

Capstone Project-1

Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

Name: -Ankur Vishwakarma

Email ID:- ankurvish1920@gmail.com

- notebook for data cleaning, data manipulation, and in EDA Visualization (Booking Cancellation analysis and Customer centered analysis).
- PPT, finalizing notebook making sure all the points to be covered.
- Write Summary of Project.
- Write conclusion for project.

Please paste the GitHub Repo link. :- <https://github.com/ankurvish1920>

please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

The goal of data science is to construct the means for extracting business- focused insights from data. This requires an understanding of how value and information flows in a business, and the ability to use that understanding to identify business opportunities.

The analysis of hotel data-set aids to build more customer oriented business module. This would be very useful for hotelier to strategize upcoming business focused on certain discussed categories in this analysis. This analysis will help in increasing business opportunities with precise and accurate objectives.

In the initial phase, we have focused more on the data cleaning and analyzed data in various categories. In later part we have tried to come out with conclusion for give problem statement. From this we have tried to bring out best results out of our analysis.

We have faced major challenge in data cleaning. After data exploration we found that 31994 duplicate rows and 94786 null values. In Data Cleaning we have performed few steps to ensure the data quality such as replacing null values with 0 and removing duplicate rows. We have converted columns to appropriate data types and also added appropriate columns.

From Exploratory Data Analysis we got to know that few columns which have categorical data having numerical type. With the cleaned data, we have performed Exploratory Data Analysis to understand our dataset in five different categories like 1. Univariate analysis, 2. Bi-variate analysis, 3. Distribution channel wise analysis, 4. Booking cancellation Analysis, 5. Customer centric analysis. 6. Special

request analysis.

Under univariate analysis. We have explored . August, may, October and july are months where customer tend to move to hotels more as compared to other months, so inorder to improve profits these months are best for increasing the revenue. From other categories analysis, we can see that given dataset is given for two types of hotel that City hotel and Resort hotel. City hotel is most preferred by guests and also it has significantly longer waiting time than Resort hotel. For longer stay, Resort hotel is more preferred than city hotel. It was also seen that weekdays nights were spent more than weekends nights so accordingly business can be targeted.

From the pie chart, it is clear that 72.51% people have not cancelled the bookings and 27.51 % people have cancelled bookings in past as per the given data. And Almost more than 1/4th of the customer cancelled bookings , to keep the bookings cancellation or there should be certain dues for cancelled bookings.

most customers stay in hotel for 1 to 4, there are hardly some customer who stays for more that 12 days. interest of customer and hotel industry stringent measures should be taken to reduce none stayed for more than 28 days according to past collected data. the data we analyzed that, prt - 27453, gbr - 10433, fra - 8837, esp - 7252, deu - 5387, ita - 3066 are the top 5 countries in terms of hotel booking indicates people move to those hotel frequently rather than rus - 561 nor - 515 rou - 458 no.s of bookings as they might be more remote from the city place.

Our motive in whole project was to analyze the data and find out components that affect hotel business and also customer's reluctance to book in hotel.

Months from May to Aug were in high peak in terms of booking and few guests were in winter. City hotel has longer waiting time as well as higher lead time than resort hotel. City hotel has more guests during spring and autumn, when the prices are also highest but in July and August we can see that there are less visitors although prices are lower. Thus, customers can get good deal on bookings in July and August in city hotel.

Most of bookings was from Travel Agent and least bookings were received from Corporate. Hotels need to do marketing and give special incentives for direct bookings which will help to establish personal relationships.

For Resort hotel, guest numbers is going down slightly from June to Sept which is also when prices are highest. Thus, these months should be avoided for bookings.

Also, we analyzed that huge number of visitors from Western Europe namely Portugal, UK and France being the highest where marketing team can target people

of this region.

We also found that average daily rate is more in Resort hotel than City hotel. For longer time of stay, large families can prefer Resort hotel type. Since lead time for Resort hotel is slightly higher than city hotel it means customers plan their trips very early as reserved does not affect adr.

It was also observed that total of special requests is higher in City hotel than Resort hotel. City hotel has more revenue than Resort hotel type.

The dataset contains immense possibilities to improve business values and have a positive impact. It is not limited to the problem taken into consideration for this project. Many other interesting possibilities can be explored using this dataset.

From the results and process we have implemented and we can conclude that we have achieved this group project objective which is analyzing the Hotel Booking data set and determine trends of two types of hotel type which is Resort type and City type hotel.

as reserved does not affect adr.

It was also observed that total of special requests is higher in City hotel than Resort hotel. City hotel has more revenue than Resort hotel type.

The dataset contains immense possibilities to improve business values and have a positive impact. It is not limited to the problem taken into consideration for this project. Many other interesting possibilities can be explored using this dataset.

From the results and process we have implemented and we can conclude that we have achieved this group project objective which is analyzing the Hotel Booking data set and determine trends of two types of hotel type which is Resort type and City type hotel.