Hotel Booking Cancellation Analysis Report

Business Problem:

In recent years, City Hotel and Resort Hotel have been seen high cancellation rates. Each Hotel is now dealing with a number of issues as a result, including fewer revenues and less than ideal hotel room use. Consequently, lowering cancellation rates in both hotel's goal to increase their efficiency in generating revenue, and for us to offer through 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.

Objective Of the Analysis:

This analysis aims to delve into the dynamics of hotel booking cancellations for both City Hotel and Resort Hotel. By exploring various factors influencing cancellations and those unrelated to their core business and yearly revenue generation, our goal is to provide actionable insights and business recommendations that will empower the hotels to curtail cancellation rates and optimize their revenue streams.

About Data:

The dataset used for this analysis was obtained from Kaggle, a platform that hosts a diverse range of datasets for public use. The specific dataset, titled **Hotel Bookings**, was selected due to its relevance to the hotel industry and comprehensive coverage of hotel booking information.

This dataset contains 119390 observations for a City Hotel and a Resort Hotel. Each observation represents a hotel booking between the 1st of July 2015 and September 2017, including booking that effectively arrived and booking that were cancelled.

Assumptions:

1. No unusual Occurrence between 2015 and 2017 will have 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 advanced 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. Clients make hotel reservations the same year they make cancellations.

Research Questions:

1. What are the variables that affect hotel reservation cancellations?
2. How can we make Hotel reservations cancellations better?
3. How will hotels be assisted in making pricing and promotional decisions?

Hypothesis:

1. More cancellations occur when prices are higher.
2. When there is longer waiting list, customers tend to cancel more frequently.
3. Majority of clients are coming from offline travel agents (TA) to make their reservations.

Exploratory Data Analysis (EDA):

## Data Characteristics:

Dataset has 119390 rows and 32 columns.

Data Cleaning:

*   The 'reservation_status_date' column, initially represented as an object data type, has been transformed into the datetime data type. This conversion allows for more effective handling of date-related operations.
*   Dataset contains company, agent, country, and children column has null values but company and agent columns have large number of missing values and are not be the part of analysis so Dropped both columns. (now dataset has 30 columns)
*   Now removing null rows because remaining two columns country and children has some null values.
*   Null rows were removed from the dataset due to the presence of null values in the country and children columns. This step was taken to enhance the dataset's completeness for subsequent analysis, ensuring that any missing values in these columns did not compromise the integrity of overall dataset.
*   only one value of ADR (Average Daily Rate) is more than 5000 (outliner) so we can remove or can use it. one value of ADR = 5400 has removed and remaining are less than 5000. Dataset is now cleaned and ready for Analysis and visualization.
*   We found an unusual value in the 'adr' (Average Daily Rate) column, specifically one value exceeding 5000. Since this seemed like an outlier, we decided to remove it. After this cleaning step, all the remaining 'adr values are now below 5000.
*   The cleaned dataset comprises 118,897 entries and 30 columns, rendering it well-prepared for subsequent analysis and visualization.

Analysis and Findings:
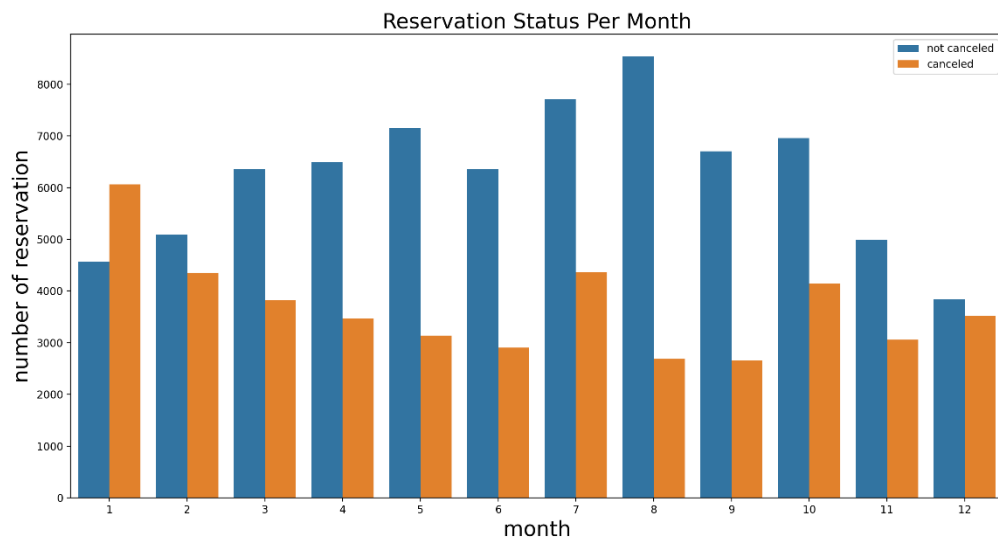
Reservation Status Count
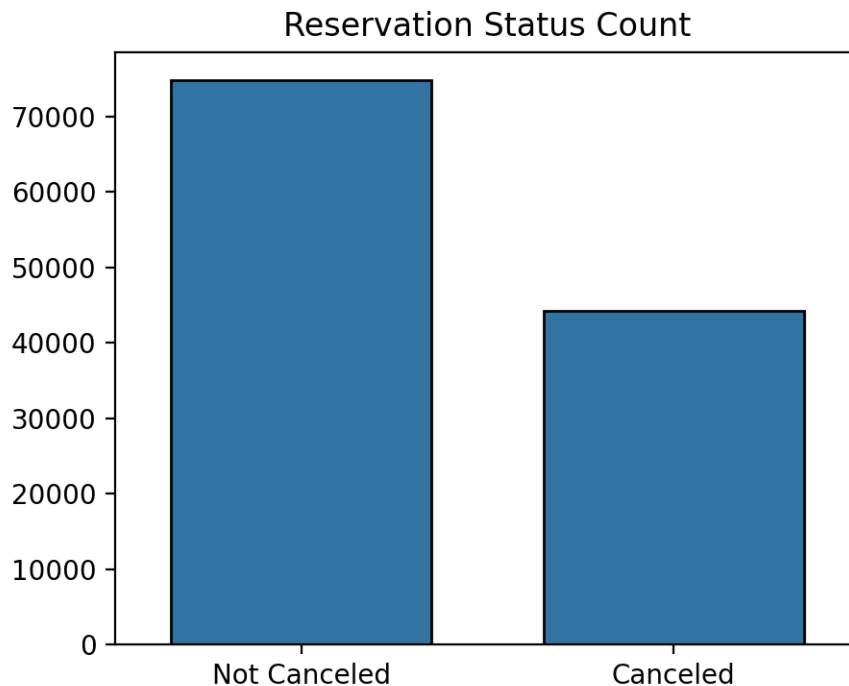
The accompanying bar graph shows the percentage of reservations that are cancelled and those that are not. It is obvious that are still a significant number of reservations (approximately 62%) that have not been cancelled. There is still 37% of clients who cancelled their reservation, which has significant impact on the hotel's earnings.

Reservation status in different hotels
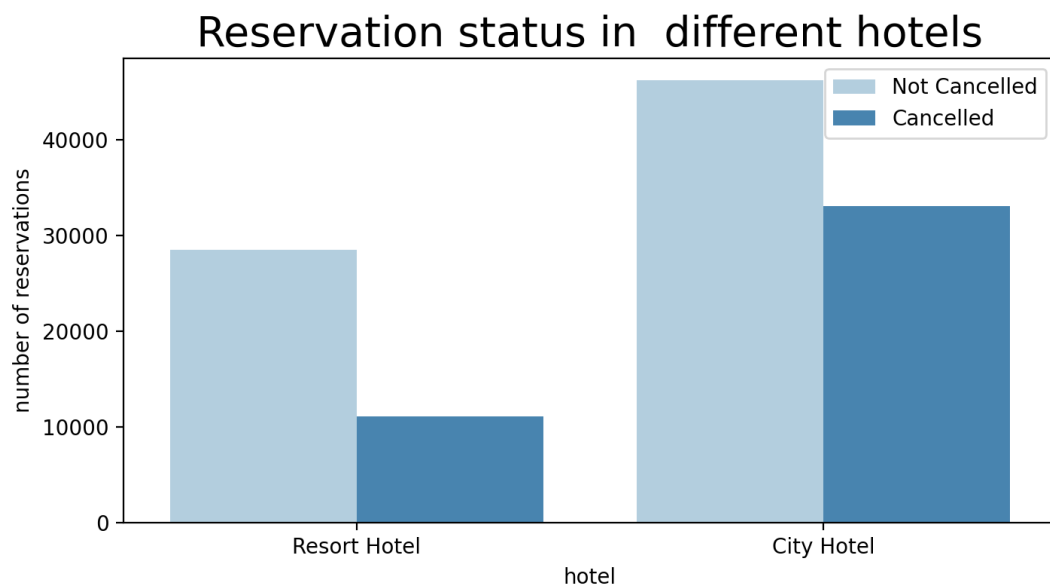
In comparison to Resort Hotels, City Hotels have more bookings. It is possible that Resort Hotels are more expansive than those in Cities. Resort Hotel has 27.97% cancellations and City Hotel has 41.70% from this analysis we can conclude that City Hotel has more cancellation rate compared to Resort Hotel.

Average Daily Rate in Resort Hotel and City Hotel
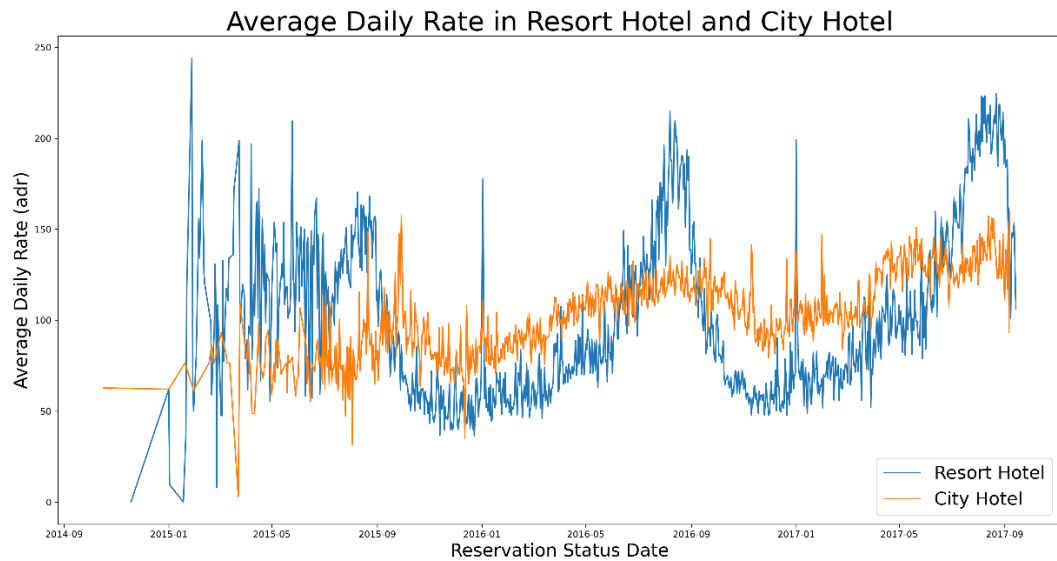
The line graph above shows that, on certain days, the average daily rate for City Hotel is less than that of Resort Hotel, and other days, it is even less. Notably, an upward trend is evident on weekends and holidays, suggesting potential price surges for City Hotel during these periods.

**Reservation Status Per Month**

We have developed the grouped bar graph to analyze the months with highest and lowest reservation levels according to reservation status. The analysis reveals that August stands out with the highest number of non-cancellations, while January displays the least. Conversely, January experiences the highest number of cancellations, while August demonstrates the lowest. This pattern suggests a seasonal influence on booking behaviors throughout the year.

ADR per month
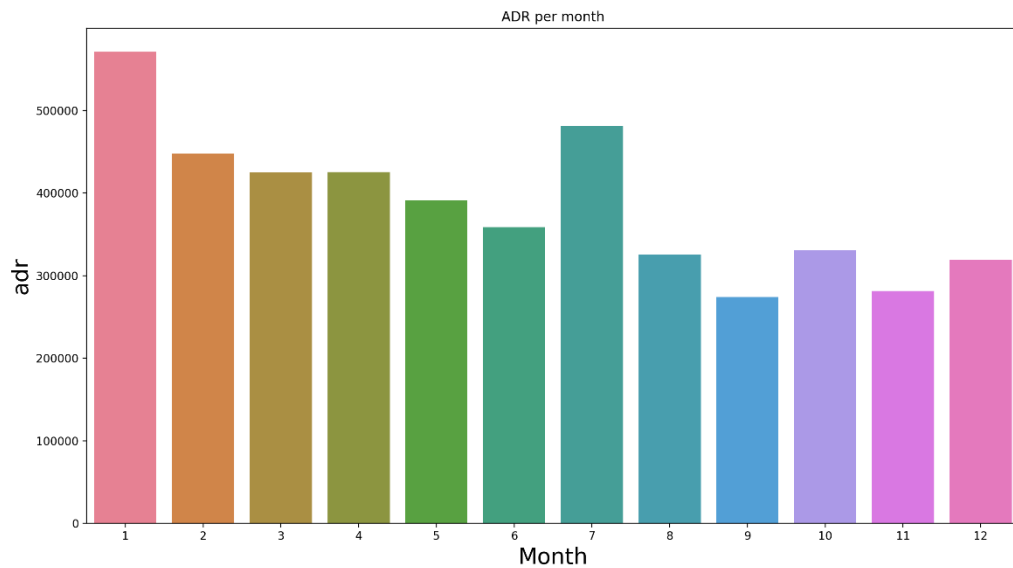
The plot supports the hypothesis that lower prices, particularly observed in August, contribute to higher non-cancellations. Conversely, the evidence suggests a correlation between increased Average Daily Rate (ADR) and higher cancellations. This finding underscores the significant impact of pricing on guest behavior, affirming that elevated prices tend to correlate with a higher likelihood of cancellations. The observed trend emphasizes the importance of pricing strategies in shaping reservation outcomes.

Now, let see which country has the highest number of reservations cancelled. Portugal (PRT) exhibits the highest cancellation rate at 70.07%. To mitigate this, consider implementing the following strategies: Improve Facilities, offer competitive and reasonable prices to attract more bookings, Promotional Discounts, Marketing Campaigns, gather feedback from guests, especially cancellations, to inform improvements. Implementing these strategies can contribute to reducing cancellation rates and enhancing guest satisfaction.

Top 10 countries with reservation cancelled
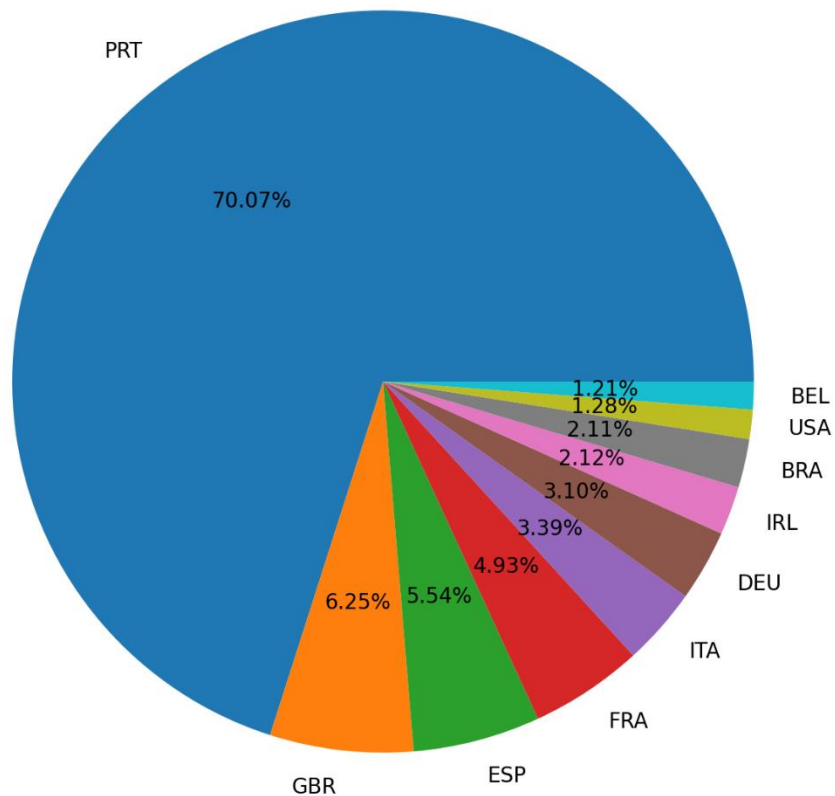
Contrary to the initial hypothesis, the data suggests that most customers are coming through online travel agents (OTAs) rather than offline channels. This shift in understanding emphasizes the importance of accurately assessing and adapting to the predominant customer acquisition channels for effective marketing and business strategies.

Let's check the area from which guests are visiting the hotels and marketing reservations. Is it coming from Direct or Groups, Online or Offline travel agents? The majority of clients, around 47%, opt for online travel agencies (OTAs), while 20% choose offline travel agencies, and only 10% book directly. Notably, 46% of reservations made through online travel agents experience cancellations. This observation prompts a deeper exploration of factor contributing to cancellations in the online TA segment. Addressing these factors could enhance the booking experience for customers, especially those coming through OTAs.

Average Daily Rate
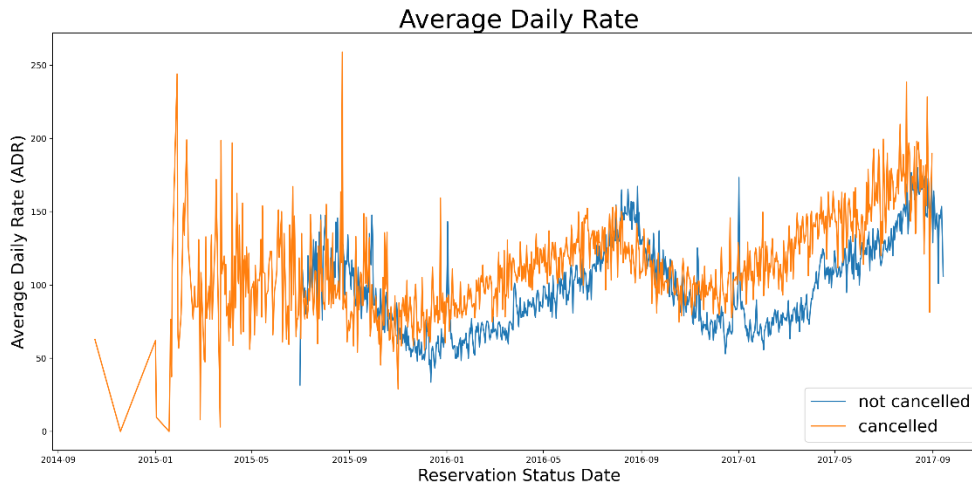
Focusing on the period from 2016 to September 2017 due to data inconsistency, our plot highlights a noteworthy correlation: as prices (ADR) increase, cancellations tend to rise. This suggests a discernible connection between pricing and cancellation patterns.

Additionally, intriguing spikes in the plot occurring on weekends or holidays indicate a potential influence of timing on reservation patterns. These insights underscore the importance of considering both temporal and pricing factors in understanding and managing reservation behaviors.

As seen in the graph, reservations are cancelled when the average daily rate is higher than when it is not cancelled. It clearly proves all the above analysis, that the higher price leads to higher cancellations.

Suggestions:

1. Cancellation rates rise as the price does. In order to prevent cancellations of reservations, hotels could work on the 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 and not cancellation of the Resort Hotel is higher in Resort Hotel 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 cancellation 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. Strengthen partnerships with online travel agents (OTAs) and tailor strategies to reduce cancellations.
6. Monitor the impact of implemented strategies and adjust based on ongoing analyses.
7. Establish a robust feedback mechanism to gather insights from guests, especially cancellations.