

Investigate Business Hotel using Data Visualization

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



Created by:

Muchammad Malik

muchammad.malik@gmail.com

<https://www.linkedin.com/in/muchammad-malik/>

“Currently working as business and system development at GESITS, the leading EV brand in Indonesia. Having a bachelor degree from engineering physics , I possessed balanced skill in engineering and management discipline.

As a data science and business analyst enthusiast, I developed skillset in Business Accumen, SQL, Pyhton, Tableau, and machine learning. I have spent 2 years to learn and maintain these skill by taking several bootcamp and online course.

I am a highly-motivated learner to keep me stay relevant, have good analytical thinking, have creative problem solving skill, and able to work in team. During college life, I developed my soft skill by having experienced in leading a commitee, participating in international competition (Model United Nation and business case competition), and also participating in extra-campus organization, such as AIESEC and StudentsCatalyst. ”

“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.”

PROSES 1 – HANDLING MISSING DATA

Terdapat 4 kolom yang memiliki nilai kosong, yaitu:

- **Kolom company:** presentase nilai null mencapai 94,3%
- **Kolom city:** presentase nilai null dibawah 1%,
- **Kolom children:** memiliki nilai null dibawah 1
- **Kolom agent:** presentase nilai null sebesar 13,6%

Dilakukan **penghapusan kolom**

kolom tetap dipertahankan namun dilakukan **penghapusan pada row** yang kosong

dilakukan **imputasi** dengan nilai rata-rata

Sebelum proses 1

(119390 rows, 29 columns)



Setelah proses 1

(118898 rows, 28 columns)

Untuk selengkapnya, dapat melihat jupyter notebook disini:

https://colab.research.google.com/drive/1DZGh2LU8gnT3F7uta4hgzh7pcVP4c_?usp=sharing

PROSES 2 – MENGGANTI VALUE YANG TIDAK SESUAI

- Value dengan tulisan 'Undefined' dianggap bernilai null
- Nilai null perlu dihapuskan

```
df['meal'].value_counts()
```

Breakfast	91863
Dinner	14434
No Meal	10638
Undefined	1165
Full Board	798

Name: meal, dtype: int64



```
df['meal'].value_counts()
```

Breakfast	91863
Dinner	14434
No Meal	10638
Full Board	798

Name: meal, dtype: int64

PROSES 3 – MEMBUANG DATA YANG TIDAK DIPERLUKAN

Kolom adults, children, dan babies digabungkan menjadi satu kolom, yaitu 'guest'

Code:

```
df['guest'] = df['adults'] + df['children'] + df['babies']  
df_1 = df.drop(['adults','children','babies'], axis = 1)
```

Sebelum proses 2
(118898 rows, 28 columns)

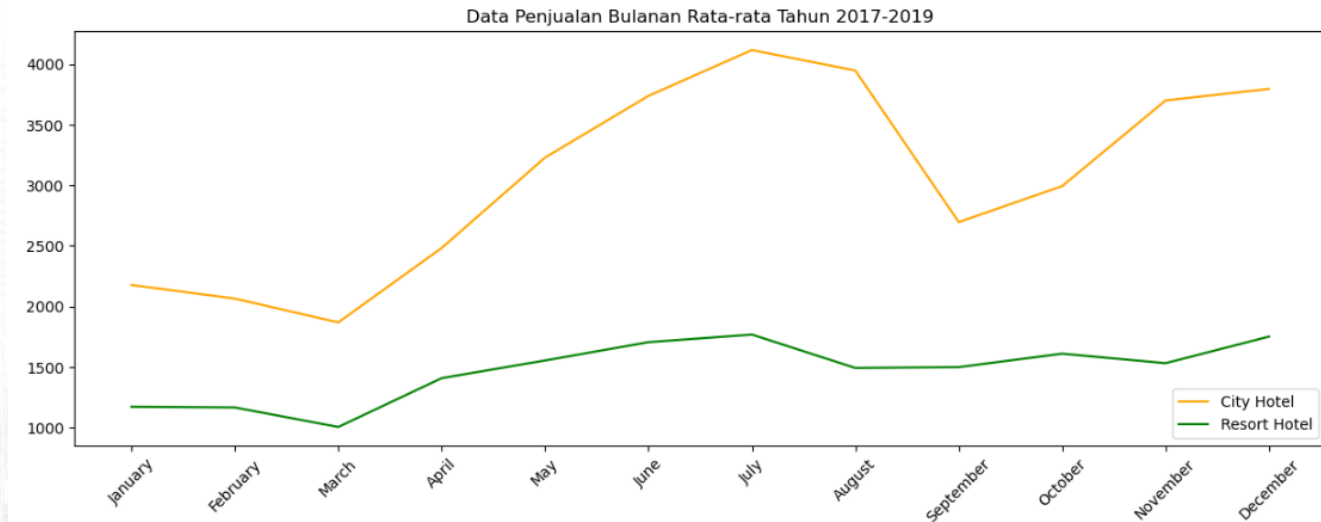


Setelah proses 2
(117733 rows, 28 columns)



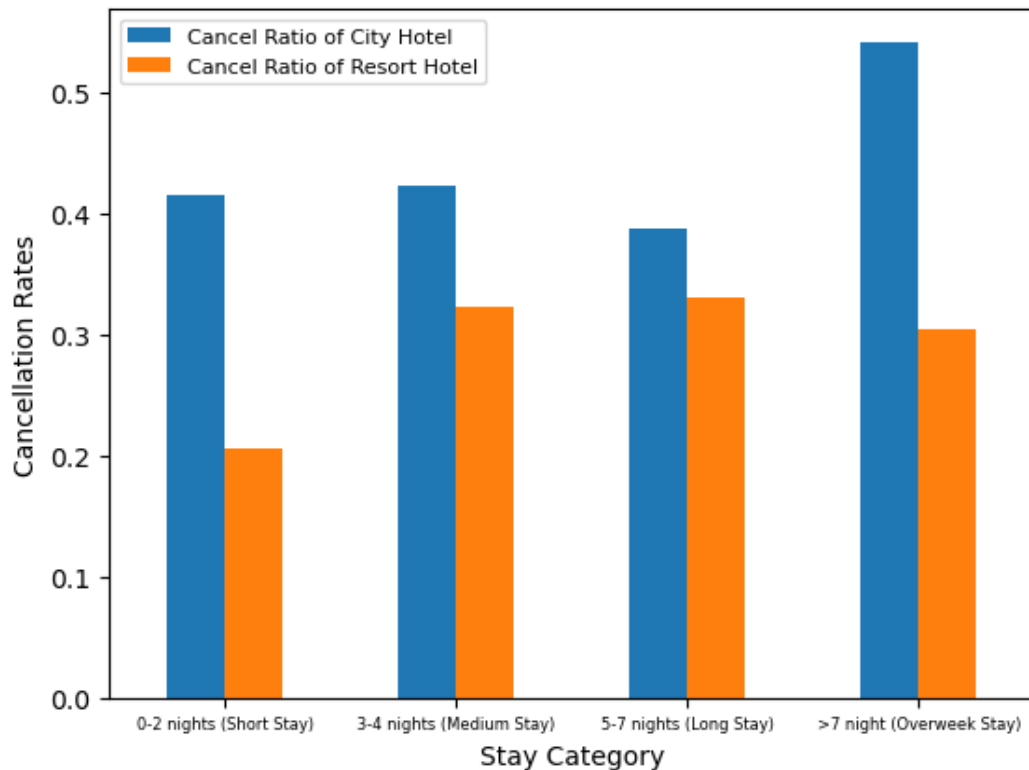
Setelah proses 3
(117733 rows, 26 columns)

Monthly Hotel Booking Analysis Based on Hotel Type



Insight:

- Penjualan hotel berada pada puncaknya di bulan Juni, Juli dan Desember. Hal ini bertepatan dengan bulan libur akhir sekolah di pertengahan tahun dan libur Natal di akhir tahun.
- Tren tersebut kemudian menurun hingga penjualan mencapai titik terendah dua bulan setelah titik tertingginya.
- Jenis hotel yang paling mengalami lonjakan signifikan adalah hotel bertipe city hotel



Pada hotel bertipe 'City Hotel' :

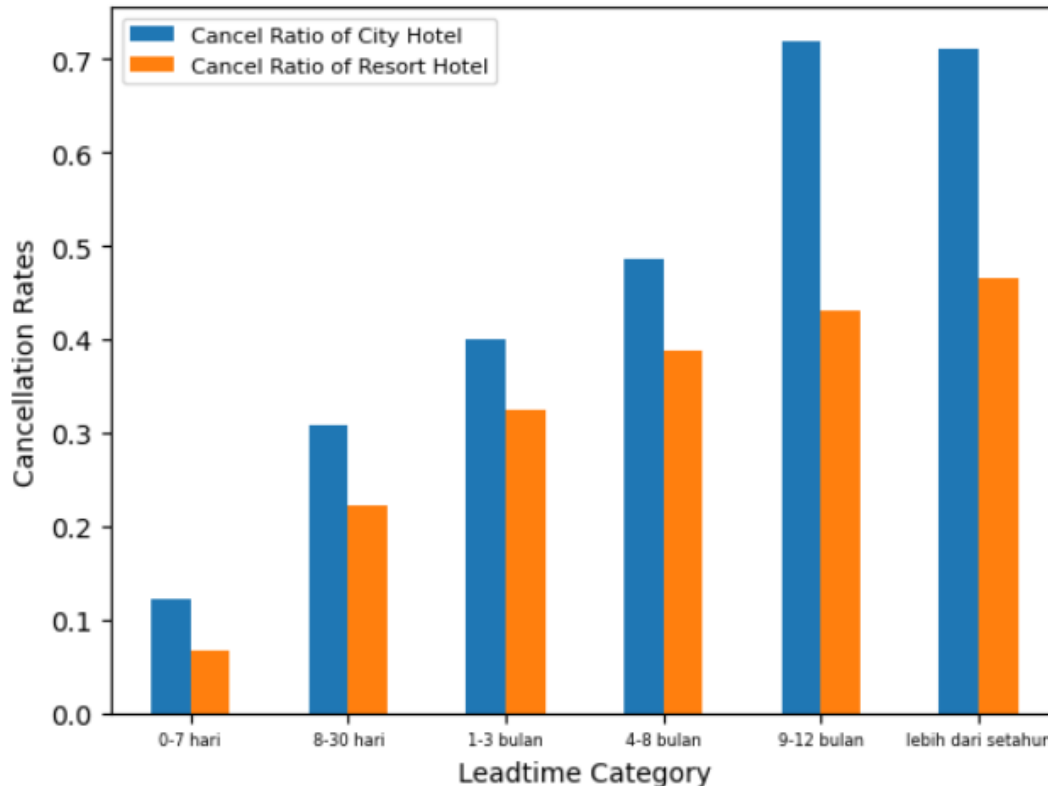
- cancel ratio yang memiliki perbedaan signifikan adalah untuk durasi menginap lebih dari tujuh malam, yaitu ratio pada kategori tersebut jauh lebih besar dibanding kategori lainnya
- untuk hotel dengan durasi menginap 0-7 malam, nilai cancel ratio tidak jauh berbeda satu sama lain

Pada hotel bertipe 'Resort Hotel' :

- cancel ratio yang memiliki perbedaan signifikan adalah untuk durasi menginap 0-2 malam, yaitu ratio pada kategori tersebut jauh lebih kecil dibanding kategori lainnya
- untuk hotel dengan durasi menginap lebih dari 2 malam, nilai cancel ratio tidak jauh berbeda satu sama lain

Untuk selengkapnya, dapat melihat jupyter notebook disini:

https://colab.research.google.com/drive/1DZGh2LU8gnT3F7uta4hgzh7pcVP4c_?usp=sharing



Perbandingan tipe city hotel dengan resort hotel

- Secara keseluruhan, nilai cancel ratio pada City Hotel lebih tinggi dibanding Resort Hotel di kategori leadtime manapun.
- Selisih cancel ratio terbesar berada pada leadtime category 9-12 bulan

Pola Cancel Ratio

Pola cancel ratio pada kedua tipe hotel juga serupa, yaitu cancel ratio sangat kecil pada kategori lead time 0-7 hari, lalu semakin membesar seiring bertambahnya lead time