**Capstone Project Submission**

|  |
| --- |
| **Name, Email and Contribution:** |
| 1. Mayank Dubey ([Mayankdubey2494@gmail.com](mailto:Mayankdubey2494@gmail.com))  * Data Wrangling   Loading and Discovering Data  Preprocessing  Enhancing Data  Data Classification  Structuring data  Enriching data  Data Validation  Preprocessing  Data Cleaning  Summarizing data  Enriching Data   * Code Evaluation * Data Analysis * Hotel and Its Type * Preferred and its type * Outlier * Booking Agents * Number of Countries * Special Request * Bookings – Month/Year/Daily * Average Daily Rate and Hotel * Stay Length in Hotel * Cancellation Rates * Distribution Channel and Booking % * Waiting time * Visualizations * Bar graphs and Pie charts * Line Chart * Line Plot and Scatter Plot * Heat Map(Correlation Graph) * Box Plot and Horizontal Bars * Scatter Plot * Debugging * Segmentation * Observations * Summarization * Conclusions |
| **Please paste the GitHub Repo link.** |
| Github Link:- <https://github.com/aksfasale/eda_hotel_booking.git>  Drive Link:- |
| **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 hotel industry is one of the most lucrative in the world.** With**international mobility and travel** becoming more and more widespread, the hotel market is turning increasingly competitive, and thus continuously expanding. wefrom 2015 to 2017. In this project I am going to analyze Hotel Booking Data in order to find out valuable insights and give suggestions to increase revenue of hotels.  As a first step I explored through all the files in the dataset to understand the content of dataset that I have. The dataset was needed to be analyzed properly as there were many null values, some data type mismatch, hence I started the first process that is data pre-processing and got clean up with all the duplicate, missing values and datatype with proper format.  Once, the data file was ready without any ambiguity in it I am ready to go for the **Data Analysis** part. Data analytics is the process of examining data sets in order to find trends and draw conclusions about the information they contain. Data analytics initiatives can help businesses increase revenue, improve operational efficiency, optimize marketing campaigns and bolster customer service efforts. Analytics also enable organizations to respond quickly to emerging market trends and gain a competitive edge over business rivals. The ultimate goal of data analytics, however, is boosting business performance.  The analysis was had two approach Univariate analysis and Multivariate Analysis, Univariate analysis allows us to understand the distribution of values for one variable while multivariate analysis allows us to understand the relationship between several variables.  For **Hotel**, I found that City Hotel was having the maximum bookings, having an upper hand over the Resort hotel, but in contrast to this those who chose to stay long prefer Resort hotel for the Stay. The hotels need to look at their retention rate, it is very low which is a negative story, needed to be improved.  For **People and Country**, I am able to derive that the large population comes for the European countries and that to most of them coming from Portugal. The number of people visiting the hotel rapidly increases after May till September, which is also the best period for the hotel’s revenue generation.  For **Hotel Rooms and Average Daily rate**, I see that most of the people preferred Room Type A which was having affordable average daily rate. For those whose average daily rate is higher i.e. (Type C, G, F, H) it's seen that preference is also less. Hotel should reconsider this and should make the necessary changes.  From the analysis I can conclude that there are lot of things which can be taken into consideration and hotel should make the necessary changes according to the analysis made. The average daily rate should be made less for some types of room so that people can seek them and return. Retention Rate should be one of the things to be taken care of as it was significantly very low. Most guests come from western European countries. I should spend a significant amount of our budget on those area. Around 61% bookings are for City hotel and 39% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel. Majority of the hotels booked are city hotel need to spend the most targeting fund on those hotels. I should also target months between May to September. |
|  |