**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| Nehal Shrikrushna Jambhulkar ([njambhulkar35@gmail.com](mailto:njambhulkar35@gmail.com))   * Getting the data * Data Cleaning * Exploratory Data Analysis: * Working on Outliers * Heatmap * Univariate Analysis:   1)which agent makes most no bookings?  2)which type of hotel is mostly preferred by the guest?  3)what is the percentage of cancellation?  4)what is percentage of repeated guest?  5)what is percentage distribution of “customer Type”?  6) what is the percentage distribution of required car parking spaces?  7)what is the percentage of booking changes made by the customer?  8)what is percentage distribution of deposite type?  9) Which type of food is mostly preferred by the guests ?   * Bivariate and Multivariate Analysis   1)Which Hotel Type has the highest ADR?  2) Which Hotel Type has the more lead time?  3)Which hotel has highest percentage of booking cancellation?  4) Which hotel has longer waiting time?  5) Which Hotels has the most Repeated guest?  7) Which distribution channel contributed more to adr in order to increase the income?  8)Which distribution channel has the highest cancellation rate?  9)Which Market Segment has the highest Cancellation rate?  10)Does the guest alloted with the same room type which was reserved by them?  11)ADR across diffrent market segment  12)What is the Optimal stay length in both types of hotels ?   * Make PPT and Technical Document |
| **Please paste the GitHub Repo link.** |
| Github Link:- https://github.com/nehal35/EDA-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)** |
| For my first EDA Capstone project I choose Hotel Booking Analysis as our project. This choose because I am familiar with hotel booking and had some experience about what factors look for while booking a room in a hotel. So this way I choose Hotel Booking analysis as our project.  The hotel and tourism industry plays an essential role in the economic growth of any country. In this project, we will do exploratory data analysis on the hotel booking dataset, in order to assist hotels in making more profits. Through feature selection, data analysis, and prediction with data visualization, we can understand what could be the factors behind the data This data set contains booking 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, among other things. All personally identifying information has been removed from the data.  For this project, we used Python libraries like Pandas, Matplotlib, Seaborn, etc. to examine, clean, and analyze the "Hotel Booking Analysis" dataset.  Before cleaning the data, we checked the total number of rows and columns in the data set, and we have 119390 rows and 32 columns. Our dataset contains a large number 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.  After data was cleaned then we plotted charts to draw inferences and conclusions.   * City hotels are the most preferred hotel type by the guests. We can say City hotel is the busiest hotel. * 27.5 % bookings were got cancelled out of all the bookings. * Only 3.9 % people were revisited the hotels. Rest 96.1 % were new guests. Thus retention rate is low. * The percentage of 0 changes made in the booking was more than 82 %. Percentage of Single changes made was about 10%. * Most of the customers (91.6%) do not require car parking spaces. * 79.1 % bookings were made through TA/TO (travel agents/Tour operators). * BB( Bed & Breakfast) is the most preferred type of meal by the guests. * Maximum number of guests were from Portugal, i.e. more than 25000 guests. * Most of the bookings for City hotels and Resort hotel were happened in 2016. * Average ADR for city hotel is high as compared to resort hotels. These City hotels are generating more revenue than the resort hotels. * Booking cancellation rate is high for City hotels which almost 30 %. * Average lead time for resort hotel is high. * Resort hotels have the most repeated guests. * Waiting time period for City hotel is high as compared to resort hotels. That means city hotels are much busier than Resort hotels. * Almost 19 % people did not cancel their bookings even after not getting the same room which they reserved while booking hotel. Only 2.5 % people cancelled the booking. * Optimal stay in both the type hotel is less than 7 days. Usually people stay for a week. |