

# Project Report: Hotel Booking Data Analysis

## Introduction:

The dataset used for this analysis contains information about hotel bookings, including various features such as booking details, guest demographics, and reservation status. The goal of this project is to perform data analysis, clean the dataset, and derive insights to help understand trends in hotel bookings.

## Libraries Used

- pandas
- matplotlib
- seaborn
- warnings

## Dataset

The dataset used for this analysis is loaded from a CSV file named "hotel\_bookings 2.csv." It contains 119,390 rows and 32 columns.

## Data Analysis and Data Cleaning

### Initial Exploration

Displayed the first 5 and last 5 rows of the dataset using `Hotel.head()` and `Hotel.tail()`.

Checked the shape of the dataset using `Hotel.shape`.

Displayed the column names using `Hotel.columns`.

Obtained information about the dataset using `Hotel.info()`.

## **Data Cleaning**

Converted the 'reservation\_status\_date' column to datetime format using `pd.to_datetime`.

Checked for missing values using `Hotel.isnull().sum()` and handled missing values appropriately.

Dropped unnecessary columns 'company' and 'agent' using `Hotel.drop(columns_to_drop, axis=1, inplace=True)`.

Removed outliers in the 'adr' column by plotting a box plot and filtering values below 5000.

## **Data Analysis and Visualization**

### **Reservation Status Overview**

Checked the proportion of canceled and not canceled reservations.

Plotted a bar chart to visualize the distribution of reservation status.

### **Hotel-wise Cancellation Analysis**

Visualized the number of reservations for each hotel and their cancellation status.

Analyzed the cancellation proportions for Resort Hotel and City Hotel.

### **Average Daily Rate (ADR) Analysis**

Calculated the average daily rate for Resort Hotel and City Hotel.

Plotted the ADR trends over time for both hotels.

### **Monthly Reservation Analysis**

Analyzed the number of reservations and cancellations per month.

Plotted a countplot to visualize the distribution of reservation status per month.

### **Top 10 Countries with Canceled Reservations**

Identified and visualized the top 10 countries with the highest number of canceled reservations.

### **Market Segment Analysis**

Explored the distribution of market segments in the dataset.

Analyzed the cancellation proportions for each market segment.

### **Average Daily Rate (ADR) Trends Over Time**

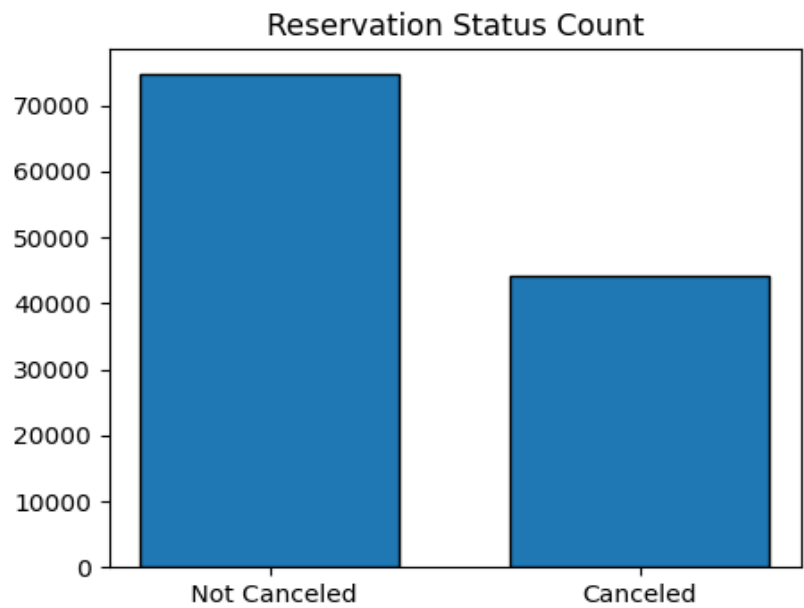
Plotted the ADR trends for canceled and not canceled reservations over time.

### **Conclusion**

This project provided valuable insights into hotel booking trends, cancellation rates, and average daily rates. The analysis can be used to make informed decisions and improvements in hotel management and marketing strategies.

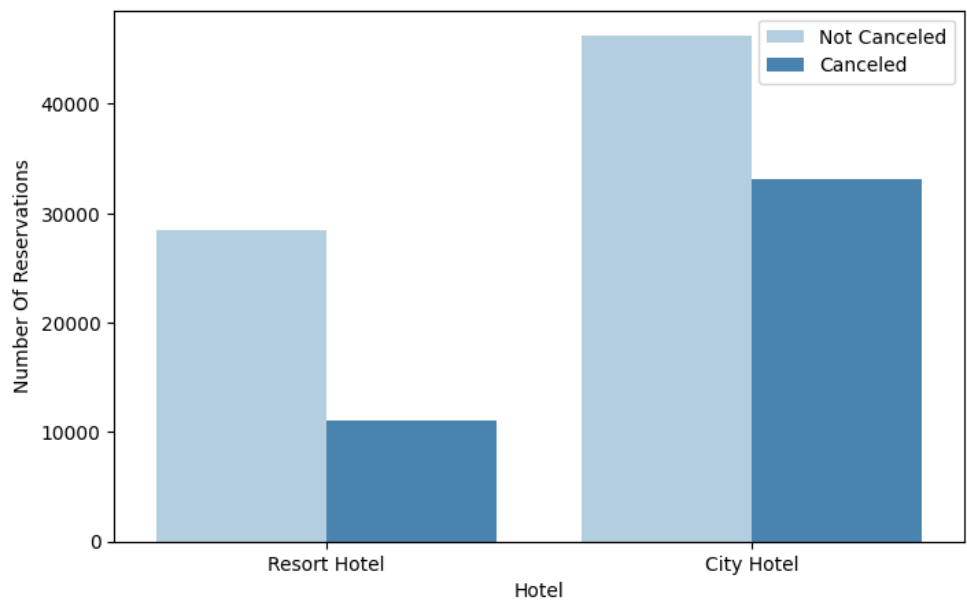
The data analysis and visualization you've performed provide valuable insights into the hotel dataset. Graphs can be particularly useful to understand patterns, trends, and relationships within the data. Here are some suggestions on where you could use graphs to enhance your analysis:

1) Cancellation Status Distribution:



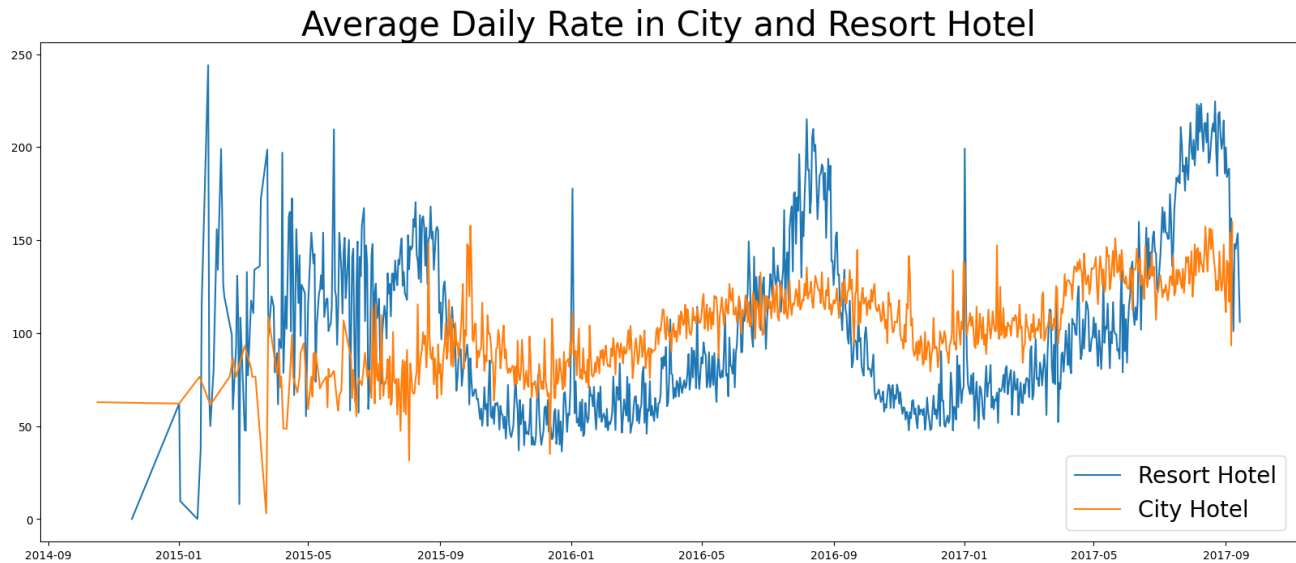
2) Hotel Type Comparison:

The count plot comparing the number of cancellations between Resort Hotel and City Hotel is informative. You might also consider creating a stacked bar plot for a more detailed view.



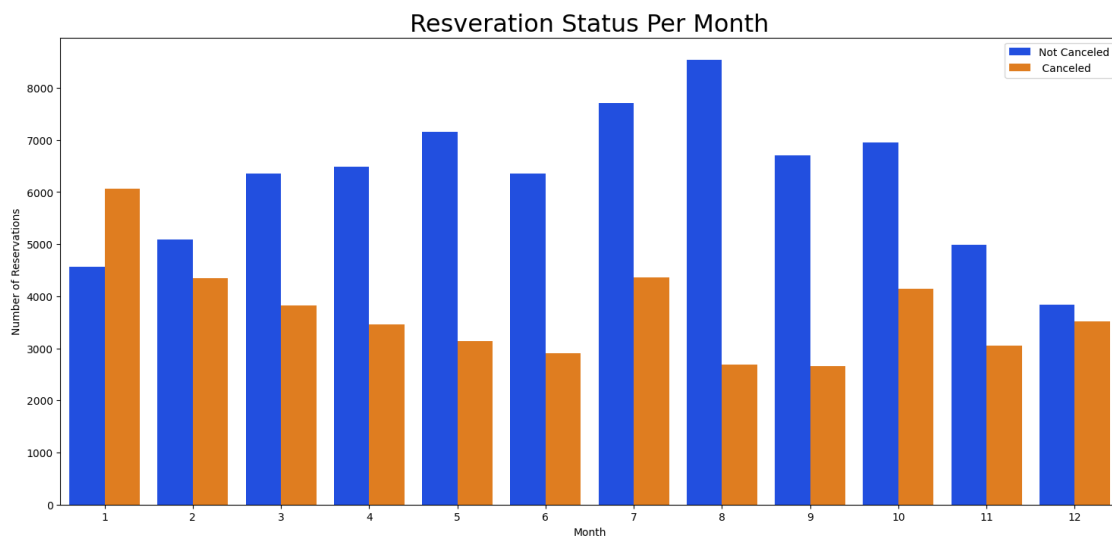
### 3) Cancellation Ratio Over Time:

The line plot showing the average daily rate over time for both canceled and not canceled reservations is useful. Consider adding more time-related visualizations, such as cancellations over months or years.



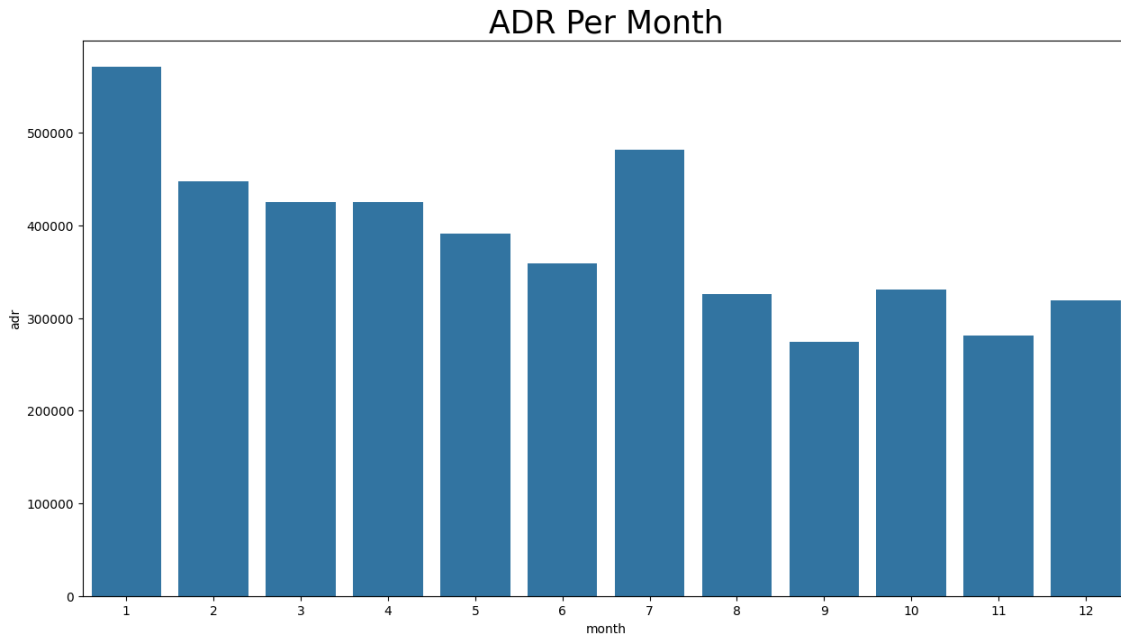
### 4) Monthly Reservation Status:

The count plot for reservation status per month is insightful. You could enhance it by using a line plot to show the trend over time.



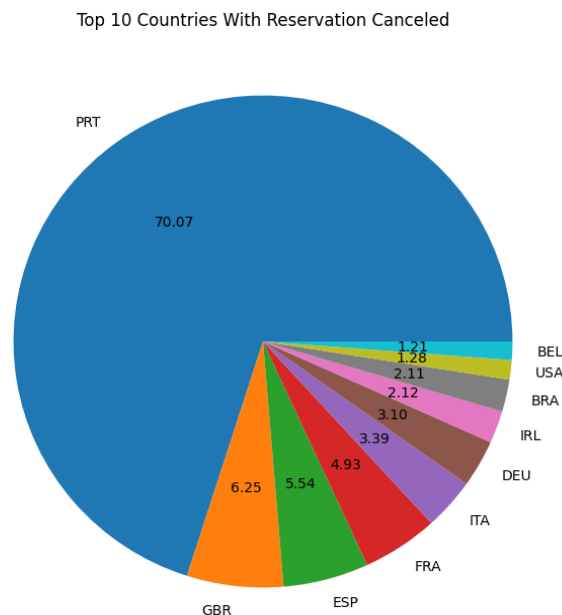
### 5) ADR Per Month:

The bar plot showing the Average Daily Rate (ADR) per month for canceled reservations is good. Consider creating a similar plot for non-canceled reservations to compare the trends.



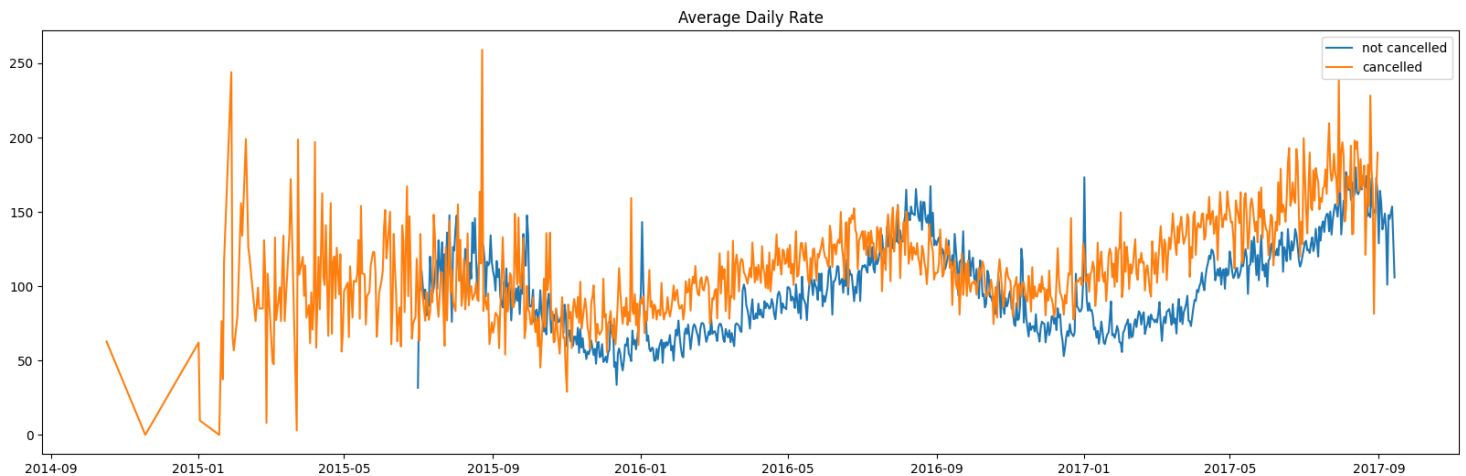
### 6) Top 10 Countries:

The pie chart for the top 10 countries with canceled reservations is effective. You might want to explore a bar plot or a horizontal bar plot for better readability.



## 7) Extended Report on Average Daily Rate Trends: Canceled vs. Not Canceled Hotel Reservations :

Compare Average Daily Rate (ADR) trends for canceled and not canceled hotel reservations over time. Additionally, visualize the ADR trends through a line graph.



## 8) Graphical Representation of Average Daily Rate Trends :

Visualize and compare Average Daily Rate (ADR) trends for canceled and not canceled hotel reservations over time.

