

# 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 less revenue 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 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 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?

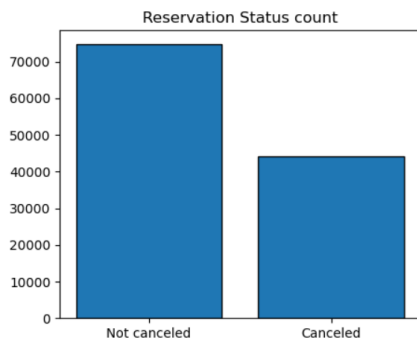
## hypomnesia

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 come from offline travel agents to make their reservations.

## Analysis and findings

```
In [56]: 1 cancelled_prec = df['is_canceled'].value_counts(normalize = True)
2 print(cancelled_prec)
3
4 plt.figure(figsize = (5,4))
5 plt.title('Reservation Status count')
6 plt.bar(['Not canceled', 'Canceled'], df['is_canceled'].value_counts(), edgecolor = 'k')
7 plt.show()

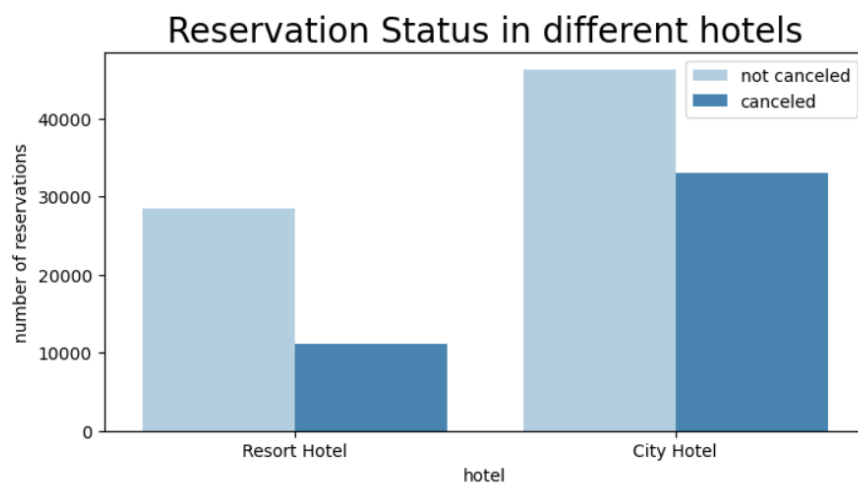
0    0.628653
1    0.371347
Name: is_canceled, dtype: float64
```



The accompanying bar graph shows the percentage of reservations that are canceled 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 clients who canceled their reservation, which has a significant impact on the hotels' earnings.

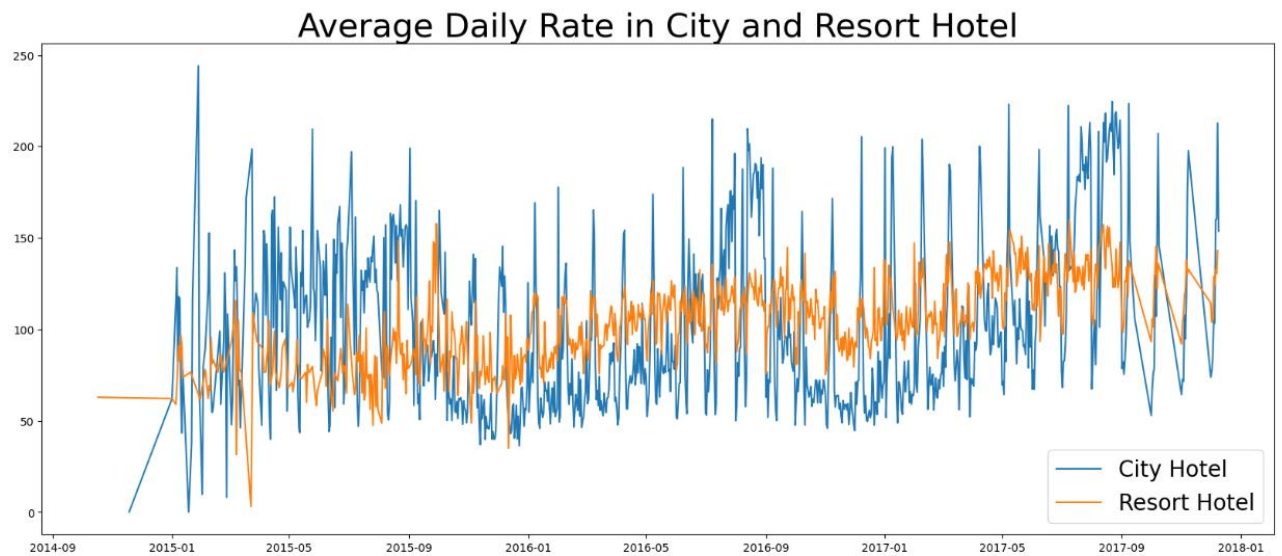
## Reservation status in different hotels

```
In [74]: 1 plt.figure(figsize = (8,4))
2 ax1 = sns.countplot(x = 'hotel', hue = 'is_canceled', data = df, palette = 'Blues')
3 legend_labels = ax1.get_legend_handles_labels()
4 ax1.legend(bbox_to_anchor = (1,1))
5 plt.title("Reservation Status in different hotels", size = 20)
6 plt.xlabel('hotel')
7 plt.ylabel('number of reservations')
8 plt.legend(['not canceled', 'canceled'])
9 plt.show()
```



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

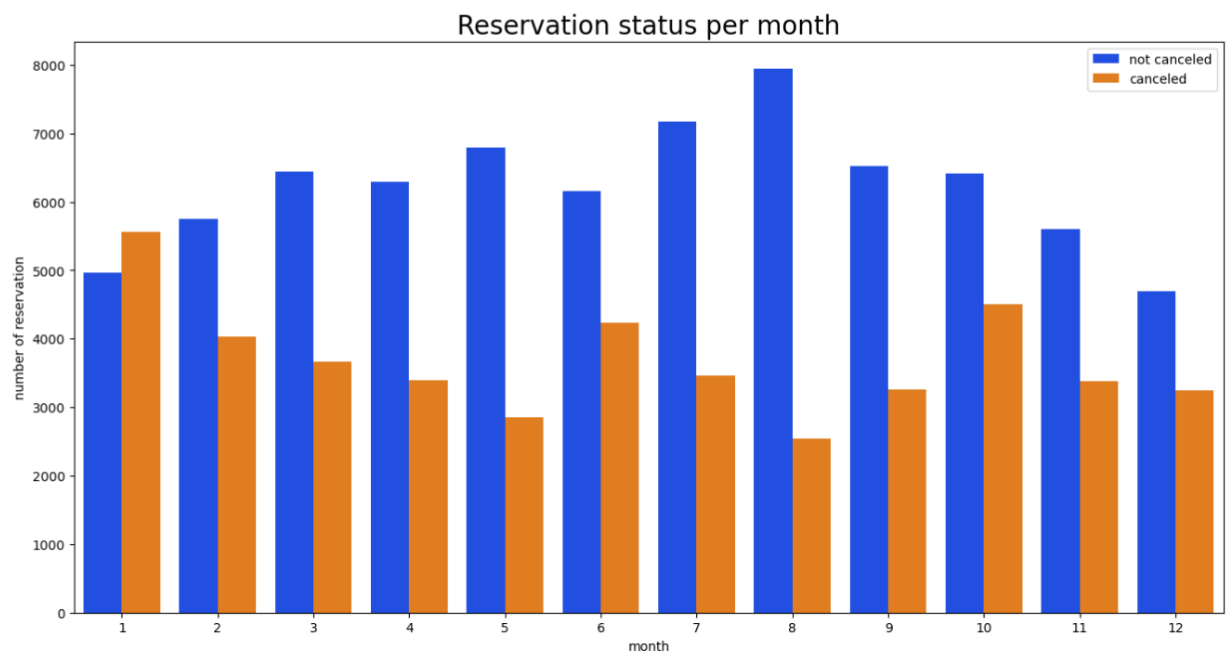
```
In [87]: 1 plt.figure(figsize = (20,8))
2 plt.title('Average Daily Rate in City and Resort Hotel', fontsize = 30)
3 plt.plot(resort_hotel.index, resort_hotel['adr'], label = 'City Hotel')
4 plt.plot(city_hotel.index, city_hotel['adr'], label = 'Resort Hotel')
5 plt.legend(fontsize = 20)
6 plt.show()
```



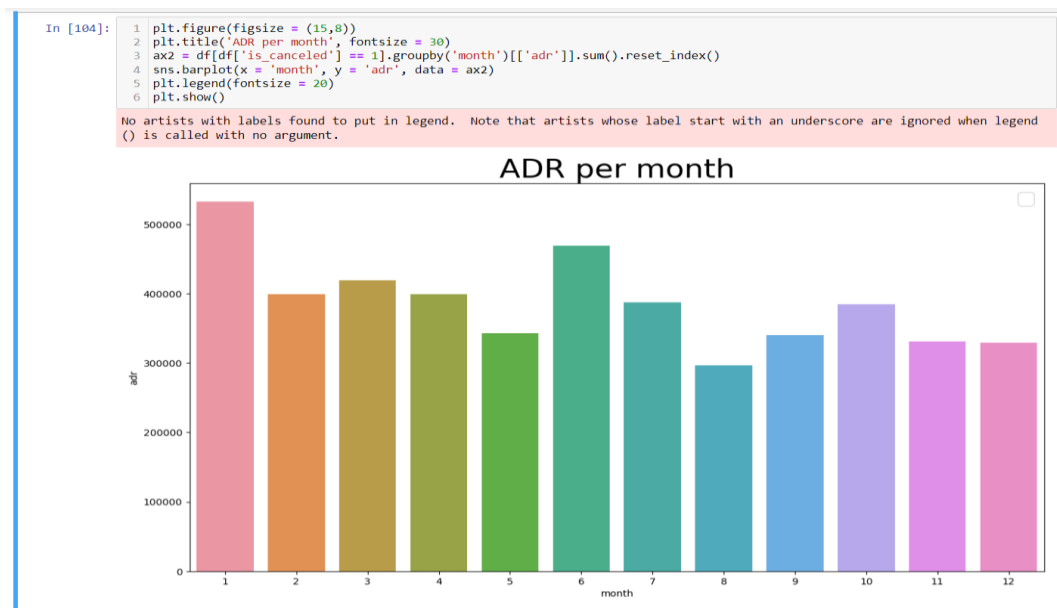
The line graph above shows that, on certain days, the average daily rate for a city hotel is less than that of a resort hotel, and on other days, it is even less. It goes without saying that weekends and holidays may see a rise in resort hotel rates.

Reservation status per month

```
In [92]: 1 df['month'] = df['reservation_status_date'].dt.month
2 plt.figure(figsize = (16,8))
3 ax1 = sns.countplot(x = 'month', hue = 'is_canceled', data = df, palette = 'bright')
4 legend_label, _ = ax1.get_legend_handles_labels()
5 ax1.legend(bbox_to_anchor = (1,1))
6 plt.title('Reservation status per month', size = 20)
7 plt.xlabel('month')
8 plt.ylabel('number of reservation')
9 plt.legend(['not canceled', 'canceled'])
10 plt.show()
```



We have developed a 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 is largest in the month of August. whereas January is the month with the most canceled reservations.

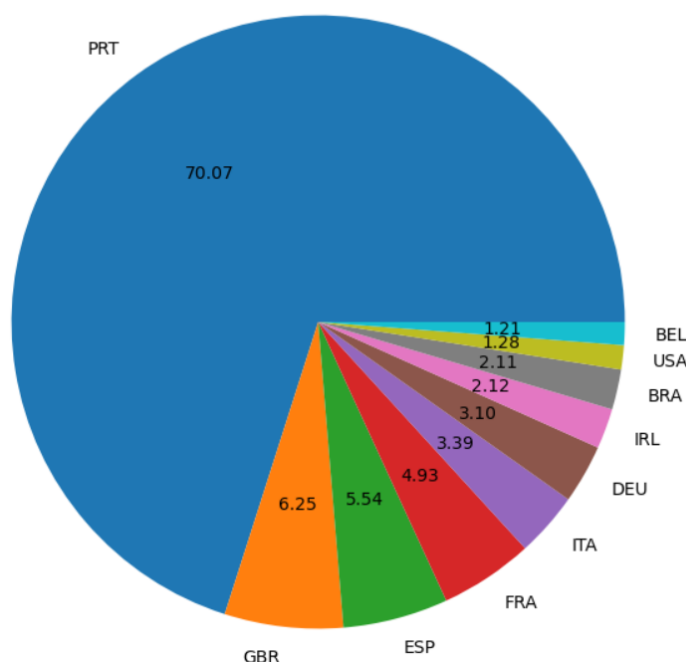


This bar chart demonstrates that cancellations are most common when prices are greatest and are least common when they are lowest. Therefore, the cost of the accommodation is solely responsible for the cancellation.

Now, let's see which country has the highest reservation canceled. The top country is

```
In [107]: 1 cancelled_data = df[df['is_canceled'] == 1]
2 top_10_country = cancelled_data['country'].value_counts()[:10]
3 plt.figure(figsize = (8,8))
4 plt.title('top 10 countries with reservation cancelled')
5 plt.pie(top_10_country, autopct = '%.2f', labels = top_10_country.index)
6 plt.show()
```

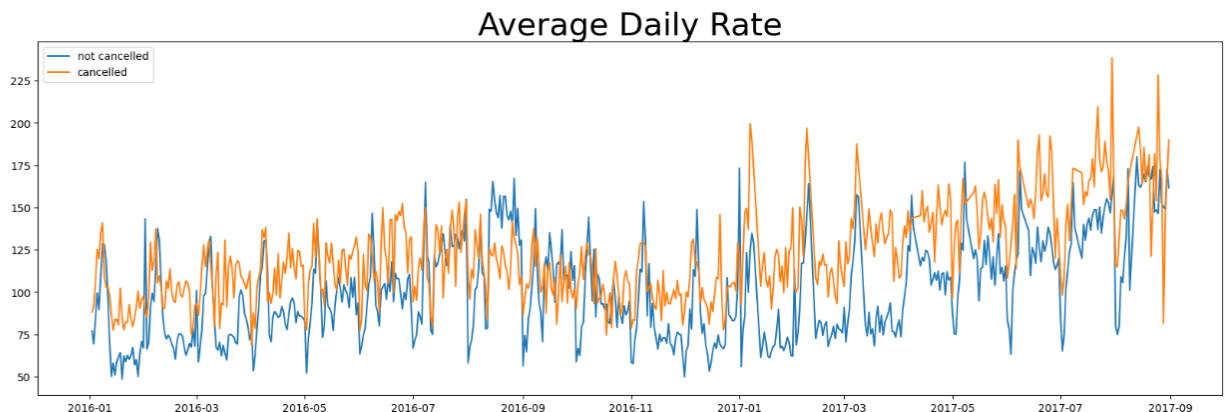
top 10 countries with reservation cancelled



Portugal with the highest number of cancellations.

Let's check the area from where guests are visiting the hotels and making reservations. Is it coming from Direct or Groups, Online or Offline Travel Agents? Around 46% of the clients come from online travel agencies, whereas 27% come from groups. Only 4% of clients book hotels directly by visiting them and making reservations.

```
In [122]: 1 plt.figure(figsize = (20,6))
2 plt.title('Average Daily Rate', size = 30)
3 plt.plot(not_cancelled_df_adr['reservation_status_date'], not_cancelled_df_adr['adr'], label = 'not cancelled')
4 plt.plot(cancelled_df_adr['reservation_status_date'], cancelled_df_adr['adr'], label = 'cancelled')
5 plt.legend()
6 plt.show()
```



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

## 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 location. They can also provide some discounts to the consumers.
2. As the ratio of cancellation and not cancellation of the resort hotel is higher in the 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 reason highest 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.