**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| |  |  |  |  | | --- | --- | --- | --- | | **No** | **Name** | **Email** | **Contribution** | | 1 | **Sohail Mohammad** | mohammadsohail825@gmail.com | Data Processing, Data Cleaning, Bar graph, pie chart | | 2 | **Laveti. Syam Sagar** | [shyamsagar651@gmail.com](mailto:shyamsagar651@gmail.com) | Data Wrangling, Checking missing values, Heat map for check correlation | | 3 | **Bisetty Kodanda Nagesh** | [dpsnnagesh38@gmail.com](mailto:dpsnnagesh38@gmail.com) | Removing total null values in the data set, Finding the traffic areas, conclusion. | |
| **Please paste the GitHub Repo link.** |
| **Github Link:**   1. https://github.com/mohammadsohail825/Hotel-Booking |
| **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 data set contains booking information for city hotel and a resort hotel, and includes information such as when the booking was made, length of stay, the number of adults, children or babies and the number of available parking spaces, among other things.  First of all we will try to understand the meaning of all columns of the data frame. For this we will see the unique values attained by each column whose meaning we are unable to understand.  Cleaning data it removed the ambiguous data that can affected the outcome of data set. While cleaned the data we Removed duplicate rows, Handled missing values, Converted columns to appropriated data types, Added important columns.  Data set can also be used to look univariate analysis, Hotel wise analysis, Distribution Channel wise analysis, Booking cancellation wise analysis and time wise analysis’  For different types analysis we used Matplotlib, seaborn library for making define data in graph format like scatterplot, barplot, boxplot, countplot, pychart, kdeplot, lineplot and seaborn. seaborn used exploring the correlation between total\_stay and lead\_time and adr and total\_people, other graphs also very uses for analysis worked in this data set.  After all analysis we found lots of result for different wise analysis we mention some conclusion like Mostly guests stay for less than 7 days in hotel and for longer stays Resort hotel is preferred, Most of the guests came from european countries, with most number of guest coming from Portugal, Around 66% bookings are for City hotel and 34% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel. Also the overall adr of city hotel is slightly higher than Resort hotel, Almost 30% of bookings via Travel Agent/Travel Operator are canceled, For customers, generally the longer stays (more than 15 days) can result in better deals in terms of low adr, Bookings made via complementary market segment and adults have on average high no. of special request. |
|  |