Hotel Booking Analysis

Shreedarsh M, Sachin S Panchal, Ranjith, Rohan A G, Kshipra Praihar
Data science trainees,
AlmaBetter, Bangalore

Abstract:

Business from any domain mainly focusses on Profit, keeping that in mind we have done the analysis how various aspects would impact on the Business.

Insights obtained from the analyzation and implementing the needed change will improve the business drastically and solve many kinds of problems.

This dataset contains the information about various hotels across the globe which contains the information about type of hotel, number of bookings across various timelines, length of the stay etc., etc.,

We have obtained meaningful insights on how strong aseasonality impact on the business of hotel industry specially resort type of hotels. And we can get the idea of why there is a high rate of booking cancellation.

1.Problem Statement:

The project contains real world data record of hotel booking of a city and a resort hotel containing details like booking cancellation, area details etc. from 2015 to 2017. Main aim of the project is to understand and visualize meaningful insights from this dataset and analyse from both hotel and customer point of view.

Description of variable used in our dataset is as given in table below:

Variable	Description
is_canceled_counts	Count number of bookings which are canceled by the customers.
booking_percentage	Percentage of booking canceled by the customer versus percentage of customers checked-in to the hotel .
deposit_type	Count number of different types of deposit policies accepted by the customer.
year_counts	Total number of bookings across different years
country_counts	Top 10 countries of maximum customers.
total_nights_stays	Number of nights customers want to stay in hotel.
arrival_date_months_count	Number of customers arrived in the hotel across all the month.
market_segment_customers	Number of bookings across various market segments.
meal_	Different types of meal preferred by the customers.
customer_type_	Different types of booking done by the customers. 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 another transient booking
room_type_booking	Different types of room booked by the customers.
ADR	Average Daily Rate of Hotels across different months

2. Introduction:

- The hotel industry is one of the most important components of the wider service industry, catering for customers who require overnight accommodation. It is closely associated with the travel industry and the hospitality industry, although there are notable differences in scope.
- Overall, sales from hotel accounts 87.4% of industry revenue (in India 53.9% and 8% total employment rate)
- This makes analyzing the patterns available in the past data more important to help the hotels plan better. Using the historical data, hotels can perform various campaigns to boost the business.
- We will be using the data available to analyze the factors affecting the hotel bookings. These factors can be used for reporting the trends and predict the future bookings.

3. City Hotels and Resort Hotels:

It is Observed that out City Hotel has a greater number of bookings compared to Resort Hotel. The obvious assumption for this trend would be because of monetary reasons. City hotels would be more affordable than resort hotels due to accessibility, reliability and lavishing factors.

4.Peak Seasons:

We can see that 2016 seems to be the year where hotel booking is at its highest. We also see an increasing trend in booking around the middle of the year, with August being the highest followed by July and May. Summer ends around August, followed straight by autumn. It seems that summer period is a peak period for hotel booking**.** We can also see that January month has lowest number of customers followed by November and December. It seems Winter period is at lowest peak for hotel booking.

5.Best time to book hotel:

Since the month of January has the lowest amount of booking, it can be the best time of year to book a hotel room. Due to less demand of rooms, the cost for room on daily basis also minimum as compared to other months where month of August has high demand for room so it is obvious that the cost of room is also at peak.

6.Reasons for cancellation:

Out of 119000 customers ,75000 of customers checked-in the hotel while 44000 of customers cancelled their bookings. If we talk about percentage, 37% of bookings got cancelled Whereas 63% of customers did check-in. So, we realise that the high rate of cancellations can be due to no deposit policies.

7. Steps involved:

- **Explore The Dataset:**

  After loading the dataset, we explored the data and divided the project into three different categories as 1. Hotel wise, 2. Booking wise, and remaining part 3. Type of rooms, meal, customer, market segment, countries etc.

- **Null values Treatment:**

  Our dataset contains a large number of null values which might tend to disturb our operations hence we replace them at the beginning of our project in order to get a better result.

- **Exploratory Data Analysis :**

  In this section, we try to making some insights, finding out reasons for variation of bookings across different years with types of hotels, different countries and so on.

7. Conclusion:

After exploring the dataset deeply and analyzing the data by various means trained by the lectures. We have set a very meaningful insights from the dataset we have given and finalized the root causes of some problems occurred in the hotel industry, and have obtained possible solutions to tackle it, we have got to know peak seasons, best time to book and so on.

8. References:

1). https://pandas.pydata.org/
2). https://matplotlib.org/
3). https://seaborn.pydata.org/
4). Geek for geeks
5). Almabetter Sample deliverable.