

# Investigate Business Hotel using Data Visualization



**Created by:**

**Fawwaz Nurmansyah**

[fawwaznurmansyah@gmail.com](mailto:fawwaznurmansyah@gmail.com)

[linkedin.com/in/fawwaz-nurmansyah/](https://www.linkedin.com/in/fawwaz-nurmansyah/)

“A graduate Bachelor of Biology from Airlangga University who is interest in pursuing a career related to data after completing some bootcamp. Able to work well independently and as part of a team, even under tight deadlines and pressure.

Have less than a year experience in data science with understanding of data processing, data analysis, and machine learning. Proficient in extracting insights from complex datasets and transforming them into profitable recommendations for business.”

“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.”

```
1 # Read data .csv
2 df = pd.read_csv('hotel_bookings_data.csv')
3 df.head(5)
4
5 ## Mengatasi Null Values
6
7 # Mengisi null value yang ada dalam kolom 'city' menjadi 'Other'.
8 # diasumsikan bahwa pelanggan tersebut berasal tidak dari Kabupaten atau Kota yang disediakan pada kolom 'city'
9 df2['city'] = df2['city'].fillna('Other')
10 # Mengisi null value yang ada dalam kolom 'agent' dan 'company' menjadi nilai 0
11 df2['agent'] = df2['agent'].fillna(0)
12 df2['company'] = df2['company'].fillna(0)
13 # Mengasumsikan bahwa null value yang terdapat pada 'children' merupakan pelanggan tanpa membawa anak
14 df2['children'] = df2['children'].fillna(0)
15 # Cek kembali untuk memeriksa apakah ada null values yang masih tersisa
16 df2.isnull().sum()
17
18 ## Memperbaiki datatype yang tidak sesuai
19
20 df2['children'] = df2['children'].astype('int64')
21 df2['agent'] = df2['agent'].astype('int64')
22 df2['company'] = df2['company'].astype('int64')
23
24 ## Memperbaiki value yang tidak sesuai
25
26 # Value 'Undefined' pada kolom 'meal' diasumsikan sebagai pelanggan yang memiliki value 'No Meal'
27 df2['meal'] = df2['meal'].replace(['Undefined'], 'No Meal')
28 df2['meal'].value_counts()
29
30 ## Drop data yang tidak diperlukan
31
32 # Menghapus data yang tidak diperlukan sesuai dengan jumlah 0 guest dan 0 night
33 df2['total_guests'] = df2['babies'] + df2['children'] + df2['adults']
34 print('Data dengan 0 tamu: {0} / {1}'.format(df2[df2['total_guests'] == 0].shape[0], df2.shape[0]))
35 df2['stay_duration'] = df2['stays_in_weekend_nights'] + df2['stays_in_weekdays_nights']
36 print('Data dengan 0 malam: {0} / {1}'.format(df2[df2['stay_duration'] == 0].shape[0], df2.shape[0]))
37 # Drop data yang memiliki nilai 0 pada 'total_guests' dan 'stay_duration'
38 df3 = df2[(df2['total_guests'] > 0) & (df2['stay_duration'] > 0)]
```

- Handling Null Values

Null values yang tersedia ada pada kolom 'city', 'agent', 'company', dan 'children'. Missing values pada 'city' akan diisi 'Other' dan lainnya diisi dengan value '0'.

- Fixing data types

'children', 'agent', dan 'company' merupakan datatype float di dataset. Oleh karena itu diubah menjadi 'int64'.

- Fixing incorrect values

Nilai 'Undefined' pada kolom 'meal' tidak sesuai dengan klasifikasi yang ada pada kolom tersebut. Solusinya, value 'Undefined' diubah menjadi 'No Meal' karena diasumsikan pelanggan tidak memesan makanan.

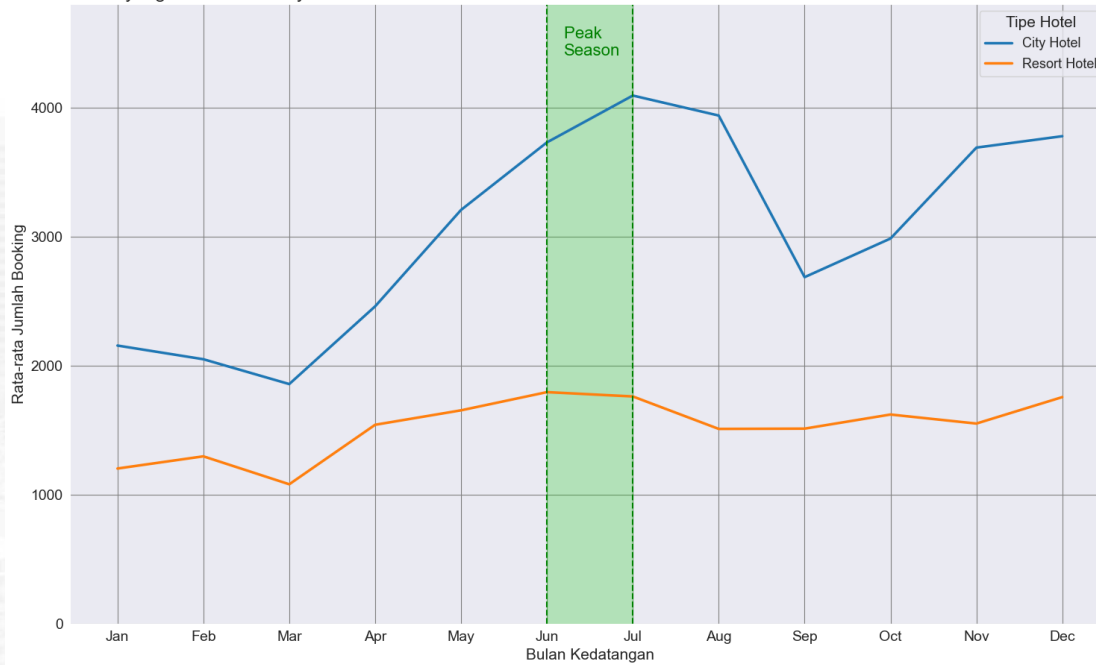
- Drop unnecessary data

Data dengan jumlah 0 guest dan 0 inap malam akan dihapus karena data tersebut tidak dibutuhkan.

# Monthly Hotel reservasi Analysis Based on Hotel Type

## Bulan yang ramai untuk penginapan hotel

Bulan Juni merupakan bulan yang ramai untuk Resort Hotel sedangkan pada Bulan Juli merupakan bulan yang ramai untuk City Hotel.

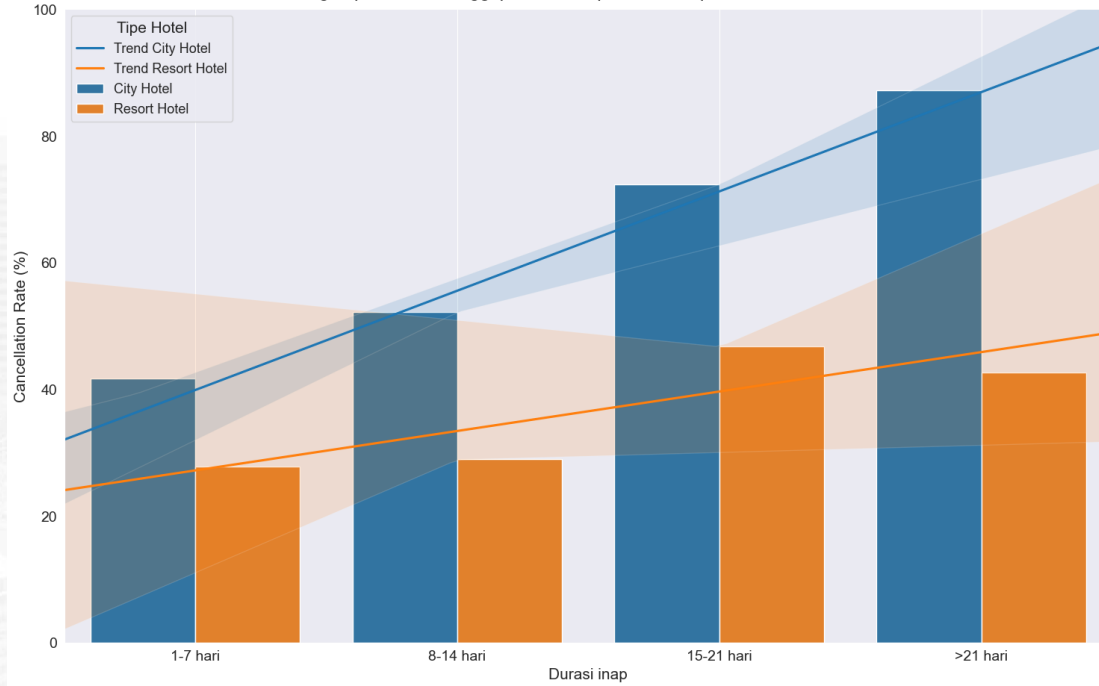


Pada rentang bulan Juni - Juli ini biasanya merupakan hari libur bagi para pelajar sehingga banyak pelanggan yang memesan reservasi hotel di bulan tersebut terutama pemesanan hotel di kota.

# Impact Analysis of Stay Duration on Hotel reservasis Cancellation Rates

## Perkembangan Positif pada Jumlah Malam dan Tingkat Persentase Pembatalan

Semakin lama total waktu menginap, semakin tinggi persentase pembatalan pemesanan.



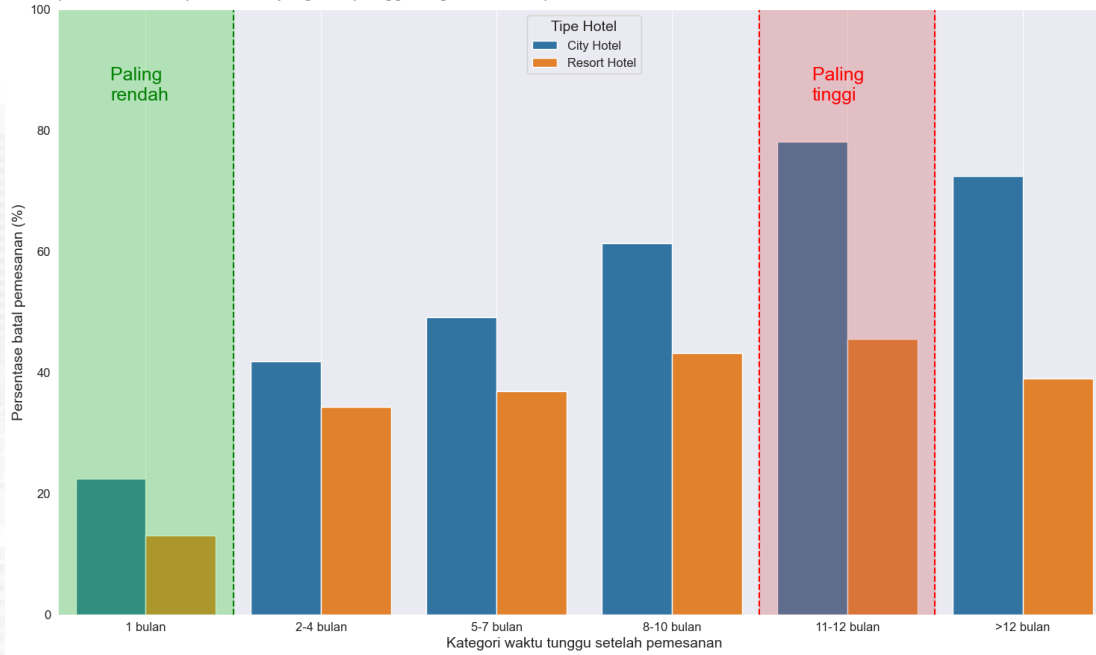
Pada grafik yang disediakan, ditunjukkan bahwa semakin lama durasi pelanggan reservasi hotel, maka semakin tinggi persentase pembatalan reservasi hotel. Hal ini bisa dilihat dengan grafik garis trend kedua tipe hotel.

Untuk mencegah terjadinya pembatalan reservasi hotel dengan durasi yang lama, hotel harus menerapkan kebijakan dimana semakin lama durasi pelanggan reservasi hotel, maka semakin mahal biaya pembatalan reservasi hotelnya.

# Impact Analysis of Lead Time on Hotel reservasis Cancellation Rate

Kedua tipe hotel memiliki tingkat batal pemesanan hotel yang rendah dengan rentang lead time 1 bulan

Hotel resort memiliki persentase batal pemesanan yang stagnan, sedangkan pada hotel kota memiliki persentase batal pemesanan yang cukup tinggi dengan lead time pemesanan sekitar 11-12 bulan.



Pada grafik yang disediakan, ditunjukkan bahwa semakin lama hotel resort dan hotel kota memiliki tingkat pembatalan yang tinggi dengan lead time sekitar 11-12 bulan. Hal ini bisa disebabkan oleh pelanggan yang mungkin lupa sudah memesan hotel tersebut jika lead time terlalu lama.

Untuk mencegah terjadinya hal tersebut, reservasi hotel harus memberi pengingat pada pelanggan agar mereka tidak membatalkan reservasi. Menerapkan kebijakan pembatalan reservasi juga bisa mencegah hal ini terjadi.