

# Capstone Project Submission

## Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

### **Team Member's Name, Email and Contribution:**

## STAR SHAPIONS

### Contributors Roles:

#### 1. Syed Faruk([syedfaruk730@gmail.com](mailto:syedfaruk730@gmail.com))

- i) Data understanding
- ii) Data cleaning
- iii) Exploratory data analysis
- iv) Checking duplicate values with help of visualization.
- v) Checking relation between column with the help of pairplot.
- vi) Most preferred hotel by customers
- vii) PPT
- viii) Technical Documentation
- ix) Summary
- x) Year got best sales

#### 2. Shraddha Gujar([gujar\\_shraddha@yahoo.com](mailto:gujar_shraddha@yahoo.com))

- i) Data understanding
- ii) Data cleaning
- iii) Exploratory data analysis
- iv) Percentage distribution of deposit type
- v) Finding unique values
- vi) Data visualization
- vii) Most visited months
- viii) Technical Documentation
- ix) PPT
- x) Summary

#### 3. Arun Palle([aruun3000@gmail.com](mailto:aruun3000@gmail.com))

- i) Data understanding
- ii) Data cleaning
- iii) Exploratory data analysis
- iv) Data exploration
- v) Finding null values
- vi) Dropping columns
- vii) Type of room highly preferred / booked by customers.
- viii) Technical Documentation
- ix) PPT
- x) Summary

#### 4. Tejas Patil([tejaspatil9984@gmail.com](mailto:tejaspatil9984@gmail.com))

- i)Data understanding
- ii)Data cleaning
- iii)Exploratory data analysis
- iv)Checking duplicate values
- v)Relationship analysis
- vi)Which hotel has highest number of cancellation by customers
- vii)Type of customers highly visited on both hotels
- viii)Technical Documentation
- ix)PPT
- x)Summary

#### 5. Rishav kumar([rishavarya001@gmail.com](mailto:rishavarya001@gmail.com))

- i)Data understanding
- ii)Data cleaning
- iii)Exploratory data analysis
- iv)Correlation visualization by Heat map
- v)Checking outlier with help of Box plot
- vi)Percentage of Repeated guests
- vii)Data discription
- viii)Technical Documentation
- ix)PPT
- x)Summary

**Please paste the GitHub Repo link.**

Github Link:- <https://github.com/Aruun003/CAPSTONE-PROJECT-1>

**Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**

## Introduction:

i) This summary is about the project that we are provided with a hotel booking analysis dataset. ii) we completed this project with 5 team members and every one contributed equally in it.

This data set has a single file which compares various booking information between two hotels. a city hotel and a resort hotel.

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. A total of 32 variables.

All personal naming information has been removed from the data.

Both hotels are assumed to be in various locations, however their exact location and name are unknown.

The dataset holds a total of 119390 entries.

The dataset has a total of 32 columns.

As the first step we imported the csv dataset from the drive. Then we checked the top ten rows from hotel booking csv dataset. and for more clarity we also checked the last ten rows of hotel booking csv dataset. then our next step was the Data exploration in which we gathered the information from the hotel booking csv. Then we moved toward the data description part where we got clarity of dataset with the help of column names.

In this hotel booking analysis we drew the the following conclusions:

i) Guests mostly preferred city hotels as they had the maximum number of bookings.

ii) Month of August has the greatest number of visitors. as most of the bookings were in month of august.

iii) code A rooms are mostly preferred by customers as code A type of room has the greatest number of bookings.

iv) we also drew conclusions on which year got best sales.as we were provided with 3-year dataset 2015,2016, 2017.In these three years of dataset we found out that sales in 2016 is higher than 2015 and 2017.

v)After booking, city hotels were canceled the maximum number of times by

customers. vi)Transient types of customers have the greatest number of visits to both

hotels. vii) In the given dataset 3.2 % of customers are repeated guests.

viii)87.6% is deposit type of both hotels.

In the Entire hotel booking analysis we covered a range of factors that affects the hotel booking using a given dataset and we estimated the cancellation rate. Booking ratios according to months. most preferred hotel.most preferred type of room. Percentage of repeated guests. And the year which had the highest sales etc. With the help of this hotel booking analysis, we helped the hotels to estimate the BEST and WORST of their business. and we also tried to predict that whether or not a hotel is likely to receive a disproportionately high number of special guests/requests.