

# **Hotel Booking Analysis – Exploratory Data Analysis Report**

**By Amol Pakhare**

**August 2025**

## **◆ 2. Business Problem / Objective**

Hotel booking cancellations can lead to loss of revenue and poor resource planning. The objective of this analysis is to explore the hotel booking dataset to identify key factors affecting cancellations, guest preferences, and provide actionable suggestions to reduce cancellation rates and improve customer satisfaction.

## **◆ 2. Requirements:**

Dataset, Python, Numpy, Pandas, Scipy, Matplotlib, Seaborn, Statistics

## **◆ 3. Dataset Description:**

### **Dataset:**

The dataset contains booking information for a city hotel and a resort hotel, including details such as booking dates, number of guests, special requests, lead time, cancellation status, country of origin, etc. The dataset has 119,390 rows and 32 columns.

### ◆ 3. Data Exploration:

General Data Overview

Missing Values Check

Duplicate Values Check

Outliers Check

Wrong Values Check

Wrong Data Type Check

Check the Categorical, Count, Continuous and Time Series Variable

### ◆ 4. Data Preparation

**Data Cleaning Steps (Summary)**

**Data Cleaning Summary:**

- Removed duplicate records
- Handled missing values in columns like children, country, agent
- Converted date columns into proper datetime format
- Replace Wrong values to original
- Drop the Unnecessary Column/ Features

## ◆ 5. Exploratory Data Analysis (EDA) / Data Analysis

Univariate Analysis

Bivariate Analysis

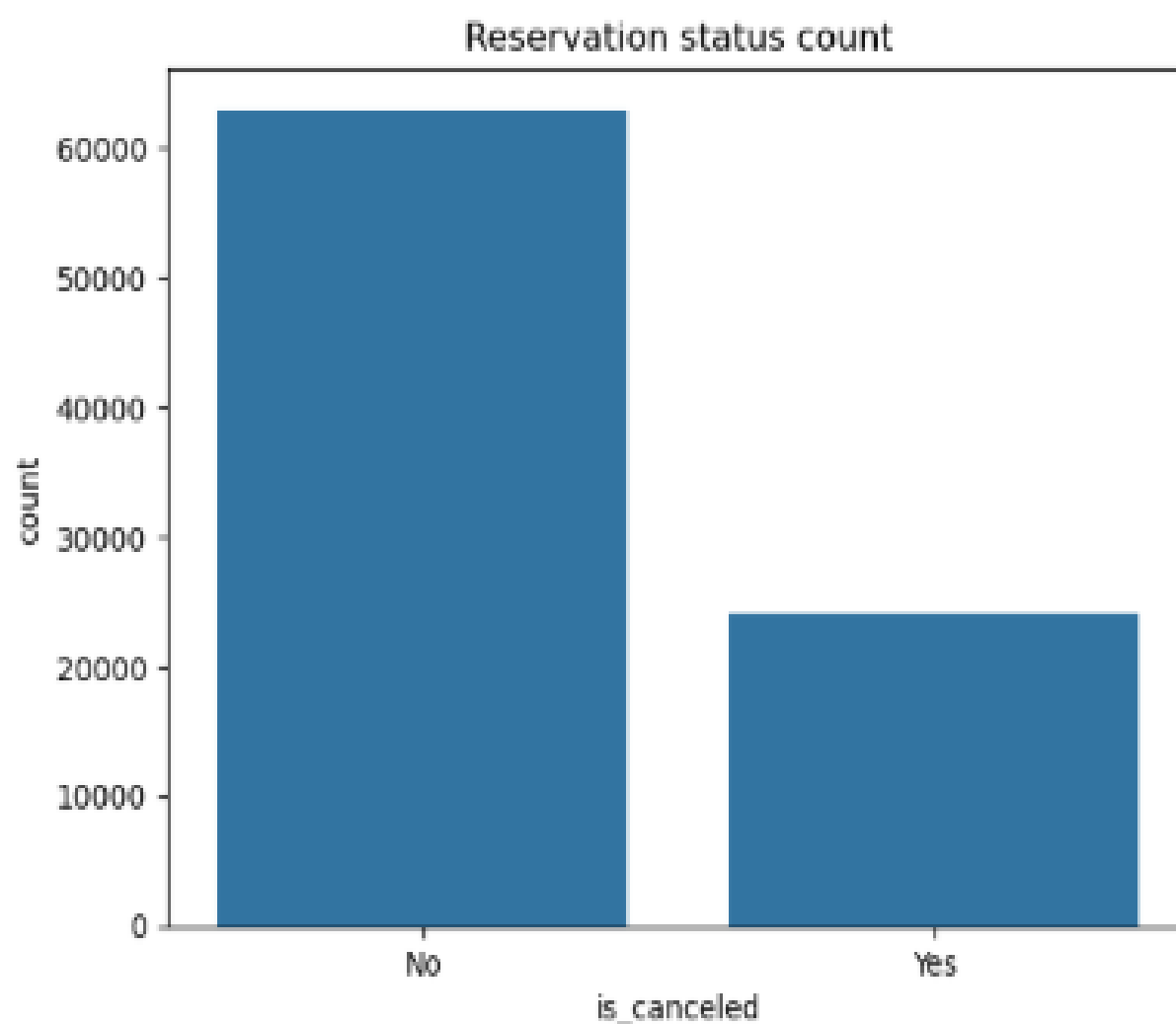
Multivariate Analysis

Chart Titel: Reservation status Count

----The chart clearly shows that:

**~63,000 bookings were not canceled**

**~25,000 bookings were canceled**



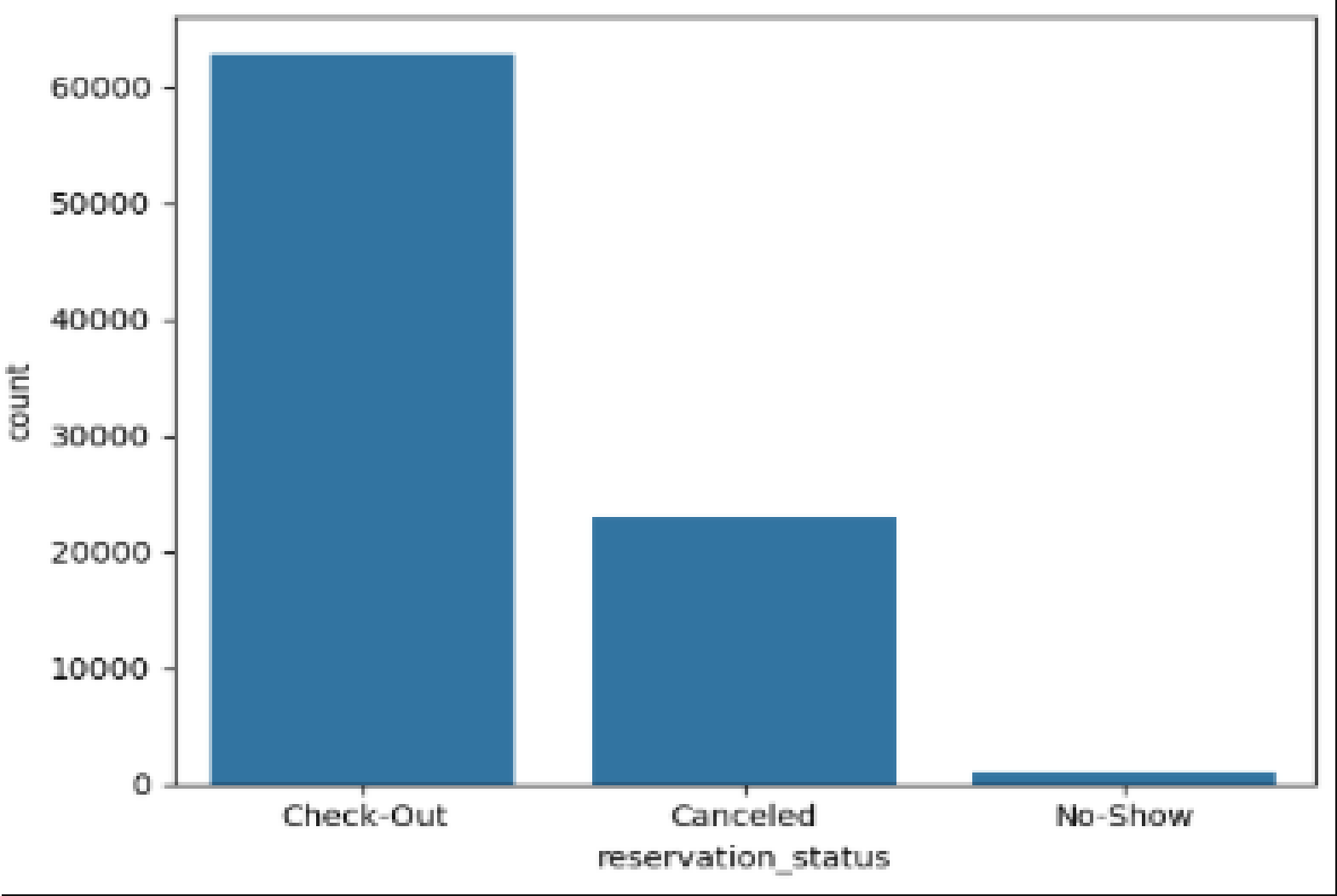
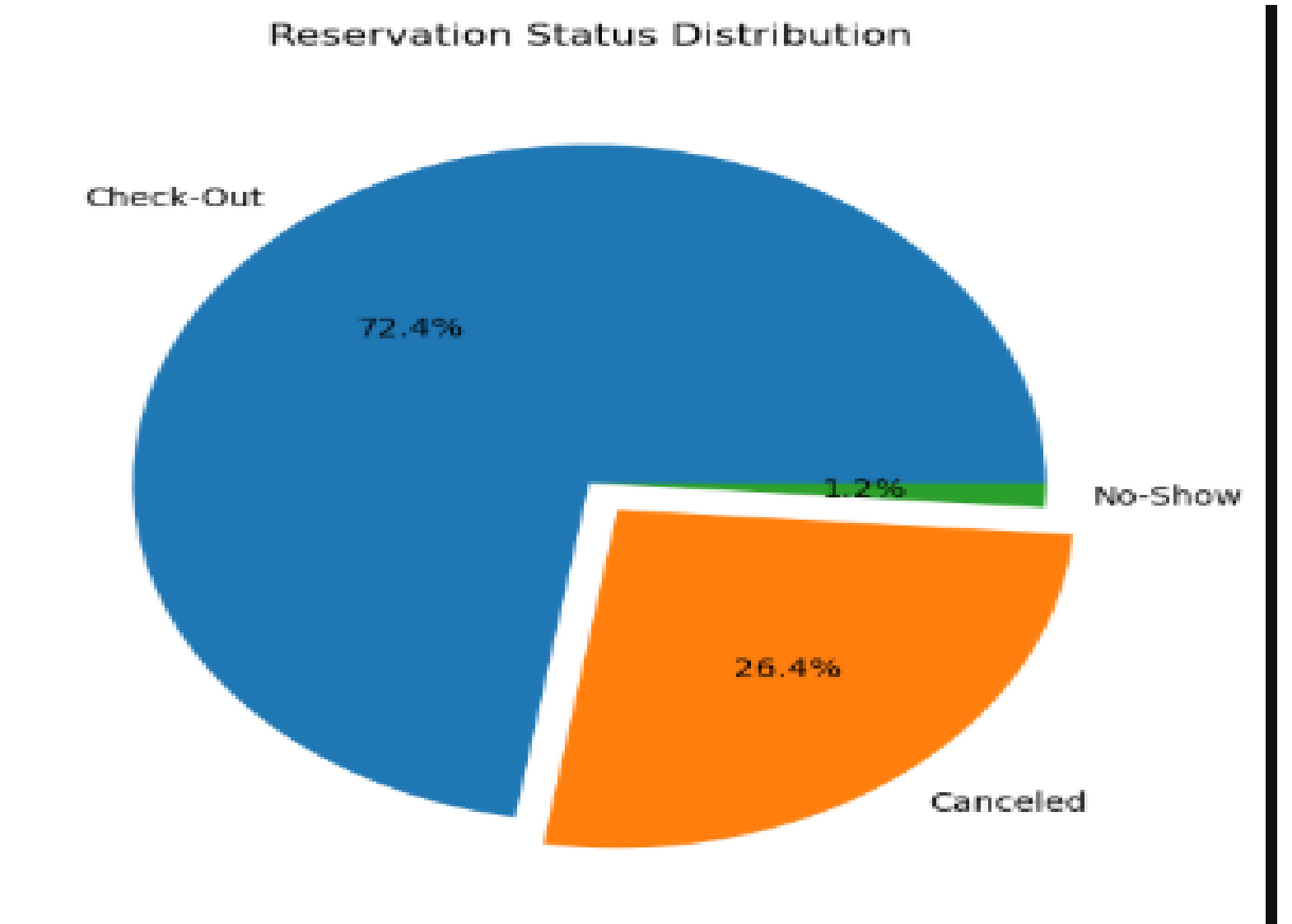
Con:

A significant number of bookings (~29%) were canceled. This indicates that **cancellations are a major issue** for the hotel, and understanding what factors lead to cancellations is important for reducing lost revenue.

**Pie Chart – Reservation Status Distribution**

**Check-Out – 72.4%**

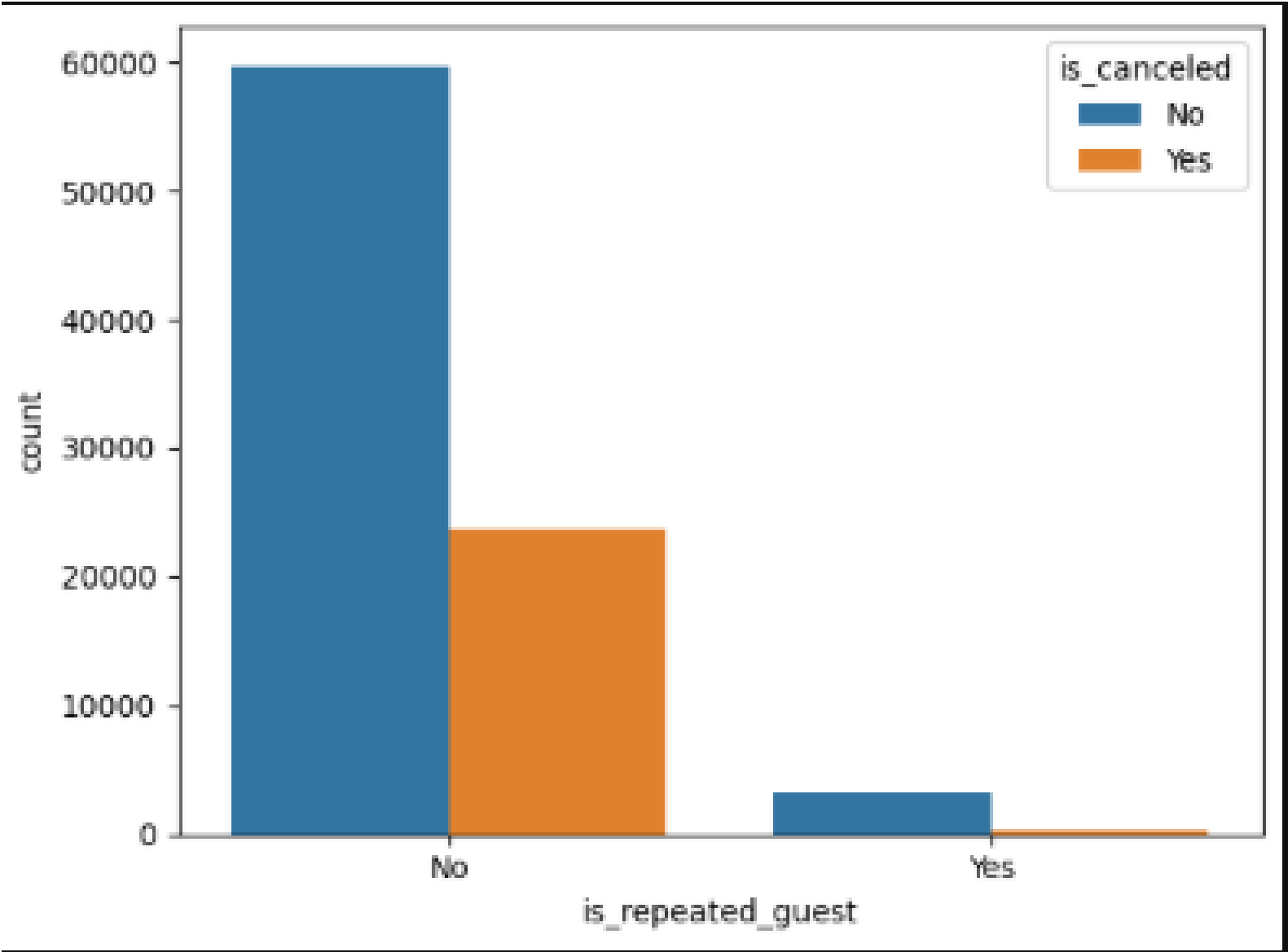
- **Canceled – 26.4%**
- **No-Show – 1.2%**



**Conl:**

Most reservations (72.4%) were successfully completed (guests checked out).  
More than 1 in 4 bookings were cancelled, which is a substantial loss potential.

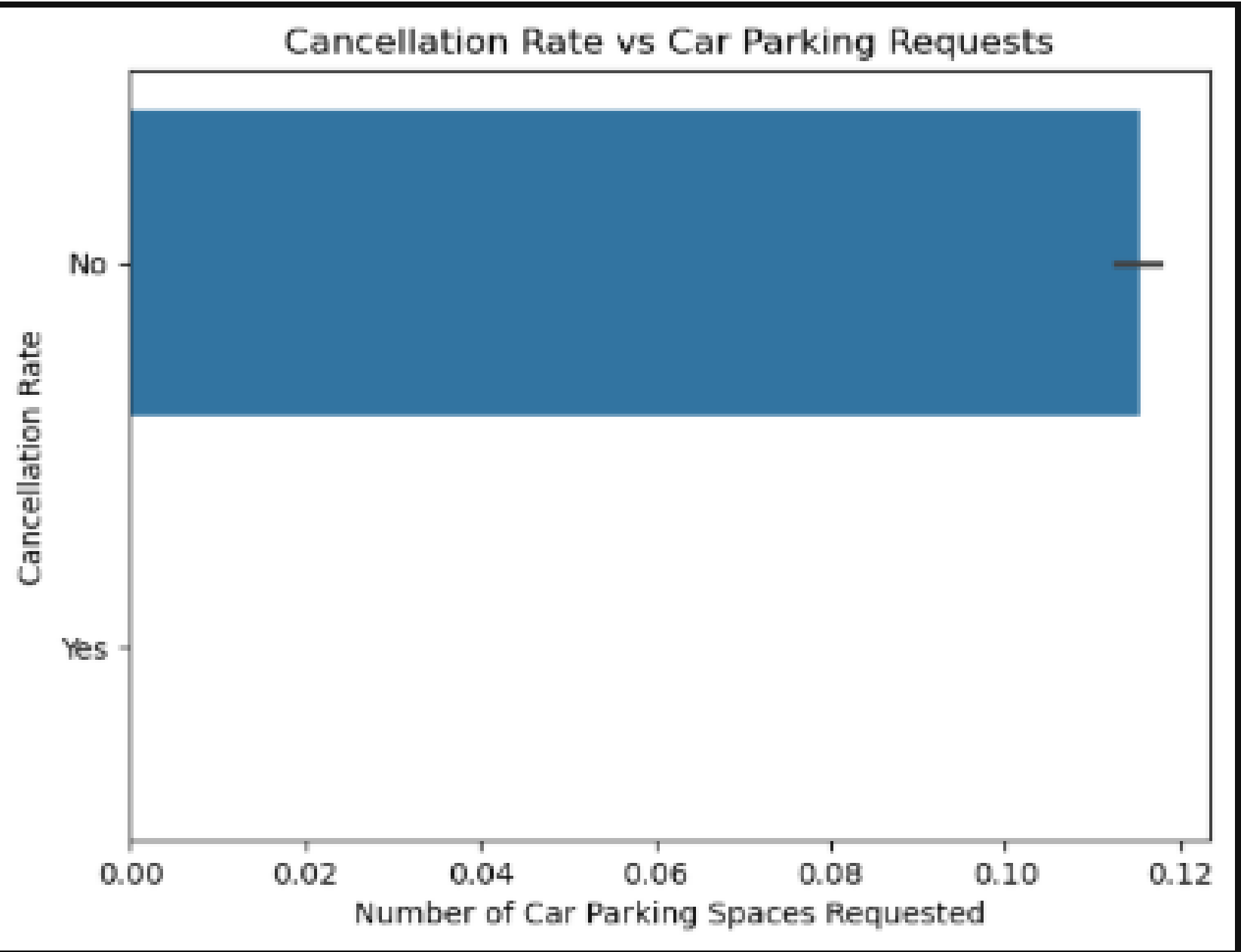
**Title: Grouped Bar Chart**  
**Guest Behaviour Analysis → Cancellation Rate by Repeat Guest**



**Conclusion from Chart:**

- Repeat guests are far less likely to cancel their bookings compared to first-time guests.
- Building loyalty programs or offering benefits to repeat customers can help **reduce cancellation rates**.

Chart Title: Cancellation Rate vs Car Parking Requests

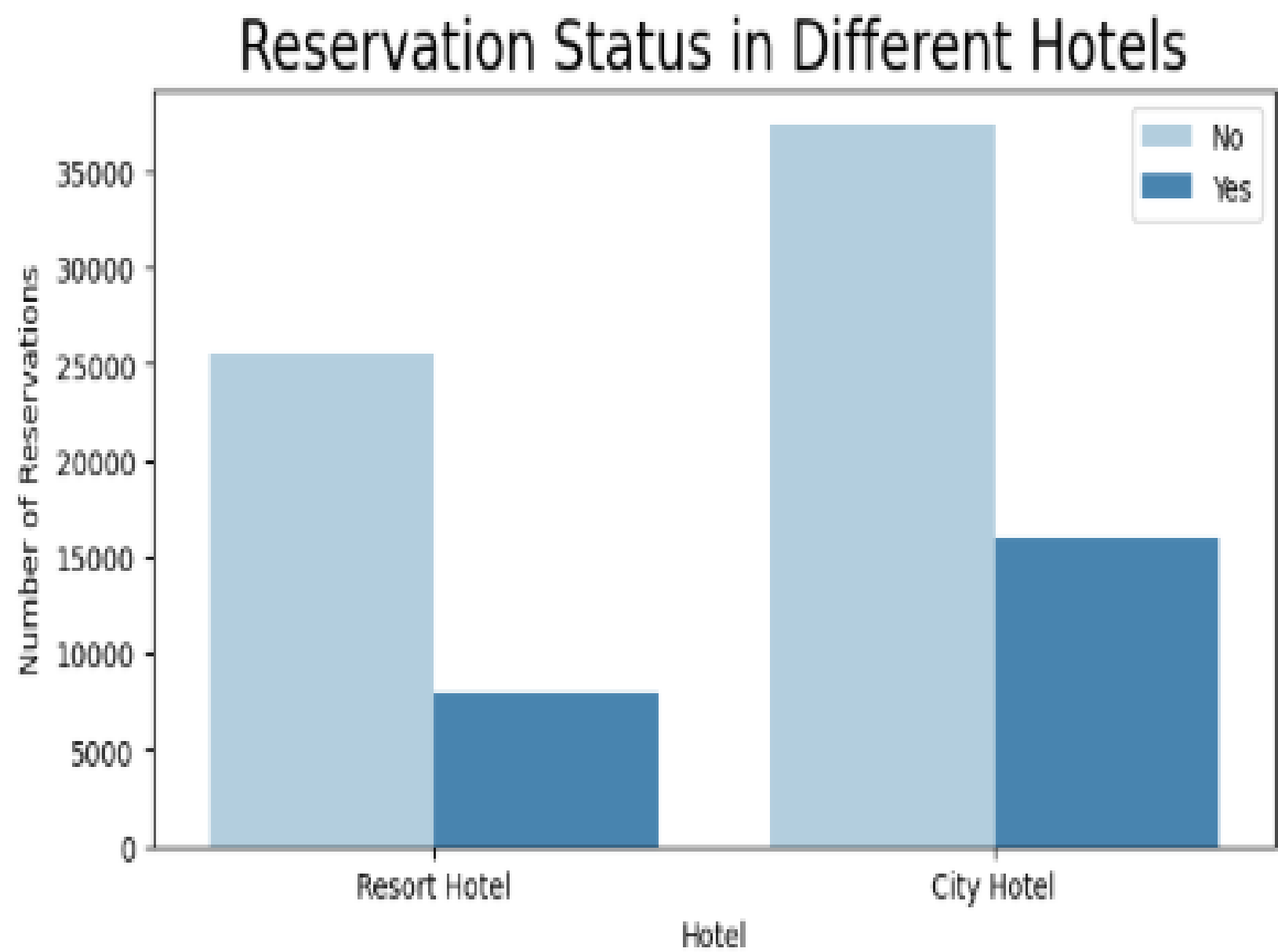


Conclusion:

Guests who request car parking are **less likely to cancel** their reservations. This implies that guests who plan for car parking are **more committed or serious** about their stay.

**Chart Title Reservation status in different hotels**

- **City Hotel** has a noticeably higher number of cancellations compared to Resort Hotel.
- Although both hotel types have more confirmed bookings than cancellations, the **City Hotel** has a higher cancellation rate.



**Con:**

**City Hotels face more cancellations than Resort Hotels.**

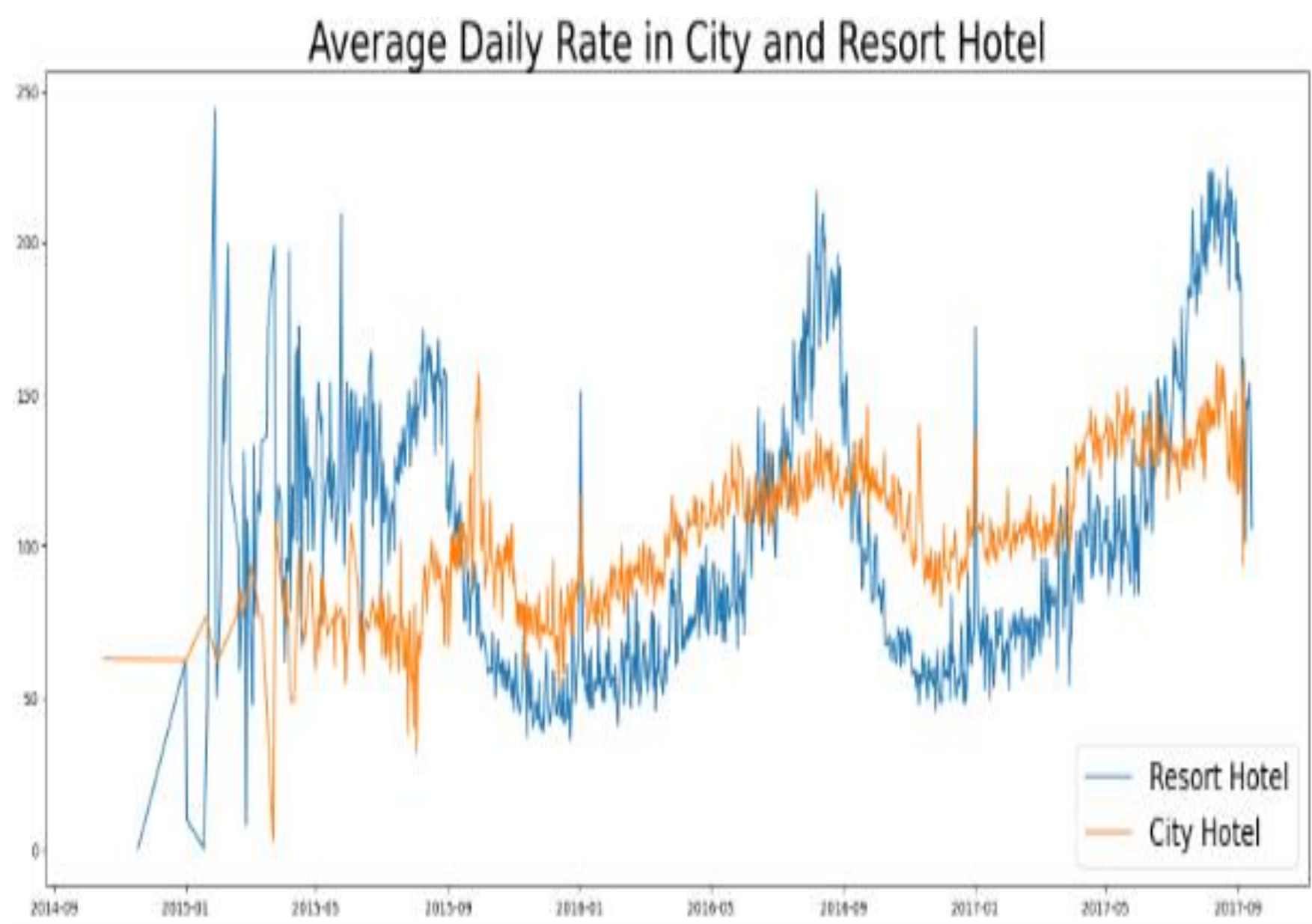
The count plot shows that both City and Resort Hotels have a significant number of cancellations, but City Hotels have a higher total and a greater number of cancelled bookings compared to Resort Hotels. This indicates that City Hotels face more booking volatility.

## line plot titled: **Average Daily Rate in City and Resort Hotel**

Resort Hotels show higher fluctuations in price (ADR), with sharp peaks, especially during mid-2016 and early 2017.

City Hotels have more stable pricing, with smoother increases and decreases.

Both hotel types exhibit seasonal patterns, but Resort Hotels are more sensitive to seasonal spikes.



## Con:

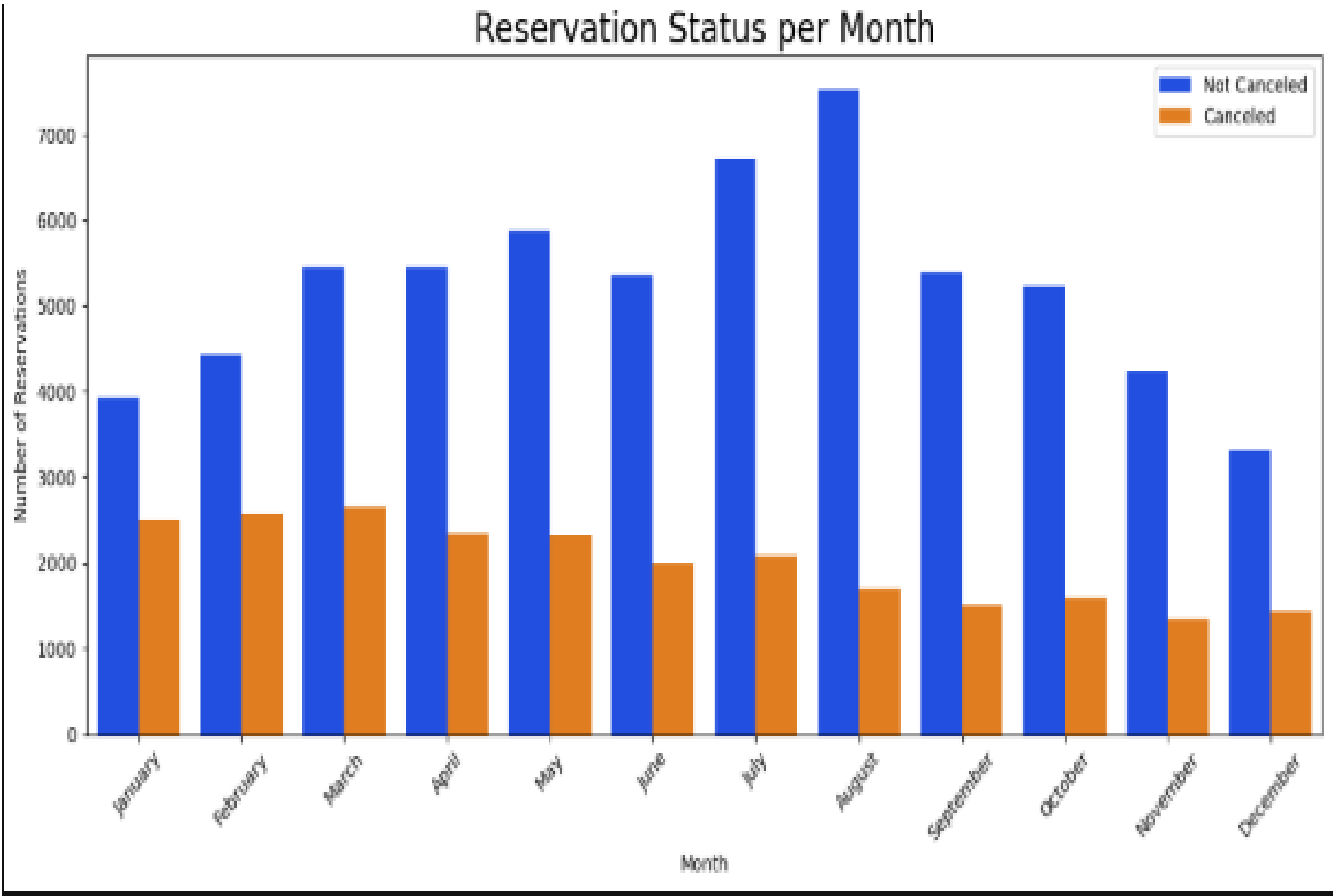
Resort Hotels experience more volatile pricing, likely due to seasonal demand (holidays, weekends, peak vacation months).

City Hotels maintain stable pricing, possibly due to business-travel demand and shorter stay durations



**Grouped bar chart**

Titled - Reservation Status per Month", showing monthly distribution of bookings split by Cancelled and Not Cancelled.

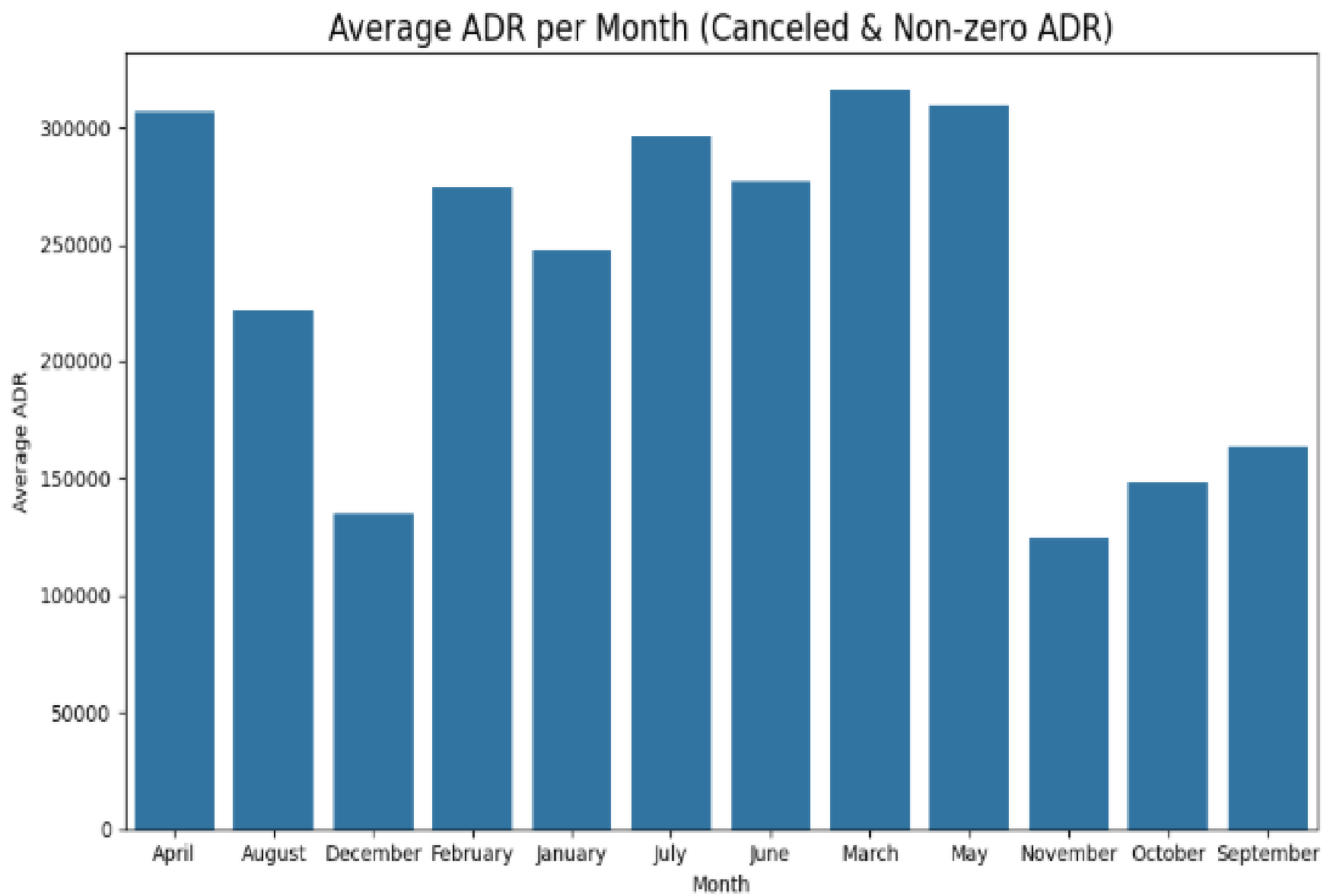


**Con:**

Non-cancelled bookings peak during July and August, followed by October and June.

Cancellations are relatively steady from January to May, and then gradually decline until December.

August has the highest number of overall reservations, but not the highest cancellations.

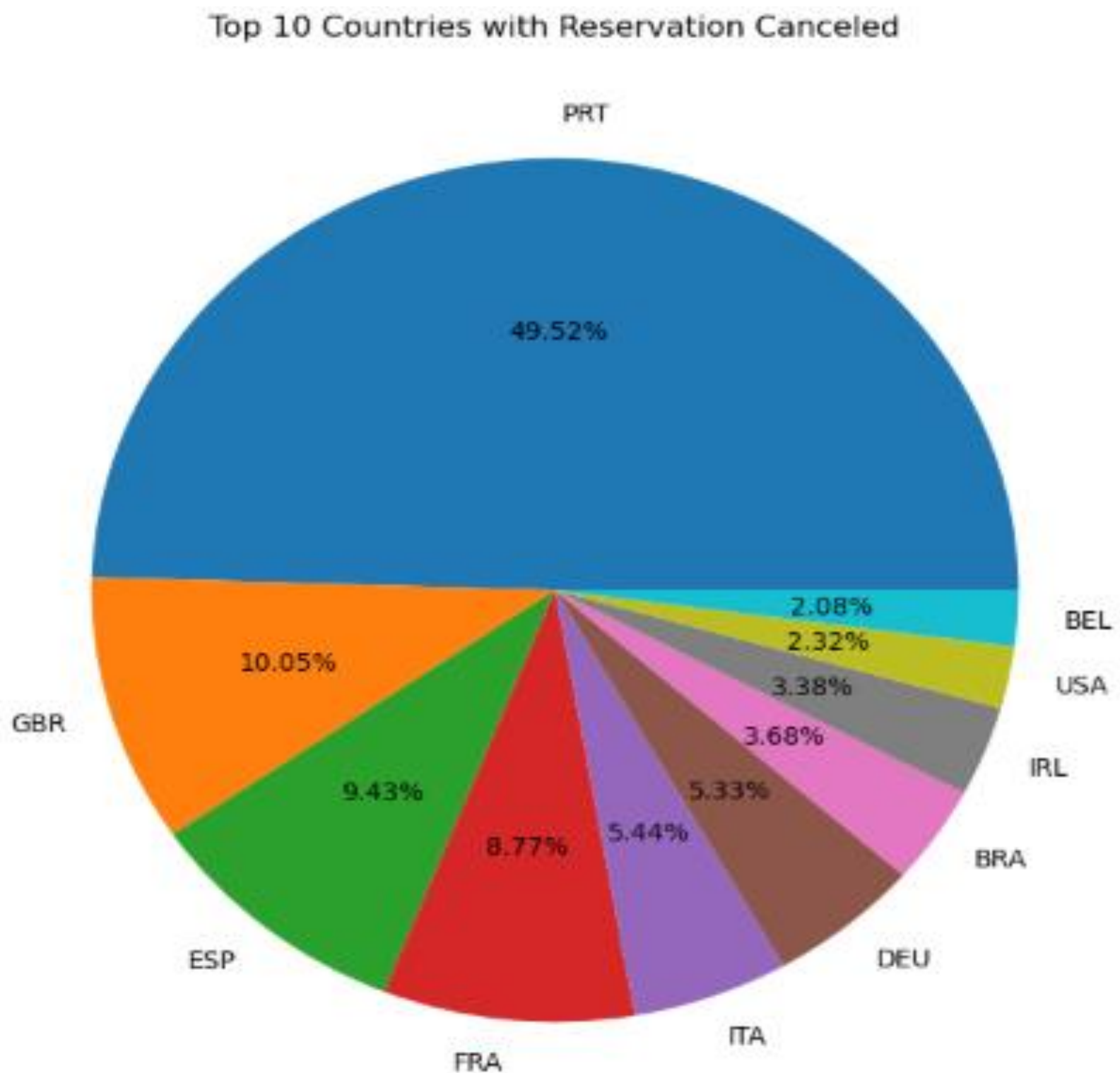


### Con:

- March, May, and April have the highest ADRs among cancelled bookings, hinting that customers tend to cancel more when prices are high.
- November, October, and December show lower ADRs, possibly indicating off-season periods.
- High ADR months coincide with more cancellations, suggesting that pricing may influence cancellation behaviour.

## Pie Chart

### Top 10 Countries with Reservation Cancelled



### Con:

**Portugal (PRT)** dominates cancellations **with 49.52%** of all **cancelled bookings**.

Other notable countries:

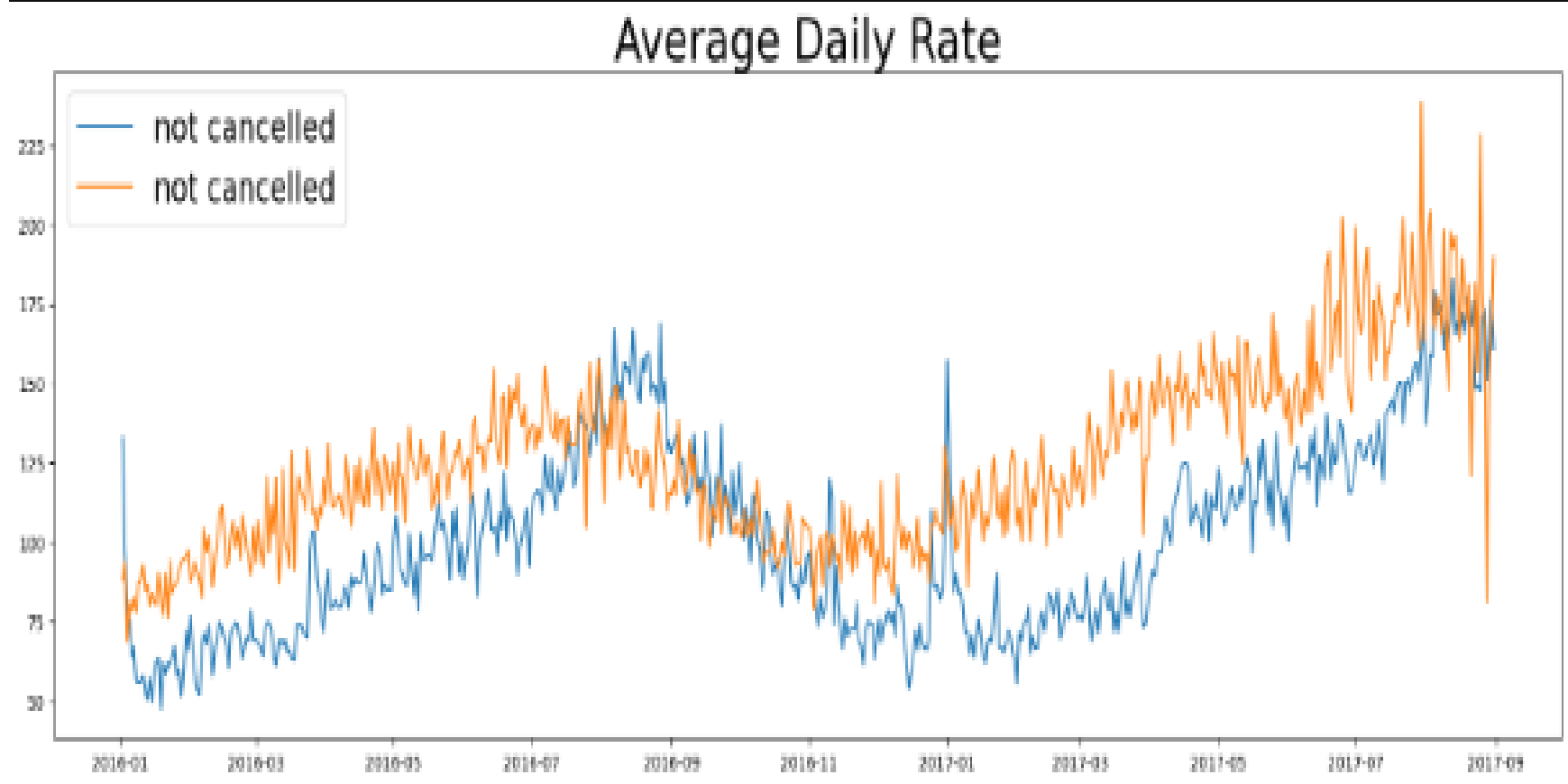
- **GBR (UK): 10.05%**
- **ESP (Spain): 9.43%**
- **FRA (France): 8.77%**

These top 4 countries together contribute over 77% of the cancellations.

Indicates a strong regional pattern in booking behaviour, possibly affected by marketing, policies, or customer expectations.

## Line Plot-

Average Daily Rate with Cancellation rate



### Con:

Average Daily Rate (ADR) significantly influences the cancellation rate — higher ADRs are often associated with a greater likelihood of cancellations.

### Key Observation:

1. City hotels have a significantly higher cancellation rate compared to resort hotels.
2. Resort hotel prices fluctuate seasonally, while city hotel rates remain relatively stable.
3. Cancellations are more common during high-price months like March, April, and May.
4. Portugal accounts for nearly 50% of all reservation cancellations.
5. Bookings with more children or babies show a higher tendency to get canceled.
6. Higher average daily rates are associated with an increased likelihood of cancellation.
7. ADR is highest during spring and summer, and lowest in late fall and winter.
8. Introducing flexible pricing, family offers, and country-specific strategies can help reduce cancellations and boost revenue.

- **Suggestions:**

1. **Control Prices to Reduce Cancellations:**

When prices go up, cancellations also increase. Hotels should adjust their pricing based on location and offer discounts to attract more confirmed bookings.

2. **Offer Weekend or Holiday Discounts (Especially for Resort Hotels):**

Since resort hotels have more cancellations than city hotels, they can provide special deals on weekends or holidays to encourage more people to book and stay.

3. **Promote More in January:**

January has the highest number of cancellations. Hotels can run special campaigns or limited-time offers during this month to increase bookings and reduce cancellations.

4. **Improve Services in Countries Like Portugal:**

Since Portugal has a high cancellation rate, hotels should focus on improving room quality and customer service there. This can help build trust and reduce cancellations.

5. **Control Room Prices:**

Don't keep prices too high in busy seasons — high prices lead to more cancellations

6. **Focus on Good Customers by Country:**

Try to get more guests from countries where people don't cancel much. Avoid depending too much on countries like Portugal.

7. **Change Prices Based on Month:**

Set room prices smartly each month — charge more in busy months, less in quiet ones, but don't overcharge.

8. **Offer Discounts in Off-Season:**

Give offers or packages in low months like November or December to fill empty rooms.