# Introduction:

Hotel Reservations Cancellation

Analysis

The online hotel reservation channels have dramatically changed booking possibilities and customers’ behavior. A significant number of hotel reservations are called-off due to cancellations or no-shows. The typical reasons for cancellations include change of plans, scheduling conflicts, etc. This is often made easier by the option to do so free of charge or preferably at a low cost which is beneficial to hotel guests but it is a less desirable and possibly revenue-diminishing factor for hotels to deal with.

In this project, we will delve into the domain of Hotel Reservation Cancelation classification, focusing on a binary classification problem with the goal of predicting whether an Customer is Cancelling Or Not Cancelling

#### **Task: The task is to predict the churn in hotel reservation as follows**:

1. **Binary Classification**: The primary objective of this project is to develop a model that Customer Is Churning or Not Churning .This binary classification task simplifies the prediction of Churning.
2. **Feature Set**: The features used for classification encompass a variety of **Adults , Children’s , Meal plan , Lead time ,Average Price** and more. The selection of these features is driven by the belief that they hold valuable information for predicting Churn.
3. **Dataset Source**: The dataset used in this project is sourced from Kaggle. This dataset has been curated to serve as the foundation of our analysis and modeling efforts. It offers a real-world representation of income-related factors, making our project both practical and applicable.
4. **Modeling Techniques**: Throughout this project, we will explore and implement various deep learning techniques and algorithms to construct a robust predictive model. The selection of methods will be influenced by our pursuit of identifying the most effective approach for the Churning task. Deep learning, with its ability to capture complex patterns and relationships in data, holds the potential to offer valuable insights into Churning task, and we aim to harness this power in our analysis.
5. **Data Preprocessing and Analysis**: Data preprocessing will involve tasks such as handling missing values, encoding categorical variables, and scaling features. We will also perform data analysis to gain insights into the dataset's characteristics and distributions.
6. **Performance Evaluation**: We will evaluate the model's performance using standard binary classification metrics, including Accuracy, F1 score, Precision and Recall, and Confusion matrix. This will enable us to assess the model's effectiveness in predicting income levels.

# Data Source:

The dataset was collected from Kaggle, a popular platform for data science and machine learning datasets.

**Dataset name:** Hotel Reservations **Dataset link:** [Hotel Reservations Dataset (kaggle.com)](https://www.kaggle.com/datasets/ahsan81/hotel-reservations-classification-dataset)

**Data Description:**

* Totally there are 17 input features and target class label in the dataset:

**Booking Status:** Cancelled or Not Cancelled

**no \_of\_adults**: Number of adults

**no\_of\_children**: Number of Children

**no\_of\_weekend\_nights**: Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel

**no\_of\_week\_nights**: Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel

**type\_of\_meal\_plan**: Type of meal plan booked by the customer:

**required\_car\_parking\_space**: Does the customer require a car parking space? (0 - No, 1- Yes)

**room\_type\_reserved**: Type of room reserved by the customer. The values are ciphered (encoded) by INN Hotels.

**lead\_time**: Number of days between the date of booking and the arrival date

**arrival\_year**: Year of arrival date

**arrival\_month**: Month of arrival date

**arrival\_date**: Date of the month

**market\_segment\_type**: Market segment designation.

**repeated\_guest**: Is the customer a repeated guest? (0 - No, 1- Yes)

**no\_of\_previous\_cancellations**: Number of previous bookings that were canceled by the customer prior to the current booking

**no\_of\_previous\_bookings\_not\_canceled**: Number of previous bookings not canceled by the customer prior to the current booking

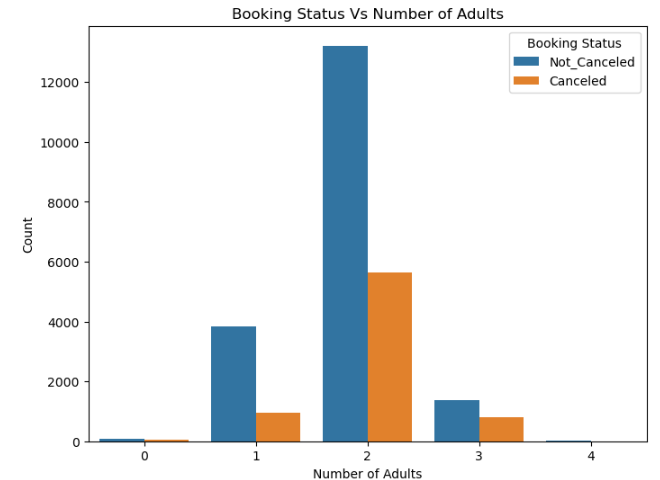
**avg\_price\_per\_room**: Average price per day of the reservation; prices of the rooms are dynamic. (in euros)

**no\_of\_special\_requests**: Total number of special requests made by the customer (e.g. high floor, view from the room, etc.)

# Exploratory Data Analysis (EDA):

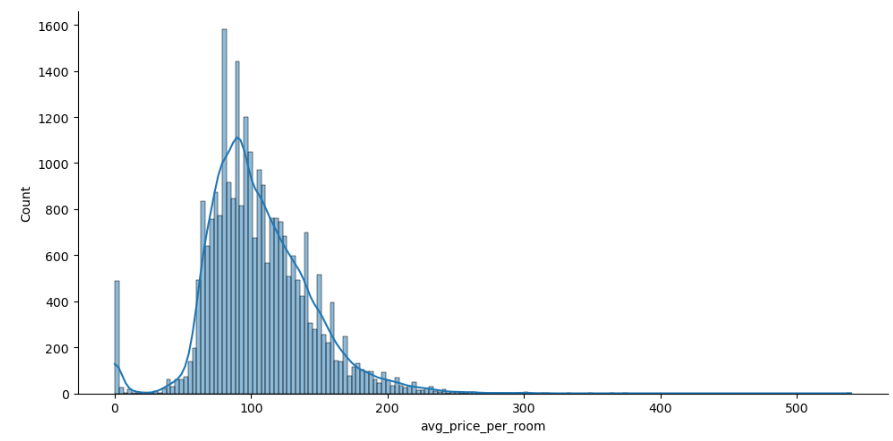
## Booking Status feature:

* The Target feature (Income) is having two classes
* Cancelled or Not Cancelled
* 71% are Not Cancelling 29% are Cancelling.



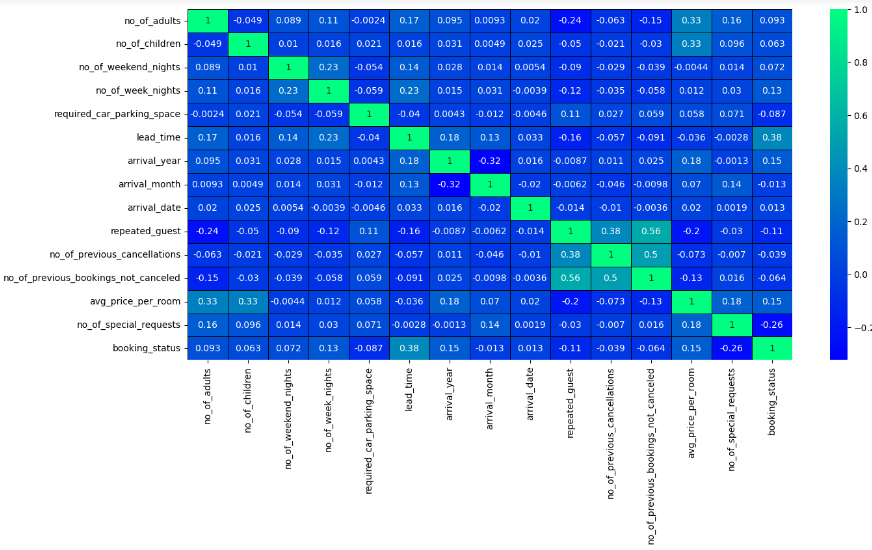
**Booking Status-Adults Correlation:**

* Here Mostly two adults are Cancelling and as well as Not Cancelling

****

**Average Price:**

* Here Average Price per Room (Euros) is normally Distributed

****

**Insights and Observations:**

* Lead time highly correlated with Booking status
* Prices, Adults and Week Nights, Arrival Year, Month are also correlated with Booking Status.
* Car parking space, Arrival month, Special requests, features are not effecting income feature.