**INSY-5339-003-PRIN OF BUSINESS DATA MINING**

**PROJECT TOPIC PROPOSAL**

**UNIVERSITY OF TEXAS AT ARLINGTON**

**SPRING 2022**

RUCHI SHUKLA

PRANATI CHAUHAN

JESSICA MENDEM

TEAM 09

SUBMITTED ON: 3/24/22

**Topic** – Hotel Booking Demand

**Data Source -** <https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand>

**Overview -** Our team is interested in analyzing the hotel industry and that’s the business domain we want to go forward with for our Data Mining Project. Running a hotel business is difficult. There are numerous risks to be aware of. However, one can always attempt to anticipate and plan for future events. Hotels should aim to predict customers' future preferences in a systematic way and coordinate their operations accordingly. Demand forecasting helps to predict future scenarios that hotel industry generally faces.

In today's world, hotel demand forecasting is critical. Most hotels struggle to analyse and predict their customers' behavior. Customers frequently modify their booking criteria, such as the number of people staying increases or decreases, the length of stay, or room type preferences.

This dataset comprises booking information for two hotels- a city hotel and a resort hotel, including dates of booking, duration of stay, number of adults, children, and/or newborns, and available parking space, etc.

**Business Problem - We'll do our best to respond to the following questions\*.**

1. How many reservations were canceled?
2. What is the difference between a Resort Hotel and a City Hotel in terms of bookings?
3. What is the booking percentage for each year?
4. Which month is the busiest for hotels?
5. Which country has the most visitors?
6. What is the average length of stay in a hotel?
7. Which lodging type (single, couple, or family) was the most popular?
8. Which guests tend to cancel booking the most- single, couple or a family?
9. And create a predictive model to predict whether the booking will be canceled in the future.

\*If needed, additional data (datasets with hotel booking and cancellation) can be investigated for this study.

**Variable Description –**

|  |  |  |
| --- | --- | --- |
| Variable | Description | Type |
| is\_canceled | If booking was cancelled | Categorical |
| lead\_time | difference between booking date & arrival date | Integer |
| arrival\_date\_year | arrival date of the year | Integer |
| arrival\_date\_month | arrival month - Jan to Dec | Categorical |
| arrival\_date\_week\_number | week of arrival | Integer |
| arrival\_date\_day\_of\_month | date and day of aarival | Integer |
| stays\_in\_weekend\_nights | number of nights stayed on weekends | Integer |
| stays\_in\_week\_nights | number of nights stayed in a week | Integer |
| Adults | Number of adults | Integer |
| Children | Number of children | Integer |
| Babies | Number of babies | Integer |
| Meal | type of meal booked | Categorical |
| Country | country of origin of the guest | Categorical |
| market\_segment | source of booking - self or through travel agent or other sources | Categorical |
| distribution\_channel | source of booking - self or through travel agent or other sources | Categorical |
| is\_repeated\_guest | If the guest | Categorical |
| previous\_cancellations | bookings cancelled by the customer at the same hotel before the current booking | Integer |
| previous\_bookings\_not\_canceled | number of booking not cancelled by guest, prior to current booking | Integer |
| reserved\_room\_type | which type of room was booked by the customer | Categorical |
| assigned\_room\_type | which room was assigned | Categorical |
| booking\_changes | count of changes made to the reservation ahead of check-in | Integer |
| deposit\_type | Category addressing whether a deposit was made during reservation or not | Categorical |
| Agent | travel agency identifier/code | Categorical |
| Company | company identifier/code | Categorical |
| days\_in\_waiting\_list | measurement of time when booking was n waitlist | Integer |
| customer\_type | type of cutomer | Categorical |
| Adr | Average daily rates at the time of booking | Numeric |
| required\_car\_parking\_spaces | number of parking spaces requested in the booking | Integer |
| total\_of\_special\_requests | count of special requests associated with the reservation | Integer |
| reservation\_status | status of reservation at the end of stay | Categorical |
| reservation\_status\_date | date of reservation\_status | Date |