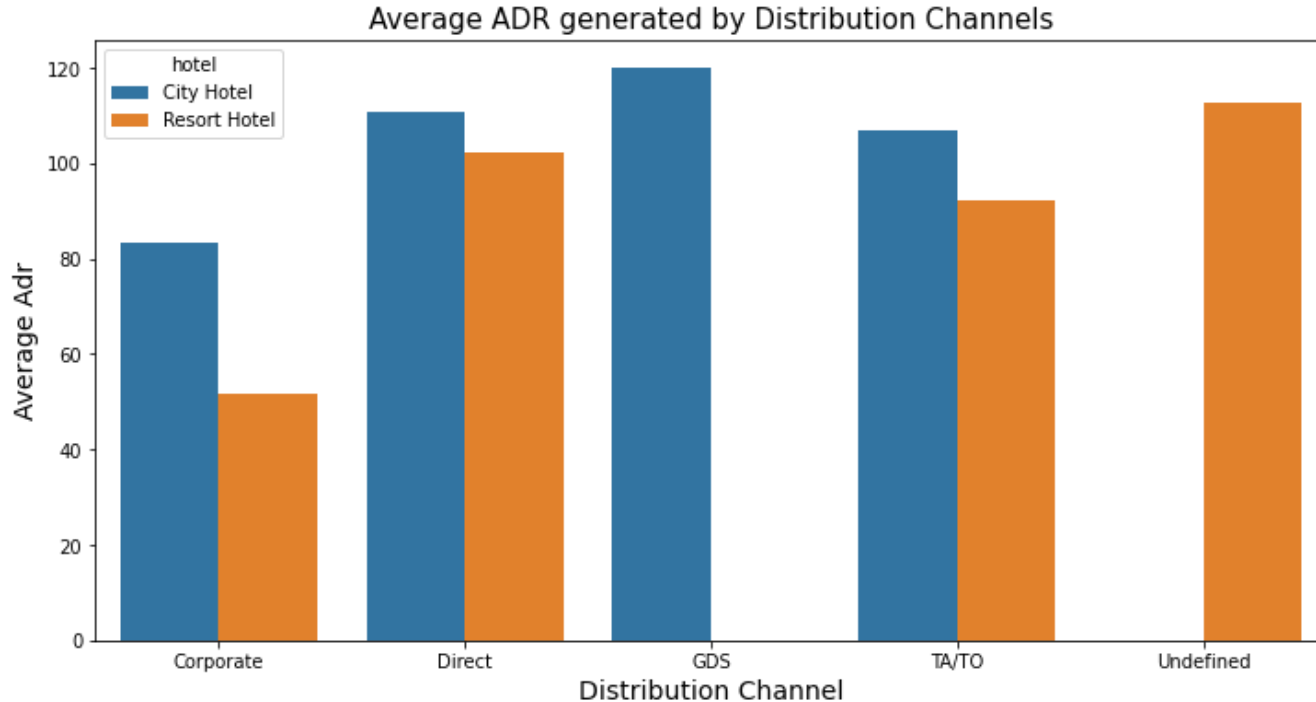
EDA & Observations (Cont.)

Waiting time of hotels

Observation
•City hotel has the higher waiting time.

EDA & Observations (Cont.)

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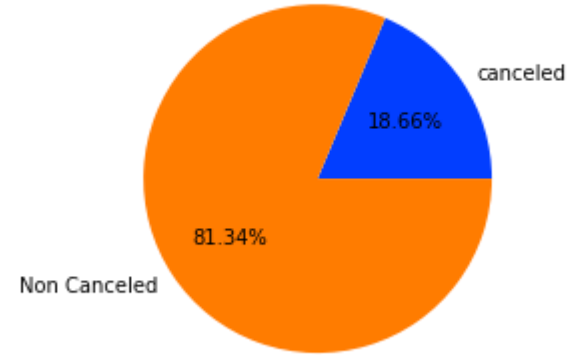
Observation:
- 'Direct' and 'TA/TO' has almost equally contributed in ADR in both types of hotels.
- GDS has highly contributed in ADR in 'City Hotel' type.
- GDS need to increase Resort Hotel Bookings, for increasing its ADR.
- Resorts made high ADR by Undefined mode of booking also.

EDA & Observations (Cont.)

Special requests made with bookings
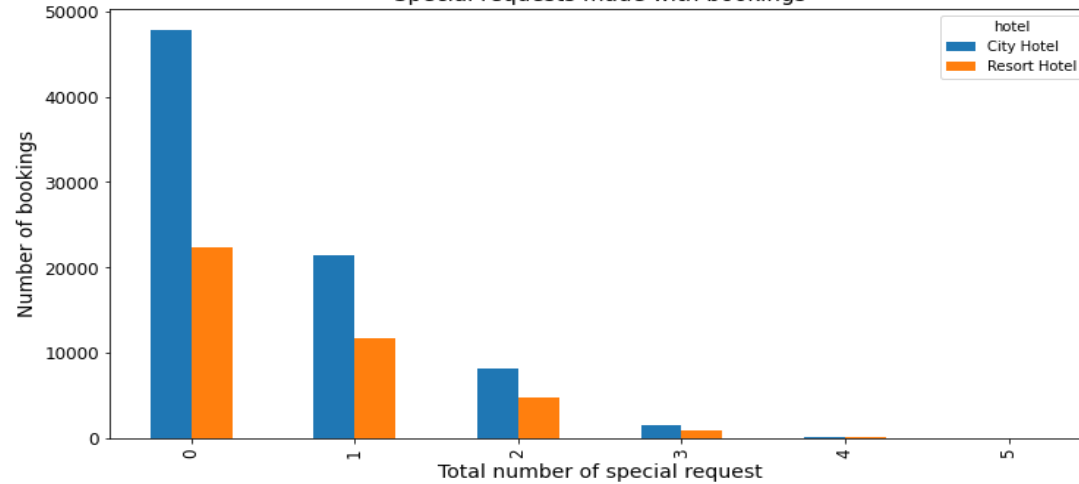
Canceled bookings when the same room was not alloted

Observations:
- Maximum customer do not made any special requests alongwith 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.
- It is noticed that the majority(81.34%) customer do not canceled their booking, when they don't get the desired room.
- 18.66% bookings canceled due to this, so hotels need to take little care about this.

EDA & Observations (Cont.)

Correlation heatmap between all the variables

AI
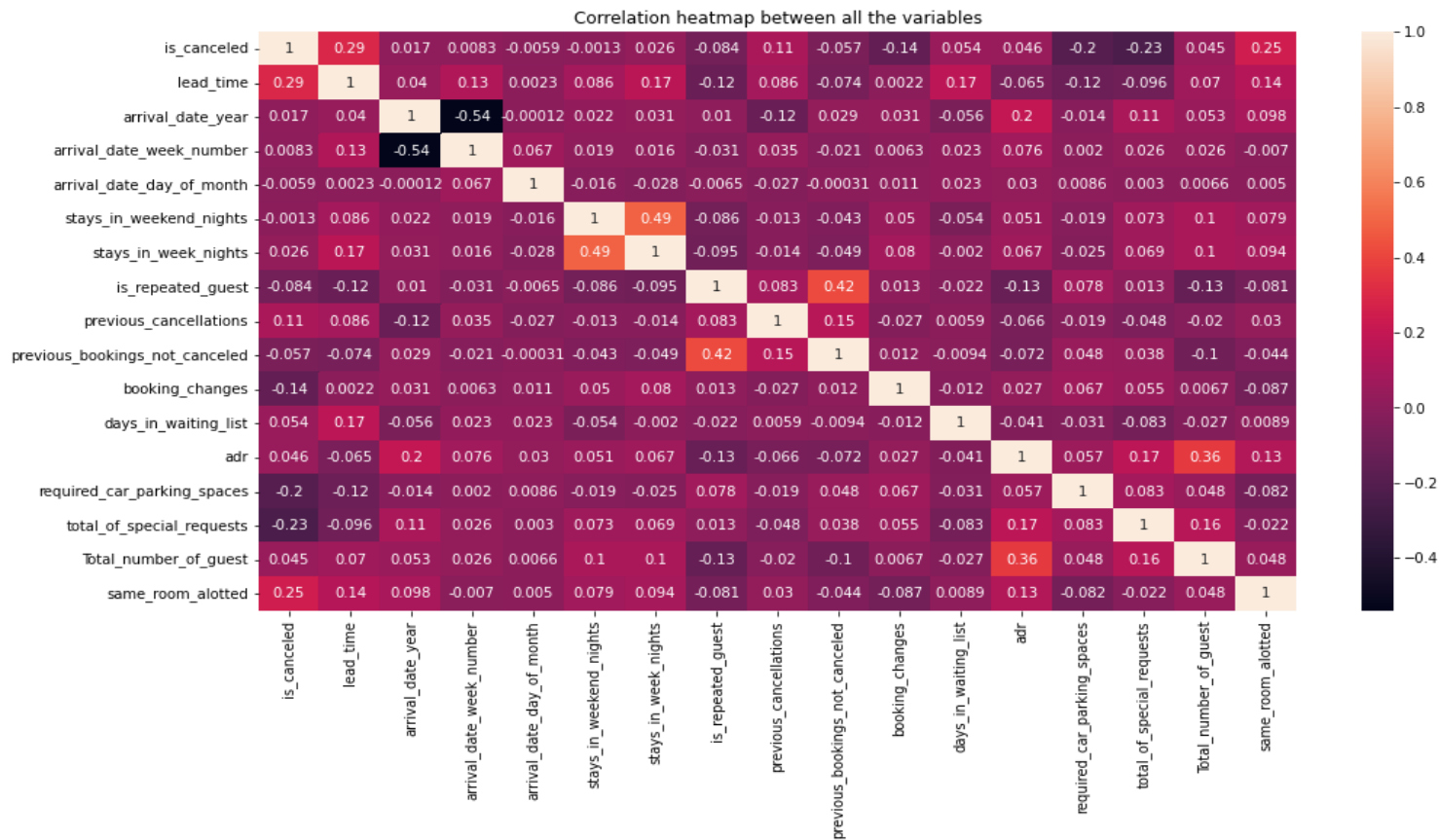
Observation:

• is_canceled and same_room_alloted_or_not are positively corelated. 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 corelated. That means more is the stay of customer more will be the lead time.

• total number of guest is positively corelated to each other. That means more the people more will be adr.

• is_repeated guest and previous bookings not canceled has strong corelation. May be repeated guests are not more likely to cancel their bookings.
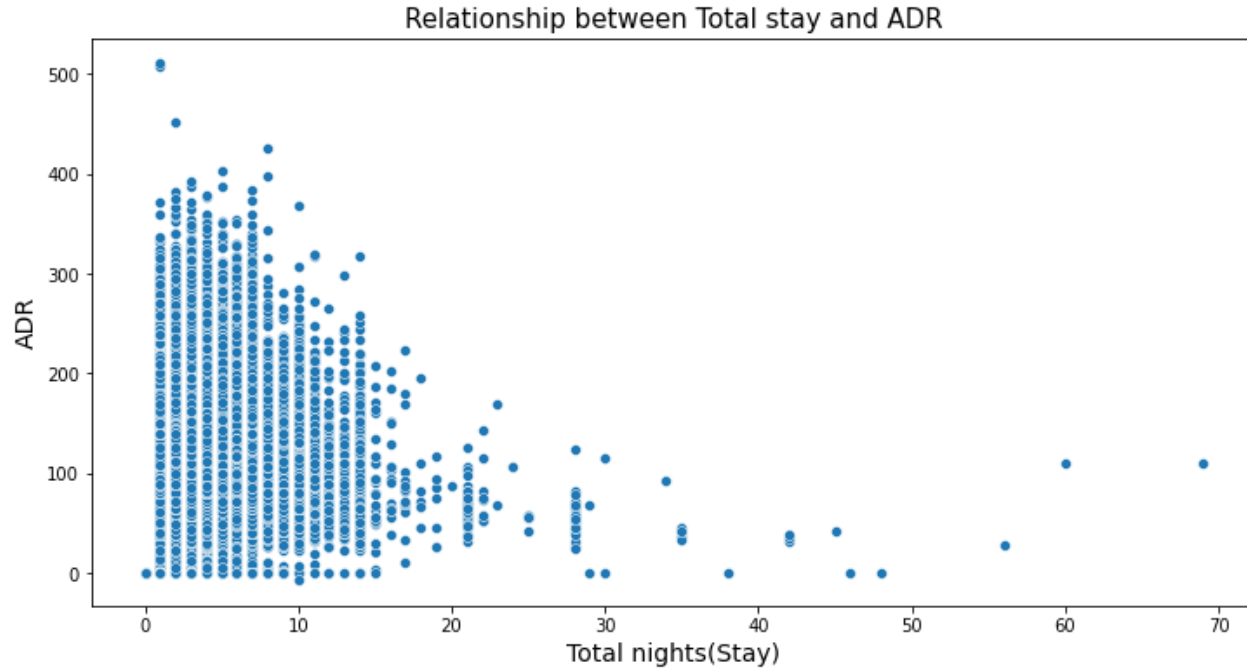
EDA & Observations (Cont.)

Relationship between Number of guest and ADR
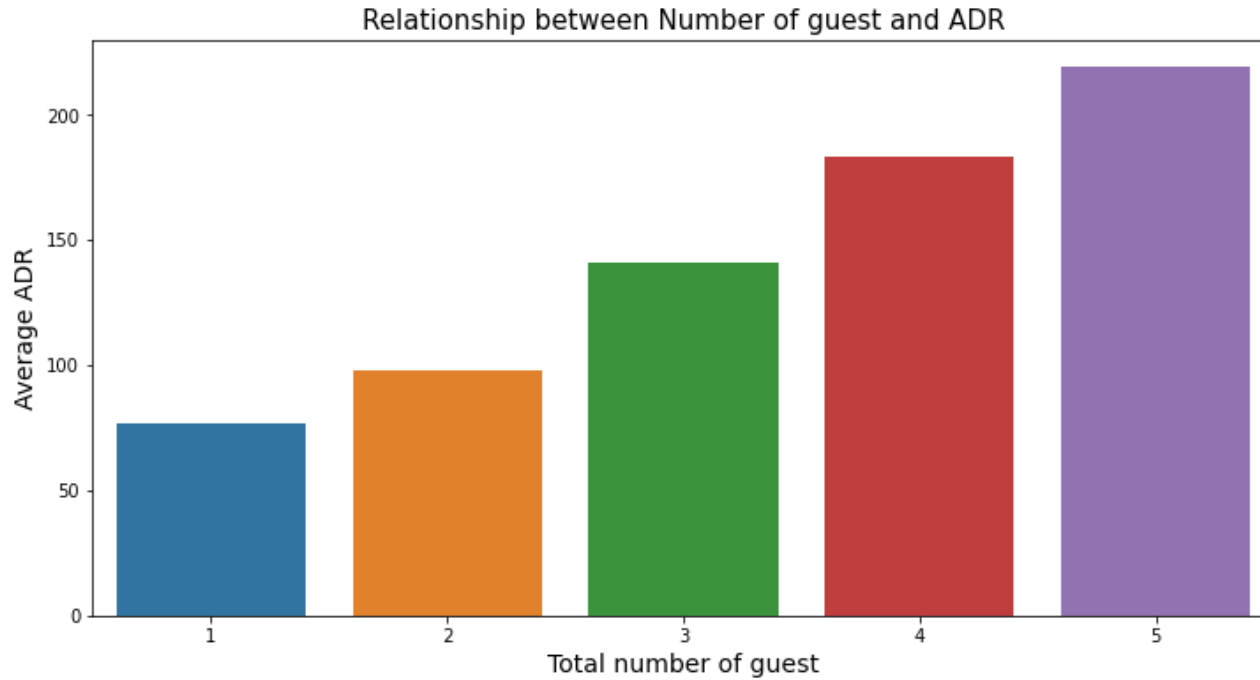
Observation:
•As the total number of people increase adr also increases. Thus adr and Total people are proportional to each other.

EDA & Observations (Cont.)

Relationship between Total stay and ADR

Observation:
• As the total stay increases adr is decreasing. Thus, for longer stays customer can get good adr(Price).

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!