

Investigate Business Hotel using Data Visualization



Created by:

Dwi Muhammad Nurafli

dwirafli13@gmail.com

[Dwi Muhammad Nurafli](#)

“An experienced employee who is proficient in using Microsoft Word, and Microsoft Excel with deep interest in data related role ”

“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.”

- Yang dilakukan pertama adalah membaca datanya agar bisa di olah di jupyter notebook.

```
data = pd.read_csv('hotel_bookings_data.csv')
data.head()
```

- Mengecek kesesuaian tipe data dan mengecek apakah ada data yang null atau tidak

```
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)
memory usage: 26.4+ MB
```

- Untuk semua tipe data sudah sesuai
- Untuk data yang null saya melakukan proses sebagai berikut :
 - Menghapus kolom **agent** dan **company** karena jumlah data yang null terlalu banyak, jika hanya menghapus baris pada kolom tersebut takutnya akan menghilangkan banyak informasi dari data.
 - Mengisi nilai null pada kolom **city** dengan value *unknown* karena tidak diketahui asal kota dari customer
 - Menganggap nilai null pada kolom **children** berarti tidak ada tamu anak kecil. Oleh karena itu diisi dengan nilai 0(zero).

- Mengganti value 'Undefined' menjadi 'No Meal' pada kolom **meal** karena sebetulnya artinya sama.

```
data['meal'].unique()  
  
array(['Breakfast', 'Full Board', 'Dinner', 'No Meal', 'Undefined'],  
      dtype=object)
```

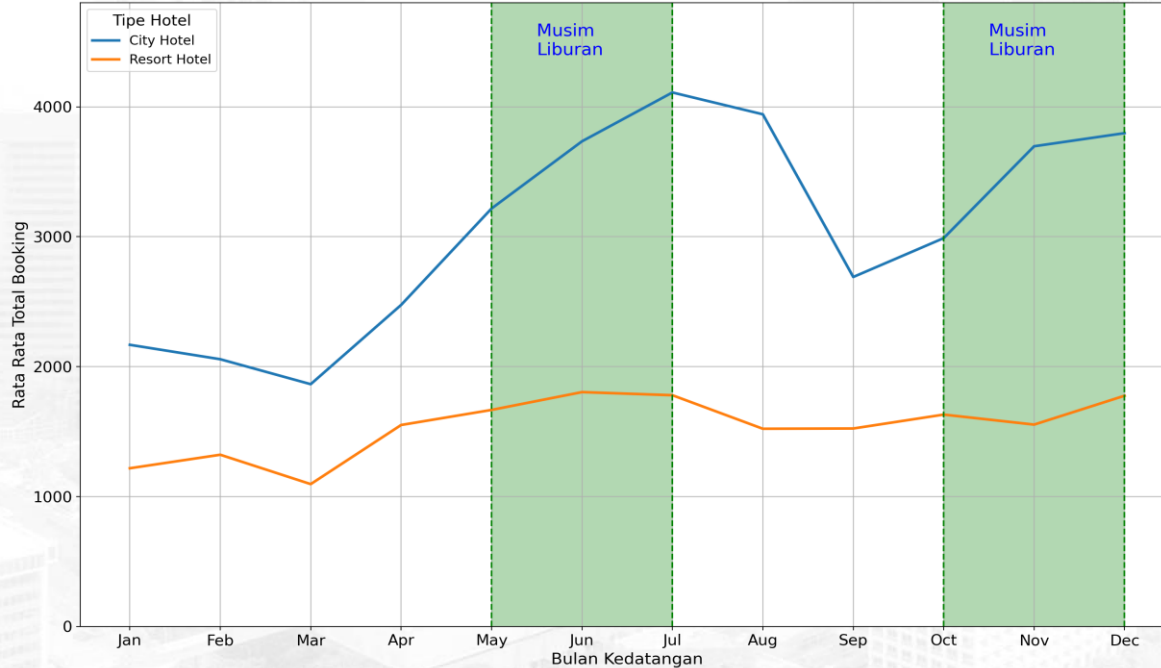
- Menghapus data dengan jumlah tamu yang 0(zero).

```
data_prep['guest'] = data_prep['adults'] + data_prep['children'] + data_prep['babies']  
  
data_prep['guest'].unique()  
  
array([ 2.,  1.,  3.,  4.,  5., 12., 40., 26., 50., 27., 55.,  0., 20.,  
        6., 10.])
```

Monthly Hotel Booking Analysis Based on Hotel Type

Kedua Tipe Hotel Memiliki Lebih Banyak Pengunjung Pada Musim Liburan

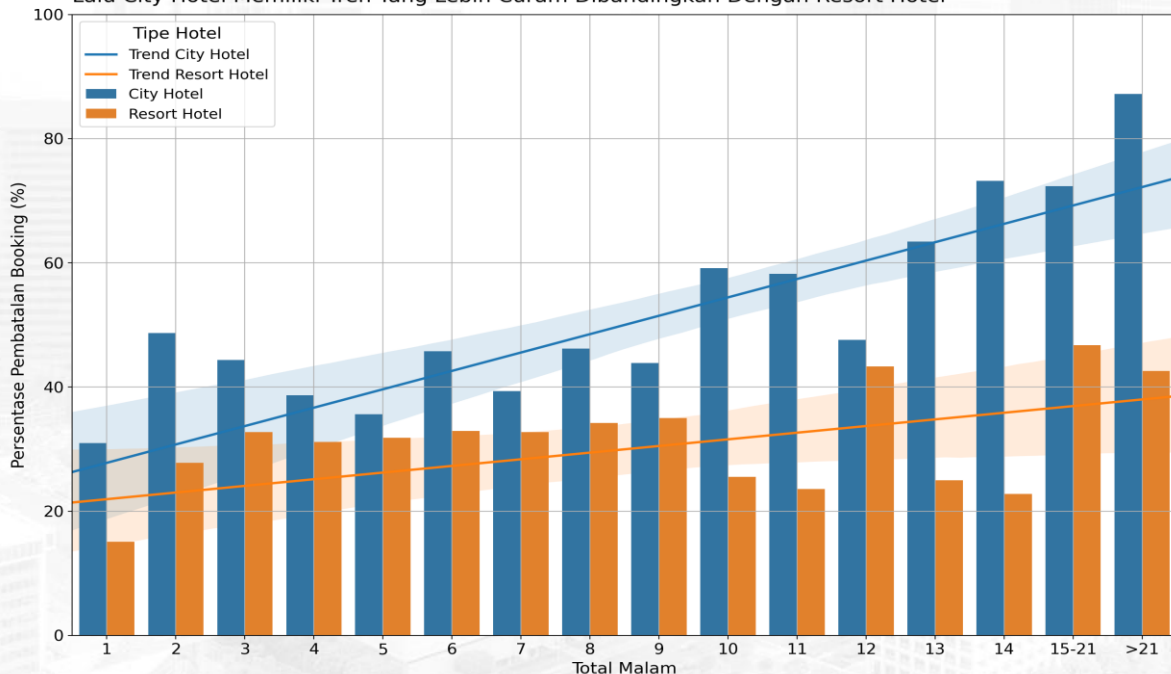
City Hotel Mengalami Penurunan Jumlah Pengunjung Yang Drastis Di Bulan Agustus Ke September, Sedangkan Kedua Hotel Memiliki Lebih Sedikit Pelanggan Selama Bukan Musim Liburan (Jan-Mar)



- Untuk kedua hotel baik *City Hotel* maupun *Resort Hotel* memiliki pengunjung yang lebih banyak pada saat musim liburan
- Ternyata *City Hotel* memiliki lebih banyak pengunjung di bandingkan dengan *Resort Hotel*
- *City Hotel* mengalami penurunan jumlah pengunjung yang drastis pada rentang bulan Agustus-September

Tren Positif Pada Total Malam Dan Persentase Dari Pembatalan Pesanan

Semakin Lama Total Malam Yang Dipesan Maka Persentase Pembatalan Juga Akan Semakin Tinggi, Lalu City Hotel Memiliki Tren Yang Lebih Curam Dibandingkan Dengan Resort Hotel

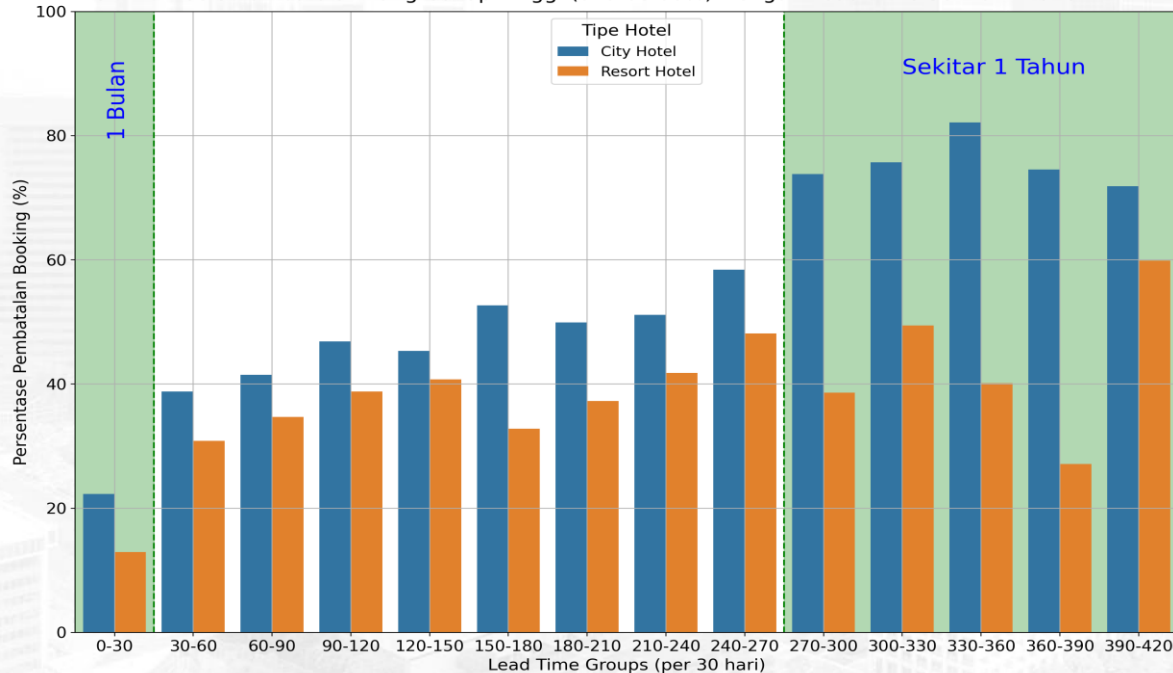


- Semakin lama total malam pada pesanan tamu maka persentase di batalkannya pesanan juga akan semakin tinggi
- *City Hotel* memiliki tren yang lebih curam jika di bandingkan dengan *Resort Hotel*

Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate

Kedua Hotel Memiliki Rasio Pembatalan Paling Rendah Dalam 1 Bulan Lead Time

Resort Hotel Memiliki Rasio Pembatalan Yang Cukup Stagnan (sekitar 40%), Sedangkan City Hotel Memiliki Rasio Pembatalan Yang Cukup Tinggi (di atas 60%) Dengan Lead Time Sekitar 1 Tahun



- Kedua hotel, baik *City Hotel* maupun *Resort Hotel* memiliki rasio pembatalan yang paling rendah pada Lead Time 1 Bulan
- Pada Lead Time sekitar 1 Tahun, *Resort Hotel* memiliki rasio pembatalan yang cukup stagnan yaitu sekitar 40%, sedangkan *City Hotel* memiliki rasio pembatalan yang cukup tinggi yaitu diatas 60%