**Capstone Project Submission**

**Team Member’s Name, Email and Contribution:**

1. **SHUBHAM JAISWAL** [**(sgjswl95@gmail.com**](mailto:sgjswl95@gmail.com)**)**
   * Reading the Dataset
     + Comprehending each column in the dataset
     + Raising the bar of questions from the dataset
     + Gaining insightful inference from the dataset for analytical conclusions.
   * Cleaning the clusters in Data
     + Dropping the null value columns
     + Replacing the minimum number of null values with str.
   * Exploratory Data Analysis
     + Statistical Based Analysis
     + Market Based Analysis
     + Categorical Analysis
     + Profitability Analysis
   * Data Visualization
     + Plotting the insights in the graph using:

Bar Plot, Box plot, Scatter Plot, Pie charts, Heat maps, etc.

* + Inference Gained
  + Conclusion

1. **MRITYUNJAY SINGH CHANDEL (mrityunjaychandel98@gmail.com)**
   * Reading the Dataset
     + Comprehending each column in the dataset
     + Raising the bar of questions from the dataset
     + Gaining insightful inference from the dataset for analytical conclusions.
   * Cleaning the clusters in Data
     + Dropping the null value columns
     + Replacing the minimum number of null values with str.
   * Exploratory Data Analysis
     + Statistical Based Analysis
     + Market Based Analysis
     + Univariate and Bivariate Analysis
     + Categorical Analysis
     + Profitability Analysis
   * Data Visualization
     + Plotting the insights in the graph using:

Bar Plot, Box plot, Scatter Plot, Pie charts, Heatmaps, etc.

* + Inference Gained
  + Conclusion

**Please paste the GitHub Repo link.**

Github Link: <https://github.com/MrityunjaySingh18/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)**

This data article describes two datasets with hotel demand data. One of the hotels (H1) is a resort hotel and the other is a city hotel (H2). Both datasets share the same structure, with 31 variables describing the 40,060 observations of H1 and 79,330 observations of H2. Each observation represents a hotel booking. Both datasets comprehend bookings due to arrive between the 1st of July of 2015 and the 31st of August 2017, including bookings that effectively arrived and bookings that were cancelled. Since this is hotel real data, all data elements pertaining hotel or costumer identification were deleted. Due to the scarcity of real business data for scientific and educational purposes, these datasets can have an important role for research and education in revenue management, machine learning, or data mining, as well as in other fields.

**CONCLUSION**

Around 60% bookings are for City hotel and 40% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel. Also the overall adr of City hotel is slightly higher than Resort hotel. Mostly guests stay for less than 5 days in hotel and for longer stays Resort hotel is preferred. Both hotels have significantly higher booking cancellation rates and very few guests less than 3 % return for another booking in City hotel. 5% guests return for stay in Resort hotel. Guests use different channels for making bookings out of which most preferred way is TA/TO. Almost 30% of bookings via TA/TO are cancelled. Not getting same room as reserved, longer lead time and waiting time do not affect cancellation of bookings. Although different room allotment do lowers the adr. July- August are the most busier and profitable months for both of hotels .Within a month, adr gradually increases as month ends, with small sudden rise on weekends. Couples are the most common guests for hotels; Hence hotels can plan services according to couples needs to increase revenue. More number of people in guests results in more number of special requests. Bookings made via complementary market segment and adults have on average high no. of special request. For customers, generally the longer stays (more than 15 days) can result in better deals in terms of low adr.

PROBLEM

Data is not cleaned first of all we have to prepare the data and after there are various null values are given in this set of data

Other various weather factor should also be consider in this dataset

APPROACHES

We have use exploratory data analysis and in this we consider various factor and after then we take the correlation of the data and consider what factor should we consider so that data is presentable and we can use this for future planning so that this help them to consider there business to grow so we use Univariate and Bivariate analysis and after then we consider single factor how each hotel perform on same factor with their resources.