Investigate Hotel Business using Data Visualization

Fadillah Akbar

Data scientist with a strong mathematical foundation and problem-solving abilities. Experienced in building projects in data mining, data processing, business performance analysis, data visualization, and predictive modeling across multiple industries. Motivated to use data science to improve business impact through analytics, statistics, and machine learning.

www.linkedin.com/in/fadillahakbar

Business Overview

It is crucial for a firm to continuously evaluate its operations. This time, we'll look into the hospitality industry's business operations. Our main objective is to determine how our consumers behave while making hotel reservations and how that behavior relates to the cancellation rate for hotel reservations. To make the findings of the insights we discover more comprehensible and persuasive, the results will be presented as data visualization.

Table of Content

www.linkedin.com/in/fadillahakbar

Data Preprocessing

Make raw data into clean data and ready to be processed

Handle Missing Values

There are 4 columns that have null values

```
df['children'] = df['children'].fillna(0)
```

Fill the children column with 0 indicating there are no children in the hotel booking data

```
df['city'] = df['city'].fillna('unknown')
```

Fill the city column with "unknown" because the country of origin of the hotel booker is unknown

```
df['agent'] = df['agent'].fillna(0)
df['company'] = df['company'].fillna(0)
```

Fill the agent and company columns with 0 as an "undefined" value.

Handle Invalid Values and Remove Unnecessary data

```python
df["meal"].replace("Undefined", "No Meal", inplace=True)
```

The "meal" data contains the value "Undefined", which is actually the same as "No Meal"

```python
df["guest"] = df["adults"]+df["children"]+df["babies"]
```

Create total guest data which is the sum of adults, children, and babies

```python
df_preprocessing = df[df["guest"] > 0].copy()
```

There are 180 data that do not have the number of guests, so we will drop these data from the analysis.
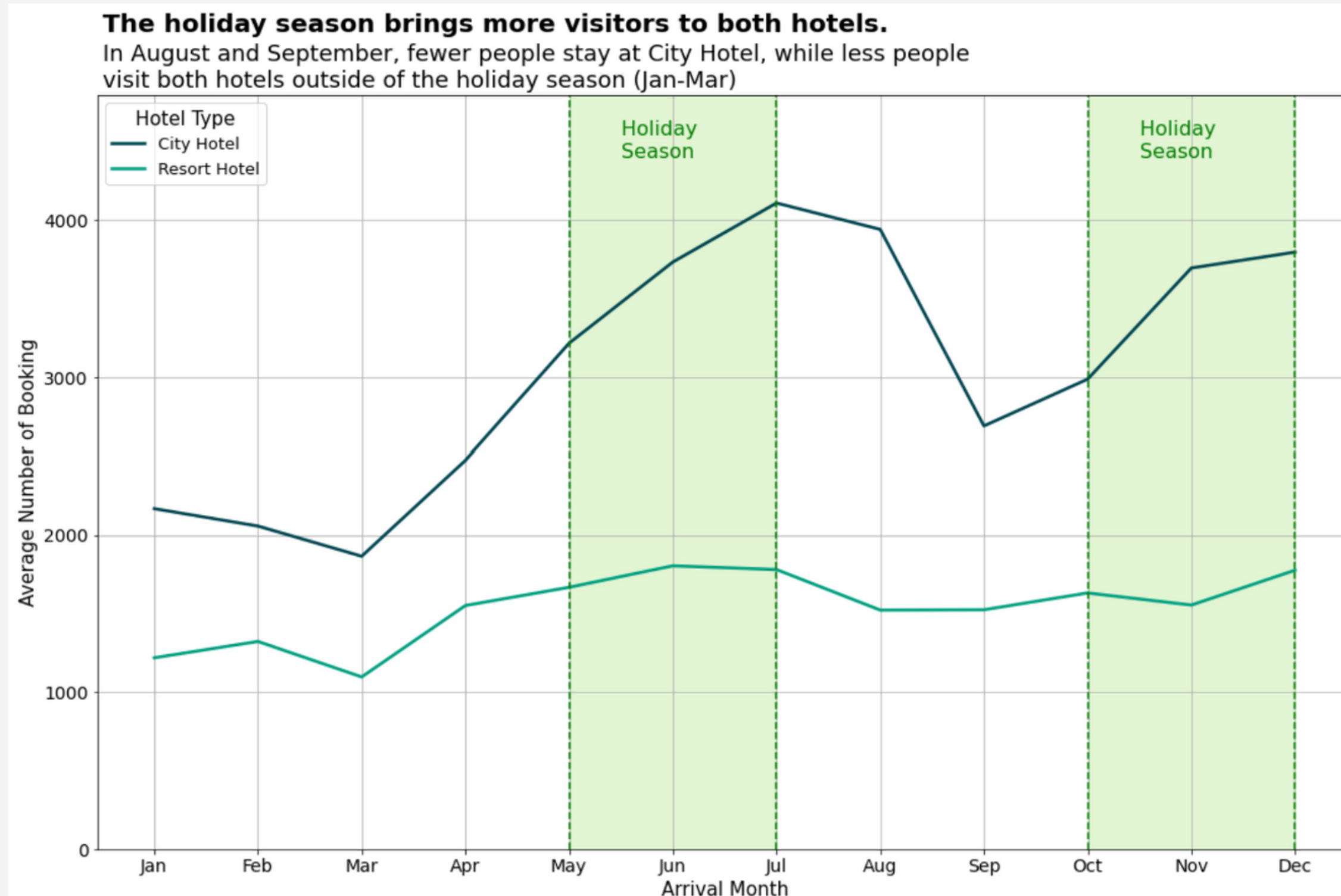
Monthly Hotel Booking Analysis Based on Hotel Type

Observe and analyze in which months or seasons there is an increase and decrease in the number of hotel reservations for each type of hotel.

Insight/analysis

The holiday season brings more visitors to both hotels.

In August and September, fewer people stay at City Hotel, while less people visit both hotels outside of the holiday season (Jan-Mar)

This analysis focuses on examining the pattern of hotel reservations for each type of hotel.

The peak booking season for both kinds of hotels is between June and July. Further investigation reveals that these times fall throughout the Hari Raya and school holidays.

Additionally, it appears that reservations are substantially lower in other months that do not overlap with any holidays. Thus, in August and September, business at Urban Hotels significantly decreased.
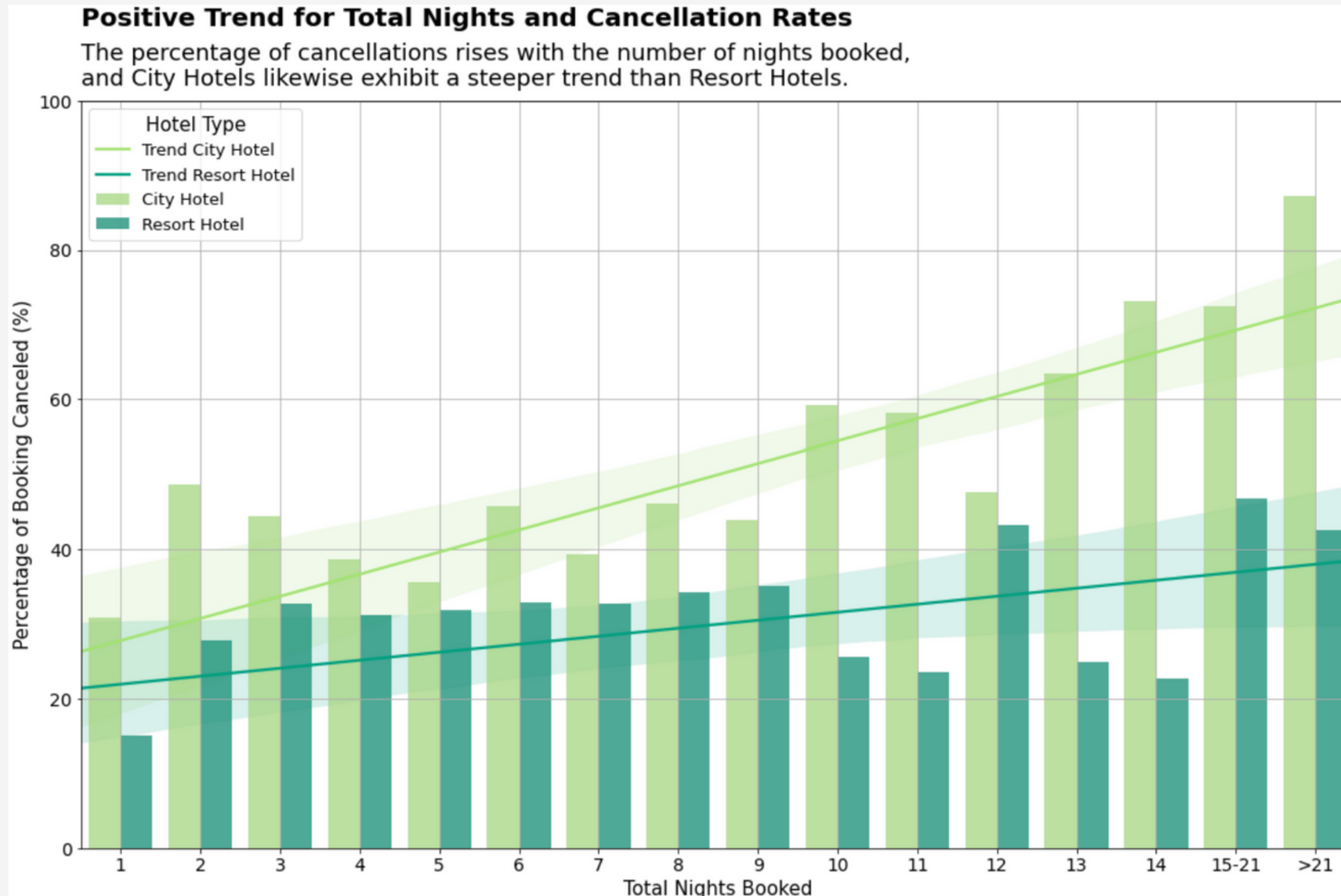
Impact Analysis of Stay Duration on Hotel Bookings Cancellation Rates

Analyze how the correlation between the duration of stay and the hotel cancellation rate of hotel reservations

Insight/analysis

Positive Trend for Total Nights and Cancellation Rates
The percentage of cancellations rises with the number of nights booked, and City Hotels likewise exhibit a steeper trend than Resort Hotels.

This analysis focuses on examining the relationship or trend between the length of stay and the frequency of hotel cancellations.

The longer the stay, the greater the chance that the reservation would be canceled, which is a positive trend for both types of hotels.

Urban hotels exhibit a steeper (significant) positive trend than resort hotels, it can be seen.
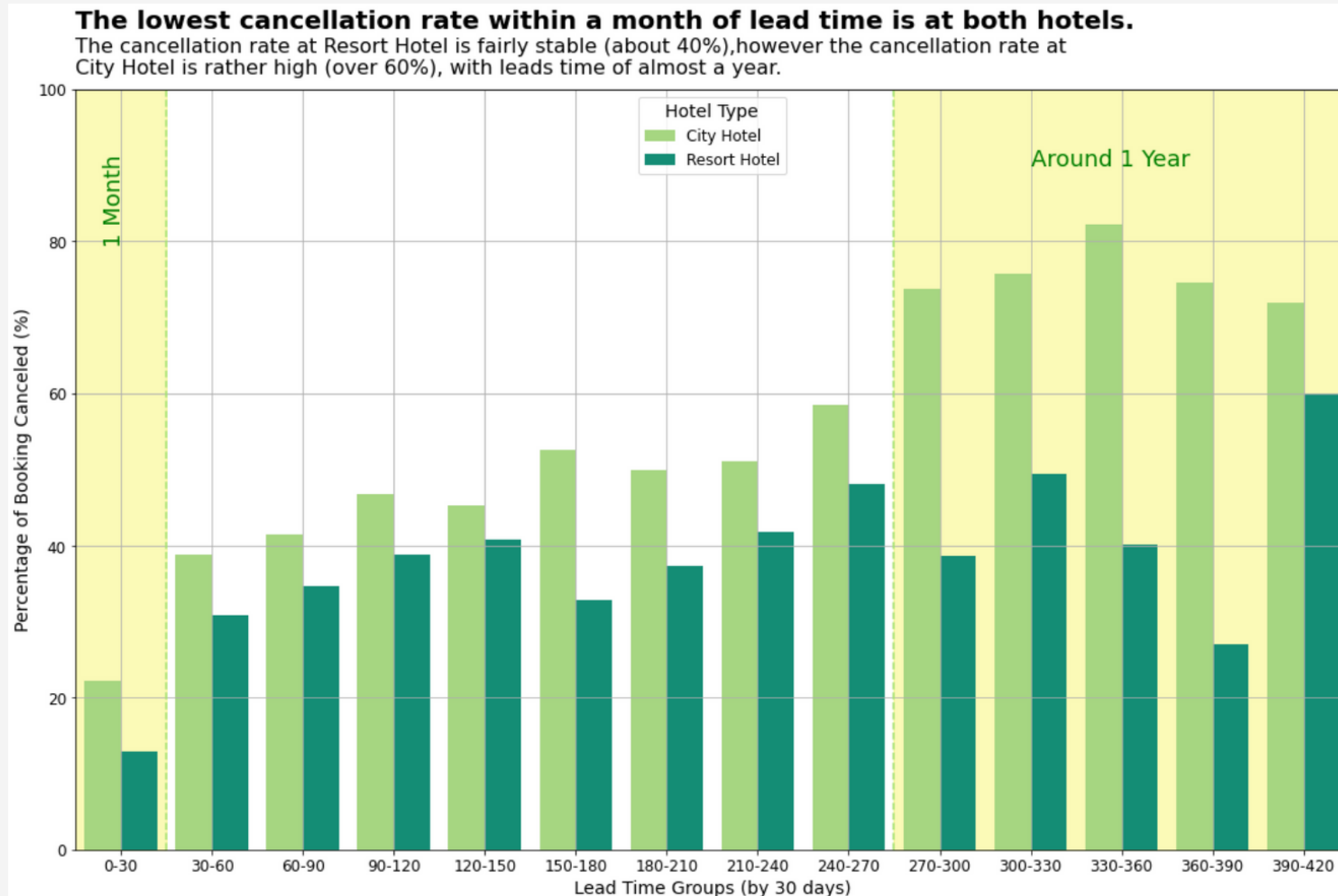
Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

Analyze how the correlation between hotel booking time distance and hotel booking cancellation rates

Insight/analysis

The lowest cancellation rate within a month of lead time is at both hotels.
The cancellation rate at Resort Hotel is fairly stable (about 40%),however the cancellation rate at City Hotel is rather high (over 60%), with leads time of almost a year.

This analysis focuses on examining the relationship or trend between waiting time (the interval between the time of hotel reservation and the time of arrival) and the percentage of hotel reservations that are canceled.

For both categories of hotels, the lowest cancellation rate is for reservations made with a lead time of fewer than 30 days. With a 40% cancellation rate, resort hotels are rather stable, whereas urban hotels have a high rate (60%) when the lead period is close to a year.

Thank You

For more details, you can check out the jupyter notebook here.

www.linkedin.com/in/fadillahakbar