

# PROJECT REPORT ON HOTEL BOOKING ANALYSIS



## TEAM MEMBERS:

1. MS. VARSHA SAXENA
2. MR. ANUP SINGH
3. MR HRITHIK VINODIA



## **HOTEL BOOKING ANALYSIS:**

The project contains the real-world data record of hotel bookings of a city and a resort hotel containing details like bookings, cancellations, guest details etc. from 2015 to 2017. In this project we are going to analyze Hotel Booking Data in order to find out valuable insights and give suggestions to increase revenue of hotels.

- **Programming Language:** Python
- **Libraries used:** Pandas, NumPy, Matplotlib, Seaborn
- **Notebook:** Google Collab
- **Dataset Source:** Provided by Almbetter

## **OBJECTIVE:**

We are provided with a hotel bookings dataset.

The main purpose of this study is to perform EDA on the given dataset and draw useful conclusions about the trends in hotel bookings and how factors that control hotel bookings influence each other.

## **BUSINESS PROBLEM:**

In recent years, City Hotel and Resort Hotel have seen high cancellation rates. Each hotel is now dealing with several issues as a result, including fewer revenues and less than ideal hotel room use. Consequently, lowering cancellation rates is both hotels' primary goal in order to increase their efficiency in generating revenue, and focus to offer thorough business advice to address this problem.

The analysis of hotel booking cancellations as well as other factors that have no bearing on their business and yearly revenue generation are the main topics of this report.

## **ABOUT THE DATASET:**

This dataset contains 119390 observations for a City Hotel and a Resort Hotel. Each observation represents a hotel booking between the 1<sup>st</sup> July 2015 and 31<sup>st</sup> August 2017, including booking that effectively arrived and booking that were cancelled. Since this is hotel real data, all the elements pertaining hotel or customer identification were deleted. Four



Columns ,name, email, phone no. and credit-card have been artificially created and added to the dataset.

Field	Description
Hotel	H1= Resort Hotel H2=City Hotel
is_cancelled	If the booking was cancelled(1) or not(0)
lead_time	Number of days that elapsed between the entering date of the booking into the PMS and the arrival date
arrival_date_year	Year of arrival date
arrival_date_month	Month of arrival date
arrival_date_week_number	Week number for arrival date
arrival_dat_day	Day of arrival date
stays_in_weekend_nights	Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel
stays_in_week_nights	Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel
adults	Number of adults
children	Number of children
babies	Number of babies
meal	Kind of meal opted for
country	Country code
market_segment	Which segment the customer belongs to

## **ASSUMPTIONS:**

1. No unusual occurrences between 2015 and 2017 will have a substantial impact on the data used.
2. The information is still current and can be used to analyze a hotel's possible plans in an efficient manner.
3. There are no unanticipated negatives to the hotel employing any advised technique.
4. The hotels are not currently using any of the suggested solutions.
5. The biggest factor affecting the effectiveness of earning income is booking cancellations.
6. Cancellations result in vacant rooms for the booked length of time.
7. Clients make hotel reservations the same year they have make cancellations.

## **RESEARCH QUESTIONS:**

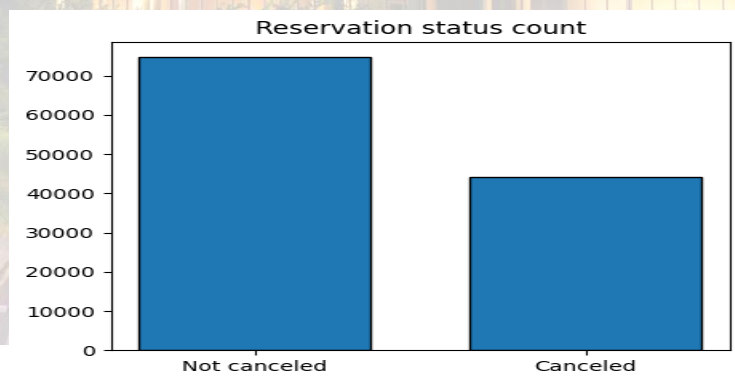
1. Which factor has significant impact on the revenue of the hotels?
2. Is cancellation more in city or resort hotel?
3. What are the variables that affect hotel reservation cancellations?

4. In which month the cancellation is the highest and in which month it is the lowest?
5. In which country the cancellation rate is the highest? Also show top 10 countries whose cancellation rate is highest?
6. How many guests are there who have repeatedly cancelled the bookings?
7. In which hotel people like to stay and spend more time depending on the market segment?
8. How many bookings were cancelled? Calculate its percentage depending on the hotel type.
9. What are the effects on cancellations by market segments.
10. What are the various types of the customers in the data set and also show how much percentage of customers are coming from which type?
11. How can we make hotels reservations cancellations better?
12. How will hotels be assisted in making pricing and promotional decisions?

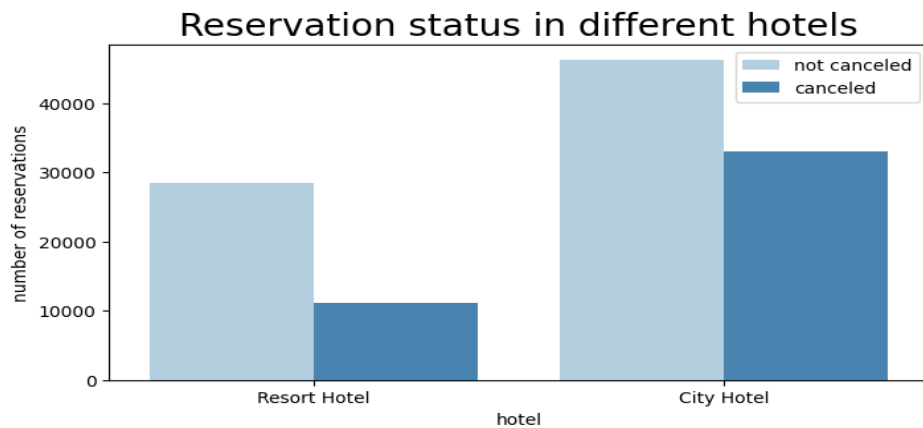
### **HYPOTHESIS:**

1. More cancellations occur when the prices are higher.
2. When there is a longer waiting list, customers tend to cancel more frequently.
3. Most clients are coming from offline travel agents to make their reservations.
4. Transient Groups must be giving a large number of customers.
5. Hotels must be busy during May-July and December since more holidays are there.
6. Aviation people must not be going to resort more since they city hotels are closer to airport therefore there stay must be in cities hotels rather than resort hotels.

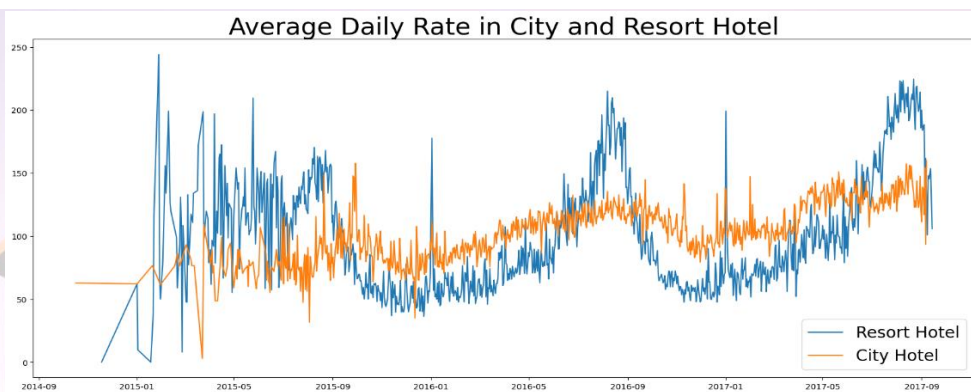
### **ANALYSIS AND FINDINGS:**



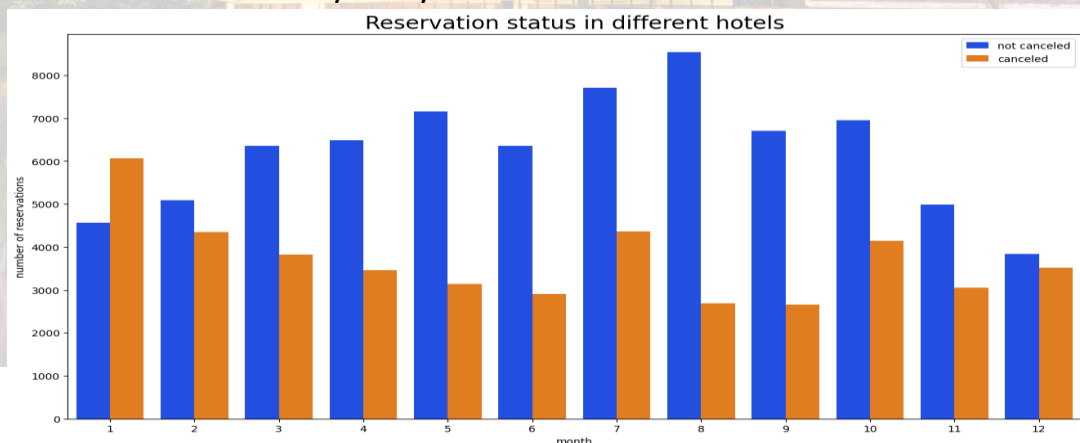
**Findings:** The accompanying bar graph shows the percentage of reservations that are cancelled and those that are not. It is obvious that there are still a significant number of reservations that have not been canceled. There are still 37% of the clients who canceled their reservation, which has a significant impact on the hotel's earning.



**Findings:** In comparison to resort hotels, city hotels have more bookings. It's possible that resort hotels are more expensive than those in cities.

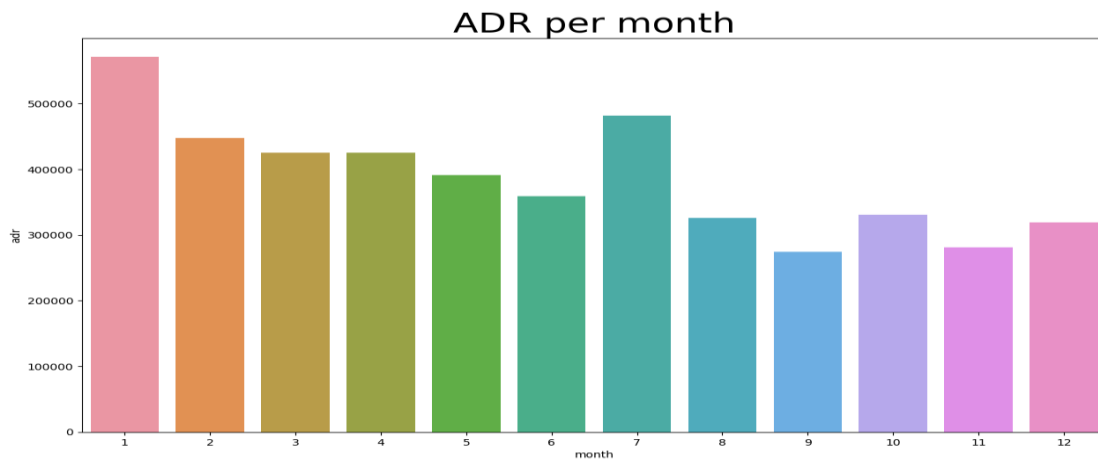


**Findings:** The line graph above shows that on certain days, the average daily rate for a city hotel is less than that of a resort hotel, and on the other days, it is even less. It goes without saying that weekends and holidays may see a rise in resort hotel rates.

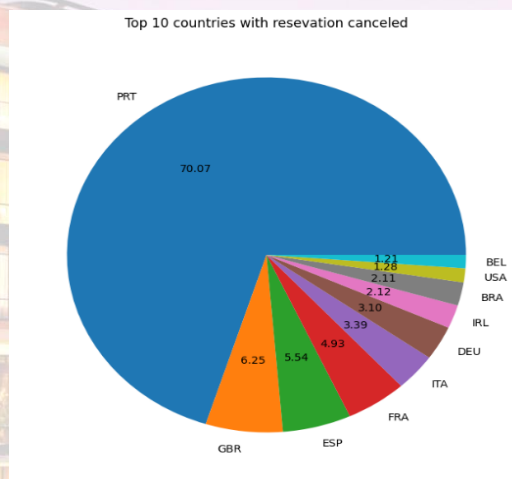


**Findings:** We have developed the grouped bar graph to analyze the months with the highest and lowest reservation levels according to reservation status. As can be seen, both the number of confirmed reservations and the number of canceled reservations are largest in the month of August, whereas January is the month with the most canceled reservations.

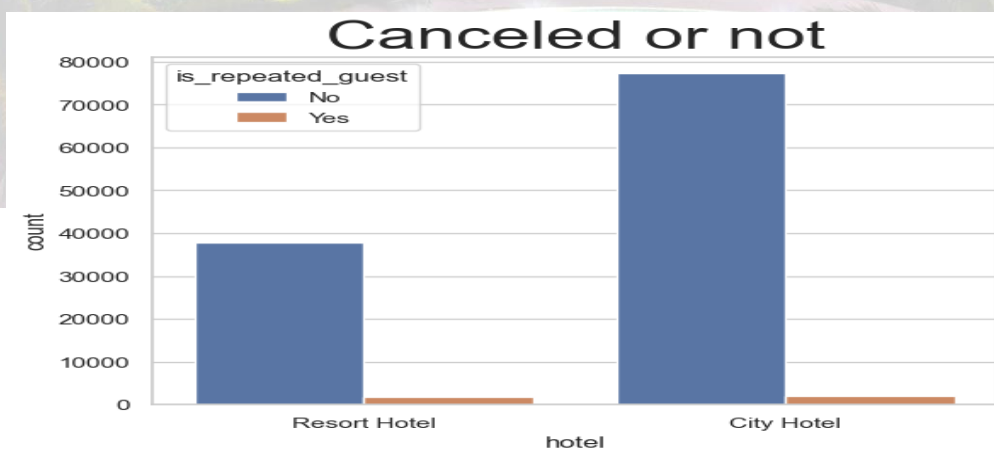




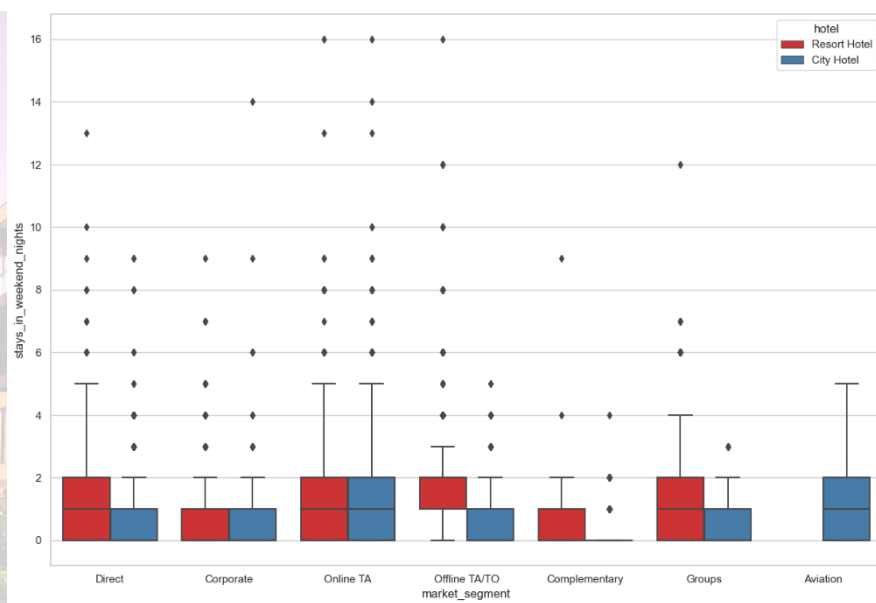
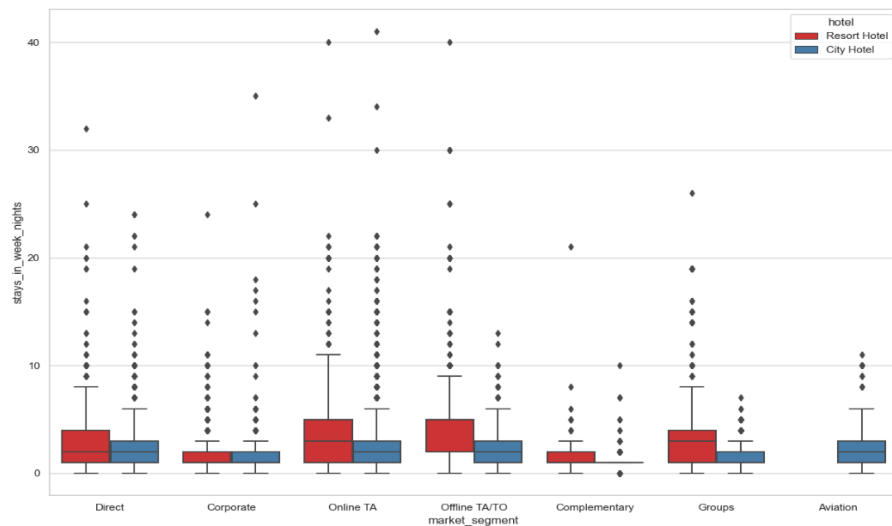
**Findings:** This bar graph demonstrates that cancellations are most common when the prices are greatest and are least common when they are lowest. Therefore the cost of the accommodation is solely responsible for the cancellation



**Findings:** The above pie chart demonstrates the top 10 countries with highest rate of cancellations. The top country is Portugal with the highest number of cancellations.



**Findings:** The above graph shows how many no. of guests are there who have repeatedly canceled the bookings.

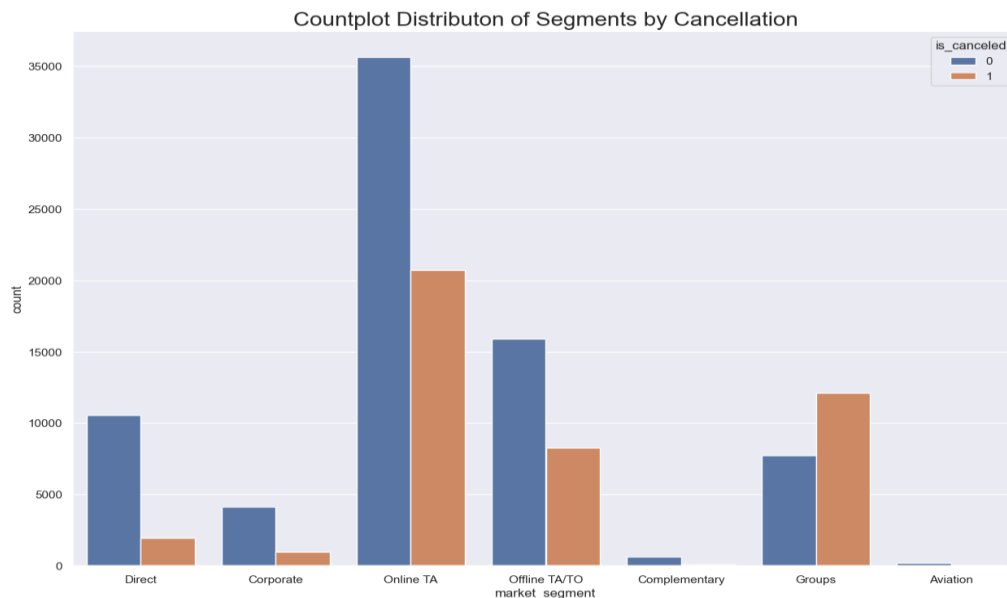


**Findings:** The above 2 graphs shows that in which hotel people like to stay and spend more time depending on the market segment. the first one is for week nights and the second one is for weekend nights

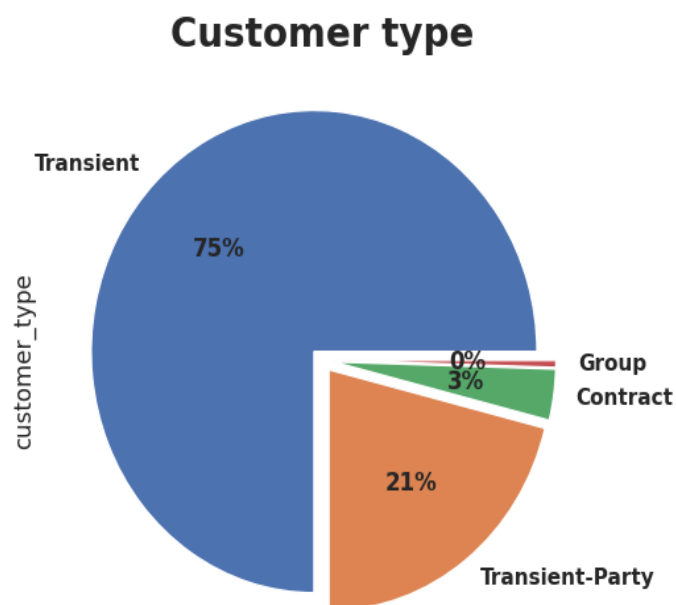
It can be seen that most of the groups are normally distributed, some of them have high skewness. Looking at the distribution, it can be concluded that most people do not seem to prefer to stay at the hotel for more than 1 week. But it seems normal to stay in resort hotels for up to 12–13 days.

As it turns out, customers from Aviation Segment do not seem to be staying at the resort hotels and have a relatively lower day average. Apart from that, the weekends and weekdays averages are roughly equal. Customers in the Aviation Segment are likely to arrive shortly due to business. Also probably most airports are a bit away from sea and its most likely to be closer to city hotels.

It is obvious that when people go to resort hotels, they prefer to stay more.



**Findings:** The cancellation rate for groups is greater than 50%.The cancellation rate for offline TA/TOs (Travel Agents/Tour Operators) and online TAs is greater than 33%.Direct segments have a lower rate of cancellation. Online reservation cancellation rates are typical for a dynamic environment with significant circulation.



**Findings :** From the above graph we can easily see that maximum customers are from transient category i.e nearly 75% whereas from groups and contract we are having a negligible customers.

## **SUGGESTIONS:**



1. Cancellation rates rises as the prices does. In order to prevent cancellations of reservations, hotel could work on their pricing strategies and try to lower the rates for specific hotels based on locations. They can also provide some discounts to the customers.
2. As the ratio of the cancellation to the non-cancellation of the resort hotels is higher than the city hotels. So the hotels should provide a reasonable discount on the room prices on weekends or on holidays.
3. In the month of January, hotels can start campaigns or marketing with a reasonable amount to increase their revenue as the cancellations is the highest in this month.
4. They can also increase the quality of their hotels and their services mainly in Portugal to reduce the cancellation rate.
5. Repeated guests do not cancel their reservations. Of course, there are some exceptions. So we can offer some discount or privilege in terms of giving one meal free or any services free can be given to attract them back.

