

Investigate Business Hotel using Data Visualization



Created by:

Dinda Galuh

dindagaluhg@gmail.com

[linkedin.com/in/dinda-galuh-guminta](https://www.linkedin.com/in/dinda-galuh-guminta)

Statistics graduate with passion in data visualization, data analysis, and reporting. Dedicated and hard working person.

I joined the data science bootcamp because I am more interested and have a passion for learning and working in the data field. I like doing data analysis and data visualization. I want to apply my knowledge in the real world.

Supported by:
Rakamin Academy
Career Acceleration School
www.rakamin.com

It is very important for a company to always analyse its business performance. **The intended focus is to find out how customer behaviour in making hotel reservations and its relationship to the level of cancellation of hotel bookings.** The results of the insights found will be presented in the form of visualization data to make it easier to understand & more persuasive.

The dataset consists of

- 119,390 rows
- 29 columns

With research unit : hotel bookings.

0	hotel	15	distribution_channel
1	is_canceled	16	is_repeated_guest
2	lead_time	17	previous_cancellations
3	arrival_date_year	18	previous_bookings_not_canceled
4	arrival_date_month	19	booking_changes
5	arrival_date_week_number	20	deposit_type
6	arrival_date_day_of_month	21	agent
7	stays_in_weekend_nights	22	company
8	stays_in_weekdays_nights	23	days_in_waiting_list
9	adults	24	customer_type
10	children	25	adr
11	babies	26	required_car_parking_spaces
12	meal	27	total_of_special_requests
13	city	28	reservation_status
14	market_segment		

Check & handle
missing values



Check value in
categorical
columns



Remove
unnecessary data

```
# check missing value percentage
df.isnull().sum().sort_values(ascending=False)/len(df)
```

```
company      0.943069
agent        0.136862
city         0.004087
children     0.000034
```

- There are many missing values in the company and agent columns so this feature is excluded because it is not needed in the analysis
- Missing value in the city column is filled with "unknown"
- Missing value in the children column is filled with 0 (zero)

```
Breakfast      91867      Online TA      56403
Dinner         14434      Offline TA/TO  24160
No Meal        10638      Groups       19806
Undefined      1165      Direct       12449
Full Board      798      Corporate    5111
Name: meal, dtype: int64      Complementary  734
TA/TO           97730      Aviation     237
Direct         14483      Undefined     2
Corporate       6491      Name: market_segment, dtype: int64
GDS             193
Undefined       5
Name: distribution_channel, dtype: int64
```

- Replace value "Undefined" to "No Meal" in 'meal' column ([based on external source](#))
- Replace value "Undefined" to "Online TA" in 'market_segment' column (based on mode value)
- Replace value "Undefined" to "TA/TO" in 'distribution_channel' column (based on mode value)

```
# Hint: Perhatikan total jumlah tamu untuk setiap pemesanan
df['total_num_of_guests'] = df['adults']+df['children']+df['babies']
```

- Create a new column: total number of guests
- Deletes rows with a value of 0 (zero) total number of guests

Data with 0 guest: 180 out of 119390 all data
(119210, 28)

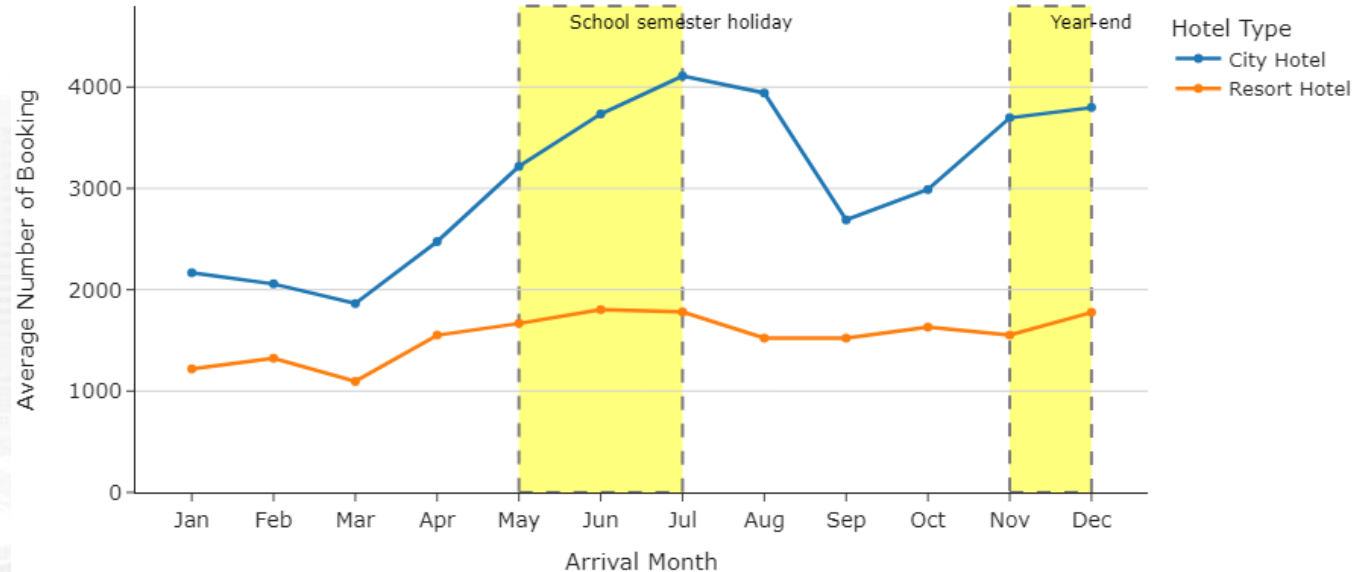
Before preprocessing: Total row : 119,390

After preprocessing: Total row : 119,210

Monthly Hotel Booking Analysis Based on Hotel Type

Peak Booking Season : Hotel Reservations Soar During Holiday!

March saw the fewest bookings overall, with city hotels significantly dropping in reservations during August-September.



The increase in the number of hotel bookings is due to :

- **Eid al-Fitr in Jun.**
- **Christmas and year-end holidays in Dec.**
- **The end of semester student holidays occur in Jun, Jul, and Dec.**

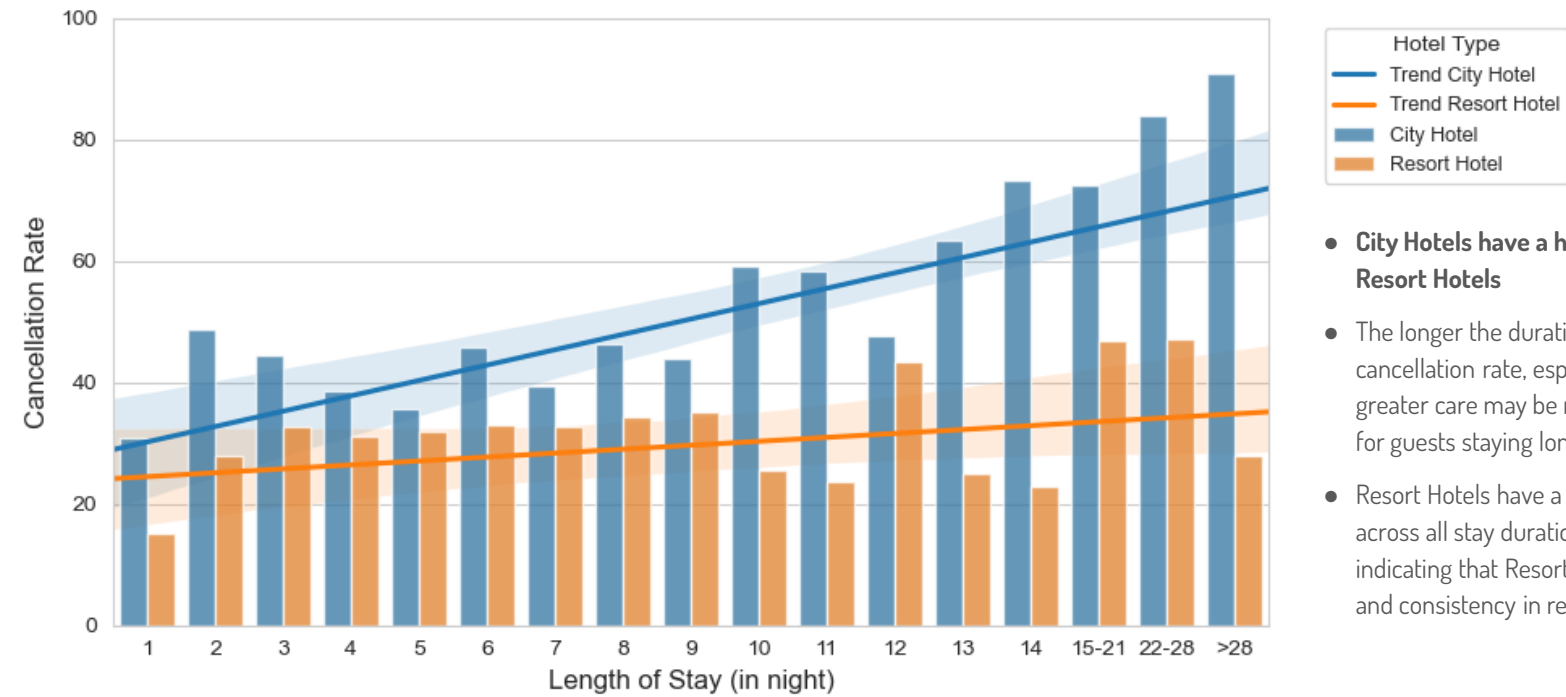
The decrease in the number of hotel bookings in Sep – Oct could be due to weather factors. It is known that September is a transitional season between summer and rainy season in Indonesian territory, thereby reducing the interest of tourists to visit.

- The number of bookings at city hotels is higher than resort hotels allegedly due to a more strategic location and close to the city center or main tourist attractions, making it more accessible for tourists. City Hotels usually have more availability than Resort Hotels, so price competition is tougher at City Hotels. Prices at Resort Hotels are also usually more expensive because they offer a more exclusive & quality experience.
- Many people prefer city hotels because they offer a variety of complete facilities and services such as fitness centers, restaurants, spas, internet access, and business facilities that are suitable for the needs of business travelers and tourists..

Impact Analysis of Stay Duration on Hotel Bookings Cancellation Rates

Prioritizing Long Stay Guests: Unveiling the High Cancellation Rate Trend in Hotel Bookings

Hotel booking cancellation rates based on length of stay and hotel type in Indonesia, 2017-2019

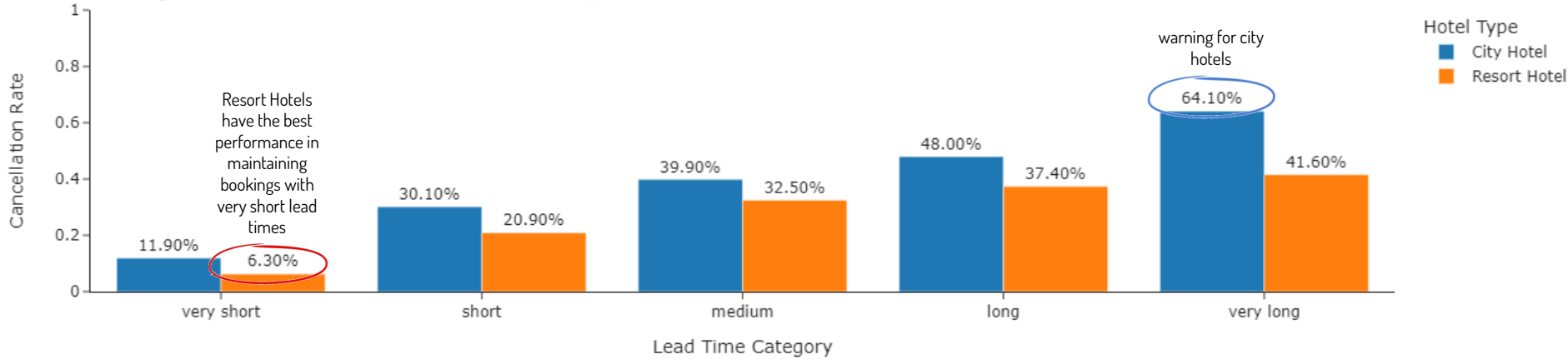


- **City Hotels have a higher cancellation rate than Resort Hotels**
- The longer the duration of the stay, the higher the cancellation rate, especially at City Hotels, so that greater care may be needed in managing bookings for guests staying longer.
- Resort Hotels have a more stable cancellation rate across all stay durations than City Hotels, indicating that Resort Hotels have greater stability and consistency in retaining bookings.

Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

Long Lead Time Increases Cancellation Rate: A Critical Insight in Hotel Bookings

Hotel booking cancellation rates based on lead time and hotel type in Indonesia, 2017-2019



Lead Time (in Days)	Category
<= 6	very short
7 - 30	short
31 - 90	medium
91 - 180	long
> 180	very long

- Orders with very long lead times tend to have a higher cancel rate.** City hotels and resort hotels need to understand cancellation patterns based on lead times and adjust booking management strategies and cancellation policies accordingly.
- The gap cancel rate between city hotels and resort hotels is getting wider as the lead time increases,** with a significant difference occurring in the very long lead time category. This shows that city hotels need to pay attention to cancelling bookings at very long lead times. Meanwhile, Resort Hotels show better performance in maintaining bookings with very short lead times
- The very short lead time category, which is less than a week, shows a lower cancel rate for both types of hotels.** This indicates that orders with that lead time are more stable and not prone to cancellation. City hotels and resort hotels can strengthen their efforts to promote bookings with lead times in this category to increase business stability and profitability.