Capstone Project - 1

EDA on Hotel Booking Analysis

By- Aswini Kumar Jena

Topics of Discussion:-

- Problem Statement & Objective of analysis
- Data information & summary
- Data Preparation processes
- Analysis based on types of Hotels
- Market Segments, Distribution channels, Customer Types
- Popular room types for guests
- Analysis on Stay length and Total people during a stay
- Effect of lead time , advance deposits on bookings
- Year wise , month wise and day wise bookings analysis
- Food of Hotels and special requests by guests
- Importance of repeated guests
- Conclusion and Solution for the client

- **Problem Statement –**
  - Explore and analyse the Hotel Bookings dataset to discover important factors that govern the bookings.

- **Objective of analysis –**
  - To provide the Hotel Business client with suggestion and solution by which they can increase the number of hotel bookings , decrease booking cancelations and therefore increase revenue and profits of their hotel business

Data information & summary -

- We have been provided with Hotel Booking Dataset for analysis which had 119390 rows & 32 features.

- Some of the main features of the dataset include:-
'lead_time', 'adr', 'total_of_special_requests', 'total_stay','total_people', 'arrival_date_year', 'arrival_date_month', 'market_segment' ,'distribution_channel' ,'is_repeated_guest', 'assigned_room_type', 'deposit_type', 'customer_type', 'reservation_status', 'adults', 'children' ,'babies'.

Data Preparation processes-

- Cleansing produces consistent, structured, accurate data, which allows for informed, intelligent decisions.

- While cleaning data we have performed following steps:-
  - Remove duplicate rows.
  - Handling missing values.
  - Adding important columns/ Creating important columns
  - Checking for outliers and imputation of outliers
  - Making the dataset EDA ready.

Analysis based on types of Hotels:-

- People are prefering city hotels (61.1%) more than resort hotels (38.9%).
- City Hotels Cancelation % = 30.1%, Resort Hotels Cancelation % = 23.5%
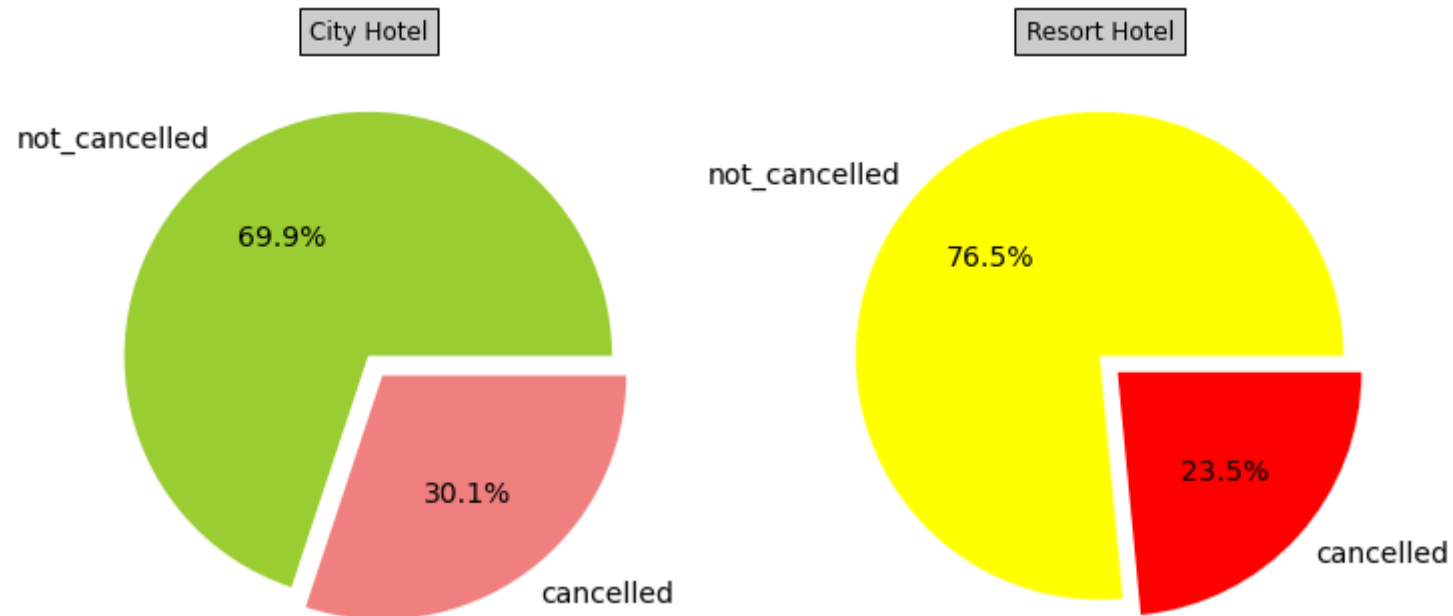
- More no of City hotels have no advance deposit rule than Resort Hotels.

- More Resort Hotels have non refundable deposit money rule than city hotels.

- City hotels are generating more mean revenue/adr than resort hotels.

Types of Deposit type

ADR by distribution channels of both hotels
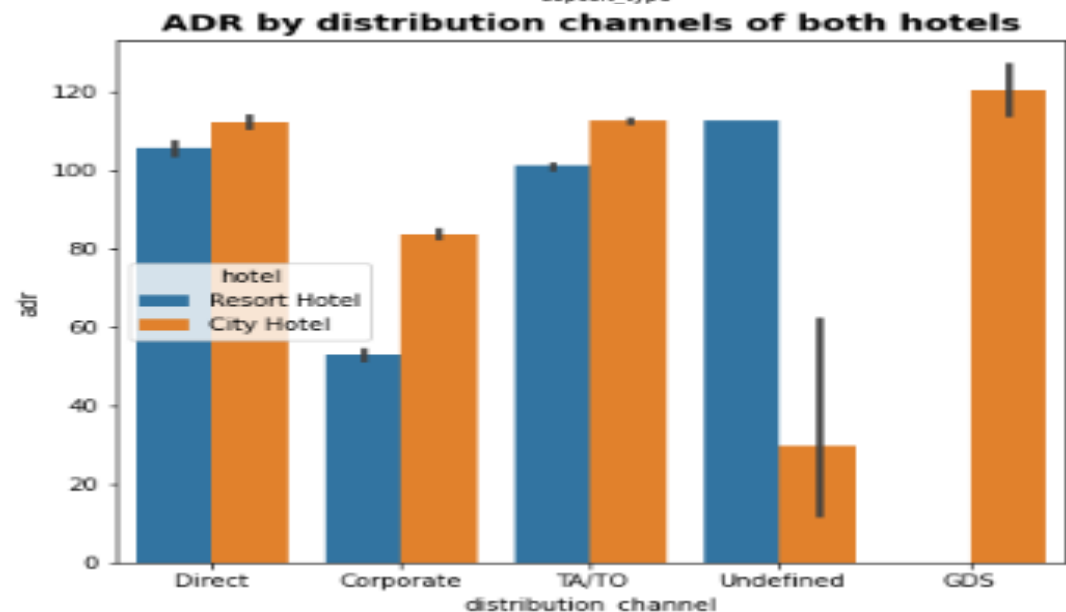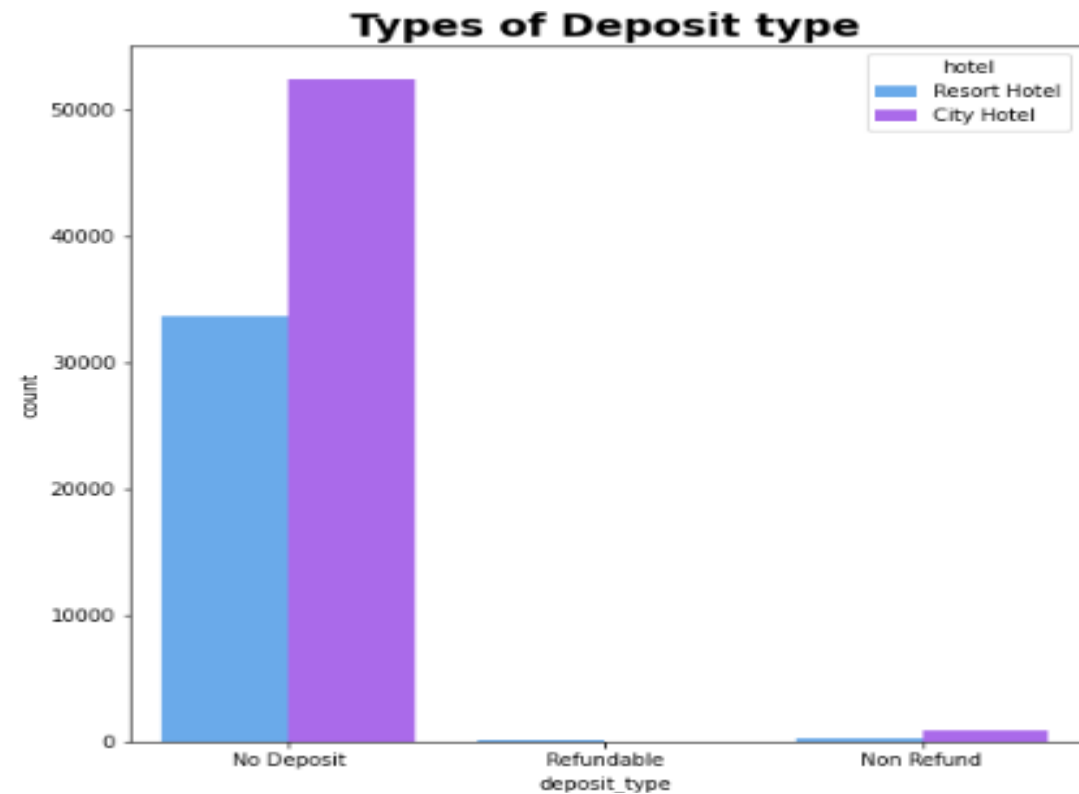
Market Segments, Distribution channels, Customer Types :-

- Hotels are getting its most customers from online travel agents followed by direct bookings and offline travel agents.

- 82% are Transient(guest for short time) customer types followed by Transient - Party(guest for short time but in groups ,maybe for some special occasions).
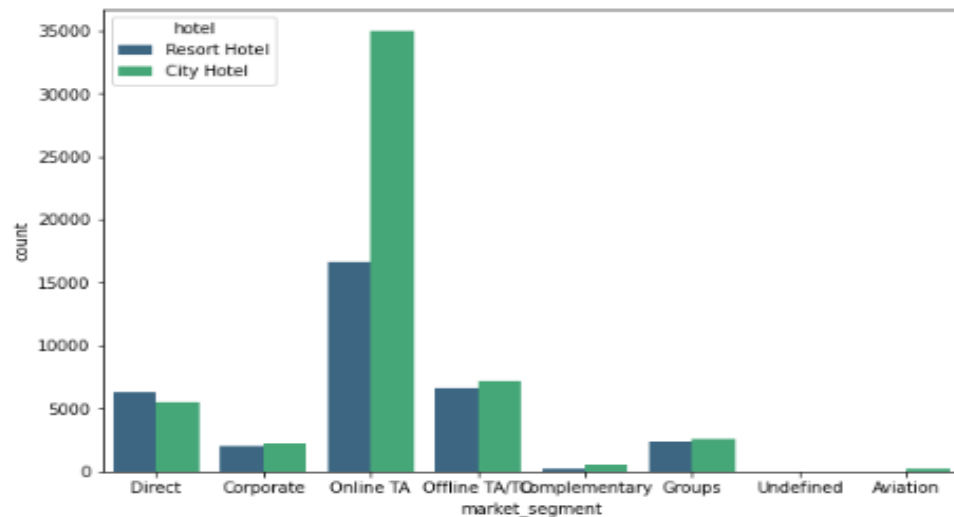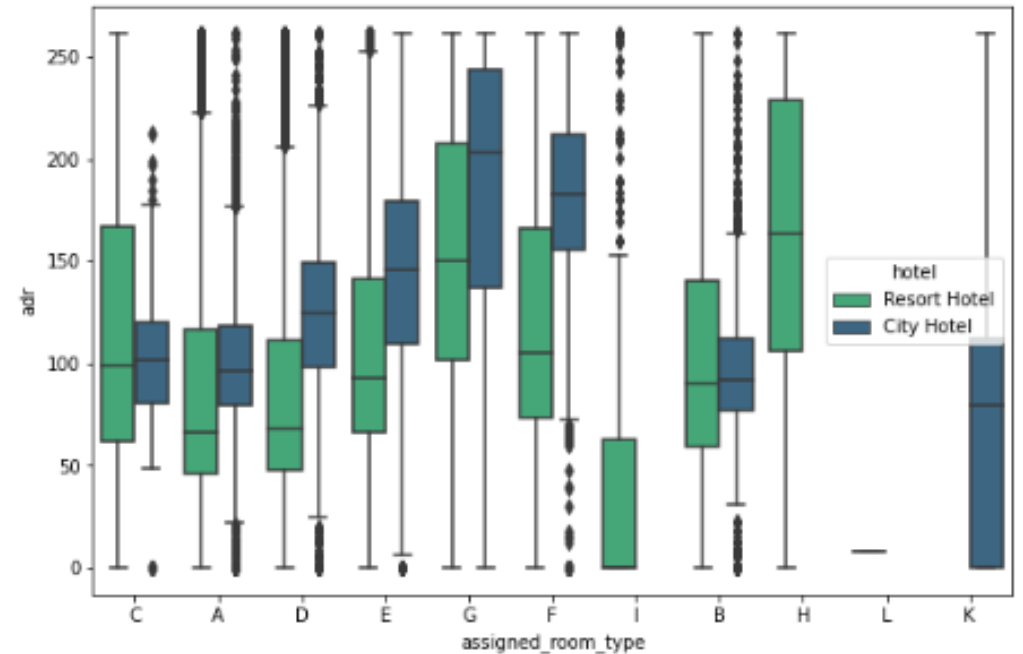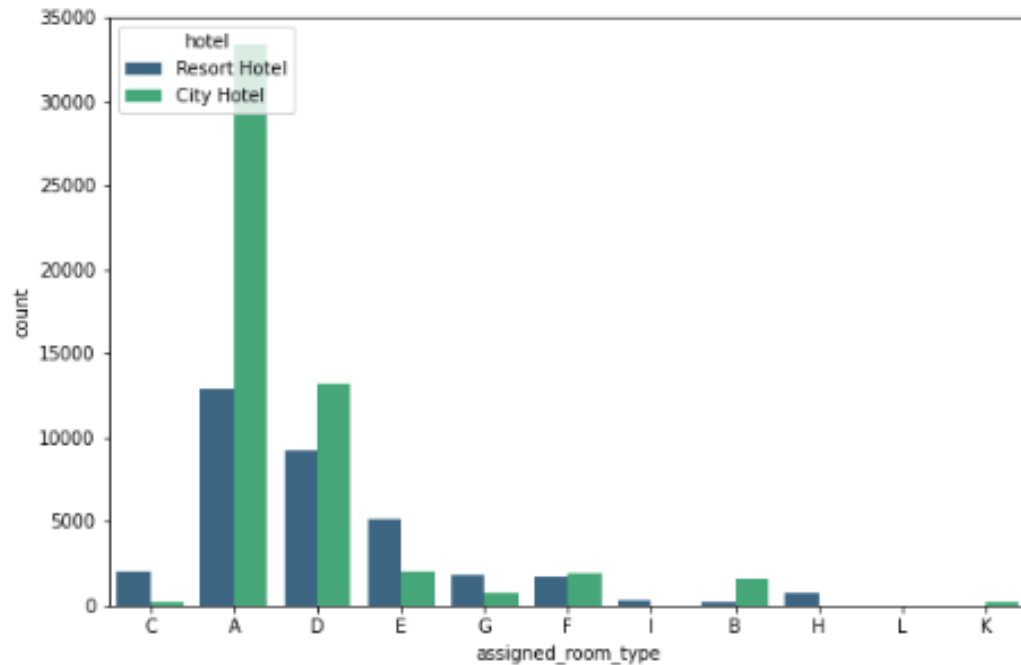

Contribution of customer type towards adr wrt type of hotels
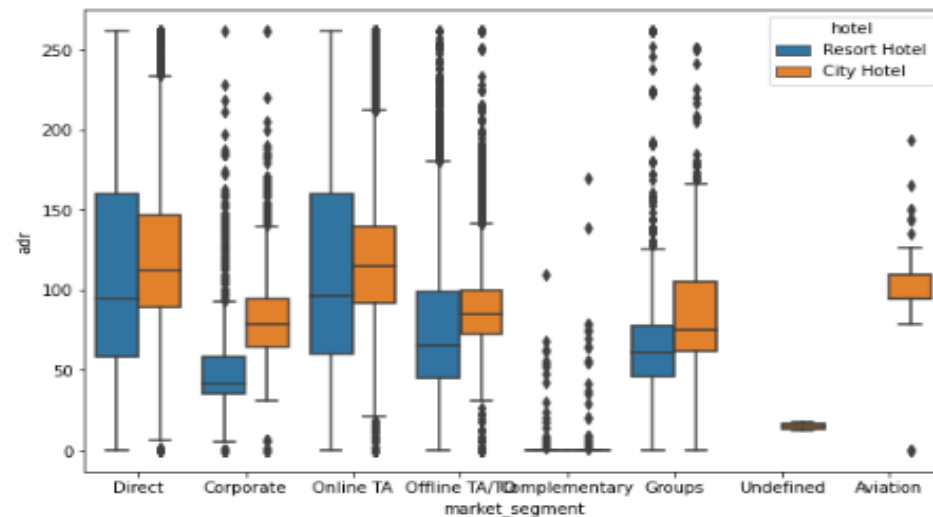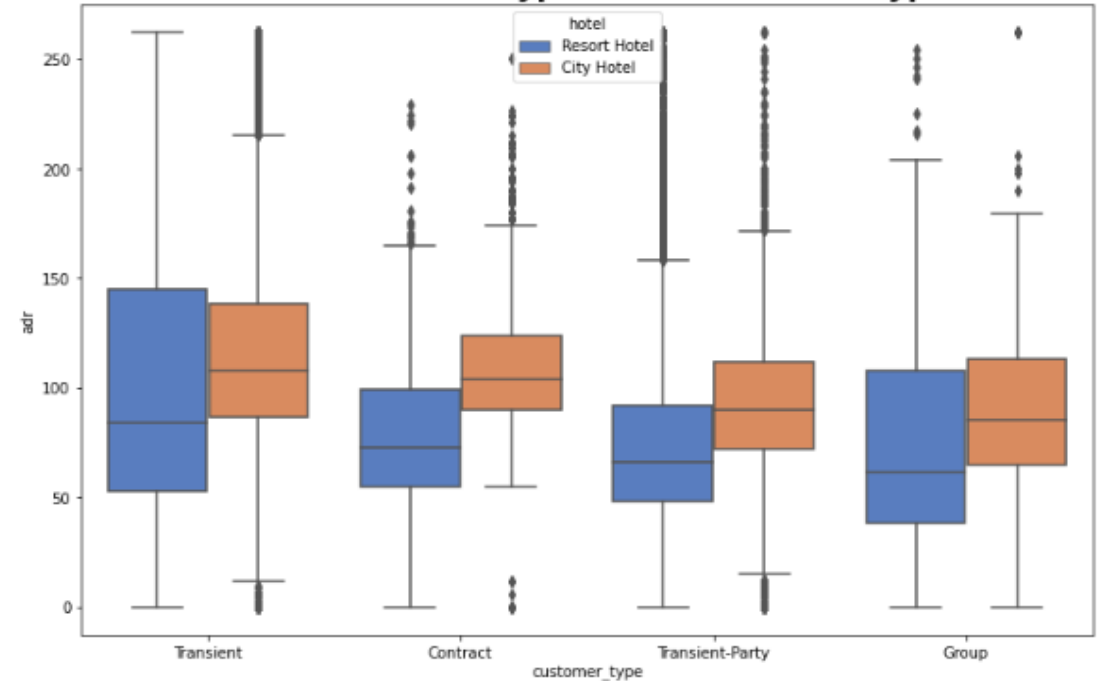
Popular room types for guests :-

- A, D, E are top 3 most popular type of rooms for both type of hotels.
- C, G, H are top 3 room type in case of adr for resort hotels.
- K, G, E are top 3 room type in case of adr for city hotels.
- Hotels should focus on increasing those type of rooms who generate more adr.
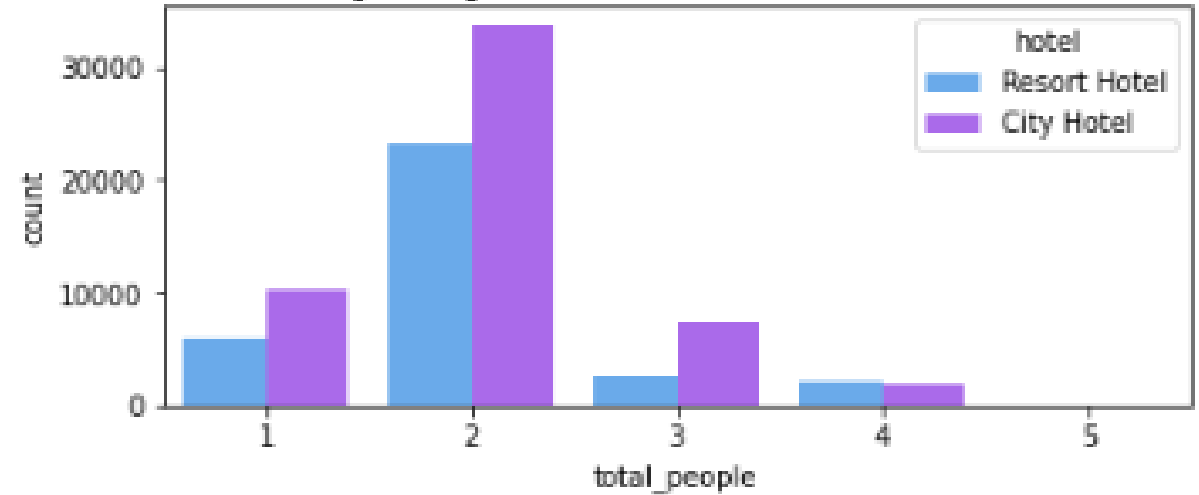
Analysis on Stay length and Total people during a stay :-

- When stay duration is less than 5 days people are preferring city hotels and when duration is more than 6 days people prefer resort hotels.
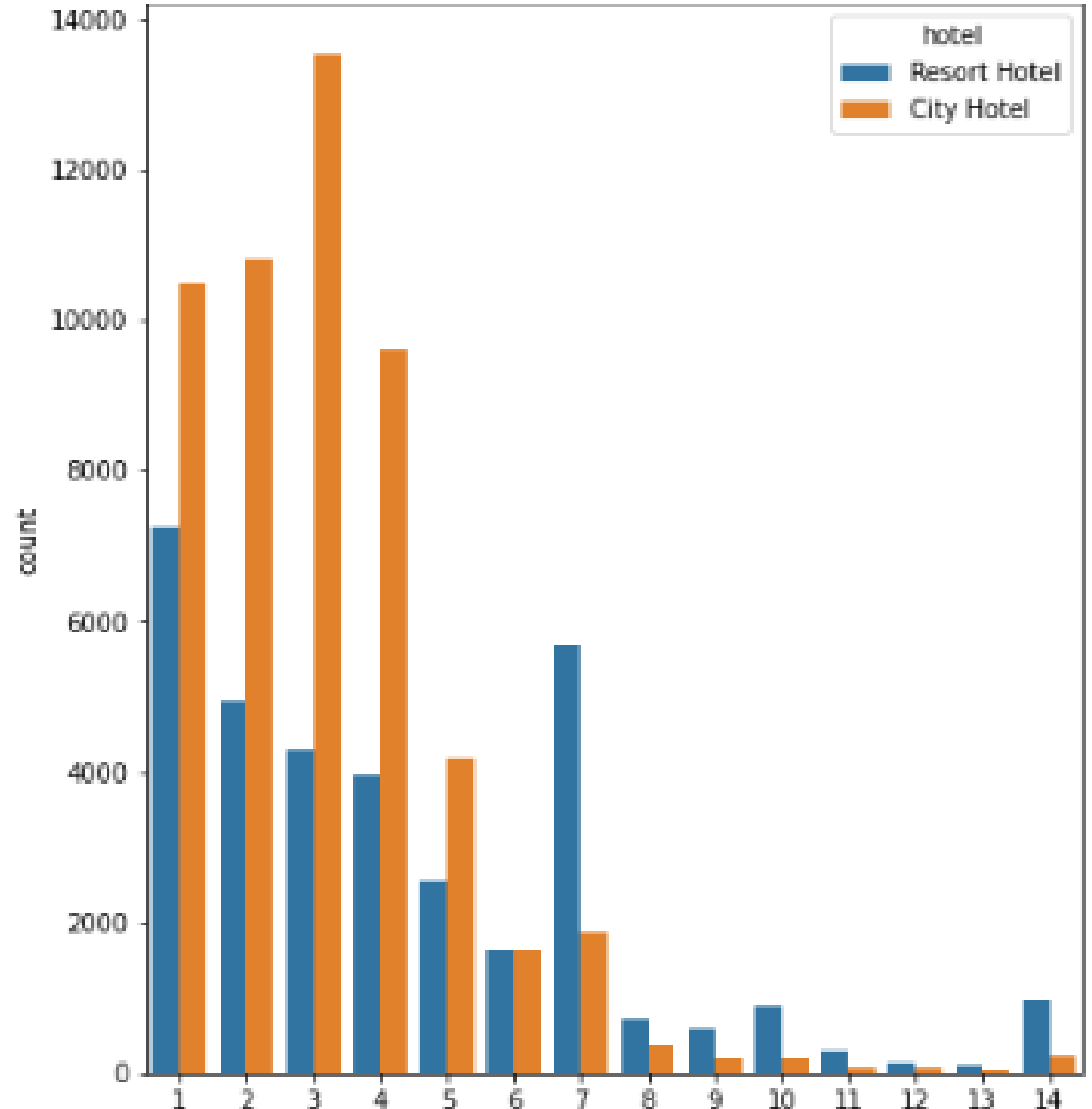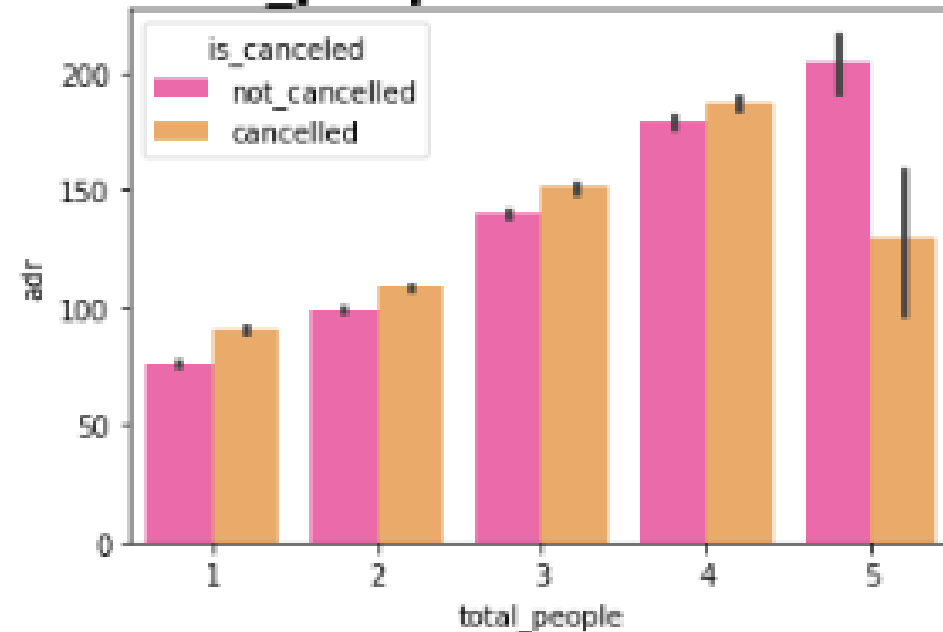
- Mostly when there are 2 people , they are preferring city hotels

- Bookings are not cancelled mostly in case of 1 or 2 people bookings
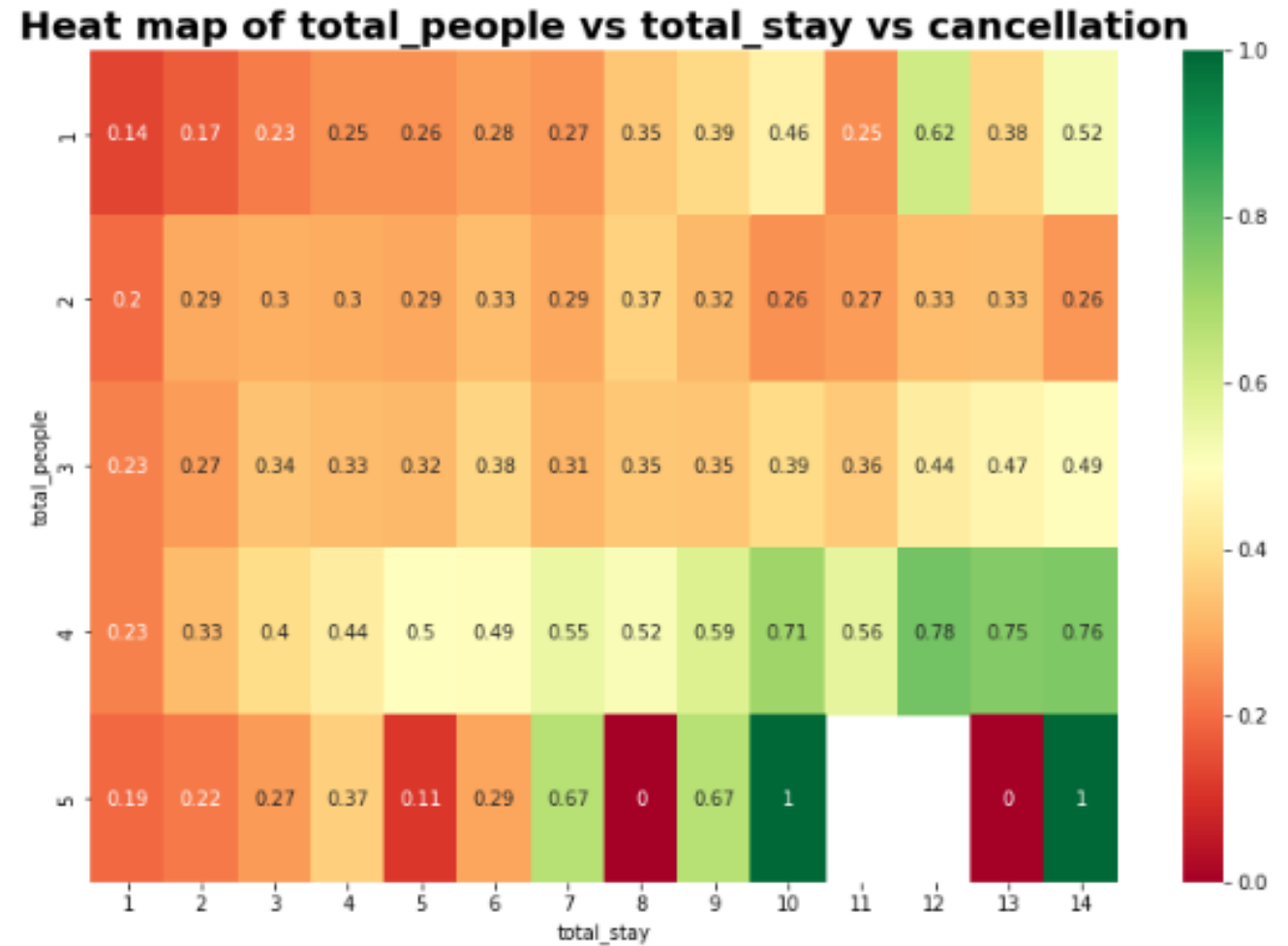

Total people wrt different hotels


Total_people vs Cancelation

- There is a higher chance of cancellation when the booking is for more than 4 people and stay is for more than 8 days
- Hotels should follow up with the customers with these type of bookings so that other genuine bookings are not blocked

Heat map of total_people vs total_stay vs cancellation

Effect of lead time , advance deposits on bookings:-

- Resort hotels get more advance bookings or bookings with higher lead time

- Bookings with avg lead time above 80 leads to cancellations

- Whenever the adr is under 190 and lead time is under 220 there are very rare chance of cancelations.
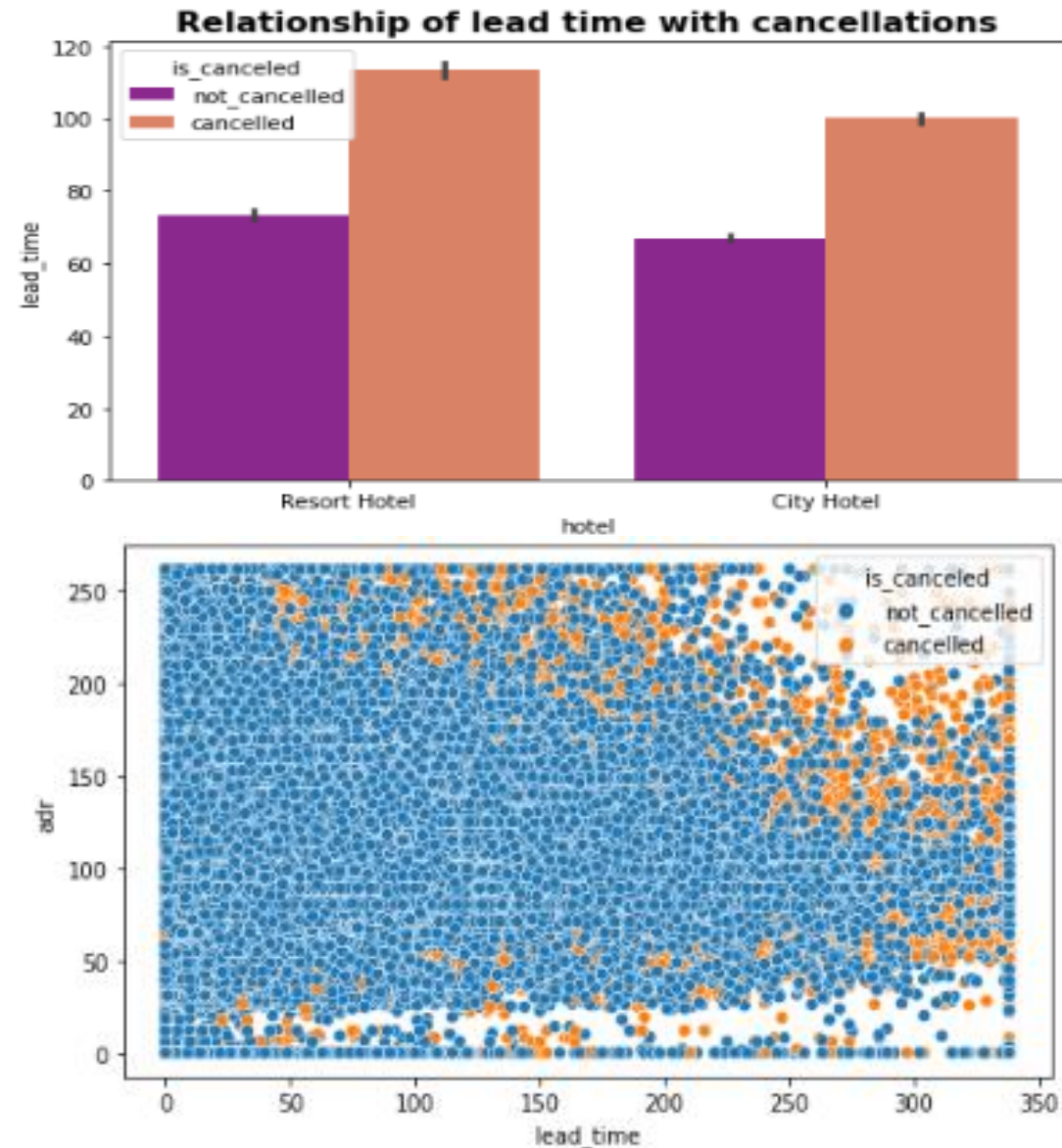
- Hotels with no deposit rule are getting nearly 90% of bookings.
- No deposit bookings are working well and giving good adr also, but it seems bookings with higher price are getting cancelled


ADR vs Deposit type vs is_canceled
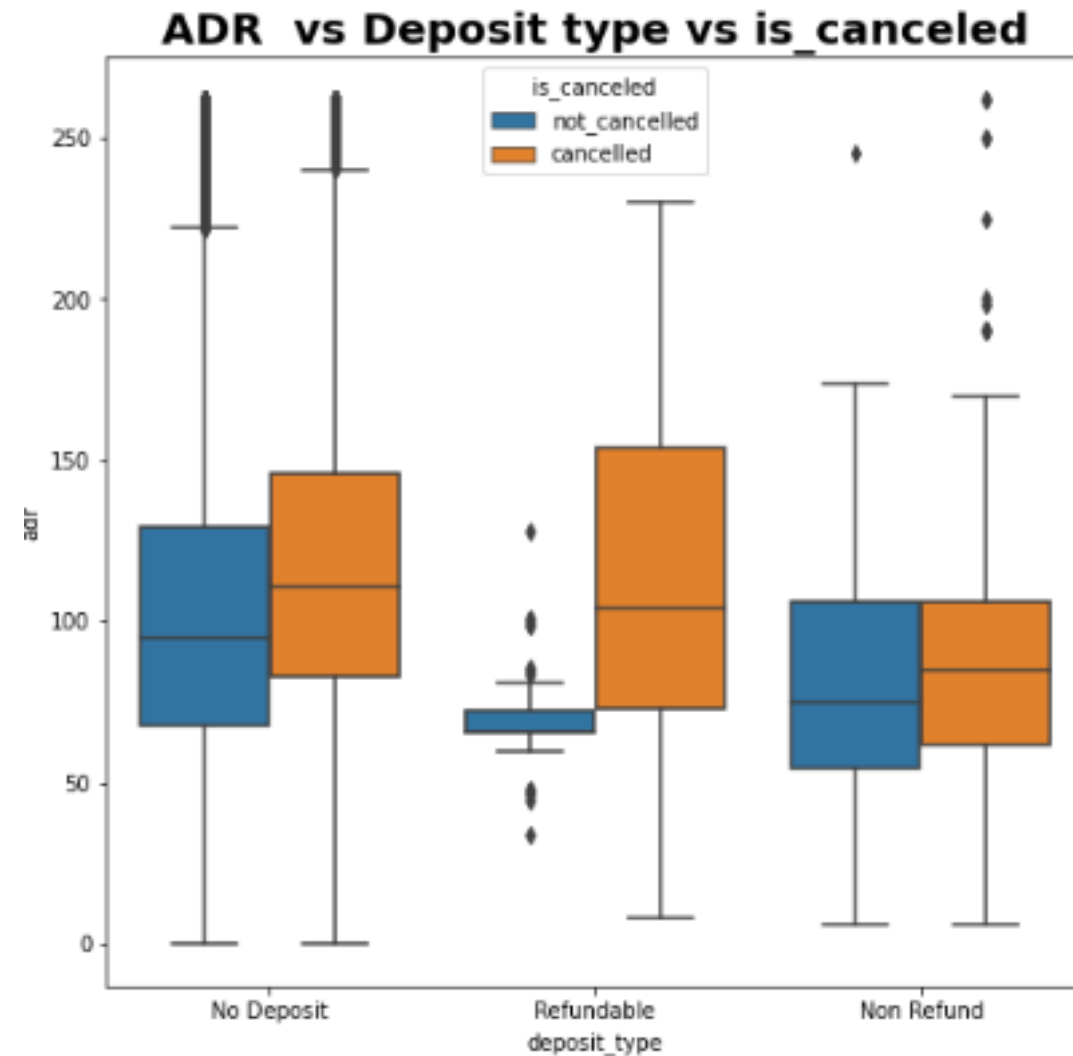
Year wise , month wise and day wise bookings analysis :-

- Hotel bookings increased drastically in the year 2016 , then slightly decreased in 2017

- Year wise adr has increased continuously in 2017 in comparison to 2015
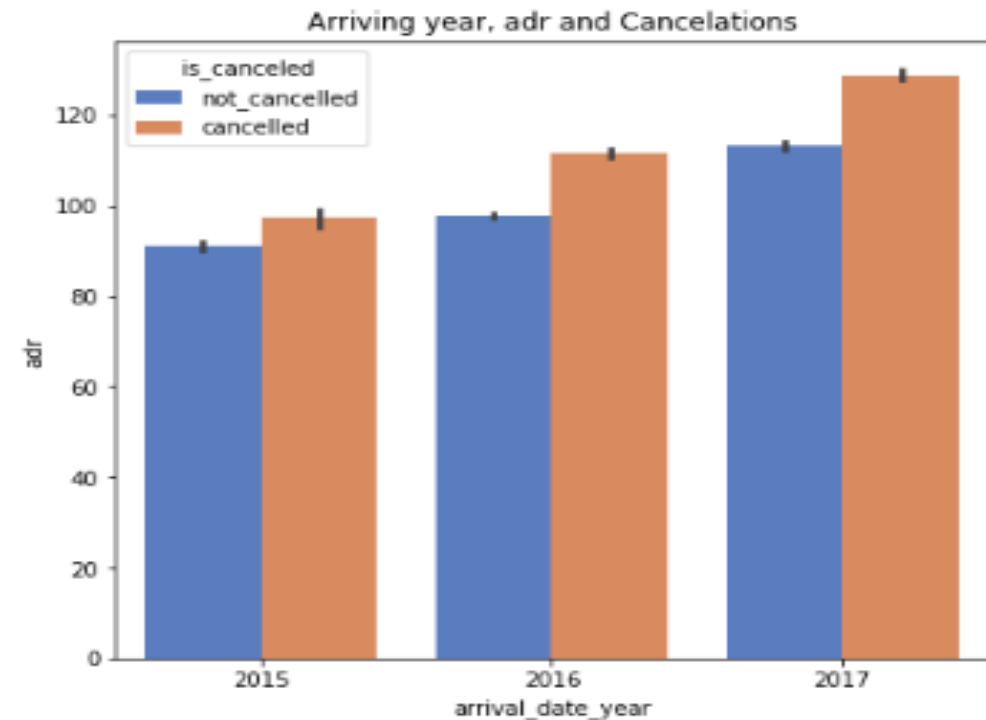
- The time of the year between March and October (mainly during summers and rainy season time) hotels get good amount of business.

- In season time adr stays between 90-150 , but in off season it goes down up to 50 .So people can book hotels between November to February to get cheaper deals
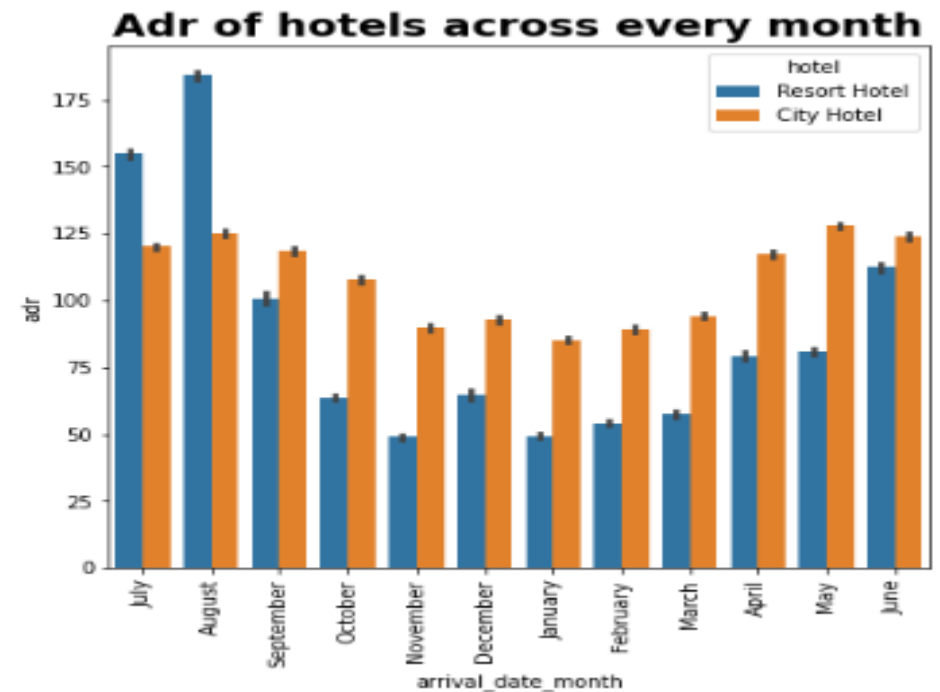
- 1st week of the month gets more bookings and booking percentage drastically decreases during last week or month end time.
- Average adr for all day of month is quite same but small peaks are during the weekends where hotels charge a little more money.
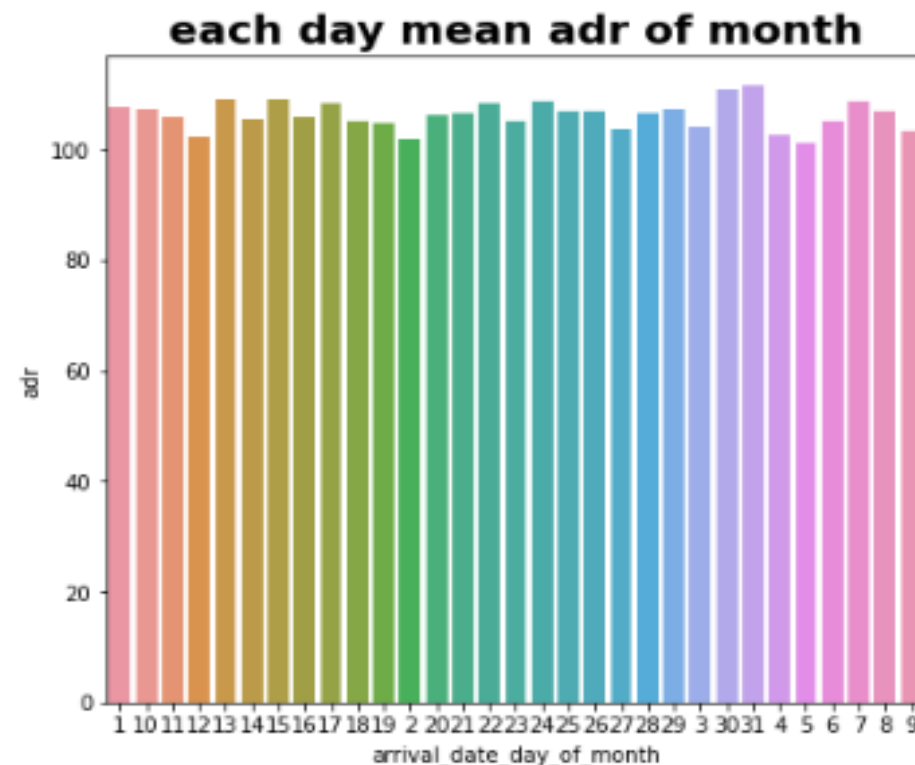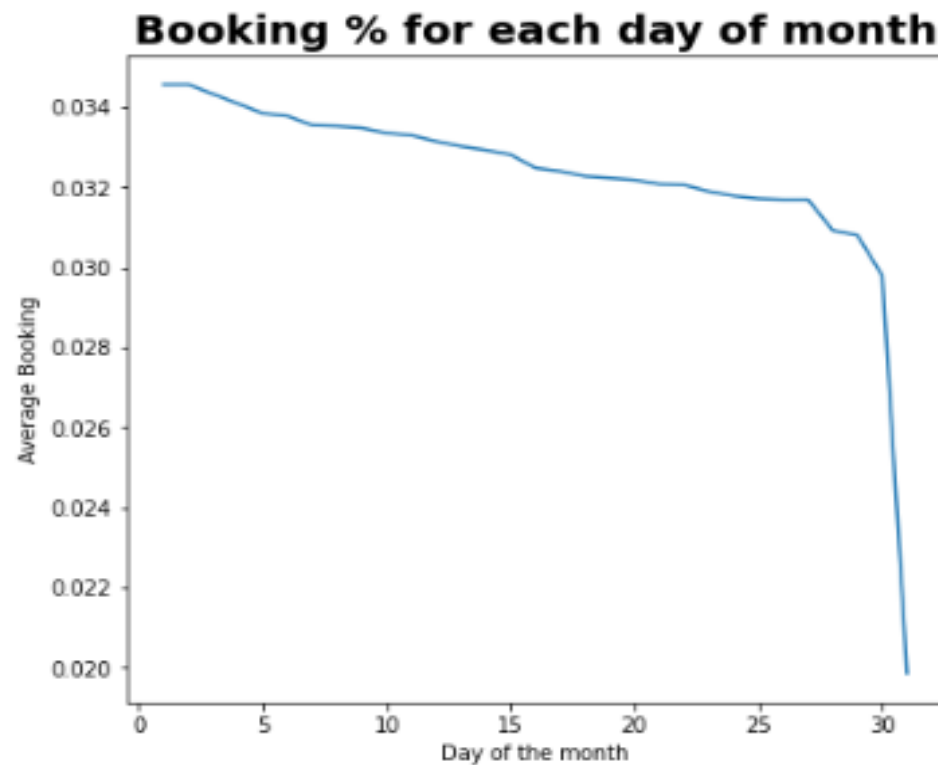
Food of Hotels and special requests by guests:-

- BB (Bead and breakfast) is the most preferred meal for guests which is comes complimentary with the booking.

- FB (Full Board- Lunch,Breakfast,Dinner), HB (Half board- breakfast, dinner),SC(Self catering) are less popular because hotels charge extra price for the meal , which is clearly seen in its adr.
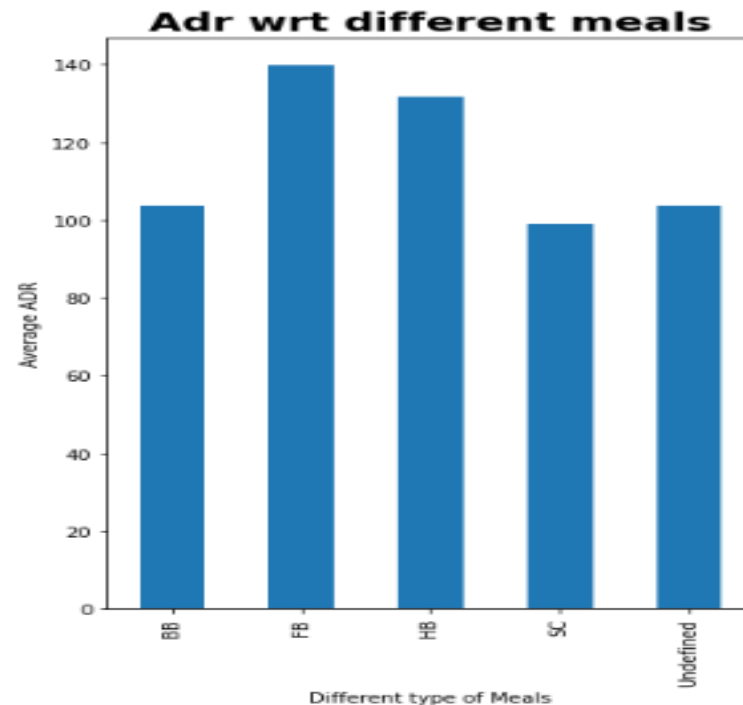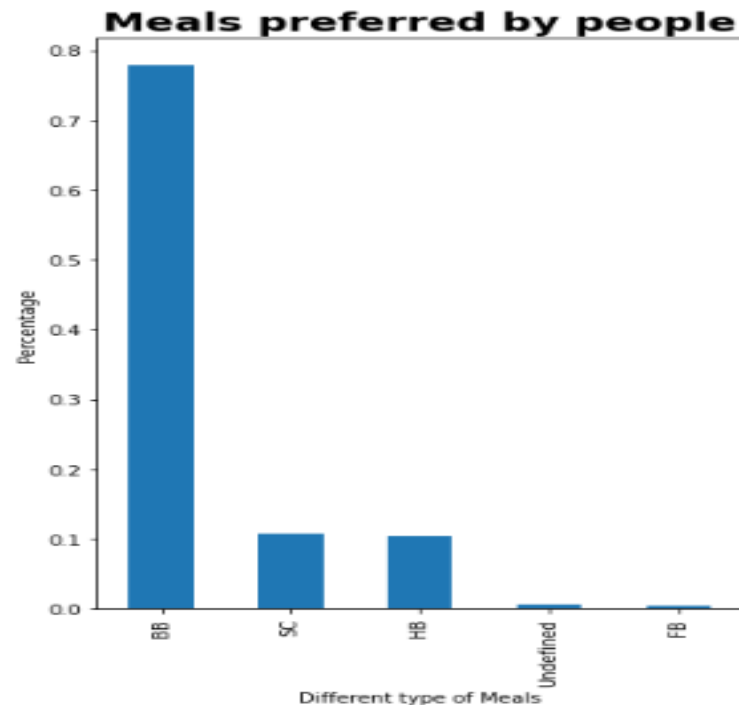
- Whenever there is more no of special requests from the guests there is a high chance that the booking will not be cancelled
- Around 50% of guest don't have any special request.
- Hotels charge extra money for special requests mainly with resort hotels may be for special occasions like birthdays or weddings
- Hotels gets most special requests in the month of August, July and May. So hotels should be ready to fulfil the requests in those months accordingly.

ADR according to special requests by guests

total_people vs total_stay vs cancellation
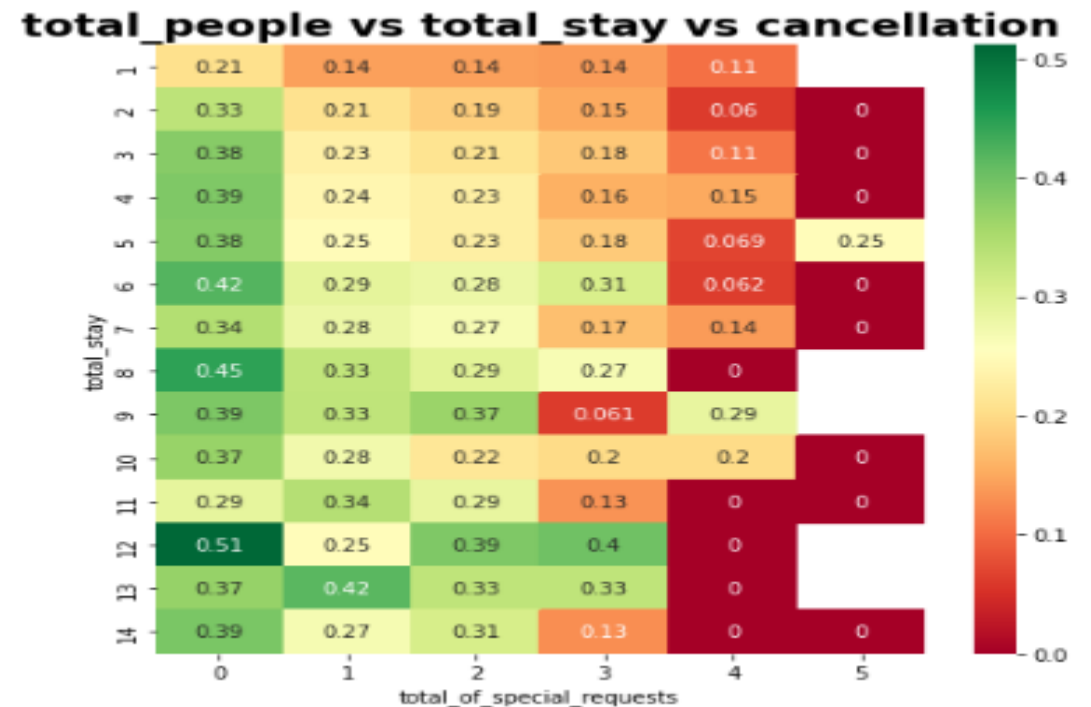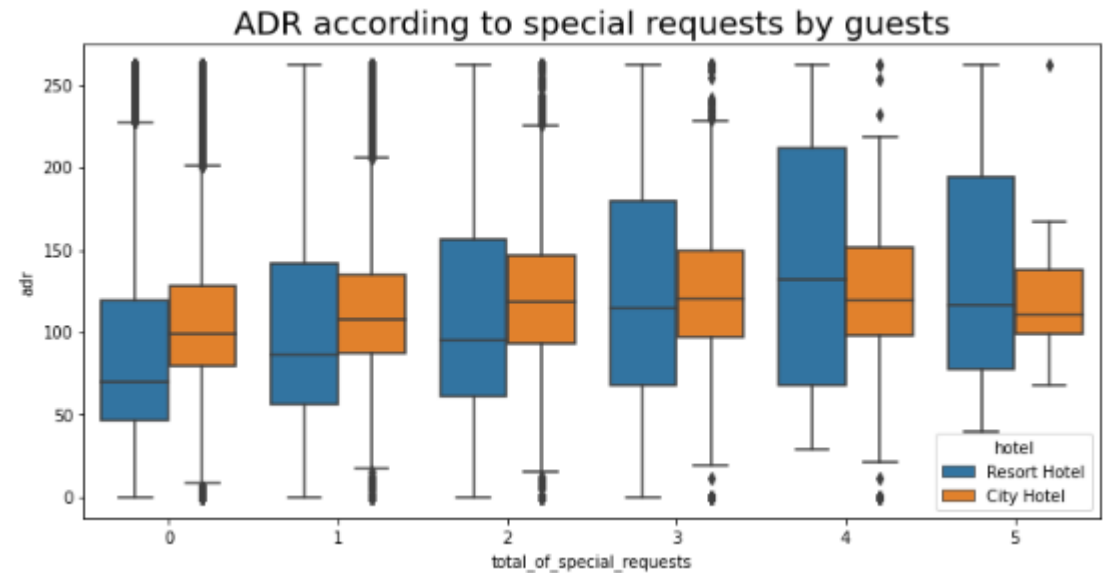
Importance of repeated guests:-

- Hotels are having very less number of repeated guests. They should work for customer retention and building customer loyalty .

- Repeat customers are more likely to upgrade and purchase your most expensive products or services. It all comes down to the development of trust.

- Attracting new customers costs five times more than keeping an existing customer. This is largely due to advances in technology and the associated costs and the increasing cost of marketing, including cuts given to online booking sites, agents etc.

- Repeat customers spread the word-of-mouth & recommendations to friends and family.

Graph showing whether guest is repeated guest
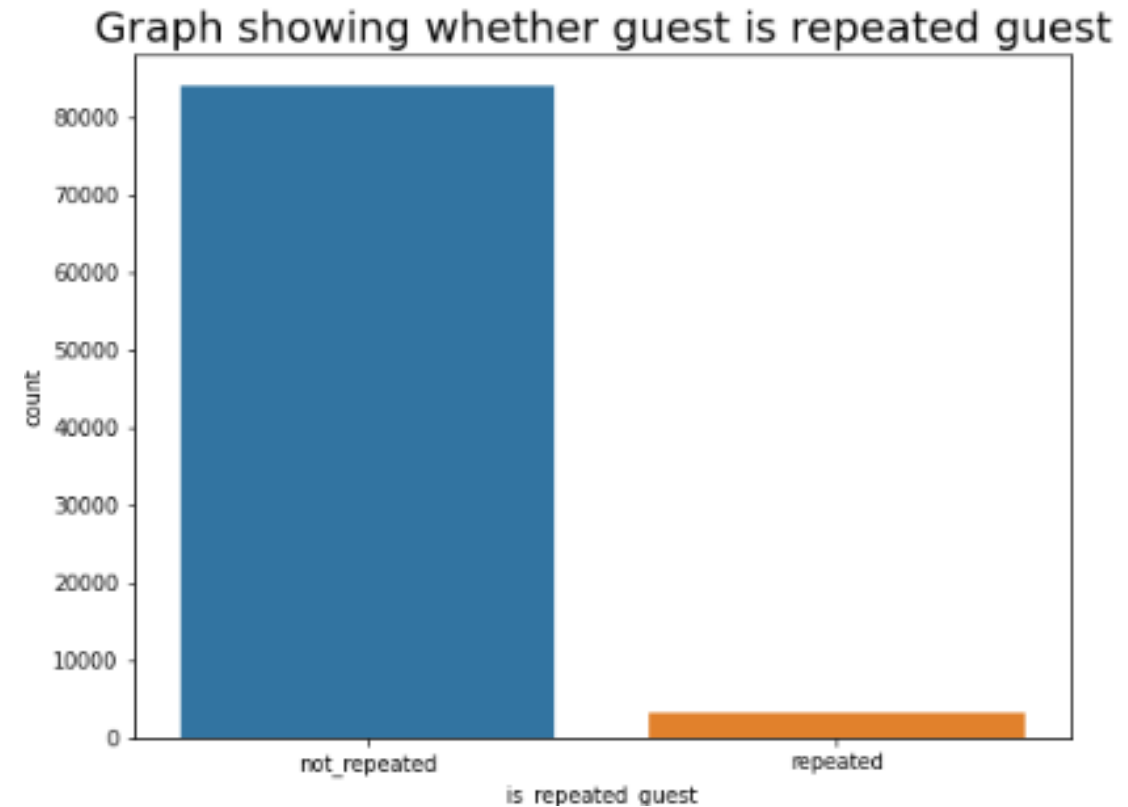
Conclusion and Solution for the client:-

- Hotel management should give highest priority to give the guests a good stay experience so that no of repeated guest will increase . In case of repeated guests, they book rooms directly with the hotels without any interference of any online travel agents . This way hotel booking rates will be cheaper and hotels can save the commission money of agents.

- Develop a loyal, profitable customer base. Reduce guest acquisition costs. Upsell and cross-sell your services to those who already value you.

- Hotel can take help of social media for marketing purposes, can hire travel influencers in between March and October peak months. Invest in creating a good personal branding of the hotel.

- Hotel can hire good individual agents who can pitch with corporate clients. Hotel can also collaborate with event planners .

- Hotel should be ready for any number of special requests to fulfil for guests, for which they can charge extra price also.

- Hotel can experiment with their meals. Give some extra complementary food items. Foods should be tasty .

- Hotel should give extra efforts on cleanliness, room service and hire good professional staff.

- Try to keep adr between 90-110, because it is more affordable for guests and also attracts more non cancelled bookings.

- In offseason months hotels can run special offers or discounts for hotels bookings and approach the people through online marketing strategy

- Hotel should keep no advance deposit policy. But in case of higher lead time bookings, hotels can charge up to 20%-30% cost of stay as non-refundable advance deposit because bookings with lead time more than 80 days have more chances of getting cancelled.

- Try to keep lead time less than 70 days, more than that have more chances of getting cancelled.

- Hotel should follow up actively with the high lead time bookings so that genuine bookings are not put in the waiting list.

Strategies for customer retention:-

- Target guests who stayed recently, spent a lot of money, and come back often and personalize their experience and develop a strong relationship with them first.

- Note any past interests, activities, or behaviours that the guest displayed during their last stay and use that information to craft your outreach.

- Understand the motivation behind repeat bookings and try to craft guest experiences that create community, relationships, and a sense of home

- Offer the right repeat booking incentives and apply guest feedback to future stays.

- Automate follow-up to keep the conversation going.

Thank you...