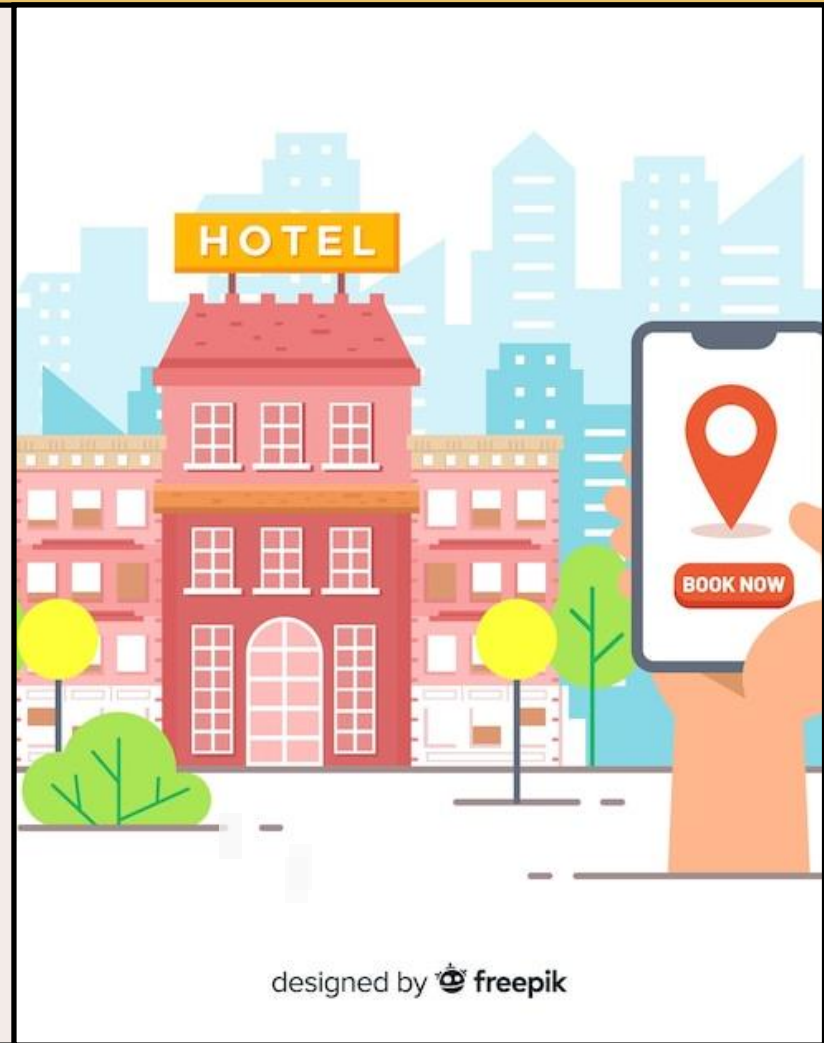
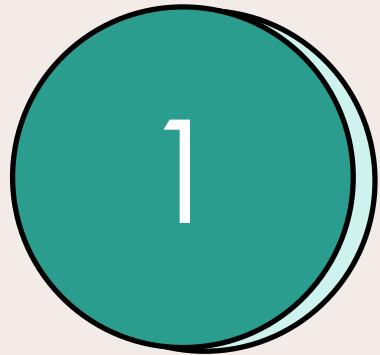


# HOTEL BOOKING EDA - PROJECT

Deepak Kumar Sahu



# Agenda



Introduction



Data  
Wrangling



EDA



OBSERVATION



Conclusion

# Introduction

A few decades ago, traveling was not a part of everyday life. But today travel is an enormous budding industry of 8.8 trillion economy. This directly affects the hotel industry which is highly competitive. We are here with a compact data to study about the hotel industry, mainly the booking. We are focussing on two types of hotels in this study. This data set contains different hotel types, countries located, guest, stays. Also the study have some factors that affect booking like wait time, lead time , months, average daily rate etc.

By the end we will conclude the study with following insights:-

- Best time for booking
- Optimal duration of stay
- Distribution Segment and market segments to be focused to increase revenue
- Factors leading to cancellation which affects the revenue.
- Factors like meals, special requests etc. which might affect in the increase of ADR and revenue.

# Work flow

1

Problem statement and Collection of raw data

2

Data cleaning and Data manipulation

3

Exploratory Data Analysis

4

Data Visualization

5

Analysing the Result

# Attributes in the study

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

- Hotel
- Is cancelled
- Arrival date year
- Arrival date month
- Arrival date week number
- Day of the month

- Stays in weekend nights
- Stays in week nights
- Adults
- Children
- Babies
- Meals
- 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
- Lead time
- Days in waiting list
- Customer type
- ADR
- Required car parking spaces
- Booking changes
- Reservation status
- reservation status date



# Data Wrangling

## Data Cleaning

- Dropping Columns with Maximum Null values– agent and company
- Dropping columns which had redundant values– Arrival date week number
- Columns with nominal null values have been manipulated by filling them with
  - Numerical column:- Median (Number of children)
  - String column:- Mode (Country)

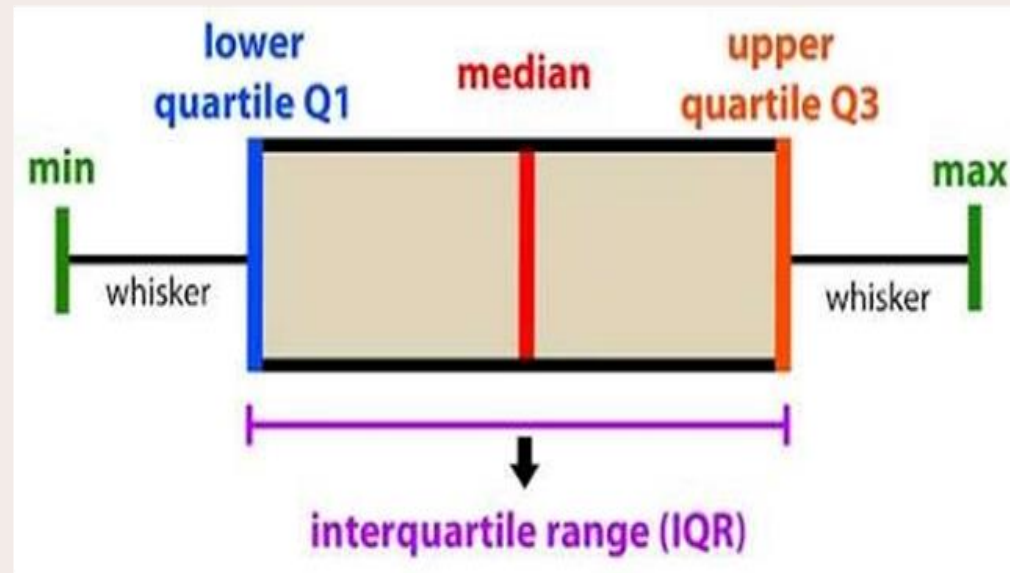
## Data Manipulation

Combining columns for an effective study

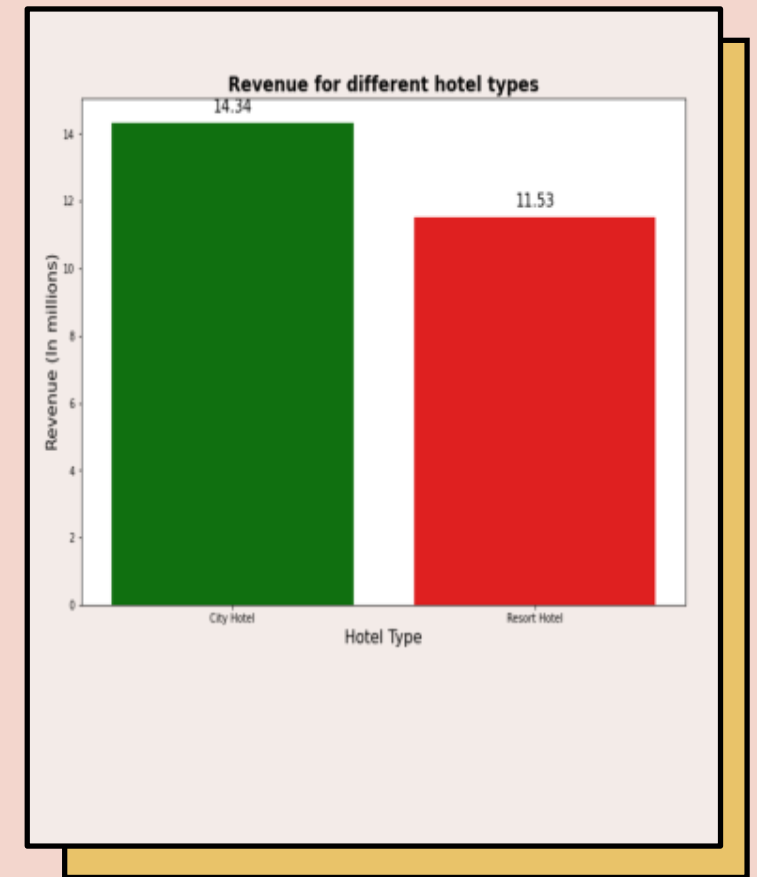
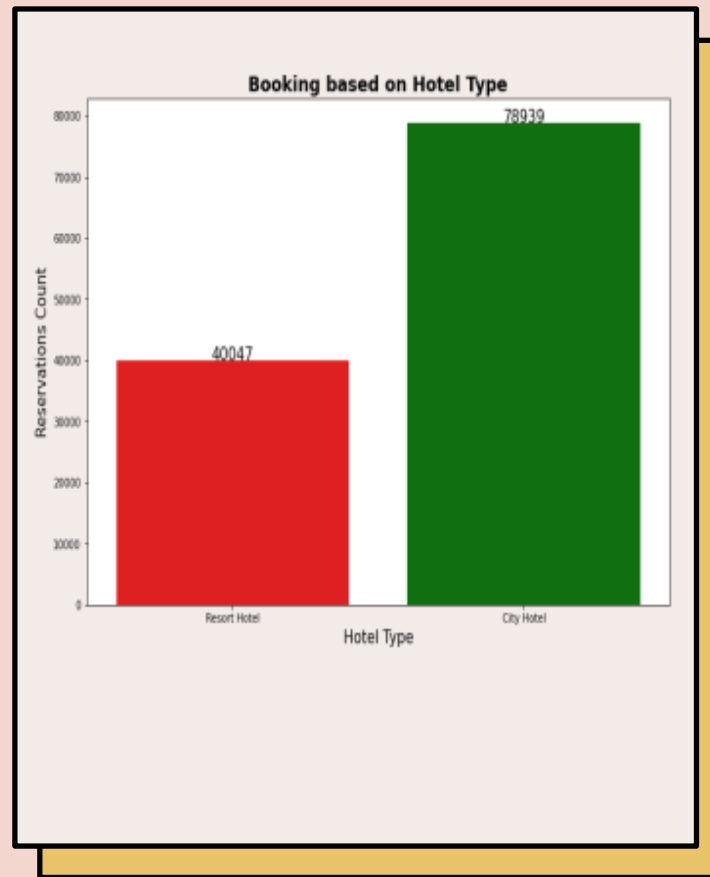
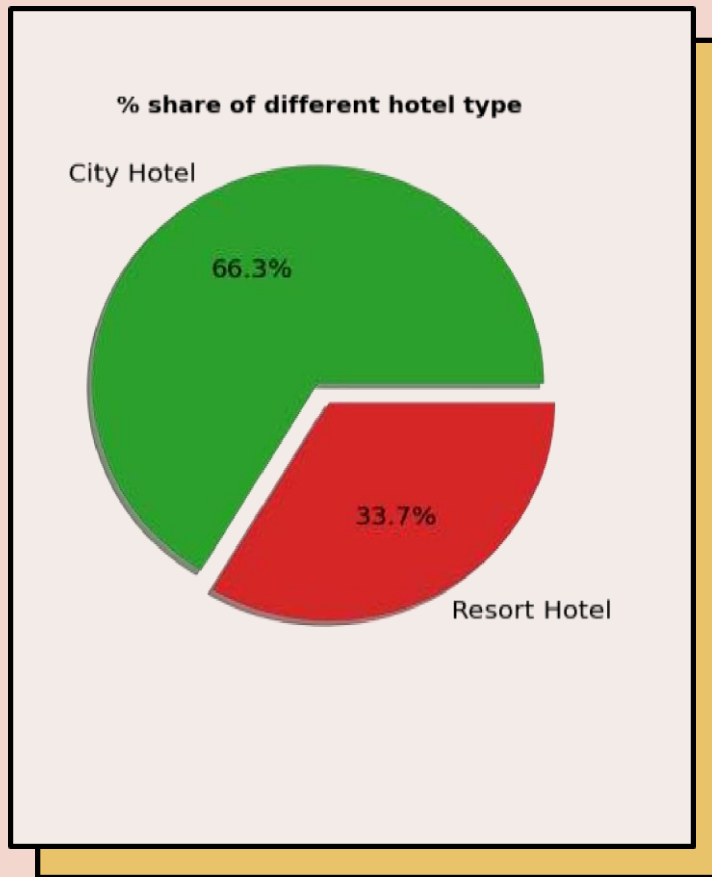
- kids = babies + children
- total stays = Stays in weekend nights + Stays in week nights
- Revenue = total stays(Non cancelled guests)\*ADR

# Handling Outliers

Interquartile range:-

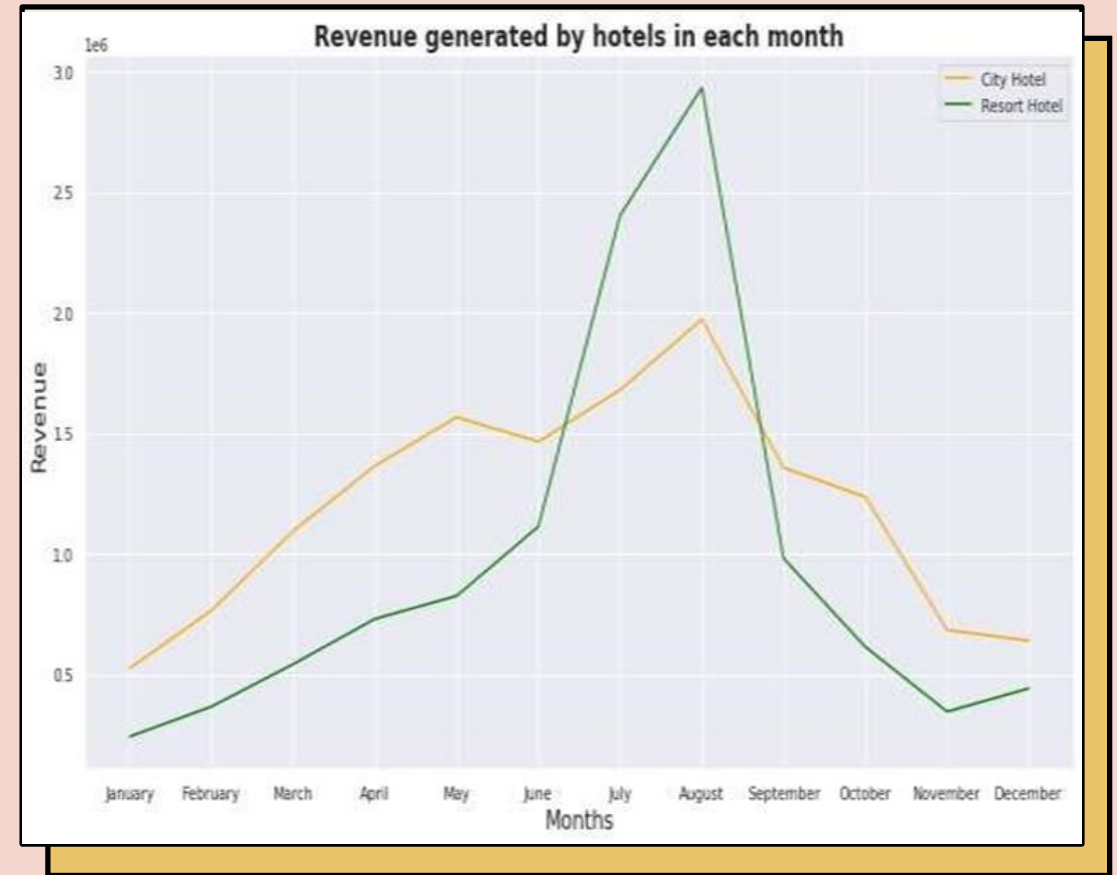


# Percentage share of different hotel type and their booking and revenue



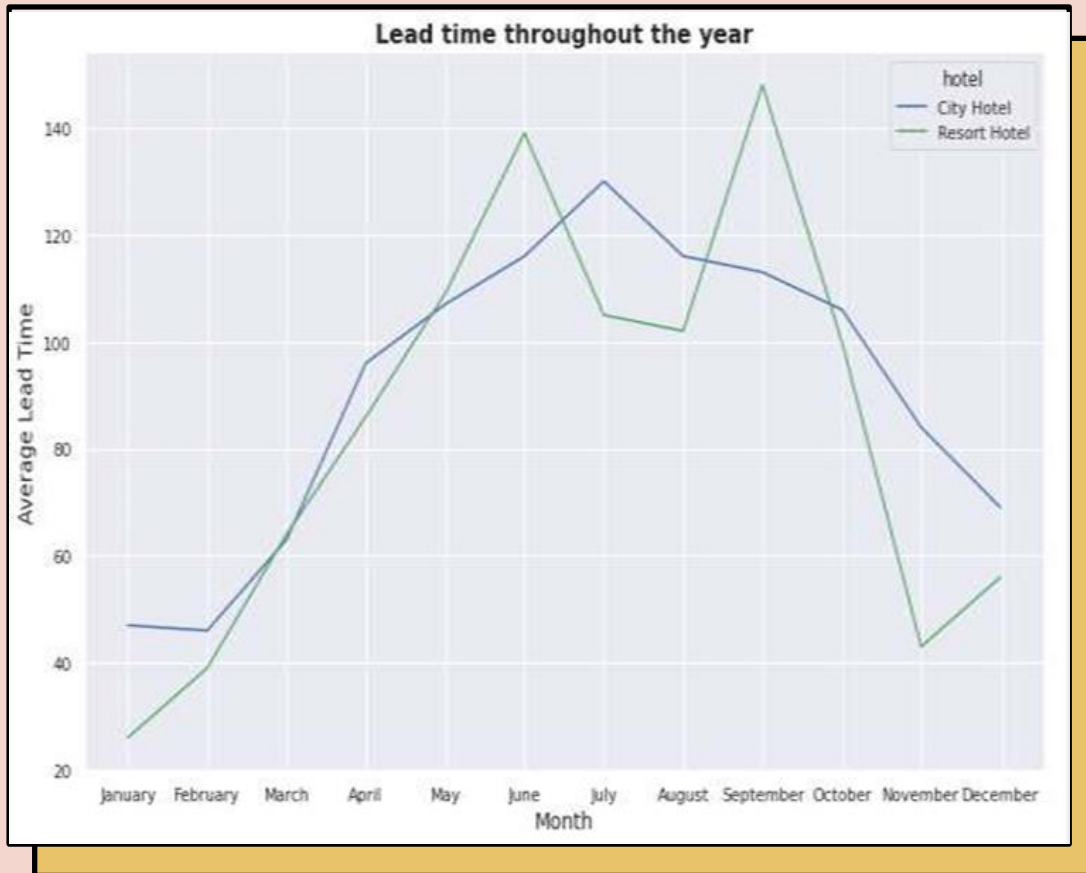
- City Hotels are most preferred hotel by guests.
- Bookings are double for City hotel but the revenue difference is not much in between them.
- This shows that Resort hotels are bit expensive as compared to City hotels.

# Analysis of ADR & Revenue for each month



- For Resort hotel-- ADR is increasing between May to September and then started falling down, so best time to book a resort hotel is from October to April as we are getting lower ADR.
- For City hotel--City hotels have nearly constant ADR from April to September and after that ADR start decreasing, so the best time to book a City hotel is from October to March.
- Resort hotels and City hotels both are getting higher revenue between June to September. This is also because at same time ADR is also high for both type of hotel as shown in first graph. Hence this period is best for hotel to generate more revenue.

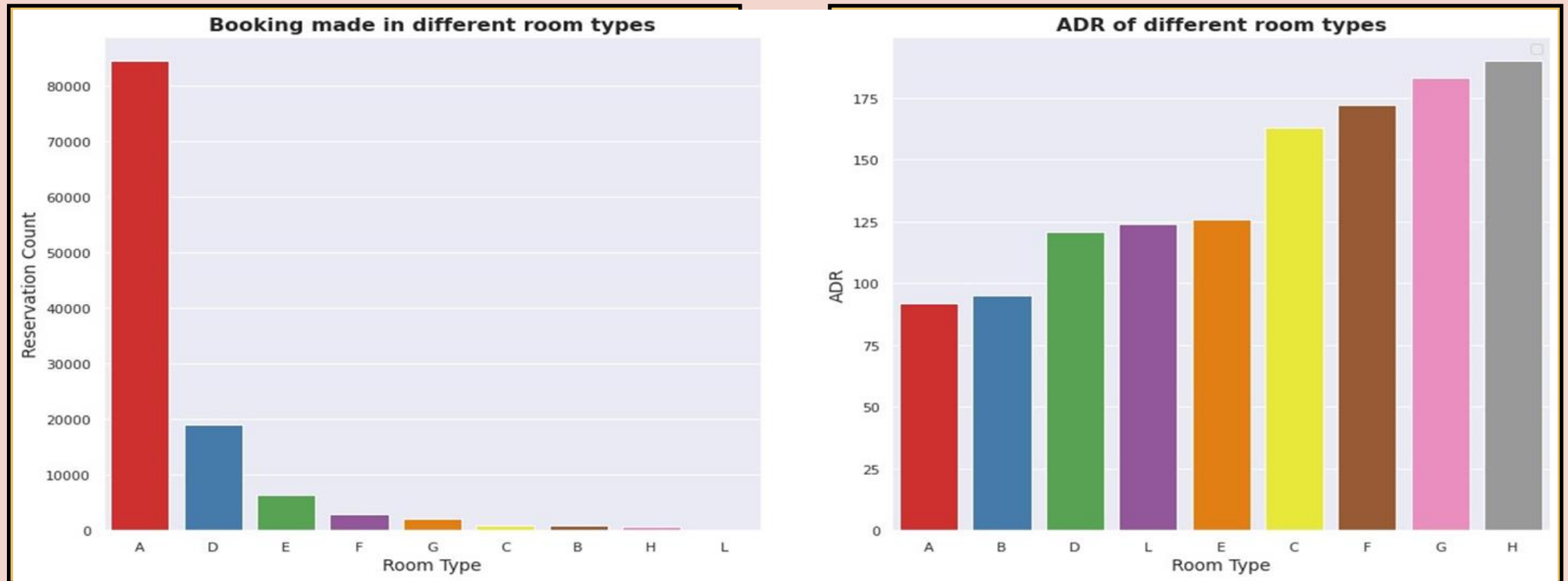
# Comparative study of reservation based on lead time and bookings in each month



From the above graph we can see in city hotels there is a peak in lead time from April to July and the booking is high in August. And in resort hotels we can see two peaks in lead time, first in June and second in September and booking is high in July , August and October, so people usually book hotel 30-60 days in advance.

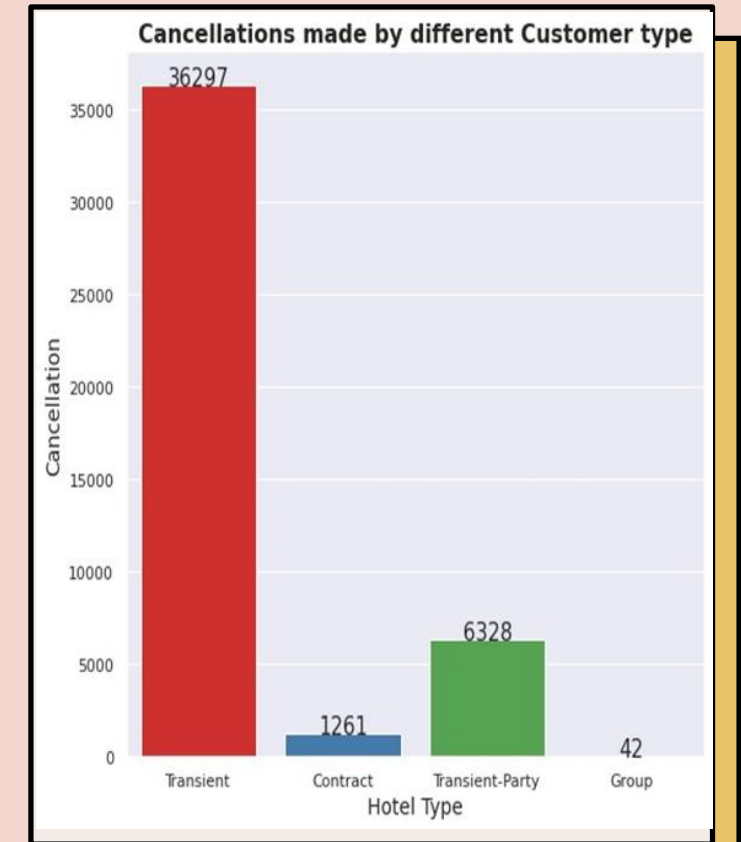
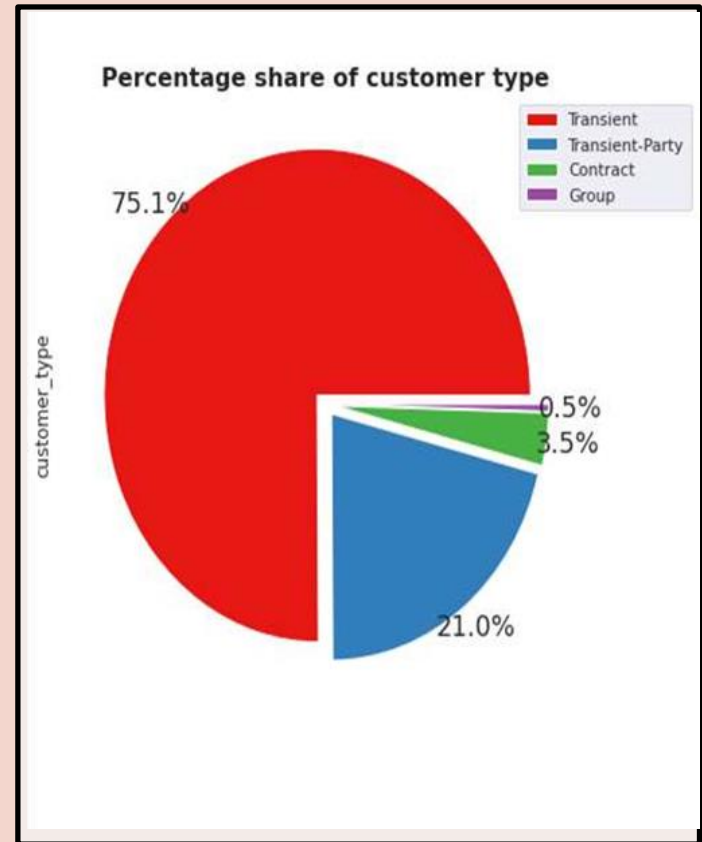
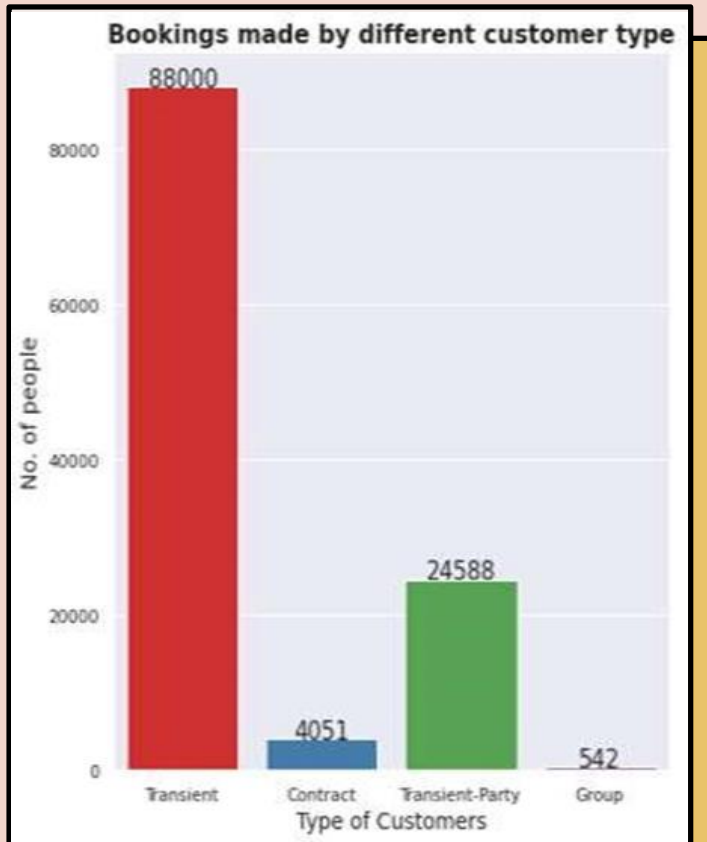


# Preference of different room type



Majority of people prefers room type-A which seems to be more economical for booking as it has the least ADR. And ADR is higher for room type-H so lesser booking are made.

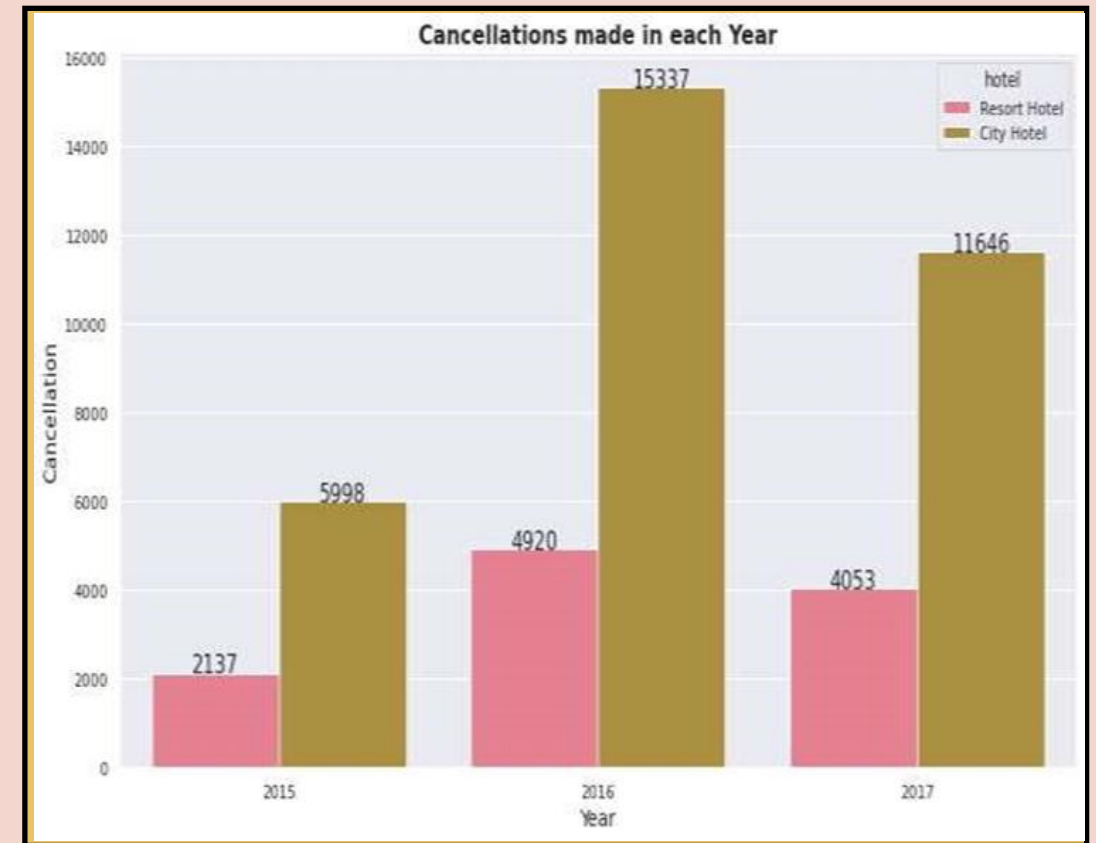
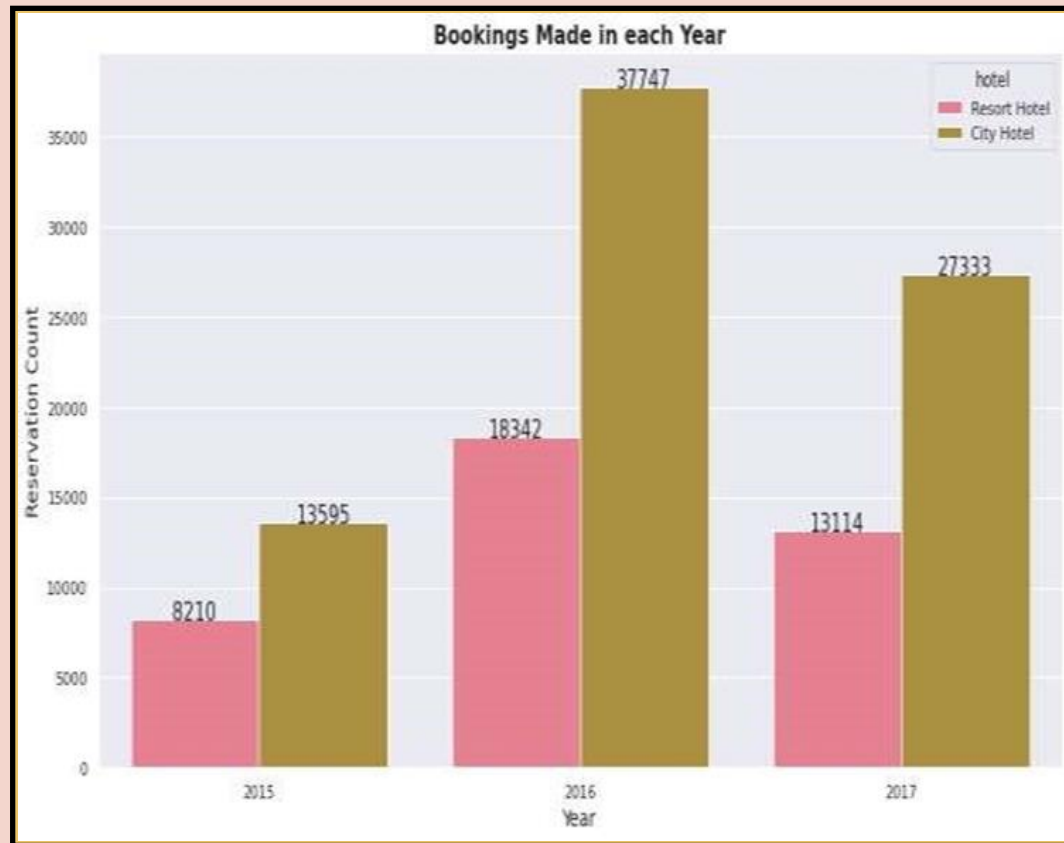
# Majority of bookings and cancellations done by which customer type



75% of bookings are done by Transient customer type.

Transient customers cancels more often but when people book in group it leads to lesser cancellations.

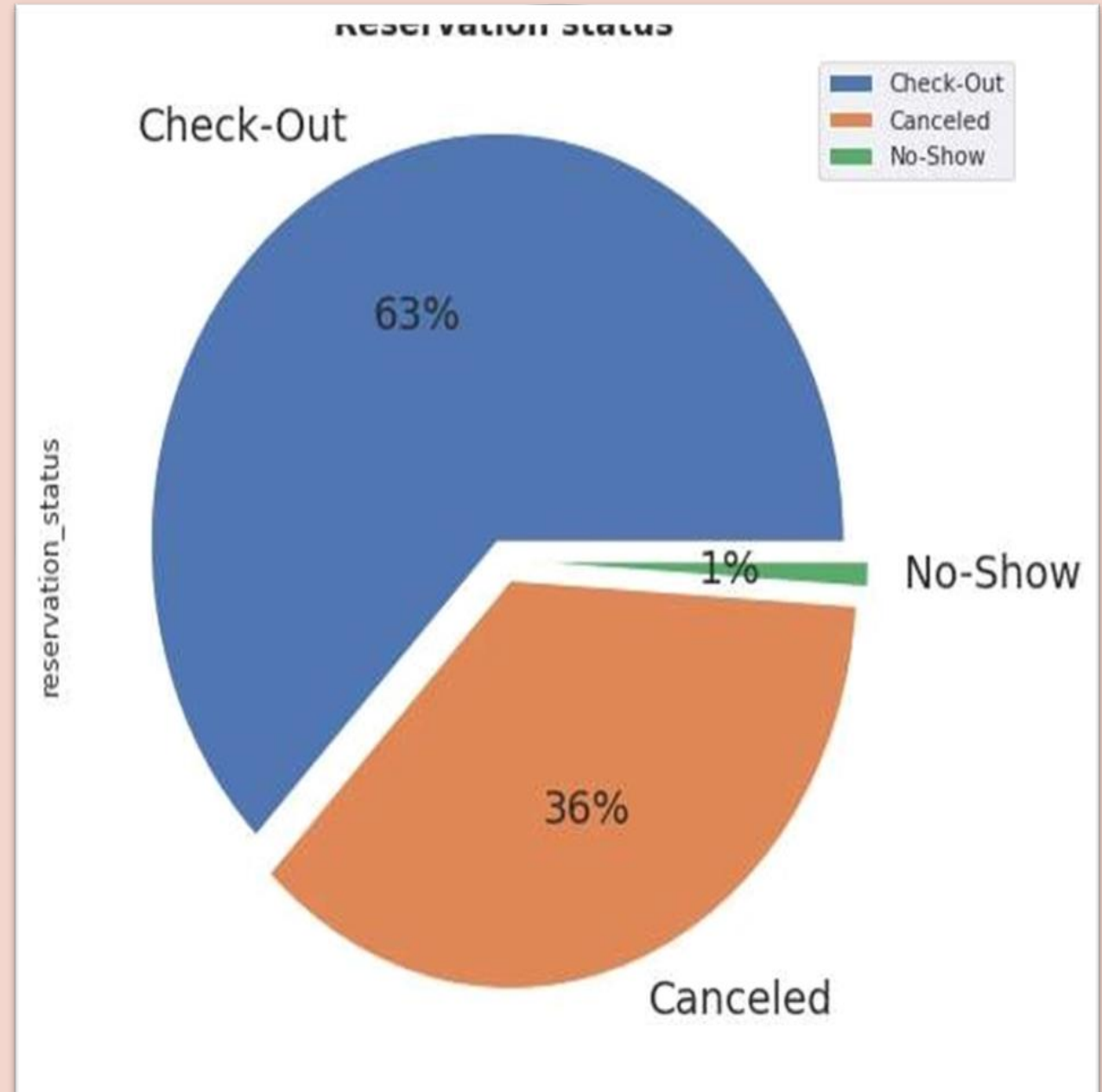
# Yearly bookings & cancellations



It seems that 2016 to be the year where the hotel bookings are highest. Every year 25-30% cancellations are received for resort hotels and 40-45% cancellations are received for city hotels.

# Analysis based on reservation status

Out of total no of reservations 63% actually show up, 36% got cancel , and only 1% reservation got No-show.



# Cancellation made in each hotel type

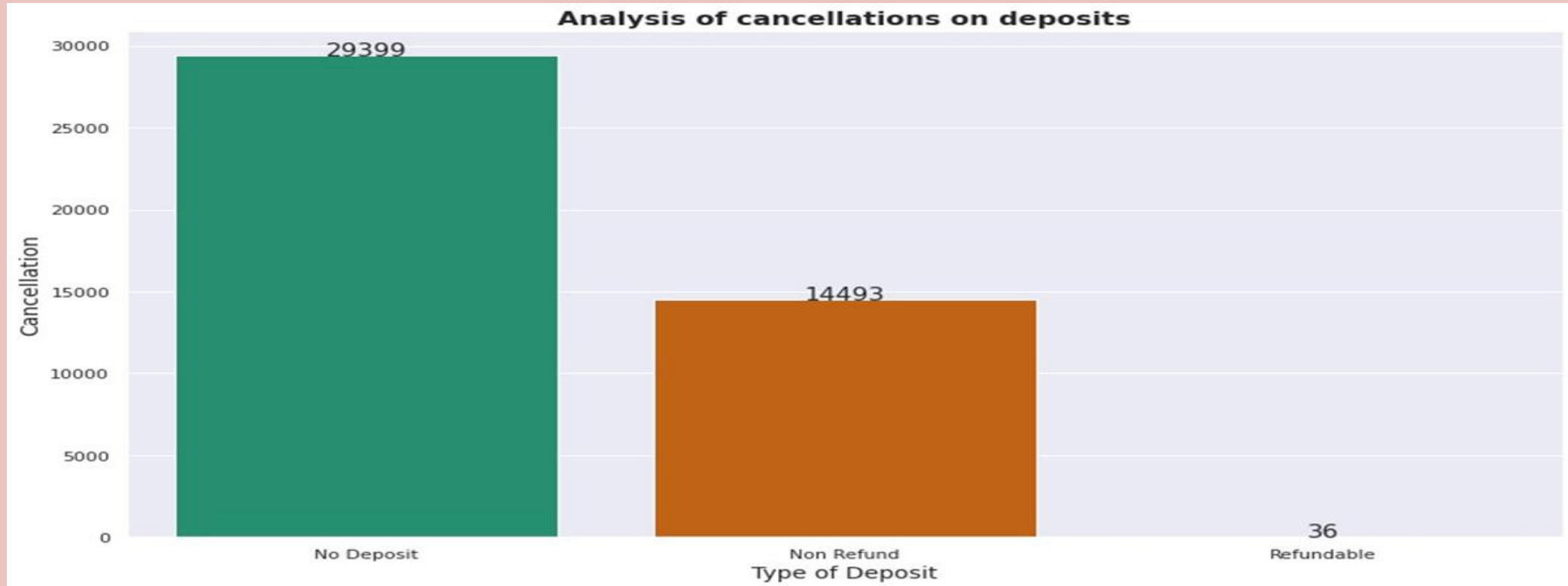




Cancellation is less in Resort hotels  
as compared to City hotels.

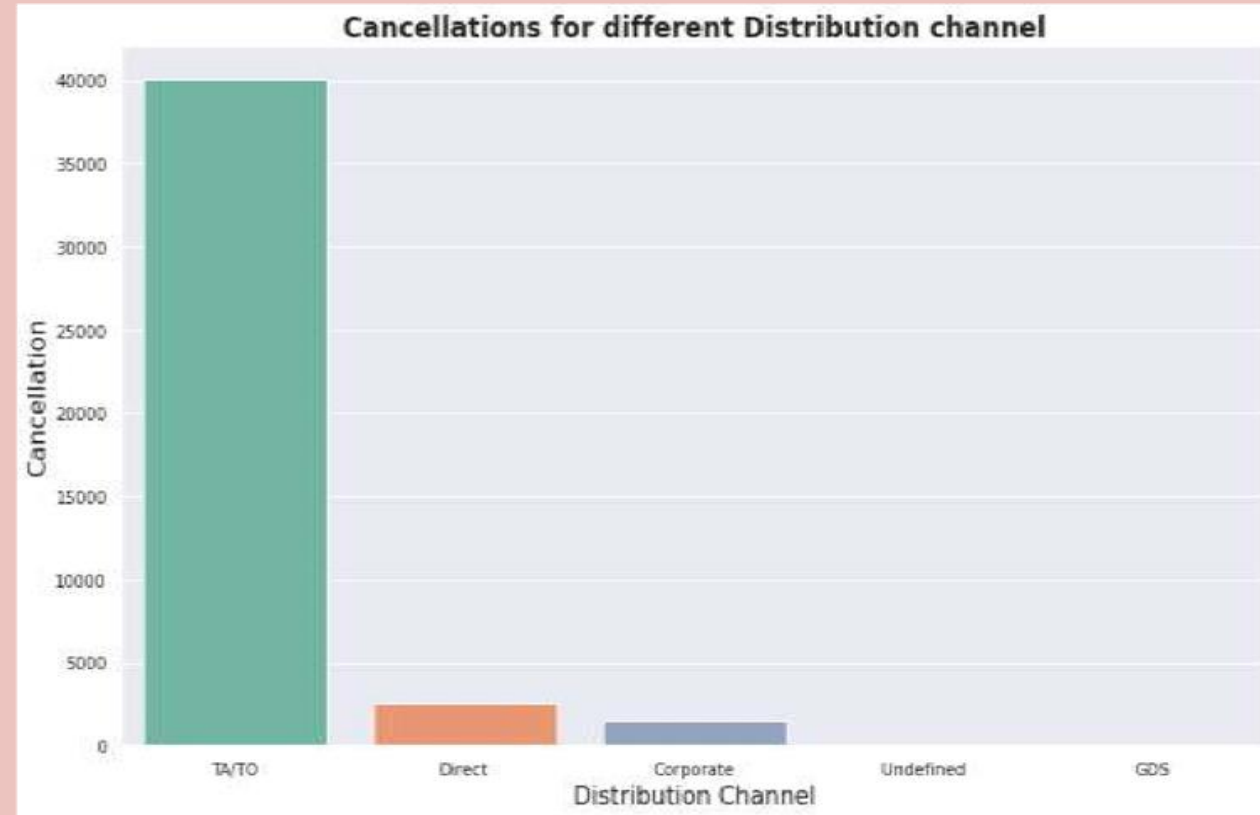
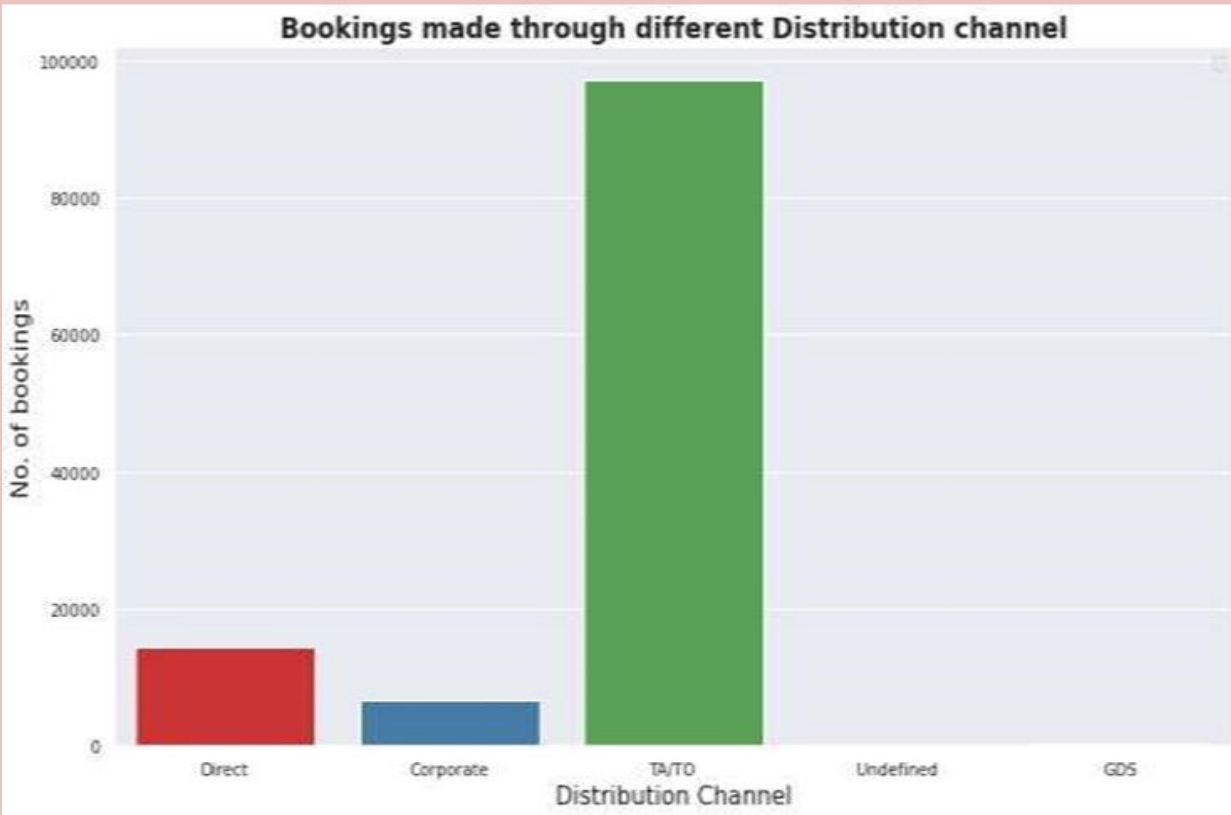
Cancellation rate is 75% in City  
hotels as compared to 25% in Resort  
hotels.

# Analysis of cancellations on deposits



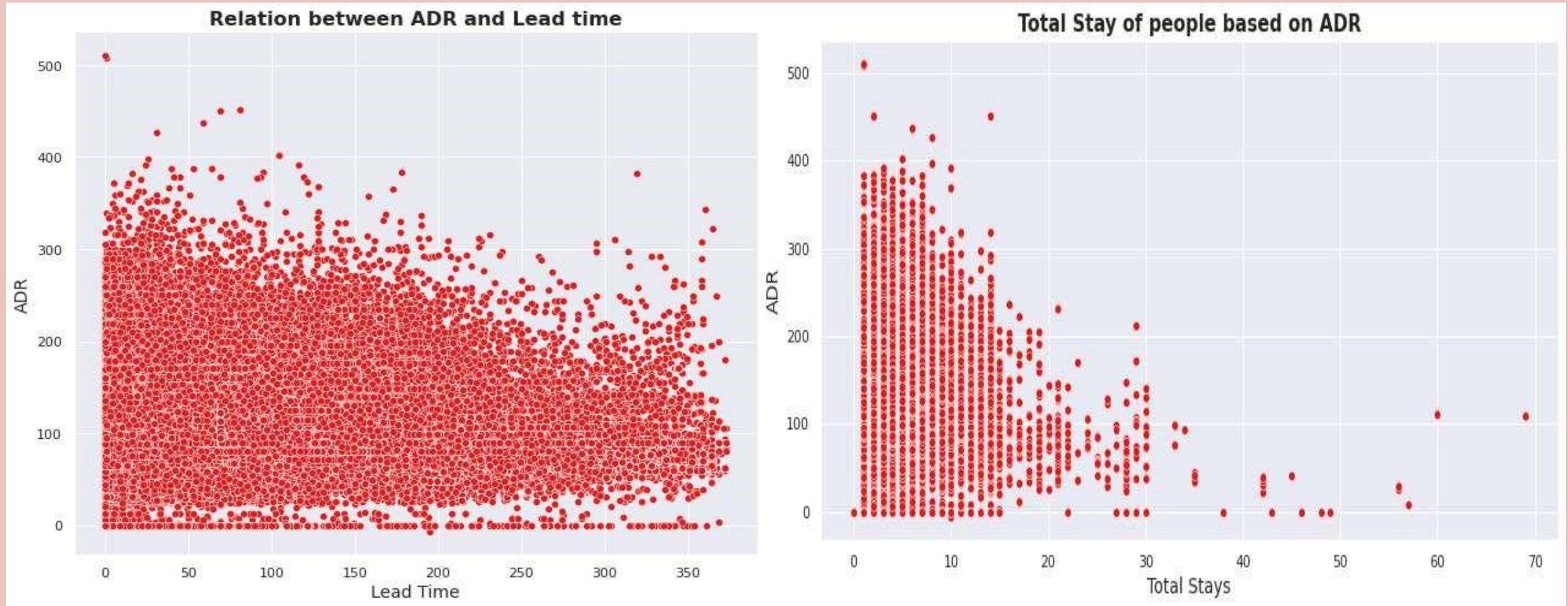
Chances of cancellation is high when there is no deposits taken by hotels. So minimum deposits should be taken by hotels to decrease the rate of cancellation.

# Analysis of Distribution channels



Majority of the bookings and cancellations are made through indirect channels i.e Travel agencies and Tour Operators compared to direct distribution channel.

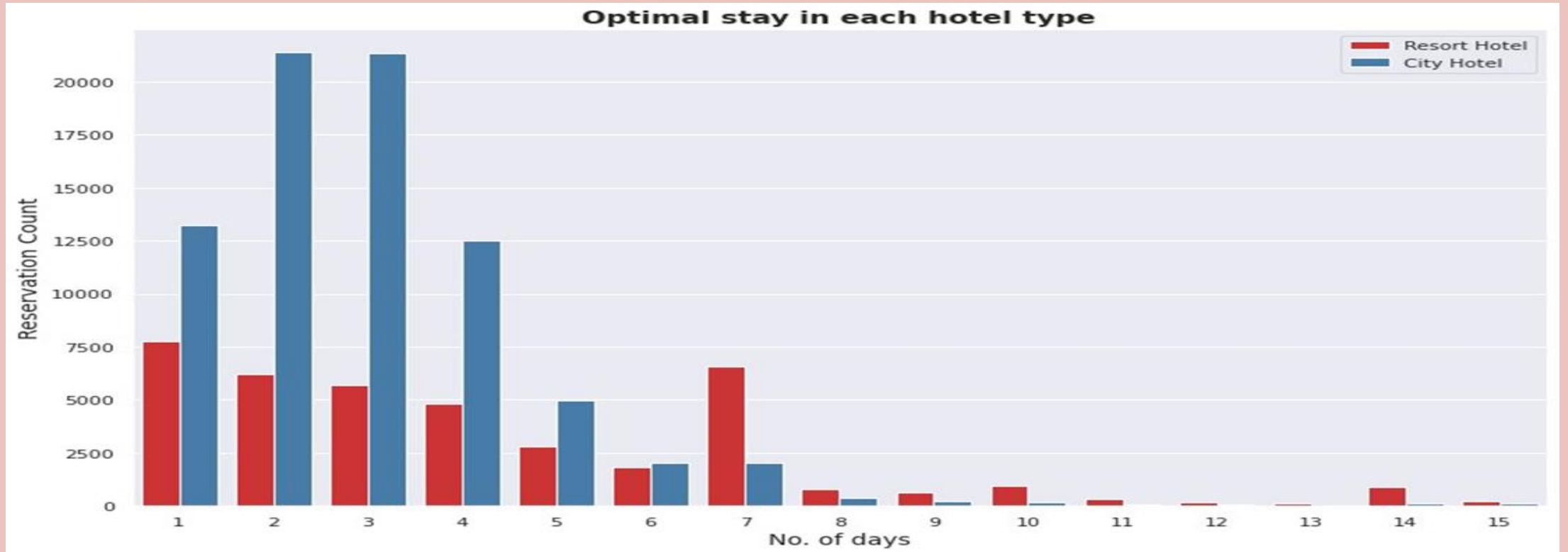
# Study of ADR on lead time and total stay



Here we can see that as the lead time increases ADR decreases. This means if a customer book a hotel in advance, he can get a better deal.

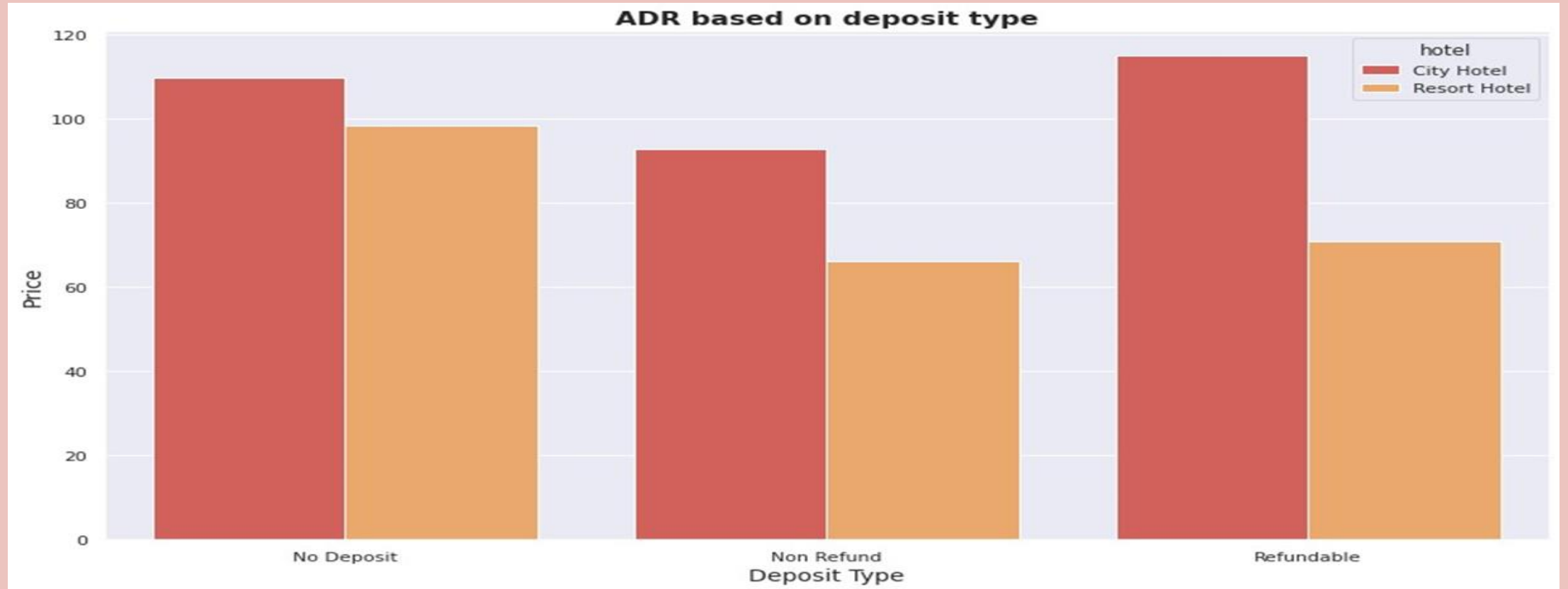
From the scatter plot we can see that as length of total stay increases as the adr decreases. This means for longer stay, the better deal for customer can be finalized.

# Preferred stay period



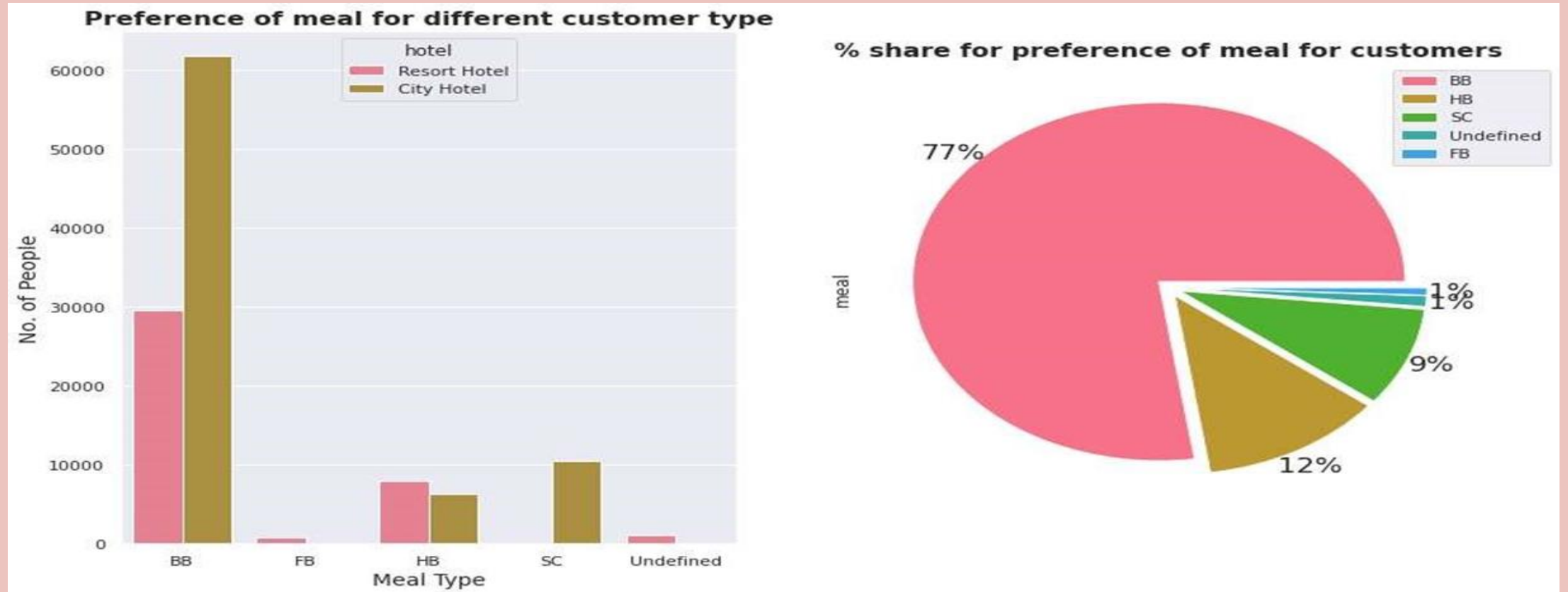
It can be stated that generally people prefer city hotels for shorter stay and resort hotels for longer stay. People prefer to stay in a city hotel for 1 to 4 days and in a resort hotel for 1 to 7 days.

# ADR based on deposit type



Hotels having refundable policy are having higher ADR where as hotel with no-refund policy are charging less.

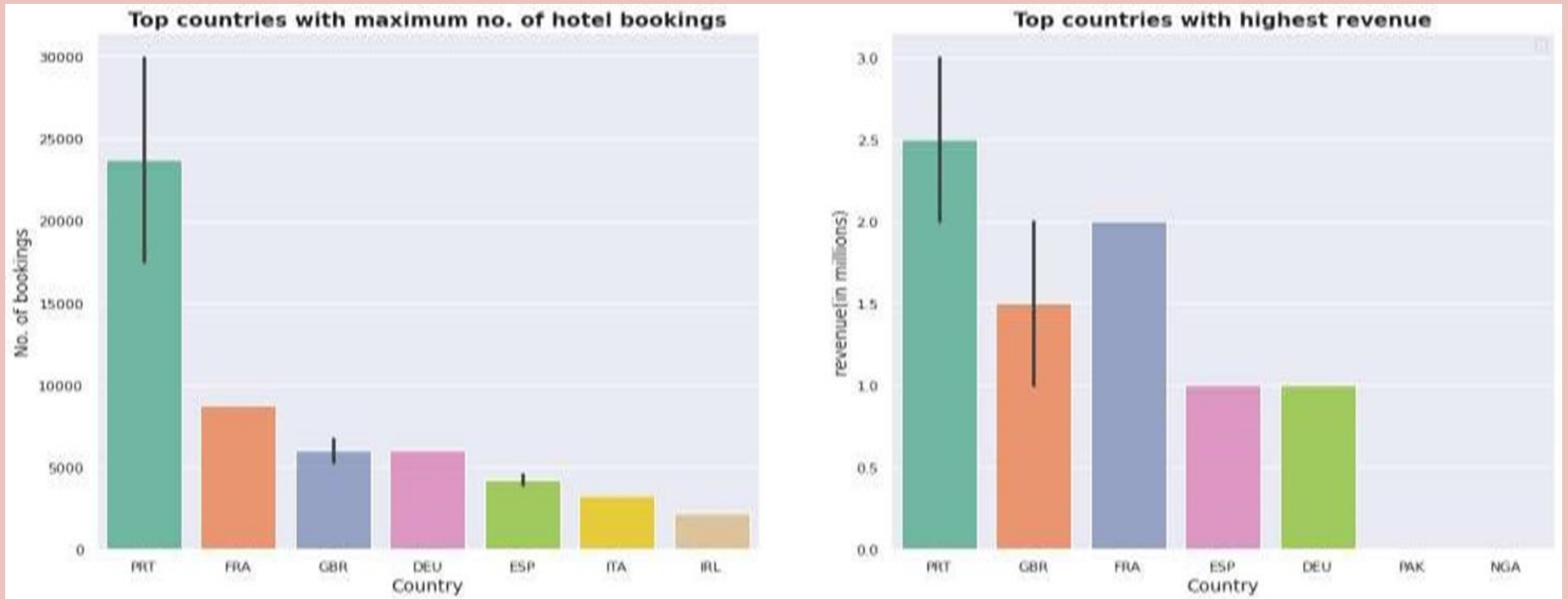
# Analysis based on Meal



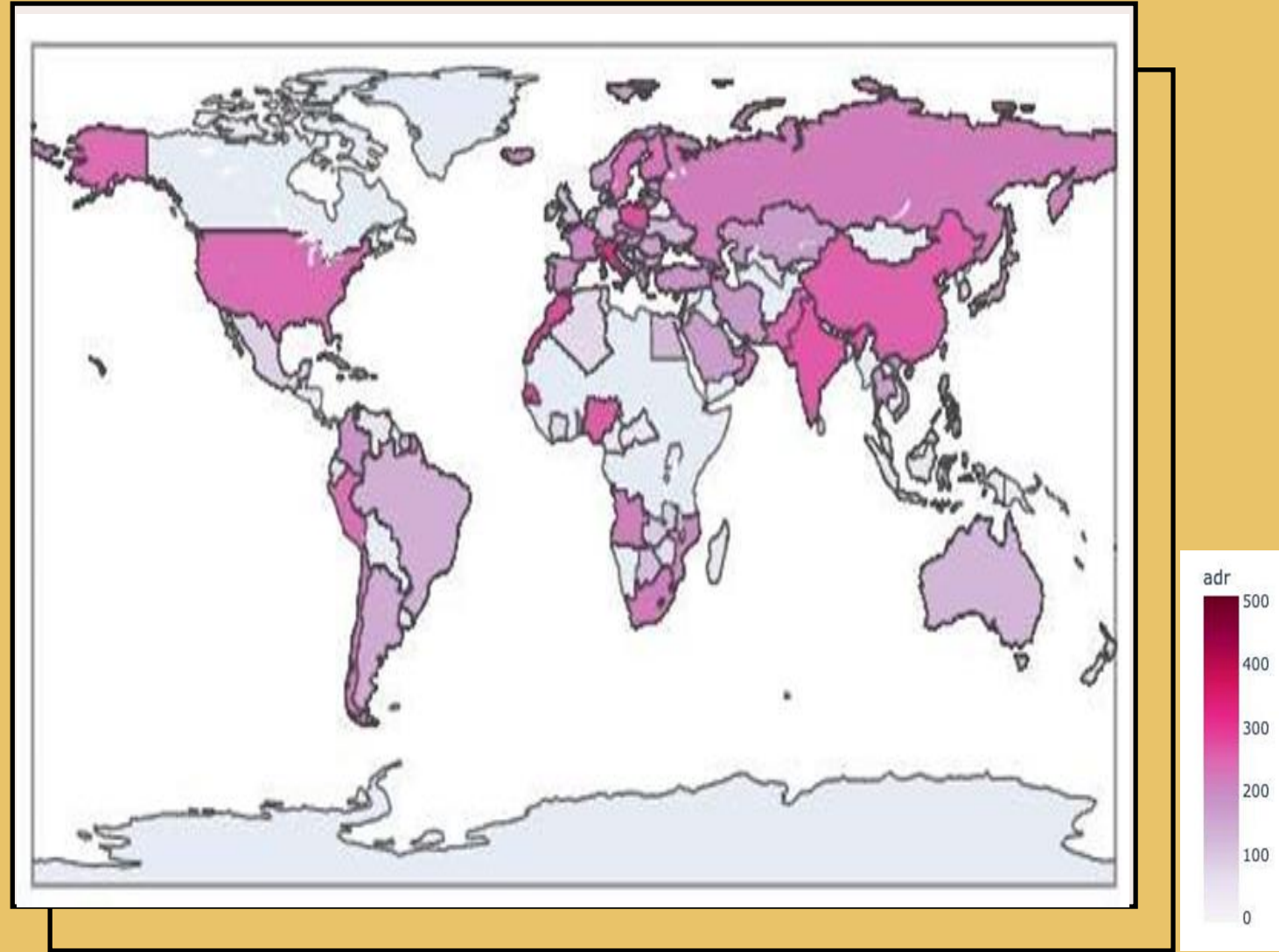
77% of the people prefer BB (bed & breakfast) meal type in both the hotel type.



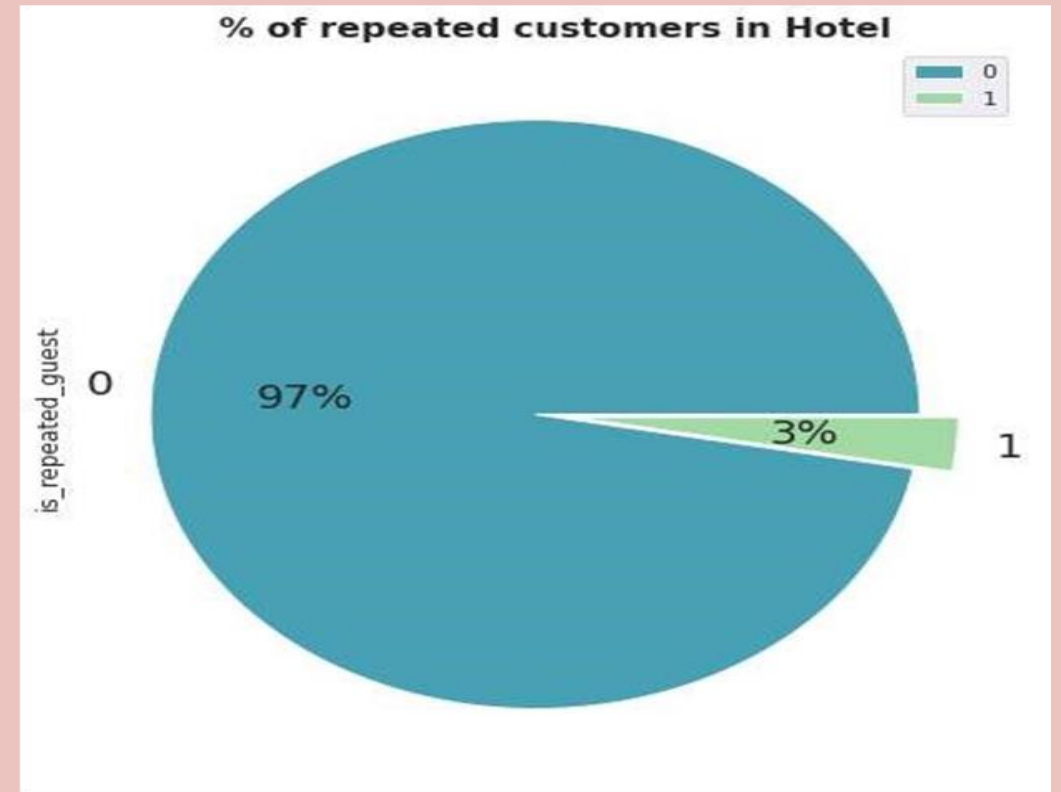
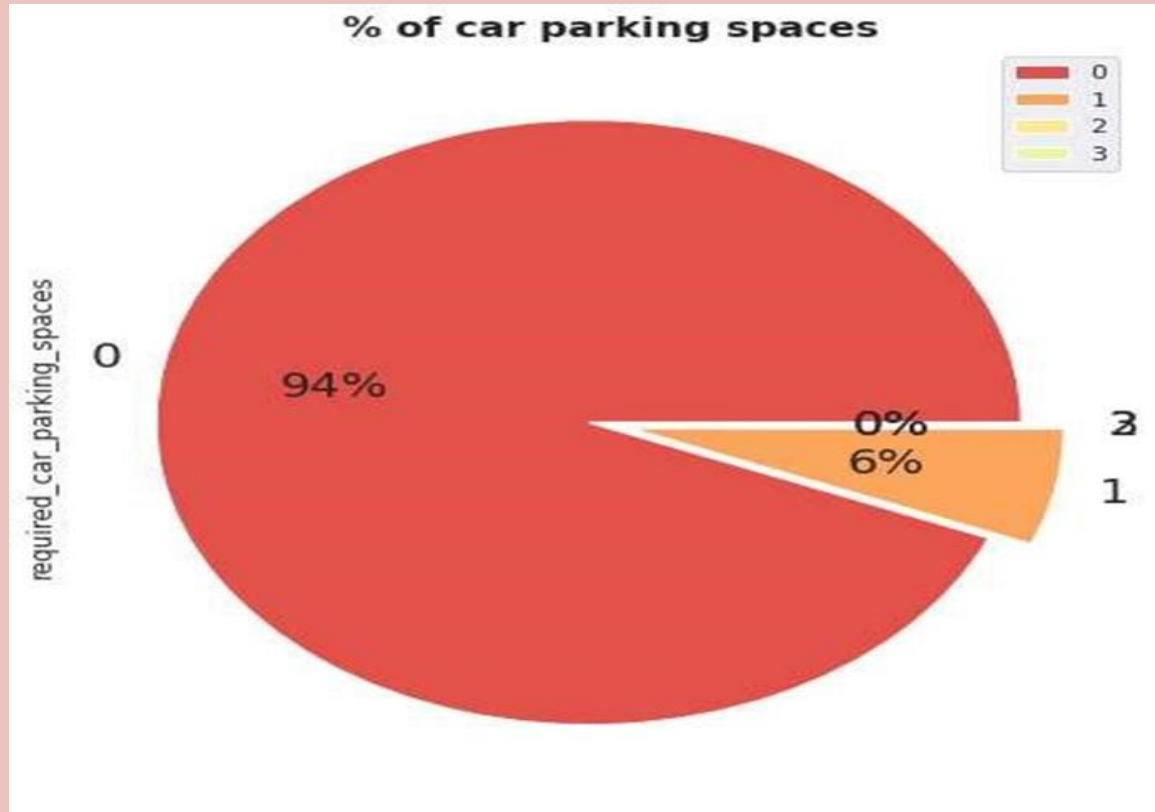
# Analysis based on Country



- We have a huge number of visitors from western Europe namely Portugal, France, UK and same countries generate highest revenue.
- European and South-East Asian countries have the highest ADR and African countries have lowest ADR.



# Car parking and Repeated customers insights



About 94% of people don't require the car Major share i.e Approx 97% are new customers. parking spaces while booking hotel.

# Observations

- Since the bookings are double for City hotel but revenue difference is less so Resort hotels are more expensive as compared to City hotels.
- People prefer city hotels for shorter stay and resort hotels for longer stay.
- Best time to book a resort hotel is from October to April and best time to book a City hotel is from November to March as we are getting lower ADR in this period.
- Maximum number of bookings and cancellations are done in the month of August. Both hotel types are getting higher revenue between June to September as ADR is high in this period.
- Hotel bookings and cancellations are highest in year 2016.
- As lead time and total stays increases ADR decreases. This means customers can get a better deal, when they book a hotel in advance, or stays for a longer duration.
  
- Hotels having refundable policy are expensive whereas hotel with no-refund policy are cheaper.
- Most bookings are done by transient customer types.
- Majority of the bookings and cancellations are made through Travel agencies (Online/Offline) and Tour Operators
- Cancellation is more in City hotels as compared to Resort hotels.

- Every year there is 25-30% cancellation for resort hotels and 40-45% cancellation for city hotels.
- Chances of cancellation is high when there are no deposits taken by hotels.
- 77% of the people prefer the BB (bed & breakfast) meal type in both the hotel types.
- About 94% of people don't require the car parking spaces while booking hotel.
- City hotels receives more special requests than Resort hotels. Most of the customers don't have special requests.

# Conclusion

1. Majority of people prefer A-room type so hotels should increase their numbers to get more revenue.
2. Chances of cancellation is high when there are no deposits taken by hotels, so hotels should take minimum deposits to minimize the rate of cancellation.
3. Transient customers cancels more often but when people book in groups it leads to lesser cancellations, hence hotels should provide some offers focusing transient customers to decrease cancellations.
4. Maximum number of bookings are in the month May to August, so hotels should provide exciting deal to customers to increase their booking in off season.
5. As hotels are getting less repeated customers so management should take customer's feedback and improve the hotel facilities to increase the count of their repeated guests.
6. We have a huge number of visitors from western Europe namely Portugal, France, UK and same countries generate highest revenue. So marketing teams should target these countries to get more customers.



THANK  
YOU

A graphic featuring the words "THANK YOU" in large, white, sans-serif capital letters. The letters are arranged in two rows: "THANK" on top and "YOU" below it. Each letter is suspended by a thin white string that passes through a small circular hole at the top of the letter. The background is a solid, vibrant orange. The entire graphic is set against a dark teal background, which is framed by a thin white border on the right and bottom edges.