**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| 1. **Mithlesh Singh :-**   [mks937449@gmail.com](mailto:mks937449@gmail.com)     * Data wrangling:- * Finding the null values “nan”. * Replaced those null value with zeros. * Data preparation * Converting “Float” data type into “integer” * Convert the data type to “string”. * Deposit type * Repeated guest * Average daily rate * Monthly booking of different hotel  1. **Ashish Yadav :-**   Itsmeashu2000@gmail.com   * Data Wrangling :  1. Finding the null values “Nan”. 2. Replaced those null values with Zeros.  * Data Preparation  1. Converting “Float” data type into “integer”. 2. Convert the data type to “String”.  * Cancelled Bookings * Arrival Dates By Months * Year of Arrival * Market Segment   **3. Nikita Negi :-**  Nikitanegi4100@gmail.com   * Data Wrangling :  1. Finding the null values “Nan”. 2. Replaced those null values with Zeros.  * Data Preparation  1. Converting “Float” data type into “integer”. 2. Convert the datatype to “String”.  * Meals * Most preferred hotel percentage by customers * Customers type * Monthly booking of different hotel   **4. Rohit Sakharkar :-**  rohitsakharkar11@gmail.com     * Data Wrangling :  1. Finding the null values “Nan”. 2. Replaced those null values with Zeros.  * Data Preparation  1. Converting “Float” data type into “integer”.      * Merge “children”, “babies” and “adults” columns into “members” column * Confirmed booking * Customer type * Relation between price and month * Yearly booking of different hotel   **5. shubham sanklecha**  [Shubhamsanklecha400@gmail.com](mailto:Shubhamsanklecha400@gmail.com)   * Data Wrangling :  1. Finding the null values “Nan”. 2. Replaced those null values with Zeros.  * Data Preparation  1. Converting “Float” data type into “integer”. 2. Convert the data type to “String”.  * Booking changes * Reservation status * Repeated guest * Deposit type * Average daily rate * Monthly booking of different hotel |
| **Please paste the GitHub Repo link.** |
| https://github.com/Mk-singh55555/Hotel-booking-analysis.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)** |
| **Through this analysis, we were able to answer key business questions which are stated below:**   * Is there a difference in volume of bookings and revenue across the two type of hotels? * Is the change driven by a time period such as a year or a month? * Is the cancellation rate higher than the booking rate? * Are the cancellations driven by any factors such as time period or hotel type? * What is the average revenue generated per booking?   **Methodology**   1. Majority of the analysis was EDA which was digging one level deeper and getting the data to answer the above questions 2. To answer few questions I sliced the data across various cuts 3. Another methodology included using different lenses to view data across segments 4. I used classification tree to predict if the booking will go through 5. For estimating the revenue over the next 12 months, I used time-series modelling   **Insights**   * The number of bookings seems to be high in 2016 while the bookings seem to be less in 2015 and 2017. This is majorly due to more data points being available in 2016 * The ratio of bookings cancelled to the confirmed bookings seems high for City Hotels * We observe that the avg. # of bookings in a week is growing year over year * The growth is higher from 2015 to 2016 as compared to the jump from 2016 to 2017 * The growth is higher from 2015 to 2016 as compared to the jump from 2016 to 2017 * The revenue almost doubled in 2016 but it did not grow at the same rate from 2016 to 2017   **Implication of business**   * Based on the classification tree, hotels can predict if the booking will convert or not. This will allow them to make a waiting list based on the chance of cancellations to make the complete use of the capacity * The revenue forecast will help them plan promotions during the low revenue period   **Limitations**   * The data for 2015 and 2017 is for different months. Even though we have converted them to same base line of weekly numbers, there are chances that some weeks perform differently as compared to other weeks * The definition of new customers is not very well described. A new customer this year will be existing next year, or they can be existing customer from the 2nd booking. A deeper analysis in required based on definition * The weekday vs weekend analysis can be further drilled for the type of bookings |