<div align="center">**Capstone Project Submission**</div>

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

Contributor Roles:

1. **Manish Kumar** – manishkrmaurya27@gmail.com
    1. Pandas Operations.
    2. Data Wrangling.
    3. Data Visualizations.
    4. Booking analysis of resorts and city hotels.
    5. Cancellation Statistics.
    6. Bookings in different countries and their percentage analysis.

2. **Prashant Bhatt** – Prashant.Bhatt072@gmail.com
    1. Pandas Operations.
    2. Data Wranglings.
    3. Data Visualizations.
    4. NaN value handling.
    5. Bookings from different market segments.
    6. Space parking demand analysis.

Complete Jupyter Notebook is prepared by combined Collaboration.

Please paste the GitHub Repo link.

Github Link: https://github.com/Manish-kumar15/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)

Businesses around the world are waking up to the importance of data analytics. Getting the right insights gives you an advantage in a competitive business environment because of following reasons:
- Introduce better hotel pricing strategies
- Managing Hotel Booking Channels
- Demand Forecasting
- Enhancing productivity
- Marketing strategy

In our project we have to do analysis of different parameters which are as follows:

Hotels:
There are only two types of hotels Resort hotel and City hotel

Market Segments:
We have eight unique market segments from where customers are coming i.e. 'Direct', 'Corporate', 'Online TA', 'Offline TA/TO', 'Complementary', 'Groups', 'Undefined', 'Aviation'

Meal:
These are the four types of meal in the given data :
- SC: self-catering (no meals are included).
- BB: bed and breakfast.
- HB: half board.
- FB: full board.

Country:
We have data of 177 countries coded in short form.

Lead Time:
At a hotel, the time taken between when a customer makes a reservation and their actual arrival is called the Lead Time.

Cancellations:
It contains only two values 1 and 0.
- 1: booking has cancelled
- 0: booking is currently active

Stays in weekend nights:
We are getting two unique values for column 'stays in weekend nights' i.e. 1 and 2 , that means someone booked for one weekend night(Saturday or Sunday) and some booked for two weekend nights(Saturday and Sunday).

Stays in week nights:
We are getting five unique values for column 'stays in weekend nights' i.e. 1 to 5 , that means someone booked for at least one week night to maximum five week nights.

Adr:
Average daily rate for individual order.

Car Parking spaces:
Required car parking spaces: this column contains number of car parking spaces required as per customer demand.

Challenges:
- Data Cleaning
- More Data or Better Data
- Understanding the business problem
- Team coordination
- Lack of clarity in data

Overall analysis:
- Comparing the magnitude of booking we see that city hotel has always scored higher booking counts.
- Number of booking start increasing from spring to summer season and then decrease till winter season.
- TA(travel agents) are the major contributors of booking
- The number of cancelled bookings is in the month of August.
- Adr: starts decreasing from month of September and it gets its minimum value in between November and January, after that adr starts increasing and it gets its peak in august.
- BB is most preferred and FB is the least preferred meal.
- Demand of type A room is highest followed by type D room.