

# Capstone Project

## Hotel Booking Analysis

**Team Name: DARK DATA**

**Team Member: AKASH VERMA  
AYUSH SHARMA**

# **POINT TO DISCUSS**

- Agenda
- Tools Required to Analysis data set
- Hotel Booking data set Analysis
- Data Preparation process
- Exploratory Data Analysis
- Conclusion
- We Learned that!



Let's Begin

# Agenda

We discuss the analysis of given Hotel Booking Dataset From 2015- 2017.

After analysis of “Hotel Booking Demand dataset”, there are several burning questions which needs to be addressed that will form the basis of our further analysis.

For Example:

- Which is the most in-demand hotel?
- What is the busiest time of the month and year for the two hotels?
- etc....

This dataset is interesting because after finding the answers to our above problem statements, hotels will be able to manage their customer bookings more efficiently and effectively.

# Tools Required to Analysis Dataset

Python: is a programming language used by many data scientists to clean data, make visualizations and build models

## Python Library:



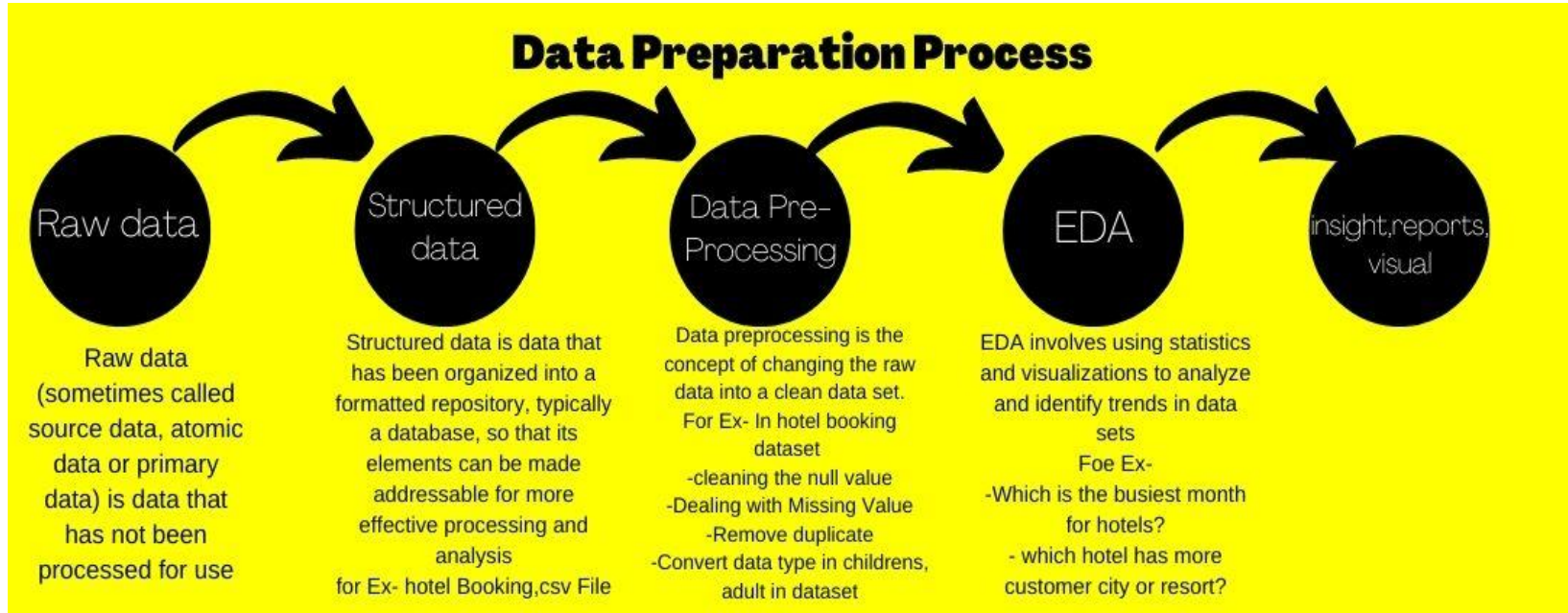
- Pandas: is an open source Python package that is most widely used for dataScience/data analysis and machine learning tasks.
- Numpy: is a Python library used for working with arrays.
- Matplotlib: is a plotting library for the Python programming language and its numerical mathematics extension NumPy.
- Seaborn: is a library for making statistical graphics in Python

# Hotel Booking Dataset Analysis

- The data set contains 119390 Rows and 32 Column.
- The data set Contain booking information for a city hotel and a resort hotel and includes information such as when the booking was made, length of stay, the number of adults, children, and/or babies, and the number of available parking spaces, among other things.
- All personally identifying information has from the data.

We will perform exploratory data analysis with python to get insight from the data.

# Data Preparation Process

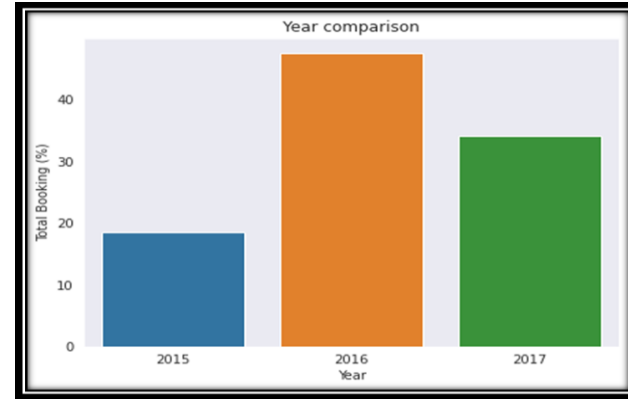


# Exploratory Data Analysis

Now let's do the fun part, extract the information from our data and try to answer our questions.



## Q1. Hotel wise yearly bookings

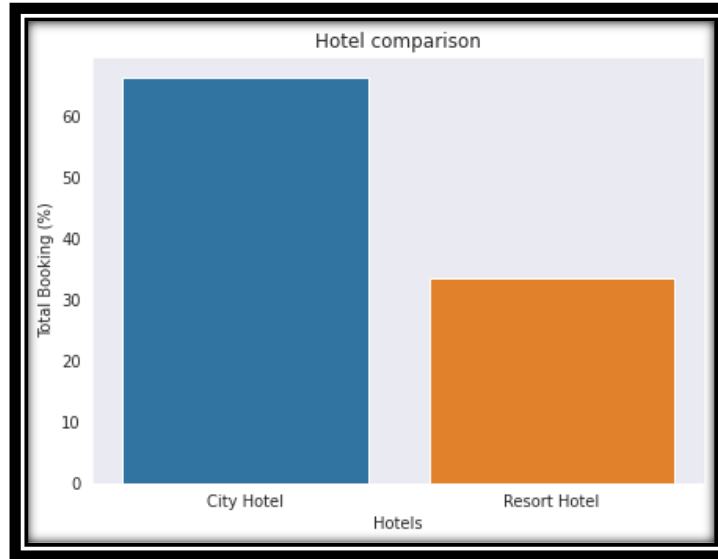


We see that bookings across years is higher for city hotel compared to resort hotel and donot increase proportionately over the years.

We See that 2016 Year is Most Booking year Comparision to the another year 2017 and 2015

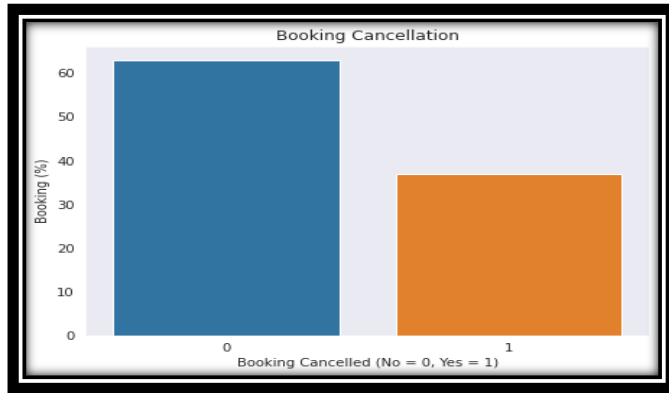


## Q2. What is the booking ratio between Resort Hotel and City Hotel?



More than 60% of the population booked the City hotel

### Q3. How Many Booking Were Cancelled?

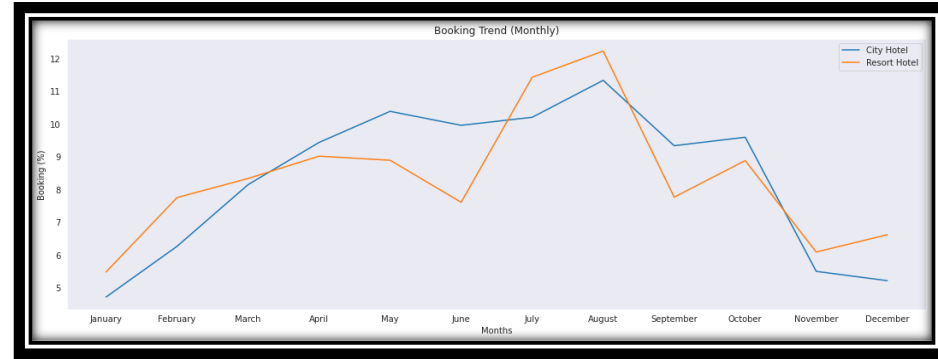
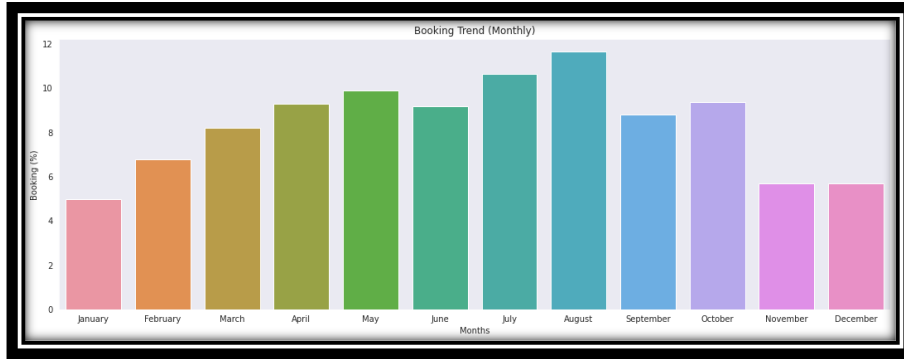


Dtaset Value(is\_canceled)have two unique values: 1 if booking got canceled, else 0.

Bookings got canceled 37% of the time. While booking guest did check-in (did not cancel the booking ) almost 63% of the time.

Around 40% bookings were cancelled in city hotel and 30% in resort hotel..

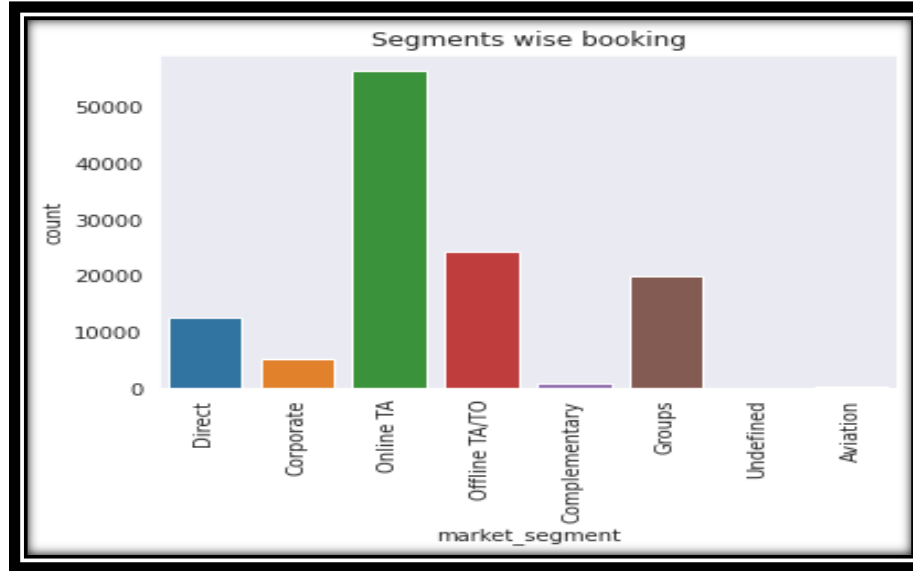
## Q4. Which is the busiest month for hotels?



As we can see most bookings were made from July to August. And the least bookings were made at the start and end of the year.

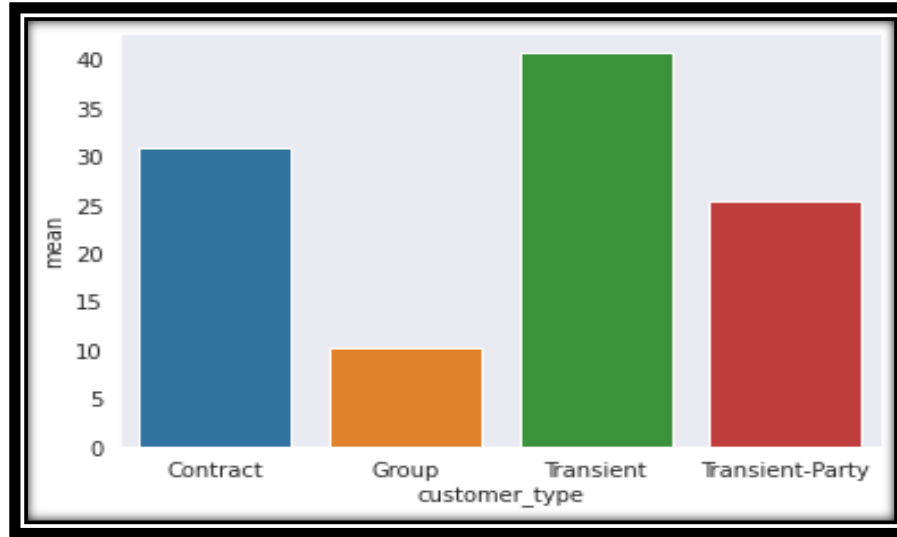
We can see the trend is kind of similar with a small difference. Resort has more bookings at the start and end of the year, and lower booking in June and September..

## Q5. Market segment wise bookings?



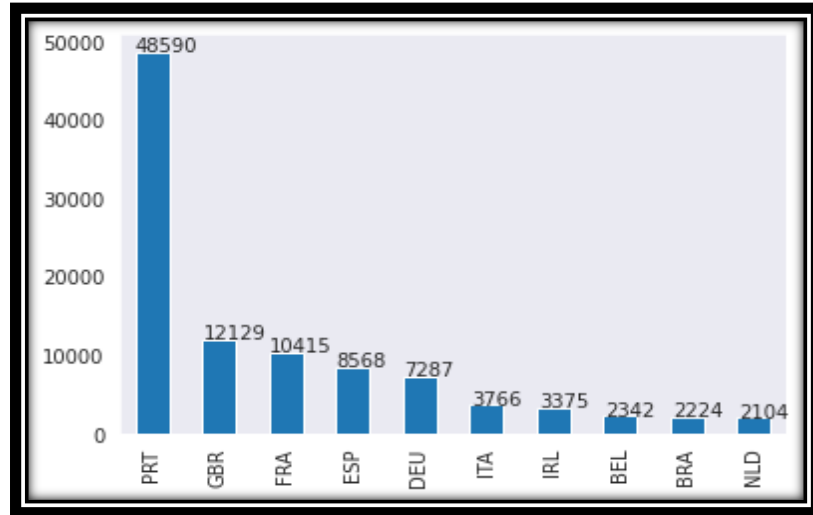
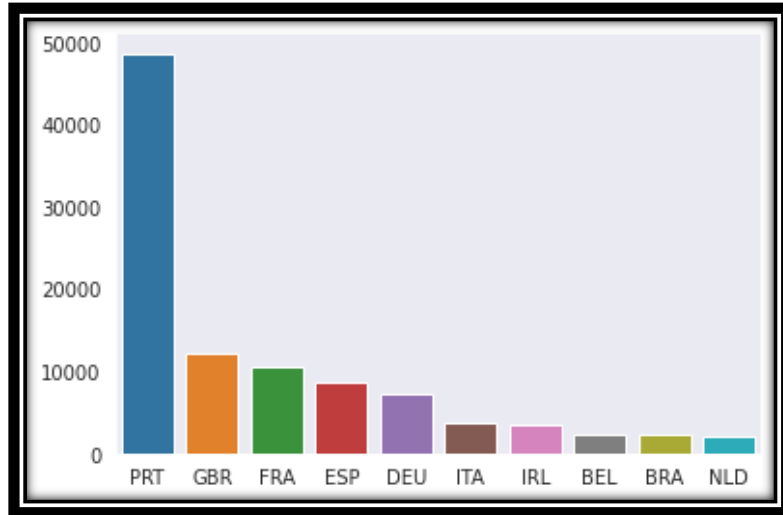
Indirect bookings through online and offline travel agents are higher compared to direct bookings and same is the case with group bookings which are also high.

## Q6. Lets Check monthly cancellations and customer type



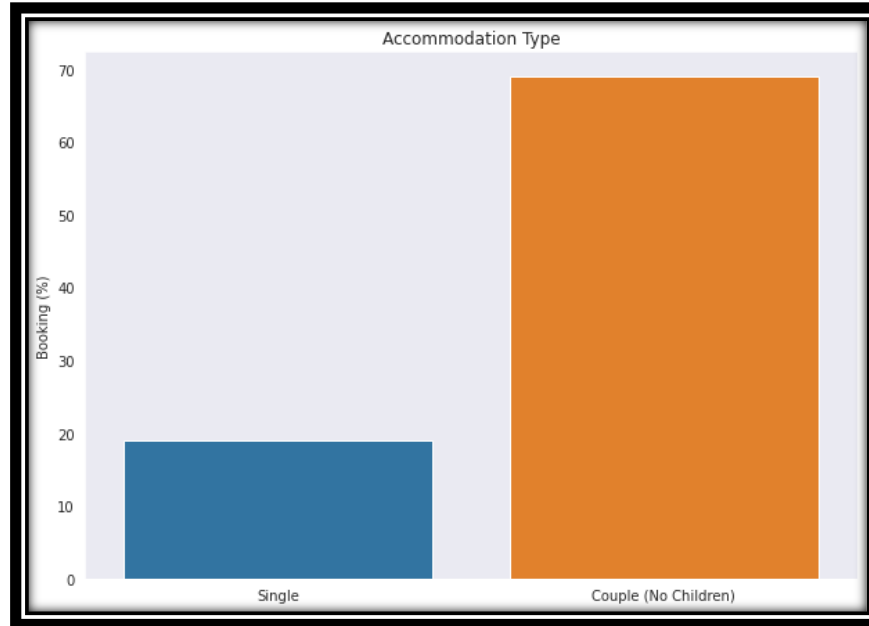
Transient customer types have higher cancellations

## Q7. From which country most guests come?



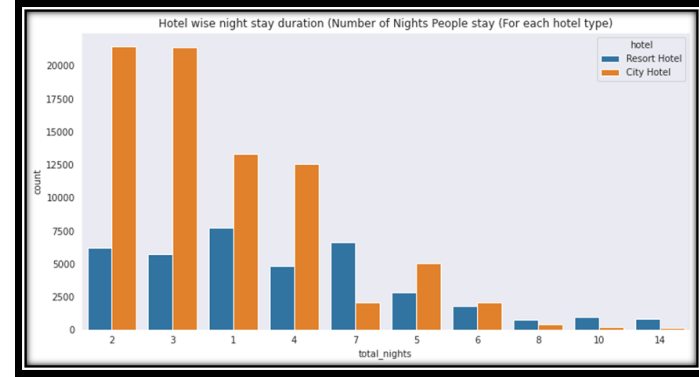
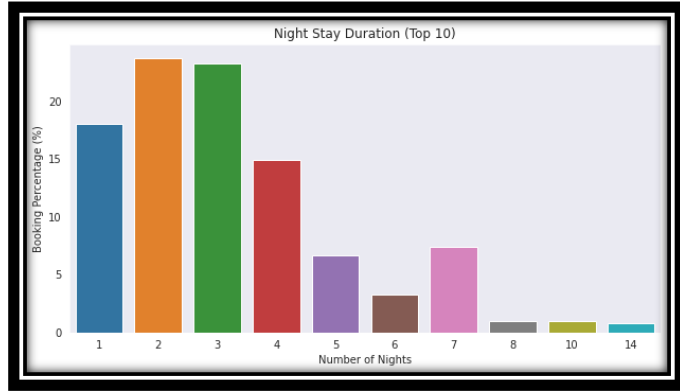
Portugal, UK and France, Spain and Germany are the top countries from most guests come, more than 80% come from these 5 countries.

## Q8. Which was the most booked accommodation type (Single, Couple)?



Couple (or 2 adults) is the most popular accommodation type. So hotels can make plans accordingly

## Q9. How Long People Stay in the hotel?

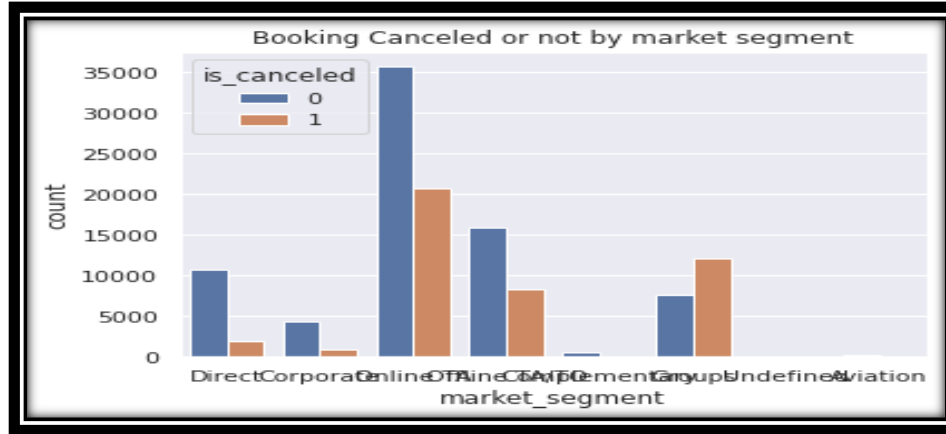


Most people stay for one, two, or three. More than 60% of guests come under these three options.

For Resort hotel, the most popular stay duration is three, two, one, and four days respectively. For City hotel, most popular stay duration is one, two, seven(week), and three respectively



## Q.10 Lets Check relationship between market segment and cancellation?



Groups segment has cancellation rate around 50%. Offline TA/TO (Travel Agents/Tour Operators) and Online TA has cancellation rate more than 33%. It is surprising that the cancellation rate in these segments is high despite the application of a deposit. The fact that cancellations are made collectively like group reservations has high cancellation rate. It is interesting to note that direct segment has cancellation rate less than 20%.

# Conclusion

We used the dataset that contains data about hotel bookings.

We cleaned and preprocessed the data and then we performed the exploratory data analysis to extract information from the data to answer the following questions.

?

# We Learned that –

- We See that 2016 Year is Most Booking year Comparision to the another year 2017 and 2015
- We see that bookings across years is higher for city hotel compared to resort hotel and donot increase proportionately over the years
- We see that More than 60% of the population booked the City hotel
- Bookings got canceled 37% of the time. While booking guest did checkd-in (did not cancel the booking) almost 63% of the time.
- We see that Around 40% bookings were cancelled in city hotel and 30% in resort hotel.
- We can see most bookings were made from July to August. And the least bookings were made at the start and end of the year

- We can see Resort has more bookings at the start and end of the year, and lower booking in June and September.
- Indirect bookings through online and offline travel agents are higher compared to direct bookings and same is the case with group bookings which are also high.
- We see Transient customer types have higher cancellations.
- Portugal, UK and France, Spain and Germany are the top countries from most guests come, more than 70- 80% come from these 5 countries.
- We see that Couple (or 2 adults) is the most popular accommodation type
- We see that More than 60% guest stay for one, two, or three