Investigate Business Hotel using Data Visualization

Created by:
Rahma Yanti
Lets Connect!
rahmars1834@gmail.com
Linkedin.com/in/rahmars

I am an individual that has experienced collecting, organizing, and transforming data to make it easier to understand and actionable insight. Has experience working in government with weather and climate data, spatial data, and water resources data in dams and rivers. Excellent understanding and proficiency with platforms for data analysis such as **Python, SQL, Excel, and ArcGIS.** I recently completed the Data Science Bootcamp Rakamin Academy with a good score and also completed SQL Bootcamp Udemy.

It is very important for a company to always analyze its business performance. On this occasion, we will go deeper into the business in **the hospitality sector**. Our objective is to know how our guests **behave in hotel reservations**, and their relation to **the rate of cancellation of hotel reservations**. The results of the insights we find will be presented in **the form of visualization** data to make it easier to understand and more persuasive.

Tools

Program Language

PYTHON

Data Processing Library

Data Visualization Library

The detailed code can be found here

```
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

This data set contains **booking information** for a city hotel and a resort hotel, and includes others information such as when the booking was made, length of stay, the number of adults, children, and/or babies, etc.

Shape

119390 data rows and **29** features

Data Type

float (4 features), **integer**(16 features), **object**(9 features)

Missing values

 4 features : children, city, company, and agent

The detailed code can be found here

Data Preprocessing

MISSING DATA HANDLING

There are 4 columns with missing value :
1. children : missing values means no children, will replace missing with 0
2. city : missing value, will replace by "Unknown" means no city name
3. agent : missing values mean not booking by agent, will replace missing values with 0
4. company : missing values mean not booking by company, will replace missing values with 0

UNSUITABLE DATA CORRECTING

On meal feature, there is meal type undefined, will replace by "No Meal" because only four Categories are presented in standard hospitality meal packages.

DATA TYPE CORRECTING

for children, agent, and company column change data type from float to integer
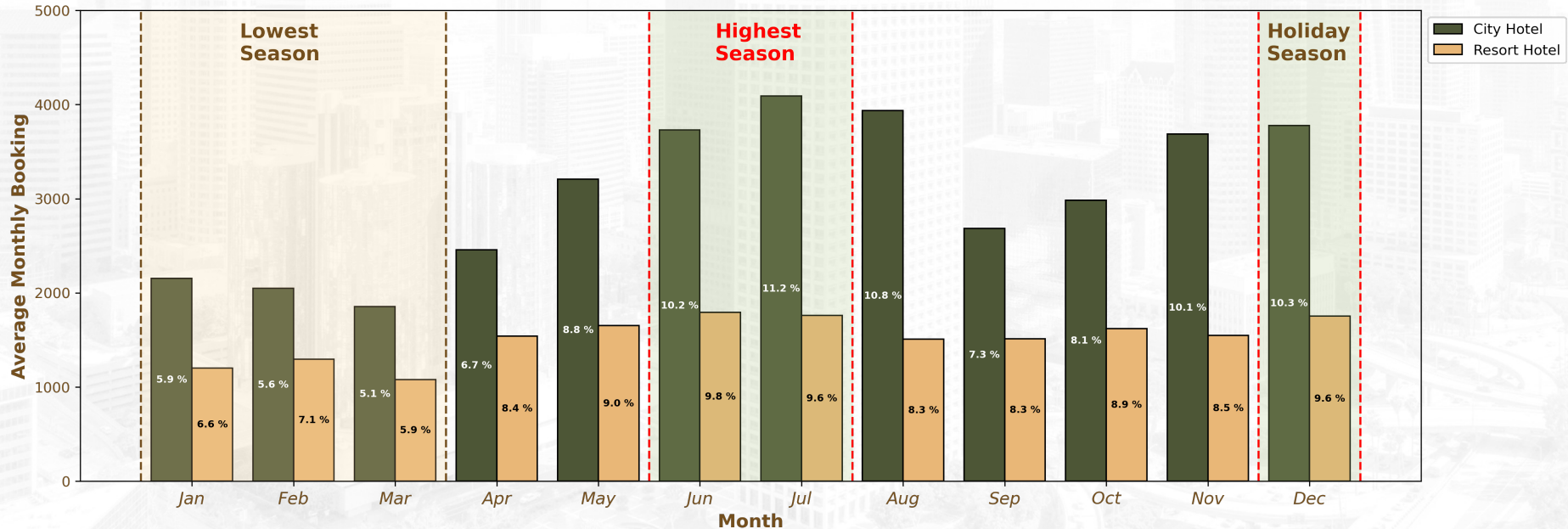
REMOVING UNWANTED DATA

Remove/drop data with zero guest and zero stay
- guest = adults + children + babies
- stay = stays_in_weekend + stays_in_weekday

The detailed code can be found here

Monthly Hotel Booking Analysis Based on Hotel Type

Rakamin Academy

Average Monthly Hotel Reservations Based on Hotel Type

June and July are the highest season for hotel reservations, because is the holiday season for the most schools in Indonesia.
The Resort hotel had the highest reservations at 9.8% in June and in July The City Hotel reached highest reservations at 11.2%.
During christmas Period and new year's eve (Holiday Season), hotel reservation reached at the second highstat at Resort Hotel and City Hotel.
The Lowest season for either type is in January through March.

Lowest Season | Highest Season | Holiday Season

Legend: City Hotel, Resort Hotel

Y-axis: Average Monthly Booking

X-axis: Month

Month	City Hotel	Resort Hotel
Jan	5.9 %	6.6 %
Feb	5.6 %	7.1 %
Mar	5.1 %	5.9 %
Apr	6.7 %	8.4 %
May	8.8 %	9.0 %
Jun	10.2 %	9.8 %
Jul	11.2 %	9.6 %
Aug	10.8 %	8.3 %
Sep	7.3 %	8.3 %
Oct	8.1 %	8.9 %
Nov	10.1 %	8.5 %
Dec	10.3 %	9.6 %

The detailed code can be found here

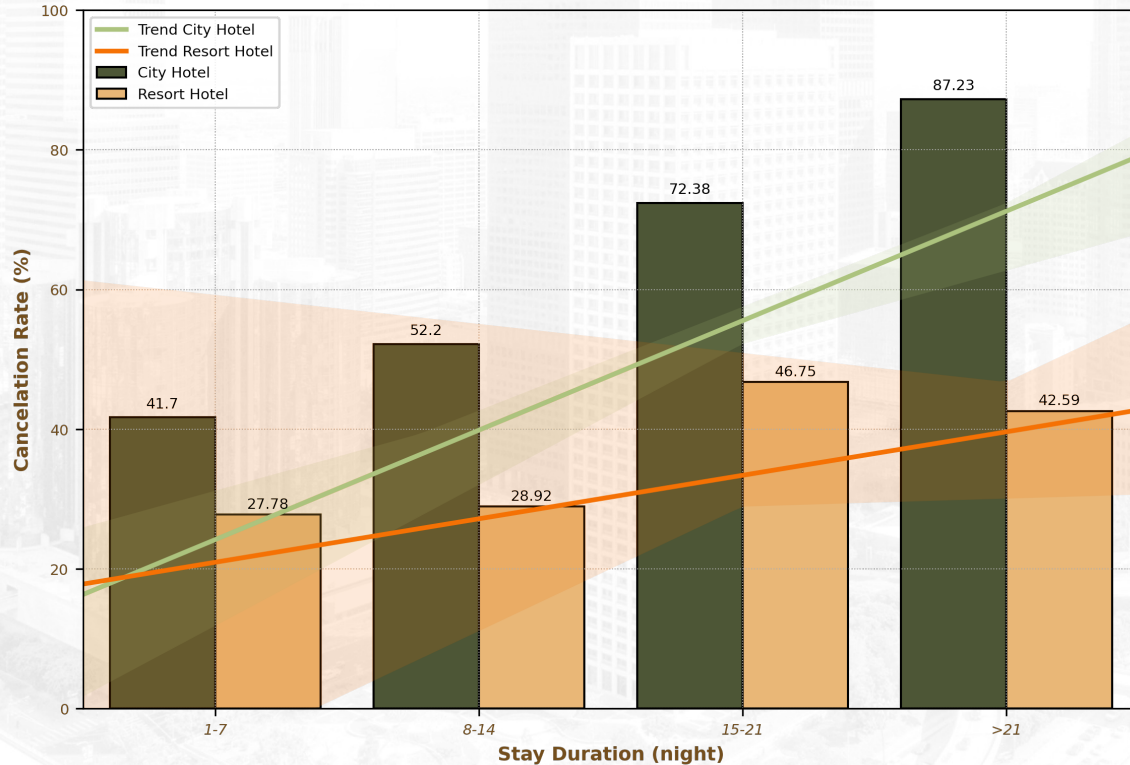# Impact Analysis of Stay Duration on Hotel Bookings Cancellation Rates

Cancellation Rate of Hotel Reservation on Total Nights Based on Hotel Type

Positive trend for all type hotel.
The longer customer stays, the higher percentage of reservation canceled.
The most canceled hotel reservation on City Hotel was on more than 21 nights reservation.
The most canceled hotel reservation on Resort Hotel was on 15-21 nights reservation.

The detailed code can be found here

Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

Rakamin Academy

Cancellation Rate of Hotel Reservation Lead Time Based on Hotel Type

The Lowest cancelation rate for any type hotel is within one month lead time.
The Highest cancelation rate for City hotel type on 11 month lead time, and for Resort hotel type occured on 10 months.
Longer lead time tend to be customer cancelled, customers have likely forgotten that they have booked a hotel if the wait time is too long.
The hotel may issue a reminder and review their cancellation policy to avoid it.

The detailed code can be found here