

Exploratory Data Analysis

Hotel Bookings

Chandrashekhar Awate

Data science trainee

AlmaBetter, Bangalore

Abstract: We know the importance of Hotel and tourism industry and its share to the economic growth of any country, In this project we are going to do exploratory data analysis on the hotel booking dataset to help hotels to grow there business and make more profits scientifically.

PROBLEM STATEMENT: Helping Hotel business studying the past records and help the Hotel industry to grow exponentially?

The data of Hotel booking Was given for Exploratory Data Analysis. It contains Total 32 columns and 119390 rows.

Lets see the column first.(Excluding irrelevant columns)

- 1)hotel – Hotel type(city hotel or Resort hotel)
- 2)is_canceled – Booking cancellation
- 3)lead_time - the number of days between the time a guest books their room and the time they are scheduled to arrive at the hotel.
- 4)arrival_date_year – year of arrival (2015,16 or 17)
- 5)arrival_date_month – Month of arrival(Jan to dec)
- 6)arrival_date_week_number – Week no of arrival in year out of 52 or 53
- 7)arrival_date_day_of_month – arrival day in moth
- 8)stays_in_weekend_nights – How many days people lived on weekend
- 9)stays_in_week_nights – How many days people lived on week nights
- 10)adults – Adult members stayed in hotel
- 11)children – child members stayed in hotel
- 12)babies – Baby member stayed in hotel
- 13)meal – Meal type

Undefined/SC -no meal package;

BB — Bed & Breakfast;

HB -Half board (breakfast and one other meal — usually dinner);

FB -Full board (breakfast, lunch and dinner)

- 14)country – Visitors country

15)market_segment - “TA” means “Travel Agents” and “TO” means “Tour Operators”

16)distribution_channel - The term “TA” means “Travel Agents” and “TO” means “Tour Operators”

17)is_repeated_guest – repeated guest in hotel
 18)previous_cancellations
 19)previous_bookings_not_canceled
 20)reserved_room_type – Categories in room type
 21)assigned_room_type – Room type allotted to customer
 22)booking_changes – Charges of booking hotel
 23)deposit_type – No deposit,refundable,non-refundable
 24)agent – Booking agent
 25)company – Booking done by company
 26)days_in_waiting_list – Waiting days
 27)customer_type –
 Contract — when the booking has an allotment or other type of contract associated to it;
 Group — when the booking is associated to a group;
 Transient — when the booking is not part of a group or contract, and is not associated to other transient booking;
 Transient-party — when the booking is transient, but is associated to at least other transient booking

28) adr - Average Daily Rate

29)required_car_parking_spaces – requirement of car parking spaces

30)total_of_special_requests – Special requests

31)reservation_status -Check in or check out

32)reservation_status_date – Check in date or check out date

- **Exploratory Data Analysis**

After loading the dataset we performed EDA on dataset and tried to find out some insights column wise and visualised them for better understanding. For our simplicity we merged the column babies and children under kids column and we again merged the column kids and adults under total members. We looked only some important column which are going make differences in growth of the Hotels.

- **Null values Treatment**

Our dataset contains a large number of null values in the column agents and company so we dropped these columns. And after that we dropped all the rows where null values were presents.

CONCLUSIONS

- 1) May –August happens to be the most occupied months.**
- 2) The majority of guests come from western Europe countries.**
- 3) The majority of reservations are for city hotels.**
- 4) The number of repeated guests is too low.**
- 5) Most of the customers are couple as 2 members have largest bookings.**
- 6) The Rates of city Hotels are less as compared to the Resort Hotels and this the reason why City Hotels have Huge bookings compared to Resort Hotels.**
- 7) Most of the bookings or in fact near about 67% bookings are done by travel agents.**
- 8) 77% Meals booked are of BB type followed by HB with 12%**
- 9) 72% people booked BB type of Meals**
- 10) Almost 82% bookings were done with the help of Travel agents or Tour operators**

Recommendations On Behalf of EDA

- 1) Hotels should target the customer during the time of May- August as they are busiest months.**
- 2) The Hotel Industry should work on spending its budget for marketing in Europe because of the customer size from those areas.**
- 3) The budget should also be spend on building advanced infrastructure in city hotels so that it can attract more customers.**
- 4) The number of repeated guest is too low so the services should be given properly and there should be some sort of room for feedback.**
- 5) Most of the customers are couple so Hotel should give them discounted rate on visiting again so that the customers will be more in future. And also Hotel can construct some Special packages for them.**
- 6) The resort Hotels should minimize their Rates so that more customers will choose them.**
- 7) As the majority bookings are done by Travel agents or Tour operator Hotels should give incentives to these Travel agents or Tour operator So that they can again engage with their hotels.**
- 8) The Quality of BB Meal type should be great as lot of customer books those and there satisfaction will lead to repeated guest.**

9)Hotels Can establish their existence on websites and take bookings directly from customers and also give some discount for first time visitors.