Capstone Project Submission

Hotel Booking Analysis

Hotel bookings are everywhere in Online service. They are easy to book and can be lucrative, because of these two factors, more and more Bookings are being developed. In this notebook, we will do a comprehensive analysis of the Hotel bookings market by comparing over thousand Bookings in Online services across different categories

The Hotel bookings dataset consists of enormous data that can be used to create effective insights. There are various key factors that play a major role in the success & engagement from the user’s end. Our problem statement is quite inevitable in comparison with the present Hotel Bookings store App market. Upon doing several pieces of research it can be seen that every day around 3000+ Hotel bookings is being added to the Bookings library. Therefore, enormous datasets & variety of insights can be concluded for business improvements.

* We are given a hotel bookings dataset. This dataset contains booking information for a city hotel and a resort hotel. It contains the following features.
  + -hotel: Name of hotel ( City or Resort)
  + is\_canceled: Whether the booking is canceled or not (0 for no canceled and 1 for canceled)
  + lead\_time: time (in days) between booking transaction and actual arrival.
  + arrival\_date\_year: Year of arrival
  + arrival\_date\_month: month of arrival
  + arrival\_date\_week\_number: week number of arrival date.
  + arrival\_date\_day\_of\_month: Day of month of arrival date
  + stays\_in\_weekend\_nights: No. of weekend nights spent in a hotel
  + stays\_in\_week\_nights: No. of weeknights spent in a hotel
  + adults: No. of adults in single booking record.
  + children: No. of children in single booking record.
  + babies: No. of babies in single booking record.
  + meal: Type of meal chosen
  + country: Country of origin of customers (as mentioned by them)
  + market\_segment: What segment via booking was made and for what purpose.
  + distribution\_channel: Via which medium booking was made.
  + is\_repeated\_guest: Whether the customer has made any booking before(0 for No and 1 for Yes)
  + previous\_cancellations: No. of previous canceled bookings.
  + previous\_bookings\_not\_canceled: No. of previous non-canceled bookings.
  + reserved\_room\_type: Room type reserved by a customer.
  + assigned\_room\_type: Room type assigned to the customer.
  + booking\_changes: No. of booking changes done by customers
  + deposit\_type: Type of deposit at the time of making a booking (No deposit/ Refundable/ No refund)
  + agent: Id of agent for booking
  + company: Id of the company making a booking
  + days\_in\_waiting\_list: No. of days on waiting list.
  + customer\_type: Type of customer(Transient, Group, etc.)
  + adr: Average Daily rate.
  + required\_car\_parking\_spaces: No. of car parking asked in booking
  + total\_of\_special\_requests: total no. of special request.
  + reservation\_status: Whether a customer has checked out or canceled,or not showed
  + reservation\_status\_date: Date of making reservation status.
* Total number of rows in data: 119390
* Total number of columns: 32
* On This Data set we have perform the following steps to get desirable Analysis results
* **Data Cleaning and Feature Engineering**
  1. Removing Duplicate rows
  2. Handling null values
  3. Converting columns to appropriate data types
  4. Removing outliers
  5. Creating new columns

**Exploratory Data Analysis**

* Performed EDA and tried answering the following questions:

Q1)Which agent makes most no. of bookings?

Q2)Which room type is in most demand and which room type generates highest adr?

Q3) Which meal type is most preffered meal of customers?

Q4) What is percentage of bookings in each hotel?

Q5) which hotel seems to make more revenue?

Q6)Which hotel has higher lead time?

Q7) What is preferred stay length in each hotel?

Q8) Which hotel has longer waiting time?

Q9) Which hotel has higher bookings cancellation rate?

Q10) Which hotel has high chance that its customer will return for another stay?

Q11) Which is the most common channel for booking hotels?

Q12) Which channel is mostly used for early booking of hotels?

Q13) Which channel has longer average waiting time?

Q14) Which distribution channel brings better revenue generating deals for hotels?

Q15)From where the most guests are coming ?

Q16)How long do people stay at the hotels?

**Contributor Roles**

1. **Rakesh Sahoo (** rakesh8580@gmail.com**)**
   * + Upload dataset to Google colab and explain dataset to team members.
     + Analyze null values and filter them.
     + Data cleaning.
     + Correction of data types
     + Data wrangling
     + Data Visualizations
     + Technical Write up
     + PowerPoint presentation
     + Project summary
   1. **Mahesh Patki** ([mpatki479@gmail.com](mailto:mpatki479@gmail.com))
      * Upload dataset to Google colab and explain dataset to team members.
      * Analyze null values and filter them.
      * Data cleaning.
      * Correction of data types
      * Data wrangling
      * Data Visualizations
      * Technical Write up
      * PowerPoint presentation
      * Project summary
   2. **Lubna Zarin**( [lubnazarin2001@gmail.com](mailto:lubnazarin2001@gmail.com))
      * Upload dataset to Google colab and explain dataset to team members.
      * Analyze null values and filter them.
      * Data cleaning.
      * Correction of data types
      * Data wrangling
      * Data Visualizations
      * Technical Write up
      * PowerPoint presentation
      * Project summary

**Please paste the GitHub Repo link:**

[**https://github.com/sahoo-rakesh/Hotal\_booking\_analysis**](https://github.com/sahoo-rakesh/Hotal_booking_analysis)