Capstone Project Submission

Github Link:
https://github.com/RAHULRAJ666DX/Capston-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)

This project helps by analyzing all characteristic data and help the business owner how to cope up with business requirements and how to better the service to increase the value of the business and efficiency

We have been provided a dataset of the hotel industry of the region in which our startup aims to begin its initial operations. Our primary goal is to understand the market of this particular region and try to use our data analysis techniques to draw out key features of the market. Our secondary goal is to draw out actionable insights from our analysis and give conclusions about key aspects of the market such as cancellation rate, distribution channels, and alike

Based upon the initial assessment we found that the data was pretty much clean except for some missing values in a few columns. Upon using the info() method, we draw out the following key insights about the data

The dataset has a shape of (119390, 32) which means that it contains approximately 1.2 lakh rows and 32 columns.

Our Dataset has 4 columns with float64 dtype, 16 columns with int64 dtype, and 12 columns with object dtype.In our Dataset, we observed null values in the following columns, null values in the children column, 488 null values in the country column, 16,340 null values in the agent column 112,593 null values in the company column

1. Understanding the need for EDA to solve the problem:
This is the most important step of the analysis because we cannot solve a problem if we don't know what is the need for the solution and in this step we study what is the need of doing a EDA analysis for Hotel Booking analysis
To get ideas many creative thinking techniques such as brainstorming is done

2. Studying the Data and brainstorming how to work on it:
In this step the dataset provided by Almabetter is studied thoroughly and we make some assumptions based several selection criteria and what to be done on the data
various required questions are written down from a customer and business perspective and look for answers to those questions

3. Cleaning and preparing the Data:
For working on the data the data is to be cleaned and unwanted null values has to be removed or replaced with meaningful data
There of 4 columns present in the dataset with null values we have removed two columns with the most null values and for the other columns we replaced the Nan values with mode and new data element
So we have 31994 duplicate rows in our data. So we will drop the duplicate rows from our dataset Dropping duplicates since we cannot derive meaningful insights.

4. **EDA:**

1. City hotels were more proffered by the guests because it is more easily accessible.

2. Portugal country has most guests with 31.41% followed by Britain=11.93 & France=10.11

3. 27.48% of total bookings were canceled which is 24025 cancellations

4. City hotels have 16049 cancellations which is more then Resort hotels which had 7976

5. Type 'A' room was booked 64.70% with 56552 bookings followed by type 'D' with 17398 bookings and type E with 6049 bookings Room type 'L' and 'P' is booked very few times with only 6 booking

6. Most of the Guests are using 'Online TA' which is 59.06% with 51618 bookings Followed by 'Offline TA' and 'direct' are almost same at 15.89% and 13.5%

7. Most of the ADR is contributed by GDS and TA and Direct reservations

8. Set Non-refundable Rates, Collect deposits, and implement more rigid cancellation policies.

8.1 98.68% of the guests did not make a Deposit and very few made Non refundable deposits

9. 77.31% of the guests order Bed & Breakfast and around 10% of the guests select Half Board and 10% don't opt for meal packages

10. Only 3.19% of the guests were repeated guests with only 3810 of total guests

10.1. Resort hotels had more repeated guests then City hotels

11. 75.05% of the guests are Transient guests and 21.04% are Transient-Party guests

12. August month had most travelers with 12.88% which is 13877 guests

12.1. July and May has the next most travelers with 10057 and 8355 guests

13. 50.26% of total guests stay in Hotels during Weekend

13.1. Most of the guests stay 2 days or 1 day during Weekend

14. 25.3% of the guests stay only 1 day during Weekdays and 23.73% guests stay for 2 days and 18% stay 3 days during Weekdays, Most of the guests stay 1 day Very negligible percentage of guests stay above 3 days

15. 19% of the bookings made were for Single Rooms

16. Guests who has not canceled their previous booking has more probability to be a Repeated Guest

Contributor Roles: Rahul Raj Individual Project

Github Link:

https://github.com/RAHULRAJ666DX/Capston-Hotel-Booking-Analysis-

Drive Link:
https://drive.google.com/drive/folders/1bFw9JAsQ66F6n5BdJBtJtGR2HXuesfaW?usp=share_link

Please paste the drive link to your deliverables folder. Ensure that this folder consists of the project Colab notebook, project presentation and video.