## Business Problem

In recent years, City Hotel and Resort Hotel have 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 is both hotels' primary goal in order to increase their efficiency in generating revenue, and for us 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.

## 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 analyse 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 make cancellations.

## Research Question

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?

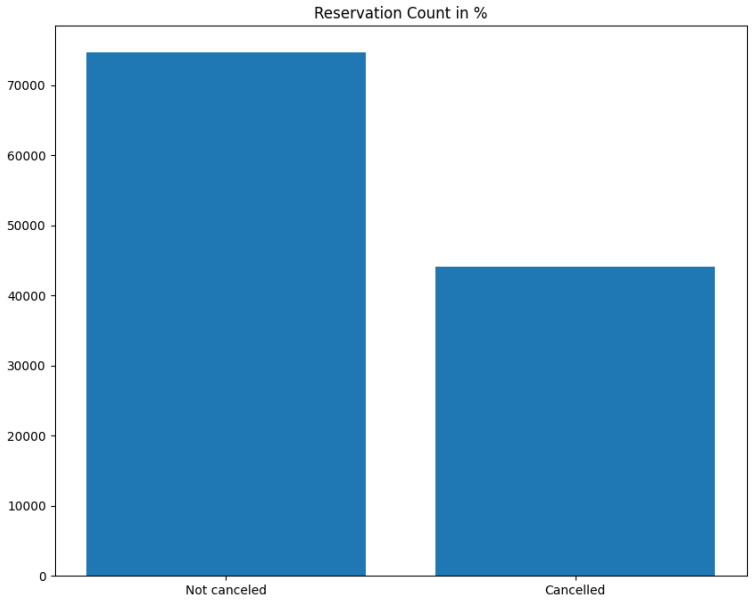
## Hypothesis

1. More cancellations occur when prices are higher.

2. When there is a longer waiting list, customers tend to cancel more frequently.

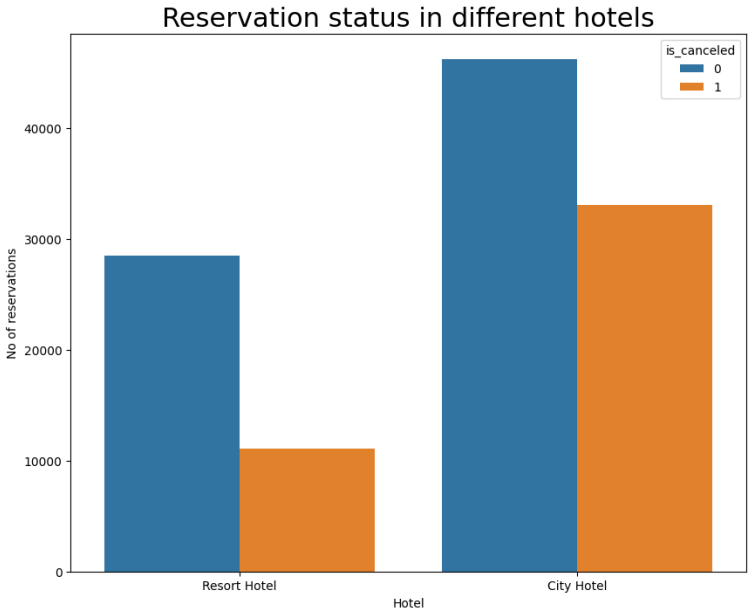
3. The majority of clients are coming from offline travel agents to make their reservations.

## Analysis and Findings

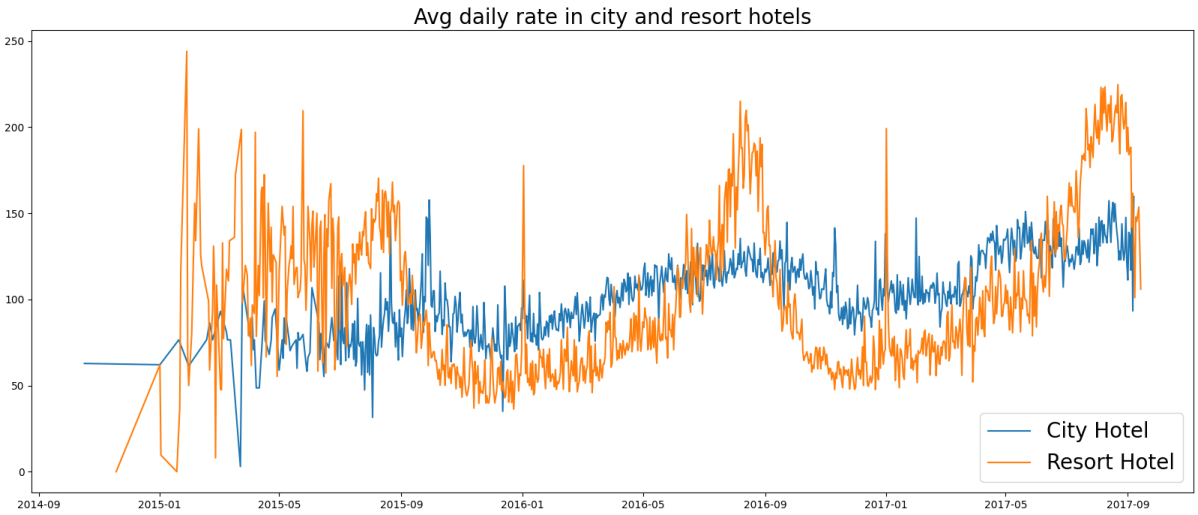



```
is_canceled
0    0.628653
1    0.371347
Name: proportion, dtype: float64
```

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 cancelled. There are still 37% of clients who cancelled their reservation, which has a significant impact on the hotels' earnings.



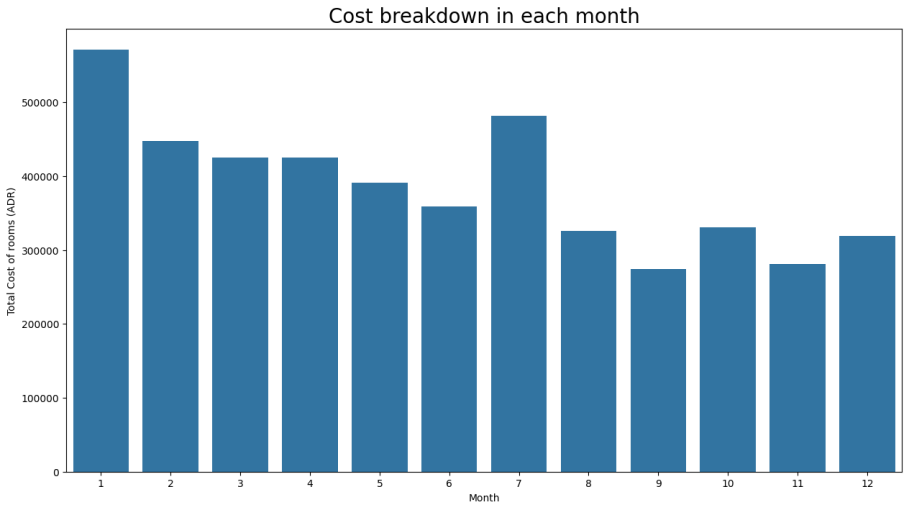

In comparison to resort hotels, city hotels have more bookings. It's possible that resort hotels are more expensive than those in cities.




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 other days, it is even less. It goes without saying that weekends and holidays may see a rise in resort hotel rates.

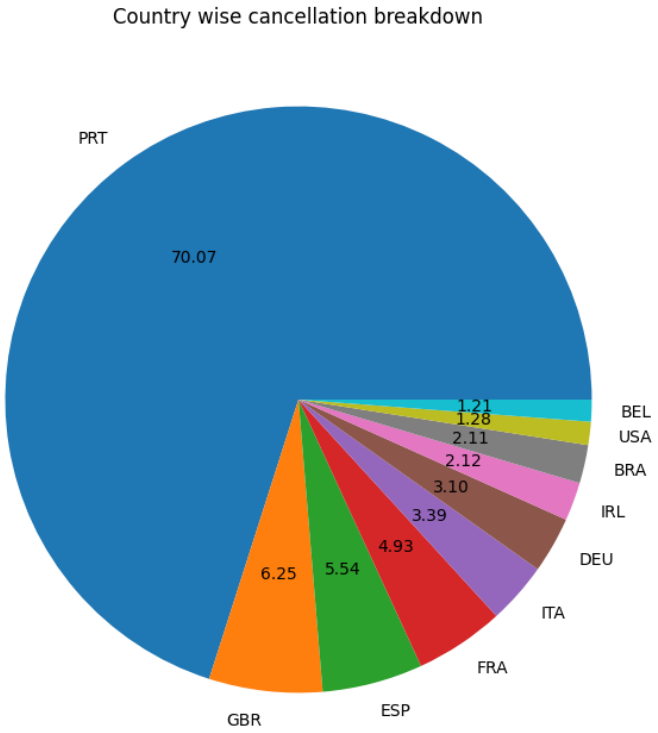

We have developed the grouped bar graph to analyze the months with the highest and lowest reservation levels according to reservation status. As can be seen, both the number of confirmed reservations and the number of cancelled reservations is largest in the month of August, whereas

January is the month with the most cancelled reservations. So ideally, the next step would be to check whether it was the cost of rooms which was playing a major role in cancellation rates in such months.




Now, let's see which country has the highest reservation cancelled.

```
country
PRT    27514
GBR     2453
ESP     2177
FRA     1934
ITA     1333
DEU     1218
IRL      832
BRA      830
USA      501
BEL      474
Name: count, dtype: int64
```




The top country was found out to be **Portugal** with the highest number of cancellations (appx 70%).

Now, Let's check the market segment from where guests are visiting the hotels and making reservations. Is it coming from Direct or Groups, Online or Offline Travel Agents?

Around 47.4% of the clients come from online travel agencies, whereas 20.3% come from Offline travel agents. Only 10% of clients book hotels directly by visiting them and making reservations.

```
market_segment
Online TA        0.474377
Offline TA/TO    0.203193
Groups           0.166581
Direct           0.104696
Corporate        0.042987
Complementary    0.006173
Aviation         0.001993
Name: proportion, dtype: float64
```

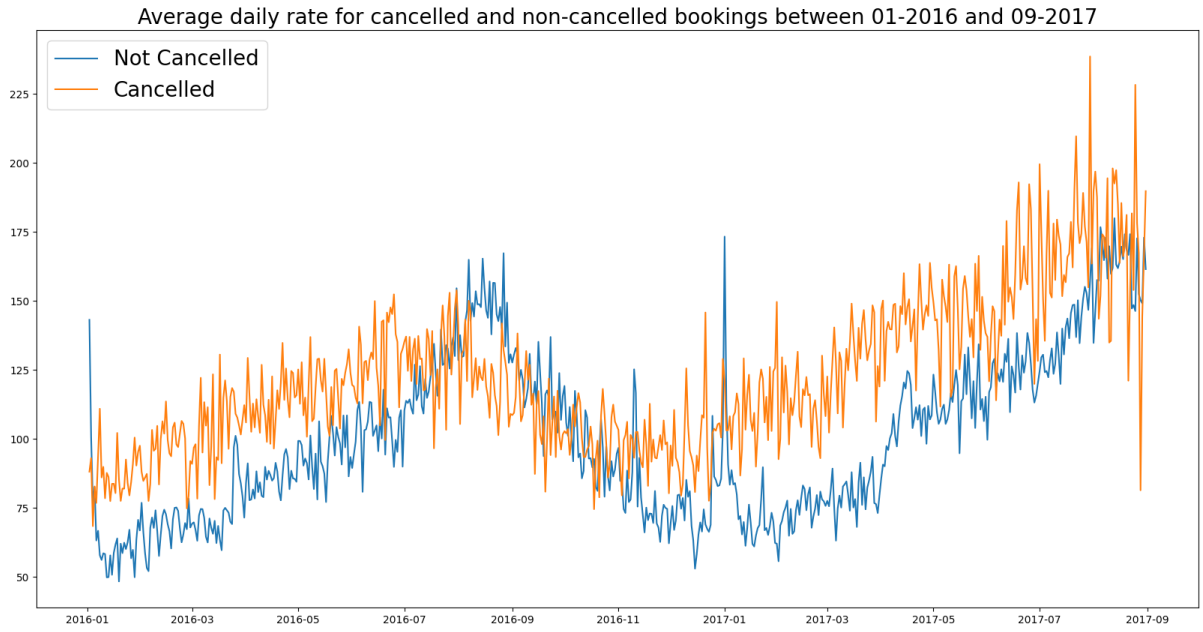
It is important to note that for bookings which were cancelled, 46.9% of the clients cancelled when they booked from online travel agencies, whereas 27.3% cancellations came when booked from Groups. Only 4% of client cancellations came from customers who visited hotels directly and made reservations.

```
df_canceled['market_segment'].value_counts(normalize=True)
```

[68] ✓ 0.0s

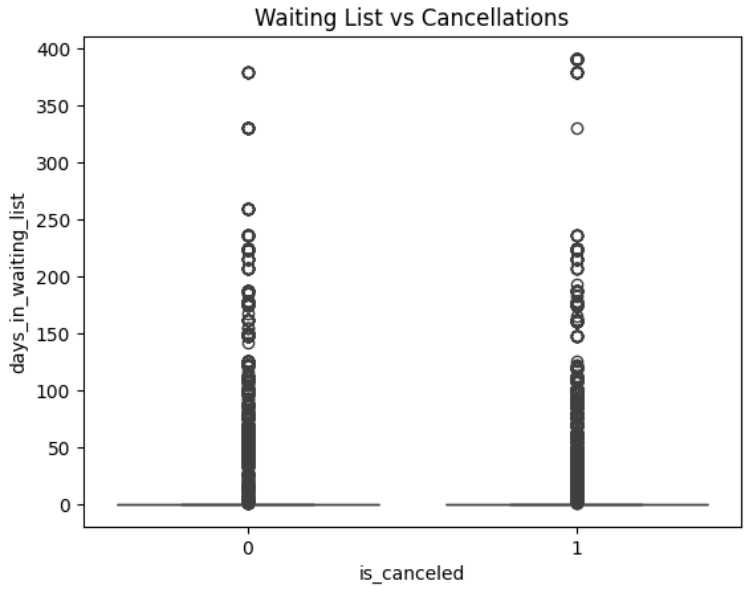
```
market_segment
Online TA        0.469696
Groups           0.273985
Offline TA/TO    0.187466
Direct           0.043486
Corporate        0.022151
Complementary    0.002038
Aviation         0.001178
Name: proportion, dtype: float64
```

Now, we deep dive into average daily rate for cancelled and non-cancelled bookings. To get a better picture, we analyse data between 01-2016 to 09-2017 as we have consistent data between these timelines to make better assessment of our results.




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 is the price, higher is the cancellation rate of bookings.

Contrary to expectation, customers on **short waiting lists** show higher cancellation rates compared to those on **medium or long waiting lists**. This suggests that customers willing to wait longer are more committed, while short-wait bookings may reflect opportunistic or tentative reservations.




On performing Corelation (Spearman Rank) test on waiting list and cancellation status variables, we observed following outcomes:

- $p < 0.05$, implying that the correlation is statistically significant; and
- A positive correlation suggests longer waiting lists → higher chance of cancellation

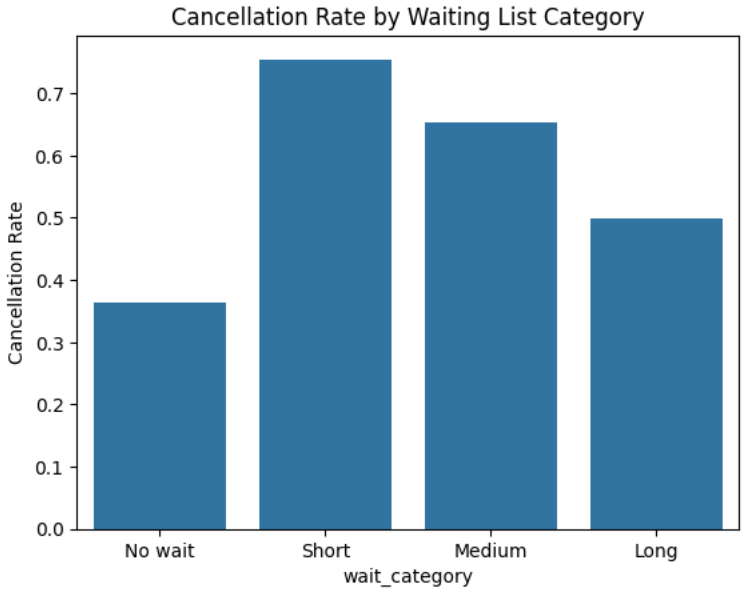
```
from scipy.stats import spearmanr

corr, p = spearmanr(df['days_in_waiting_list'], df['is_canceled'])
print("Spearman correlation:", corr, "p-value:", p)
```

✓ 10.5s Python

```
Spearman correlation: 0.0980478300755052 p-value: 9.470158509459084e-252
```

Furthermore, categorising wait times into bins, performing a Chi-Square test and generating visualisation gave us some intriguing results.

Chi2: 1261.432224521698 p-value: 3.43707254016225e-273




So,

- Short waiting lists → higher cancellation rates
- Medium & long waiting lists → lower cancellation rates

This might be happening because of several reasons:

1. **Customer commitment effect:**
   - People willing to wait longer may be more committed to actually staying.
   - Customers with short waits might be more "opportunistic" and cancel easily if they find a better deal.
2. **Inventory management:**
   - Hotels may prioritize high-demand bookings on long waiting lists (corporate or group bookings), which are less likely to cancel.
3. **Behavioural aspect:**
   - Short waits could represent "last-minute" bookings → higher volatility.
   - Longer waits often = planned trips (vacations, events) → cancellations less likely.

## Suggestions

1. Cancellation rates rise as the price does. In order to prevent cancellations of reservations, hotels could work on their pricing strategies and try to lower the rates for specific hotels based on locations.

2. As the ratio of the cancellation and non-cancellation is higher in the city hotel than the resort hotels. So, the hotels could provide a reasonable discount on the room prices on weekends or on holidays.

3. In the month of December - January, hotels can start campaigns or marketing with a reasonable amount to increase their revenue as the cancellation is the highest in this month.