**Predicting Cancellations of Hotel Bookings**

**Abstract**

This comprehensive dataset provides a wealth of information on bookings made at a city hotel and a resort hotel, offering valuable insights for both travelers and hotel operators. With data points including booking dates, length of stay, and party size, as well as details on parking availability and special requests, this dataset enables in-depth analysis of hotel booking trends. Whether you are looking to identify the most advantageous time of year to book a room, determine the optimal length of stay for the best daily rate, or predict the likelihood of receiving special requests, this dataset is an essential resource. With its diverse range of information, this data set has the potential to transform our understanding and analysis of hotel bookings, providing valuable insights for anyone interested in the hospitality industry.

**Data Source and Description**

We have taken the data source from Kaggle (<https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand>). The data set booking information for a city hotel and resort hotel and includes information such as when the booking was made, length of stay, number of adults, children, babies and the number of available parking spaces, among other things.

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Attribute** | **Description** |
| 1 | hotel | Hotel (H1 = Resort Hotel or H2 = City Hotel) |
| 2 | is\_canceled | Value indicating if the booking was canceled (1) or not (0) |
| 3 | lead\_time | Number of days that elapsed between the entering date of the booking into the PMS and the arrival date |
| 4 | arrival\_date\_year | Year of arrival date |
| 5 | arrival\_date\_month | Month of arrival date |
| 6 | arrival\_date\_week\_number | Week number of year for arrival date |
| 7 | arrival\_date\_day\_of\_month | Day of arrival date |
| 8 | stays\_in\_weekend\_nights | Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel |
| 9 | stays\_in\_week\_nights | Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel |
| 10 | adults | Number of adults |
| 11 | children | Number of children |
| 12 | babies | Number of babies |
| 13 | meal | Type of meal booked. Categories are presented in standard hospitality meal packages: Undefined/SC – no meal |
| 14 | country | Country of origin. Categories are represented in the ISO 3155–3:2013 format |
| 15 | market\_segment | Market segment designation. In categories, the term “TA” means “Travel Agents” and “TO” means “Tour Operators” |
| 16 | distribution\_channel | Booking distribution channel. The term “TA” means “Travel Agents” and “TO” means “Tour Operators” |
| 17 | is\_repeated\_guest | Value indicating if the booking name was from a repeated guest (1) or not (0) |
| 18 | previous\_cancellations | Number of previous bookings that were cancelled by the customer prior to the current booking |
| 19 | previous\_bookings\_not\_canceled | Number of previous bookings not cancelled by the customer prior to the current booking |
| 20 | reserved\_room\_type | Code of room type reserved. Code is presented instead of designation for anonymity reasons. |
| 21 | assigned\_room\_type | Code for the type of room assigned to the booking. Sometimes the assigned room type differs from the reserved room type due to hotel operation reasons |
| 22 | booking\_changes | Number of changes/amendments made to the booking from the moment the booking was entered on the PMS |
| 23 | deposit\_type | Indication on if the customer made a deposit to guarantee the booking. This variable can assume three categories: No |
| 24 | agent | ID of the travel agency that made the booking |
| 25 | company | ID of the company/entity that made the booking or responsible for paying the booking. |
| 26 | days\_in\_waiting\_list | Number of days the booking was in the waiting list before it was confirmed to the customer |
| 27 | customer\_type | Type of booking, assuming one of four categories: Transient, Transient-Party, Contract, Other |
| 28 | adr | Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights |
| 29 | required\_car\_parking\_spaces | Number of car parking spaces required by the customer |
| 30 | total\_of\_special\_requests | Number of special requests made by the customer (e.g. twin bed or high floor) |
| 31 | reservation\_status | Reservation last status, assuming one of three categories: Check-Out, Canceled, Other |
| 32 | reservation\_status\_date | Date at which the last status was set. |

**Goal of the Project:**

The goal of the project is to predict whether the guest will actually come or cancel his booking. This can help the hotel plan personal and food requirements using Classification and we will perform . We hoped that the analysis with the data exploration and model building will answer the following questions:

Data Exploration – We will use summary statistics and visualization tools to examine the following questions:

* Where do the guests come from?
* How much do guests pay for a room per night?
* How does the price per night vary over the year?
* Which are the most busy month?
* How long do people stay at the hotels?
* Bookings by market segment
* How many bookings were canceled?
* Which month have the highest number of cancelations?

Prediction Model – We will apply different classification models to predict the severity of the severity level of the accident.

1. Data Understanding and Preparation: In this context, we can perform data cleaning, exploring, normalization, removing/fixing missing values, feature engineering, etc.
2. Modelling: We will utilize classification algorithms to anticipate the cancellation of reservations (indicated by the column "is\_canceled") by utilizing significant attributes extracted through data exploration.
3. Validation: Validate the model performance on the test data, and generate performance metrics.
4. Conclusion: What insights can we gain from the Hotel bookings cancellation model? And how can we use the model to assist hotel owners in predicting booking cancellations?