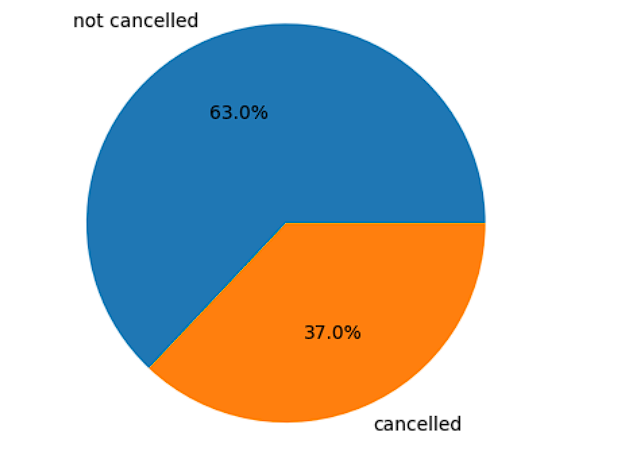
# **Project Report: Hotel Booking Demand Analysis & Insights Generation**

**TASK 1: Data Overview:**

* There are 33 columns in the Hotel Booking dataset.
* Some columns such as children, agent, and company had missing values.
* The column company had 94% missing values which is not useful in our analysis.
* Columns like hotel, is\_canceled, customer\_type, etc., had low cardinality and are good candidates for categorical analysis.

**TASK 2 : Reservation Cancellation Status:**

The analysis revealed that out of the total hotel reservations, approximately 37% were canceled, while around 63% were not canceled. Specifically, there were about 44,140 canceled bookings and 75,166 confirmed (non-canceled) bookings, indicating that the majority of guests followed through with their reservations.



**TASK 3:** **Cancellation Percentages by Hotel Type:**

| **Hotel Type** | **% Canceled** |
| --- | --- |
| City Hotel | 41% |
| Resort Hotel | 27% |

The cancellation analysis by hotel type revealed that City Hotels experienced a higher cancellation rate compared to Resort Hotels. Approximately 41% of the reservations made at City Hotels were canceled, whereas only about 27% of the bookings at Resort Hotels were canceled. This indicates that guests were more likely to cancel their reservations at City Hotels than at Resort Hotels.

**TASK 4:** **Filter for Non-Cancelled Reservations:**

A new DataFrame df\_checkout was created to include only non-canceled bookings (is\_canceled == 0). This subset was used for subsequent analysis.

**TASK 5: Monthly Reservation Analysis:**



**A screenshot of a calendar

AI-generated content may be incorrect.**

The analysis of monthly reservation trends highlighted that the peak month for City Hotels was August, while for Resort Hotels, it was July. This slight variation in peak months suggests that the two hotel types experience different seasonal booking patterns, reflecting distinct preferences and travel behaviours of their guests.

A graph of a line graph

AI-generated content may be incorrect.

**TASK 6: Arrival Date Column Creation**

A new arrival\_date column was created in yyyy-mm-dd format and converted to a **datetime** object. This enabled time-based analysis.

**TASK 7: Daily and Weekly Reservation Trends:**

The analysis of daily and weekly reservation trends revealed meaningful booking patterns over time. By aggregating the total number of reservations per day, we observed how booking volumes fluctuated on a day-to-day basis. Furthermore, by computing the average number of daily reservations for each week of the year using ISO week numbers, we identified broader trends in weekly booking behavior. These insights help understand guest demand cycles, even though visualizations were not included in this analysis.

**TASK 8: Average Daily Rate (ADR) Analysis**



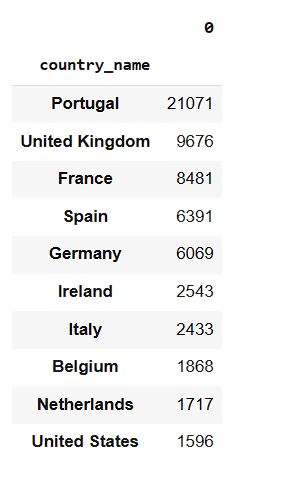
* Highest ADR in City Hotel: Transient customers (110.42)
* Highest ADR in Resort Hotel: Transient customers (96.00)
* Lowest ADR in City Hotel: Group customers (87.40)
* Lowest ADR in Resort Hotel: Transient-Party customers (77.20)
* City Hotels consistently have higher ADRs across all customer types compared to Resort Hotels.
* Contract customers in City Hotels also pay significantly more (108.93) than those in Resort Hotels (78.58).

A graph of different colored squares

AI-generated content may be incorrect.

**TASK 9 : Top Countries by Reservations**

* Merged df\_checkout with a countries code dataset.
* Top 10 countries by reservation count :



A graph of blue bars with white text

AI-generated content may be incorrect.

**TASK 10: Guest Analysis**

* Average guests per reservation: 2
* Maximum guests in a single reservation: 55