**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| Team Member’s Name, Email and Contribution: |
| Name :Sneha Owandkar  Email : [sowandkar21@gmail.com](mailto:sowandkar21@gmail.com)  Contribution : I am contributing as a team player to learn, understand the data and prepare questions individually. shared my view in group discussions. I complete the data understanding part & library codes.  Also created a ppt . |
| Please paste the GitHub Repo link. |
| Github Link:- https://github.com/snehao5698/Hotel-booking-analysis/upload |
| Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words) |
| Data set name - Hotel Booking Analysis  Components:  Shape – 119390 rows × 32 columns  Columns - 'hotel', 'is\_canceled', 'lead\_time', 'arrival\_date\_year', 'arrival\_date\_month', 'arrival\_date\_week\_number', 'arrival\_date\_day\_of\_month', 'stays\_in\_weekend\_nights', 'stays\_in\_week\_nights', 'adults', 'children', 'babies', 'meal', 'country', 'market\_segment', 'distribution\_channel', 'is\_repeated\_guest', 'previous\_cancellations', 'previous\_bookings\_not\_canceled', 'reserved\_room\_type', 'assigned\_room\_type', 'booking\_changes', 'deposit\_type', 'agent', 'days\_in\_waiting\_list', 'customer\_type', 'adr', 'required\_car\_parking\_spaces', 'total\_of\_special\_requests', 'reservation\_status', 'reservation\_status\_date'  Problem Statement :   * For this project we will be analyzing Hotel Booking data. This data set contains looking information for a city hotel and a resort hotel , and includes information such as when the booking was made, length of stay ,the number of adults ,children and/or babies, and the number of available parking spaces. * Hotel industry is a very volatile industry and the bookings depends on above factors and many more. * The main objective behind this project is to explore and analyze data to discover important factors that govern the bookings and give insights to hotel management , which can perform various compaigns to boost the business and performance.   Approaches :  1. Understand the data.    Data understanding focuses on the comprehension of the information available in the project. In this step we basically check on the kind of variables provided with the dataset, dtype of the columns, shape of the data frame.  2. Basic cleaning.  Our dataset contains numbers of null values which might tend to disturb our accuracy hence we dropped them at the beginning of our project in order to get a better result.  Pandas isnull() and notnull() methods are used to check and manage NULL values in a data frame  Conclusion :    Some conclusions drawn from the analysis are as follows.  ● Customers preferred City Hotel more than  Resort Hotel.  ● The maximum length of stay is higher (than  city hotel) in resort type as the resort is  mostly used for vacation purposes.  ● Median value of staying days of both the  hotels are approximately equal.  ● Bookings in the month of August are highest  and January found lowest number of  bookings. |