**Hotel Booking Analysis –Technical Documentation**

**Roshana Patil**

**Table of Content:-**

1. Abstract

2. Introduction

3. Problem Statement

4. Steps Involved

5. Data Visualization

6. Conclusion

**Abstract:**

In our project on “Hotel Booking Analysis”

, we have tried to show

bookings on types of hotels. Exploring a hotel booking dataset in hopes

to find some interesting trends and insights in this industry that can

help us understand the business better and help us improve.

Our dataset describes two types of hotels, one is Resort hotel and the

other is a City hotel. The dataset contains booking data for both hotels

and has some 1, 19,390rows and 32columns.

Each observation describes all the relevant information about that

particular instance of booking.

We perform Exploratory Data Analysis with python(using some

libraries) to get insights from the data, to help the hotel industry take key

decisions which will help them to improve their efficiency as well as

their revenue and possibly help reduce their operational costs.

**1.Problem Statement**

The Objective of this exercise is to perform Exploratory Data Analysis

on the given dataset which contains booking information for two hotels.

We will be using the data available to analyze the factors that govern the

hotel bookings.

We will use the following steps in our pursuit improvement and solution

of the mentioned problem statements.

Step1: Data Cleaning

Step2: Dealing with outliers

Step 3: Visualizing the numerical and categorical columns

Step 4: Performing univariate and multivariate analysis.

Step5: Gathering useful insights

**2. Introduction**

## Hotel industry is a very volatile industry and the bookings depend on

## variety of factors such as type of hotels, seasonality, days of week and

## many more. This makes analyzing the patterns available in the past data

## more important to help the hotels plan well. Using the historical data,

## hotels can perform various campaigns to boost the business.

* **Procedure for Data Analysis**

We have to follow various procedures such as – Importing packages,

data cleaning, and giving useful insights.

1. Importing Packages

We are importing Numpy and Pandas library to prepare the data

and importing Seaborn and Matplotlib for graphical representation

of data or Data Visualization.

**2. Data Cleaning**

Data Cleaning involves basic operations performed on the dataset

such as

1)Remove duplicate rows**,**

2) Handling missing values.

3) Convert columns to appropriate datatypes.

4) Adding important columns

After following these basic steps, we get clean dataset which we

are using for further analysis.

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

We checked the null values in all the columns and then we found

that agent and company children country,column have maximum null values. So,

we delete them as it won’t affect our dataset**.**

**4. Detecting Outliers**

We detect outliers using Box plot and replace them with some

specific nearby values**.**

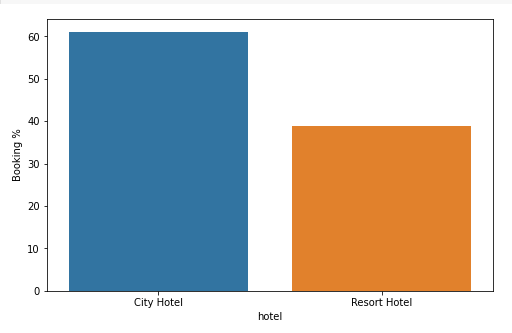
1. Data Visualization

Now, we move forward to the most important content i.e. DataVisualization, it is nothing but the graphical or pictorial representation of our dataset which result in the conclusion

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

visualization

**Percentage of bookings for each hotels**

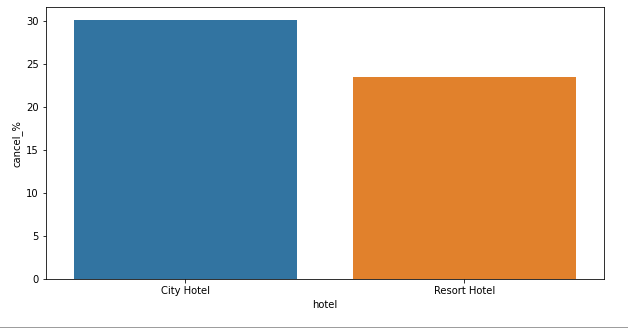


The above graph states that most customers prefer to book City

hotel rather than the Resort hotel.

Around 60% bookings are for City hotel and 40% bookings are for Resort hotel.

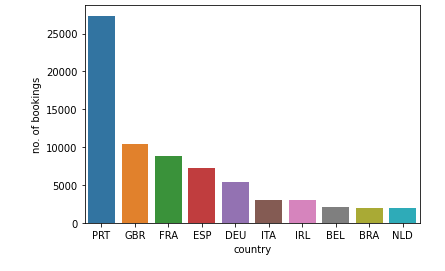
**Highest cancellation of the hotel**



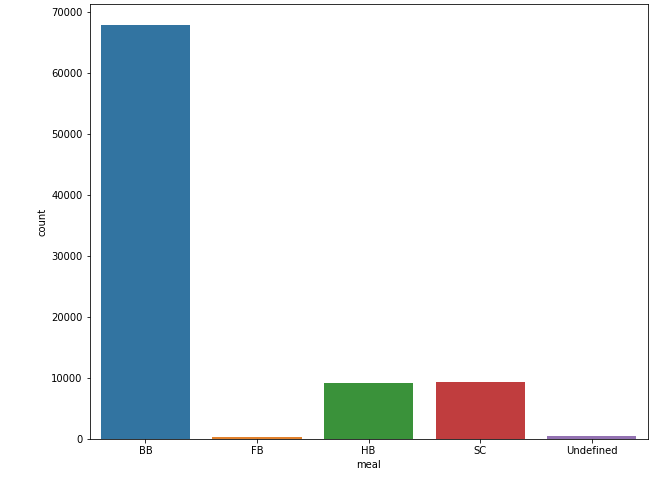
From above bar plot, we see that City hotel has more number of

cancellations than Resort hotel.

**Top 10 countries with highest bookings**



***Preferred Meal by the customer***



Form the above graph*s,we can say that the heigest number of booking are from PRT(Portugul),GBR(Great Britain),* *and FRA(France).*

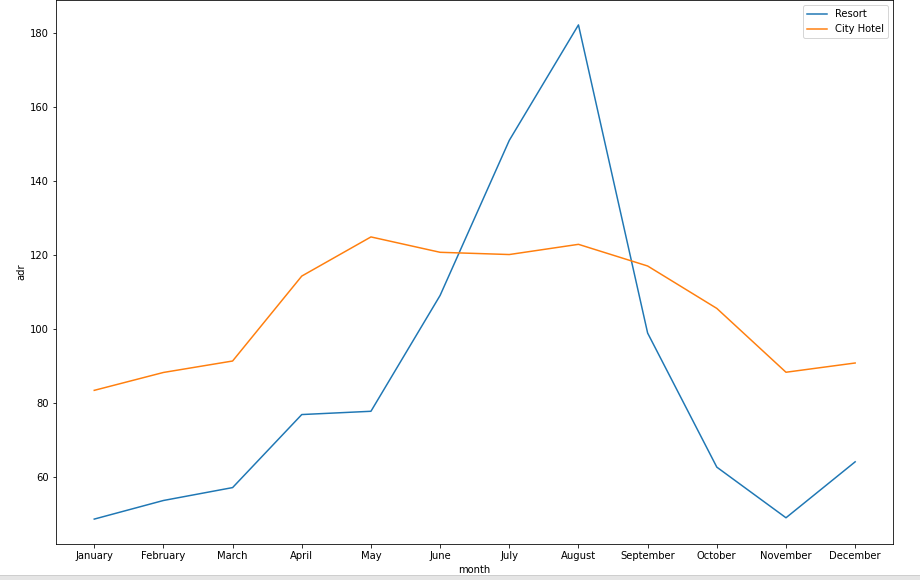
*From above graph we see that most preferred meals are BB (Bed*

*and Breakfast) followed by SC (self-catering), HB (Half Board).*

*FB (Full Board) is less preferred meal type in comparison to other*

*categories.*

**Variation of price per night according to year**

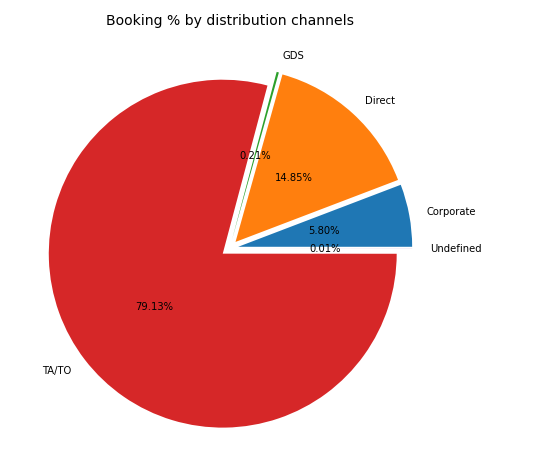


The above line graph concludes that, the ADR (average daily rate)

is higher during July and August for Resort hotels and for City

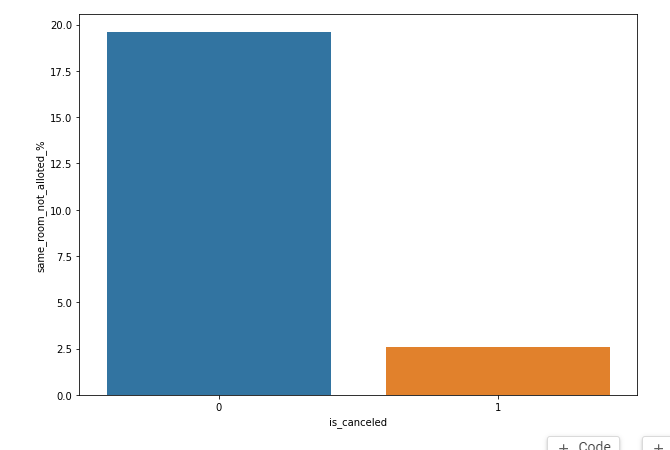
hotel the ADR is higher for the months of May, July and August.

**Booking Percentage By Distribution Channels**



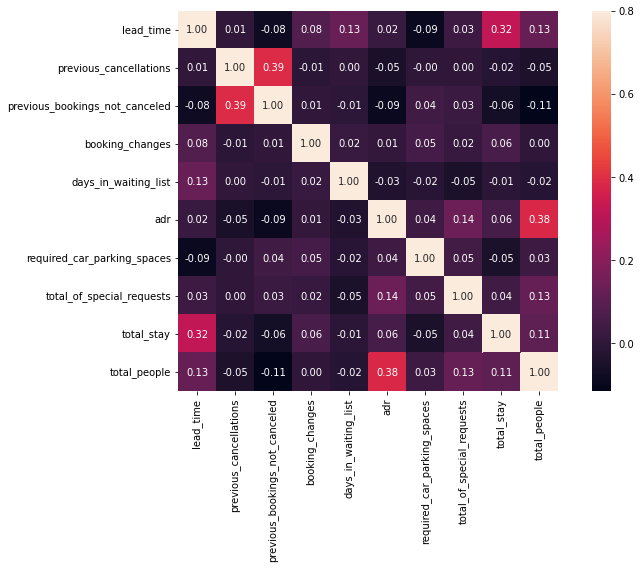
From above the Pie Chart we can see that most of the booking are TA/TO (Travelling agent, tour operators)

* **Room Not Allotted Percentage Vs Cancellation**



We see that not getting same room as demanded is not the case of cancellation of rooms. A significant percentage of bookings are not cancelled even after getting different room as demanded.

* **Correlation Heatmap**



 Total stay length and lead time have slight correlation. This may means that for longer hotel stays people generally plan little before the the actual arrival.

ADR is slightly correlated with total\_people, which makes sense as more no. of people means more revenue, therefore more adr.

**5. Conclusion:**

1. From the month-wise booking data, we see that some months

like July and August have much higher bookings than other.

1. BB (Bed and Breakfast) meal type is the most affected which is cancelled by guest.
2. After analyzing the ADR (Price per night) for both hotels, we have drawn the conclusion that ADR is high during May-

August. These months have highest number of bookings and it could be the reason for high ADR during this period.

1. Online Travelling Agent which is having highest number of bookings.
2. Around 25% bookings were cancelled at the Resort and

30% bookings were cancelled at the City hotel.

1. Same Room Not Allotted The Customer not much affected the city hotel. A significant percentage of bookings are not cancelled even after getting different room as demanded.

**Some prediction from above data visualization so as to help the**

**hotel industry to take key decisions and improve their efficiency,**

**performance and profits:**

1.Hotel can reduce ADR to attract more guests and increase the

number of bookings during this period.

2. Majority of the guests are from Western Europe. So target this area

for more customers.

3.Hotel may use some strategies to increase his bookings every

year. They can offer discounts and figure out when things are

busy, steady and slow and then adjust their price accordingly.

**References-**

1. GeeksforGeeks
2. Kaggle