Capstone Project:1

EDA Hotel Booking

Amitha Sridhar

Introduction

Hotel rooms are your second home on your journey.

The global hotel industry is a $525bn industry, made up of around 18 million rooms.

But this is a very massive competitive market and to be the front runners the hotels will have to keep up with the customer demands and needs. The study of the trend will help these players to upgrade themselves.

In this project we will be using our knowledge of python to study the compressed data of hotel booking. We will be studying about 2 hotel types i.e the city hotel and resort hotel. We will be studying factors like lead time, wait time, deposit, meal types etc to understand what's more in demand.

We will be deducing conclusions such as:-

> Optimal stay
> Distribution channels who can get maximum revenue
> Market segment which gets most revenue
> Customer type who can get in most business
> Does meal types, deposit and car parking be a factor
> Finally the most successful country who get the most revenue

Workflow

AI

Problem Statement and collection raw data

Data cleaning and Data Manipulation

Exploratory Data Analysis

Data visualization

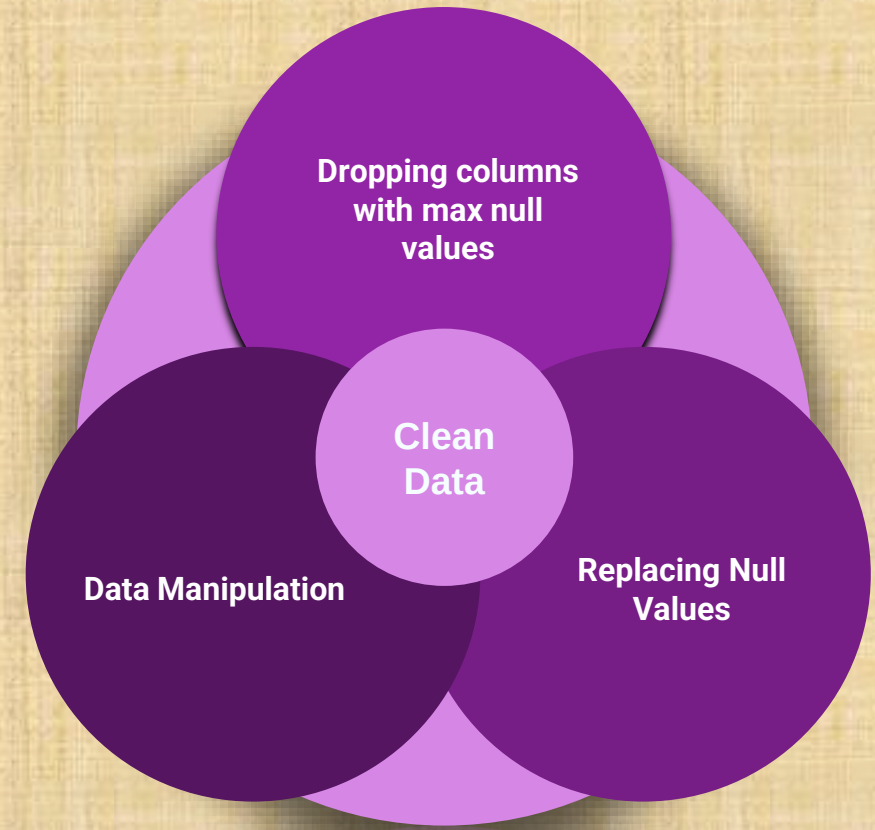
Analyzing the result

Attributes in the study

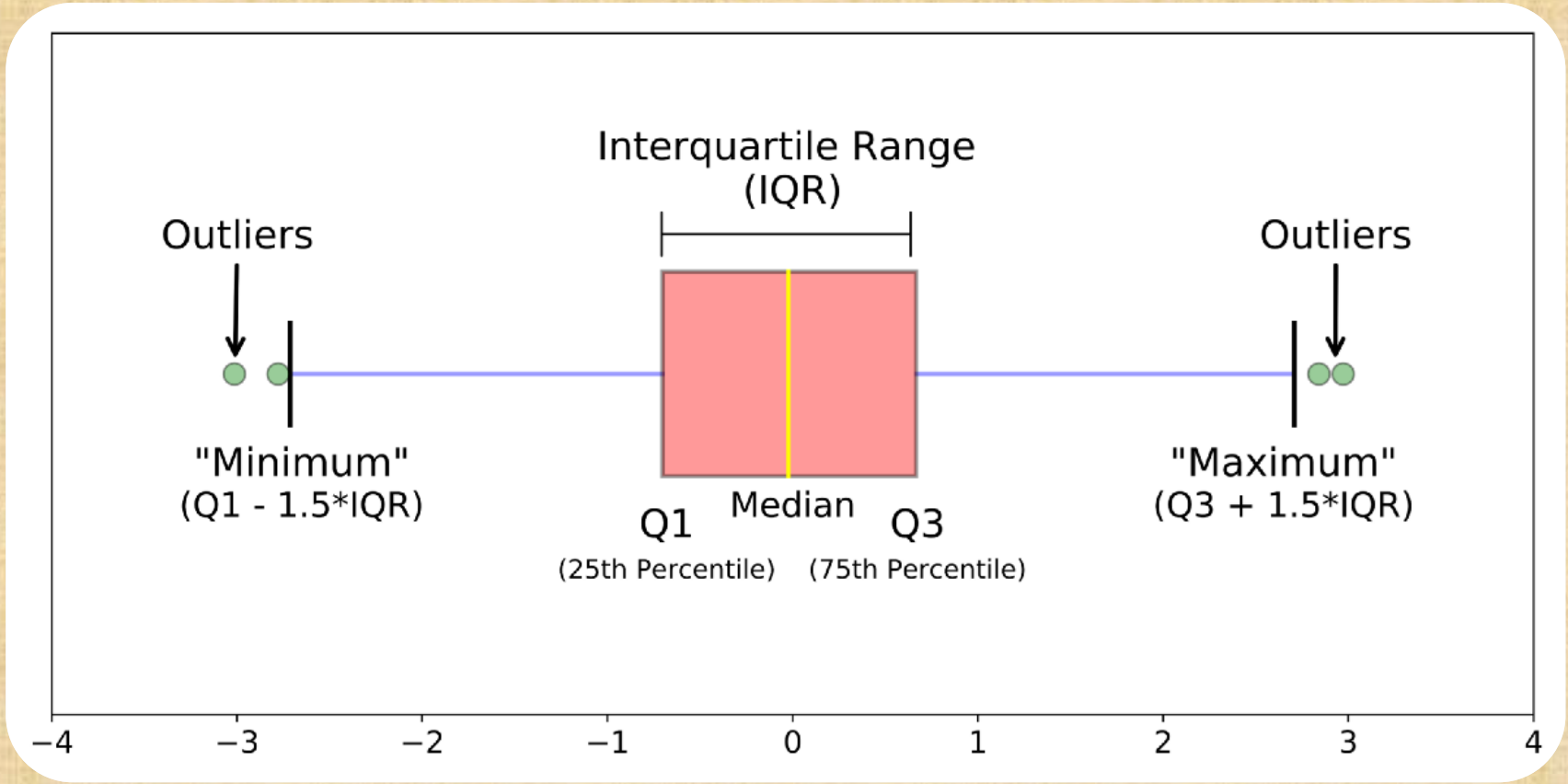
➔ This data has 119390 rows and 32 columns. Here are the columns

- hotel
- is_cancelled
- lead_time
- arrival_date_year
- arrival_date_month
- arrival_date_week_number
- arrival_date_day_of_month
- stays_in_weekend_nights
- stays_in_week_nights
- adult
- children
- babies
- meal
- country
- market segment
- distribution channel
- is_repeated_guest
- previous_cancellations
- previous_bookings_not_canceled
- reserved_room_type
- assigned_room_type
- booking_changes
- deposit_type
- agent
- company
- days_in_waiting_list
- customer_type
- adr
- required_car_parking_spaces
- total_of_special_requests
- reservation_statue
- reservation_status_date

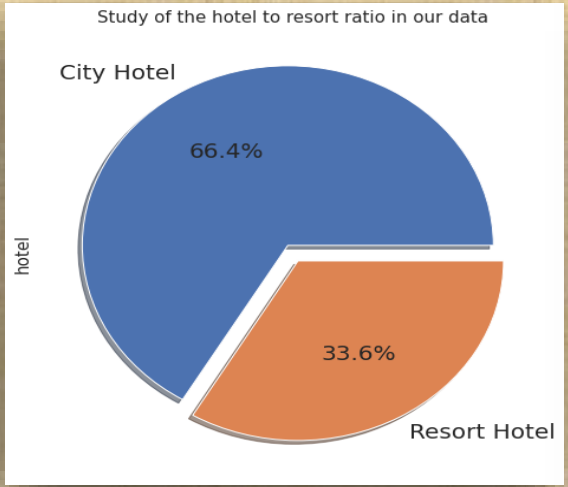
Data Wrangling



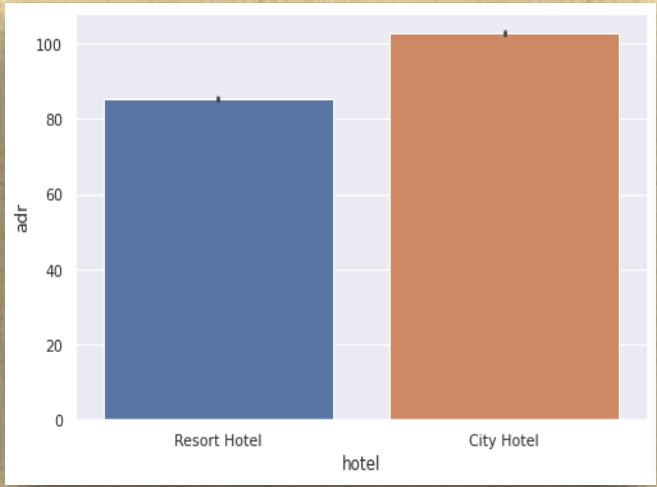
Handling outliers



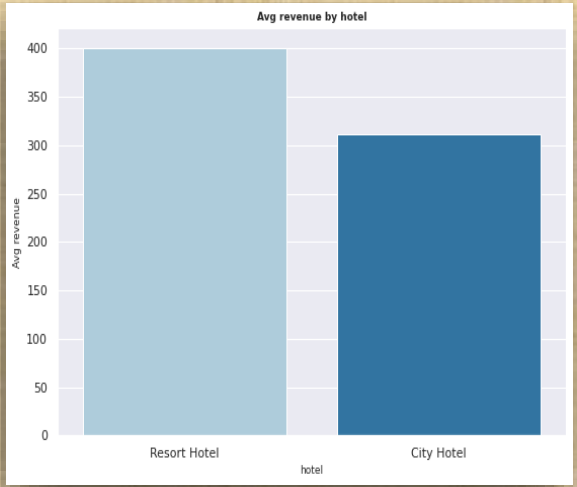
Hotel type studies



Study of market share in our dataset. Around 66% of our dataset is a study for city hotel and the remaining 33% is about the resort hotel.

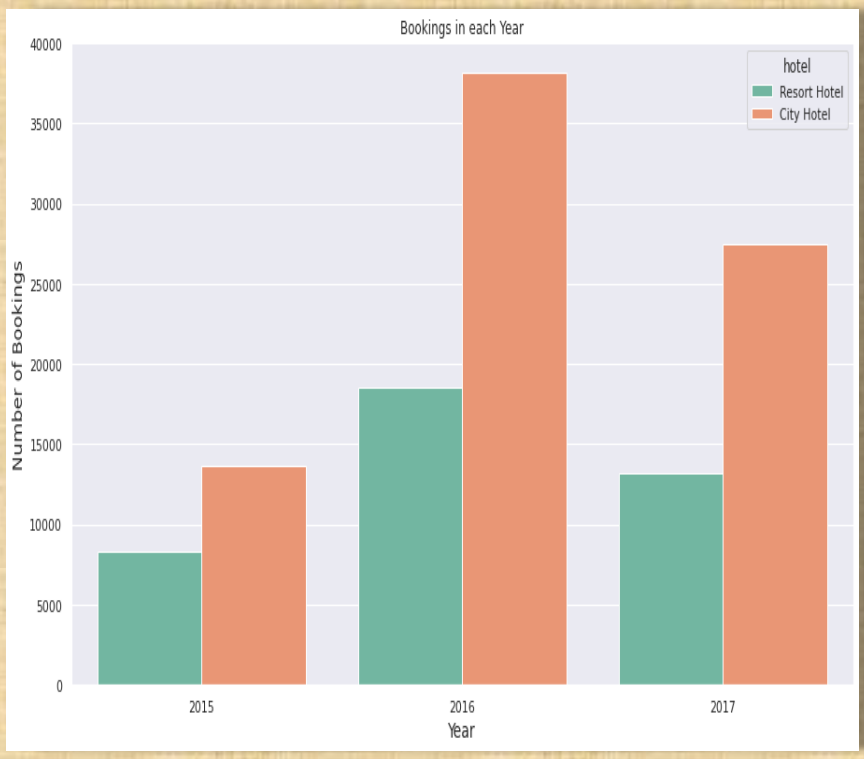


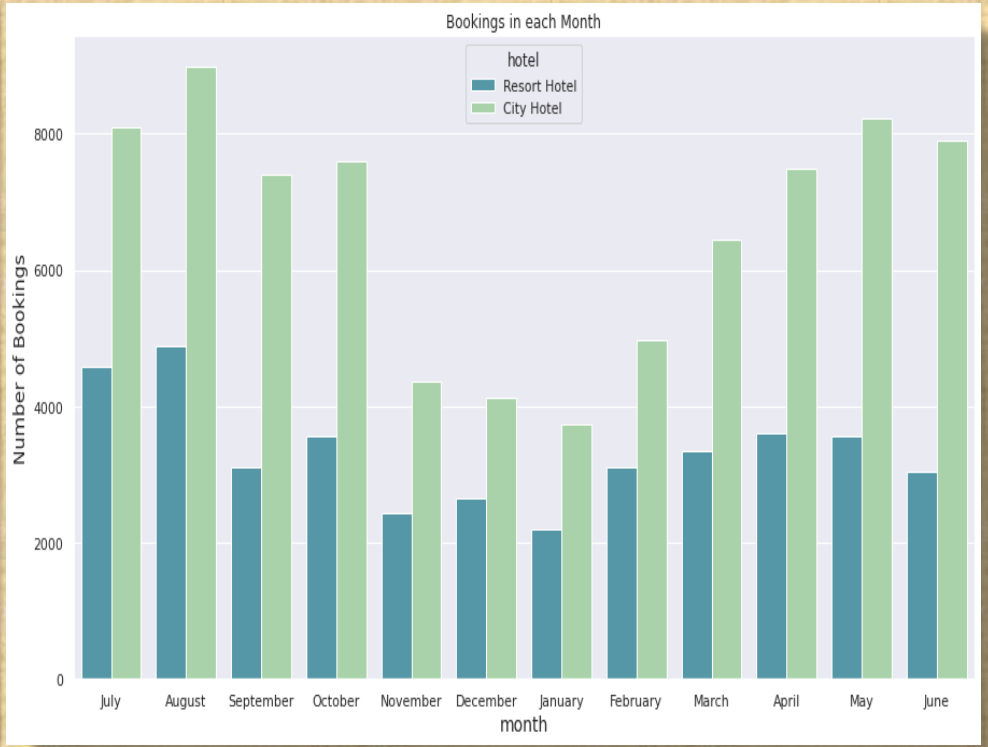
Average daily rate of resort hotel vs city hotel. Resort hotel have ADR around 80 and then the city have a higher ADR of around 100.



Average revenue earned on hotel type basis. Though the ADR is lesser for the resort type the revenue earned by resorts are higher.

Booking based on hotel and month

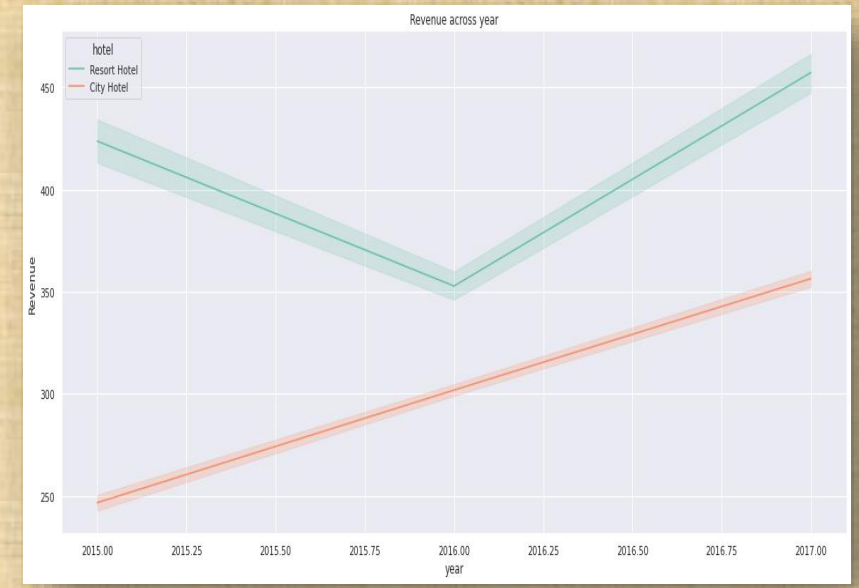


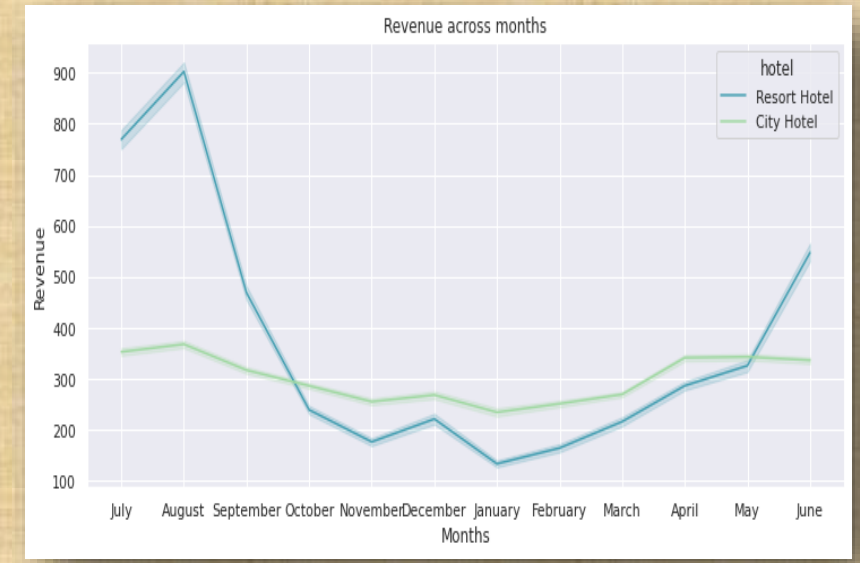


Booking throughout the years for study. 2016 sees a peak in booking number in both types of hotel.

Booking throughout the year. Resort Hotels sees peak booking in July and August. Whereas City hotels sees increase high grow in booking between April to October.

Revenue study based on year and month

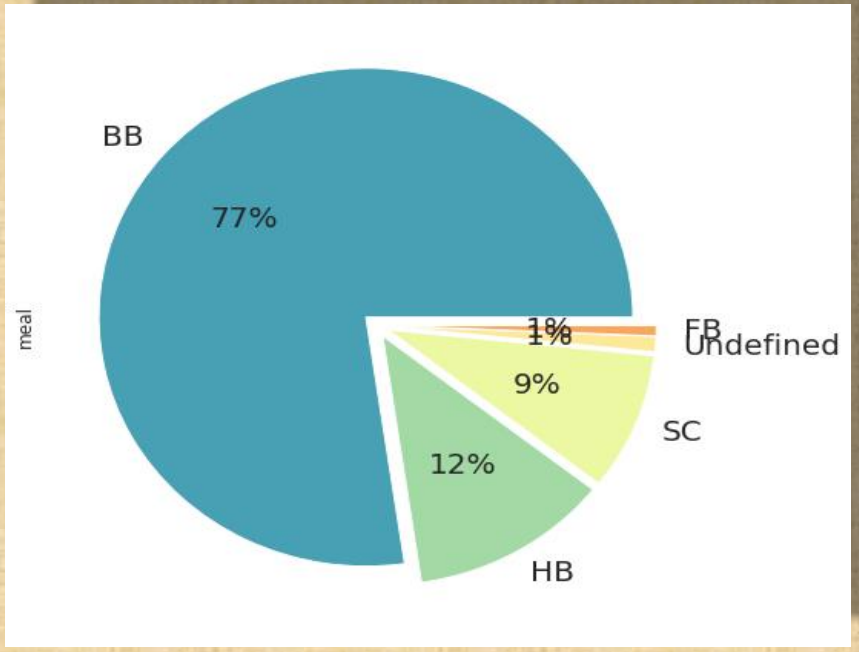


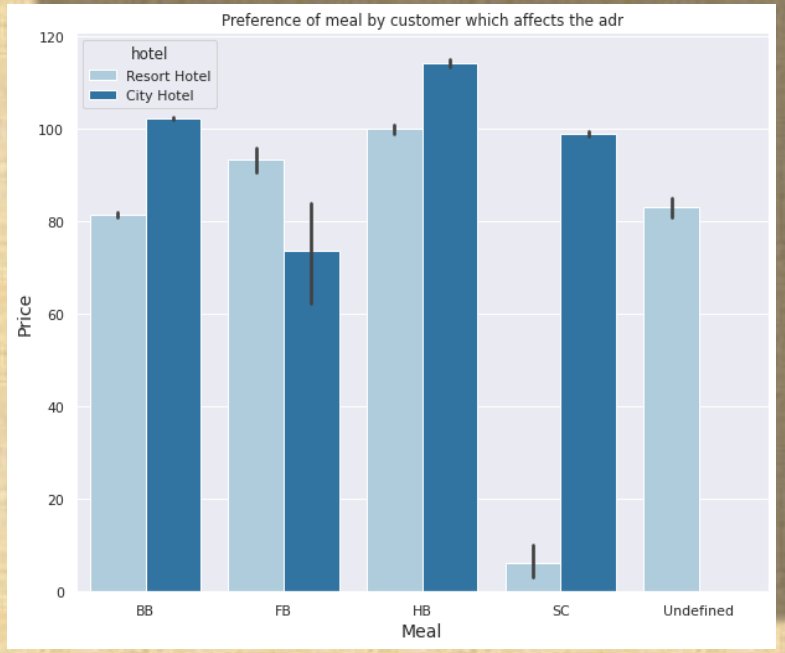


Revenue earned over 3 years. City hotel have constantly increased their earnings over 3 years. Resort hotel type have dipped in between in 2016 but then again improved.

Revenue earned by hotels over months. As we have earlier seen the bookings are highest in July and August for the resort hotel so the revenue in those months are highest whereas in city hotel its highest between April to august.

Meal Type Studies

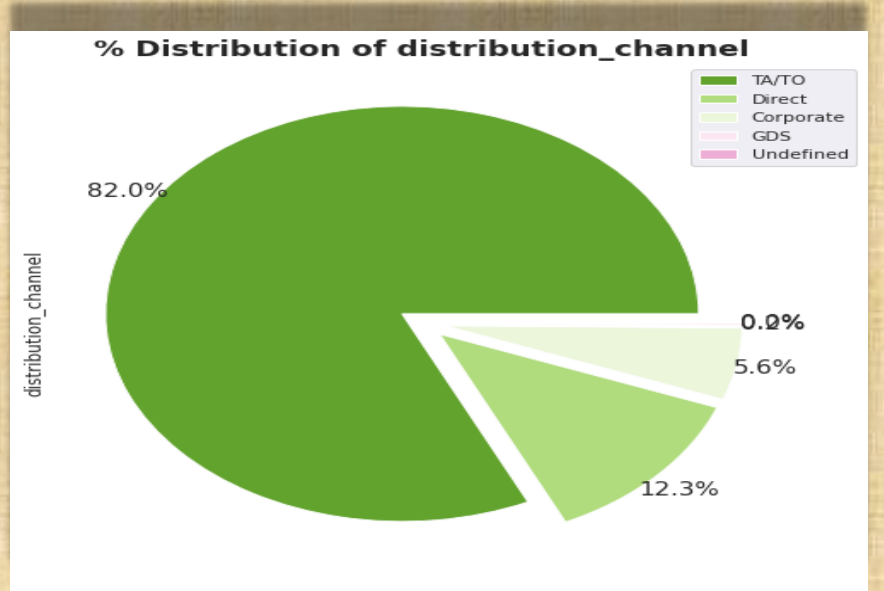


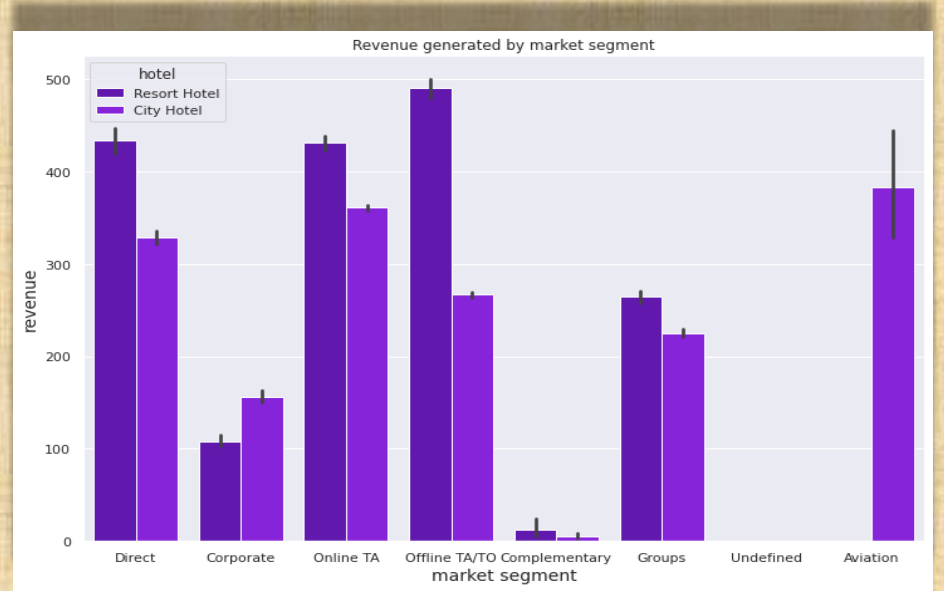


Most preferred meals in this study is BB around 77%.

Half Board has highest ADR followed by BB.

Study of distribution channels and market segments.

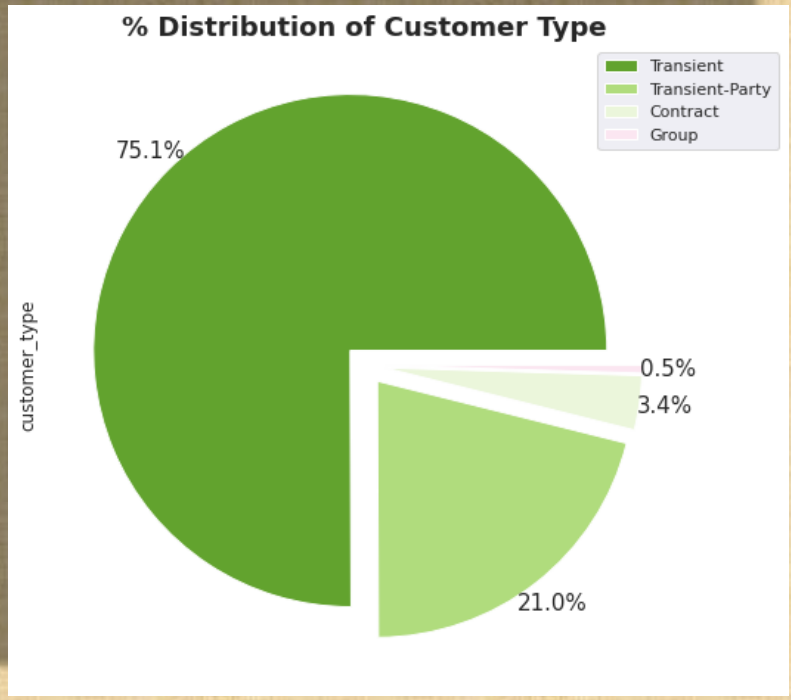


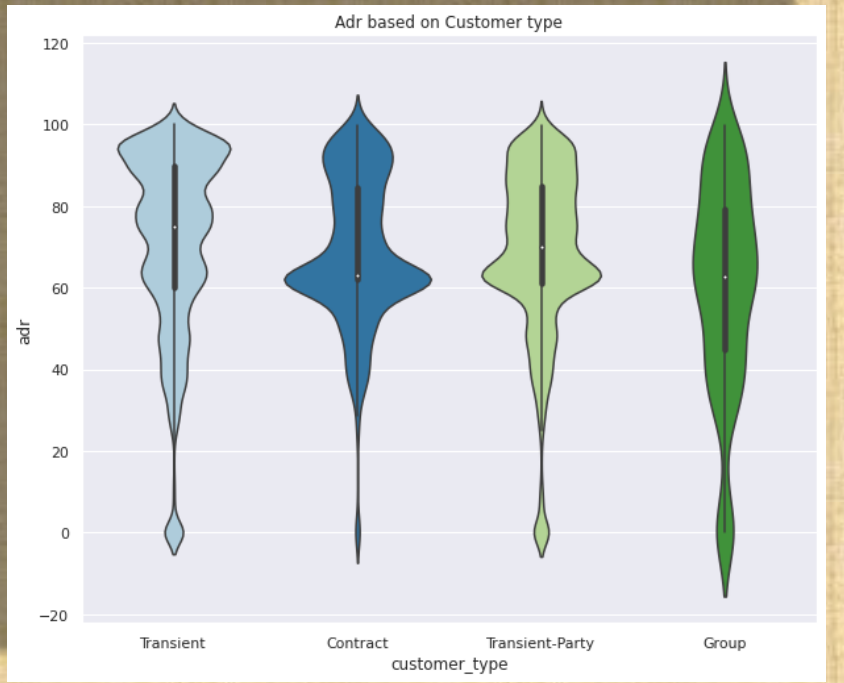


82% of the booking has been done through TA/TO

Most revenue generated Resort Hotel is through Offline TA/TO followed by online TA and Direct whereas for city hotel type is through Aviation

Customer Type Studies

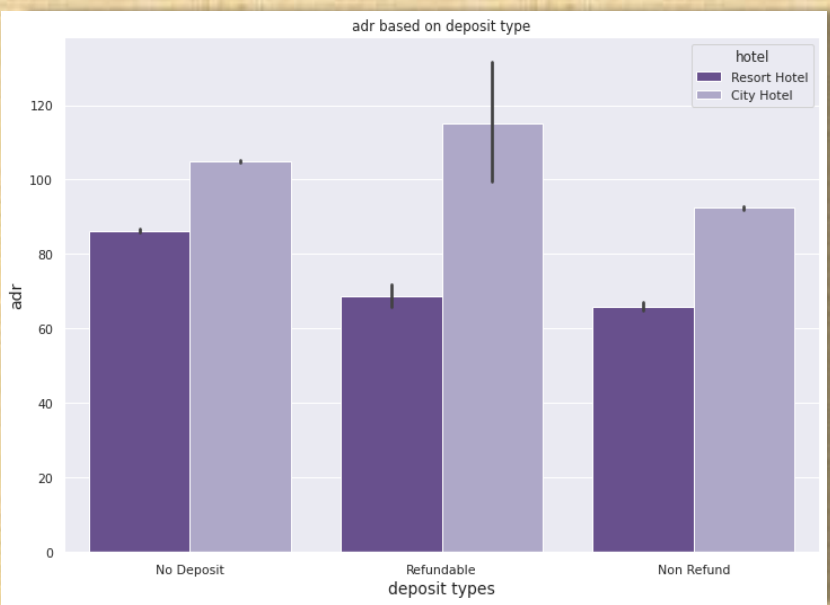


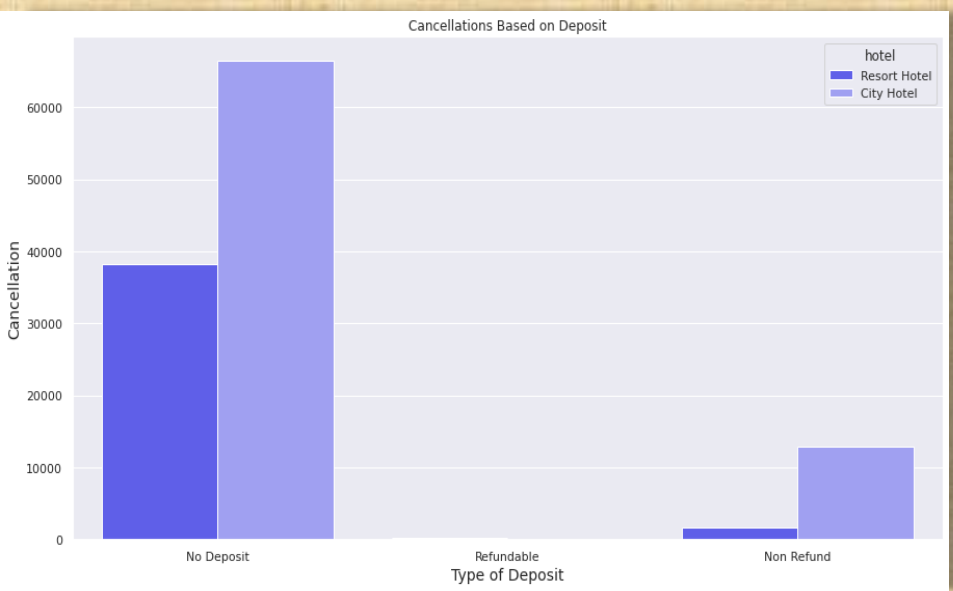


Around 75% of customers are transient

Transient has higher ADR whereas contract has the lowest ADR with maximum distribution around 60

Deposit type studies

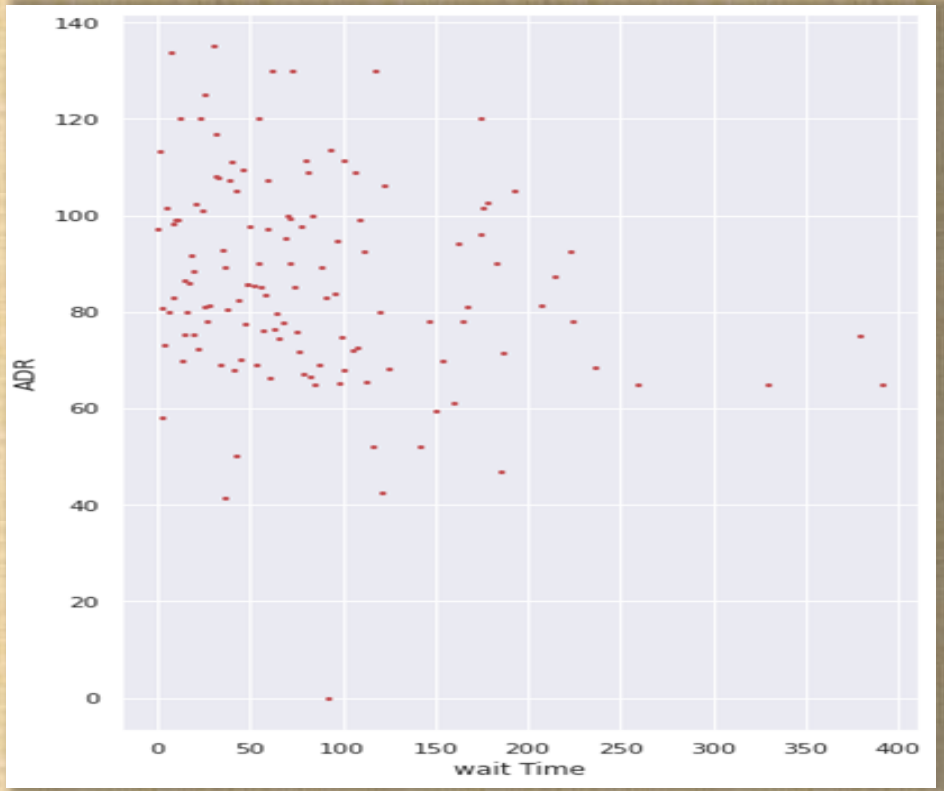




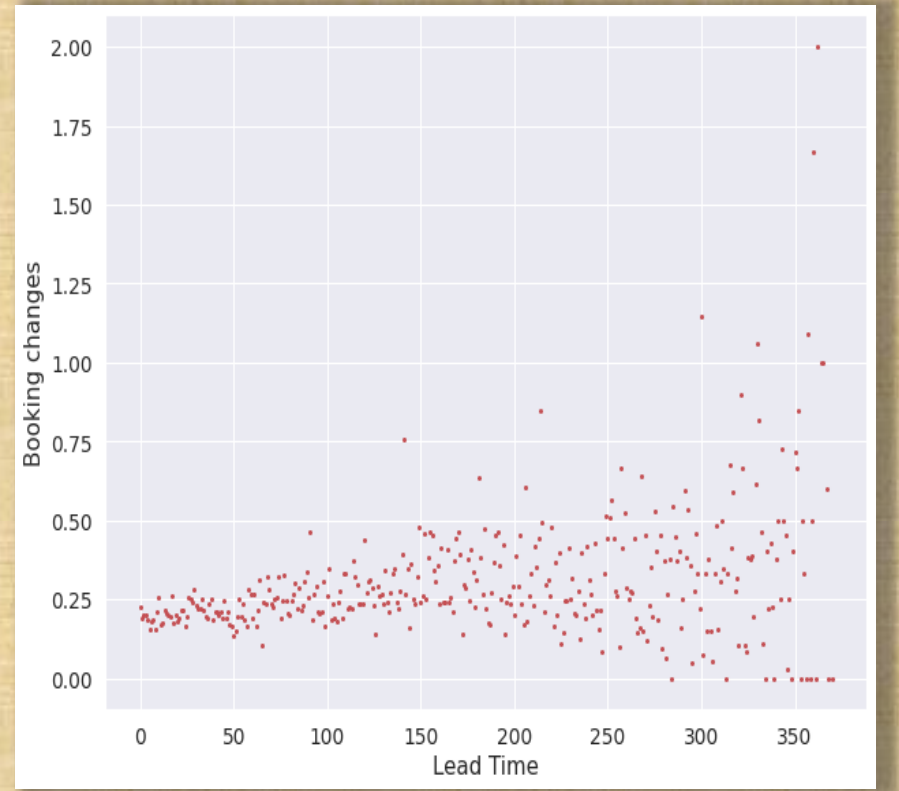
ADR of refundable hotels are more in resort hotels whereas in city hotel No deposit has higher ADR.

Cancellations in more when there is no deposit.

Lead Time Studies

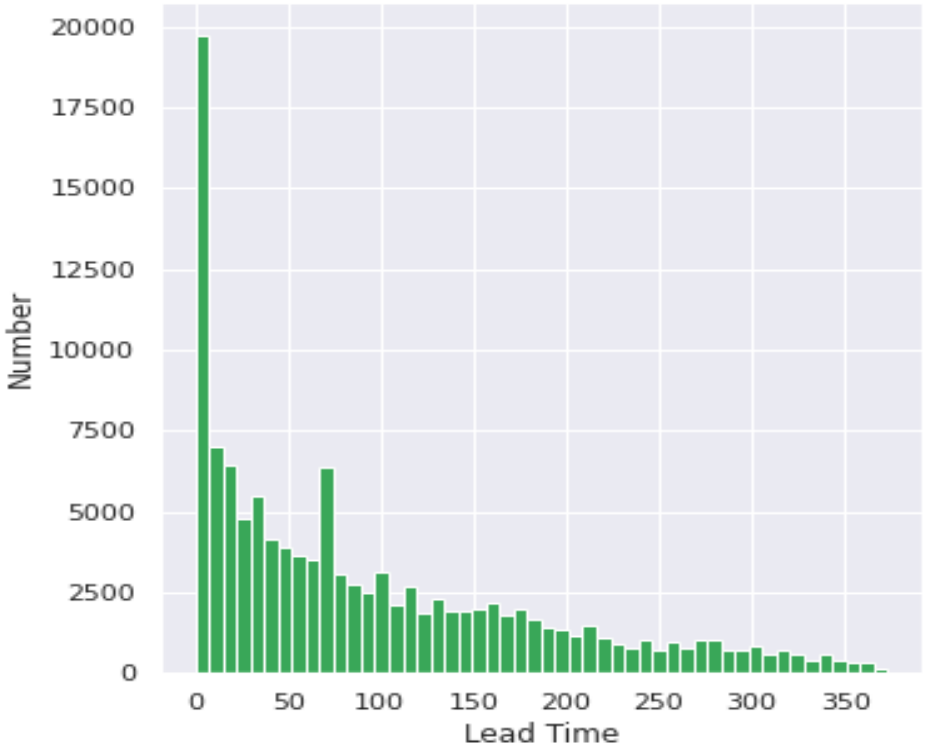


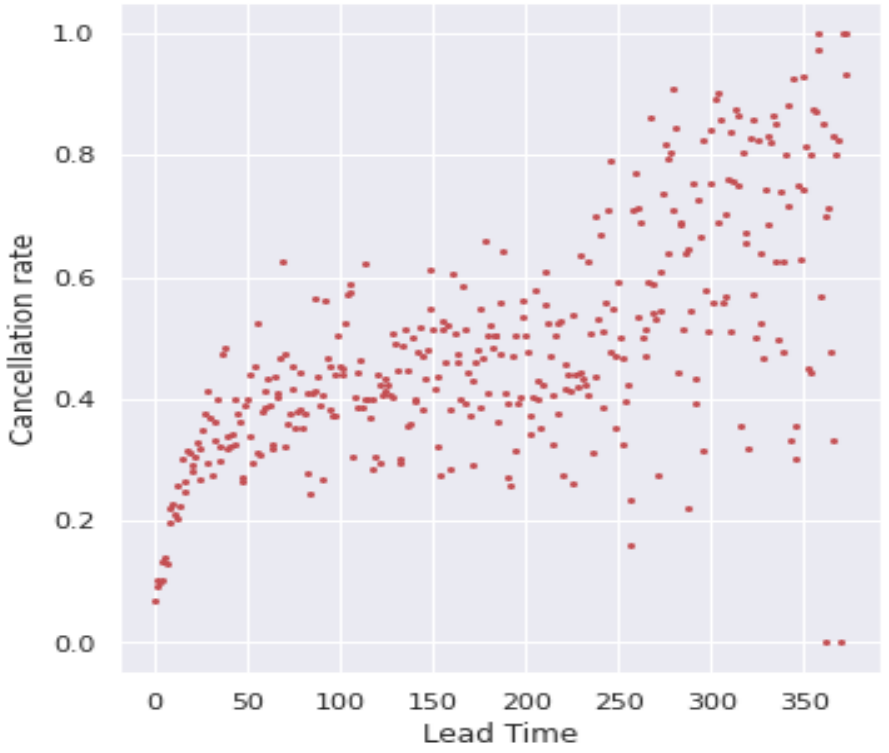
ADR reduces with wait time



Advance booking has more booking changes

Lead time Studies

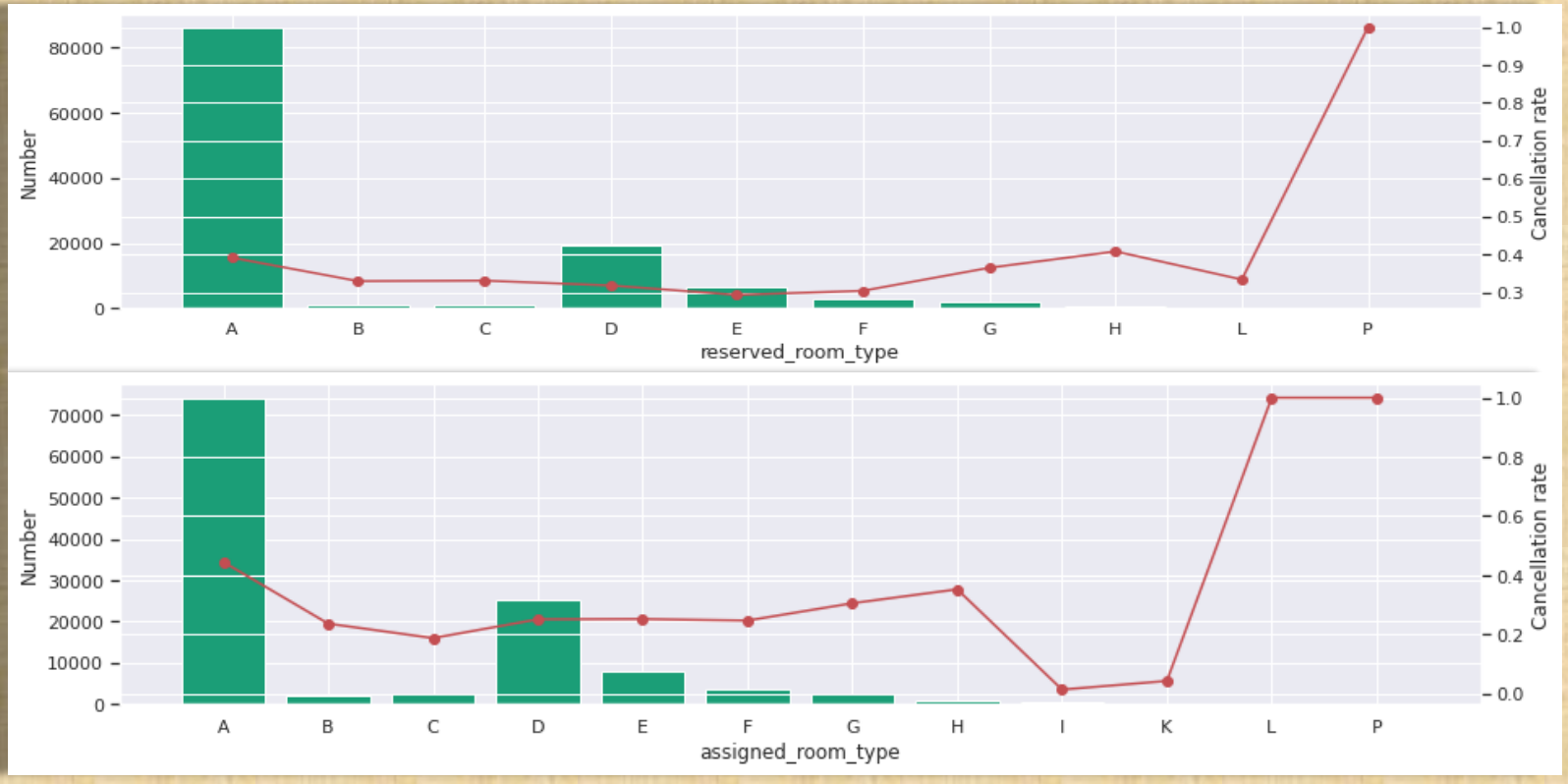




Most booking is done in the last minute

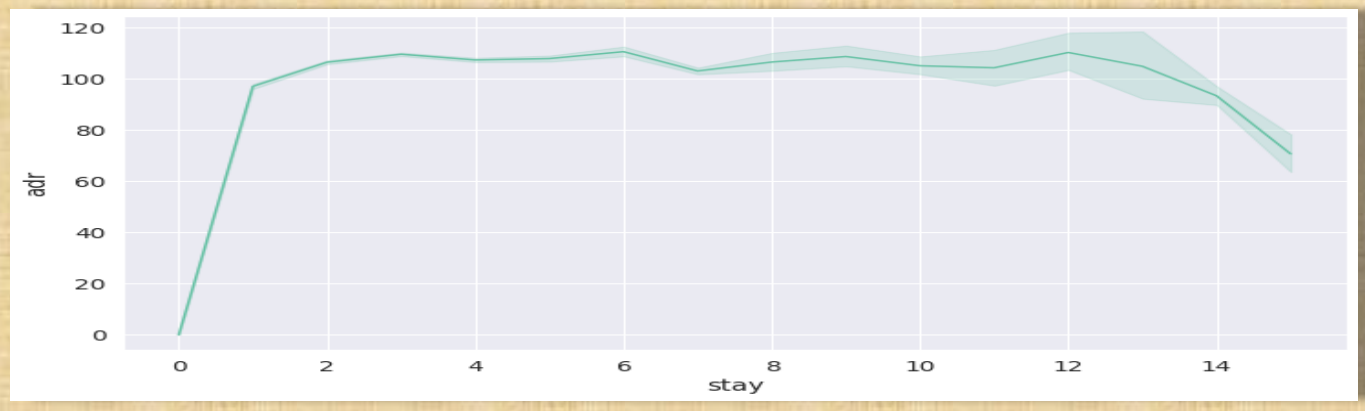
If the booking is more in advance chances are higher in cancellations

Room Type study

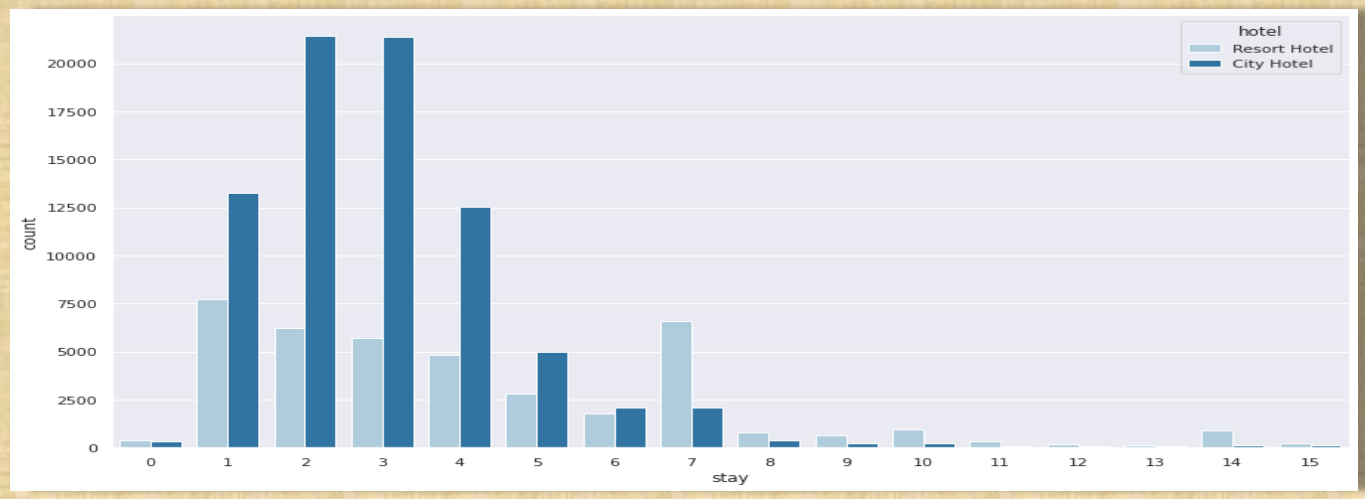


Around 80000 of the room types booked are for room type A whereas 70000 room types are for room type A but the change in the room type does not play a role in cancellations

Optimal Stay studies

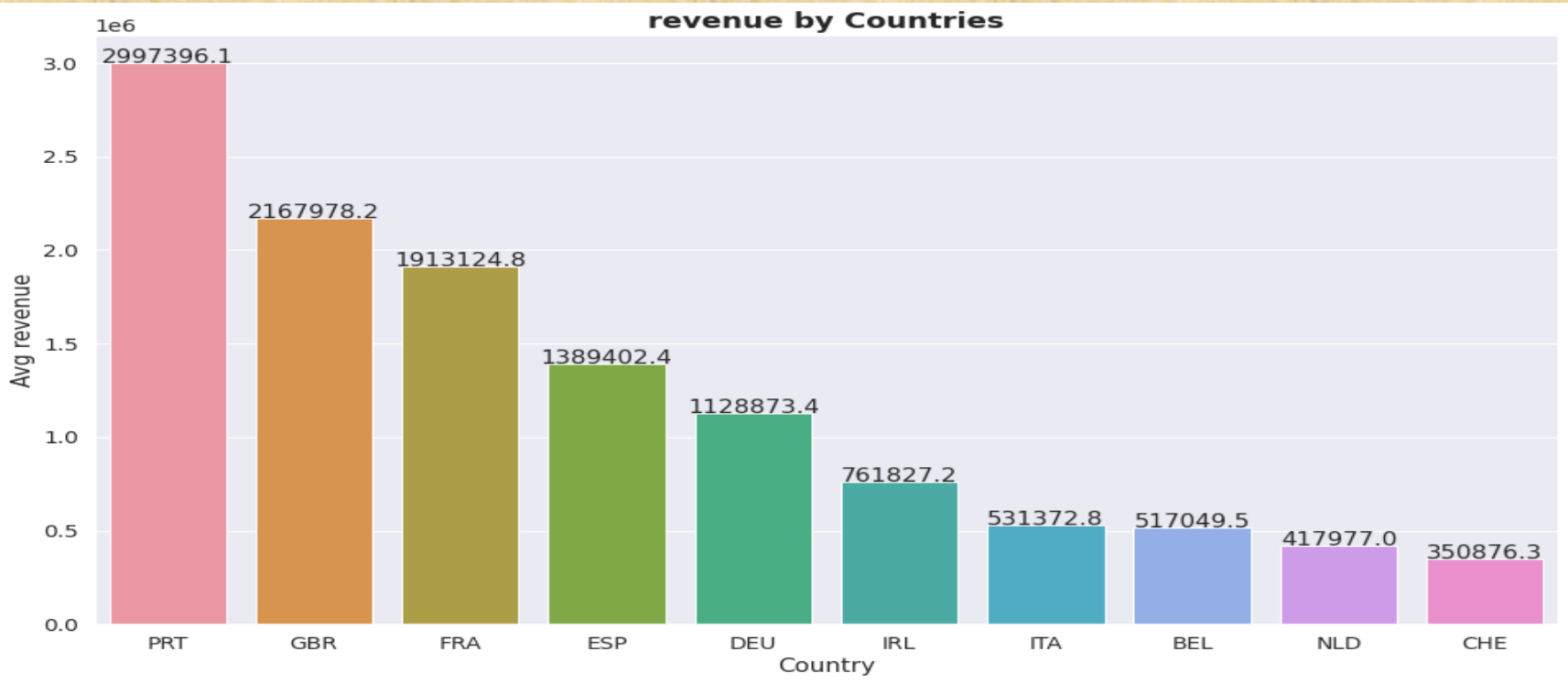


The ADR for shorter stays are higher and as the stay is more the ADR reduces



City hotel are preferred for shorted stays around 1/3rd of the bookings are in 2-3 nights whereas resort hotels are distributed between 1 night upto a weeks stay.

Revenue across top 10 countries



The top countries had mostly the European Countries which have generated the highest revenue across the 3 year study.

Cumulative Conclusions

- 25% of the booking in resort and 37% of the bookings on the city hotels gets cancelled.
- Only around 1% of repeated customers
- Usually, customers with a group size of 2 are more prone to cancel and kids don’t play a major role in cancellations
- Car parking don’t play a major role for customer booking.
- Lead time in city hotels are higher around 3.5 months(100 days) whereas the resort hotel lead time is around 3 months.
- City hotels are in higher demand thus the wait time for it is high. Most customers tend to book with 0 wait time.Also, its seen that the ADR is lesser if booking is done ahead of time.
- Resort hotels get a greater number of special requests.
- The most remote islands like United States Minor Outlying Islands have pretty expensive rooms.
- Portugal has around 28% of the hotel share which makes it the country with the most hotels in our data with highest revenue earned. Followed by Great Britain with around 12.9% of the hotels

THANK
YOU !