Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
1. **Poonam Dattu Shevkar** **Email-id :-** dattupoonam@gmail.com ❖ **Contribution -** • Checking the presence of NAN values by using the tool missing value matrix. • Visualization of the Best Time of Hotel Bookings in the year. • Visualization of Special Requests Received from the Guests to Hotels. • Overview of the country for Hotel Bookings using Word Cloud , Area & Folium map. • Deposit Type for Hotels. • Overview of Market Segment and Distribution Channel. • Correlation between the variables using Heat map. • Scatter plot for lead time and stays for week nights to observe the pattern of variables. 2. **Sanjay Ramkishan Verma** **Email-id :-** shankyverma1998@gmail.com ❖ **Contribution-** • Most Busy month for Hotel Bookings. • Visualization of Guests pay for a room per night using Box Plot and checking the prices vary per night over year. • Optimal Stay of the guests in the Hotel using Bar Plot. • Bookings and Cancelled Bookings by different types of Customers. • Meal Preference by Customers. • Compilation of Team Collab Notebook.
Please paste the GitHub Repo link-

Github Link:-

https://github.com/Sara19598/Hotel-Booking-Analysis

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

Hotel industry is a very volatile industry and the bookings depend on a variety of factors such as type of hotels, seasonality, days of week and many more. 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.

The Hotel Booking Dataset is an open-source database including information about the booking of hotels with different factors associated with it around the world from 2015 to 2017.These Dataset includes systematic data structure with 32 variables and 119390 rows to visualise the patterns of variables with analytical graphs & Charts.

In this EDA project,we were provided with the dataset containing information about the hotel booking factors as like Hotel,arrival Date/Month/Year,number of repeated guests, Stays at hotel,prices of hotel for per night and among others also.Our team mainly focus on the hotel booking factors which are responsible to show trend in the booking as per the availability of resources for Hotel type,Special requests from guests, Best time to book the hotel, prices variation over the year,overview of country for hotel booking by customers and meal preference.

We perform the data wrangling on the raw data to get the useful data without NAN values and observe the summary statistics of the dataset.Further, we divide the project into different parts such as the best time of hotel booking , optimal stay and price variation over the year,special requests received from the guests, overview of country and meal preference by customer.

In the analysis of best time for Hotel bookings,we observed that in the month of August City Hotel has the huge bookings compared to Resort Hotel.In conclusion,most of the busy months are July and August for City and Resort Hotel.

For the analysis of optimal stay and price variation over the year , we saw that Average daily rate per person per night is 94.95 in the Resort Hotel Whereas Average daily rate per person per night is 105.30 in the City Hotel and the average stay in City Hotel is nearly 3 days compared to Resort Hotel where guests stay for nearly 5 days.In conclusion, the average price per night

depends on room type and its standard deviation and majority of people stays for 5 nights in hotels.

Moreover, analysis of Special requests received from guests gave an appropriate view of hotel booking cancellation rate.In conclusion, as more special requests of customers are fulfilled by hotels the cancellation rate will decrease i.e. customers are more likely to stay in hotel and City hotel has high special requests compared to Resort hotel.

In the analysis of country wise Hotel Bookings , we observed that PRT has more frequency of hotel bookings than other countries like GBR,FRA,etc.In conclusion,people from all over the world stay in City and Resort Hotels and most guests are from Portugal & other countries from Europe.

Reviewing the Meal preference by customers,we saw that among the 3 types of meal ‘BB’-Bed & Breakfast option is most popular with a frequency of 77.26%.In conclusion, breakfast in bed is mostly preferred by the guests/customers.

In the end, we have visualized the correlation between the variables to observe the highly correlated variables in the dataset and scatter plot between two variables lead time and stay in week nights.In conclusion, variable “is_cancelled” is highly correlated with 5 variables which can be used to fit a model in future use and there is negative correlation relationship between lead time and stays in week nights.

Drive Link:-
https://drive.google.com/drive/u/0/folders/1Nf7k7M9GarQ0RohZDAmAW8M1Dd_EjS1e