Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Contributor Role 1. **Rahul Chouhan** (rahul.chouhan.04@gmail.com): o Data Wrangling o Hotel Booking Percentage o Find the usual stay in hotel o Overview on canceled booking o Busiest month for hotels o Find the type of customers o Helped in PPT and Technical Document o Helped in Team Notebook , Planning and approaches 2. **Pankaj Rathod** (pankaj.rathod.1007@gmail.com): o Data Wrangling o Hotel Booking Percentage o Find the usual stay in Hotel o Overview on canceled booking o Busiest month for hotels o Find the type of customers o Helped in PPT and Technical Document o Helped in Team Notebook , Planning and approaches
Please paste the GitHub Repo link.
GitHub Link: https://github.com/Rahul-Chouhan-407/Capstone-Project-1-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)
The hotel and tourism industry plays an essential role in the economic growth of any country. In this project, we will do exploratory data analysis on the hotel booking dataset, in order to assist hotels in making more profits. Through feature selection, data analysis, and prediction with data visualization, we can understand what could be the factors behind the data This data set contains booking information for a city hotel and a resort hotel and 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. All personally identifying information has been removed from the data.
For this project, we used Python libraries like Pandas, Matplotlib, Seaborn, etc. to examine, clean, and analyze the "Hotel Booking Analysis" dataset.

The first step towards understanding the data set is to import essential libraries such as numpy, pandas, matplotlib and seaborn. Once the drive has been mounted, we load the CSV file to check the data.
Before cleaning the data, we checked the total number of rows and columns in the data set, and we have 119390 rows and 32 columns.
Our dataset contains a large number of null values which might tend to disturb our accuracy hence we dropped them at the beginning of our project in order to get a better result.

The purpose of this analysis is to examine these questions using a large dataset.

- Are you curious about hotel booking trends?
- When it comes to staying, how long do people stay?
- Is there a high rate of cancellations?
- When is the busiest time of year?

Helping Hotel business studying the past records and help the Hotel industry to grow exponentially

Almost 67% bookings are for City hotel and 33% bookings are for Resort hotel, therefore City hotel is busier than resort hotel. Most of the customers prefer to stay one, two, three or four nights, where number of bookings for two and three nights has a high demand.

In total, 119000 customers checked in at the hotel, while 75000 canceled their reservations. In percentage terms, 37% of bookings were canceled, whereas 63% of customers checked in. As a result, we realize that no deposit policies may account for the high rate of cancellations.

The demand for city hotels is always higher than that for resort hotels. It means the majority of the hotels booked are city hotel. So definitely need to spend the most targeting fund on these hotels.

Hotel bookings seem to be at their highest level in 2016 so far. Bookings also increase around the middle of the year, with August being the highest followed by July and May. Most of bookings are created from Portugal and then Great Britain in second place and France at third place.t is seen that most number of customers are having 2 adults customer and those who have more than 2 either containing adults, children & babies have the lowest number of customer.