Team name
Sandeep Salunke
Sakshi Jadhav
Almabetter Banglore

GitHub repo link

https://github.com/sakshi11jadhav8/EDA-on-Hotel-Booking-Analysis-

Hotel Booking Analysis

Project Summary

The Hotel booking data is provided with the different variables. It consists about 32 variables which contain different kind of information such as hotels,is_canceled,lead_time,arrival_date_year,arrival_date_month,arrival_date_week_number,arrival_date_day_month,stays_in_weekend_nights,stays_in_week_nights,adults,children,babies,meal,country,market_segment,distribution_channel,is_repeated_guest,previous_cancellations,previous_bookings_not_canceled,reserved_room_type,assigned_room_type,booking_changes,deposit_type,agent,company,days_in_waiting_list,customer_type,adr,required_car_parking_spaces,total_of_special_requests,reservation_status,reservation_status_date.

As the first step of the project we have performed data cleaning as well as data wrangling in the next step we have performed Exploratory Data Analysis(EDA) in which we have created different questions related to the both of the hotels. Some questions based on both of the hotels and some are individually divided.

First we have done comparison between the hotels means which hotel is most preferred by peoples then likewise we have checked about cancellation, month of visit, which hotels prefer for reservation this type of analysis gives the brief information about both hotels which is most popular among the peoples

When we analysed which type of customers visits to hotel it is found that transient type of customers most visit to hotel and group type of customers less visit to hotel it means hotels should have more focus on group type of customers. When we have analysed which date is most crowded in months it is found that 18th of month is most crowded hotels should have focus to arrange necessary arrangement nearby of the date.

When we have checked the popularity of both the hotels the city hotel is most popular among the peoples we have used the bar graph to show the data visualization

Lastly we have shown correlation through the heat map it defines that is_canceled and same_room_allotted_or_not are negatively correlated. That means customer is unlikely to cancel his bookings if he don't get the same room as per reserved room. We have visualized it above that lead_time and total_stay is positively correlated. That means more is the stay of customer more will be the lead time. Adults, children's and babies are correlated to each other. That means more the people more will be adr. is_repeated guest and previous bookings not canceled has strong correlation. may be repeated guests are not more likely to cancel their bookings.

Contributors Roles:

Sandeep Salunke

1. Data wrangling:
   1. Hotel Dataset
   2. City hotel,Resort hotel
2. preferred Hotel
3. Cancellation and Non cancellation
4. Month analysis
5. Reservation analysis
6. Guests
7. Customer type
8. Market segment

Sakshi Jadhav

1. Data wrangling
   1. Hotel bookings
   2. Agent
2. Agent bookings
3. Lead time

4.Meal
5.Waiting time
6.Correlation
7.Deposit Type