**Capstone Project Summary**

| **Team Member’s Name, Email and Contribution:** |
| --- |
| **Abhishek Kumar-** [**kabhikr74@gmail.com**](mailto:kabhikr74@gmail.com)   1. **Data Cleaning**  * **Dropping Columns with Maximum Null values** * **Filling columns with minimum Null Values**  **Revenue generated by hotels in each month****Analysis based on Meal****Fare of hotels in different countries of the world****Top countries with maximum no. of hotel bookings** **Analysis based on Car parking spaces** **Analysis based on Repeated Guest.****Analysis based on reservation status****Analysis based on Special Request** **Mukesh Sablani-** [**ms24sablani@gmail.com**](mailto:ms24sablani@gmail.com)   1. **Data Manipulation**  * **Combining columns for EDA**  **Reservations in different Hotel Type****Comparative analysis of reservations and revenue of hotels**  1. **Study of reservation based on lead time and booking of hotels**  **Reservations Made in each Year**  1. **Booking made by different Distribution channel** 2. **Market share by different customer type**  **Preferred Room type****Bookings based on deposit type** **Sanjay Paul-** [**s8800117688p@gmail.com**](mailto:s8800117688p@gmail.com)   1. **Handling outliers**  * **By defining threshold values** * **By standard methods** * **Using IQR method** * **Replacing with median values**  1. **Optimal Stay Period**  **Total Stay of people based on ADR**  1. **Hotel Price distribution**  **ADR of hotels in each month****Relation between ADR and Lead time****ADR based on deposit type** **Mohita Rathour-** [**MahiSingh641@gmail.com**](mailto:MahiSingh641@gmail.com)   1. **Data Visualization**  **Histogram to show the distribution of data**  * **Plotting boxplot to find outliers**  **Cancellation made in each hotel type****Cancellations made in each Year****Cancellations made in each months****Cancellations for different distribution channels****Cancellations made due to deposits****Cancellation based on customer type** |
| **Please paste the GitHub Repo link.** |
| **Github Link:- https://github.com/sanjaypauloct1989/Project.6-Hotel-booking-EDA.git** |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| **Summary of the EDA capstone project on Hotel Booking.**  **The hotel industry is one of the most important components of the wider service industry, catering for customers who require overnight accommodation. This project discusses the key success factors of the hospitality industry. We have studied the data to get some to run a successful business.**  **In this study we have sample data about the hotel industry that is not processed for use. Unprocessed data gives inaccurate results. To process this data is called data cleaning. We have cleaned the data by handling null values, outliers and dropping unwanted columns.**  **We are then condensing the data for our study. This is Data Manipulation. We achieve this result by performing functions on the columns. We have come up with new columns such as total kids, guests, stays, revenue etc.**  **In our data study we have 2 types of hotels- the resort type and city hotel type. There are factors in the study which affect the business of the hotels. Factors such as location, ADR, Deposits charged, wait time, etc. We also have channels like distribution channel, Market segment to focus on to get more revenue.**  **We have used python to analyze the data. We have studied the factors which are affecting the booking and cancellations in hotels. With this study we can calculate the revenue. Any business is measured by its KPI. For this industry Revenue is the KPI.**  **After all this here are some prominent conclusions such as city type hotels are more preferred hotels. Cancellation is less in resort hotels. Whereas people prefer city hotels for shorter stay and resort hotels for longer stay. Bed and Breakfast is the most preferred meal type. City hotels having refundable policies and Resort hotels having no deposit are expensive. Also, Transient customers bring in most revenue and TA/TO are the most profitable distribution channels. Lastly Portugal is the most profitable destination with the highest revenue.** |