

Investigate Business Hotel using Data Visualization



Created by:

Feri Dwi Saputro

feridwisa95@gmail.com

[linkedin.com/in/ferids55](https://www.linkedin.com/in/ferids55)

Saya seorang **mantan Electrical Staff** yang telah bekerja selama 3 tahun di PLTU dan kini ingin mengubah jalur karir ke bidang data karena lebih sesuai dengan passion saya yaitu analisis dan perhitungan numerik. Tools yang biasa saya gunakan adalah **SQL dan Python** untuk analisis data dan pemodelan. Saya tertarik pada penerapan data science dalam **bidang retail, marketing, atau mungkin esport**.

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

“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.”

Hotel Bookings Dataset

119390
rows

29
columns

Tahapan Preprocessing yang dilakukan:

1. Check Missing Values
Mencari kolom yang memiliki nilai null (count data < total)
2. Check Duplicate Data
Mencari baris yang redudan lebih dari sekali
3. Check Distribution Data
Melihat nilai statistik dari tiap kolom
4. Check Inconsistent Data
Mencari data yang ambigu terutama pada kolom kategorik

Handling Missing Values

	kolom	jumlah	persentase
0	children	4	0.003
1	city	488	0.409
2	agent	16340	13.686
3	company	112593	94.307

```
[19] 1 # handling missing values
      2 df_hotel_prep['children'].fillna(0, inplace=True)
      3 df_hotel_prep['city'].fillna('Undefined', inplace=True)
      4 df_hotel_prep['agent'].fillna(df_hotel_prep['agent'].median(), inplace=True)
      5 df_hotel_prep.drop('company', axis=1, inplace=True)
```


Handling Duplicate Data

```
[20] 1 # remove duplicate rows
      2 print(f'Jumlah data sebelum difilter duplikat data adalah {len(df_hotel_prep)} baris')
      3 df_hotel_prep.drop_duplicates(inplace=True)
      4 print(f'Jumlah baris setelah difilter duplikat data adalah {len(df_hotel_prep)} baris')
```

Jumlah data sebelum difilter duplikat data adalah 119390 baris

Jumlah baris setelah difilter duplikat data adalah 86118 baris

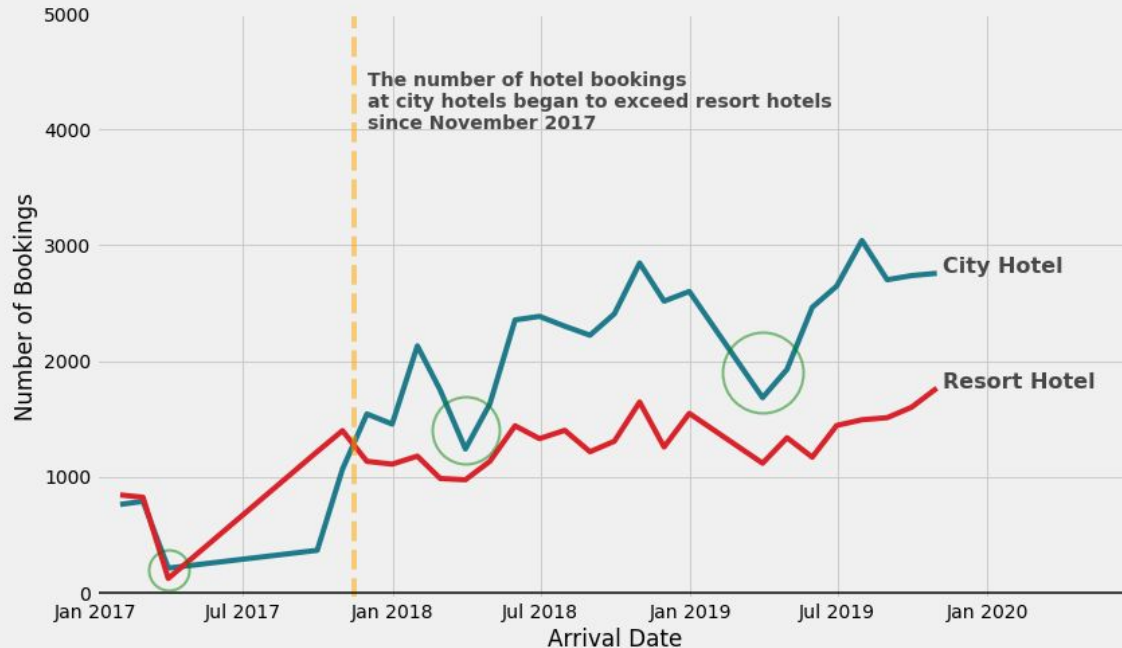
Handling Inconsistent Data

```
[24] 1 # filter inconsistent data
      2 df_hotel_prep = df_hotel_prep[df_hotel_prep['adr'] >= 0]
      3 # replace inconsistent data
      4 df_hotel_prep['meal'].replace('Undefined', 'No Meal', inplace=True)
      5 print(df_hotel_prep.shape)
```

(86117, 28)

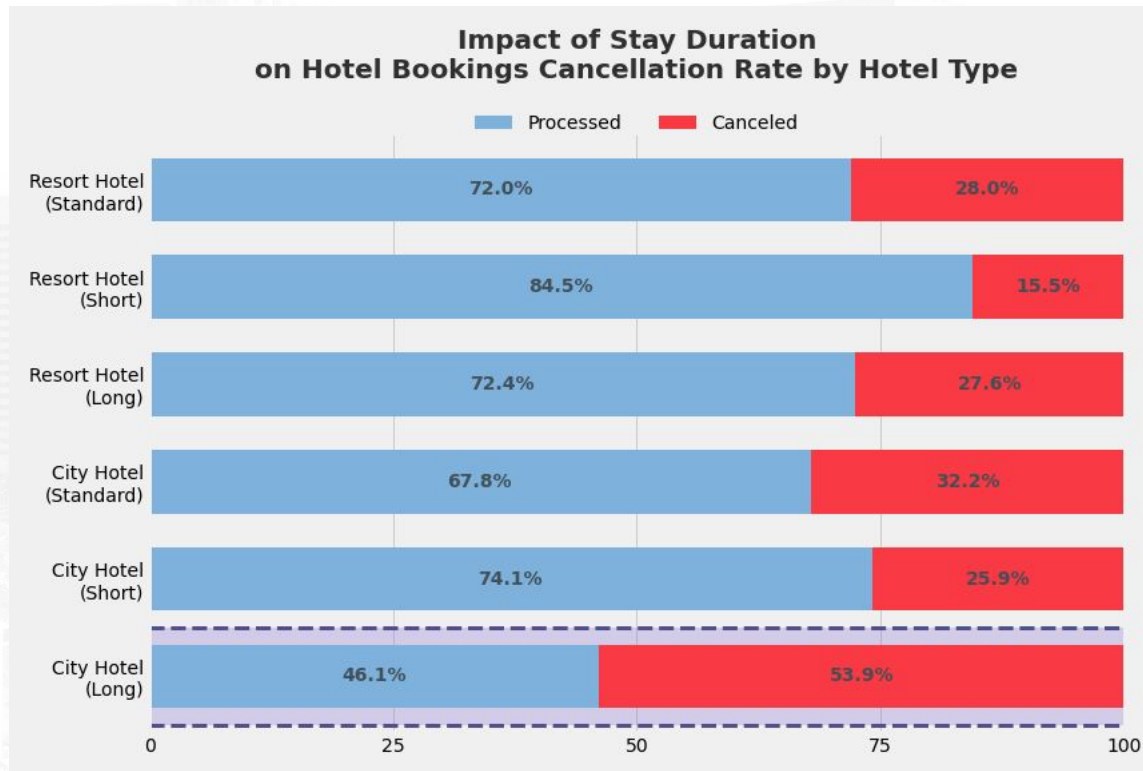
Monthly Hotel Booking Analysis Based on Hotel Type

The Number of Hotel Bookings at City Hotels Tends to Decrease when Entering March (Spring) Every Year



Insight:

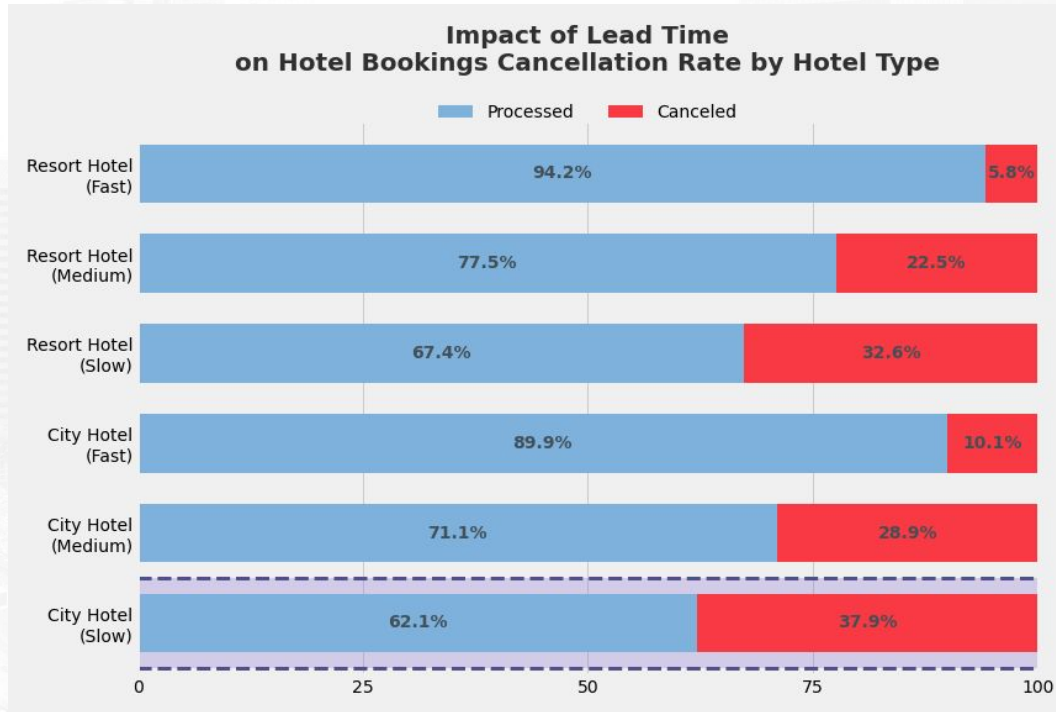
Jumlah pemesanan hotel dengan **jenis City Hotel** terlihat mulai melampaui **jenis Resort Hotel** sejak **November 2017**. Hal yang menarik adalah jumlah pemesanan di **City Hotel** cenderung akan mengalami penurunan saat memasuki bulan **Maret** atau musim semi pada tiap tahun nya.



Insight:

Pada **City Hotel** dengan durasi menginap bertipe Long (diatas 7 malam) lebih memungkinkan untuk melakukan pembatalan pemesanan hotel dengan tingkat pembatalan diatas 50%.

Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate



Insight:

Pada kedua jenis hotel, **semakin lama jarak antara waktu memesan dan tanggal kedatangan maka semakin tinggi kemungkinan tingkat pembatalan pesanan hotel** tetapi tidak sampai 50% chance nya.

Thanks !!

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