Capstone Project – 1

Exploratory Data Analysis

Hotel Booking Analysis

Akash Gawande

Contents :

- Introduction
- Data exploration
- Data cleaning
- Types of hotel
- Bookings
- Deposit policies
- Customer segmentation
- Cancellation rate
- Preferences
    - Meals
    - Room Type
    - car parking
- Guests from different countries
- ADR and Corelation Heat map
- Conclusion
- References

Hotel Industry

- The Hotel Industry is one of the most important components within the service industry, catering for customers who require overnight accommodation. It is closely associated with the travel industry and the hospitality industry, although there are notable differences in scope.
- Overall, sales from hotel accounts 87.4% of industry revenue (in India 53.9% and 8% total employment rate).

Exploring the Dataset

`df.head()`

	hotel	is_canceled	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	adults	...	deposit_type	agent	company	days_in_waiting_list	customer_type	adr	required_car_parking_spaces
0	Resort Hotel	0	342	2015	July	27	1	0	0	2	2	No Deposit	NaN	NaN	0	Transient	0.0	1
1	Resort Hotel	0	737	2015	July	27	1	0	0	2	2	No Deposit	NaN	NaN	0	Transient	0.0	1
2	Resort Hotel	0	7	2015	July	27	1	0	1	1	No Deposit	NaN	NaN	0	Transient	75.0	1
3	Resort Hotel	0	13	2015	July	27	1	0	1	1	No Deposit	304.0	NaN	0	Transient	75.0	1
4	Resort Hotel	0	14	2015	July	27	1	0	2	2	No Deposit	240.0	NaN	0	Transient	98.0	1

5 rows × 32 columns

[62] `df.tail()`

	hotel	is_canceled	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	adults	...	deposit_type	agent	company	days_in_waiting_list	customer_type	adr	required_car_parking
119385	City Hotel	No	23	2017	August	35	30	2	5	2	...	No Deposit	394.0	0.0	0	Transient	96.14	
119386	City Hotel	No	102	2017	August	35	31	2	5	3	...	No Deposit	9.0	0.0	0	Transient	225.43	
119387	City Hotel	No	34	2017	August	35	31	2	5	2	...	No Deposit	9.0	0.0	0	Transient	157.71	
119388	City Hotel	No	109	2017	August	35	31	2	5	2	...	No Deposit	89.0	0.0	0	Transient	104.40	
119389	City Hotel	No	205	2017	August	35	29	2	7	2	...	No Deposit	9.0	0.0	0	Transient	151.20	

5 rows × 32 columns

✓ 0s   completed at 2:53 PM

Checking and replacing the Null Values

Missing Values

```
df.isnull().sum()
```

```
hotel                               0
is_canceled                         0
lead_time                           0
arrival_date_year                   0
arrival_date_month                  0
arrival_date_week_number            0
arrival_date_day_of_month           0
stays_in_weekend_nights             0
stays_in_week_nights                0
adults                              0
children                            4
babies                              0
meal                                0
country                           488
market_segment                      0
distribution_channel                0
is_repeated_guest                   0
previous_cancellations              0
previous_bookings_not_canceled      0
reserved_room_type                  0
assigned_room_type                  0
booking_changes                     0
deposit_type                        0
agent                           16340
company                        112593
days_in_waiting_list                0
customer_type                       0
adr                                 0
required_car_parking_spaces         0
total_of_special_requests           0
reservation_status                  0
reservation_status_date             0
dtype: int64
```

Data Cleaning

City Hotel and Resort Hotel

➤ <u>**What is mean by City Hotel ?**</u>

A city hotel is what you probably know best. It provides accommodation and meals to traveler's. **Often times, people come from all over the world to stay at a hotel so that they can tour around the place that they are staying.** They do not spend their day in the hotel and they only come back to sleep and shower, maybe a meal or two. **In general, hotels despite having the amenities and luxuries similar to these of resorts are more a place to stay and not a destination.**

City Hotel and Resort Hotel

➢ What is a Resort ?

A resort is the destination. Oftentimes, resorts are located in scenic areas but not near a city or anything really. **People visit a resort to get away from other people and cities. They stay in the resort and probably don't travel around as they may do when staying at a city hotel.**

City Hotel and Resort Hotel

- It is Observed that City Hotel has more number of bookings compared to Resort Hotel.

- The obvious assumption for this trend would be because of monetary reasons.

- City hotels would be more affordable than resort hotels due to accessibility, reliability and lavishing factors.

Legend:
- City Hotel (red)
- Resort Hotel (black)

33.55%

66.45%
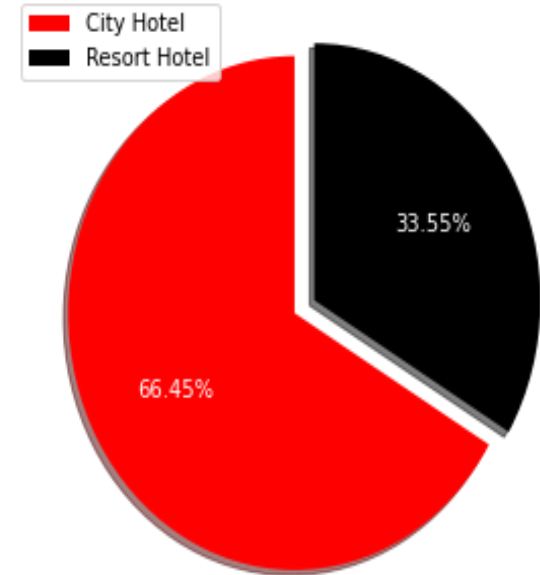
Number of Bookings Across Different Years

➢ Most Number of the bookings are done in the year 2016 following 2017 and 2015 for the dataset given comprising the data for these 3 years.

➢ When, sub plotted, we can see that City hotel has high demand rather than Resort hotels due to its reliability and price difference compared to resort hotels.

# Number of Bookings Across Different Years
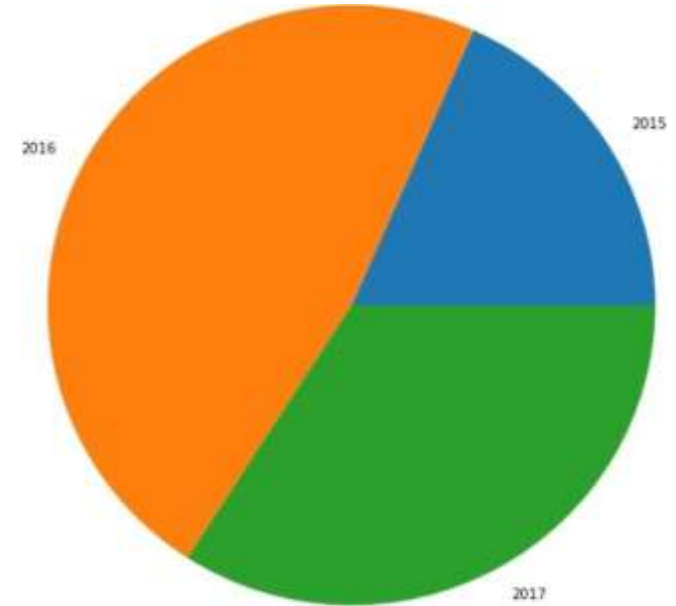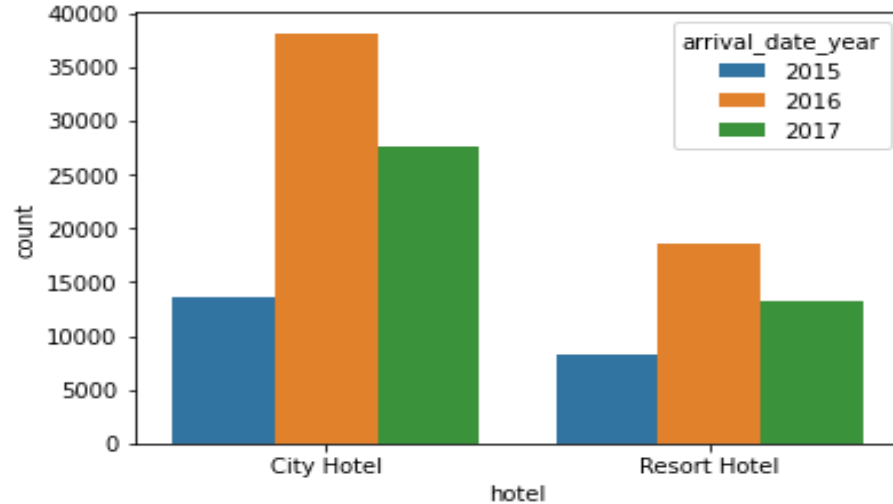
Deposit Policies of Hotel

AI

- Most number of the hotel bookings are made in '**No Deposit**, type of transaction which is the reason for the high cancelation rate.

- 'No Deposit' type of transaction means without paying any advance money for the booking.

- In order to reduce the cancelation rate, we need to find suitable alternative for the 'No Deposit' type of transaction.


Refundable
Non Refund
0.03%
16.22%
83.75%
No Deposit
City Hotel

Refundable
Non Refund
0.35%
4.29%
95.35%
No Deposit
Resort Hotel
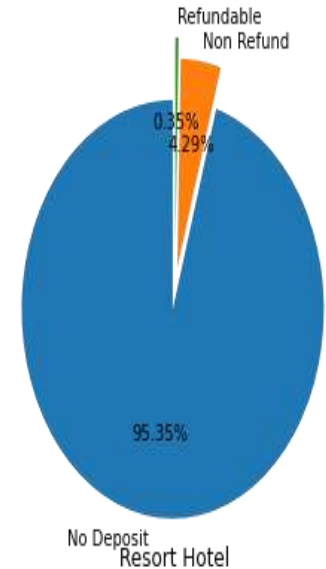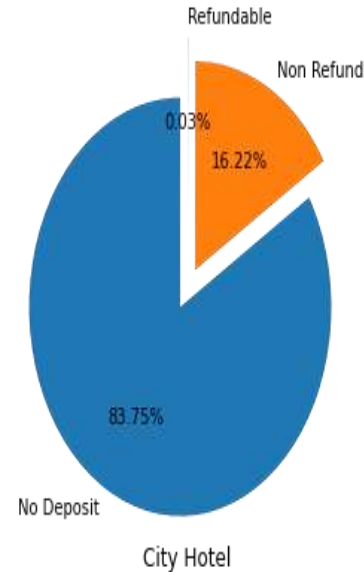
Total Number of Customers across Various Market Segments

- Online Travel agency segment gives the high amount of leads for the hotel booking than that of any other sources of Marketsegments.

- We can report that we need to target our marketing area on online TA websites or apps and focus majorly on online TA.

- The following majority market segments are offline travel agencies, groups and direct customers.


City Hotel

Resort Hotel
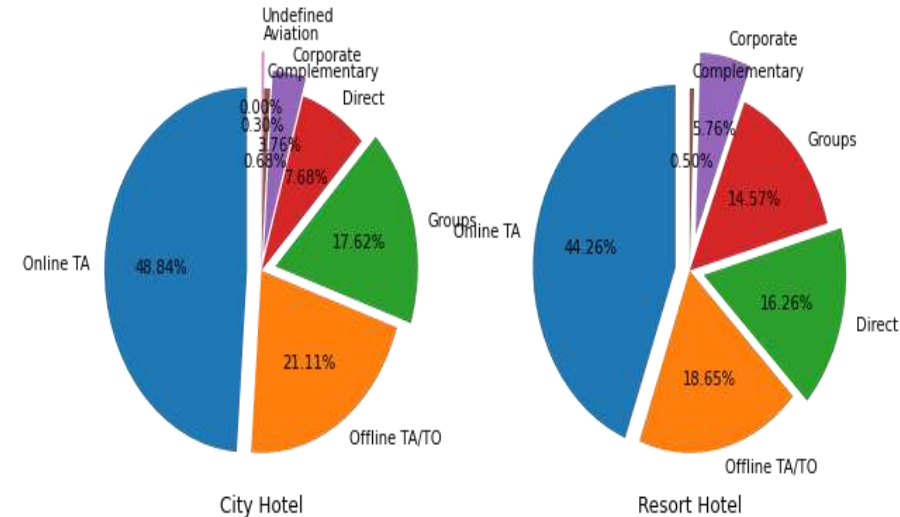
Cancelled Bookings

- It is Observed that out of 119390 bookings made, 75166 of them have checked in and 44224 of them has cancelled their bookings.

- In percentage, it is about 63% of check-in bookings and 37% of bookings were cancelled.

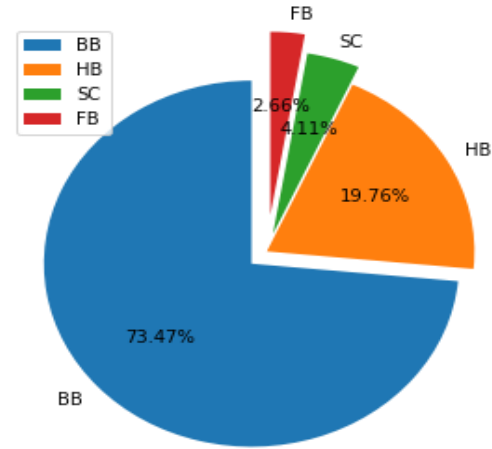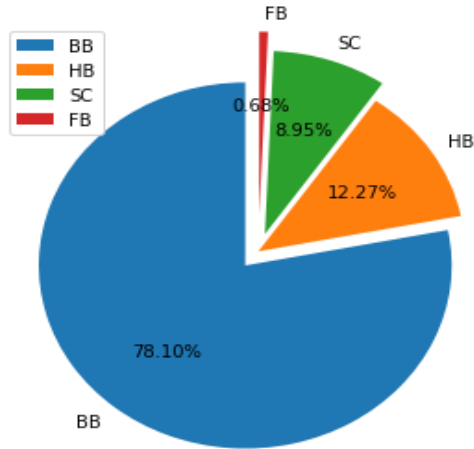- High cancellation rate will adversely effect the hotel industry which reduces the yield of profit.
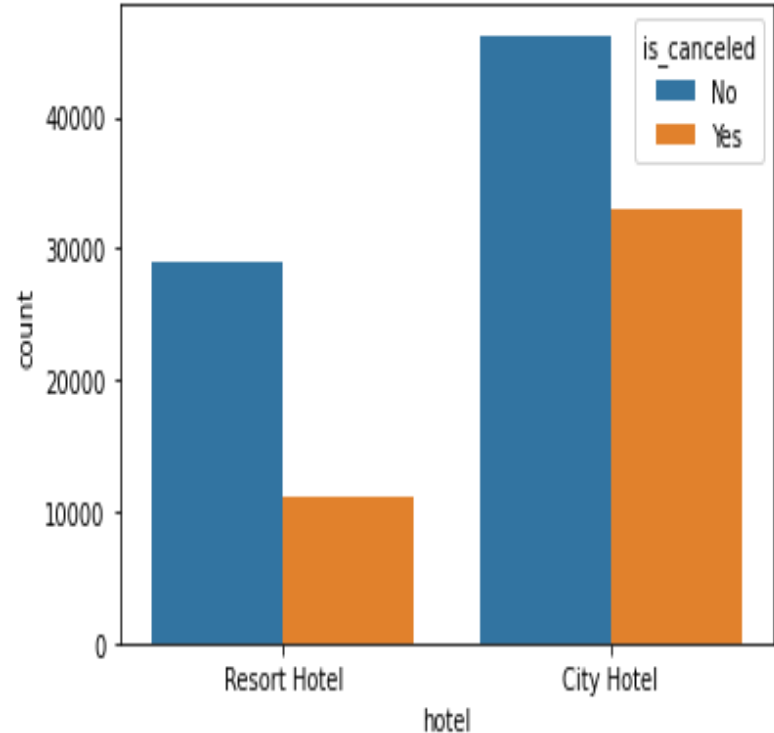
Number of Booking Cancelled By the Same Customers Repeatedly
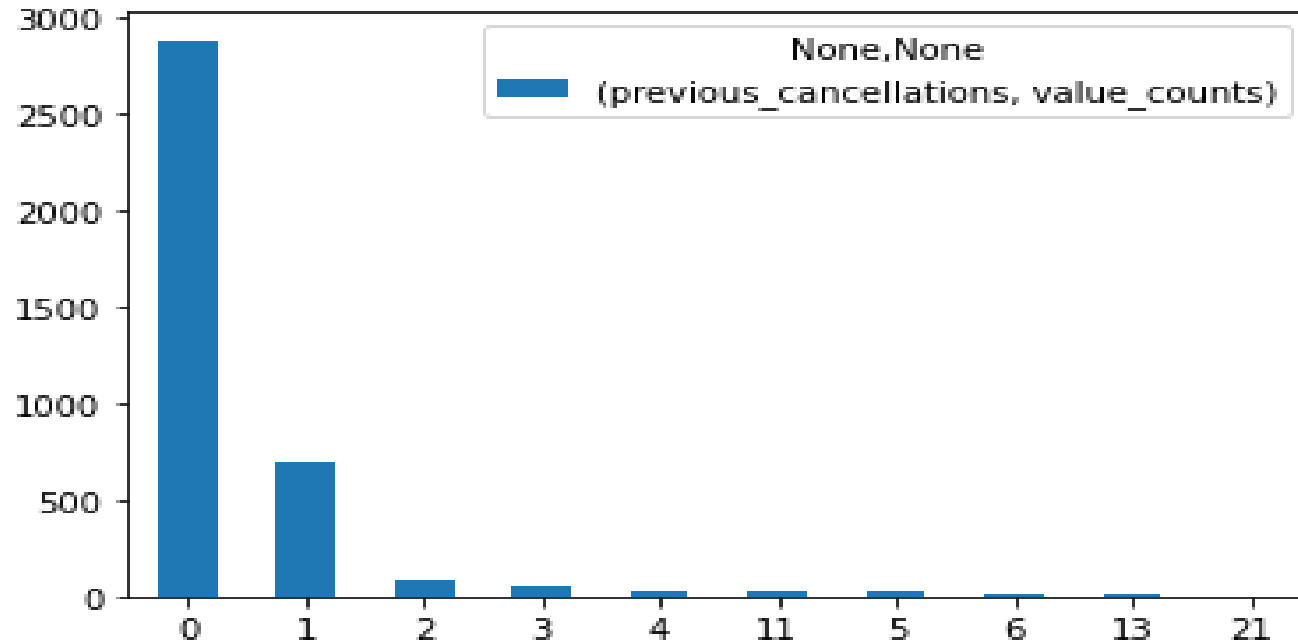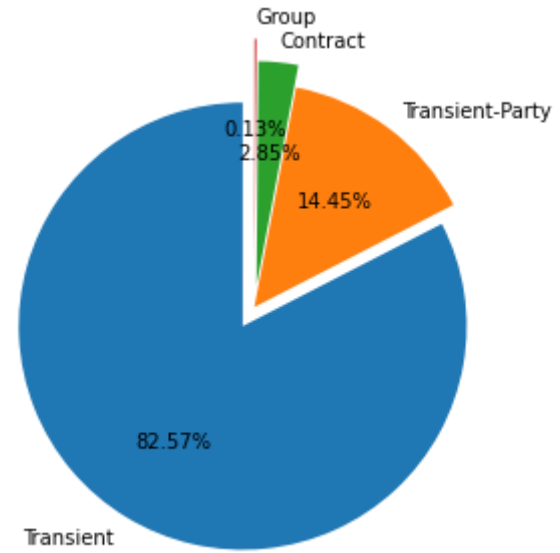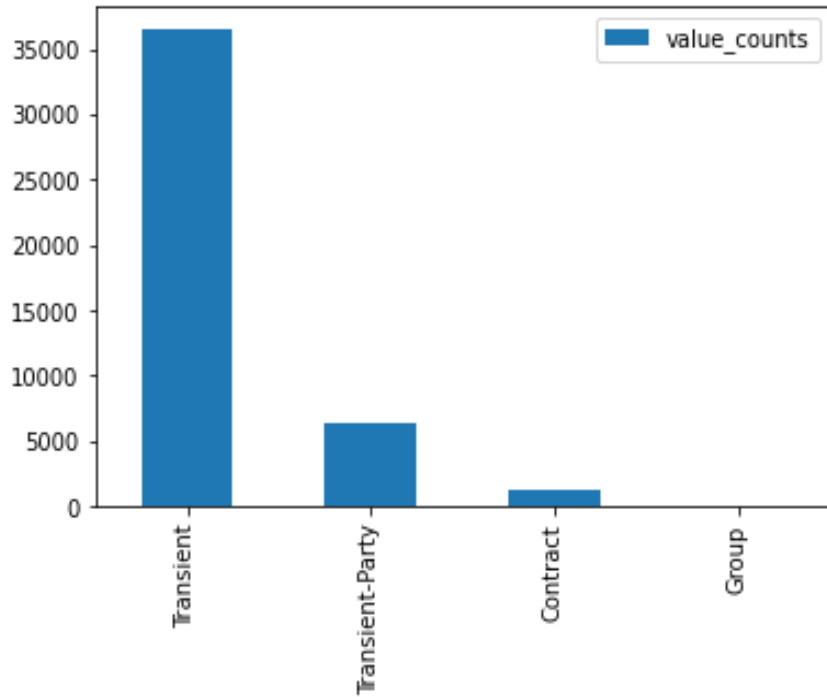
It is observed that **Transient customers** are more likely to cancel the booking.

Meals preference

As per the above Graph, we can conclude that guests for BB (i.e. Bed & Breakfast) is more than others i.e. Half Board(HB), Full Board(FB), Self Catering(SC)

Preferred room type by guests :

- **Room only**

Room only is used to indicate that **the price of accommodations in a hotel or guest house does not include the cost of food**. which is denoted in the graph by '**A**'

- **Standard rooms**

Standard rooms have a bathroom and perhaps some other simple furniture, like a desk, a chair, a dresser, and a TV. Denoted by **'D'**

- **Minimalist hotel rooms.**

While the rooms may still offer a separate bedroom and bathroom, they're **typically smaller than standard hotel rooms and are designed to accommodate guests that don't intend to spend a lot of time in their room**. And its denoted by **'E'** in the following graph.

Preferred room type by guests :

- **Deluxe hotel rooms.**

Size and price are usually the two main differences between standard and deluxe hotel rooms. Deluxe rooms are **usually larger than their standard counterparts, may include a bathtub and a shower in the bathroom, and include more high-end amenities**.it has been denoted by **'F'**

- **Studio hotel rooms**

Studio-style hotel rooms **offer living, sleeping, and often dining space all in one room**. Some studio hotel rooms include a bed while others include a pull-out sofa that doubles as both the living and sleeping furniture. Studio hotel room is denoted by **'C'**
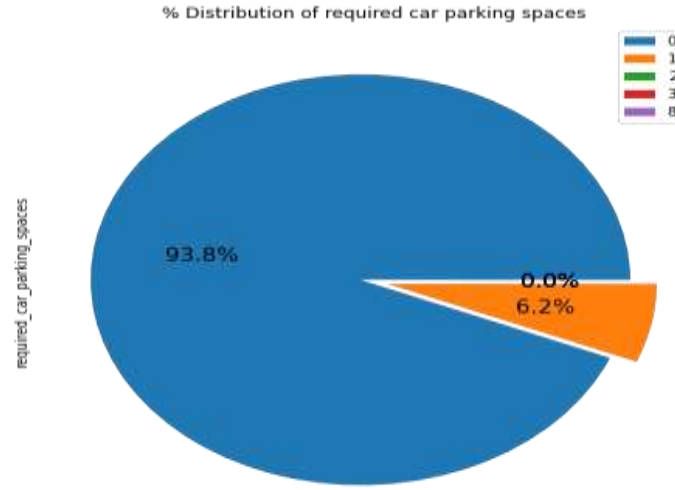
As the price of the hotels room has been increase the booking rate of the Hotels has been decreases.

Preferred room type by guests :

Maximum number of guest Preferred 'A' type rooms.

Most preferred Room type
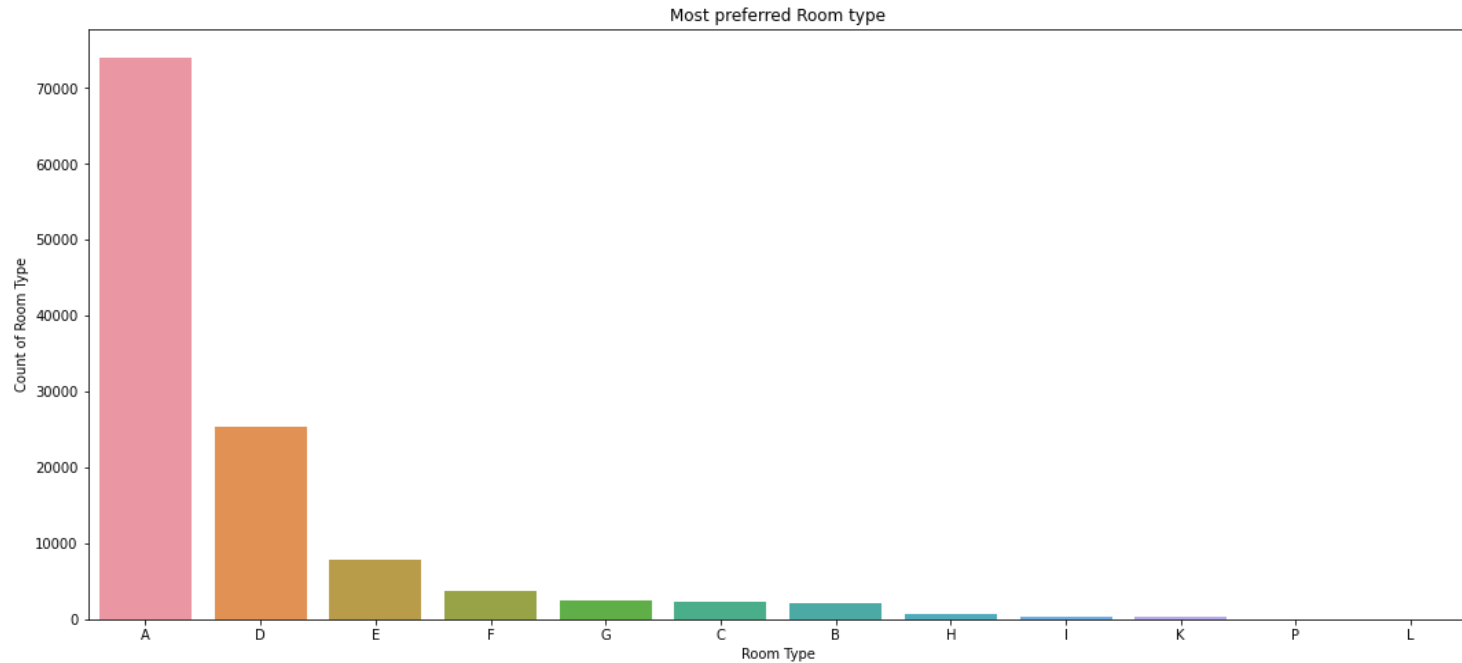
Car Parking Space Analysis

% Distribution of required car parking spaces

- Most of the hotels have zero car space parking.
- Among which have car space parking has one car spacing and the trend decreases drastically as number of car spacing increases.

Top 10 Countries with highest number of Bookings

- After analyzing the dataset, we found that **Portugal** tops the position with 48590 customers followed by **UK** with 12129, **France** with 10415, **Spain** with 8568 and **Germany** with 7287 customers.

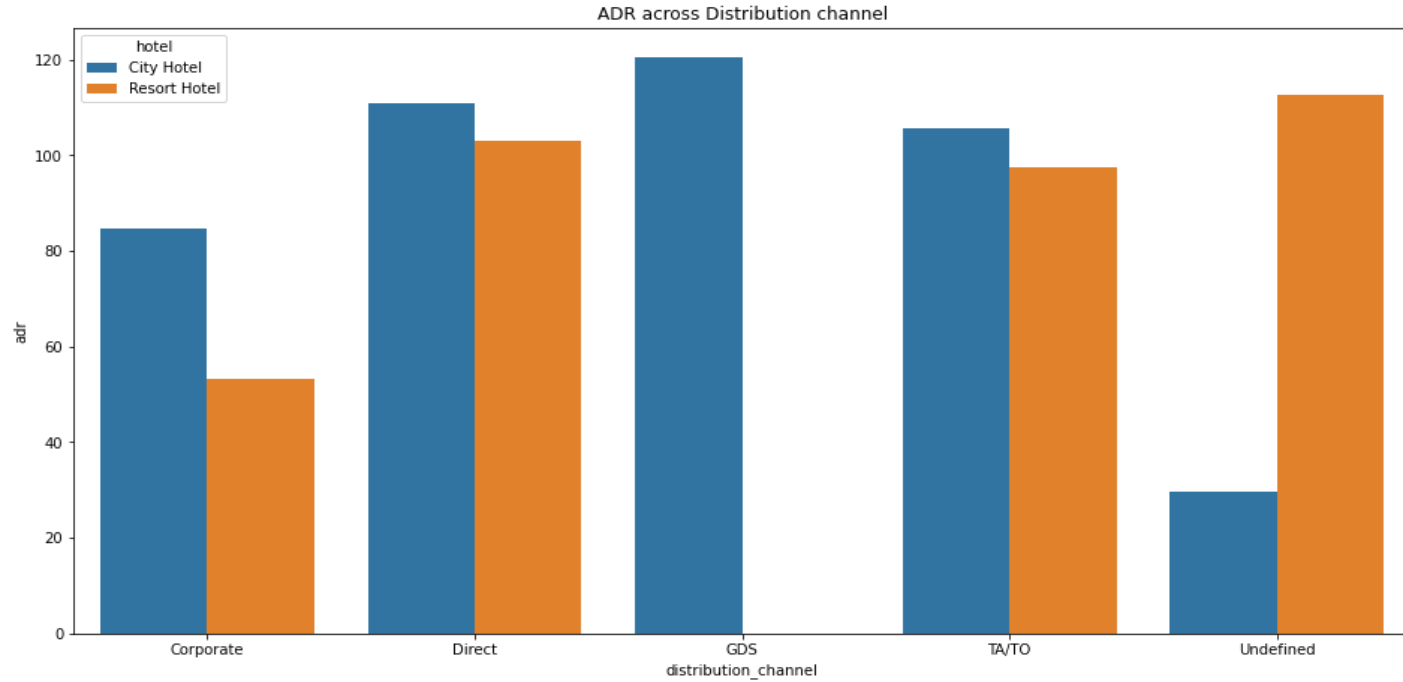- **Netherlands** sits back with least number of customers among the top 10 Countries.
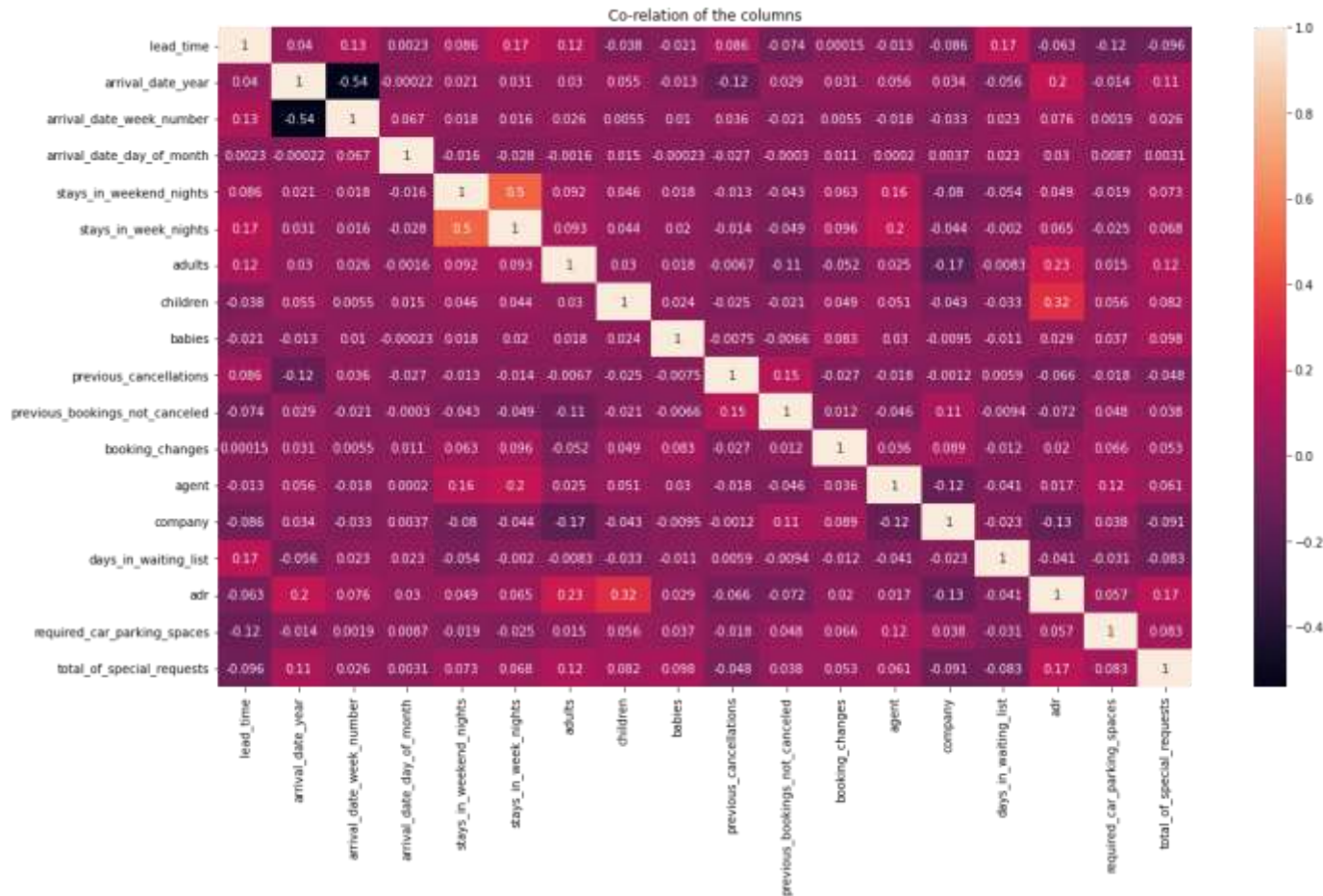
Top 10 Countries with highest number ofBookings

Number of guests from diffrent Countries
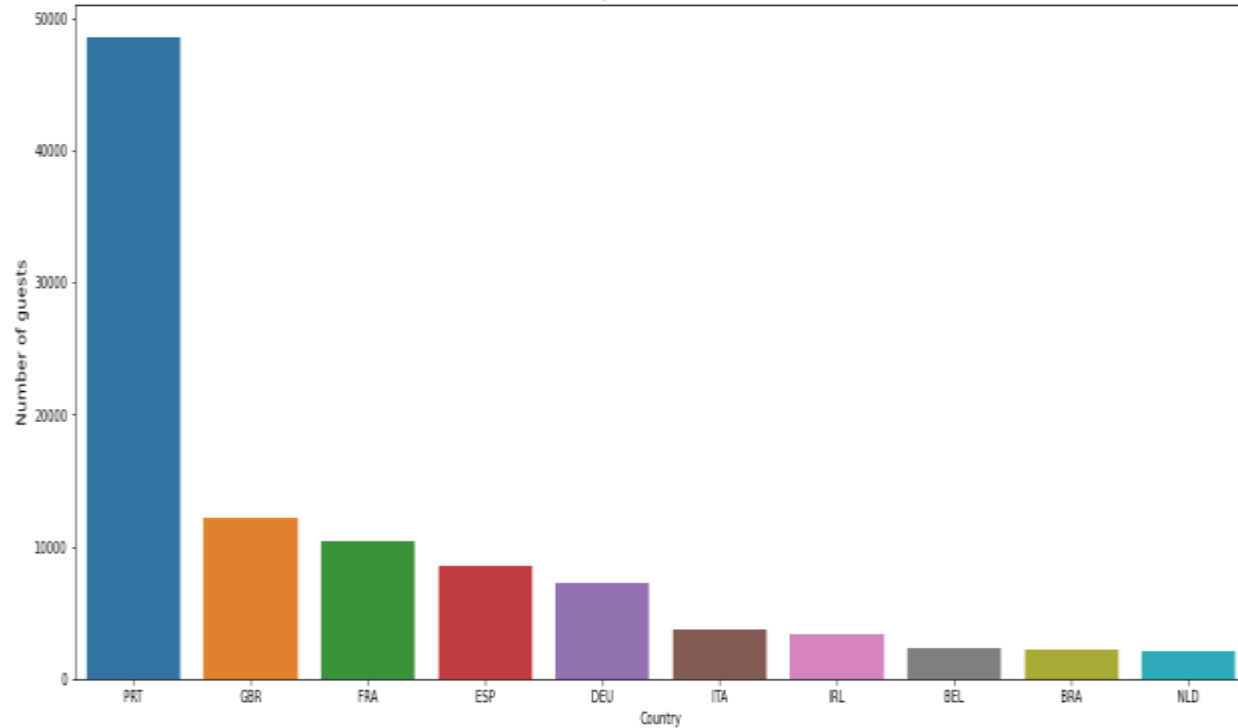
ADR Across Distribution Channel

Correlation Heatmap

Co-relation of the columns

Conclusion

1) Out of 119000 customer dataset, 75166 customers checked in hotel while 44224 customers cancelled their bookings, that is about 37% of the booking got cancelled whereas 63% customers checked in the hotel.

2) Majority of the deposit type is of **'No deposit'** type, which itself concludes the high rate of cancellation rate.

3) From the given dataset, we can see 2016 is the year in which hotel bookings are the highest.

4) We can also see the trend in the middle of the year as those season has less weather condition and holidays are more during that season. We can also infer that winter season has the lowest number of bookings around the globe and we can assume it because of the weather condition.

5) City hotels are of in high demand compared to resort in all aspects due to its reliability of majority of the population.

7) Online Travel agencies provides high amount of leads and customers following Offline TA, Groups, Direct, etc., So, we an conclude that we need to focus our marketing area on online TA since majority of the visitors tend to reach out to them.

8) After analyzing the meal data, its found that 77.3% of customers prefers BB(Bed & Breakfast).

9) Most of the hotels have 0 to 1 car parking space.

10) Among the top 10 countries with respect to number of bookings, most of them are European countries which belongs to the list of which Portugal stands first with 48.5k bookings.

11) Its observed that average ADR is incrementing every year from 2015 to 2017, which clearly states that hotel business is scaling up every year.

12) Given that we do not have repeated guests, we should target our advertisement on guests to increase returning guests.

References

1) AlmaBetter Study Materials

2) https://pandas.pydata.org/

3) Geek for geeks

4) https://matplotlib.org/

5) https://seaborn.pydata.org/

Thank You