HOTEL BOOKING ANALYSIS

Renuka Jadhav

Data Science Trainee,
Alma better.

• **ABSTRACT:** Hotel booking dataset contains booking information for city and resort hotel. Both datasets share the same structure, with 32 columns with 119390 entries describing Each observation represents a hotel booking. Both datasets contains bookings data of arrive between the year 2015 and 2017.In this booking data some bookings were cancelled. Due to the scarcity of real business data for scientific and educational purposes, these datasets can have an important role for research and education in revenue management,

PROBLEM STATEMENT: We are here to explore a hotel booking dataset which contain booking information for a city hotel and a resort hotel. We will analyze some important aspects of hotel helpful to run profitable hotel business are as follows:

•

• The time of year to book a hotel room?
• Optimal length of stay to get the best daily rate?

• **INTRODUCTION:** In tourism and travel related industries, most of the research on Revenue Management demand forecasting . hotel datasets withdemand data are shared to help in overcoming this limitation. The datasets now made available were collected aiming at the developmentof prediction models to classify a hotel booking's . And the goal here is to get

some useful insights and find correlation between respective columns.

Data set:
1. hotel: type of hotels.

2. is_canceled: canceled or not.

3. lead_time: no. of days before actual arrival in the hotel.

4. arrival_date_year: year of booking.

5. arrival_date_month: month of booking.

6. arrival_date_week_number: week number of the year in which booking.

7.arrival_date_day_of_month: arrival month date **.**

8. stays_in_weekend_nights: no. of weekends guest stayed .

9. stays_in_week_nights: no. of weekdays guest stayed.

 10. adults

11. children

12. babies

13. meal: BB – Bed & Breakfast

        HB – only two meals including breakfast meal

        FB – breakfast, lunch, and dinner

14. country

15. market_segment: TA: Travel agents.

               TO: Tour operators

16. distribution_channel

17. is_repeated_guest

18. previous_cancellations: cancellation in past **.**

19. previous_bookings_not_canceled: not cancelled in past **.**

20. reserved_room_type

21. assigned_room_type

22. booking_changes

23. deposit_type

24. agent

25. company

26. days_in_waiting_list

27. customer_type

28. adr: average daily rate **.**

29. required_car_parking_spaces

30. total_of_special_requests

31. reservation_status
32. reservation_status_date

• Steps Involved:

- **Data study:**

  - Firstly, we need to understand the dataset to make any analysis on dataset.

  - So our first step should be understand the columns and check the basic correlation between columns .

  - After that we need to imported libraries and dataset, some of the libraries used are NumPy, pandas, matplotlib, seaborn, warnings. Once the data is collected, process of analysis begins
  - **Clean the data set**
  - **Checking and deleting the duplicate values.**
  - **Analyze the data.**

    DATA PREPROCESSING: A dataset may contain noise, missing values, and inconsistent data, thus, pre- processing of data is essential to improve the quality of data and time required in the data mining.

    - **CLEANING AND MANIPULATING THE DATASET:**

CLEANING:
After completing the Data Sourcing, the next step in the process of EDA is Data Cleaning. It is very important to get rid of the irregularities and clean the data after sourcing it into our system. Irregularities are of different types of data.

MANIPULATING:

Data Manipulation: Manipulation of data is the process of manipulating or changing information to make it more organized and readable. Made some new features with the help of column present in the datasets .

CONCLUSION:

Our analysis, would be capable of helping prospective guests in choosing the right hotel, right stay duration and much more for their stay and moreover, would also be introspecting for hotel management in bringing out changes in their services for the guests.

- City Hotel is the most booked hotel .
- Resort Hotel has been preferred over City Hotel by larger group of guests or families.
- Direct bookings have very less cancellation%.
- Most preferred meal is BB (Bread and Breakfast.
- Online marketing is the best way to attract customers.
- People do not want to pre-deposit the money for booking.
- Only 10% of people require parking space.
- Resort hotel is preferred mostly for longer stay, day time stays. and when the parking space is needed.

-

stay and moreover, would also be introspecting for hotel management in bringing out changes in their services for the guests.

- City Hotel is the most booked hotel with 62 percent not_canceled bookings.
- Resort Hotel has been preferred over City Hotel by larger group of guests or families.
- One out of every three bookings are cancelled.
- Direct bookings have very less cancellation%.
- Most preferred meal is BB (Bread and Breakfast.
- Online marketing is the best way to attract customers.
- People do not want to pre-deposit the money for booking.
- Only 10% of people require parking space.
- Resort hotel is preferred mostly for longer stay, day time stays. and when the parking space is needed.
- More than 15 days advance bookings have high chances of cancellation.
- Assigning different room is not a reason for cancellation.

As for the prediction of cancellations concerns. it is clear that better results can be achieved in a more