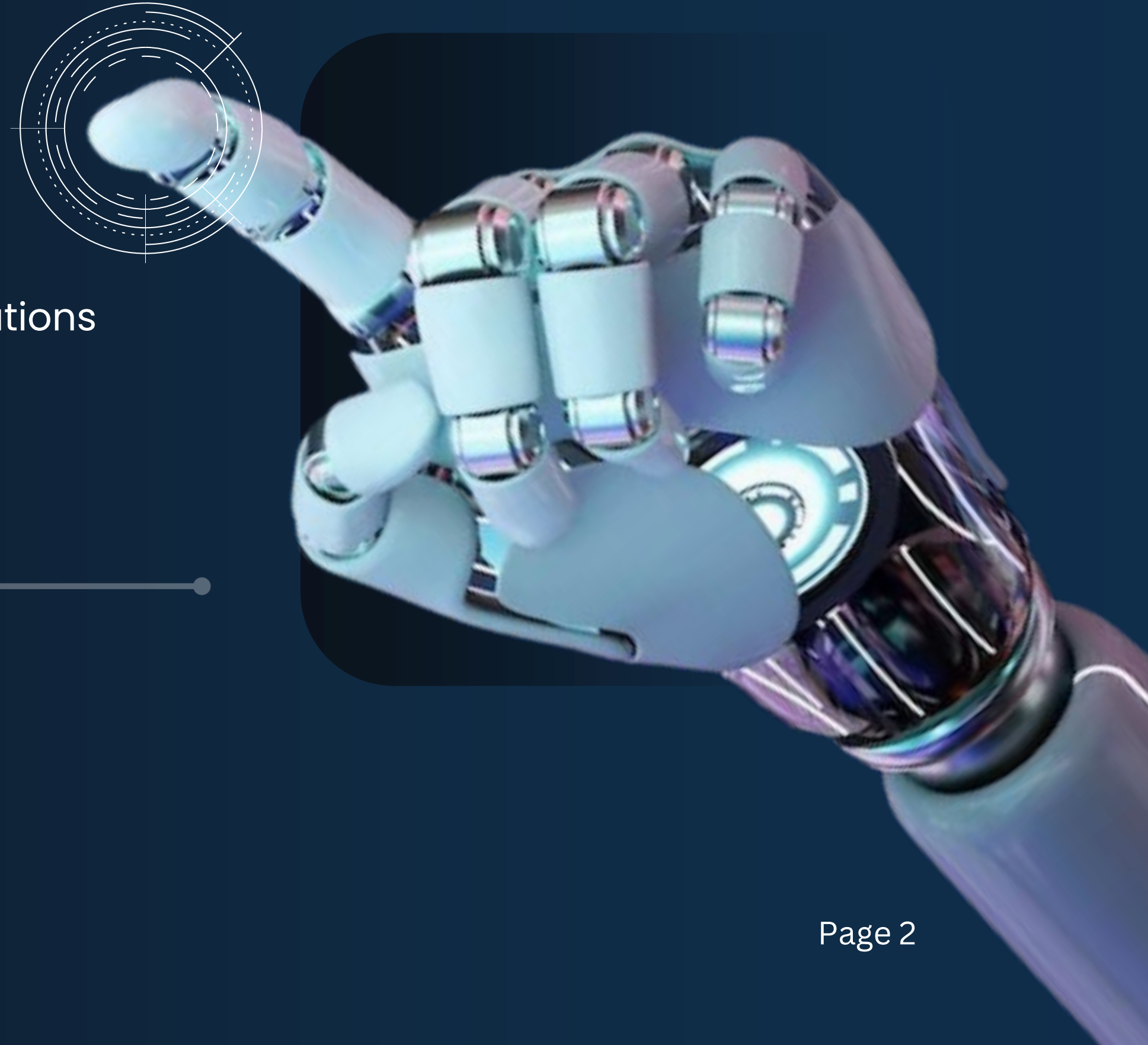# Cellula Task 1

JOHN SHOUKRY

Introduction

Performing EDA on a given Dataset for Hotel Reservations Given that Cancellations in the hotel industry are a significant challenge. They lead to revenue loss, operational inefficiencies, and wasted resources.

EDA

Exploratory Data Analysis or EDA is an approach to determine key attributes, get insight, determine patterns in data, spot anomalies, and identify relationships between variables in raw data using numbers, tables, and graphs. It is always a good practice to know about the data before you start working on it.

Dataset Overview

- Consists of (rows:36285, columns:17)
- data set doesnot contain Null values or duplicate values

- data info (data types of each column) :

```
 #   Column                  Non-Null Count   Dtype
---  ------                  --------------   -----
 0   Booking_ID              36285 non-null   object
 1   number of adults        36285 non-null   int64
 2   number of children      36285 non-null   int64
 3   number of weekend nights  36285 non-null  int64
 4   number of week nights   36285 non-null   int64
 5   type of meal            36285 non-null   object
 6   car parking space       36285 non-null   int64
 7   room type               36285 non-null   object
 8   lead time               36285 non-null   int64
 9   market segment type     36285 non-null   object
 10  repeated                36285 non-null   int64
 11  P-C                     36285 non-null   int64
 12  P-not-C                 36285 non-null   int64
 13  average price           36285 non-null   float64
 14  special requests        36285 non-null   int64
 15  date of reservation     36285 non-null   object
 16  booking status          36285 non-null   object
```

Data insights

Number of canceled Reservation: 11889
percentage of canceled Reservations is 32.77%

which is high percentage, our task is to find what are the features that influenece the customers to cancel their reservation to adjust them and help the business grow more efficiently.
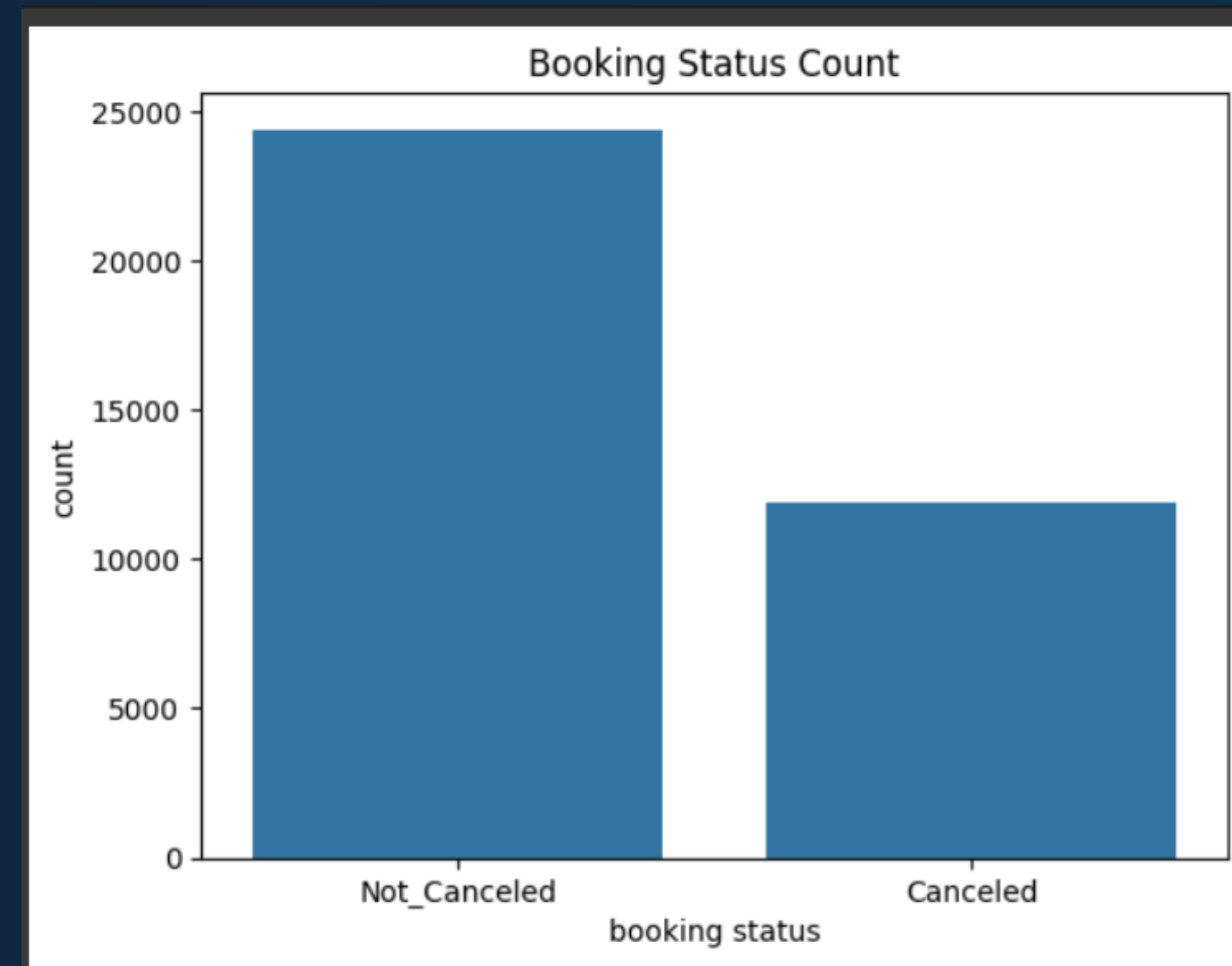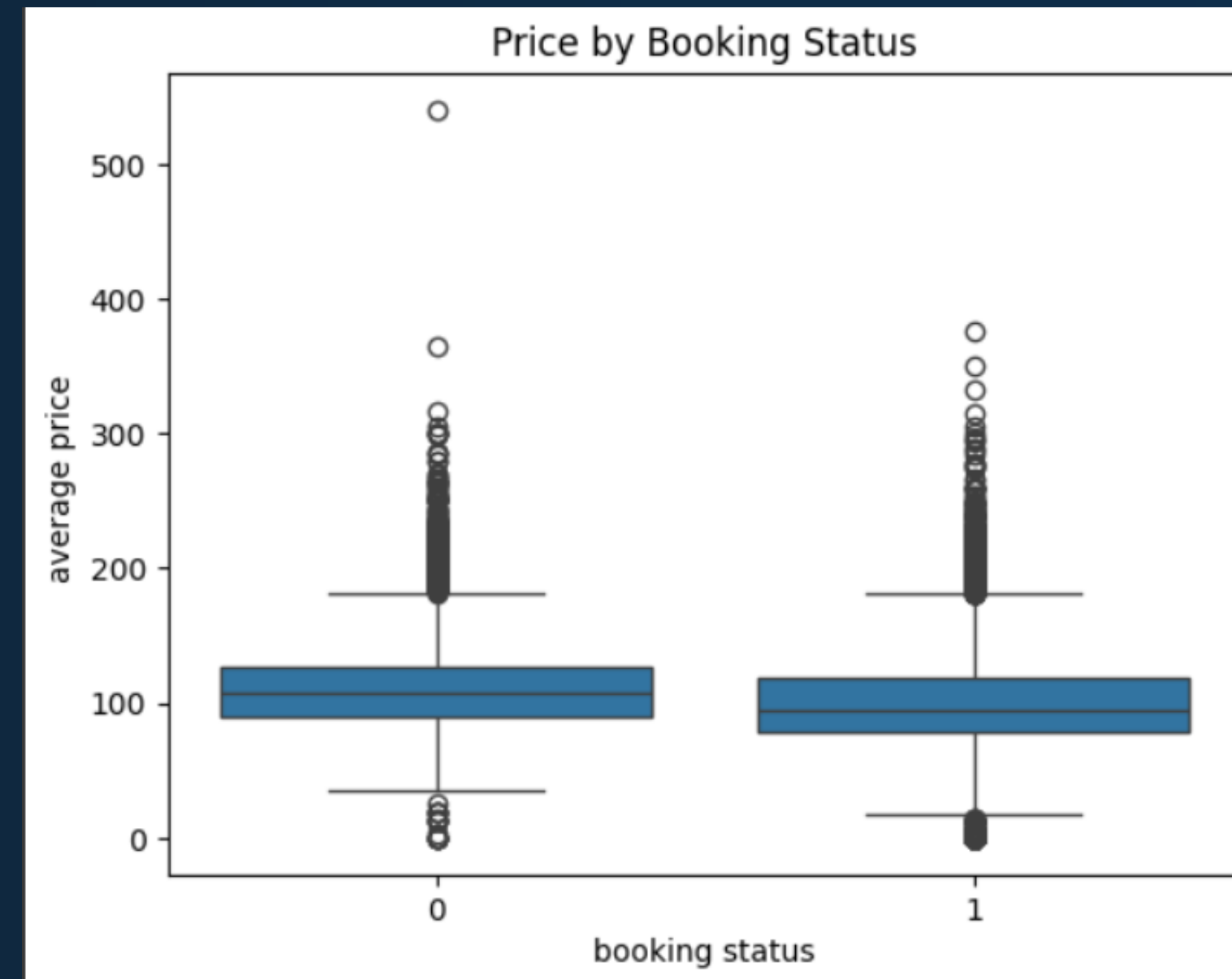
Average price influence on booking status

customers who paid higher prices were more likely to cancel their reservations.
 This might suggest:
They expected something high from the hotel and got disappointed so they canceled the reservations

Price by Booking Status

lead Time influence on booking status

Customers who Reserved far in advance were more likely to cancel their reservations.
This might suggest:
They Found a better deal or an offer from other hotels

Lead Time by Booking Status
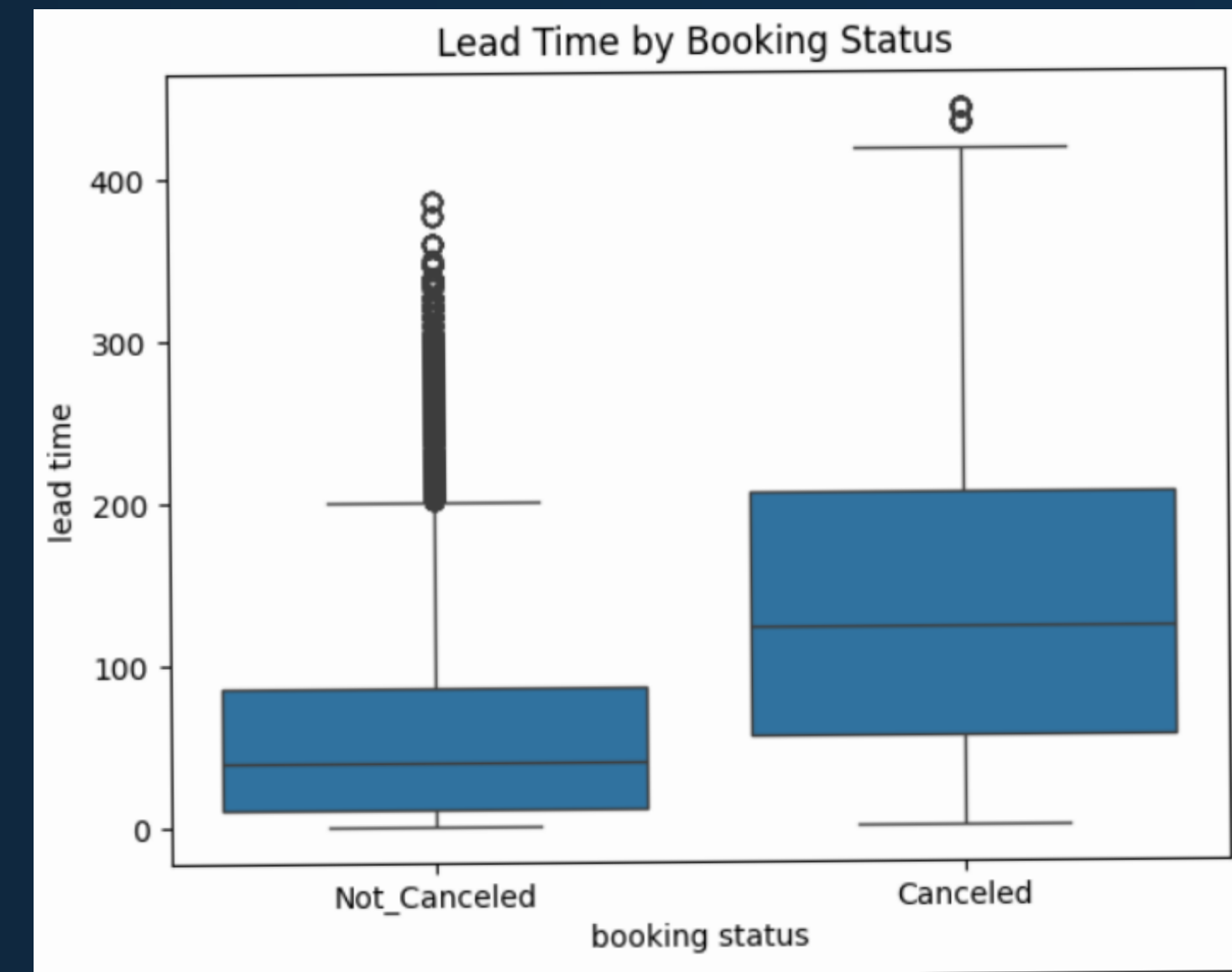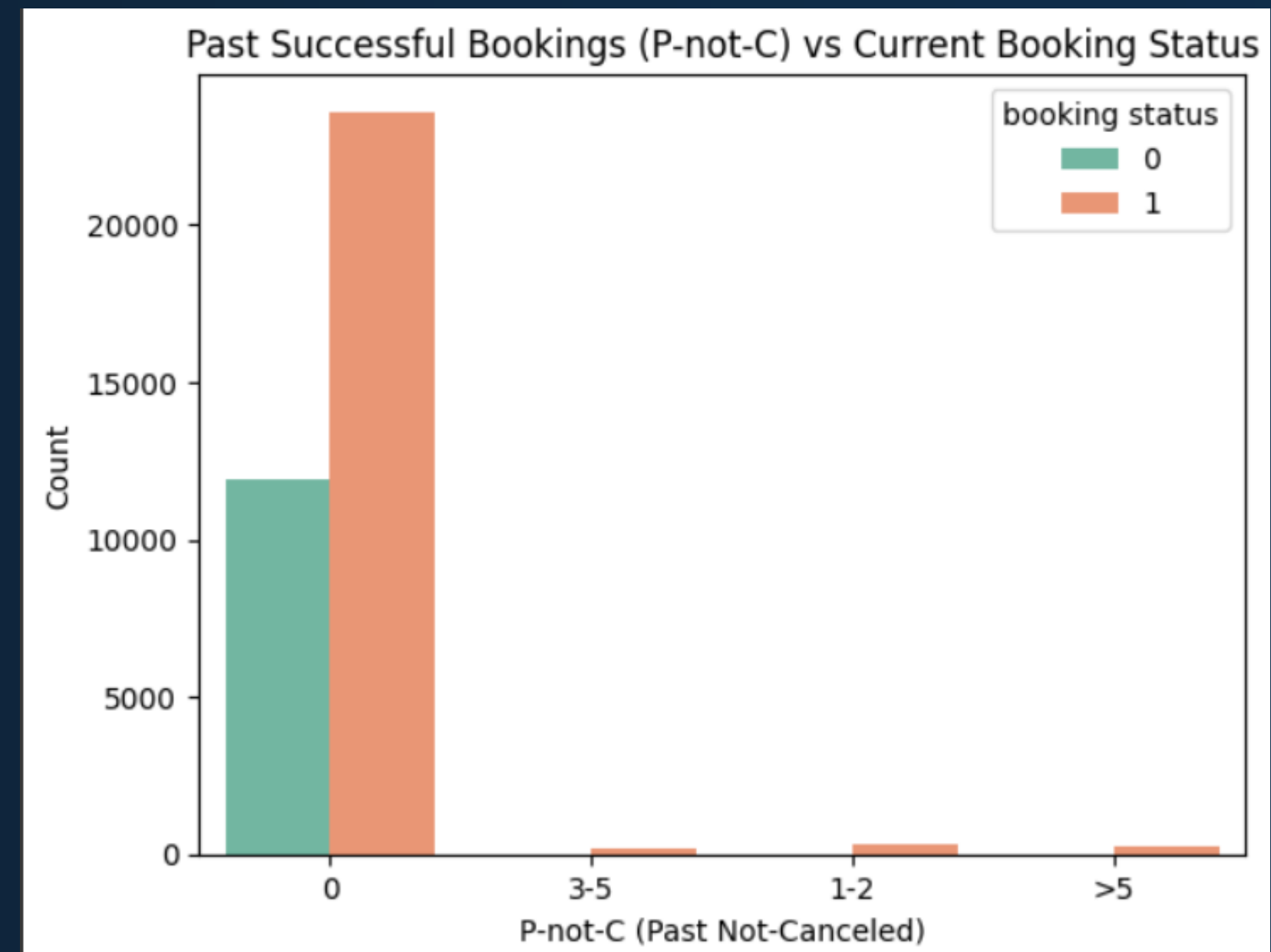
Market Type influence on booking status

- Online segment has the highest number of bookings, but also the highest number of cancellations.
- Offline bookings are more balanced but still have more cancellations.
- Corporate, Aviation, and Complementary segments show much lower volume overall.
- The Corporate segment, despite its smaller size, shows a notable cancellation rate.
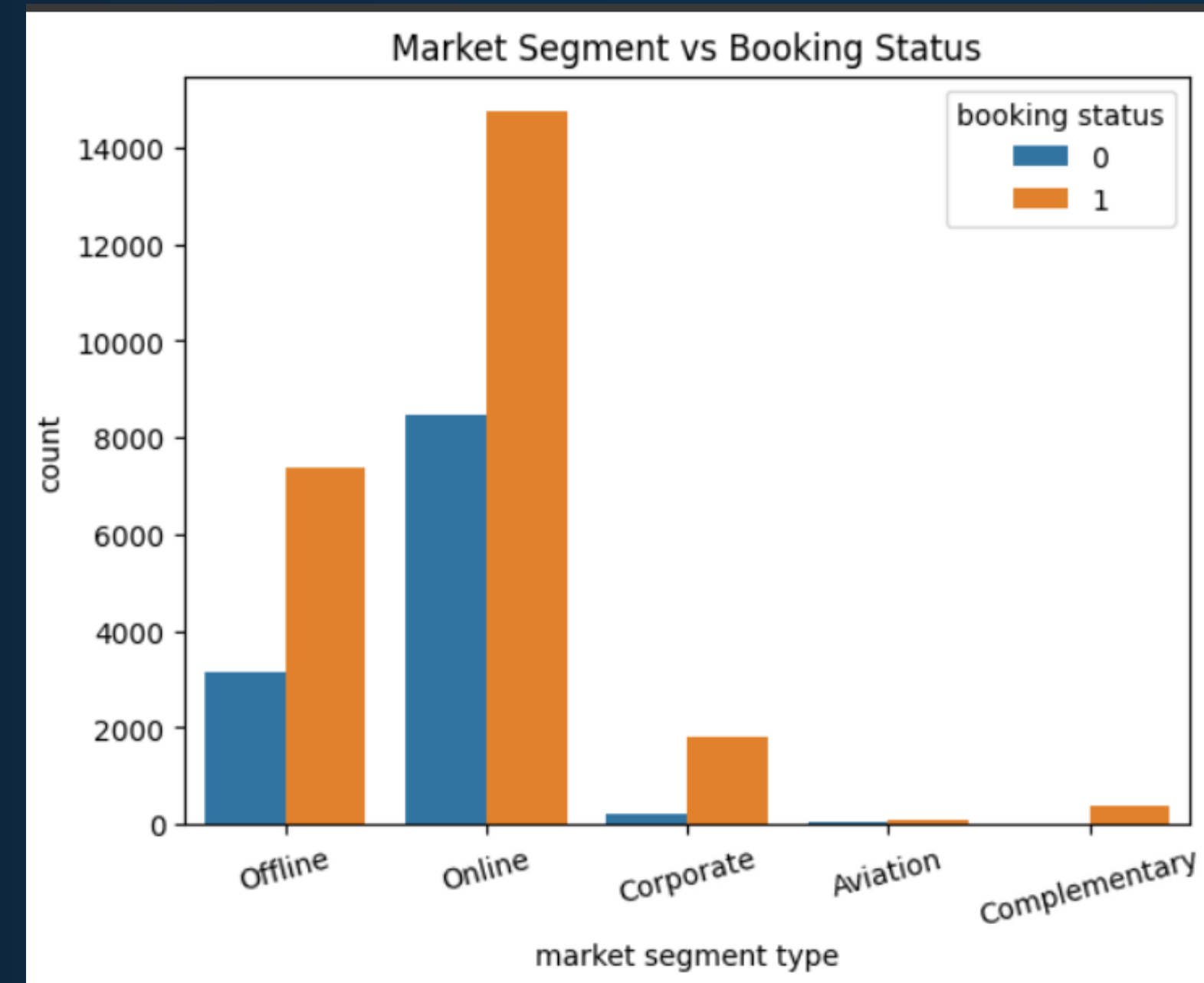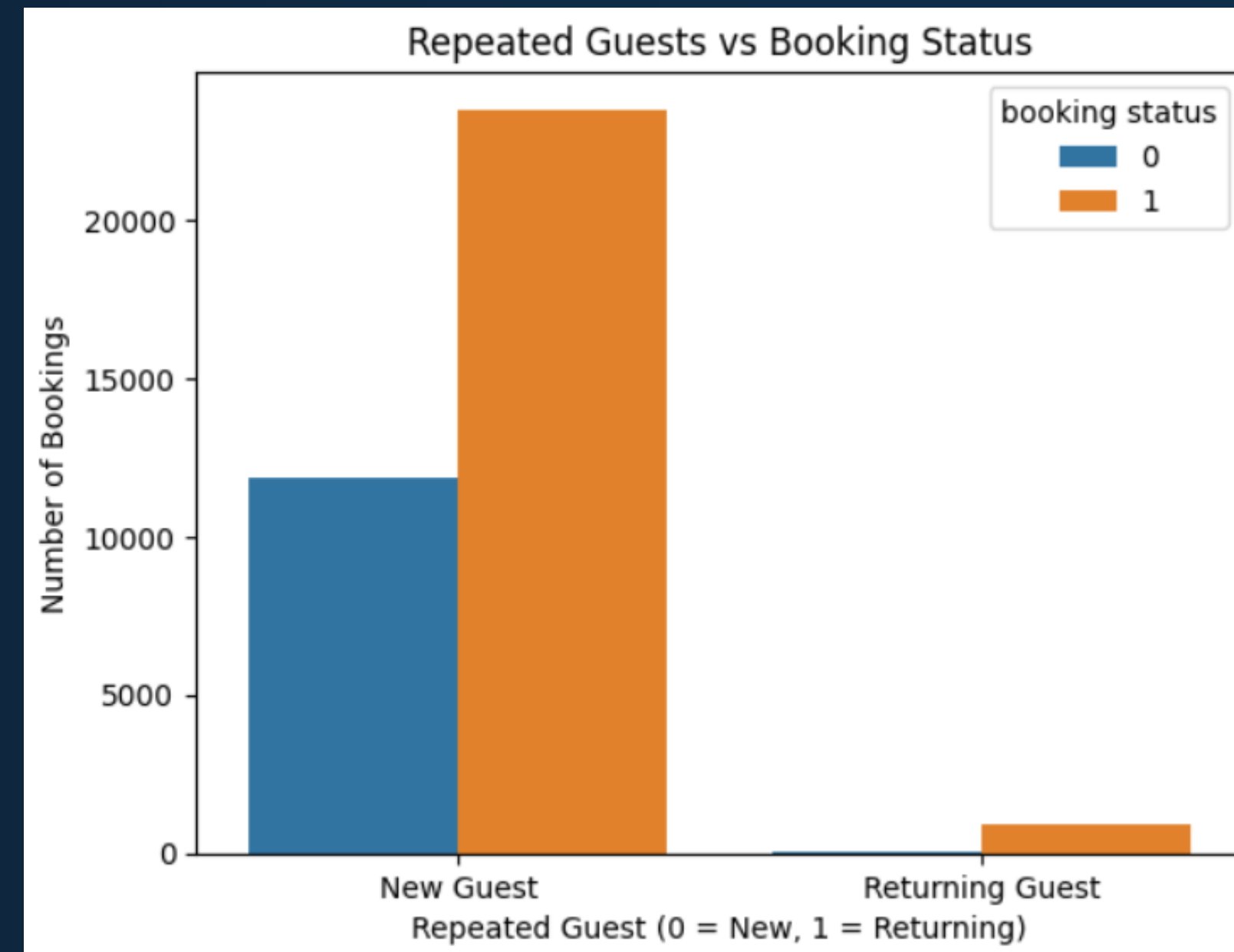
P_not_c influence on booking status

customers who Reserved before and didnot cancel shows that the are less likely to cancel reservations

Past Successful Bookings (P-not-C) vs Current Booking Status

Repeated influence on booking status

customers who Came before to the Hotel are less likely to cancel the reservation

New guests (Repeated = 0) show a higher cancellation rate, possibly due to uncertainty or better deals

Thank You

FOR YOUR ATTENTION