

# Capstone Project Submission

## Instructions:

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

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

#### Contributor Roles:

- 1. Pradeep Kumar Yadav: [krpradeep0828@gmail.com](mailto:krpradeep0828@gmail.com)

- 1. Data Wrangling
  - a. Variable/Column(s) exact meaning/clarification
  - b. Modifying the Null data
- 2. EDA
  - a. Heat correlation map results
  - b. Pi chart figure shows
- 2. Data Summary
- 3. Hotel Wise Analysis wrt country
- 4. Distribution channel wise Analysis pie chart
- 5. Distribution channel wise Analysis results/conclusions
- 6. Booking cancellation Analysis
- 7. Time-wise Analysis ADR and busy months
- 8. Optimal stay length for better deals in adr
- 9. Conclusions

- 2. Y Ishwar Rao: [raoji4676@gmail.com](mailto:raoji4676@gmail.com)

- 1. Data Wrangling
  - a. Filling null values with zero.
  - b. Modifying the Null data
- 2. EDA
  - c. Pair Plotting
  - d. Bar Graph figure shows
- 2. Data Summary
- 3. Hotel Wise Analysis wrt country
- 4. Distribution channel wise Analysis pie chart
- 5. Distribution channel wise Analysis results/conclusions
- 6. Booking cancellation Analysis
- 7. Time-wise Analysis ADR and busy months
- 8. Optimal stay length for better deals in adr
- 9. Conclusions

- 3. Ganesh Pramod Patil: [ganeshp746725@gmail.com](mailto:ganeshp746725@gmail.com)

- 1. Data Wrangling
  - a. Filling null values with zero.
  - b. Modifying the Null data
- 2. EDA
  - e. Pair Plotting
  - f. Bar Graph figure shows

2. Data Summary
3. Hotel Wise Analysis wrt country
4. Distribution channel wise Analysis pie chart
5. Distribution channel wise Analysis results/conclusions
6. Booking cancellation Analysis
7. Time-wise Analysis ADR and busy months
8. Optimal stay length for better deals in adr
9. Conclusions

4. Shashank Mishra: [m123shashank@gmail.com](mailto:m123shashank@gmail.com)

1. Data Wrangling
  - c. ETL pipeline (Big data concept)
  - d. Variable/Column(s) exact meaning/clarification
  - e. Modifying the Null data
2. EDA
  - g. Heat correlation map results
  - h. Pair-plotting
3. Data Summary
4. Hotel Wise Analysis wrt country
5. Distribution channel wise Analysis pie chart
6. Distribution channel wise Analysis results/conclusions
7. Booking cancellation Analysis
8. Time-wise Analysis ADR and busy months
9. Optimal stay length for better deals in adr
10. Conclusions

**Please paste the GitHub Repo link.**

Github Link:- [https://github.com/M123shashank/Capstone-1\\_Hotel-Booking-Analysis](https://github.com/M123shashank/Capstone-1_Hotel-Booking-Analysis)  
Google Drive Link :-  
[https://drive.google.com/drive/folders/1glq1XIH9rYfGG91IEIJcpM\\_cxvSi?usp=sharing](https://drive.google.com/drive/folders/1glq1XIH9rYfGG91IEIJcpM_cxvSi?usp=sharing)

**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 industry is any types or forms of business relating to the provision of accommodation in lodging, food and drinks and various types of other services that are interconnected and which are intended for public service, both of which use the lodging facilities or who simply use the services or the production of certain of the hotel.

Hotels offer enormous range of guests' services such as banqueting, conference and fitness, sport and facilities, beauty spas, bars, sophisticated restaurant, casinos, night clubs and casinos. The Hotel sector consists of more than 15% of all the people who worked in the hospitality sector. Hotels falls into a number of different categories which includes the glamorous five-star resort international luxury chains, trendy boutiques, country house, conference, leisure or guest houses.

For the first step, we performed Data wrangling to clean and remove ambiguity in the data (if any). Further, we defined our problem statement and then, set the agenda to work for.

In the second step, we did a proper detailed observation of our data and defined the exact meaning of our variables in the data. Then, we chose the crucial columns required for our data summary. Also, we planned our roadmap like ETL pipeline to reach our goal i.e. solution of problem statement.

During Univariate Analysis, we mainly focused on the choice and attractions of the customers for booking a room. Through which, we came to conclusion wrt preferred meal or room type demand.

During Hotel wise Analysis, we focused upon revenue generation as well booking cancellation issues. Based on this analysis, we answered things like higher booking cancellations rate, etc.

During Distribution channel wise Analysis, as the name suggests, we constrained ourselves to various info related to distribution channels and how to improve them. Similarly, in Booking cancellation analysis, we focused on cancellation reasons and how to tackle them.

During Time wise Analysis, we analyzed the bookings month-wise too. Also, we answered few more questions like optimal stay length, and special requests, etc.

### ***Few Conclusions:***

1. The majority of reservations are for city hotels.
2. Most of the bookings either in the canceled or checkout done by online TA.
3. More visitors are from western europe, namely Portugal, France, Great Britain, and Spain being the highest.
4. August and July are the most profitable and busiest months for both the hotels.
5. Data suggests that hotel business could be improved by targeting working travelers or improving daily rates for weekdays.