**Capstone Project**

**Hotel Booking Analysis – Technical Documentation**

Raavi sahu

Data Science trainee

AlmaBetter, Bangalore

**Contents**

1. Abstract

2. Introduction

3. Problem Statement

4. Steps Involved

5. Data Visualizations

6. Conclusion

7. References

**Abstract**

The Dataset consists of information of City Hotel & Resort Hotel for the years (2015-17). It includes information such as when the booking was made, arrival date year, arrival date month, children, Market segment, and cancellations per month, among other things. This relational database consists of following rows and columns.

Rows – 119390

Columns – 32

# Introduction

Hotel industry is a very volatile industry and the bookings depend on a variety of factors.

Analyticsin the hotelier world today is important, and nowadays this business cannot be run with some sensible and smart use of data.

# Problem Statement

In this analysis, we will see what are the factors that are governing the bookings and what are the steps that can be taken in order to improve the bookings.

**Steps involved**

The problem statement can be analysed by using these steps:

1. Data Wrangling
2. Data Analysis
3. Data Visualization
4. Gathering insights

**Procedure for the Data Analysis**

We have to follow various procedures such as Importing libraries and Database, Summary of Data:

Data wrangling, Data Analysis, Data visualization, Conclusion

1. **Importing Packages**

Import inbuilt libraries: NumPy, Pandas to prepare the Data and Seaborn, Matplotlib for data visualizations.

1. **Data Cleaning**

This dataset consists of null values and different data types as well. So, in order to work on our data, we need to remove the null values and in order to ger the desired result we need to change some of the data types as well.

1. **Detection and Deletion of the null values**

We have replaced null values (Nan) with zero. So, we don’t miss any of the data in the columns. We have also changed the float data types into int data type in order to get the desired result.

1. **Data Analysis**

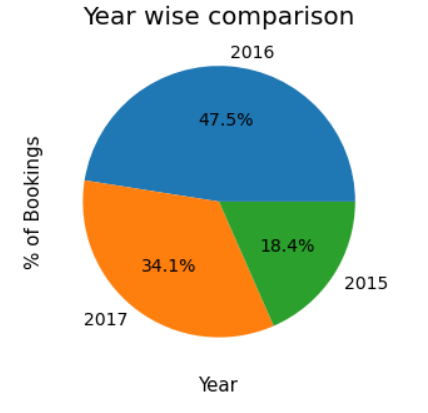
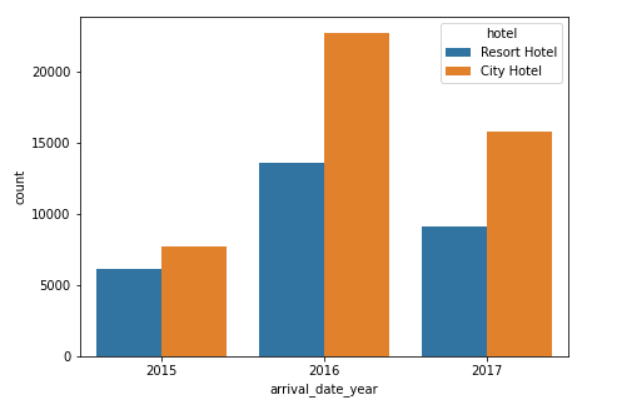
Here, we analysed our data to find factors which govern hotel bookings. So, we followed this analysis.

1. Year wise Comparison of Hotel Bookings
2. Month wise comparison of bookings
3. Total bookings cancelled
4. Booking percentage between 2 hotels
5. Repeated guests percentage between 2 hotels
6. Booking types for Market segment
7. Number of night stays in hotels
8. Total number of special requests
9. **Data Visualizations**

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

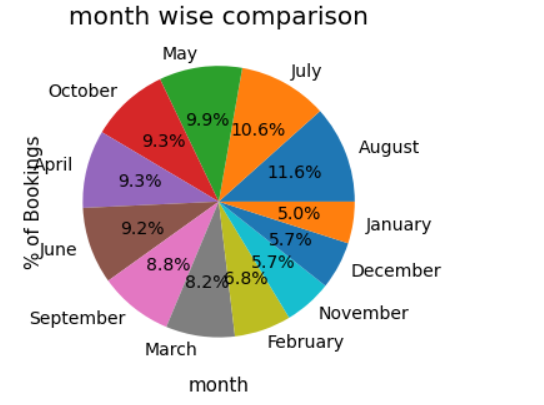
Now, we take a look at some insights of the dataset using data visualisation.

**Year wise Comparison of Hotel Bookings**

****

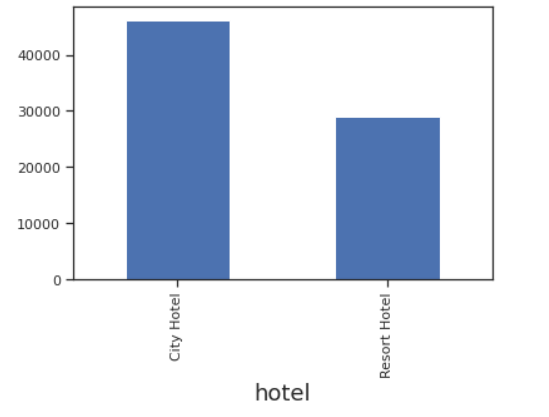
2016 is the most productive year with 47.5%.

**Month wise comparison of bookings**



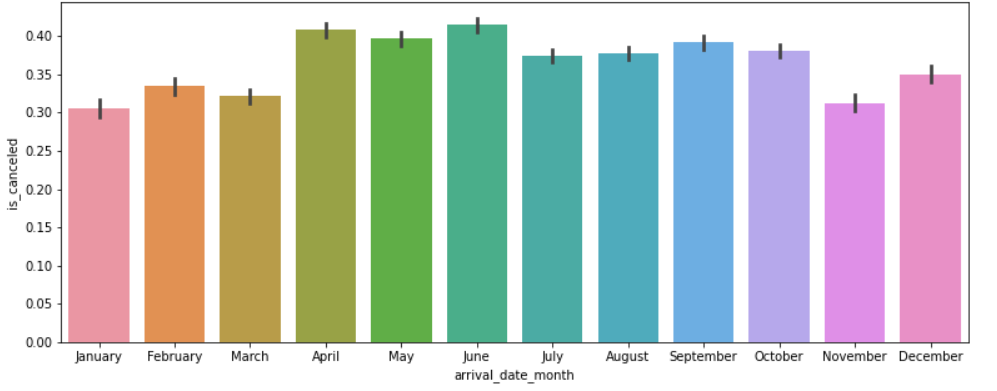
August has 11.6% number of Bookings, which is highest.

**Total bookings cancelled**



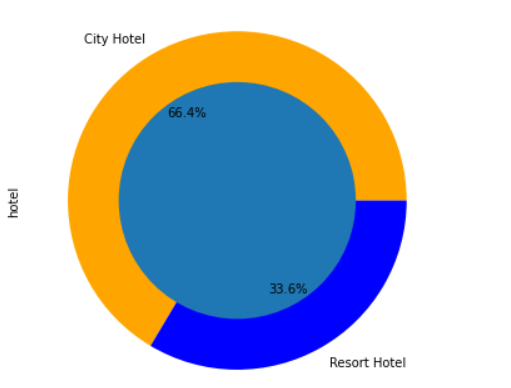
City Hotels has higher number of Cancellations.

**Monthly Cancellations**



Number of Cancellations are highest in June.

**Booking percentage between 2 hotels**

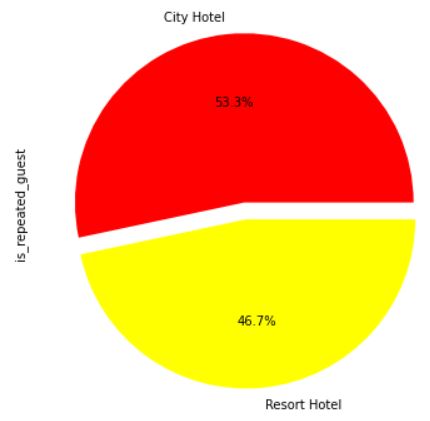


City Hotel has more booking percentage than Resort Hotel.

City Hotel – 66.4%

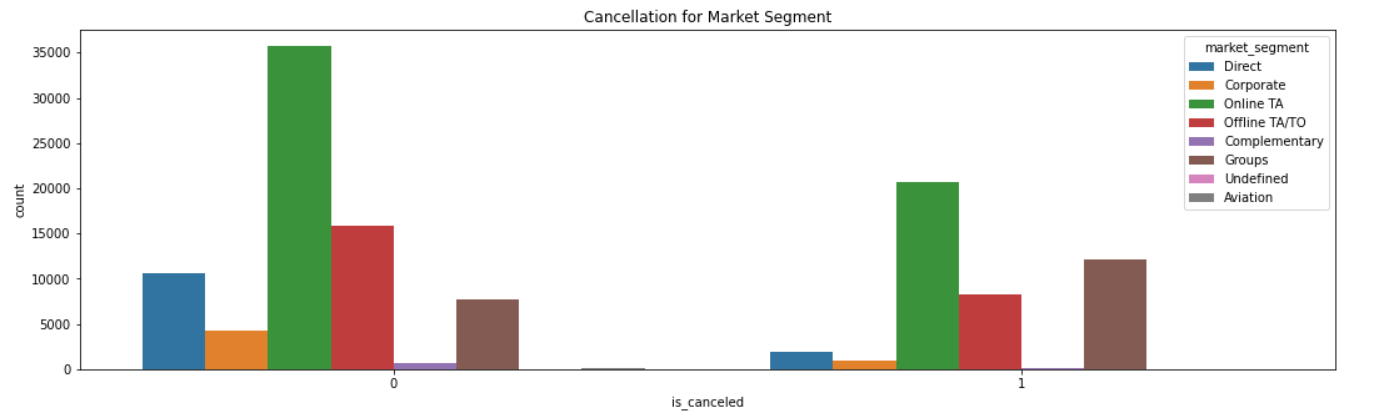
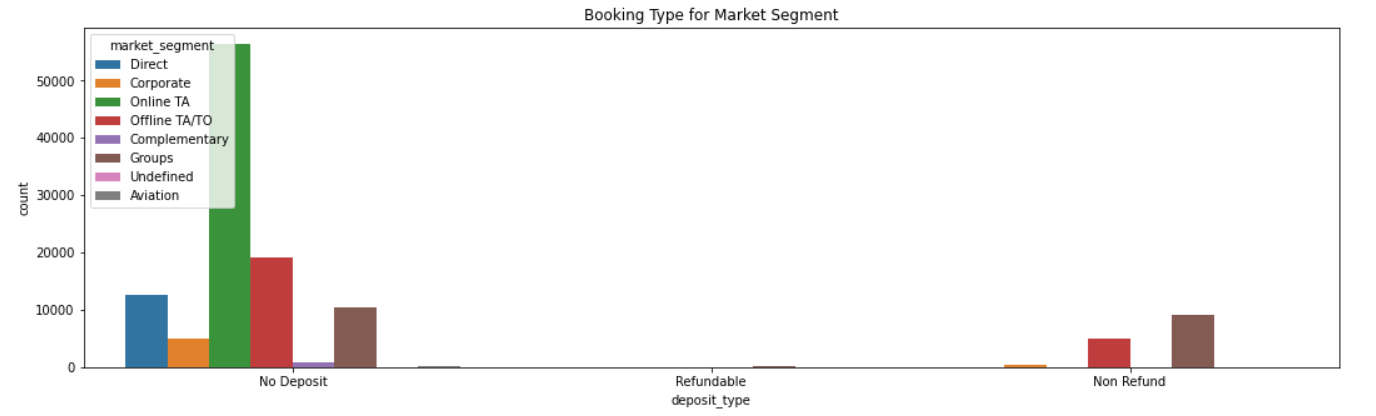
Resort Hotel – 33.6%

**Repeated guests percentage between 2 hotels**



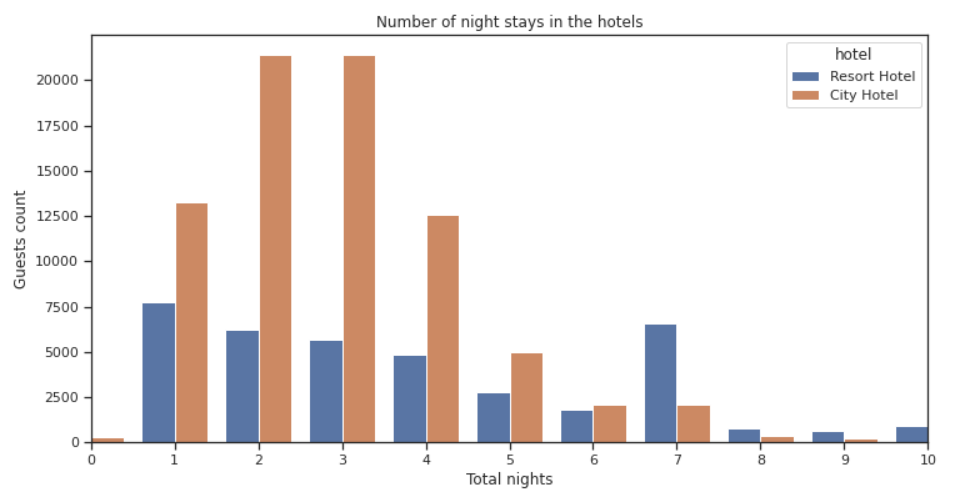
City Hotel has more repeated guests with 53.3%.

**Booking types for Market segment**



Online Bookings are the most Preferred among other Booking types.

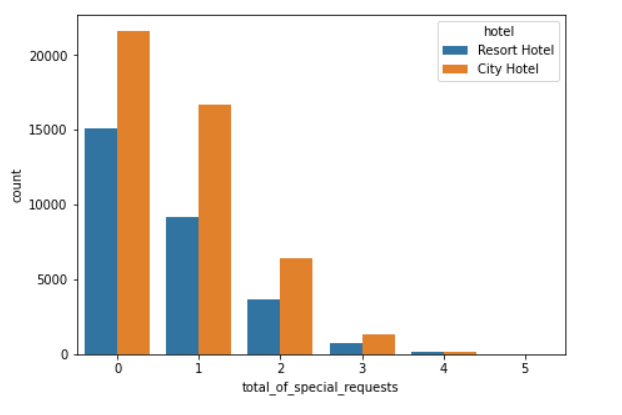
**Number of night stays in hotels**



For short stay (3-5 Days), People prefer City Hotel.

For long stay (more than 6 Days), People prefer Resort Hotel.

**Total number of special requests**



City Hotel has more special requests than Resort Hotel

**Conclusions**

* 2016 has a greater number of Bookings with 47.5% and it went down to 34.1% in 2017. We can see booking percentage decreased. So, we need to come up with better booking plans.
* Q2 & Q3 are having higher number of Bookings. Hotels can be better prepared for high volume in Q2 and Q3 and in order to increase the bookings in Q1 and Q4, can offer better packages.
* City Hotel has higher number of Bookings and Cancellations as well. Approximately, 40% of the total bookings are getting cancelled in city hotels. Hence, the percent of confirmed bookings is more in resorts.
* Repeated guests percentage is more for City Hotel. So, City hotel can avail some special discounts to the repeated guests and this can be followed by Resort hotel in order to increase their bookings as well.
* Online Booking plays crucial role in Hotel Booking. The ‘online TA’ is most used market segment and ‘TA/TO’ is maximum in all distribution channels. So, hotels can run online campaign for better reach to the customers and they can also promote their hotels in online portals to attract new customers.
* For Short stay people choose City Hotel and for long stay people choose Resort Hotel. So, City hotel can come up with plans to attract people for long stays and Resort hotels can come up with plans to attract people for short stays.

**References**

1. GeeksforGeeks
2. Analytics Vidhya
3. Stack overflow