



## Team Creation Presents

### Capstone Project EDA ON Hotel Booking Analysis

- Bipasha Zade
- Chinmay Rojatkar (Team Leader)
- Deepali Mahajan
- Kunal Gawande
- Nikhil Aggarwal



## Points to Discuss:-

While doing the complete analysis the Team has created the pointers in the form of the Index which can be find below:-

- Agenda
- Data Conclusion
- Bivariate Analysis
- Hotel wise Analysis
- Cancelation Analysis
- Time wise Analysis



# Data Collection and Understanding ÷

## **Data Description:**

- hotel :Resort Hotel or City Hotel
- is\_canceled : Value indicating if the booking was canceled (1) or not (0)
- lead\_time : Number of days that elapsed between the entering date of the booking and the arrival date
- 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
- stays\_in\_weekend\_nights : Number of weekend nights
- stays\_in\_week\_nights : Number of week nights.
- adults : Number of adults
- children : Number of children
- babies : Number of babies
- meal : Type of meal booked.
- country : Country of origin.

- market\_segment : Market segment designation. (TA/TO)
- distribution\_channel : Booking distribution channel.(T/A/TO)
- is\_repeated\_guest : is a repeated guest (1) or not (0)
- previous\_cancellations : Number of previous bookings that were cancelled by the customer prior to the current booking
- previous\_bookings\_not\_canceled : Number of previous bookings not cancelled by the customer prior to the current booking
- reserved\_room\_type : Code of room type reserved.
- assigned\_room\_type : Code for the type of room assigned to the booking.
- booking\_changes : Number of changes made to the booking from the moment the booking was entered on the PMS until the moment of check-in or cancellation
- deposit\_type : No Deposit, Non Refund , Refundable. agent : ID of the travel agency that made the booking company : ID of the company/entity that made the booking .
- days\_in\_waiting\_list : Number of days the booking was in the waiting list before it was confirmed to the customer
- customer\_type : type of customer. Contract,Group,transient,Transient party.
- adr : Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights
- required\_car\_parking\_spaces : Number of car parking spaces required by the customer
- total\_of\_special\_requests : Number of special requests made by the customer (e.g. twin bed or high floor)
- reservation\_status : Reservation last status.

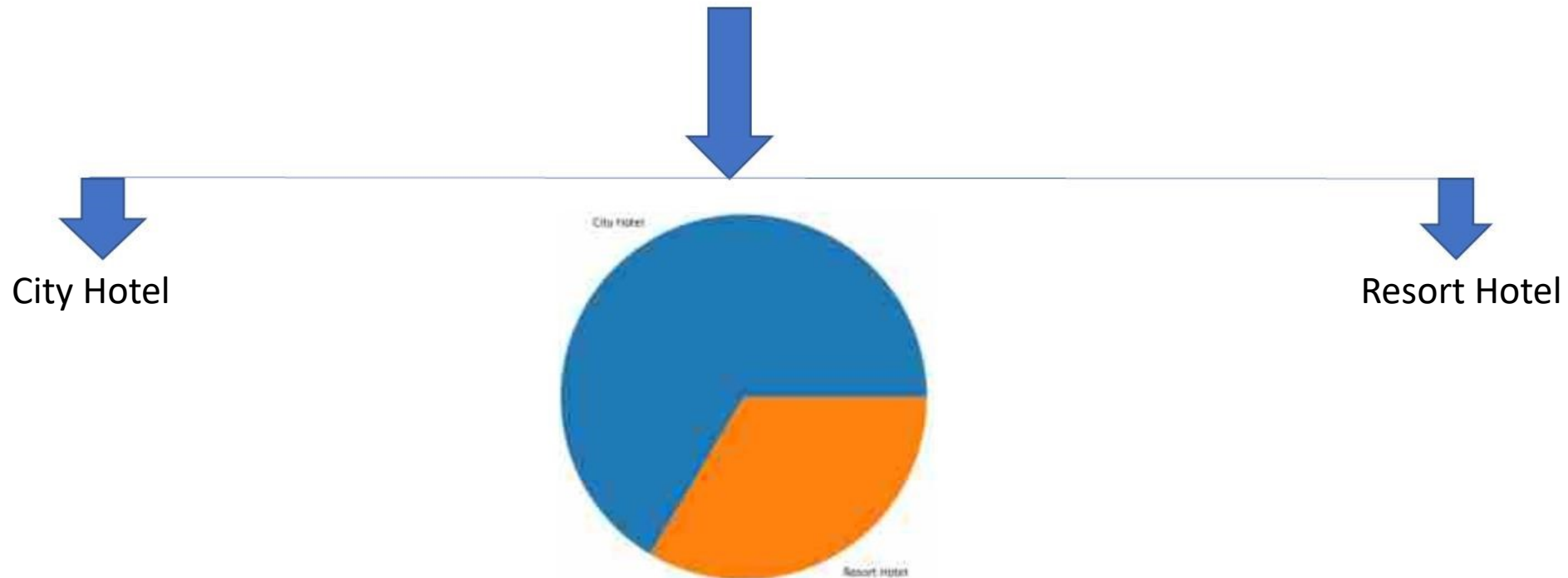
# Agenda:-

- The Main center of discussion is the analysis of given hotel bookings data set from 2015-2017.
- The Team Creation has done the analysis of given data set in following ways:-
  1. Bivariate Analysis
  2. Hotel wise analysis
  3. Distribution wise Analysis.
  4. Cancellation Analysis.
  5. Time Wise Analysis.

# Data Conclusion:-

Provided data set has different columns of variables crucial for hotel bookings. Few of them are as below:

- Hotel: The category of Hotels. There are 2 Categories of the Hotel:-



# Data Conclusion(Continued):-

Cancelled:- The value of this column shows the cancellation type. Either the booking were cancelled or not.

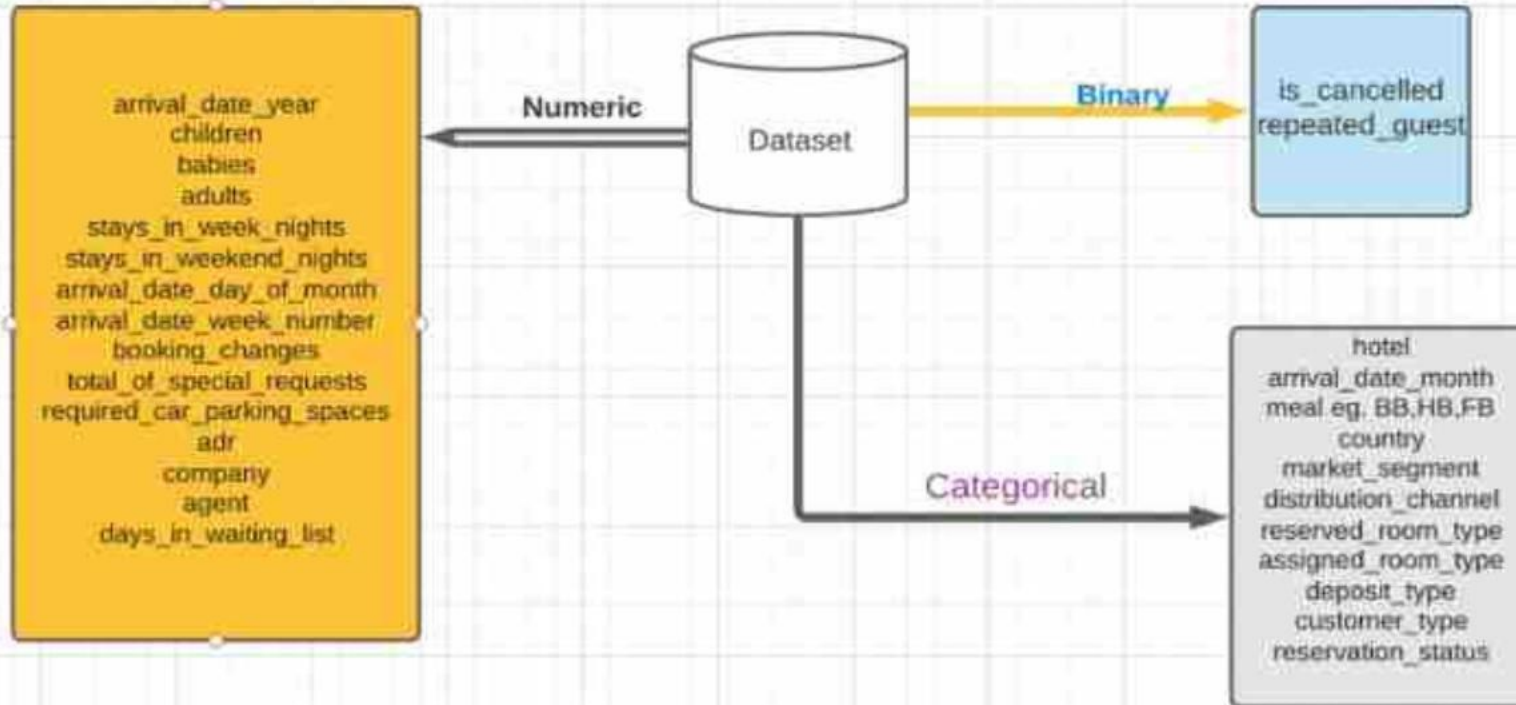
Where the values have been indicated as:-  
Values[0,1], where 0 indicates not cancelled.  
1 Indicates as Cancelled

# Data Conclusion(Continued):-

- Customer Types:- This Includes either the Customer is an Adult, Children or the Toddler(Babies).

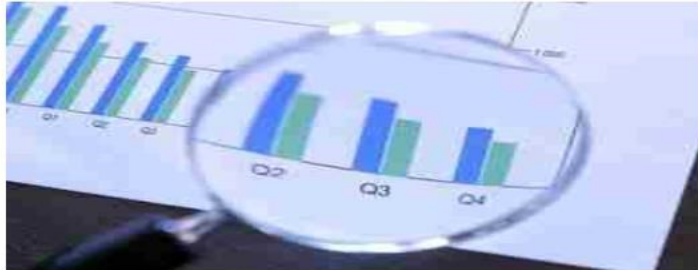


# Data Conclusion

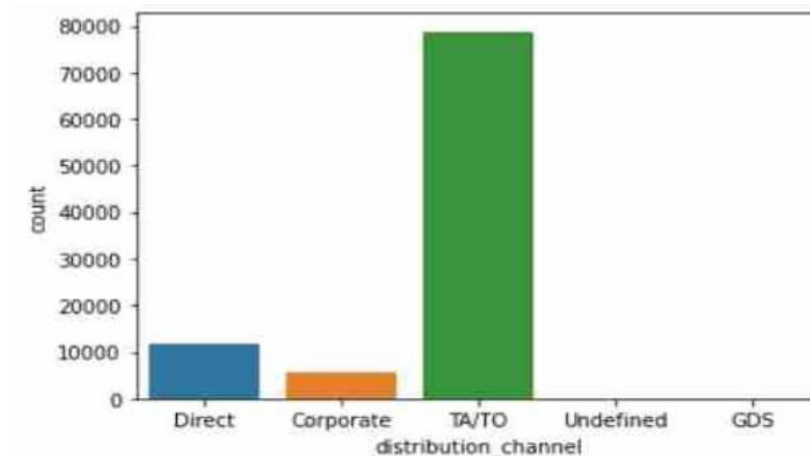
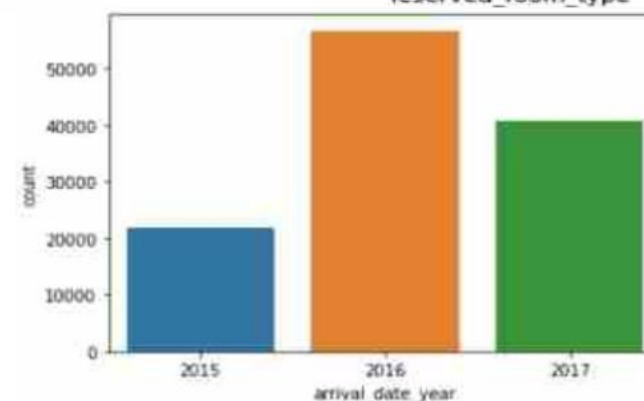
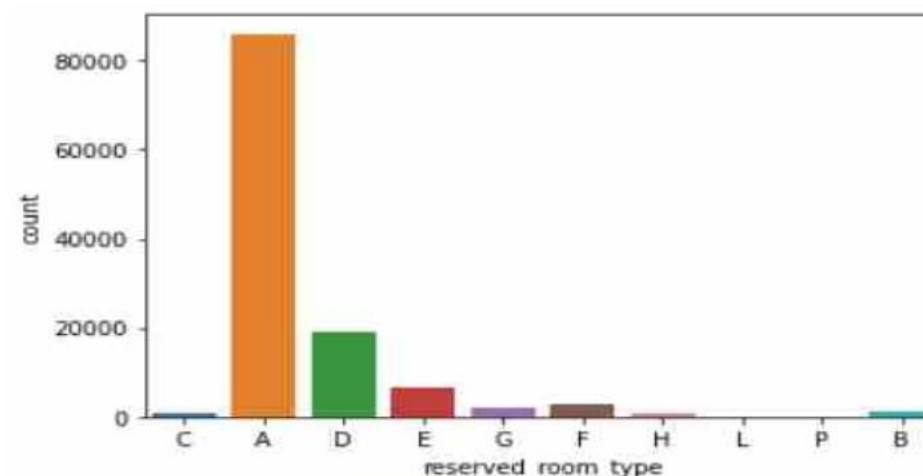


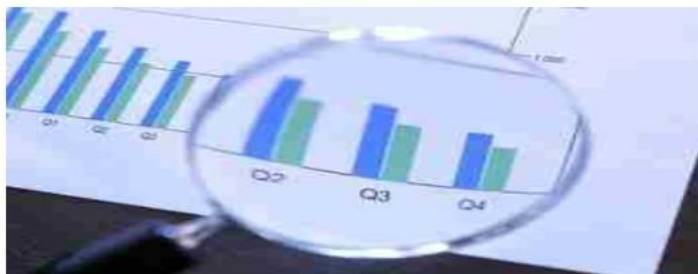
# Bivariate Analysis:-

- While doing Bivariate analysis of the given hotel booking dataset, we tried to figure out answering the following questions:
  1. Which room type is in the demand?
  2. Which room type generated highest Revenue?
  3. In Which Year the Highest Number of Rooms booked by the each type of the Hotel?
  4. Maximum Number of the Rooms booked by ?
  5. In which Month the Highest Number of Customers came?

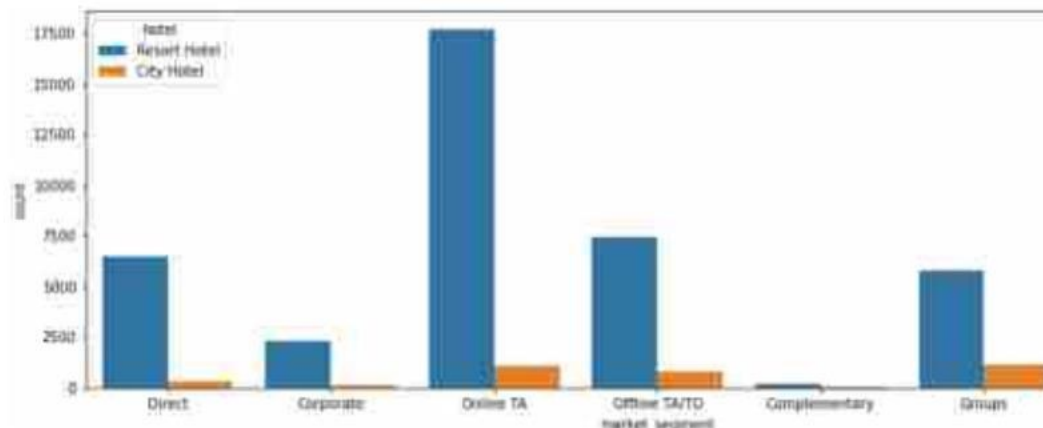


- Type of the Room **A** is in most demand by the Customers.
- Type of the Rooms D, E and F are some of the highest adr(average daily rate) generating rooms.
- Type of Room A is generating more profit to the Hotel.
- In the year 2016 most number of the Hotel Rooms were booked as in comparison of the year 2015 and 2017
- The TA/TO (Travel Agents/Travel Offices) were able to book the Highest number of the Hotel Rooms.

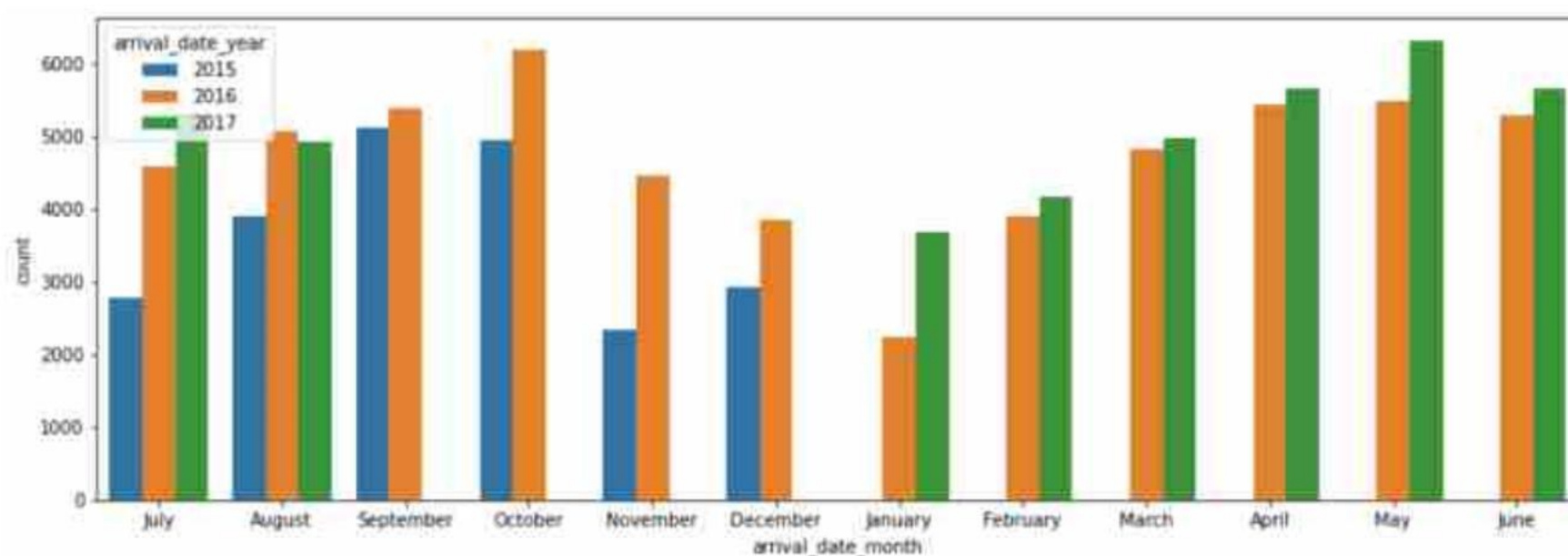




- Rooms booked by the Online Travel Agents/offices(Online TA/TO) are having Highest Number of Bookings in comparison of the offline Travel Agent/Travel Offices(Offline TA/TO).

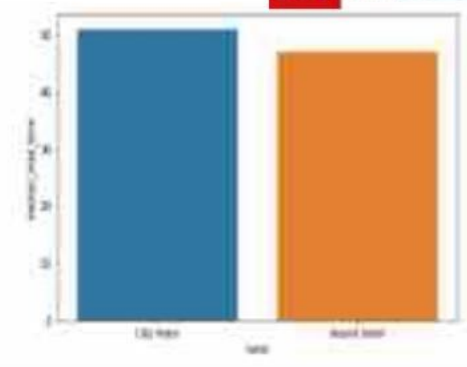
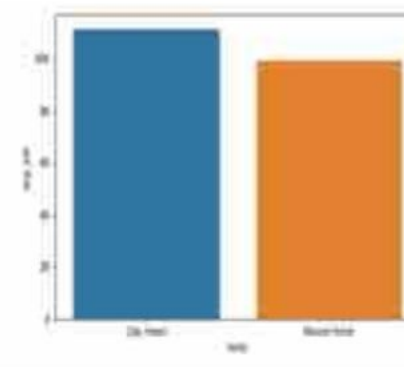
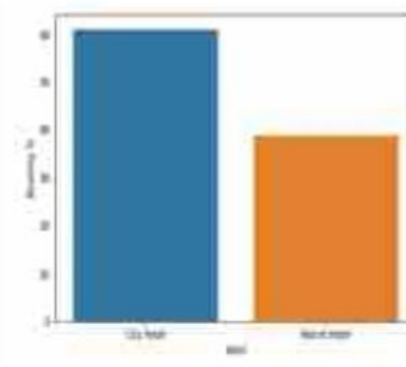
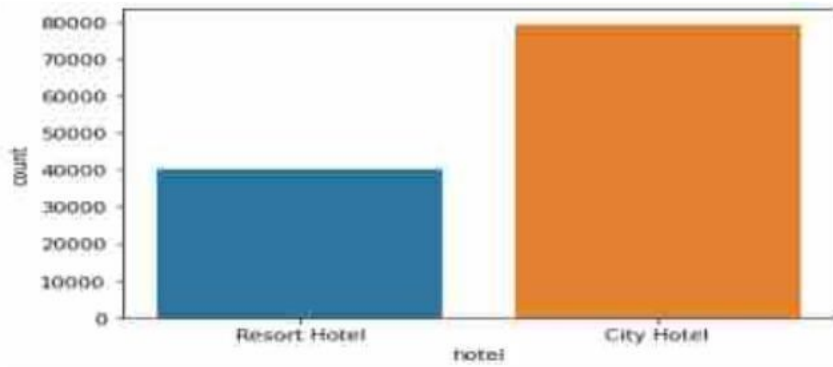


- The Customers generally likes to come in the month of October and May

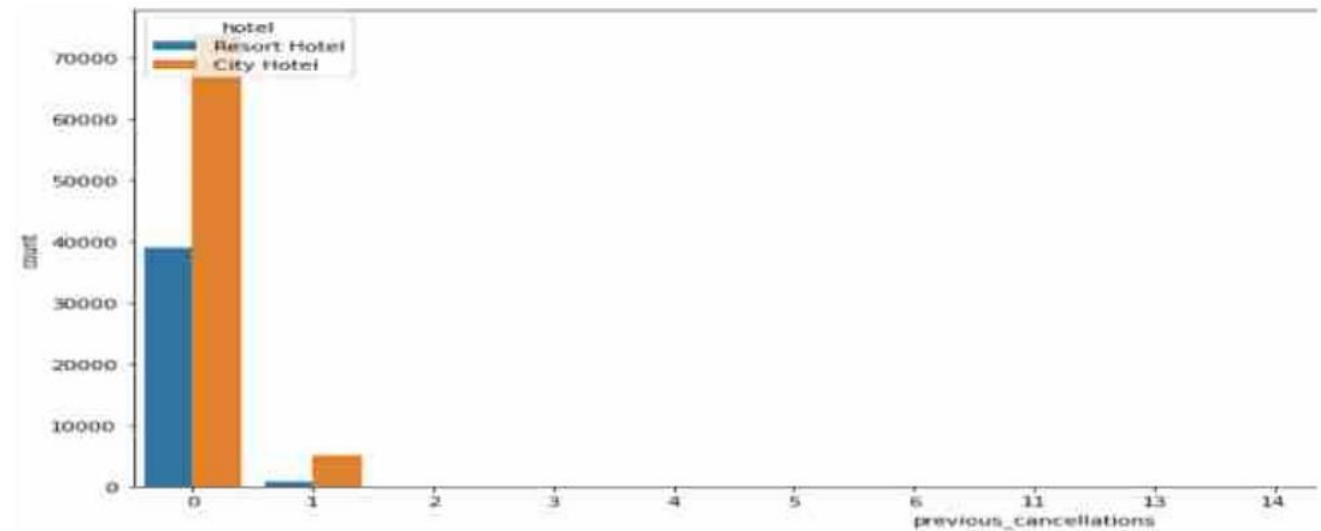


# Hotel Wise Analysis:-

- While doing the Hotel Wise Analysis of the given hotel booking dataset, we tried to figure out answering the following questions:
  1. Percentage of the bookings in each hotel?
  2. Which hotel makes more revenue?
  3. Which hotel has higher booking cancellations rates?
  4. What is most preferred stay length in each hotel?
  5. What is the preferred stay time by the Customers?

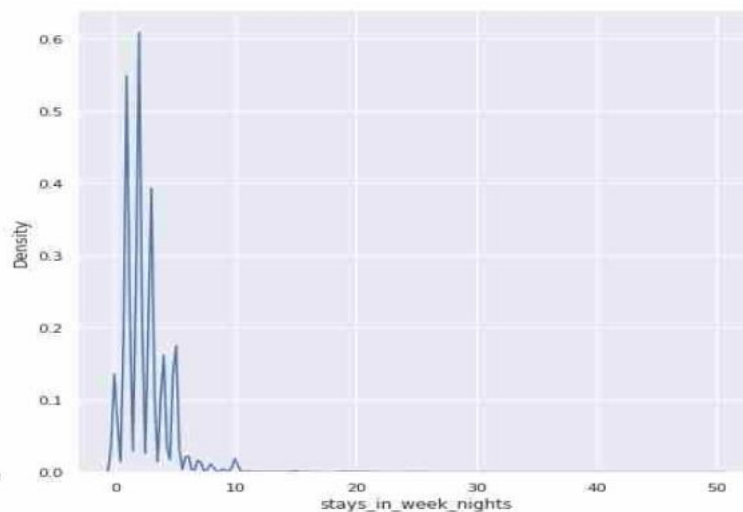
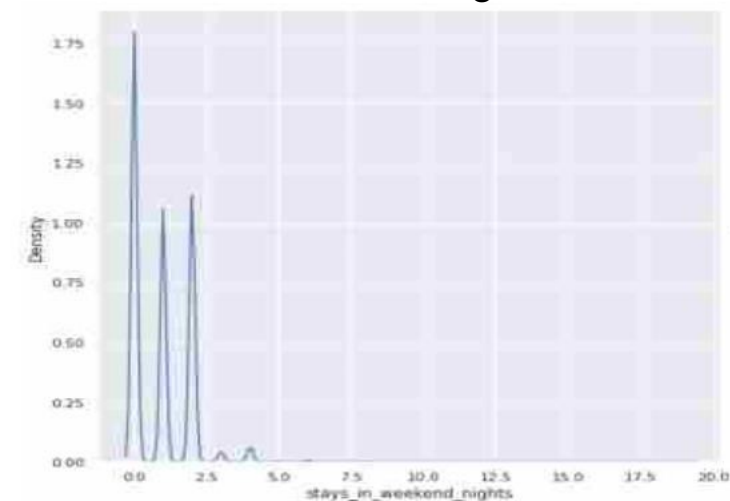
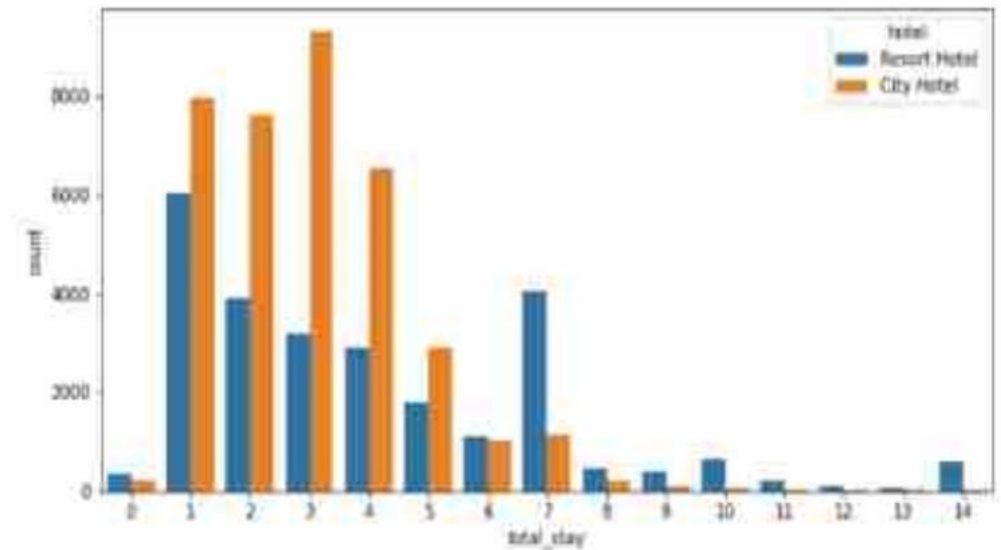
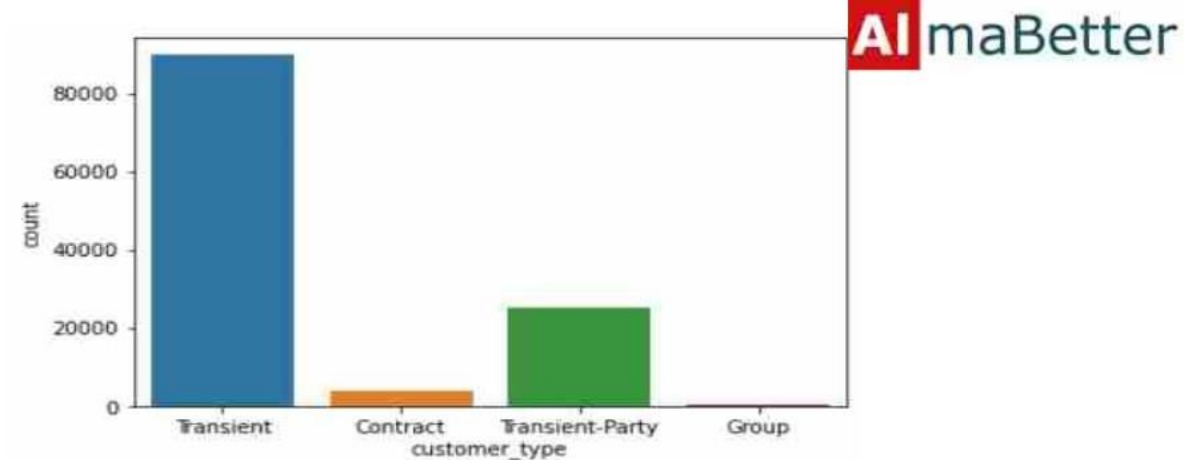


- Nearly around 60% bookings are for the City hotel and 40% bookings are for the Resort hotel.
- The City hotel has significantly higher bookings , hence City Hotel is much busier than Resort Hotel.
- The City Hotel is generating more Revenue than the Resort type Hotel by looking at the trend of booking analysis
- Approximately 30 % of the City Hotel bookings and 25 % of the Resort hotel bookings got canceled.





- There are many Transient(staying in a place for a short period of time) Customers stayed In the Hotels in comparison of other guest categories.
- As per the Trends The Majority of the stays are less than for the 5 days. There are very few long stays at hotels but Resort Hotel is preferred for long stays.
- The Customers preferers to Stay at weekend nights



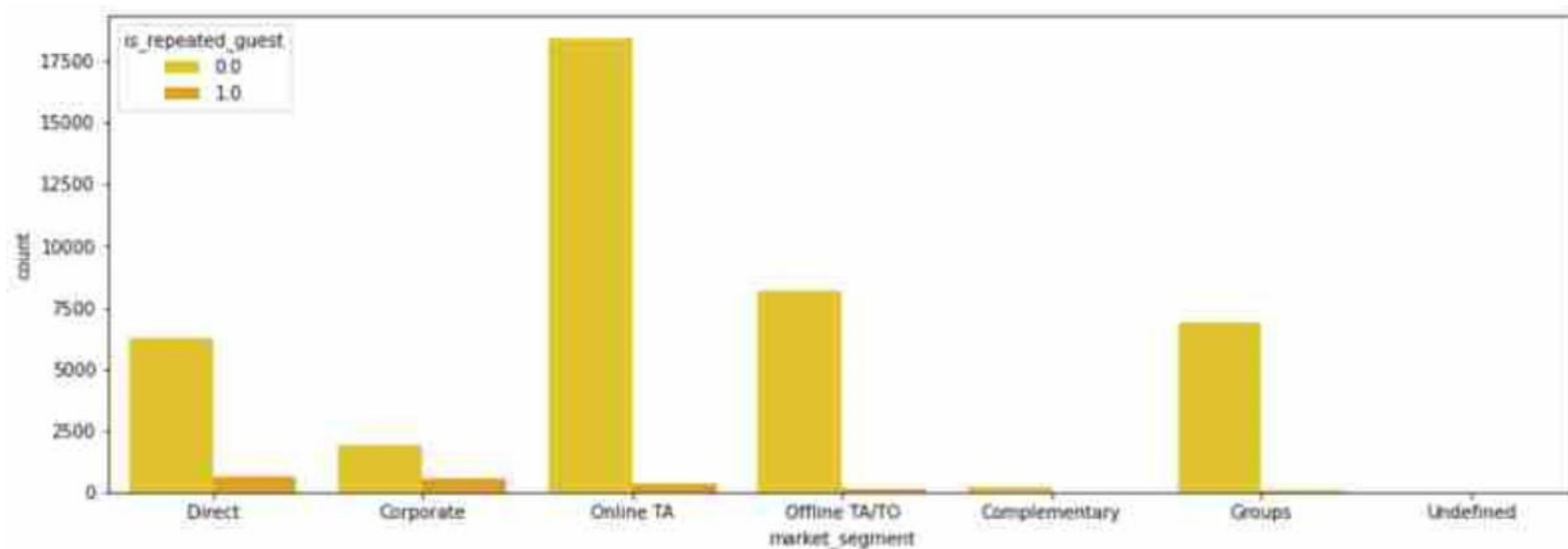
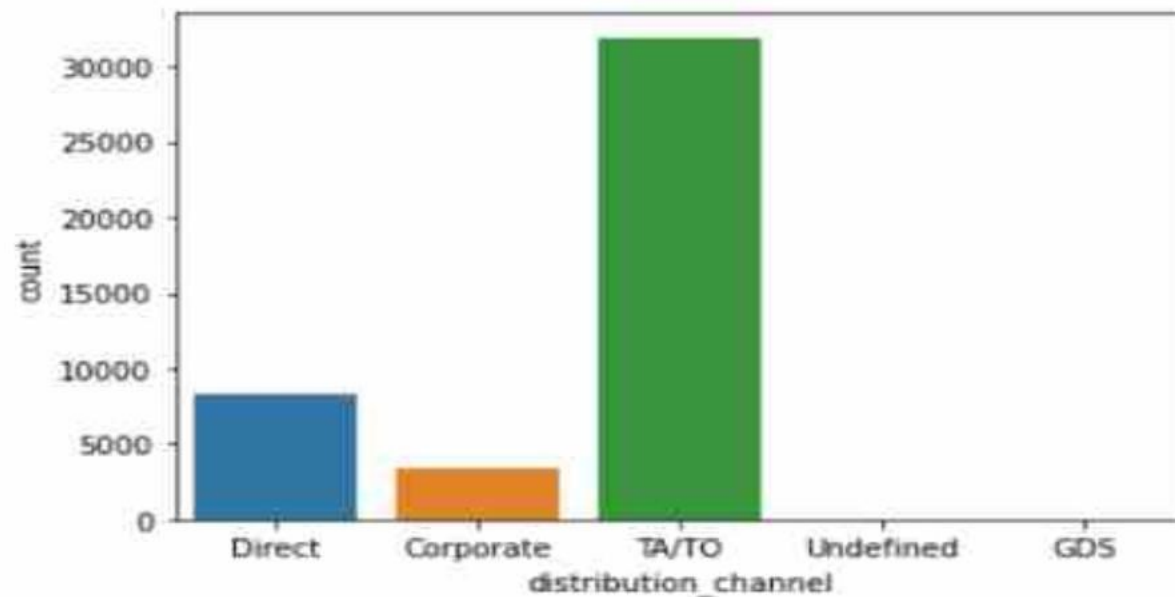
# Distribution Wise Analysis:-

While doing this set of the analysis of provided hotel booking dataset, the Team Tried to find the answer of the following questions:

1. What is the most common channel for making the booking of the hotels?
2. Which distribution channel generates more revenue deals for the hotels?
3. Which channel does have more repeated Customers?



- As per the provided data set we can clearly see that the most number of the guest are doing the reservations through TA(Travel Agency)/TO(Tour Operator).
- The Second most booking Channel is the Direct booking done by the Customers at the Time of arrival.
- Hence, The most used Channel by the Customers are the TA(Travel Agency)/TO(Tour Operator) and the TA/TO Channels are generating the more Revenues to the Hotels.
- The online and offline TA(Travel Agency)/TO(Tour Operator) are having more numbers of the repeated Customers.



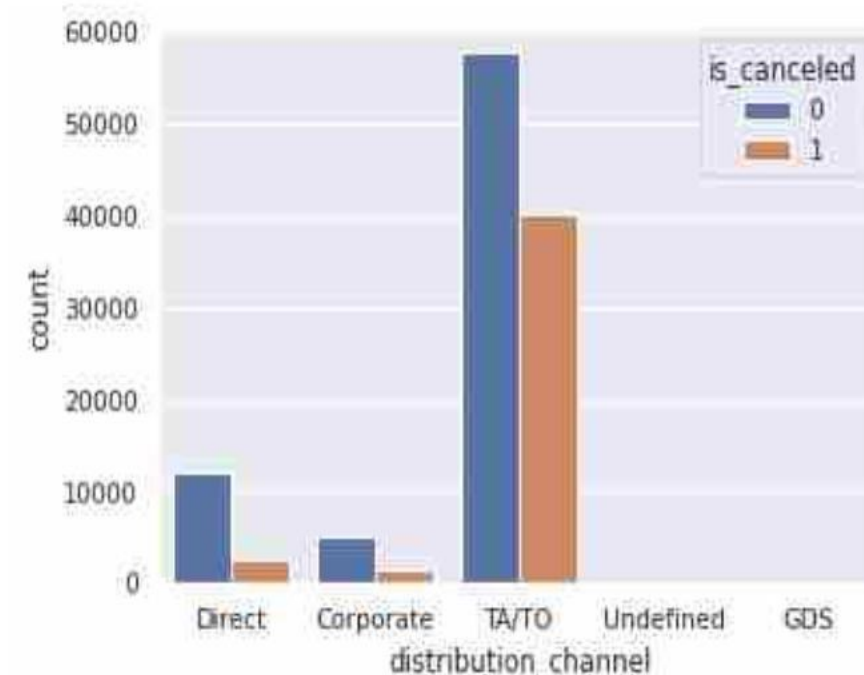
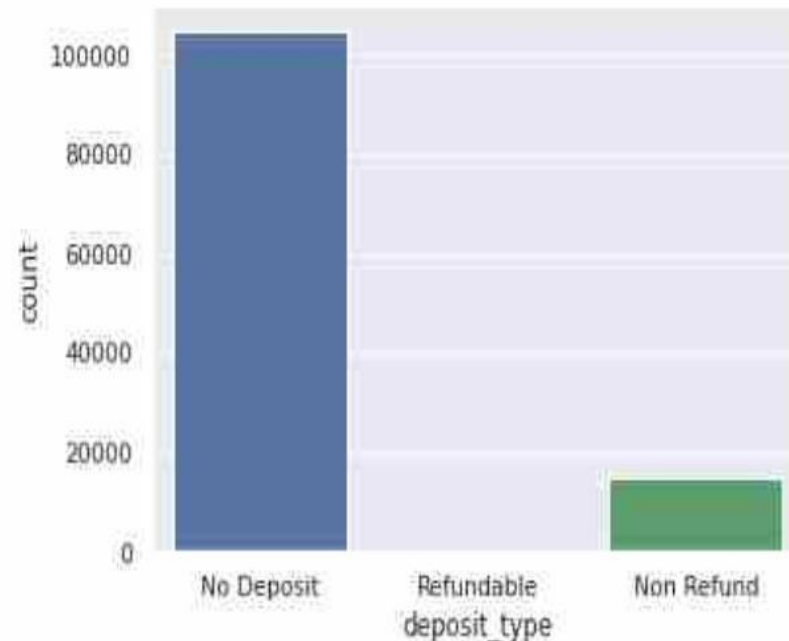
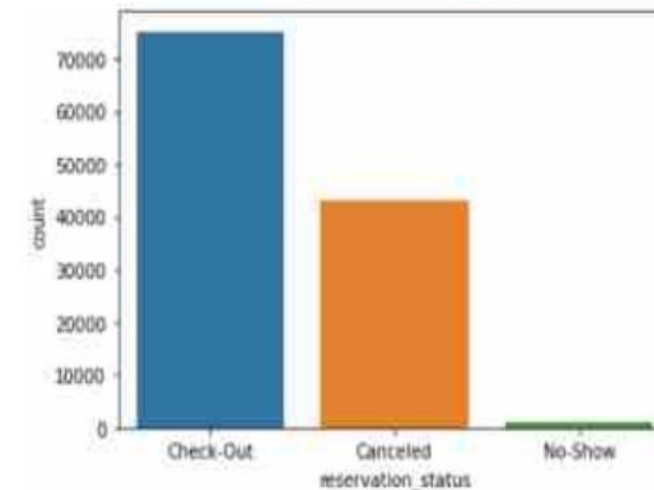
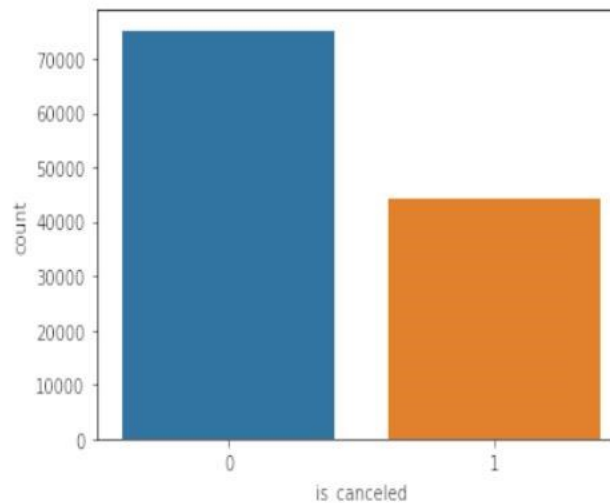
# Cancellation Analysis:-

While doing the Cancellation Analysis The team figured to do the analysis and tried to answer the following:

1. Which distribution channel has highest the cancellation percentage?
2. Reasons behind the Cancellation Occurs?
3. Which channels is affecting the most revenue of the Hotels?



- While doing the Analysis we were able to figure out that TA/TO has highest booking cancellation percentage. Therefore, a booking via TA/TO is almost 28% likely to get cancelled.
- There can be n reasons to get the Hotel rooms to be cancelled, but while looking at the Trend many cancelation occurs during check in and check outs the majority reason can be that they are not getting the same rooms which they have booked.
- Thus by looking at the Trend the TA/TO are affecting the businesses more.
- The Most number of the Cancellation are done by not paying any deposits.



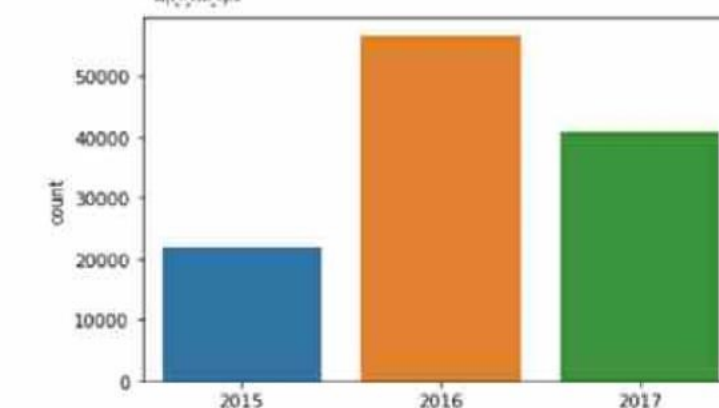
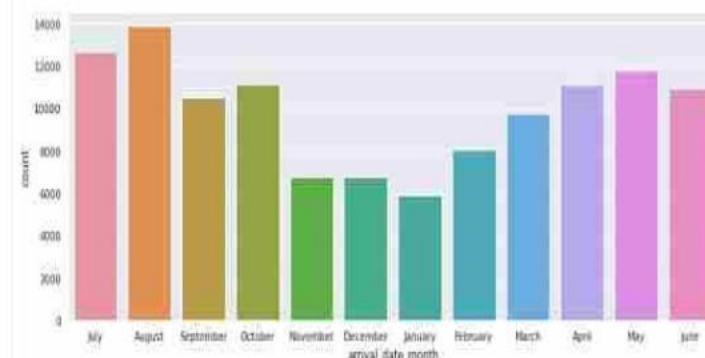
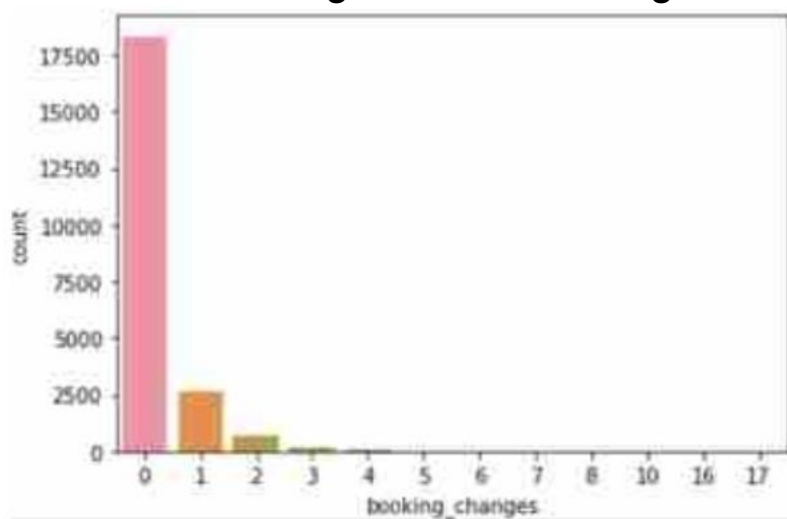
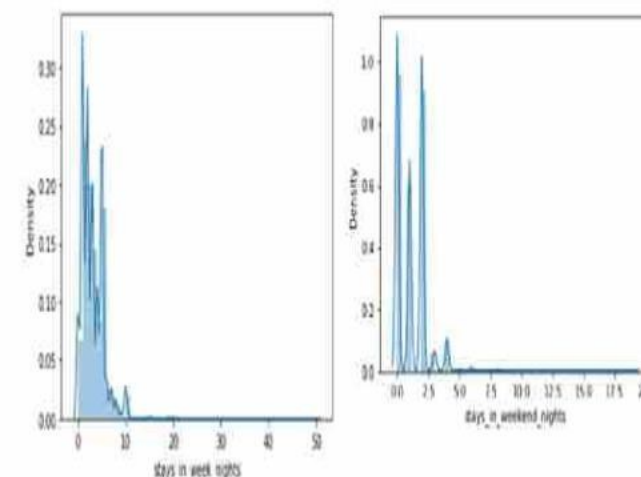
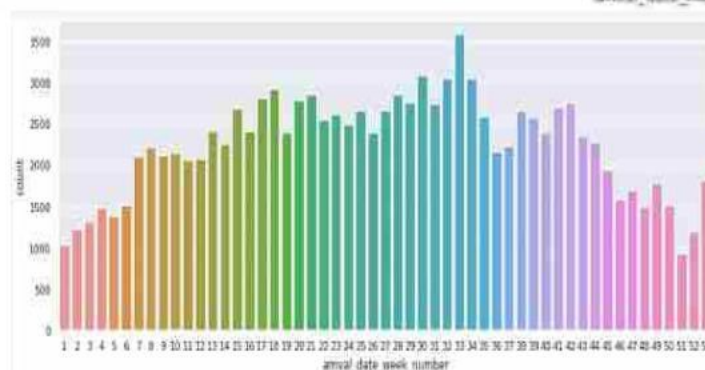
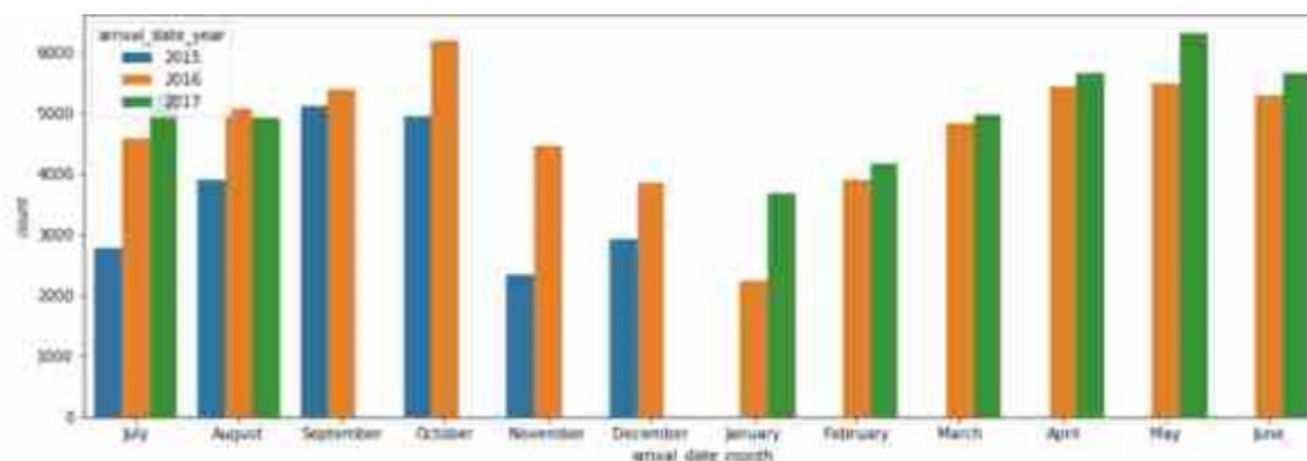
# Time Wise Analysis:-

While doing the Time Wise Analysis The team figured to do the analysis and tried to answer the following:

1. Which Months are the most busiest months for the hotels?
2. In which year the Highest Number of Bookings were Done?
3. What are the busiest days of the months in staying the Hotels?
4. Trends of booking Changes and Months of Arrivals?



- The Year 2016 most number of the Hot Rooms were booked as in comparison o the year 2015 and 2017
- The Customers generally likes to came in the month of October and May
- The Customers preferers to Stay at weekend nights
- There are few customers who prefers to make changes to the bookings



# Summarization of Complete Theory:-

- Around 60% bookings are for City hotel and 40% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel. Also the overall average daily rate and revenue of the City hotel is slightly higher than the Resort hotel.
- Both hotels have significantly higher booking cancellation rates and very few guests less than 3 % return for another booking in City hotel. 5% guests return for stay in Resort hotel.
- Customers used different channels for making bookings out of which most preferred way is TA/TO.
- July- August are the most busier and profitable months for the both of the hotels.
- Nearly 28% of bookings via TA/TO are cancelled.
- In the Year 2016 most Guest/Customers came to the Hotels.
- Customers preferred to stay in weekends rather than weekdays.
- Booking made by the TA/TO are Resulting the more repetition of the Customers.
- The Second most booking Channel is the Direct booking done by the Customers at the Time of arrival.
- City Hotels are more liked and booked by the Customers.



**Q/A**

**Thank You**