

Capstone Project Submission

Instructions:

For this EDA project, I will use the “Hotel booking Analysis” dataset. Hotel industry is a very volatile industry and the bookings depend on variety of factors such as type of hotels, seasonality, days of week and many more. 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. This makes analysing the patterns available in the past data more important to help the hotels plan better. Using the historical data, hotels can perform various campaigns to boost the business.

Team Member's Name, Email and Contribution:

1. NAME- **VIKAS KUMAR MANJHI**
Email- vmanjhi2@gmail.com
Contribution:-
 1. Make data frame of hotel booking analysis
 2. Hotel wise analysis
 3. Checked-in and cancelled
 4. Booking across different year and month
 5. Visitor analysis
 6. Market segment analysis
 7. Distribution channel analysis
2. NAME- **VINEET KUMAR**
Email- binny01.kumar@gmail.com
Contribution:-
 1. Make data frame of hotel booking analysis
 2. Analysis of hotel
 3. Removing null value
 4. Distribution channel analysis
 5. Country wise data analysis
 6. ADR analysis
3. NAME- **ANKIT UPADHYAY**
Email- upadhyayankit2003@gmail.com
Contribution:-
 1. Making data frame of hotel booking analysis.
 2. Identifying type of customers visiting.
 3. Preparing Technical document of Hotel Booking analysis.
 4. Predict the cancellation of bookings based on factors present in the data
 5. Hotel booking analysis

4. NAME- **ANKIT SHARMA**
[Email- connectwithankitsharma@gmail.com](mailto:connectwithankitsharma@gmail.com)

Contribution:-

1. Analysis of Dataframe
2. Removing Duplicate rows
3. Booking analysis
4. Deposit policies of hotel
5. Overview of guest

5. NAME- **RAGHVENDRA SINGH**
Email- raghvendra.rbmi197@gmail.com

Contribution:-

1. Analysis of Dataframe
2. Market segment analysis
3. Meal preference analysis
4. Heatmap correlation
5. Country wise data analysis

Please paste the GitHub Repo link.

Github Link:- <https://github.com/vikasmanjhi/HOTEL-BOOKING-ANALYSIS.git>

Drive link:- <https://drive.google.com/drive/folders/1yqf-d5uSicECi7Ek-gHrxt4r6GhnEEwp?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)

Conclusion:-

- (1) Around 66.40% bookings are for City hotel and 33.60% bookings are for Resort hotel, therefore City Hotel is busier than Resort hotel. Also the overall adr of City hotel is slightly higher than Resort hotel.
- (2) Mostly guests stay for less than 5 days in hotel and for longer stays Resort hotel is preferred. (3) Both hotels have significantly higher booking cancellation rates and very few guests less than 3 % return for another booking in City hotel. 5% guests return for stay in Resort hotel.
- (4) Most of the guests came from European countries, with most of guests coming from Portugal. (5) Guests use different channels for making bookings out of which most preferred way is TA/TO. (6) For hotels higher adr deals come via GDS channel, so hotels should increase their popularity on this channel.
- (7) Almost 30% of bookings via TA/TO are cancelled.
- (8) Not getting same room as reserved, longer lead time and waiting time do not affect cancellation of bookings. Although different room allotment do lowers the adr.
- (9) July- August are the most busier and profitable months for both of hotels.
- (10) Within a month, adr gradually increases as month ends, with small sudden rise on weekends.
- (11) Couples are the most common guests for hotels, hence hotels can plan services according to couples needs to increase revenue.
- (12) More number of people in guests results in more number of special requests.
- (13) Bookings made via complementary market segment and adults have on average high no. of special request.
- (14) For customers, generally the longer stays (more than 15 days) can result in better deals in terms of low adr.

Challenges:-

- (1) There was a lot of duplicate data.
- (2) Data was present in wrong datatype format.
- (3) Choosing appropriate visualization techniques to use was difficult.
- (4) A lot of null values were there in the dataset.