Capstone Project - 1
Exploratory Data Analysis (EDA) on Hotel Bookings

Arun Prasath R

Shabnam Bano

Roshan Patil

Rupali auti

Content

- About Dataset

- Problem Statement

- Data Understanding

- Data Cleaning

- Data validation and Publishing

- Conclusion

Why Hotel Booking Analysis ?

➢ Hospitality industry is big contributor to economic growth of any country

➢ With a consistently growing middle class and increasing disposable income , tourism and hospitality sector witness healthy growth.

➢ Our project we will be analyzing Hotel Booking data.

➢ This data set contains 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.

AI

Problem Statement :

- ➢ The main objective behind this project is to explore and analyze data
- ➢ To discover important factors that govern the bookings and give insights to hotel management , which can perform various campaigns to boost the business and performance.

AI

Data Understanding :

- Libraries used: pandas, Numpy, Matplotlib, seaborn, datetime

- Total Bookings: 119390

- Total Variables: 32

Data Cleaning :

- Checked for Null values

- Null values are presented in Variables such as children, country, agent and company.

- Finding Percentage of null values

- Addition and deletion of Null values

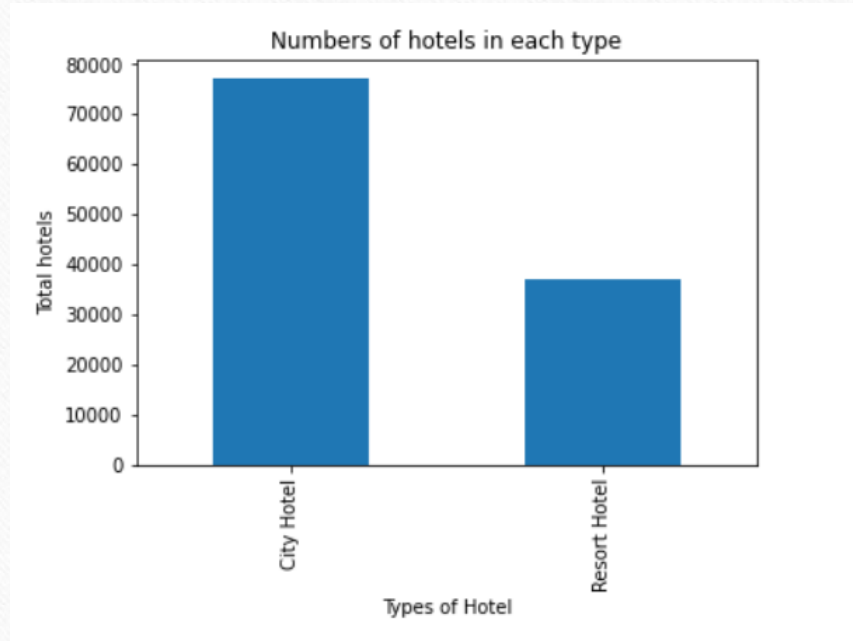- Converting to appropriate format

- Finding outliers
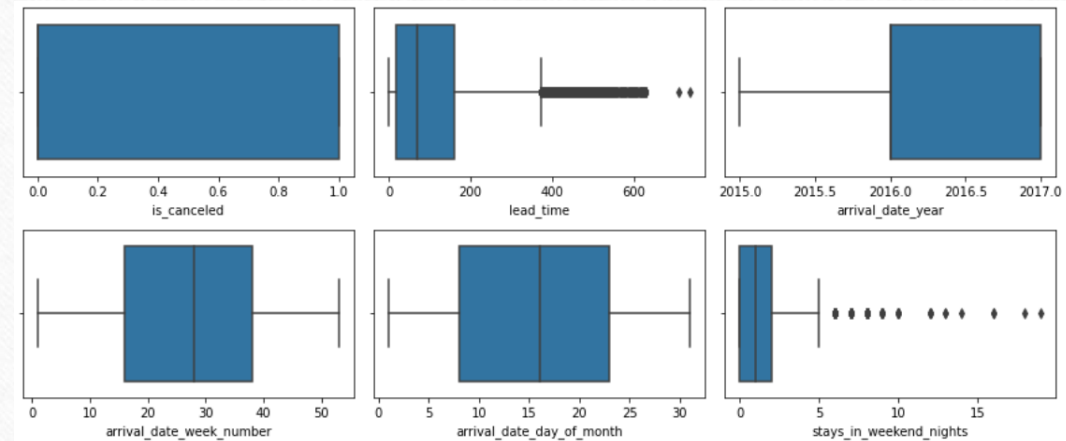
Data validation and Publishing

Meaning of Univariate, Bivariate & Multivariate Analysis of Data

- **Univariate Analysis** – In univariate analysis, one variable is analysed at a time.

- **Bivariate Analysis** – In bivariate analysis two variables are analysed together and examined for any possible association between them.

- **Multivariate Analysis** – In multivariate analysis, the concern is to analyse more than two variables at a time.
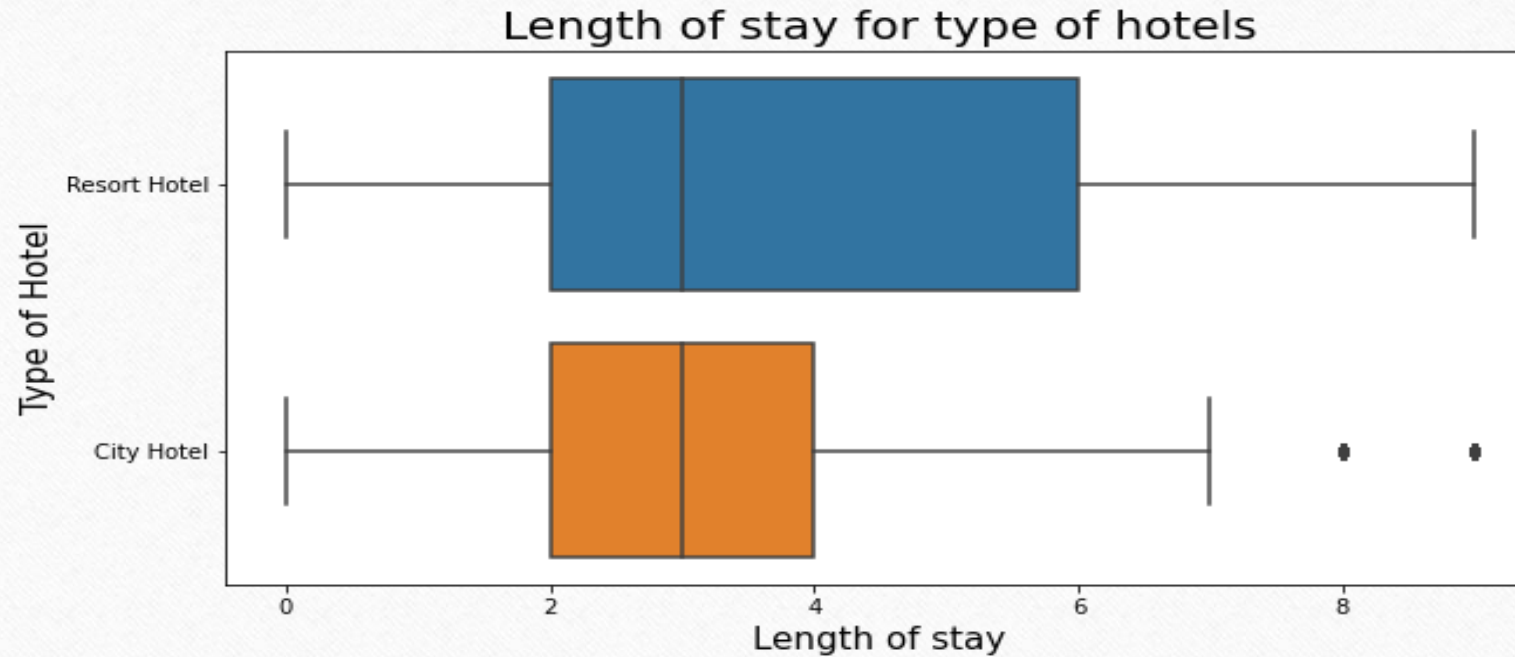
# Number of Bookings in each Hotel type   Number of Booking month-wise
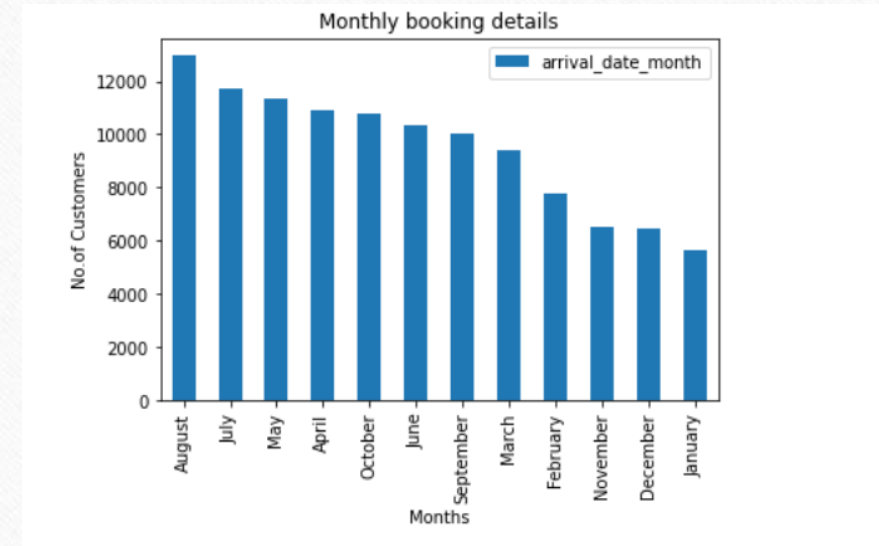
- Customer preferred City Hotel more than Resort Hotel
- Bookings in August is highest
- January found lowest numbers in booking counts.

Length of stay in each hotel type Box plot


Length of stay for type of hotels

- If we ignore the outliers, the maximum length of stay is more in resort type as resort is mostly used for vacation purpose.
- Median of both the hotels are approximately equal.

Average daily rates(adr) for each hotel type according to year
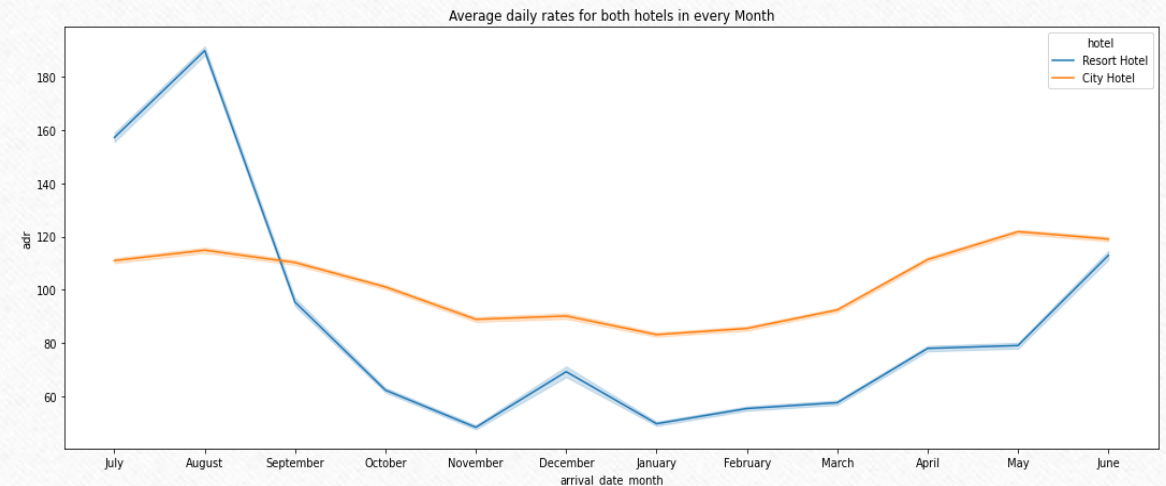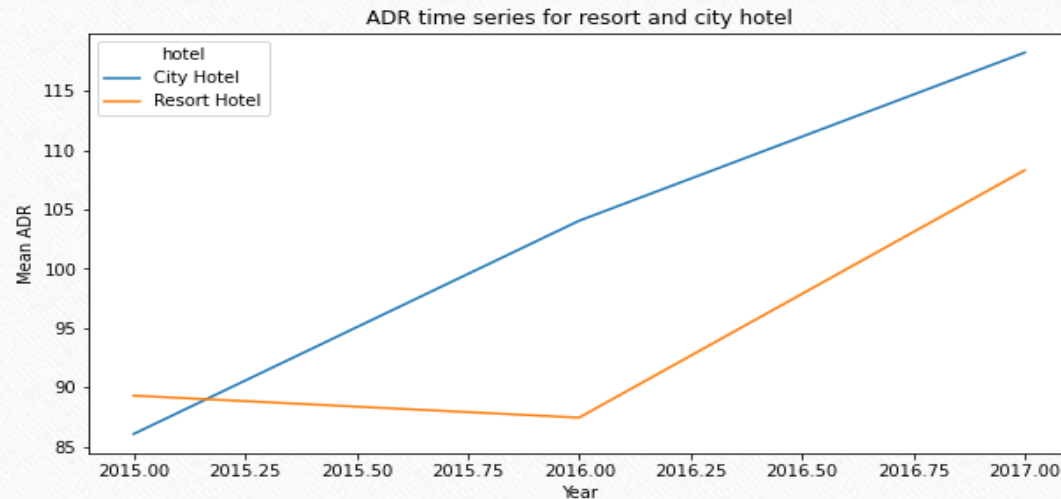
- City hotels has always higher (adr) than resort hotel.
- If the trend continues like that, resort hotel (adr) showing maximum increased inclination which means in few years resort hotel adr will cross the city hotel.
- In the month of July and till the last week of August Resort hotels received more "adr" than City hotel City hotel although dominating in 'adr' in remaining months of the year.

Stays in weekend and weekdays in hotels vs Bookings
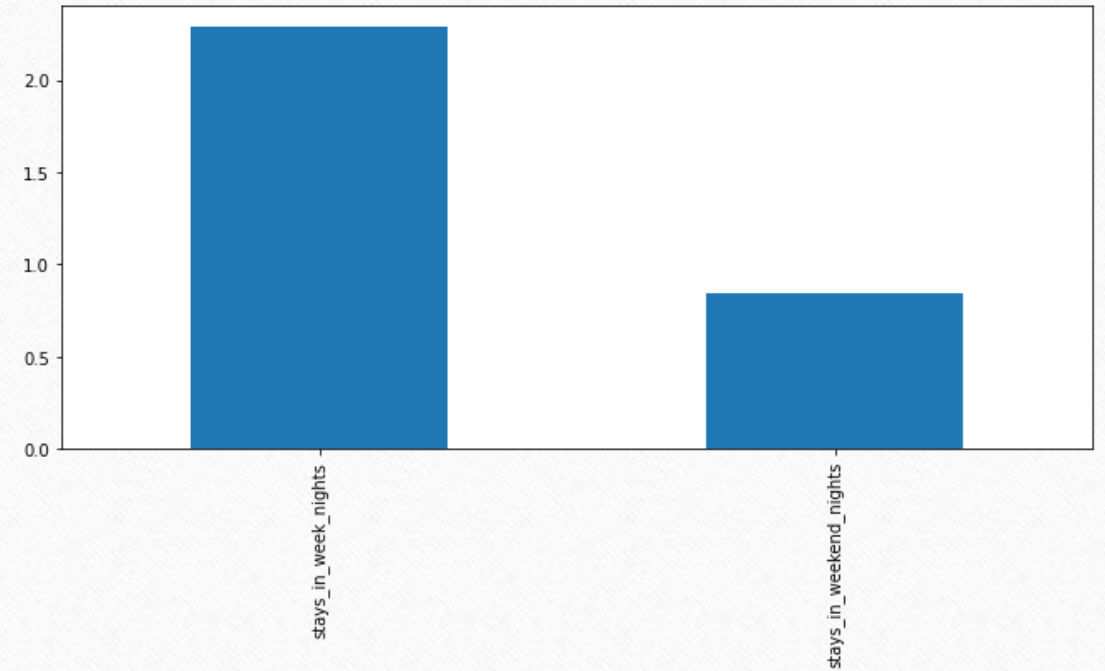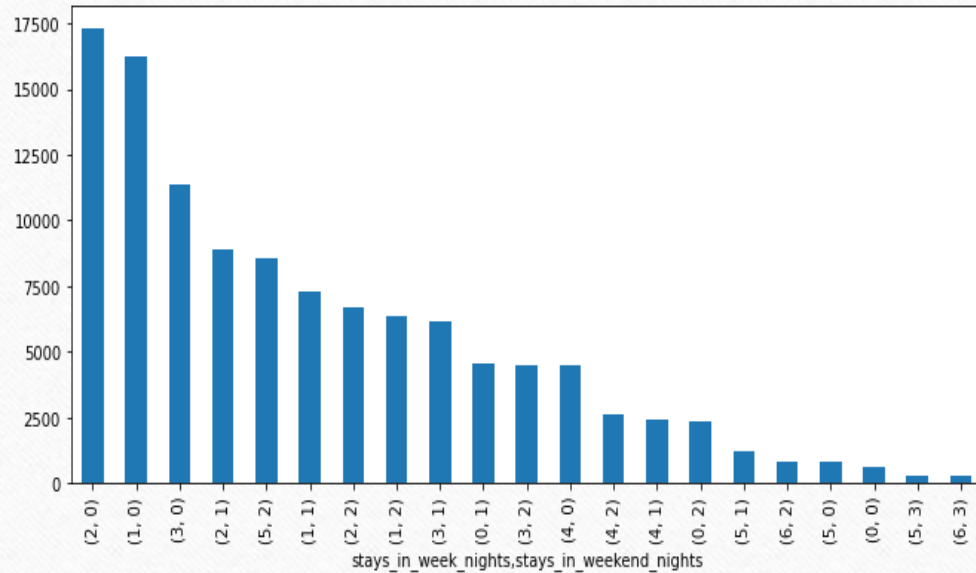
Observation:
- Maximum Booking done by customer for "2" weeknights stay and "0" weekend nights
- On average, customers booking for 2-3 week-nights stay and 1 weekend-nights stay

Count of adults , children ,babies in booking
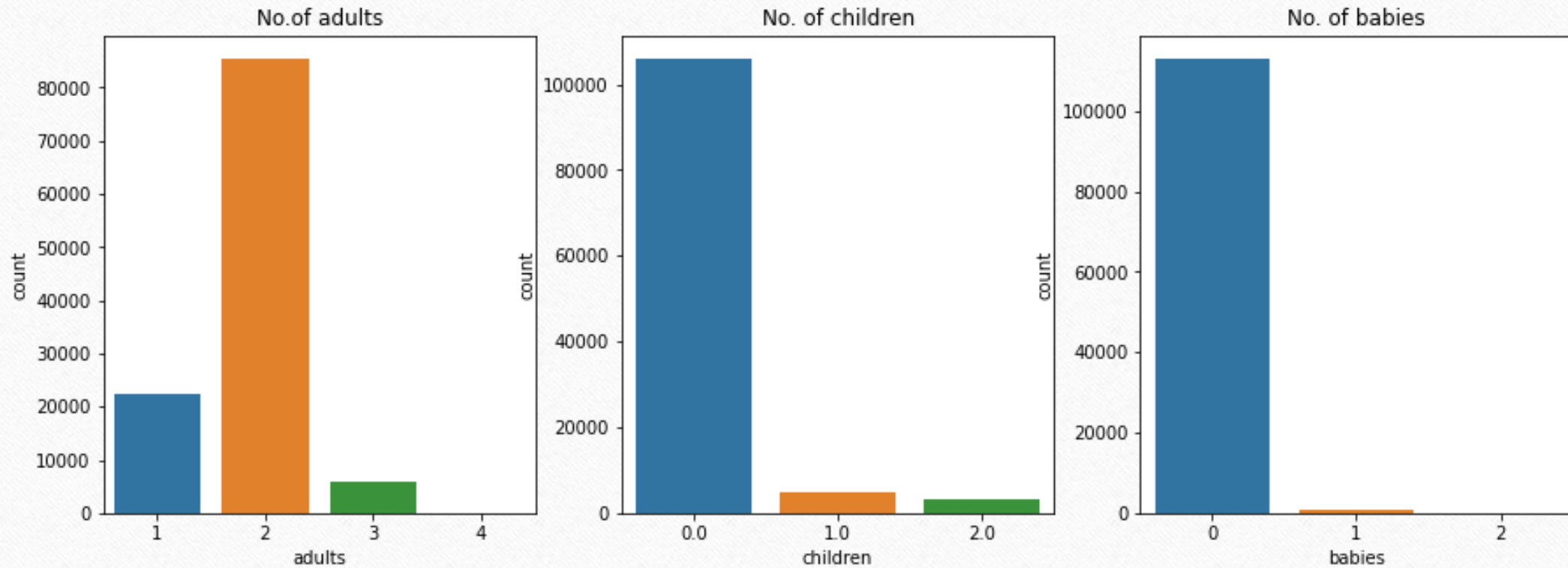
Observation:
- Bookings are mostly made for 2 adults with 1 children in combination

Preference of the meal by customer

Preferred Meal
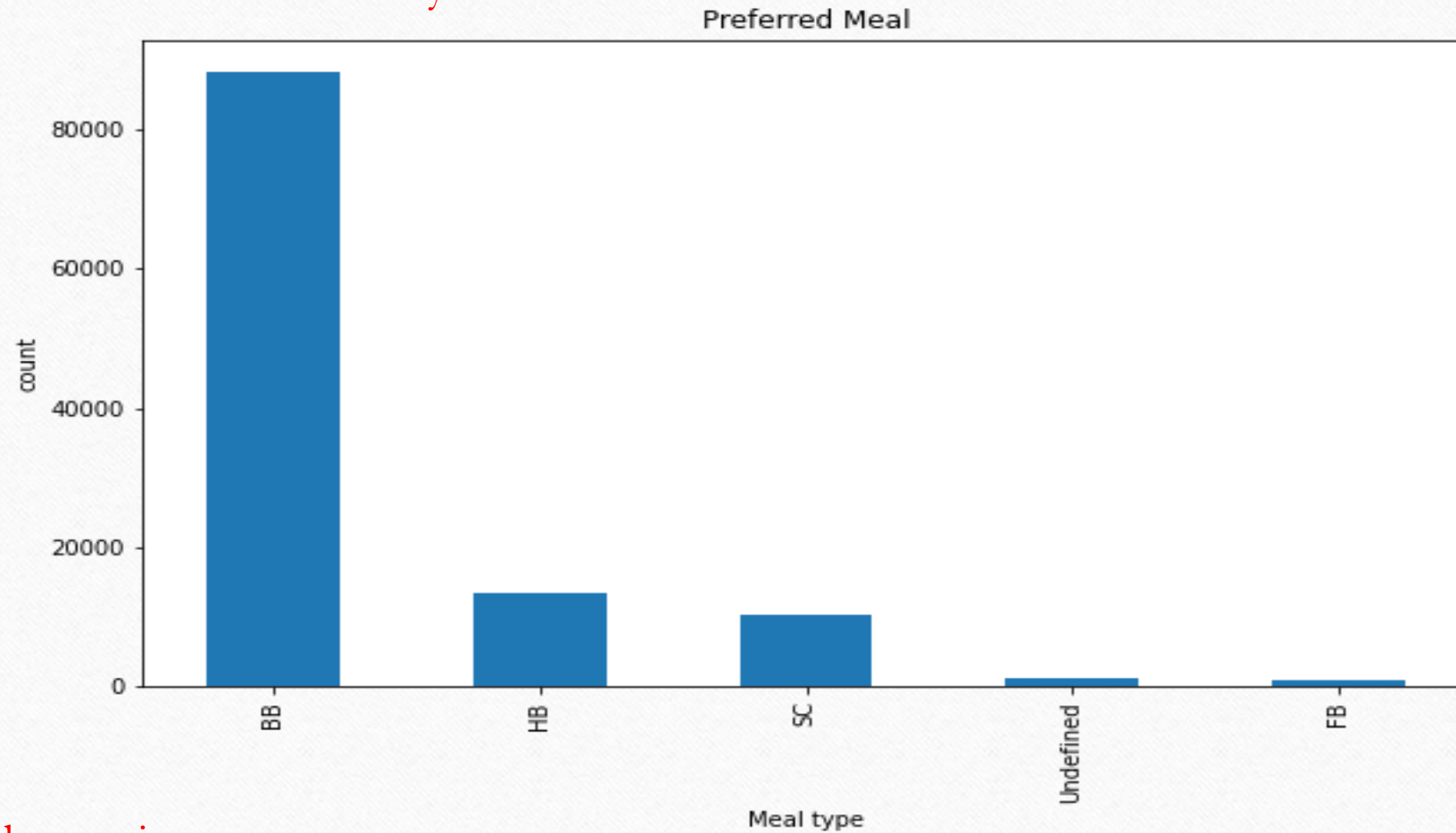
Observation:

- Customers of any type preferring "BB" Bed and breakfast type of meals.

Booking by Market-segment

Bookings by market-segment

Observation:
- Travel agency (TA) or Tour operator (TO) Plays vital role in Hotel booking
- Except "Direct bookings", all market-segment has more number of bookings in city hotel type.

Number of Weekdays booked vs market segment

No of weekdays Vs Markert segment
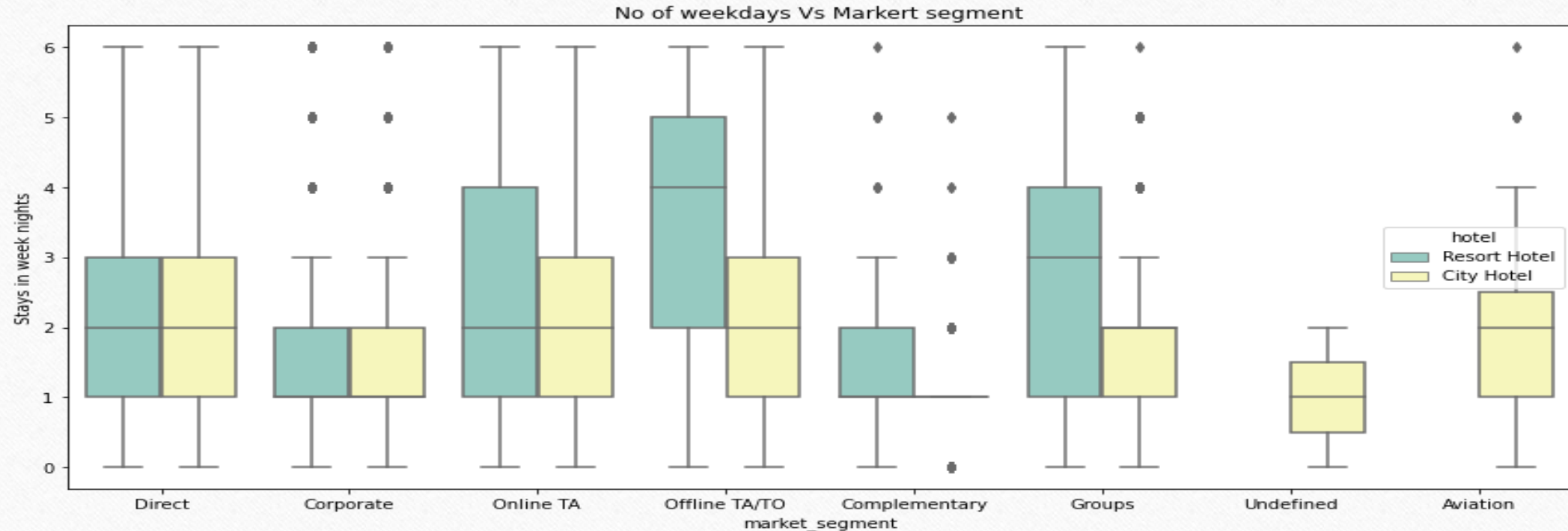
Observation:

- Customer from direct market segment staying in same range numbers of week nights.
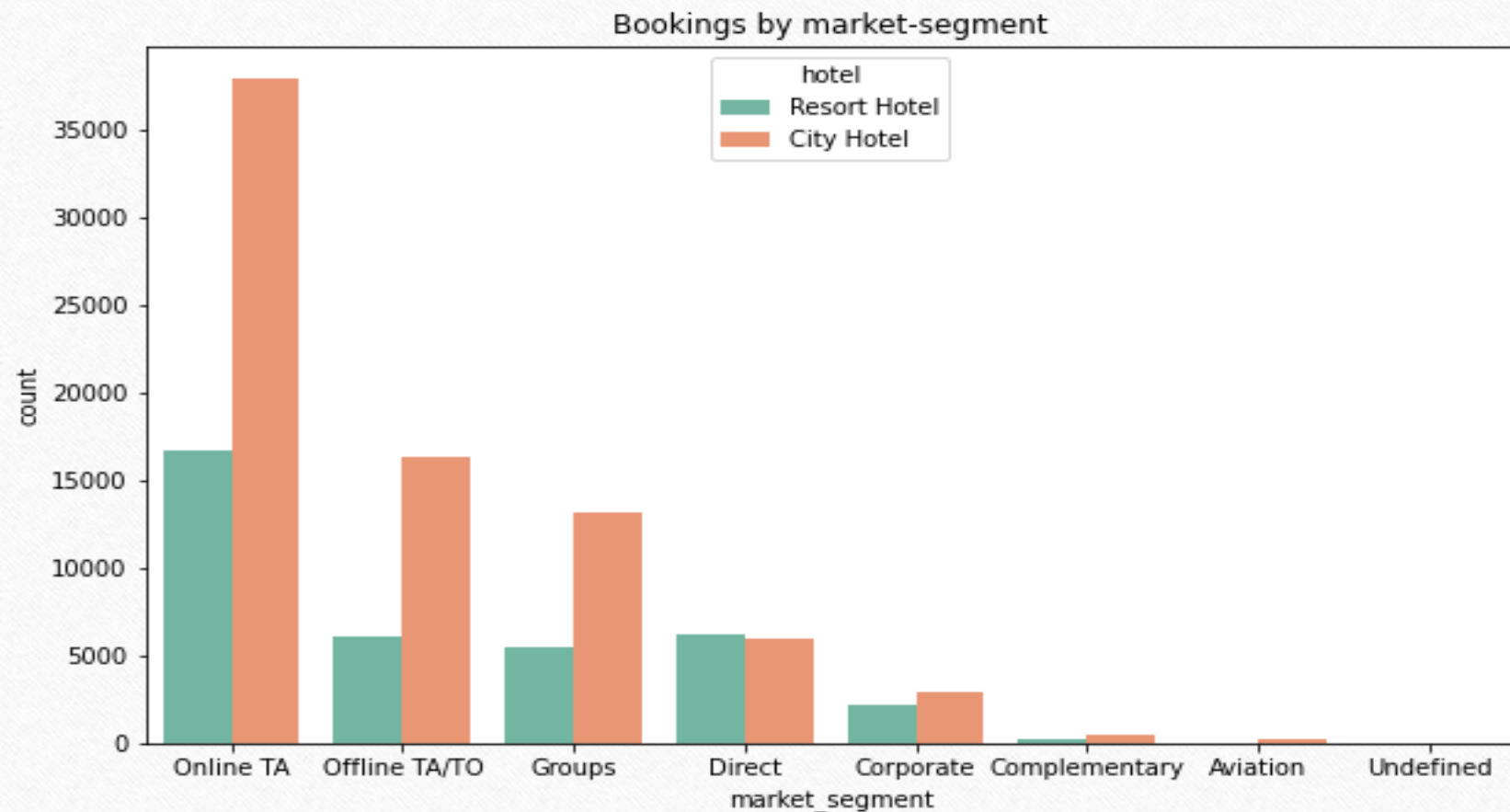- Offline TA/TO and Group market segment has some deviation over stays week-nights between Resort and City hotels
- Undefined and Aviation market segment customer had not shown interest in the Resort Hotel

Number of Weekend nights booked vs market segment

No of weekend night Vs Markert segment
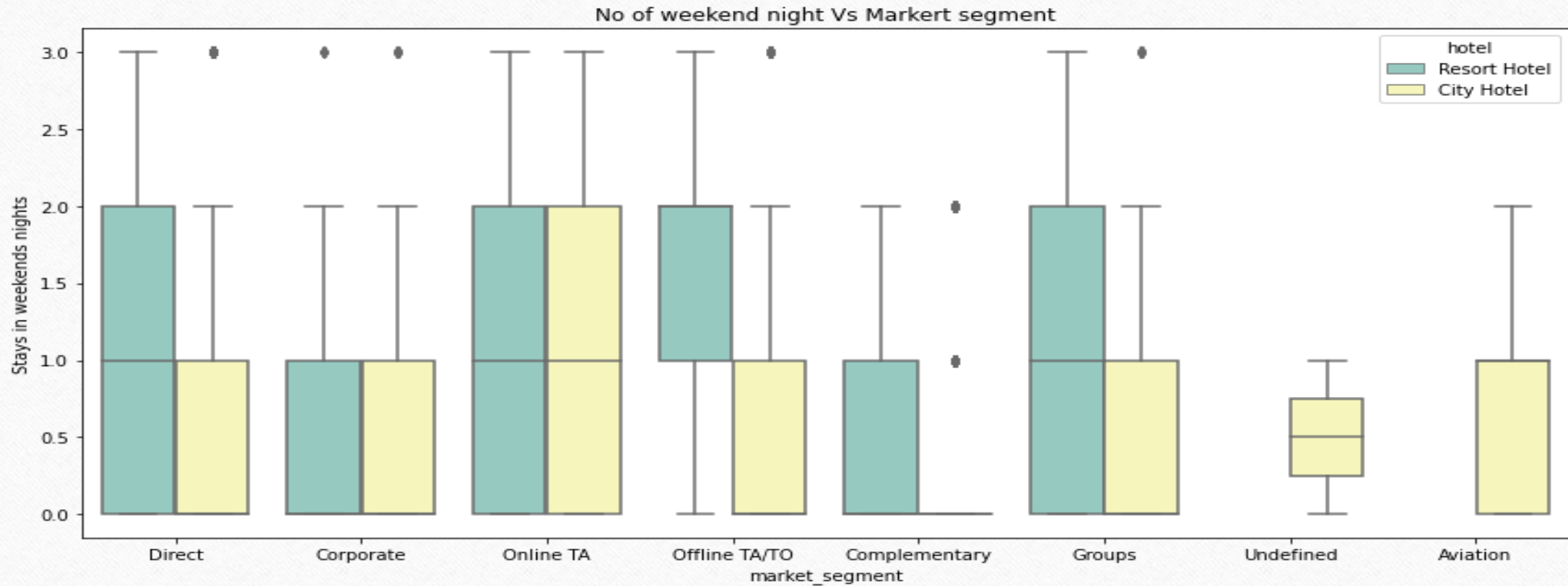
Observation:
- Direct market segment customer prefer to stay more weekend nights in the Resort Hotel type.
- Online TA customer equally preferring between Resort and City hotels.

Preference of Room types by customer

Preference of Room types by customer
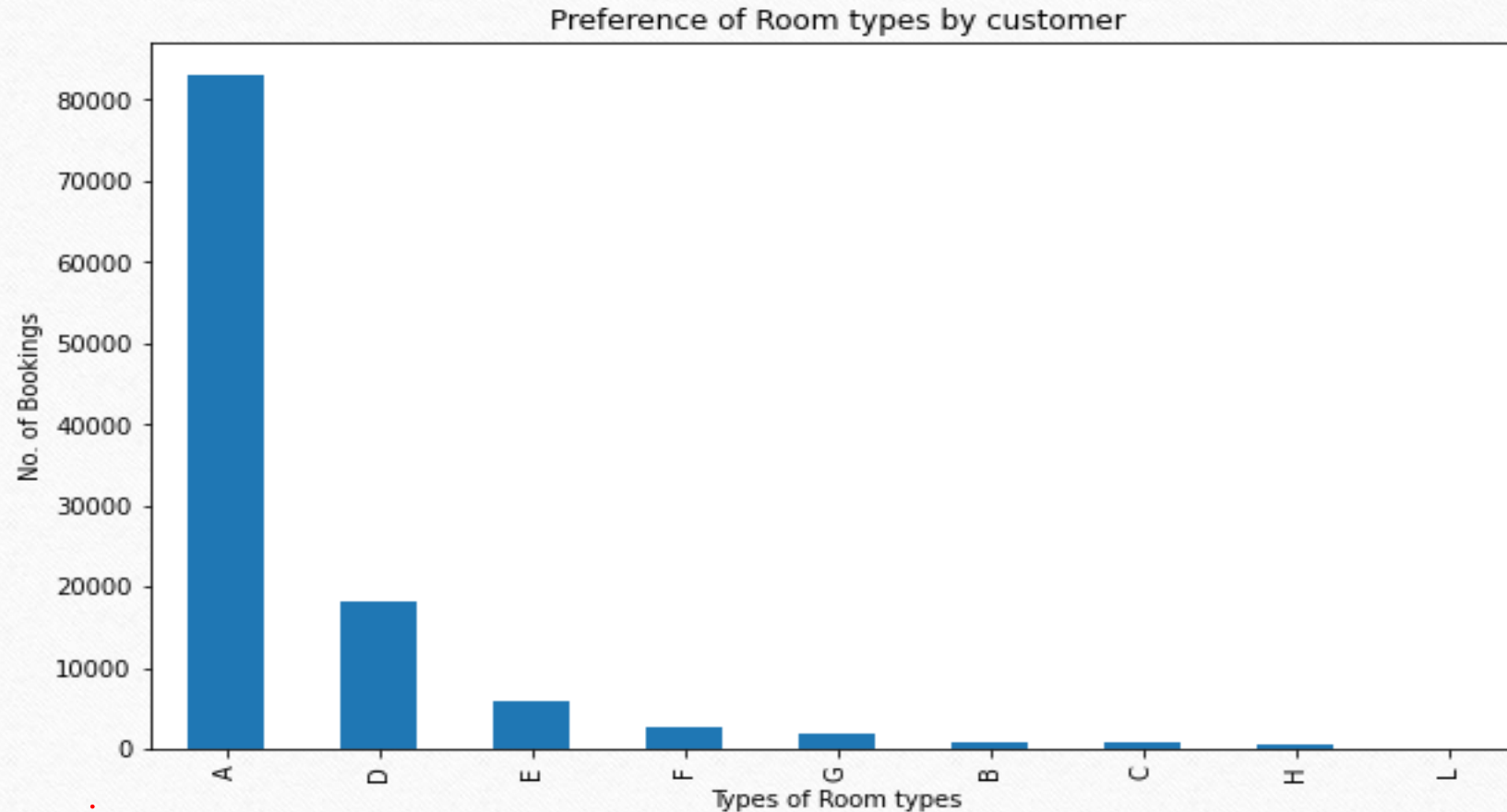
Observation:

- Room type of "A" preferred mostly by customers
- Also it should be note the maximum numbers of booking done in the rooms type of A,D,E than others with minimum number of bookings

Rooms assigned to customer vs Rooms Reserved by the customer

| assigned_room_type	A	B	C	D	E	F	G	H	I	K	L
reserved_room_type											
A	85.0	1.0	2.0	9.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
B	9.0	89.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
C	1.0	0.0	95.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0
D	2.0	0.0	0.0	92.0	4.0	1.0	0.0	0.0	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	90.0	7.0	2.0	0.0	1.0	0.0	0.0
F	0.0	0.0	0.0	0.0	1.0	94.0	4.0	0.0	0.0	0.0	0.0
G	0.0	0.0	0.0	0.0	0.0	1.0	97.0	0.0	1.0	0.0	0.0
H	0.0	0.0	0.0	0.0	0.0	0.0	2.0	97.0	1.0	0.0	0.0
L	17.0	17.0	17.0	0.0	0.0	17.0	0.0	17.0	0.0	0.0	17.0
All	63.0	2.0	2.0	21.0	6.0	3.0	2.0	1.0	0.0	0.0	0.0

Observation:

- From crosstab, relationship of reserved and assigned rooms found.
- Hotel ensured that, 97.0 % of Customer getting the same room type as reserved in room type "G" and "H"
- Lowest possibility of getting the same room type when reserved in room type "L"
- As we know, maximum number of booking done for room type "A" which ensured 85.0% of same room as reserved by the customer
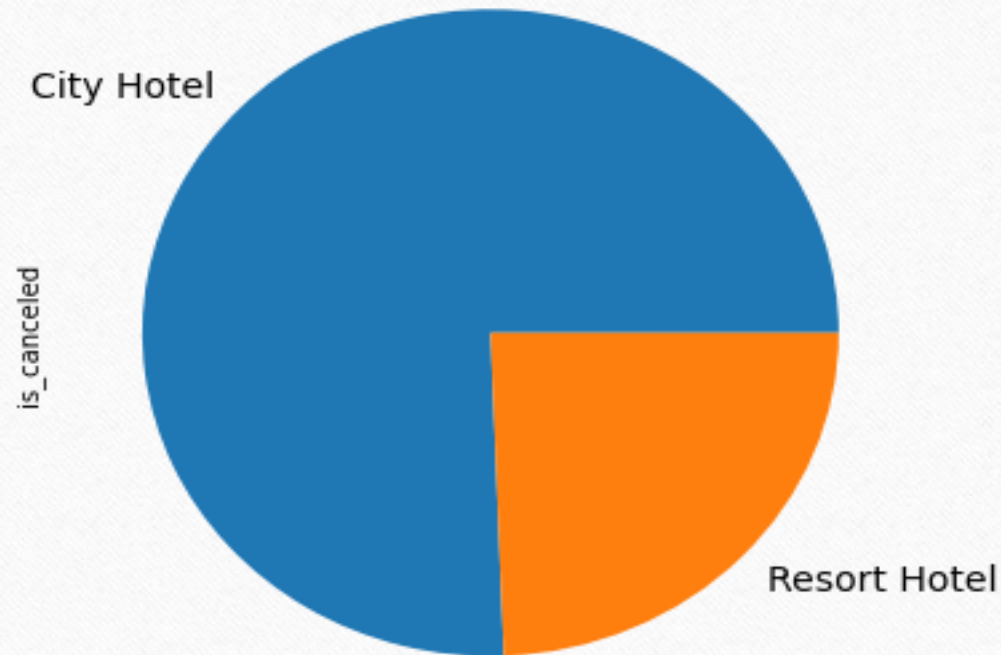
waiting time Vs cancellation
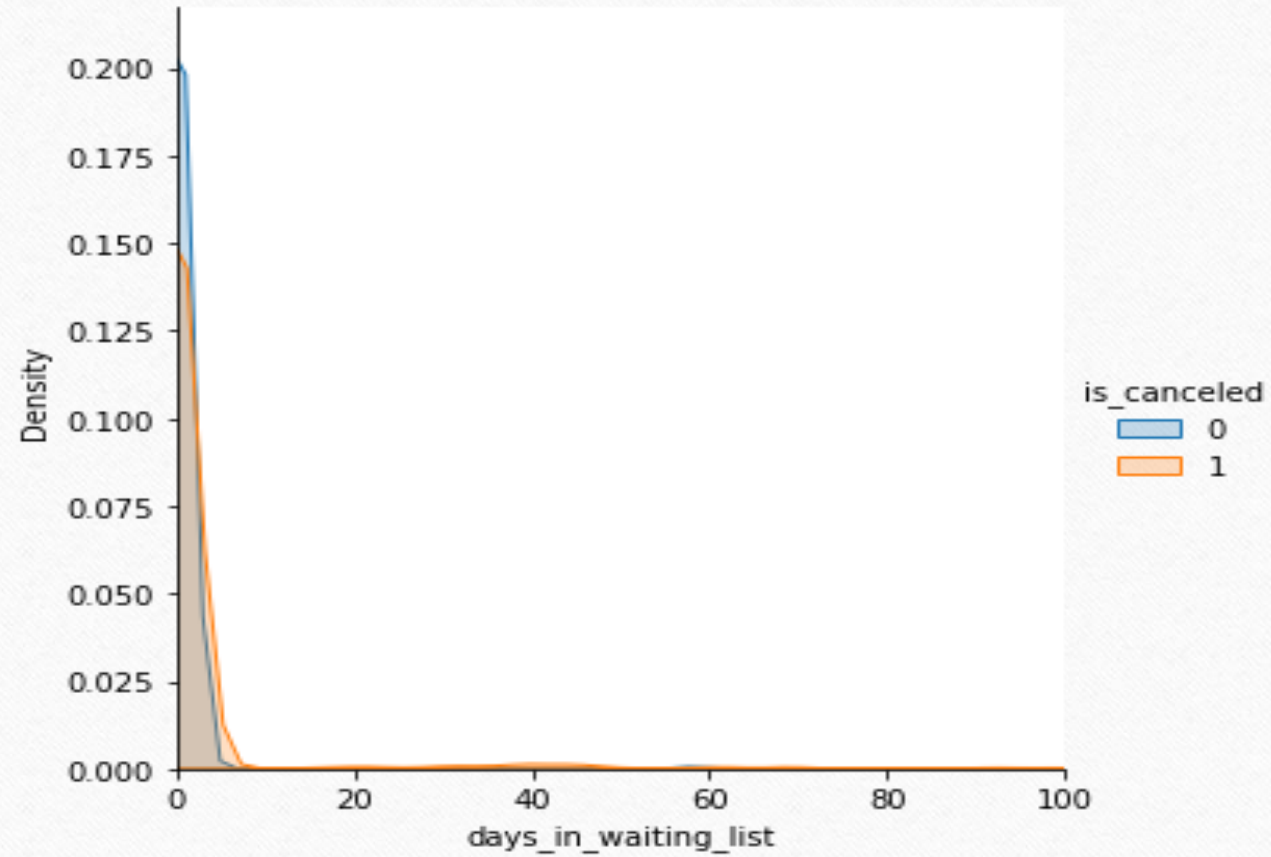
Observation:
- As found from the density,
- Increase in days in waiting list increasing the cancellation of the booking

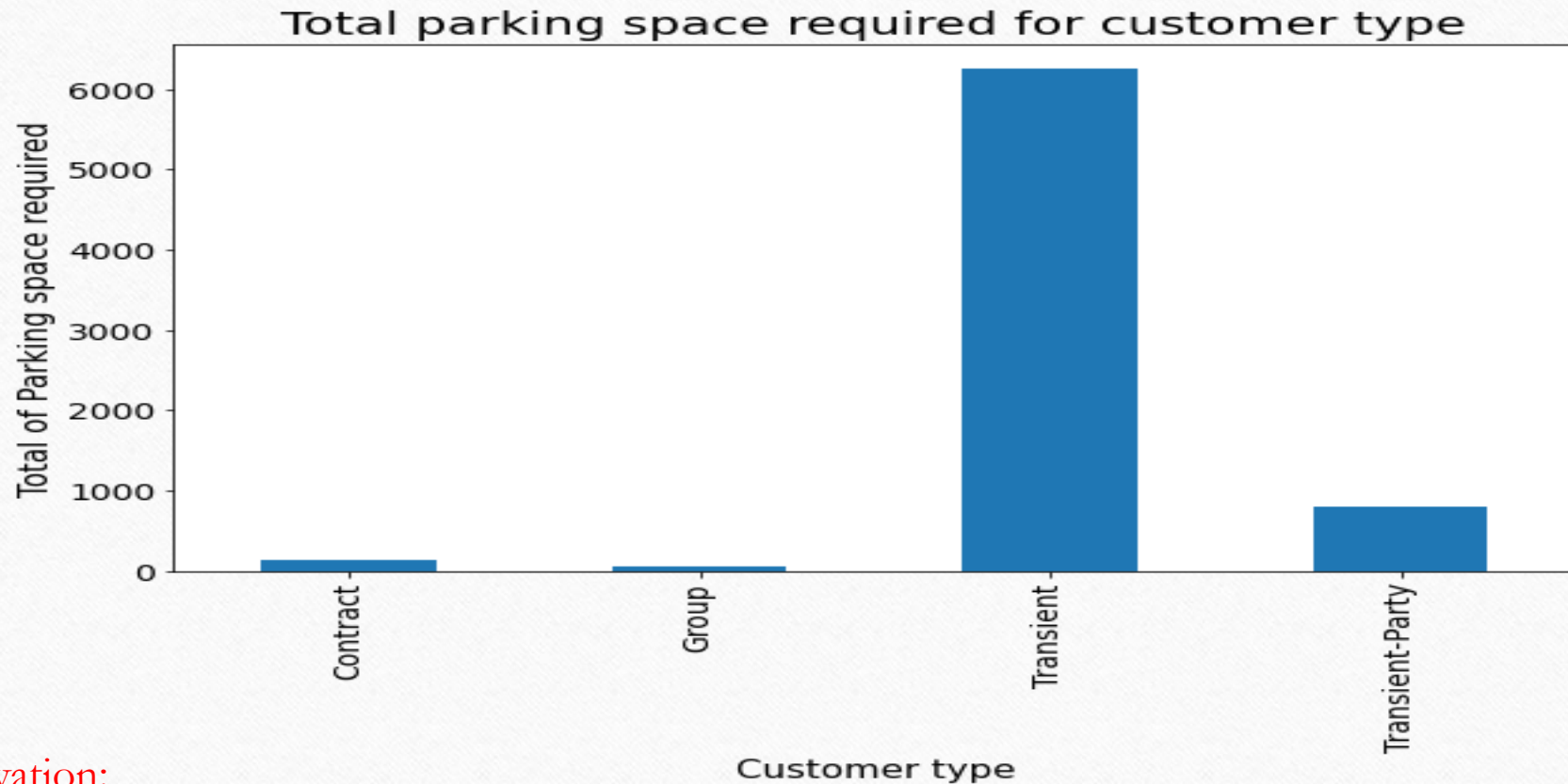Total cancellations for each hotel type

Total cancellations for each hotel type

Observation:
- Percentage cancellation in city hotel is 41.09%.
- Percentage cancellation in resort hotel is 27.66%

Sum of parking space needed according to type of booking

Total parking space required for customer type

Observation:
- When customer type is Transient that means the stay is more few days so it is possible that customer bringing his/her own vehicle that's why the parking space required it high for Transient
- In contract and group booking customer will probably take a hired vehicle form hotel or from somewhere else that's why they do not need parking space.

Conclusion

- Customer preferred City Hotel more than Resort Hotel
- But, the maximum length of stay is higher (than city hotel) in resort type as resort is mostly used for vacation purpose.
- Bookings in the month of August is highest and January found lowest number of bookings.
- Bookings are mostly made for 2 adults with 1 children in combination
- Customers of any type preferring "BB" Bed and breakfast type of meals
- Greater number of bookings done by customer from country PRT(Portugal)
- Travel agency (TA) or Tour operator (TO) Plays vital role in hotel booking
- Except "Direct bookings", all market-segment has a greater number of bookings in city hotel type.

- Room type of "A" preferred mostly by customers, which ensured only 85.0% of same room as reserved by the customer

- Customer booking in the room type "G" and "H" getting 97.0% of the reserved room.

- Undefined and Aviation market segment customer had not shown much interest in the Resort Hotel

- When customer type is Transient that means the stay is more few days so it is possible that customer bringing his/her own vehicle that's why the parking space required it high for Transient

- In contract and group booking customer will probably take a hired vehicle form hotel or from somewhere else that's why they do not need parking space.