

# Investigate Business Hotel using Data Visualization



**Created by:**

**Arief Mis Hargi**

[arief.299@gmail.com](mailto:arief.299@gmail.com)

<https://github.com/amishargi>

<https://linkedin.com/in/amishargi/>

“Professional Admin Support with experience in the banking industry. Wanted to bring this experience to the data role. Have expertise in analytical skills, planning and passionate in solving Data Analytics / Data Science problems using Python and SQL. .”

“Sangat penting bagi suatu perusahaan untuk selalu menganalisa performa bisnisnya. Pada kesempatan kali ini, kita akan lebih mendalami bisnis dalam bidang perhotelan. Fokus yang kita tuju adalah untuk mengetahui bagaimana perilaku pelanggan kita dalam melakukan pemesanan hotel, dan hubungannya terhadap tingkat pembatalan pemesanan hotel. Hasil dari insight yang kita temukan akan kita sajikan dalam bentuk data visualisasi agar lebih mudah dipahami dan bersifat lebih persuasif.”

RangeIndex: 119390 entries, 0 to 119389

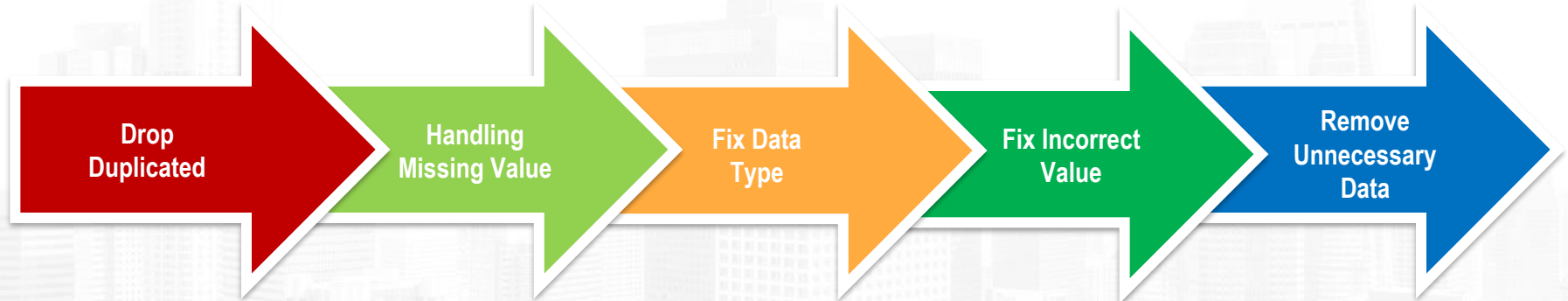
Data columns (total 29 columns):

#	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)

- Data terdiri atas 119.390 baris, dan terdiri atas 29 atribut.
- Atribut terdiri atas 9 data bertipe object, 4 data bertipe float64 dan 16 data bertipe integer.
- Terdapat 4 atribut yang memiliki nilai null seperti children, city, agent, dan company.

[Untuk selengkapnya, dapat melihat jupyter notebook disini](#)



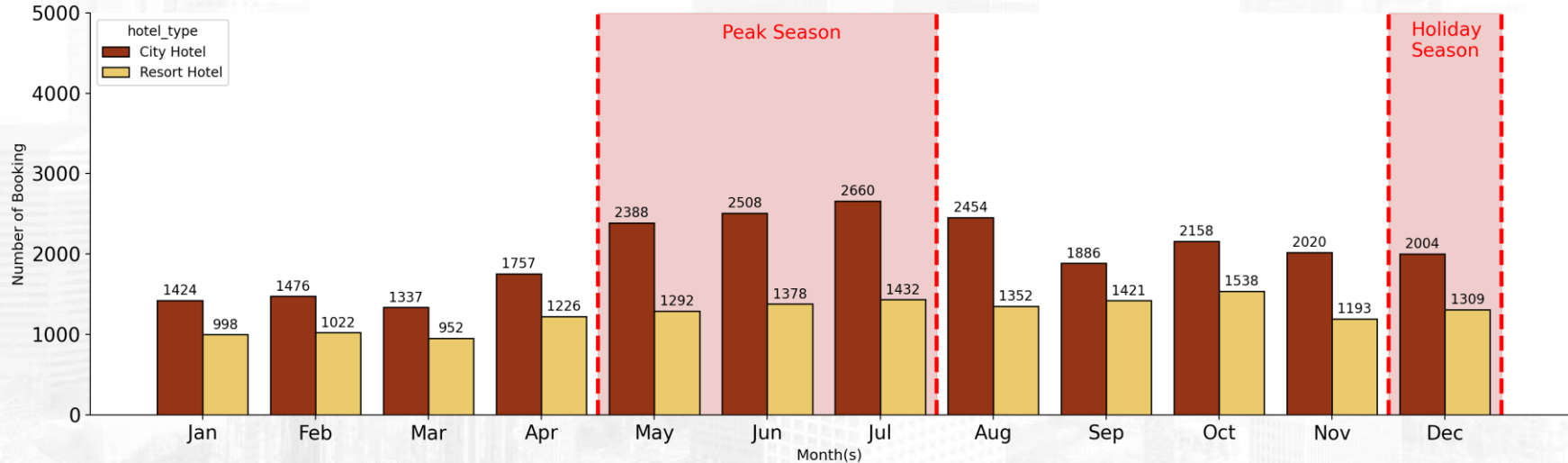
- Menghapus data duplikat sebanyak 33261
- Mengubah nilai null pada feature children dengan nilai 0
- Mengubah nilai null pada feature city dengan nilai 'Unknown'
- Mengubah nilai null pada feature agen dan company dengan nilai 0 yang berarti tidak melakukan pesanan melalui agen
- Mengubah tipe data pada feature children, agen dan company menjadi integer
- Mengubah nilai undefined pada feature meal menjadi No Meal
- Menghapus data dengan nilai tamu 0 dan jumlah malam 0

[Untuk selengkapnya, dapat melihat jupyter notebook disini](#)

# Monthly Hotel Booking Analysis Based on Hotel Type

Number of Guest by Hotel Type

May - July is the peak of holiday season. City hotel and Restort Hotel gets an increas of guest during peak season. While both hotels have less customer during not holiday (Jan-Mar).



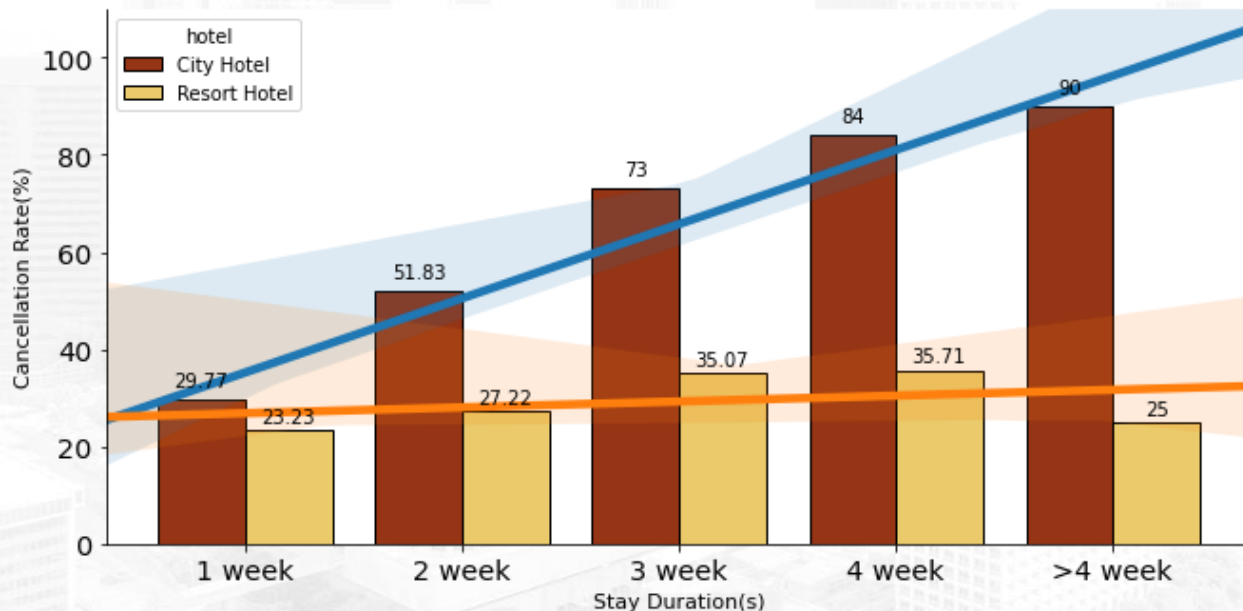
- Bulan Mei sampai Juli menjadi puncak dari liburan, dimana terdapat libur lebaran dan libur kenaikan kelas di Indonesia. Puncak kenaikan pada libur lebaran berpengaruh signifikan pada kenaikan *booking City Hotel*. Hal ini dikarenakan konsep *City Hotel* yang hanya sebagai tempat transit sementara bagi para tamu yang ingin melanjutkan perjalanan untuk pulang kampung halaman.
- Pada saat libur natal dan tahun baru *Restort Hotel* mengalami kenaikan yang cukup signifikan dibanding bulan sebelumnya, dikarenakan kebanyakan para pelancong ingin menikmati liburan yang jauh dari hiruk pikuk di ibukota.
- Penerimaan jumlah tamu diluar musim liburan cenderung stabil , dan pada periode Maret justru cenderung terjadi penurunan jumlah pengunjung baik *City Hotel* dan *Resort Hotel*.

[Untuk selengkapnya, dapat melihat jupyter notebook disini](#)



## Trends in Hotel Bookings Cancellation Rate per Stay Duration Based on Hotel Type

The longer customer stayed, the higher the percentage of booking being canceled  
The most canceled hotel bookings on city hotel was on more than four weeks stay duration (90%).  
The most canceled hotel bookings on resort hotel was on four weeks stay duration (35.71%)



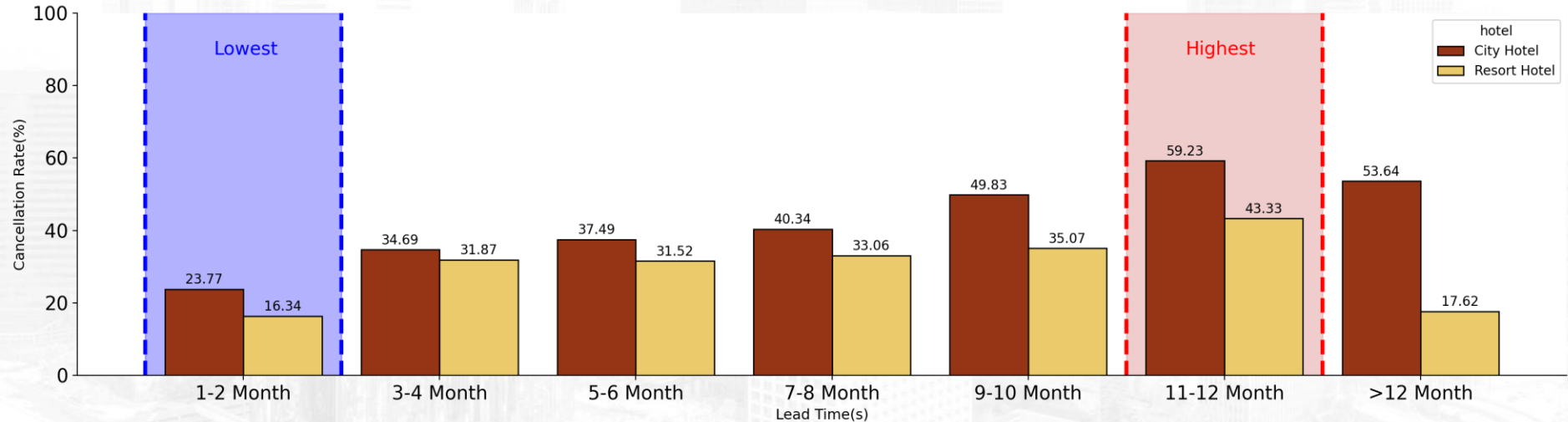
Semakin lama waktu menginap tamu maka semakin tinggi juga rasio Hotel Booking akan di batalkan, khususnya pada *City Hotel*, tamu yang menginap lebih dari 4 minggu memiliki *cancellation Rate* sebesar 90%.

Hal ini dapat dijadikan pertimbangan bagi pengusaha untuk memberikan pinalti kepada calon tamu yang membatalkan pesanannya, sehingga *Cancellation Rate* dapat diturunkan

# Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

## Cancellation Rate of Booking Hotel per Lead Times Based on Hotel Type

*Both of two Hotel Type has the lowest cancellaction ratio on 1-2 Month lead time and the highest on 11-12 Month lead time*



- Rasio pembatalan pemesanan hotel terendah terjadi pada rentang waktu tunggu 1 – 2 bulan sebelum waktu menginap, yaitu pada City Hotel sebesar 23.77% dan Resort Hotel sebesar 16.34%.
- Sementara itu rasio pembatalan pemesanan hotel tertinggi terjadi pada rentang waktu tunggu 11 – 12 bulan sebelum waktu menginap, untuk rasio pembatalan pada City Hotel sebesar 59.23% dan Resort Hotel sebesar 43.33%.
- Perubahan rencana berlibur dari calon tamu dapat meningkatkan kemungkinan rasio pembatalan pemesanan.