

Capstone Project

Hotel Booking Analysis

Team Members

Gaurav Gade

Hitesh Verma

Anand Gend

➤ Let's Analyse Hotel Booking

Data Exploration

Observe the Data

Find Missing

Data Cleaning

Replace the Null Values

Drop un-necessary columns

Visualise the data

Analyse the data

Conclusion



➤ Fast Growing Ever Green Hotel Business



- A hotel is an establishment that provides lodging, meals and other services for travellers and other paying guests. Overall, sales from hotels account for 87.4% of industry revenue and 82.0% of industry employment.

➤ Explore The Dataset



Go Through The Dataset



hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number
City Hotel	0	44	2017	August	35
City Hotel	0	188	2017	August	35
City Hotel	0	135	2017	August	35
City Hotel	0	164	2017	August	35
City Hotel	0	21	2017	August	35
City Hotel	0	23	2017	August	35
City Hotel	0	102	2017	August	35
City Hotel	0	34	2017	August	35
City Hotel	0	109	2017	August	35

```
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
```

Checking Null Values In Dataset



Replacing The Null Values With Their Mean



```
country          0
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            0
company          0
days_in_waiting_list 0
customer type    0
```

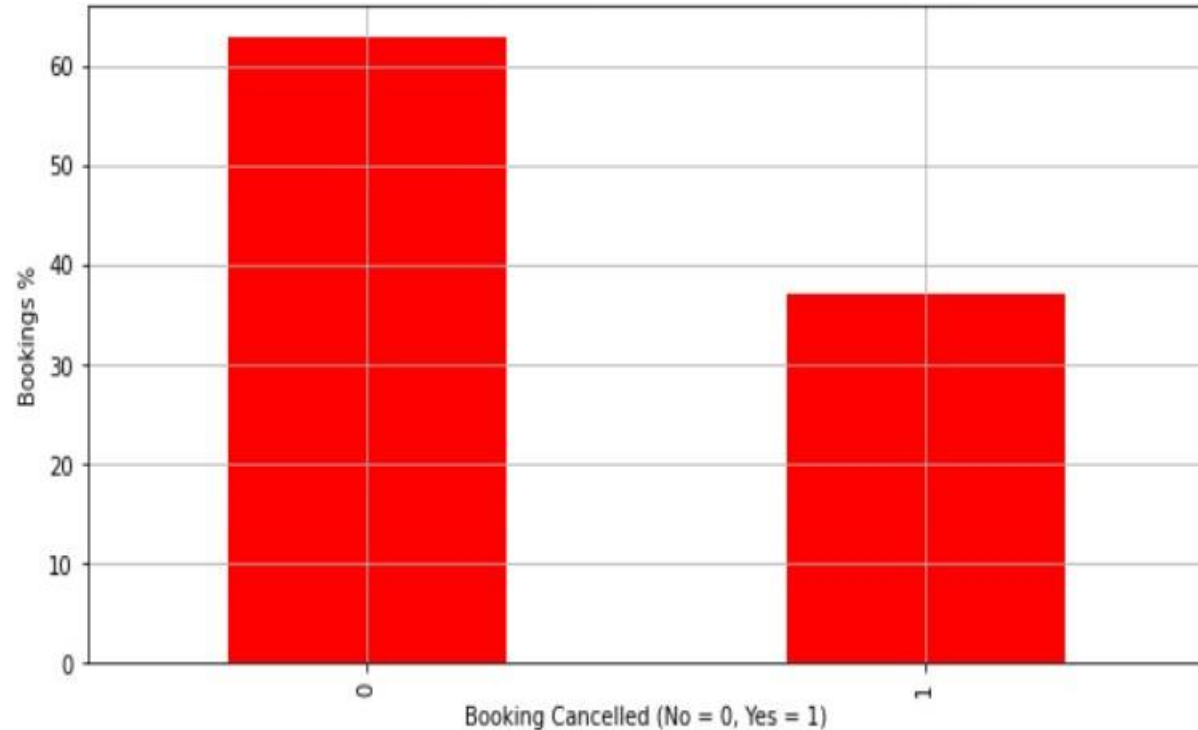
➤ Canceled & Check-Out Booking

In this plot, it is observed that 75166 customers actually checked-out from the hotel but at the same time we discovered some different data resulting in cancellation of bookings i.e. 43017 customers. Due to these cancelled bookings there will be an adverse effect on hotel business which means hotels are not able to make more profit, they are losing their customers.



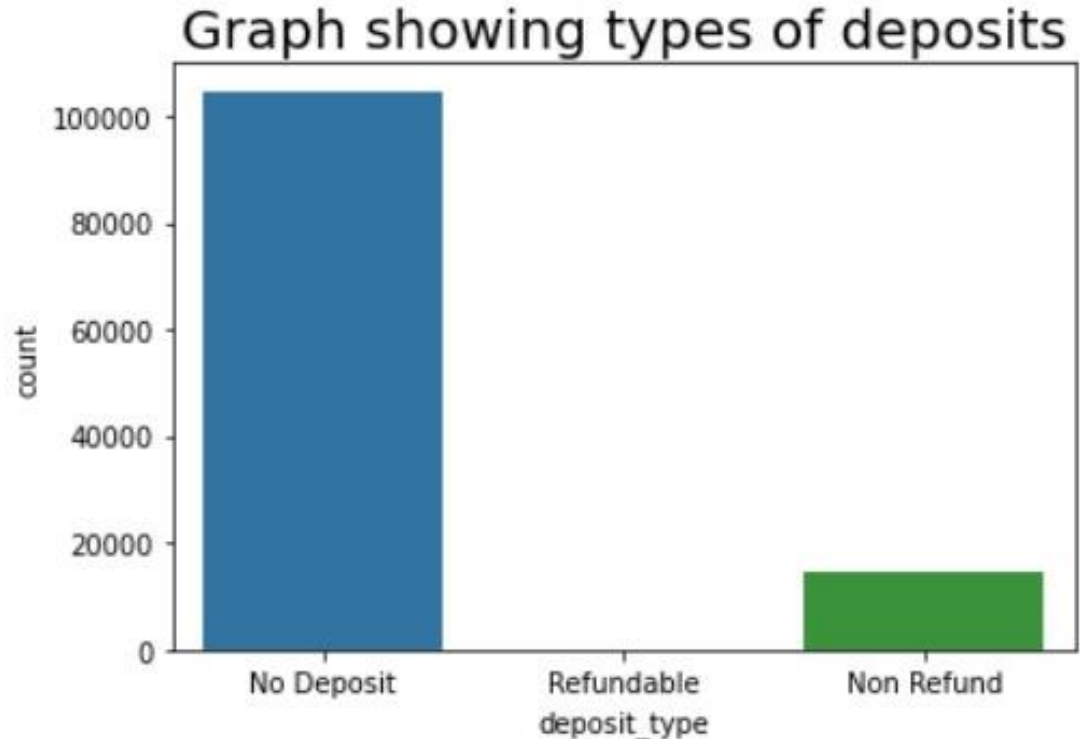
➤ Non-Canceled vs Canceled Booking Percentage

As we saw, total number of booking from our last slide, here we are going to see the same but in terms of percentage. The bar graph here represents that 63% of customers check-in to hotels whereas 37% of customers cancelled their bookings.



➤ Deposit Policies Of Hotels

Very large amount of hotels have “No Deposit” Policy. This possibly might be the reason for cancellation of high amount of bookings. To avoid this booking cancellation, in account to collect more profit and customers- “No Deposit” policy should be either removed or improvised.



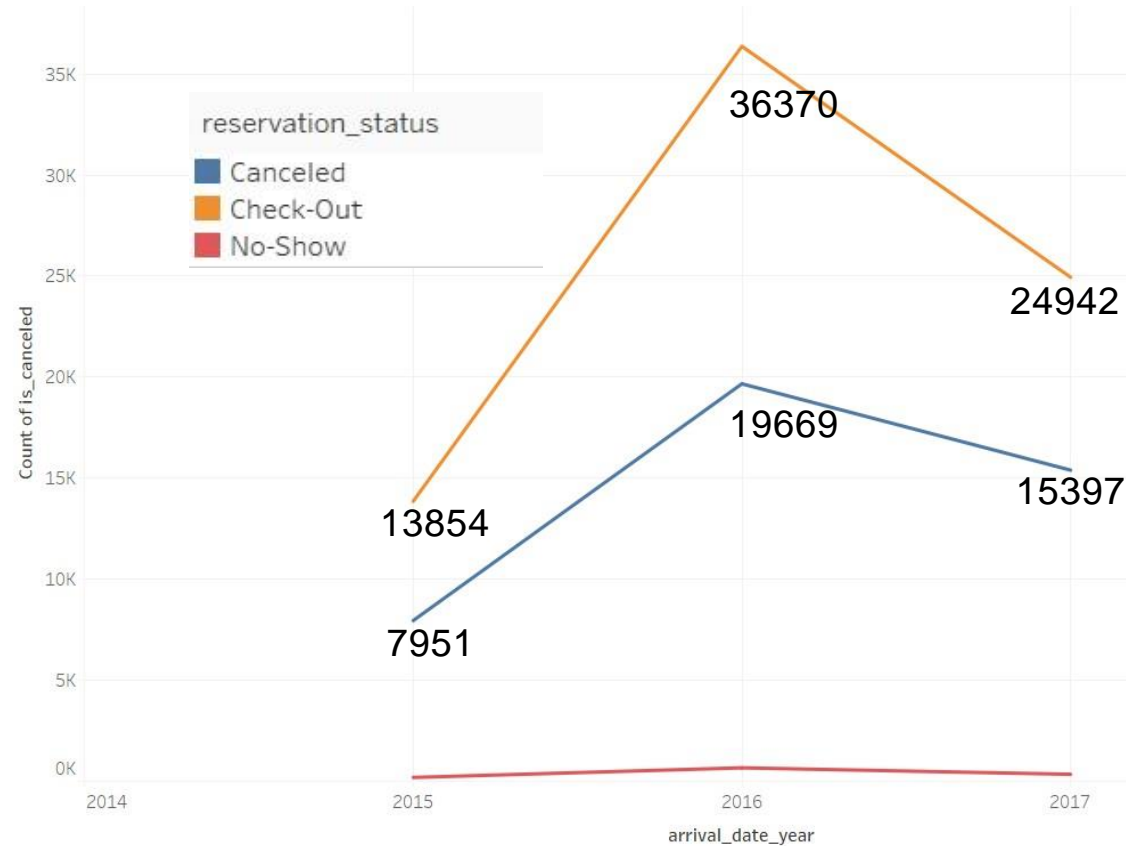
➤ Total Number Of Bookings Across Different Years

How many customers cancelled their bookings & how many customers actually checked-in to the hotel across different years?

Let us find out with simple line chart, in 2015 there are 13854 customers checking-out and 7951 customers cancelled bookings.

While we see in 2016, trend goes upside increasing check-out number to 36370 increase cancelled bookings upto 19669.

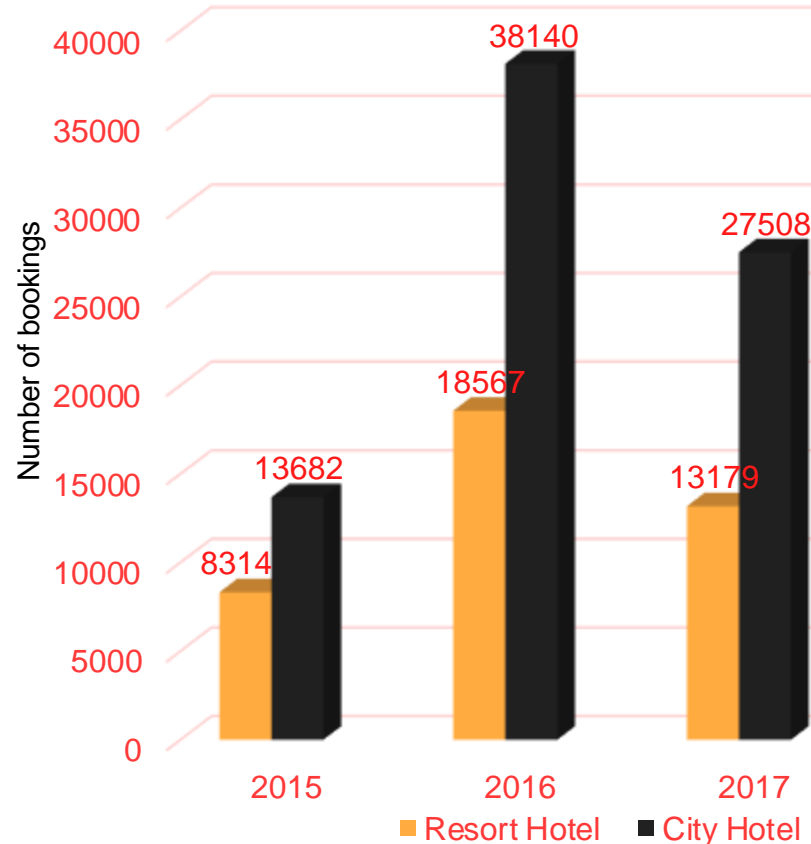
This increased trend did not sustain for more time, went downward in 2017 with 24942 of check-out & 15397 of cancelled bookings.



➤ Demand Trend Of Hotels Year wise

Which type of hotels customer preferred to stay in different years?

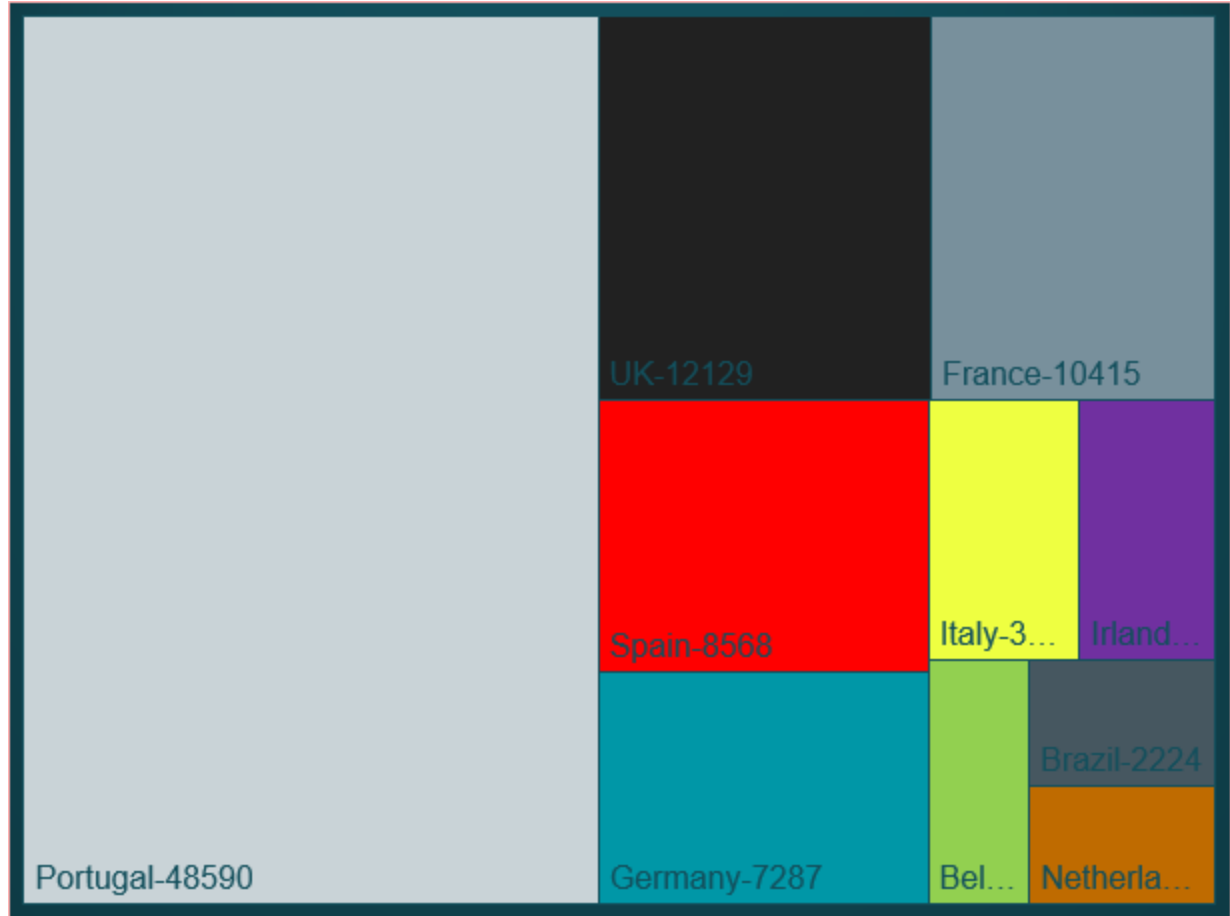
Here, we plotted a subplot for Resort hotel and City hotel. From these columns we can conclude that there is always demand of City hotels as compared to Resort hotels across three different years 2015, 2016 and 2017. As we discussed early, after increasing the booking trend it got decreased again. This happened in both cases – Resort as well as for City hotels.



➤ Top 10 Countries With Maximum Customers

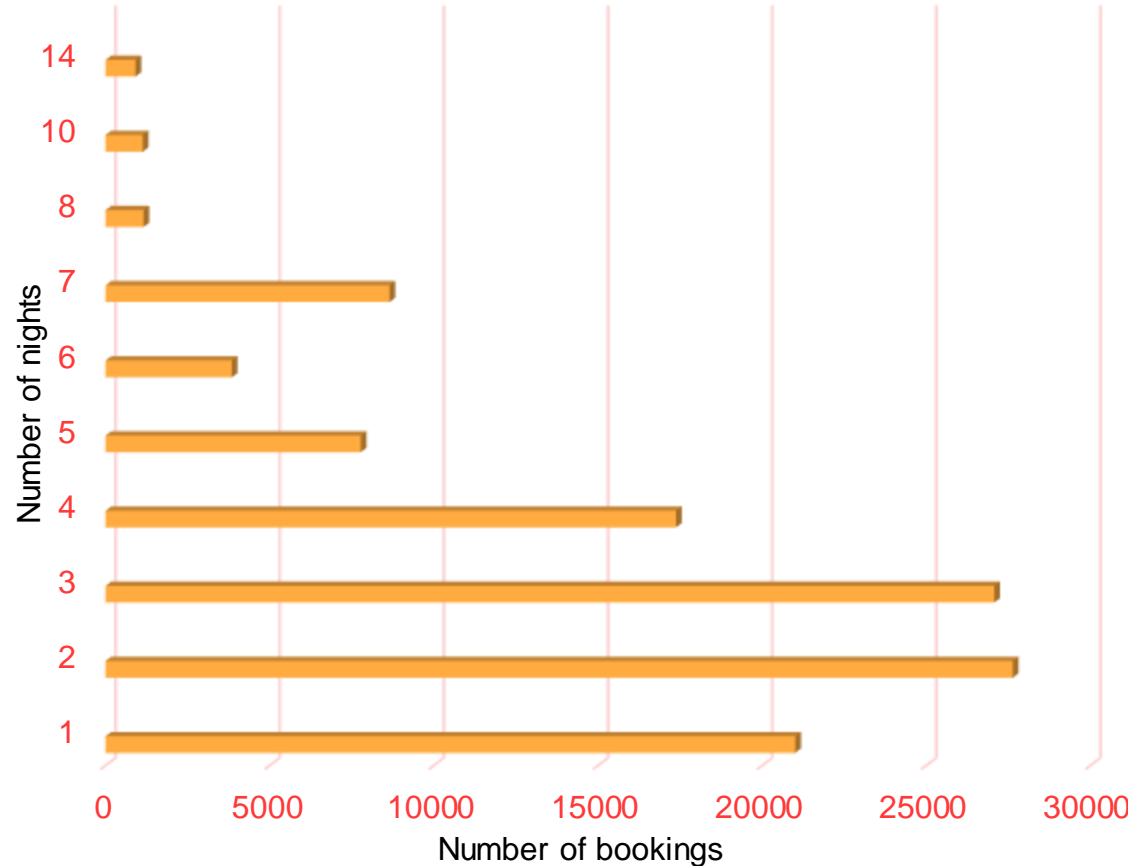
Which are the countries giving maximum customers?

So, here is the result – after analysing the dataset we found that Portugal is on Rank 1 with 48590 customers followed by UK, France, Spain, Germany with 12129, 10415, 8568, 7287 customers accordingly. After these 5 topmost countries Italy, Ireland, Belgium, Brazil and Netherlands has 3766, 3375, 2342, 2224 and 2104 customers i.e., Netherlands sits back with lowest number of customers.



➤ Night Stay Duration

By combining the two columns `stays_in_week_nights` and `stays_in_weekend_nights` we got total number of nights. Hence, we can say that more customers like to stay 2 – 3 nights where some customer prefer to stay for 1 – 4 nights. Very few customers are there who are interested to stay for more than 5 days.

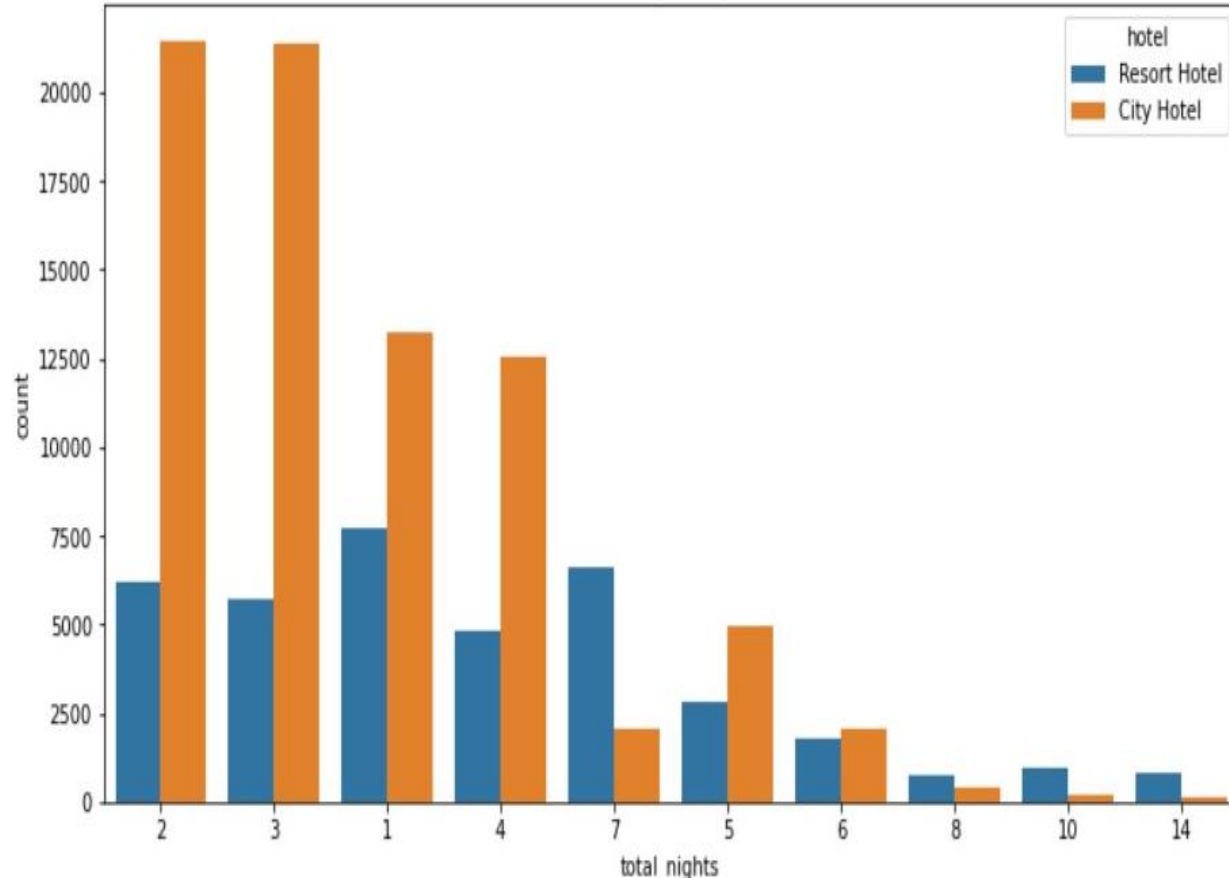


➤ Hotel wise Night Stay Duration

Now we are going to track night stay duration of customers according to Resort hotels and City hotels.

As we already aware that customers loves to stay in City hotels, here also

City hotels have large amount of bookings for 2-3 night stay and followed by less customers who stay in for 1-4 nights. In Resort hotels, 1 night stay customers are more and then 7 night stay customers comes in focus. Very few customers likely to stay for 8 night or more than it for both type of hotels.



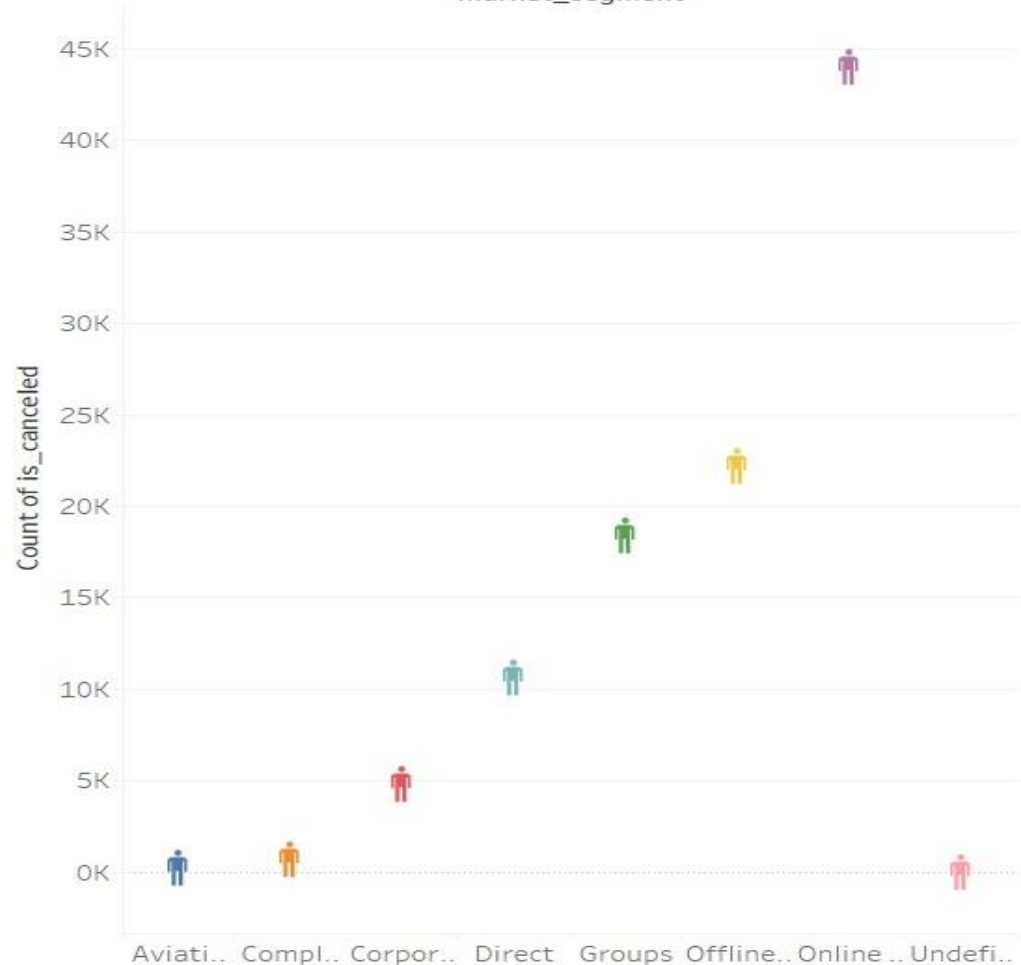
➤ Booking Trend Throught The Year

If we go through booking data along with different months, we found out that August has the highest number of bookings throughout the year with July at the second place where January has the lowest number of bookings i.e., we can assume that January will be the best month for booking to get the best rate on daily basis where booking in month of August will not be economical since it has high demand of room bookings obvious that the cost will also be high from customer shoes.



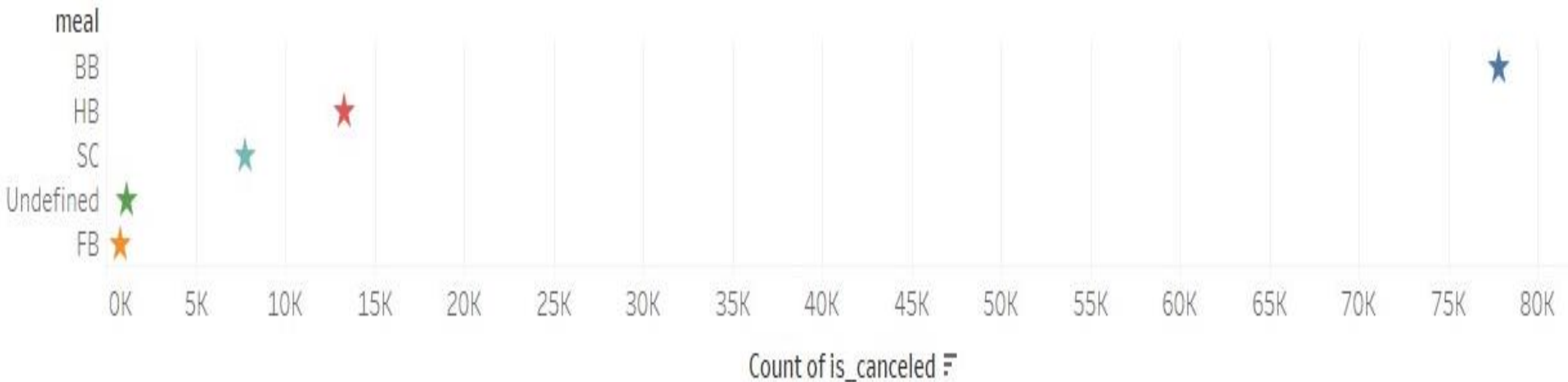
➤ Total Number Of Bookings Across Various Market Segment

7. Online TA (Travel Agency) segment gives high amount of customers and then Offline TA/TO, Groups, Direct etc. respectively. Complementary, Aviation and Undefined has the lowest amount of customers.
8. So , from this we conclude that We can target our marketing area to be focus on these travel agencies website and work with them since majority of the visitors tend to reach out to them.



➤ Meal Category vs Count Of Booking

- Undefined/SC – no meal package
 - BB – Bed & Breakfast
 - HB – Half board (breakfast and one other meal – usually dinner)
 - FB – Full board (breakfast, lunch and dinner)
- Maximum of the bookings are made with bed and breakfast .So, BB type of meal category is the most preferable in all type of customers, where negligible bookings are made with FB type of meal.



➤ Booking vs Customer Type

1. **Contract** — When the booking has an allotment or other type of contract associated to it.

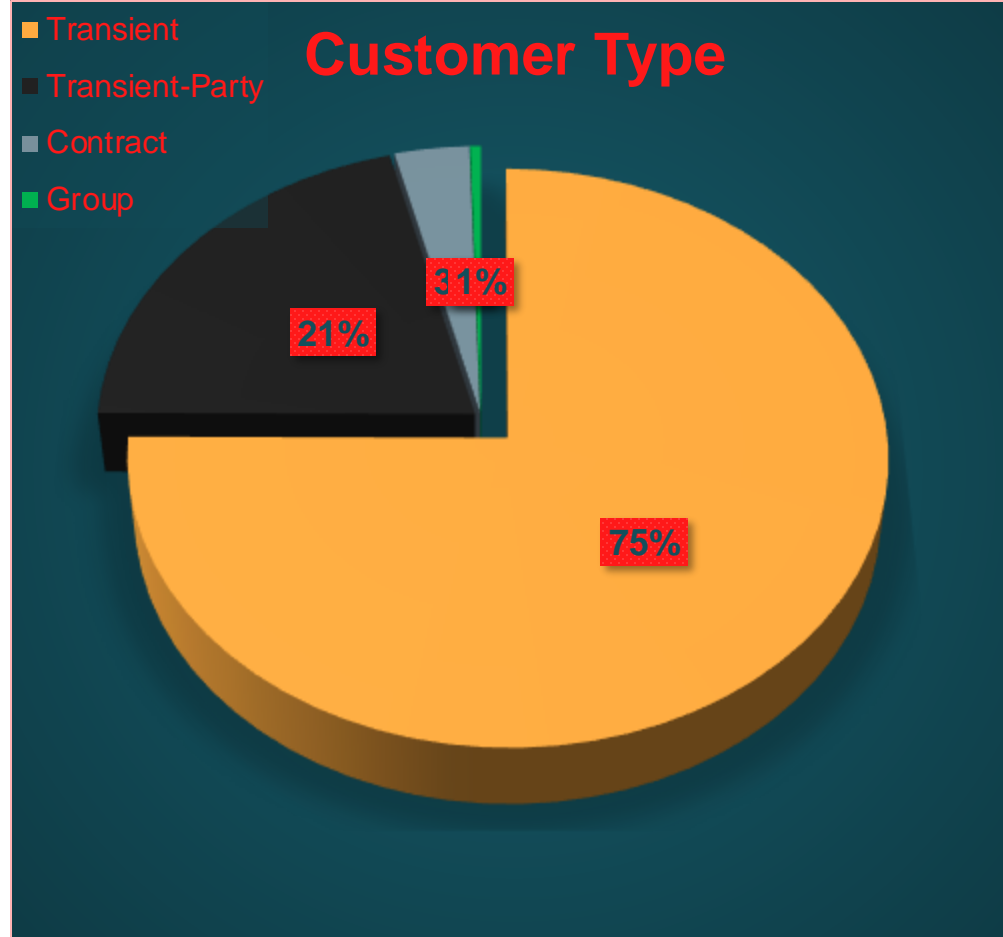
2. **Group** — When the booking is associated to a group.

3. **Transient** — When the booking is not part of a group or contract, and is not associated to other transient booking.

4. **Transient-party** — When the booking is transient, but is associated to at least other transient booking.

This means that the booking is not part of a group or contract. With the ease of booking directly from the website, most people tend to skip the middleman to ensure quick response from their booking.

Transient type of customer is the main source of booking because 75% of booking coming from this side after that Transient-Party,

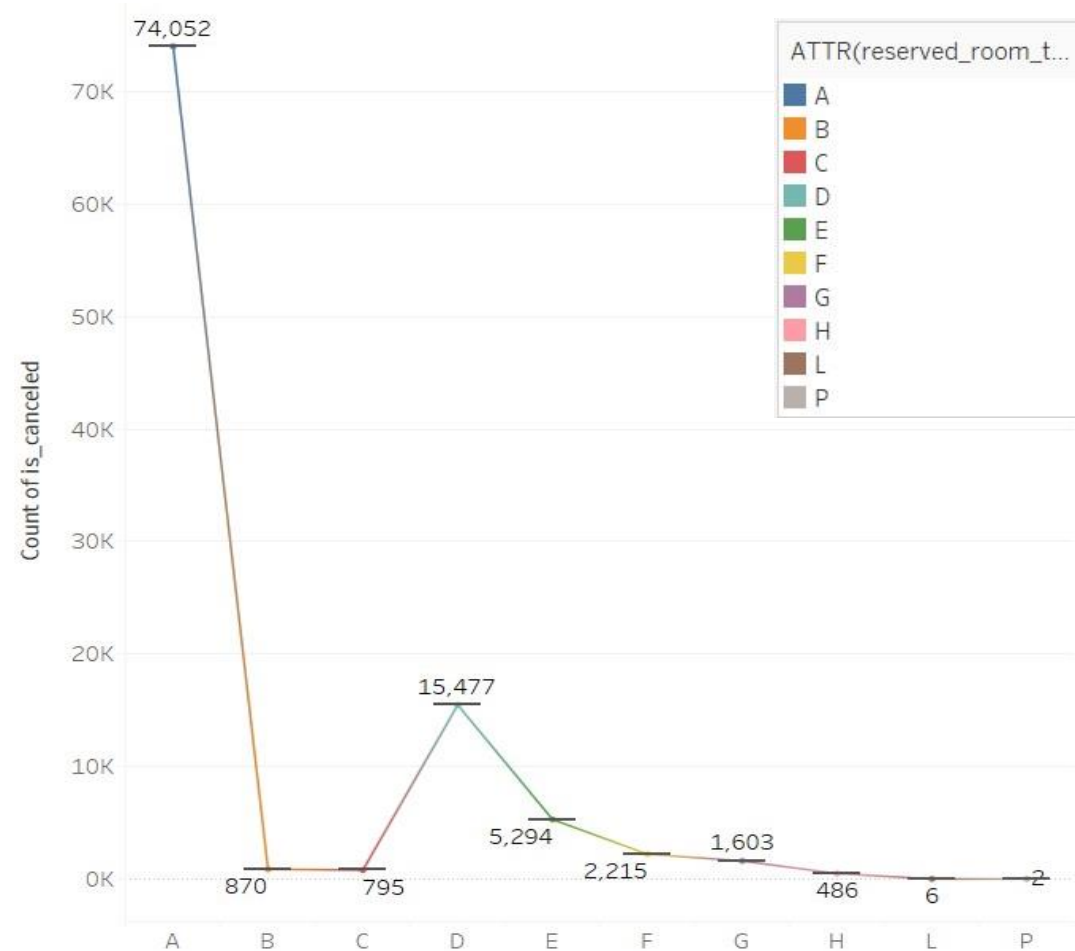


➤ Booking Trend With Respect To Room Type



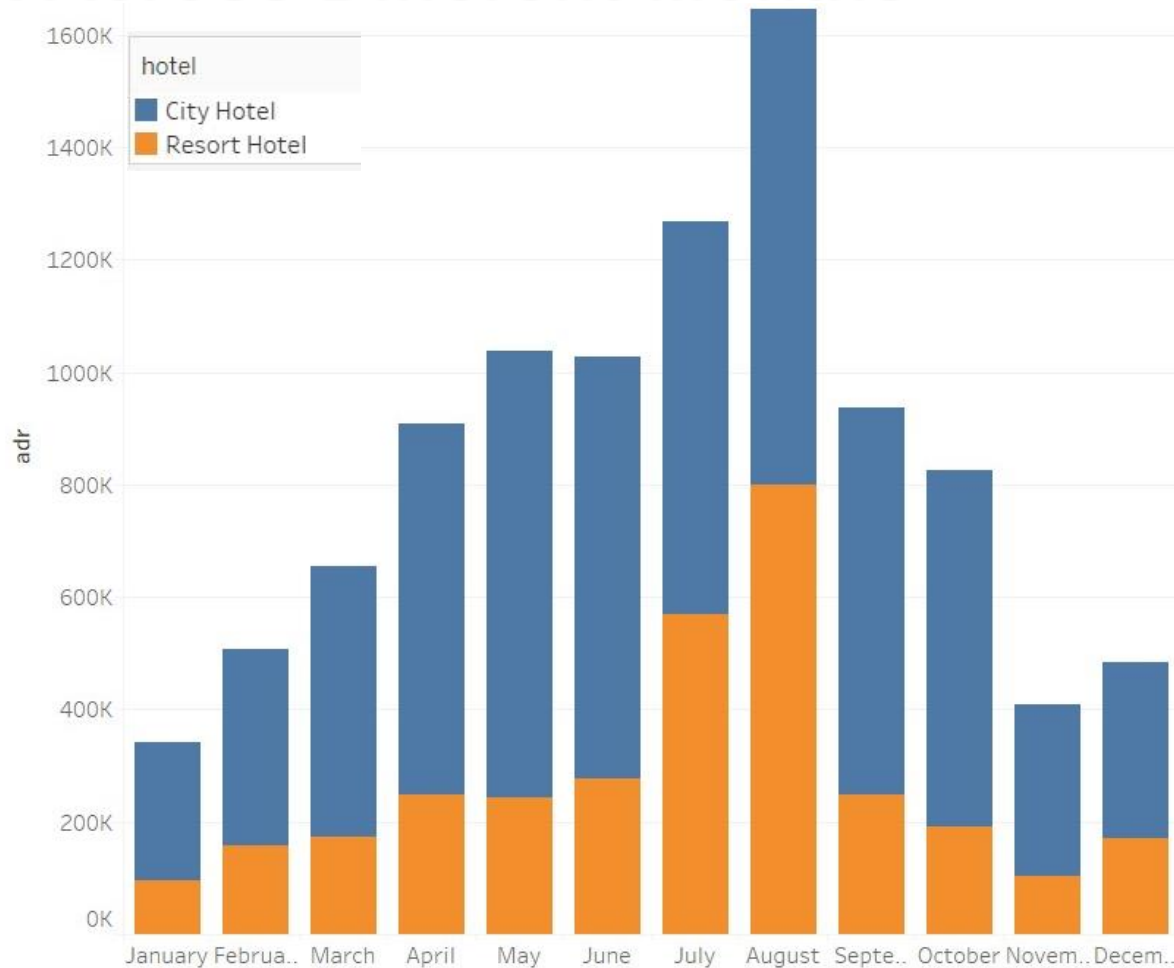
A - type of room is the most favourite in all types of customers covering all the market about more than 85% ,the D – type of room is at second place in queue while negligible customers are there which are ready to stay in L - type and P – type of room.

So we need to upgrade L- type and P – type of room to attract more customers so that no one should be in waiting list and do not search any other hotel which results in increasing the profit of hotels as more customers will book the rooms in the hotels.



➤ Variation In ADR Across Different Months

7. For resort hotels, the Average Daily Rate (ADR) is more expensive during August, July, June and September where it is lower for January and November. For city hotels, the Average Daily Rate (ADR) is more expensive during August, July, May and June where it is lower also for January and November.
8. So overall Average Daily Rate of both city hotels and resort hotels are more expensive between May and September.



➤ Conclusion

1). Out of 119000 customers ,75000 of customers checked-in the hotel while 44000 of customers canceled their bookings. If we talk about percentage, 37% of bookings got canceled Whereas 63% of customers did check-in. So, We realize that the high rate of cancellations can be due to no deposit policies.

2). We can see that 2016 seems to be the year where hotel booking is at its highest. We also see an increasing trend in booking around the middle of the year, with August being the highest followed by July and May. Summer ends around August, followed straight by autumn. It seems that summer period is a peak period for hotel booking . We can also see that January month has lowest number of customers followed by November and December .It seems Winter period is at lowest peak for hotel booking .

So for Winter season price of hotel booking should be low or they should provide some special discount so that most of the customers attract to the hotel and also it can be the best time of year to book a hotel room. Due to less demand of rooms, the cost for room on daily basis also minimum as compared to other months where month of August has high demand for room so it is obvious that the cost of room is also at peak

3). Most of the customer prefer to stay one, two, three and four nights where number of bookings for two and three nights has high demand.

So if we try to analyze the optimal length to stay in order to get best daily rate, we can say that two to four days of accommodation can be more affordable since more people prefers the same, the hotels also launches the offers for mid term accomodation because long term stay offers will consume their rooms and new customers can not avail the rooms and hotels may fail to generate extra profits.

➤ References

- 1). <https://pandas.pydata.org/>
- 2). <https://matplotlib.org/>
- 3). <https://seaborn.pydata.org/>
- 4). Tableau Visualization

➤ Challenges

- 1). Huge chunk of data was to be handled by keeping in mind not to miss anything which is even of little relevance.
- 2). Handling with too many null values and replacing it.

Thank You!