# CAPSTONE PROJECT -1

EDA on Hotel Booking Analysis
By
Tripti Shukla

# PROJECT TYPE EDA (EXPLORATORY DATA ANALYSIS)

- **Contribution** Individual
- <u>Team Member 1</u> Tripti Shukla

# HOTEL BOOKING ANALYSIS A BUSINESS PROBLEM

In recent year city hotel and Resort hotel have been high cancellation rates .Each hotel is now dealing with a number of issues as a result , including fewer revenues and less than ideal hotel room use .

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 cancellation as well as other factors that have no bearing on their business and yearly revenue generation are the main topics of this report



#### **PROJECT SUMMARY**

- The Hotel Booking Data Analysis Project consists of two hotels- 1)Resort Hotel
   2)City Hotel
- We will do an analysis of these two hotels from the year 2015-2017. In this data analysis, there are 119390 Rows and 32 column.
- It has information such as 1) Type of hotel booked 2)Average daily rate (ADR) 3)Booking details, 4)Arrival date 5)Length of the stay, 5)The number of adults, children, and / or babies booked, 6)Customer country, 7)Meal preferences, 8)Type of the customer, 9)Parking space details, 10)Reservation status 11) Channels used for booking, 12)Booking cancellation details, 13)Among other details.
- ❖ It has 3 types of analysis in this data −1)Univariate Analysis
  - 2)Bivariate Analysis
  - 3) Multivariate Analysis
- After collecting data it's very important to understand your data. So we had hotel booking analysis which had 119390 rows and 32 columns. So lets understand this 32 columns.

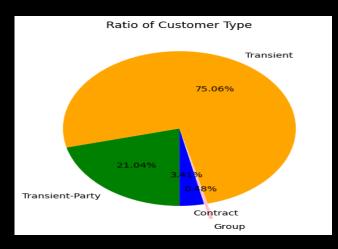
#### **VARIABLE DESCRIPTION**

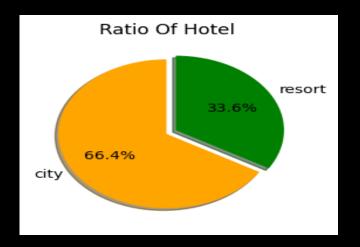
- hotel: Name of hotel(Categorical)
- \* is\_canceled: the booking is canceled or not (0 for no canceled and 1 for canceled) (Numerical)
- arrival\_date\_year: Year of arrival (Numerical)
- arrival\_date\_year: Year of arrival (Numerical)
- arrival\_date\_month: month of arrival (Categorical)
- \* arrival\_date\_week\_number: week number of arrival date. (Numerical)
- \* arrival\_date\_day\_of\_month: Day of month of arrival date . (Numerical)
- \* stays\_in\_weekend\_nights: No. of weekend nights spent in a hotel. (Numerical)
- stays\_in\_week\_nights: No. of weeknights spent in a hotel. (Numerical)
- \* adults: No. of adults in single booking record. (Numerical)
- \* children: No. of children in single booking record. (Numerical)
- \* babies: No. of babies in single booking record. (Numerical)
- meal: Type of meal chosen (Categorical)
- country: Country of origin of customers (as mentioned by them) (Categorical)
- market\_segment: What segment via booking was made and for what purpose. (Categorical)

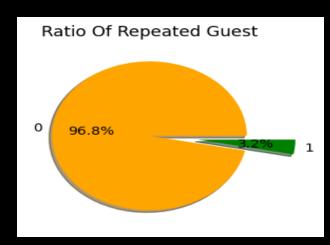
- distribution\_channel: Via which medium booking was made. (Categorical)
- \* is\_repeated\_guest: Whether the customer has made any booking before(0 for No and 1 for Yes) (Numerical)
- previous\_cancellations: No. of previous canceled bookings. (Numerical)
- previous\_bookings\_not\_canceled: No. of previous non-canceled bookings. (Numerical)
- reserved\_room\_type: Room type reserved by a customer. (Categorical)
- \* assigned\_room\_type: Room type assigned to the customer. (Categorical)
- booking\_changes: No. of booking changes done by customers (Numerical)
- deposit\_type: Type of deposit at the time of making a booking (No deposit/ Refundable/ No refund) (Categorical)
- \* agent: Id of agent for booking. (Numerical)
- \* company: Id of the company making a booking. (Numerical)
- days\_in\_waiting\_list: No. of days on waiting list. (Numerical)
- customer\_type: Type of customer(Transient, Group, etc.) (Categorical)
- \* adr: Average Daily rate. (Numerical)
- \* required\_car\_parking\_spaces: No. of car parking asked in booking. (Numerical)
- total\_of\_special\_requests: total no. of special request. (Numerical)
- \* reservation\_status: Whether a customer has checked out or canceled, or not showed. (Numerical)
- \* reservation\_status\_date: Date of making reservation status. (Numerical)

- \* Ratio of Hotel: Most preferred Hotel is city Hotel 66.4%.
- Ratio of Cancelled Booking: 37.0% booking were got cancelled out of all the booking.
- \* Ratio of Repeated Guest: 3.2% people were revisited the hotels. Rest of 96.8% were new guests. Thus ratio rate is low.
- \* Ratio of customer type: Most of the customers/guest were Transient type 75.06%.

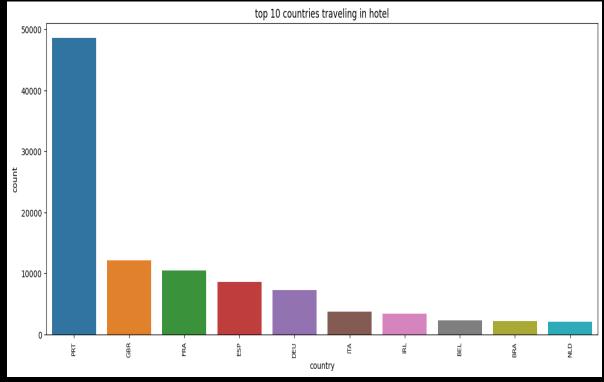




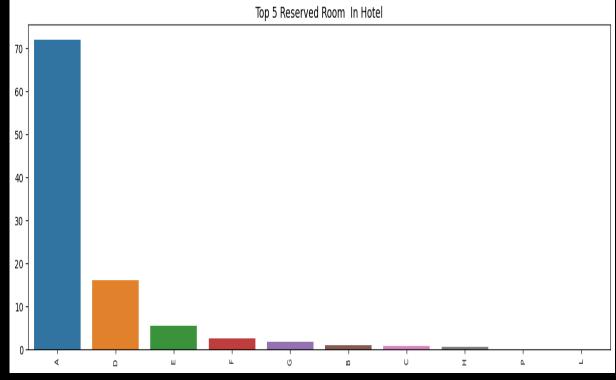




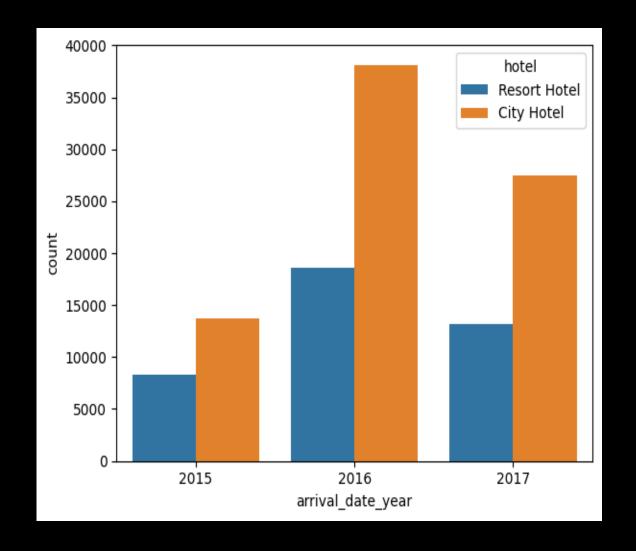
- \* Maximum Number of guest were from PRT.
- Minimum Number of guest Were from NLD



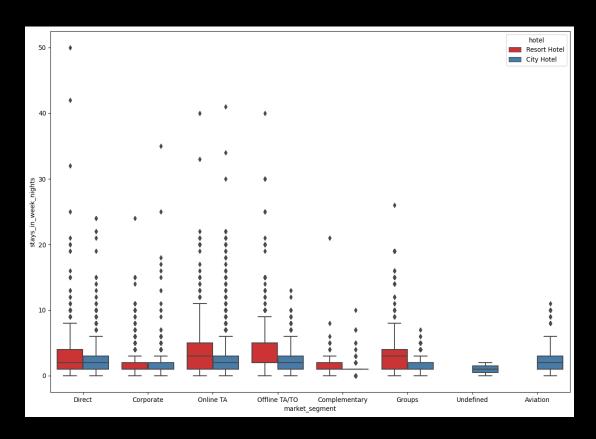
- Most Preferred Room is A type.
- Least Preferred Room is H type



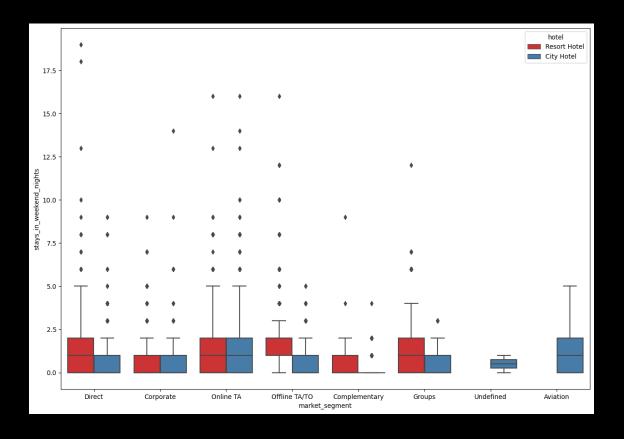
- In 2016 city hotel most preferred and less preferred in 2015.
- In 2016 Resort Hotel is also preferred and less preferred in 2015.



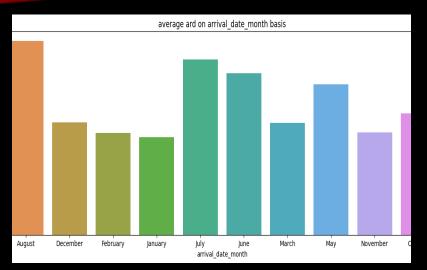
- ♦ Most reserved Hotel in week in online TA.
- Less reserved Hotel week in undefined and Aviation

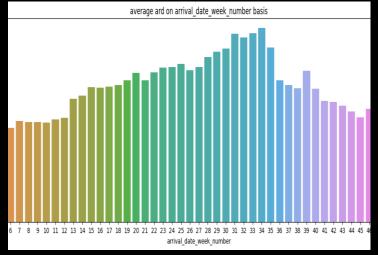


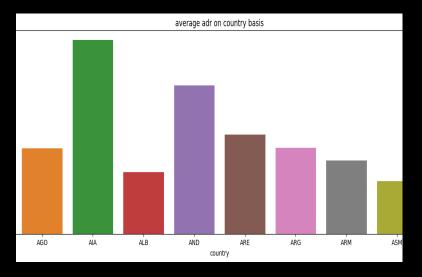
- Most Reserved Hotel in weekend in Online TA.
- Less Reserved Hotel week in undefined and Aviation



#### OBSERVATION ON ADR







#### Analysis of ADR on Month

- Most busiest month is August
- less busiest month are December , February , January

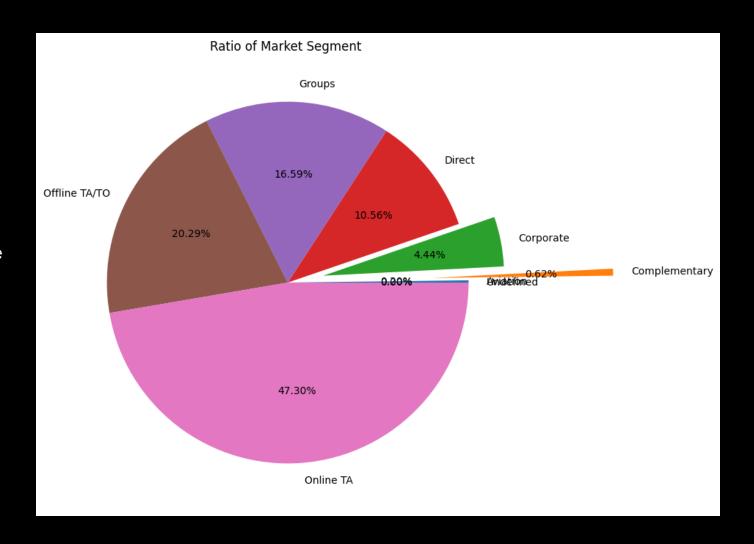
## Analysis of ADR on Week number

Most busiest week in hotel is 34 to 36

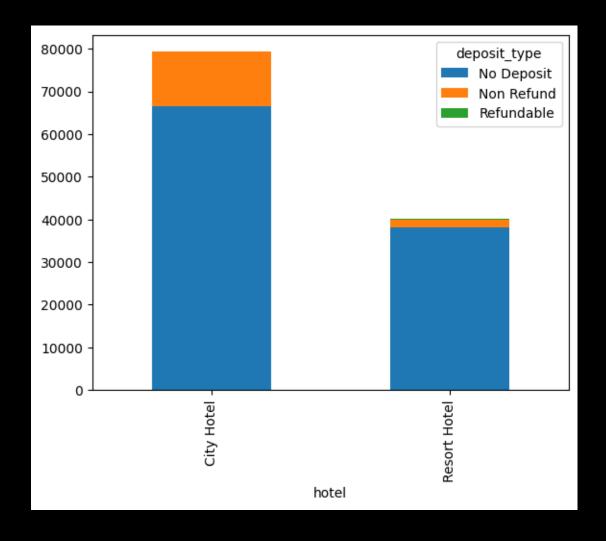
#### Analysis of ADR on Country

 Most visited country according to ADR is AIA

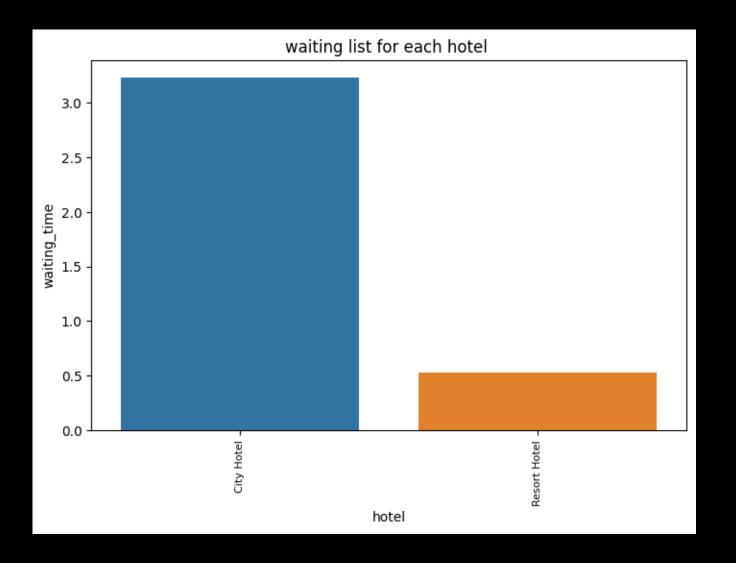
Most customer come through Online TA Market segment



- According to visualization ...no deposit type to higher in both hotel
- Refundable polices is not demanding in city hotels



- In this analysis ...most waiting booking in city type hotels ...
- Less booking in resort type hotel



## ASSUMPTIONS

- No unusual occurrences between 2015 and 2017 will have a substantial impact on the data used.
- The information is still current and can be used to analyze a hotel's possible plans in an efficient manner.
- Cancellation result in vacant rooms for the booked length of time.
- Client make hotel reservations the same year they make cancellations.