Capstone Project Submission

Team Member's Name, Email and Contribution:

Name	VARSHA RANI
Email	e.varsharani@gmail.com
Coding	Individual
Presentation	Individual
Technical Documentation	Individual
Project Summary	Individual
Presentation Video	Individual

GitHub Repo link:

GitHub Link:- https://github.com/123-VARSHA/Exploratory-Data-Analysis-Capstone

Summary of Project

Problem Statement:

Have you ever wondered when the best time of year to book a hotel room is? Or the optimal length of stay in order to get the best daily rate? What if you wanted to predict whether or not a hotel was likely to receive a disproportionately high number of special requests? This hotel booking dataset can help you explore those questions!
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.

Libraries Used:

- Pandas
- NumPy
- Matplotlib
- Seaborn

Steps Involved:

- Data Exploration
- Data Cleaning
- Data Preparation
- Visualization

Conclusion:

As per the EDA, the important factors derived from the given dataset are :

- There are two different types of hotels namely - 'Resort Hotel' and 'City Hotel'.
- Bookings across the years 2016 and 2017 are higher for City Hotels compared to Resort Hotels and do not increase proportionately over the years.
- Top 10 countries from where guests are coming are - PRT, GBR, FRA, ESP, DEU, IRL, ITA, BEL, NLD and USA. Most number of guests i.e. 17573 are coming from PRT.
- Focusing on bookings cancellation, around 30% of bookings were cancelled in City Hotels and 25% in Resort Hotels.
- The City hotel has more guests during spring and autumn, when the prices are also highest, In July and August there are less visitors, although prices are lower. Thus, customers can get good deal on bookings in July and August in city hotel.
- Guest numbers for the Resort hotel go down slightly from June to September, which is also when the prices are highest. Thus, these months should be avoided for bookings.
- Broadly, April to August is the peak season of bookings. Both hotels have the fewest guests during the winter.
- There are four different types of customers namely - Transient, Transient-Party, Group and Contract. Transient customer types have higher cancellations.
- Higher lead time has higher chance of cancellation. Also, history of previous cancellations increases chances of cancellation.
- No deposit cancellations are high compared to other categories but these should not be discouraged per se as bookings in this category are also very high compared to non-refundable type bookings.
- Cancellations are high when done through agents compared to direct bookings. Hotels need to do marketing and give special incentives for direct bookings as these may establish personal one to one relationship promoting customer loyalty.