

# Investigate Business Hotel using Data Visualization



Created by:

**Dika Kencana Putera**

Let's connect!



dikakencanaputera@gmail.com



<https://www.linkedin.com/in/dikakencanaputera>



<https://github.com/dikaaka>

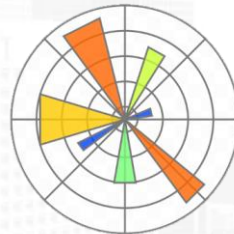
A bachelor with abilities in analyzing and solving problems through fact-based and data-driven decision making which make him proficiency in python, SQL, statistics, machine learning and also had experiences in data analytics and project management.

It is very important for a company to always analyze its business performance. On this occasion, I will explore more business in the hospitality sector. The focus for this project is to find out how hotel's customers behave in making hotel reservations, and how they relate to the cancellation rate of hotel bookings. The results of the insights will be presented in the form of visualization data to make it easier to understand and more persuasive.

## PROGRAMMING LANGUAGE



## DATA VISUALIZATION



*matplotlib*



seaborn

## PYTHON LIBRARY



## NOTEBOOK



```
RangeIndex: 119390 entries, 0 to 119389
Data columns (total 29 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   hotel                  119390 non-null object
 1   is_canceled            119390 non-null int64
 2   lead_time              119390 non-null int64
 3   arrival_date_year      119390 non-null int64
 4   arrival_date_month     119390 non-null object
 5   arrival_date_week_number 119390 non-null int64
 6   arrival_date_day_of_month 119390 non-null int64
 7   stays_in_weekend_nights 119390 non-null int64
 8   stays_in_weekdays_nights 119390 non-null int64
 9   adults                 119390 non-null int64
10   children               119386 non-null float64
11   babies                 119390 non-null int64
12   meal                   119390 non-null object
13   city                   118902 non-null object
14   market_segment         119390 non-null object
15   distribution_channel    119390 non-null object
16   is_repeated_guest       119390 non-null int64
17   previous_cancellations  119390 non-null int64
18   previous_bookings_not_canceled 119390 non-null int64
19   booking_changes         119390 non-null int64
20   deposit_type           119390 non-null object
21   agent                  103050 non-null float64
22   company                 6797 non-null float64
23   days_in_waiting_list    119390 non-null int64
24   customer_type           119390 non-null object
25   adr                     119390 non-null float64
26   required_car_parking_spaces 119390 non-null int64
27   total_of_special_requests 119390 non-null int64
28   reservation_status      119390 non-null object
dtypes: float64(4), int64(16), object(9)
```

## DESCRIPTION

Dataset contains customer behavior features who made hotel bookings

## SHAPE

119.390 data rows, 29 features

## DTYPE

Float64 (4 features), int64 (16 features), object (9 features)

## MISSING VALUE

4 features that has missing value; **company**, **agent**, **city** and **children**

## HANDLE MISSING VALUE



## CORRECTING UNSUITABLE VALUE



## CORRECTING DTYPES



## DROP UNNECESSARY DATA

There are 4 features that has missing value; **company** (94.30%), **agent** (13.68%), **city** (0.40%) and **children** (0.003%).

- Fill **company** with 0, because NaN indicates bookings didn't come from corporate
- Fill **agent** with 0, because Nan indicates bookings didn't use any agent channel
- Fill **city** with 'unknown', because the city origin is unknown
- Fill **children** with 0, because it's median is 0 or indicates that the customer didn't have any children with them

Change 'undefined' to be 'No Meal' in **meal** feature

Change **children**, **agent** and **company** dtype from float to be integer

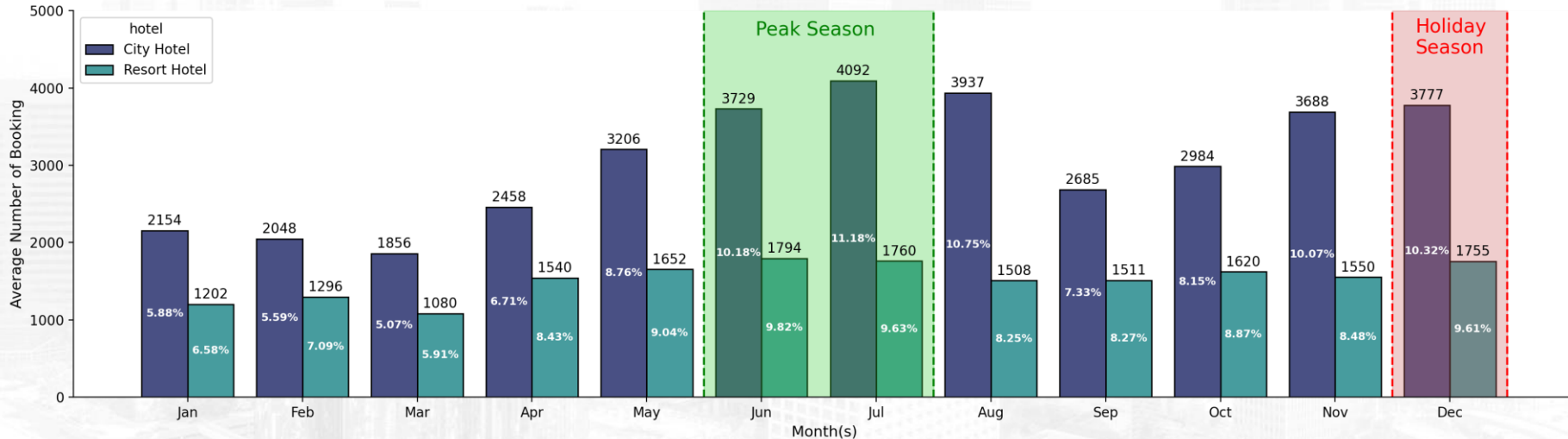
Drop data rows that has 0 **total\_guests** and 0 **stay\_duration**



# Monthly Hotel Booking Analysis Based on Hotel Type

## Average Number of Hotel Bookings per Month Based on Hotel Types

June and July are peak season for hotel bookings. City hotel reached highest average number of hotel bookings at 11.18% on July and resort hotel reached 9.82% on June. There're another growth on average number of hotel bookings in December for city hotel (10.32%) and resort hotel (9.61%). It could be caused by Christmas and New Year's Eve holiday.

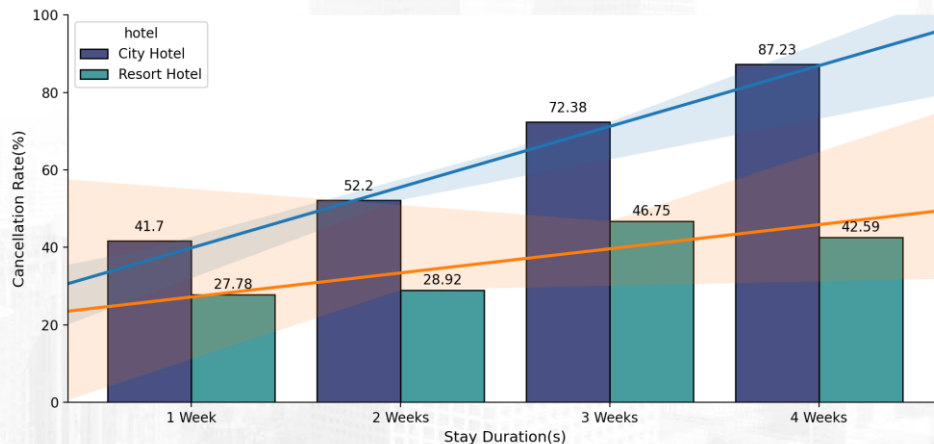


- June and July are peak month of hotel bookings.
- Both hotel types reaches highest average number of booking numbers; city hotel (4.092 or 11.18%) and resort hotel (1.794 or 9.82%). It could be happen because June and July are holiday month for students in Indonesia.
- There is another growth of average number bookings for both hotel types in December; city hotel (3.777 or 10.32%) and resort hotel (1.755 or 9.61%). It could be happen because December is end year's holiday month.

For the detail of my codes, you can see [here](#)

## Positive Trend on Cancellation Rate of Hotel Bookings per Stay Duration Based on Hotel Types

The longer customer stayed, the higher the percentage of booking being canceled.  
The most canceled hotel bookings on city hotel was on four weeks stay duration (87.23%).  
The most canceled hotel bookings on resort hotel was on three weeks stay duration (46.75%).

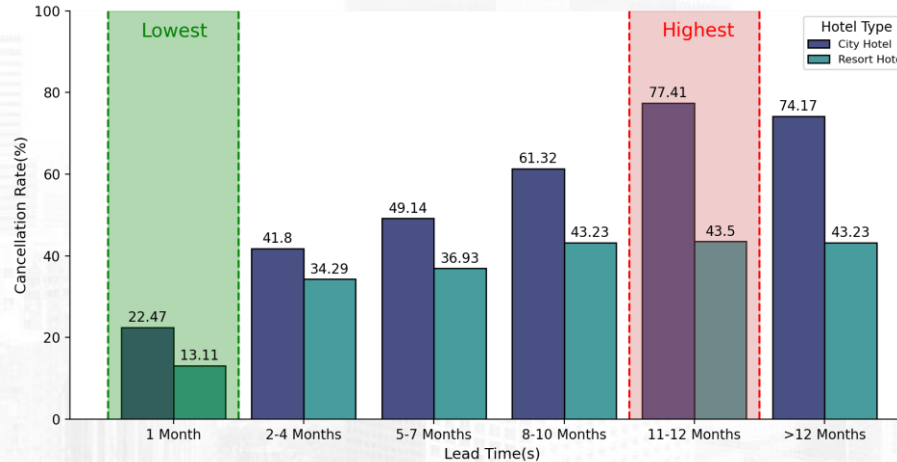


- There is positive trend of cancellation rate of hotel bookings per stay duration based on hotel types.
- The longer customer stayed, the higher cancellation rate for both hotels; city hotel on 4 weeks stay duration (87.23%) and resort hotel on 3 weeks stay duration (46.75%).
- There are a lot possibilities how this thing could be happen, it could be human error when customer did hotel reservation or their long vacation plan canceled because any urgent reasons that they couldn't turn down.

# Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

**Cancellation Rate of Hotel Bookings per Lead Time  
Based on Hotel Types**

Both hotel types has lowest cancellation rate on 1 month lead time (city = 22.47%; resort = 13.11%)  
and highest cancellation rate on 11-12 months lead time (city = 77.41%; resort = 43.5%).



- Both hotel types has lowest cancellation rate of bookings on 1 month lead time; city hotel (22.47%) and resort hotel (13.11%).
- Both hotel types has highest cancellation rate of bookings on 11-12 months lead time; city hotel (77.41%) and resort hotel (43.5%).
- Significant growth of cancellation rate for city hotels each month from around 20% to around 70%, also it happens for resort hotels each month from around 10% to around 40%.
- This growth of cancellation rate could be happen because customer vacation plan canceled.

For the detail of my codes, you can see [here](#)