Investigate Hotel Business using Data Visualization

Table of Contents

Download Dataset Here

Data Preprocessing

Data pre-processing:
1. First of all, import all the libraries you need in one cell code. It will simplify when you need to add or edit the libraries instead of import library in separate cell.
2. Load the dataset. Make sure the file path and format (csv/xlsx). I usually save the dataset in the same folder with the .ipynb so I do not need to write completed path to load the data. More effective.
3. Do descriptive statistics. After load the data, I need to look the overview of dataset itself. How many rows and columns they have, what the data type, and etc. It shows the profile of dataset.
4. Handle missing values. I found some columns had null values in descriptive statistics. So, I handle this null values with "0" for numericals and "Unknown" for categoricals.
5. Handle odd values. I found there are odd values in meal columns, so I replace the "Undefined" to "No Meal" as long as the define of this 2 values are same.
6. Handle unnecessary values. I found that some guest are 0 based on the order so in the future I need to analyze the guest > 0 only. I make a copy of dataset onto new variable.

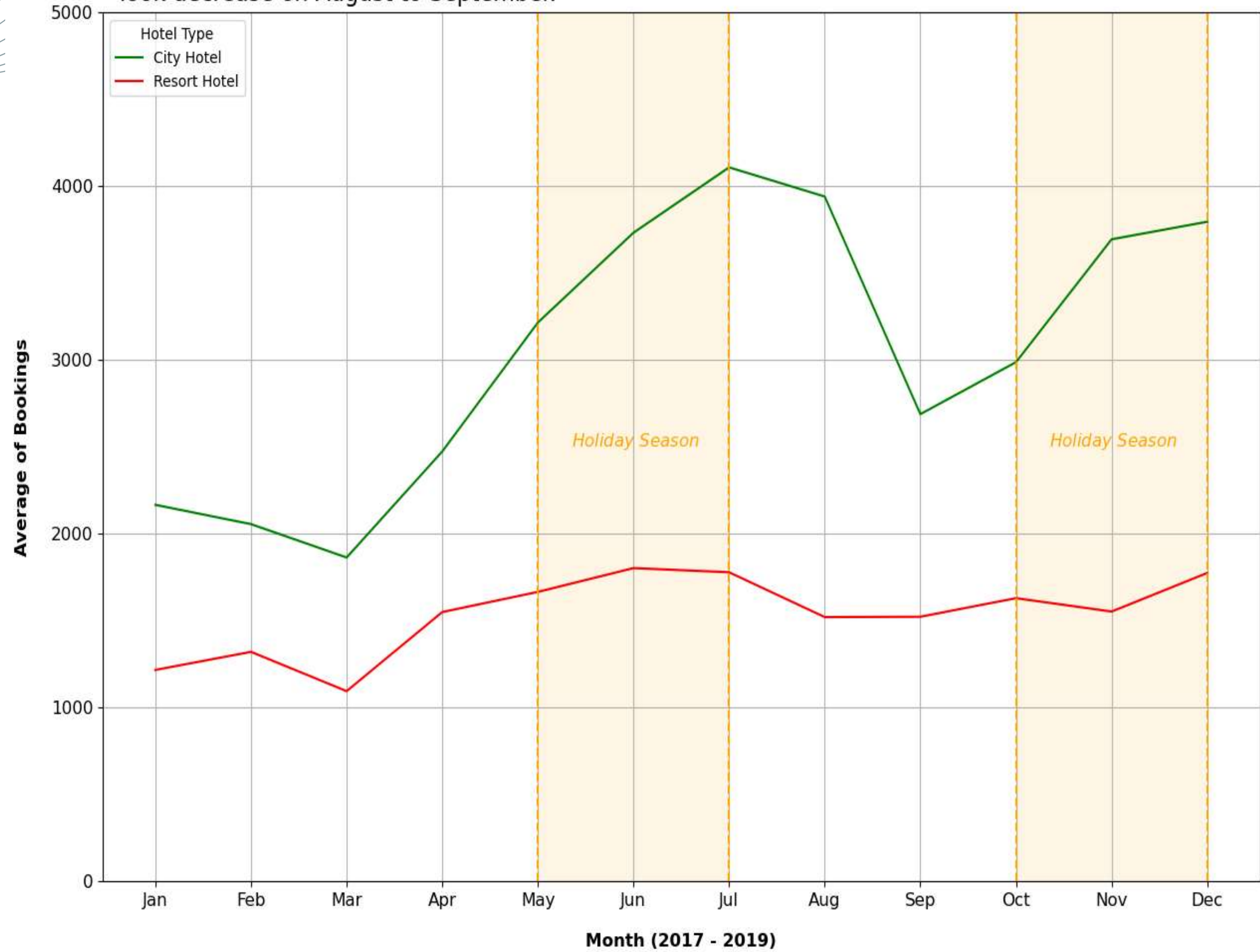For details, see jupyter notebook here
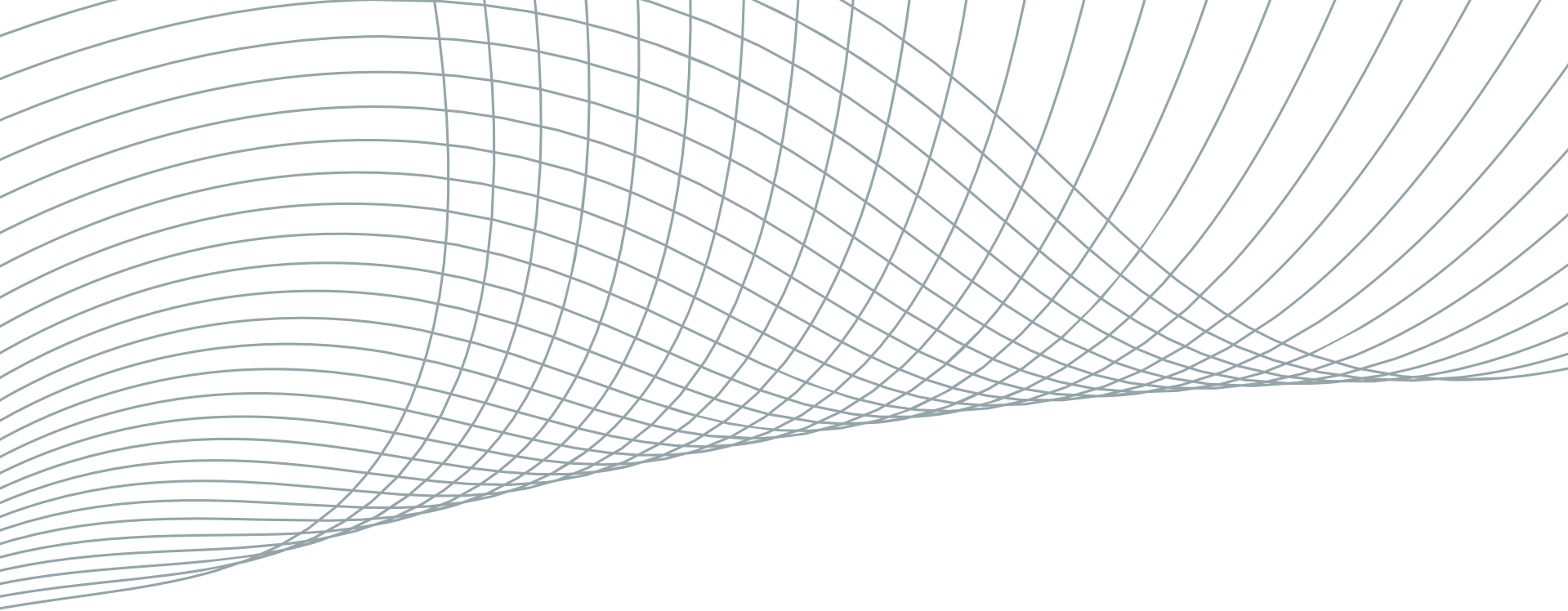
Monthly Hotel Booking Analysis Based on Hotel Type

Interpretation:

• City hotel is the most booking hotel in 3 years (2017, 2018, 2019). However, amount monthly hotel booking of City Hotel looks decrease on August to September.

• Both of hotel type tend to be increase on the holiday season

For details, see jupyter notebook [here](#)

Comparison of Monthly Hotel Booking Based on Hotel Type

Both of hotel type tend to be increase in the holiday season, however amount of City Hotel bookings look decrease on August to September.
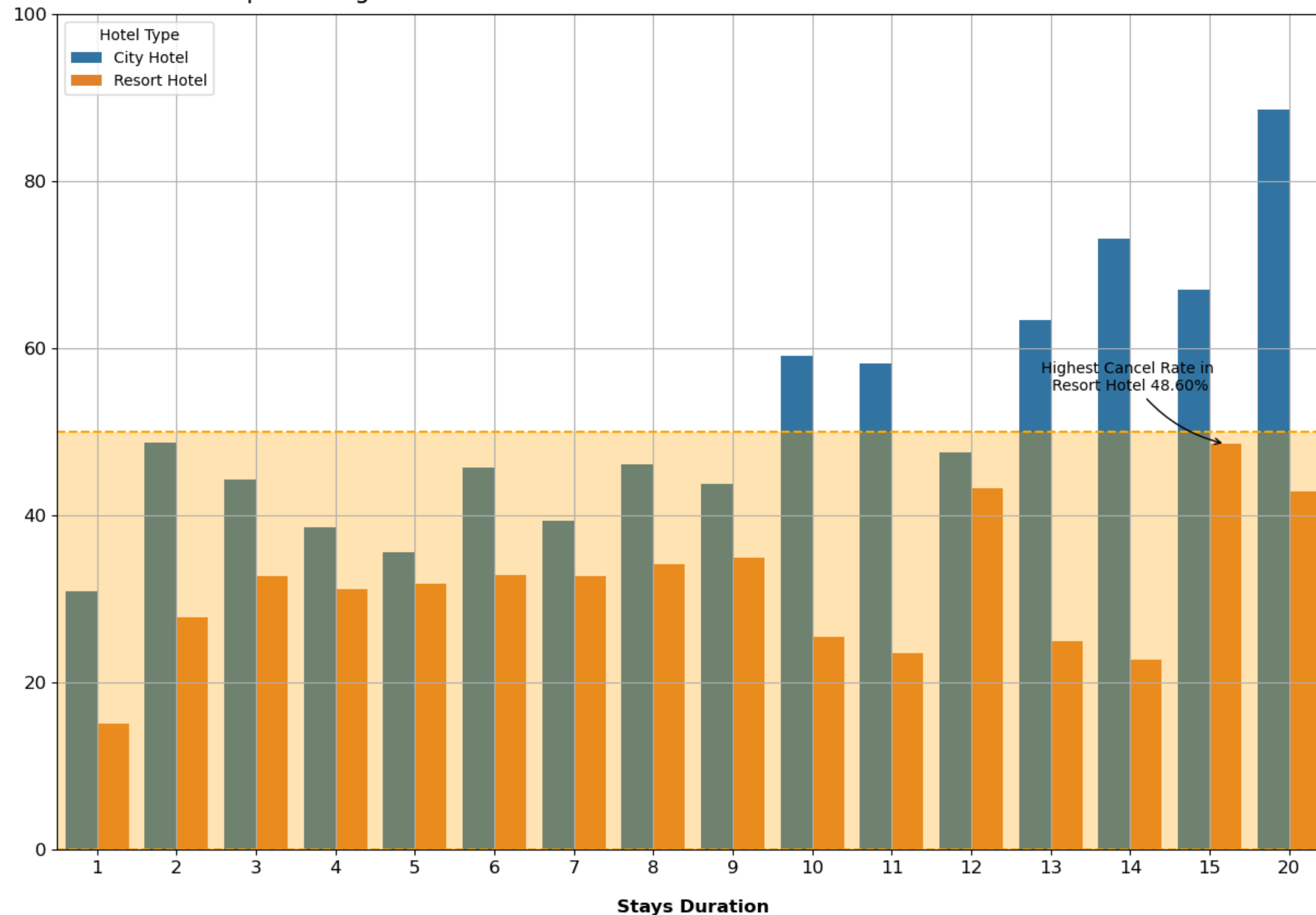
Impact Analysis of Stay Duration on Hotel Bookings Cancellation Rates

MORE LONGER STAYS OF DURATION WILL AFFECT CANCEL BOOKINGS

City Hotel was the highest cancel bookings rate than the Resort Hotel, almost 100%. However, Resort Hotel had the percentage less 50%.

Interpretation:

• More higher stays duration, more higher cancel bookings.

• City Hotel was the highest cancel bookings rate than the Resort Hotel, almost 100%. However, Resort Hotel had the percentage less 50%. Only 48.6%

For details, see jupyter notebook here

# Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

Interpretation:

- Lead time would affect percentage of cancel bookings in City Hotel.

- Resort Hotel had fluctuating distribution lead time towards percentage of cancel bookings

For details, see jupyter notebook [here](#)

BOTH HOTEL TYPE HAD DIFFERENT DISTRIBUTION OF LEAD TIME BASED ON CANCELLATION RATE

Increase of lead time City Hotel will affect to cancellation bookings. However, Resort Hotel had fluctuative distribution
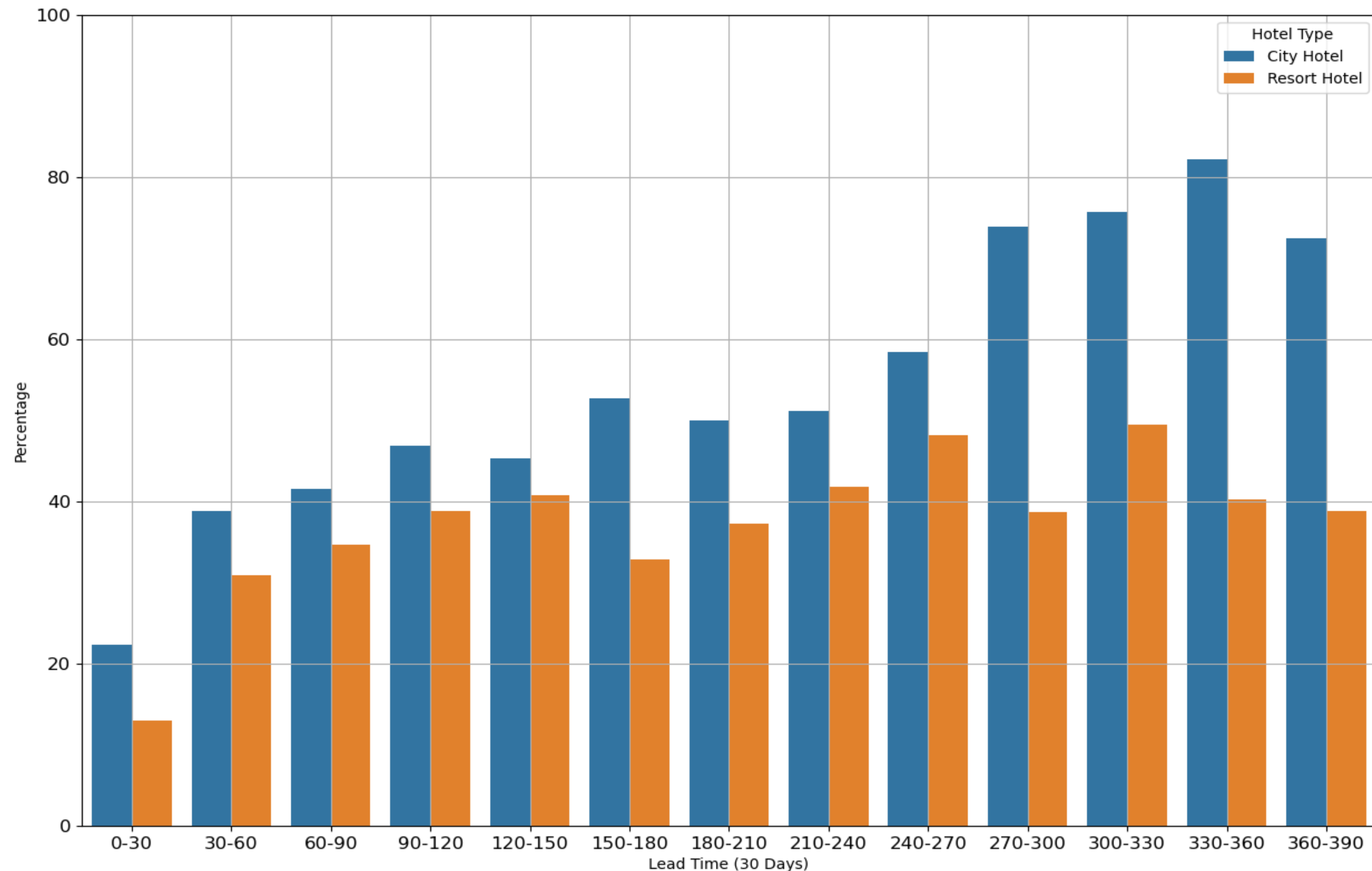
Thank you !

Let's
Connect

 Mustiadi Zakki

 Mustiadi Zakki