

PROJECT SUMMARY

Project Name: Hotel Booking Analysis

Project Type: EDA

Contribution: Individual

Name: Abhijeet Kumar

Email: abhijeet.kr7252@gmail.com / abhikr7252@gmail.com

GitHub link:

https://github.com/abhikr11/EDA_Hotel_Booking_Analysis_Project

Summary:

The Hotel Booking Analysis project aimed to conduct an exploratory data analysis(EDA) on hotel booking data to identify patterns and insights that could help hotel managers make informed decisions. The project utilized data of two hotels - City hotel and Resort hotel - over a period of three years (2015, 2016 and 2017). The data included information on bookings, cancellations, hotel type, room type, number of guests, duration of stay, arrival date, month, country, etc.

The project began loading the data and perform data cleaning, which involved removing duplicates, missing values, outliers and changing data type of required variables. New columns were added accordingly. The cleaned data was then analysed using visualization tools. The charts used for visualization were pie chart, bar chart, line plot, histogram, box plot, kernel density estimate (KDE plot), heat map and pair plot. NumPy and Pandas libraries were used for data wrangling and feature engineering. Matplotlib and seaborn libraries were used for visualization of patterns and trends.

One of the main objectives of this analysis was to identify patterns and trends in hotel booking behaviour. The exploratory analyses included basic summary statistics, such as mean, median, and mode, as well as more advanced visualizations such as histograms, scatterplots, and box plots. To accomplish this, we examined various aspects of the booking process, such as the most popular hotels, booking rates by month, booking channels, room types, distribution of booking cancellations and the reasons for cancellations.

The analysis revealed several insights that could help hotel managers improve their operations and customer service. For example, the data showed that the highest cancellation rate occurred for bookings made through TA/TO, which suggests that hotels may need to reconsider their partnerships with these agencies. The data also showed that the average daily rate was higher for the city hotels. The analysis showed that customers tend to book hotels further in advance for longer stays, suggesting that hotels could benefit from offering early bird discounts for longer stays. Additionally, the analysis revealed that certain types of customers are more likely to cancel their bookings, suggesting that hotels could benefit from offering more flexible cancellation policies for these customers.

A correlation analysis was conducted to identify any relationships between different variables in the dataset, such as the relationship between total guests and average daily rate (ADR). The pair plot was used to plot scatter plot between different variables like previous cancellations, ADR, repeated guest or first visit.

In conclusion, the Hotel Booking Analysis project was successful in conducting an exploratory analysis of hotel booking data to identify patterns and insights that could inform decision-making for hotel managers. The insights gained from this project could help hotels improve their operations and enhance customer satisfaction.