



# Prediction of hotel booking cancellations

---

**Suraj Rana Vujjini**

GJ58372

Department of Data Science

Data 603 – Platforms for Big Data Processing

Andrew Enkeboll

# Introduction

---

- Hotels thrive on advance reservation systems. Due to the current competitive market, hotels need to have an easy cancellation policy for guests.
- The average percentage of canceled reservations is 24%. Cancellations effect the revenue of the hotel will lose potential revenue customers who will not cancel.
- As a part of the project, I will be analyzing bookings data from Microtel BWI and will be creating a model to predict if the reservation will cancel or not. This can help the hotel forecast the future revenue and also help price the rooms accordingly.



# About the dataset

## No. of Rows

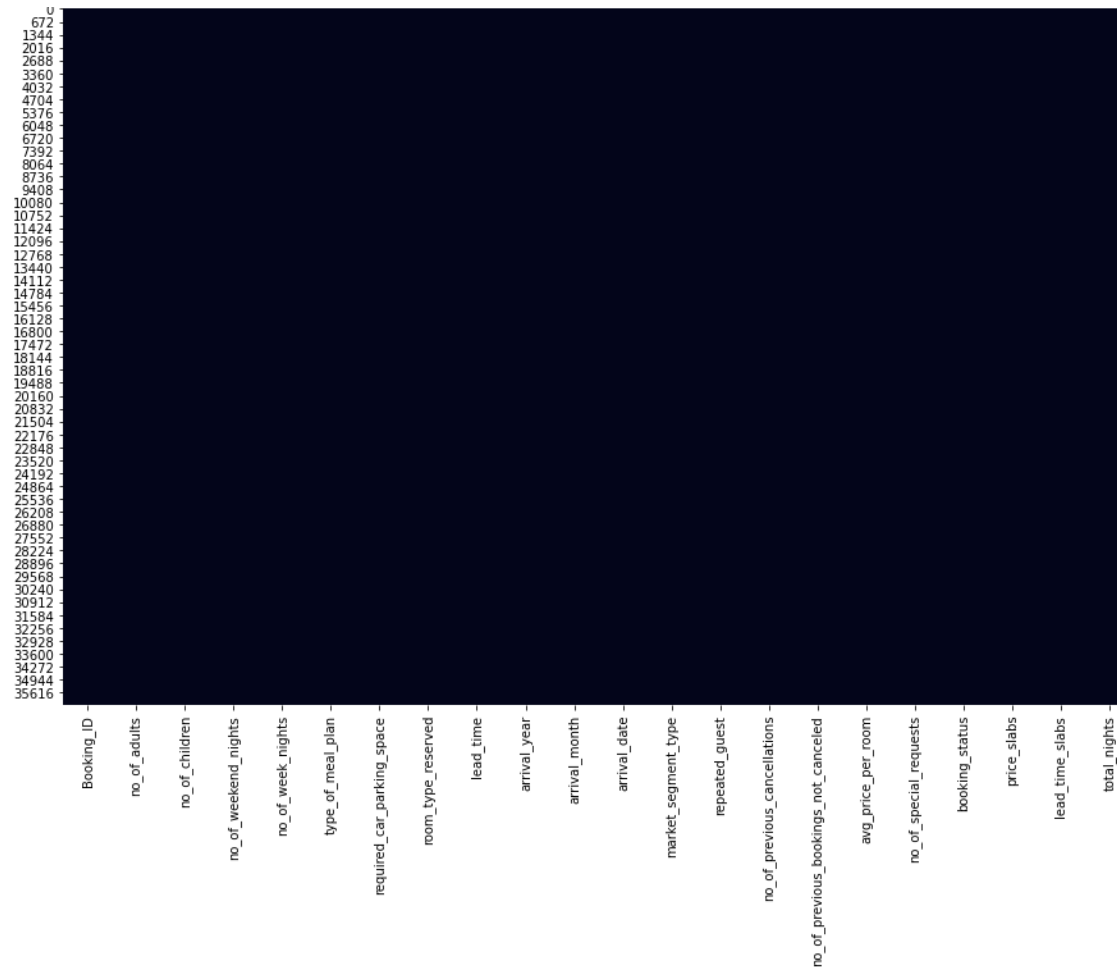
36,275

## No. of columns

19

## Total number of values

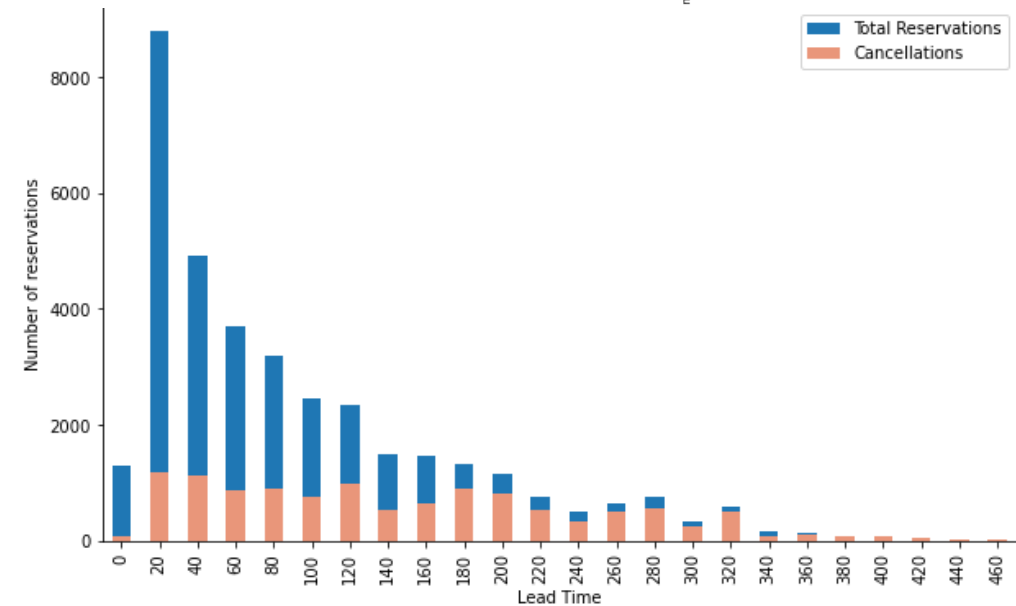
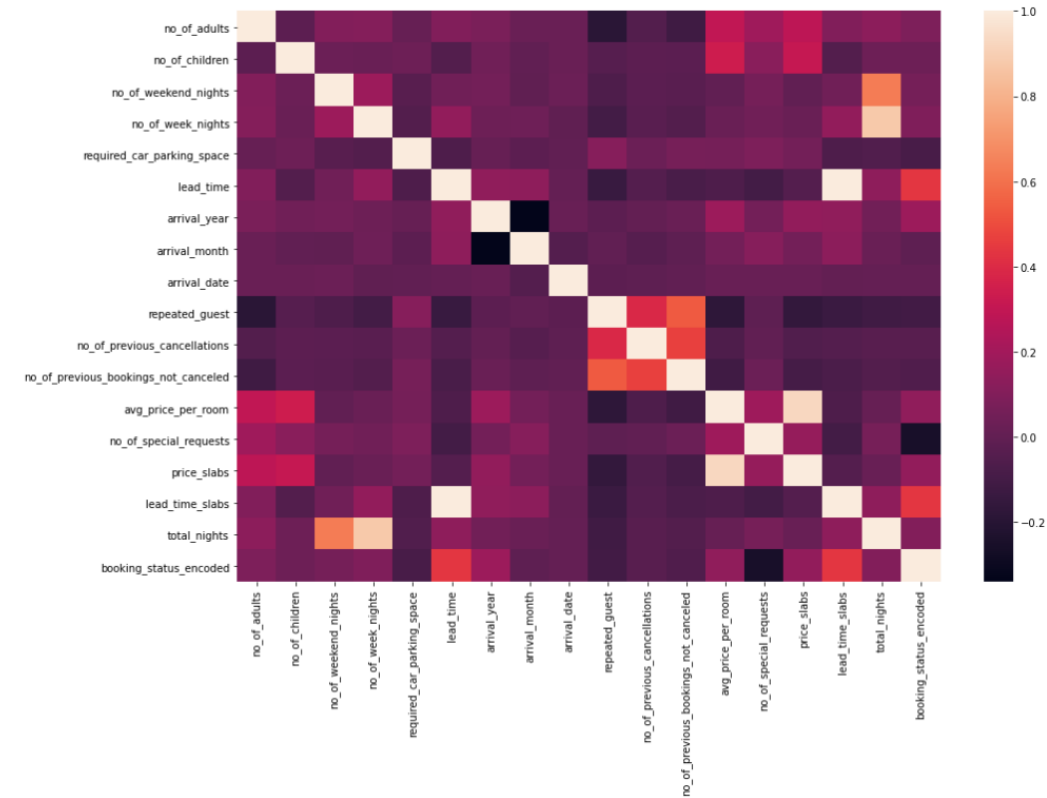
689,225



SNS heatmap of null values in the dataset

# Correlation among the columns

- I have created a new column “booking\_status\_encoded” to use in the correlation heatmap.
- Comparatively there is a good correlation of the booking status column with lead time, as the lead time increases, there is a high chance of a booking getting cancelled.



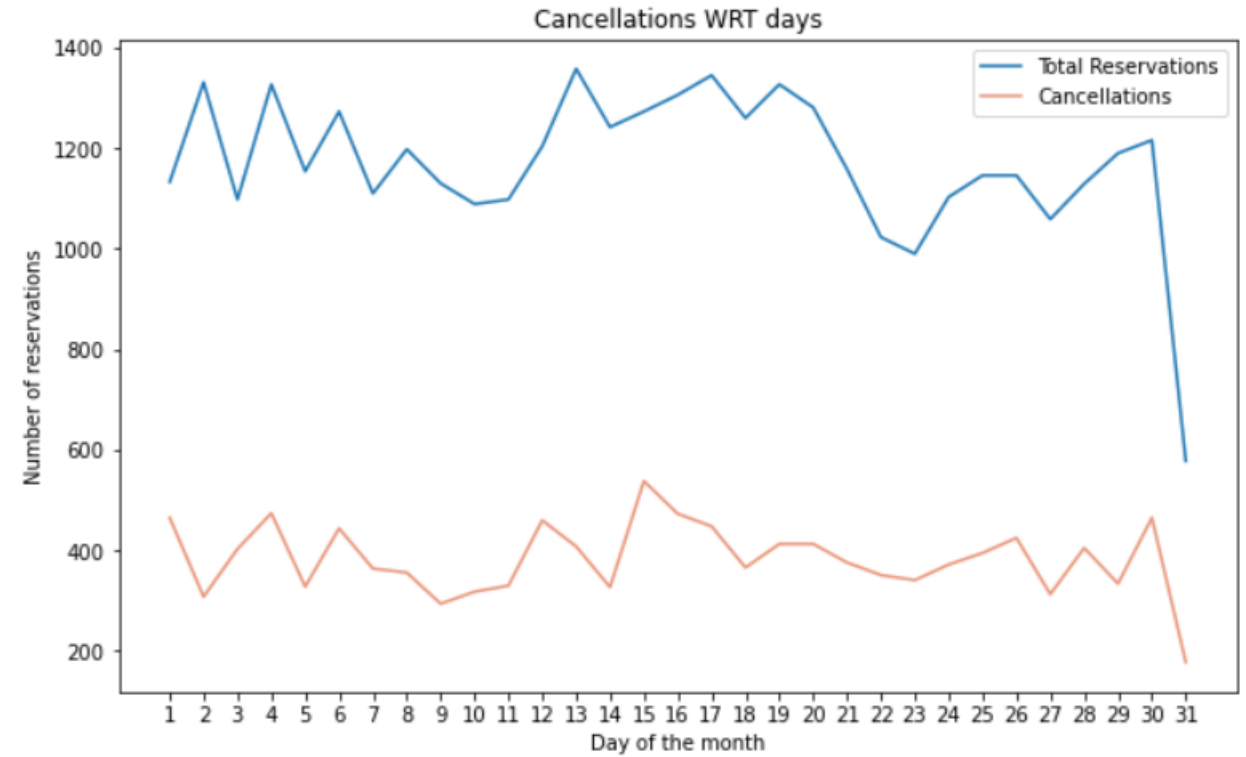
# Cancellations WRT months

- As expected, the bookings increase for summer months and fall during wintertime.
- For this hotel, there was an increase in reservations for the month of October as well.
- The cancellations tend to vary with reservations as seen from the graph.



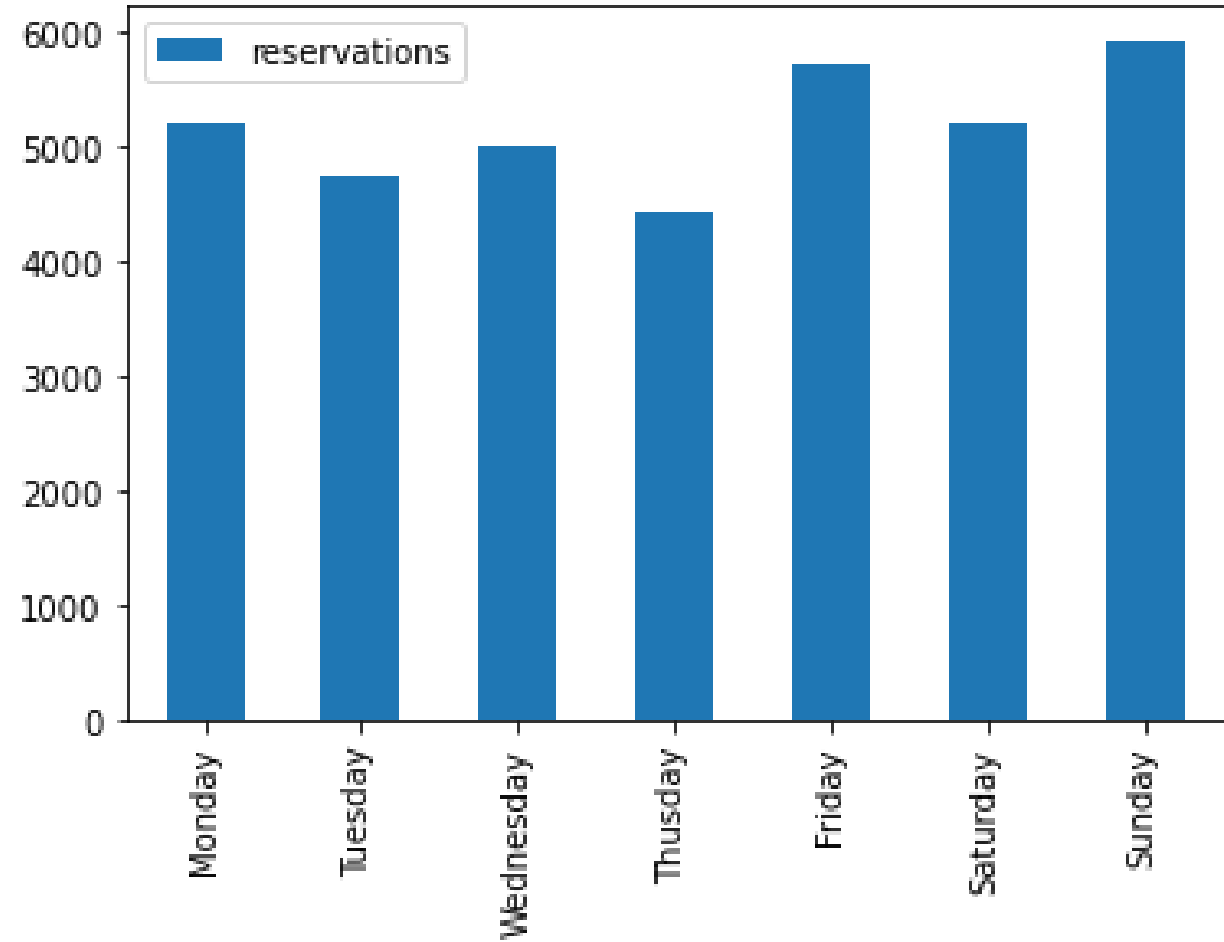
# Cancellations WRT days

- There is no useful data in this chart.
- The comparatively low number of reservations for day 31 is that only 7 months have day 31.
- The cancellations tend to vary with reservations as seen from the graph.



# Reservations WRT day of the week

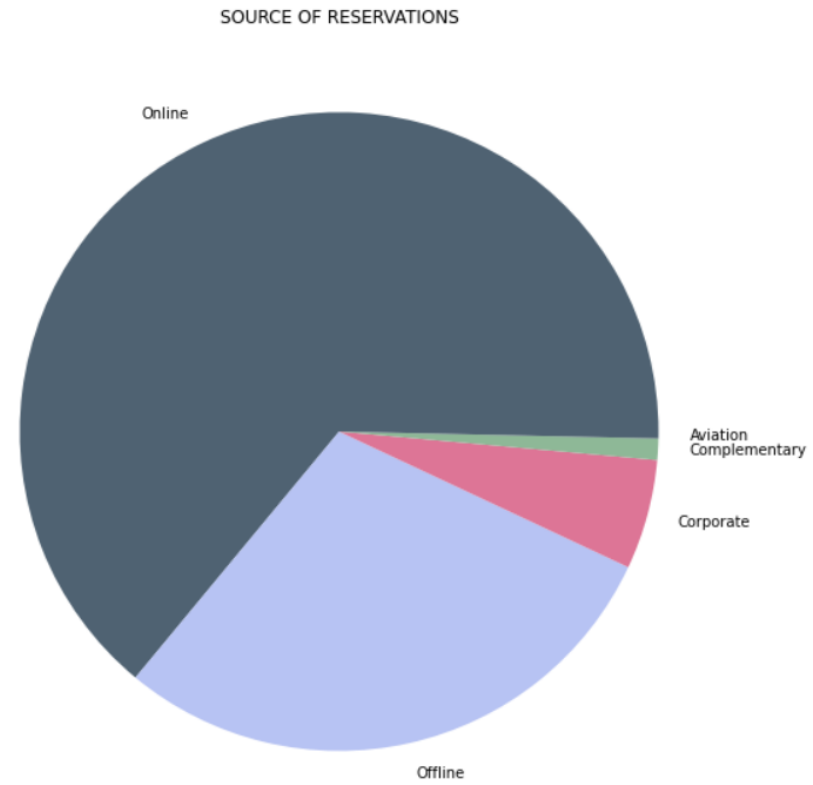
- The booking activity is highest on the weekends and lowest on the weekdays as expected.
- Friday and Sunday are the ones with highest number of reservations



# Source of reservations

---

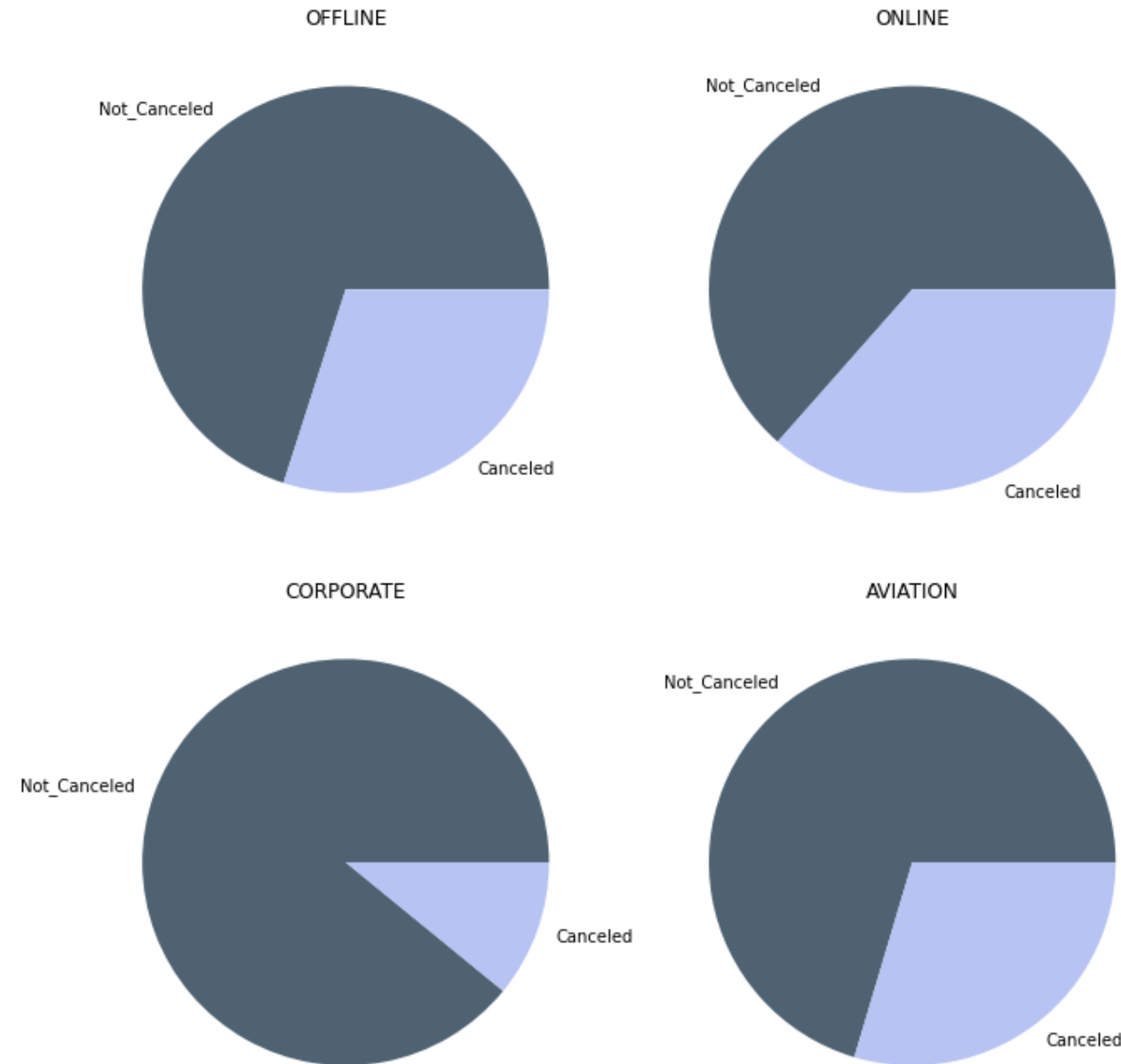
- Online is the major source for reservations at Microtel followed by reservations made offline.
- This is due to the ease of use and lucrative offers given by the booking companies.





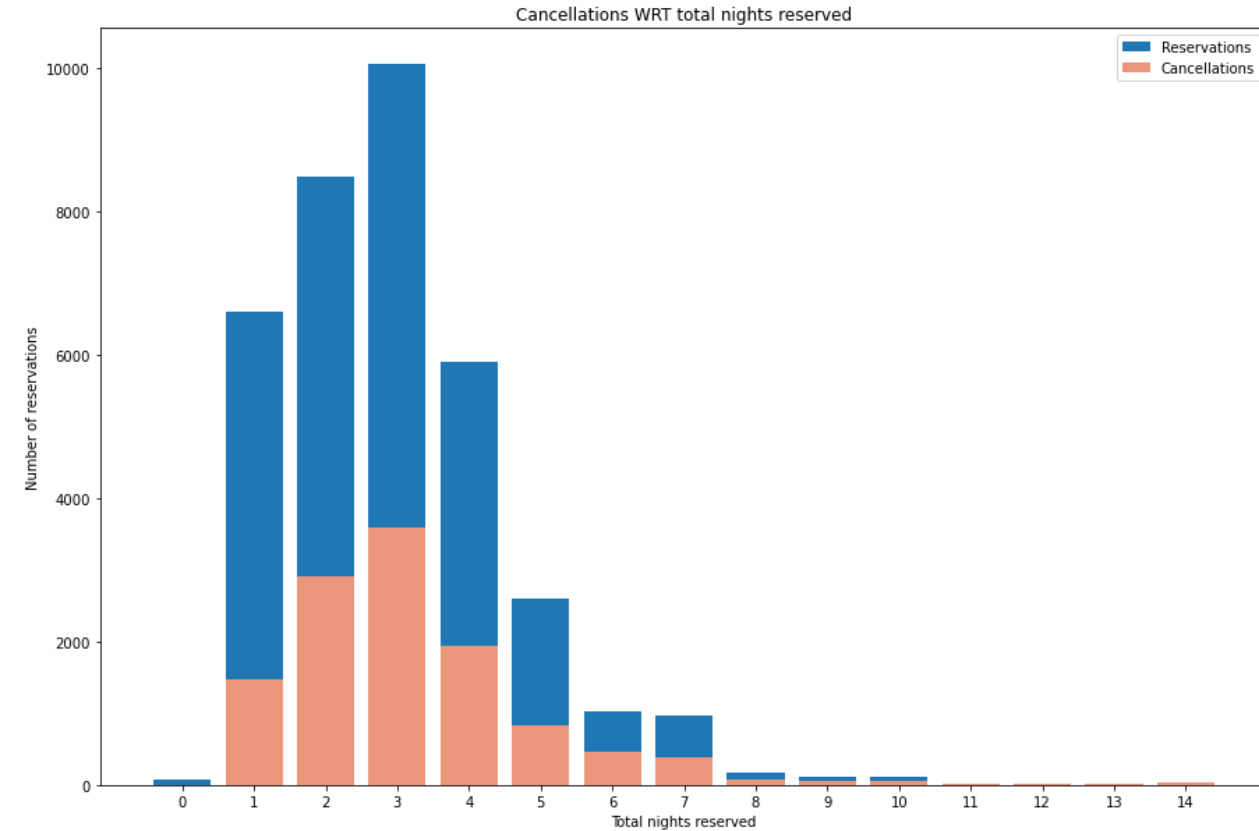
# Cancellations WRT Source of reservations

- The easy cancellation policy of online booking sites allow users to cancel at their ease.
- This can be the reason for the high cancellation ratio for online bookings.
- It is followed by offline and aviation.
- Corporate has relatively low cancellations.



# Cancellations WRT Total nights

- Most of the guest chose to stay for 3 nights.
- There is no good information for cancellations as the cancellations are higher when the reservations are high.



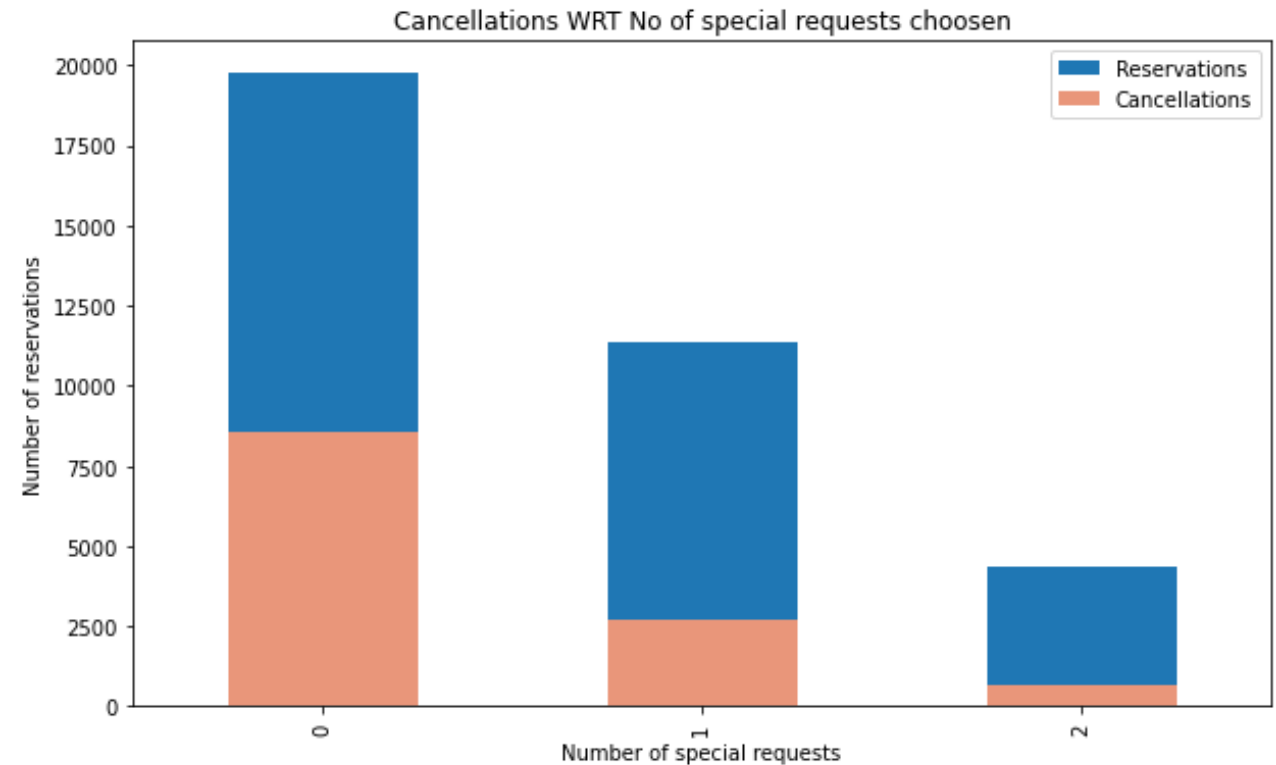
# Cancellations WRT no. of parking spots chosen

- We can see that if guests choose to reserve a parking spot, they will most likely not cancel the reservation.
- If they do not reserve a parking spot, there is a higher chance of the reservation getting cancelled.



# Cancellations WRT no. of special requests made

- Special requests are made for online order where a guest can request certain things during their stay at the hotel.
- This shows that a guest is serious about his/her stay at the hotel.
- As seen from the graph, as the number of special requests increases, there is less chance of a booking getting cancelled.



# Classification – Logistic regression

---

	precision	recall	f1-score	support
Canceled	0.75	0.70	0.72	2144
Not_Canceled	0.84	0.87	0.85	3809
accuracy			0.81	5953
macro avg	0.79	0.78	0.79	5953
weighted avg	0.81	0.81	0.81	5953

The background features several large, overlapping geometric shapes, primarily diamonds and triangles, in teal, yellow, and green colors. These shapes are arranged in a way that creates a modern, abstract design. The teal shapes are located in the top-left, top-right, and bottom-left areas. The yellow shapes are in the top-right and bottom-left areas. The green shape is a central diamond shape.

## Conclusion

To conclude, irrespective of the price, hotel reservations get cancelled due to several reasons. From these findings, it can be understood that if a guest has more requests or book a parking spot thereby validating their stay and tend to cancel less.



**Thank you**

---