**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| An exploratory data analysis is an analytical process of collecting, cleaning, interpreting, and manipulating the data to determine which variables are important and which do not play a significant role in the output.  The hotel industry is a highly competitive industry with regular advancements.   * Have you ever wondered about the optimal length of stay in a hotel? * What kind of meal is preferred more? * Which hotel type gets booked more? * Which month of the year has more incoming guests? * Which part of the world brings in more clients to the hotel? * What are the market segments in the hotel industry?   In this project, we are about to analyze a hotel dataset and try to answer more of these questions. In addition to all these questions, we do have an objective. The objective of this project is to find out the factors that contribute to booking cancellations. The one metric that measures the success of the business is revenue. We are going to dig deeper to drive more success for the hotel.  Our dataset consists of 119390 observations and 32 columns. We will have a fair amount of duplicates and Null values in the dataset, which is taken care of in the data cleaning and wrangling process. Every variable in the dataset will be checked up or visualized to derive a better business decision to drive increased bookings.  Analysis used in deriving relationships between variables:   * **Univariate analysis** - Analyzing a single variable in a dataset is Univariate analysis. * **Bivariate analysis** - Bivariate analysis is slightly more analytical than Univariate analysis. When the data set contains two variables, to know the relationship between the two variables, bivariate analysis guides us in exploring the relationship. * **Multivariate analysis** - When there is a relationship between more than two variables, multivariate analysis comes into play.   The EDA for the hotel dataset uses python. Analyzing a dataset is a hectic task that takes a lot of time. Python provides numerous open-source libraries or modules. like Numpy, Pandas, Matplotlib and Seaborn, etc., that can help through Exploratory Data Analysis by generating visualizations.  Few solutions to the business objective are     * The number of bookings and adr is high in August. We can promote bookings in other months with exciting offers while the bookings are low. * The number of booking cancellations is due to the no-deposit option in the deposit type. Implement deposit rate for bookings will reduce cancellations. * The contribution of marketing channels is low other than the online channel. Develop a marketing strategy to promote bookings. * Most of the bookings are through online channels, so effective advertisement and online presence drive more bookings |
| **Please paste the GitHub Repo link.** |
| Github link: https://github.com/srkpriyanka95/Hotel-Booking-Analysis-EDA.git |
| **Please paste the Google Drive link** |
| Drive link: |