

Investigate Business Hotel using Data Visualization

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



Created by:

M TOHAR SAGARA

Email : toharsagara@gmail.com

LinkedIn : M Tohar Sagara

Tohar adalah orang yang dapat diandalkan untuk menganalisa dan mengevaluasi suatu data. Dia dapat diandalkan untuk mencari jalan keluar terkait suatu hal dengan data.

Dia adalah pribadi yang mudah bergaul dapat bekerja dalam tim ataupun individual. Dia menyukai tantangan dan menggali hal-hal baru untuk menambah ilmu dan pengalaman saya baik di bidang saintis ataupun pembangunan dan pemberdayaan masyarakat.

Dia dapat diandalkan dalam kondisi beban kerja dengan waktu yang terbatas

“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. ”

- Setelah data dibuka dapat dilihat bahwa dataset ini memiliki 119390 data dan 29 kolom
- Data ini sendiri memiliki :
 - 4 Kolom data float64
 - 16 Kolom data integer
 - 9 Kolom data object

#	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

A. Nilai Hilang

1. Kolom '**Company**' dan '**Agent**' memiliki jumlah nilai yang hilang yang sangat tinggi yaitu **94,31%** dan **13,69%**. Kolom tersebut ingin saya hapus, tetapi saya tidak menghapusnya melainkan akan **mengganti nilai null pada tersebut menjadi nol** (karena mereka memiliki tipe data float64).
2. Di sisi lain, kolom '**city**' memiliki **0,41%** nilai yang hilang (488 baris). Dalam kasus ini, sebagai variabel kategorikal, saya memilih untuk **mengganti dengan nilai 0(nol)**.
3. Kolom '**children**' juga memiliki beberapa nilai yang hilang, tetapi mereka bahkan tidak mewakili **0,01%**. Oleh karena itu, saya memilih untuk **menghapus** baris yang terpengaruh sejumlah 4 baris.

Kolom	Null	Persentase
company	112593	94.31
agent	16340	13.69
city	488	0.41
children	4	0.00
hotel	0	0.00
is_repeated_guest	0	0.00
total_of_special_requests	0	0.00
required_car_parking_spaces	0	0.00
adr	0	0.00
customer_type	0	0.00
days_in_waiting_list	0	0.00
deposit_type	0	0.00
booking_changes	0	0.00
previous_bookings_not_canceled	0	0.00
previous_cancellations	0	0.00
market_segment	0	0.00
distribution_channel	0	0.00
is_canceled	0	0.00
meal	0	0.00
babies	0	0.00
adults	0	0.00
stays_in_weekdays_nights	0	0.00
stays_in_weekend_nights	0	0.00
arrival_date_day_of_month	0	0.00
arrival_date_week_number	0	0.00
arrival_date_month	0	0.00
arrival_date_year	0	0.00
lead_time	0	0.00
reservation_status	0	0.00

B. Data Tidak Konsisten

Pada dataset tersebut terdapat 10 data kategorikal antara lain :

- | | |
|-------------------|-------------------------|
| 1. hotel | 6. distribution_channel |
| 2. is_canceled | 7. is_repeated_guest |
| 3. meal | 8. deposit_type |
| 4. City | 9. customer_type |
| 5. market_segment | 10. reservation_status |

Pada kolom 'meal' ada 5 hasil: ['Breakfast' 'Full Board' 'Dinner' 'No Meal' 'Undefined']

Breakfast artinya sarapan

Fullboard artinya sarapan, makan siang dan makan malam

Dinner artinya hanya makan malam

No Meal artinya tanpa makanan

Undefined artinya tidak di definisikan yang dapat juga

dimaksud tidak makan yang sama dengan 'No Meal

oleh sebab itu saya **mengganti Undefined menjadi No Meal**

Sebelumnya

```
meal:  
['Breakfast' 'Full Board' 'Dinner' 'No Meal' 'Undefined']
```

Setelahnya

```
meal:  
['Breakfast' 'Full Board' 'Dinner' 'No Meal']
```

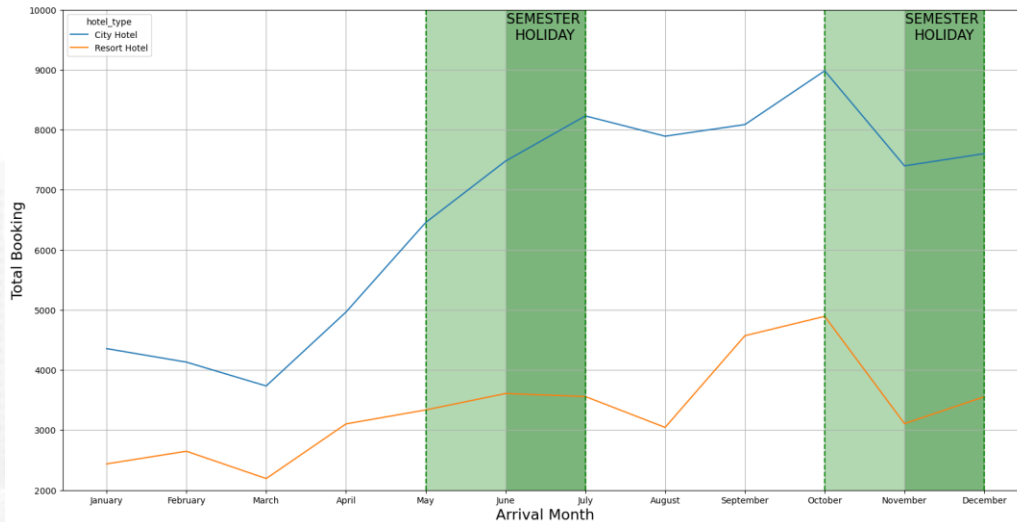

C. Data Tidak Sesuai

1. kita dapat melihat bahwa kolom **'previous_cancellations'** memiliki nilai paling tinggi 26 pembatalan, yang mengimplikasikan bahwa beberapa pelanggan melakukan 26 kali pembatalan, yang mana hal ini tidak mungkin terjadi.
2. kolom **'adult'** memiliki nilai tertinggi 55 dan terendah 0 orang. Nilai minimum ini mengasumsikan bahwa ada reservasi hotel untuk 0 orang dewasa, yang tidak mungkin karena harus ada minimal 1 orang dewasa per reservasi (jelas anak-anak tidak dapat memesan kamar hotel). Oleh karena itu, saya akan **menghilangkan baris yang jumlah orang dewasanya sama dengan 0.**

	previous_cancellations	adults
count	119386.000000	119386.000000
mean	0.087121	1.856390
std	0.844350	0.579261
min	0.000000	0.000000
25%	0.000000	2.000000
50%	0.000000	2.000000
75%	0.000000	2.000000
max	26.000000	55.000000

Monthly Hotel Booking Analysis Based on Hotel Type

Total pemesanan hotel setiap bulan berdasarkan jenis hotel

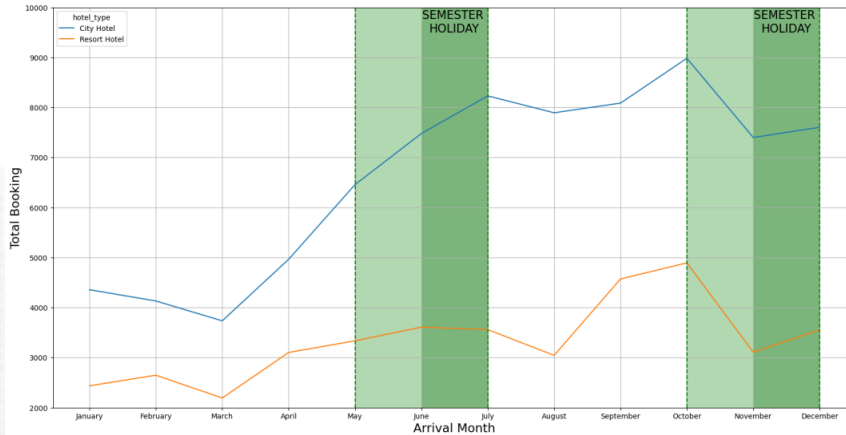


- Pada Bulan 10 (Oktober) adalah bulan dengan pemesanan tertinggi sebesar 8983 Pemesanan City Hotel dan 4894 Pemesanan Resort Hotel yang terkonfirmasi .
- Pada Bulan 3 (Maret) adalah bulan dengan pemesanan terendah sebesar 3735 pesanan City Hotel dan 2193 pesanan Resort Hotel yang tekonfirmasi

hotel_type	arrival_month	total_arrival_year	total_booking
City Hotel	January	2	4357
City Hotel	February	2	4132
City Hotel	March	2	3735
City Hotel	April	2	4965
City Hotel	May	2	6458
City Hotel	June	2	7480
City Hotel	July	2	8232
City Hotel	August	2	7894
City Hotel	September	3	8088
City Hotel	October	3	8983
City Hotel	November	2	7400
City Hotel	December	2	7604
Resort Hotel	January	2	2437
Resort Hotel	February	2	2648
Resort Hotel	March	2	2193
Resort Hotel	April	2	3103
Resort Hotel	May	2	3336
Resort Hotel	June	2	3609
Resort Hotel	July	2	3559
Resort Hotel	August	2	3045
Resort Hotel	September	3	4572
Resort Hotel	October	3	4894
Resort Hotel	November	2	3108
Resort Hotel	December	2	3555

Monthly Hotel Booking Analysis Based on Hotel Type

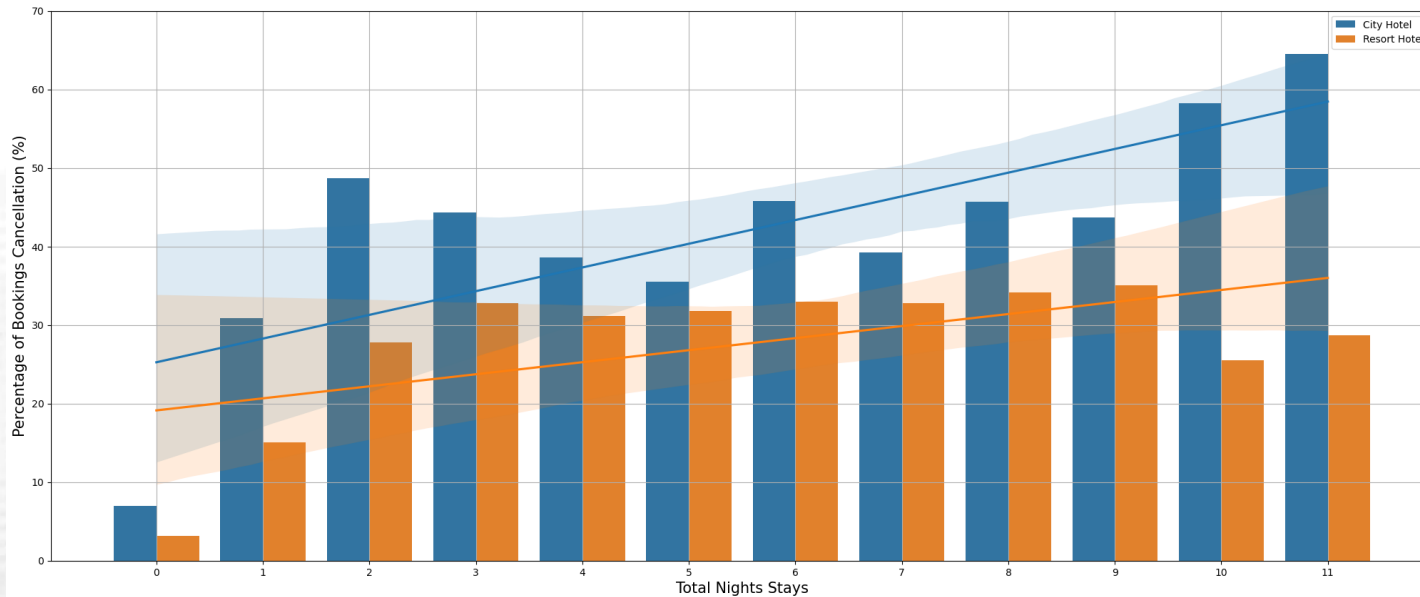
Total pemesanan hotel setiap bulan berdasarkan jenis hotel



- Pada periode Januari hingga Maret adalah periode terendah yang melakukan pemesanan hotel baik city hotel maupun resort hotel
- Pada periode libur semester awal (mei hingga juli) terjadi peningkatan pemesanan yang mungkin disebabkan oleh libur anak sekolah dan hari raya idul fitri baik untuk city hotel dan resort hotel bagi yang ingin libur bersama keluarga dan mulai ada kegiatan fullboard meeting untuk city hotel.
- Pada periode juli hingga oktober adalah periode banyaknya kegiatan fullboard meeting di city hotel, sedangkan untuk resort hotel memiliki kenaikan karena biasanya pada pertengahan bulan ini banyak pernikahan yang membutuhkan resort hotel untuk honeymoon. Oleh sebab itu memiliki kenaikan signifikan dari awal tahun
- Pada periode libur semester akhir tahun terjadi peningkatan tertinggi baik di city hotel, karena padatnya kunjungan perjalanan dinas baik pemerintahan ataupun swasta dan kegiatan – kegiatan fullboard meeting serta masa libur sekolah, natal dan libur tahun baru . Dan juga untuk resort hotel menunjukkan kenaikan karena masa libur sekolah, natal dan libur tahun baru .

Lead Time on Hotel Bookings Cancellation Rate

Impact of Stay Duration on Hotel Bookings Cancellation Rates Based on Hotel Type



Secara garis besar, baik pada City Hotel dan Resort Hotel, terlihat tren naiknya persentase cancel booking seiring bertambahnya jumlah hari menginap dari customer. Hal ini dapat menyebabkan ruginya perusahaan (hotel) karena kamar yang telah dipesan oleh customer bersangkutan ternyata dicancel yang menyebabkan hilangnya potensi revenue.

Business Recommendations

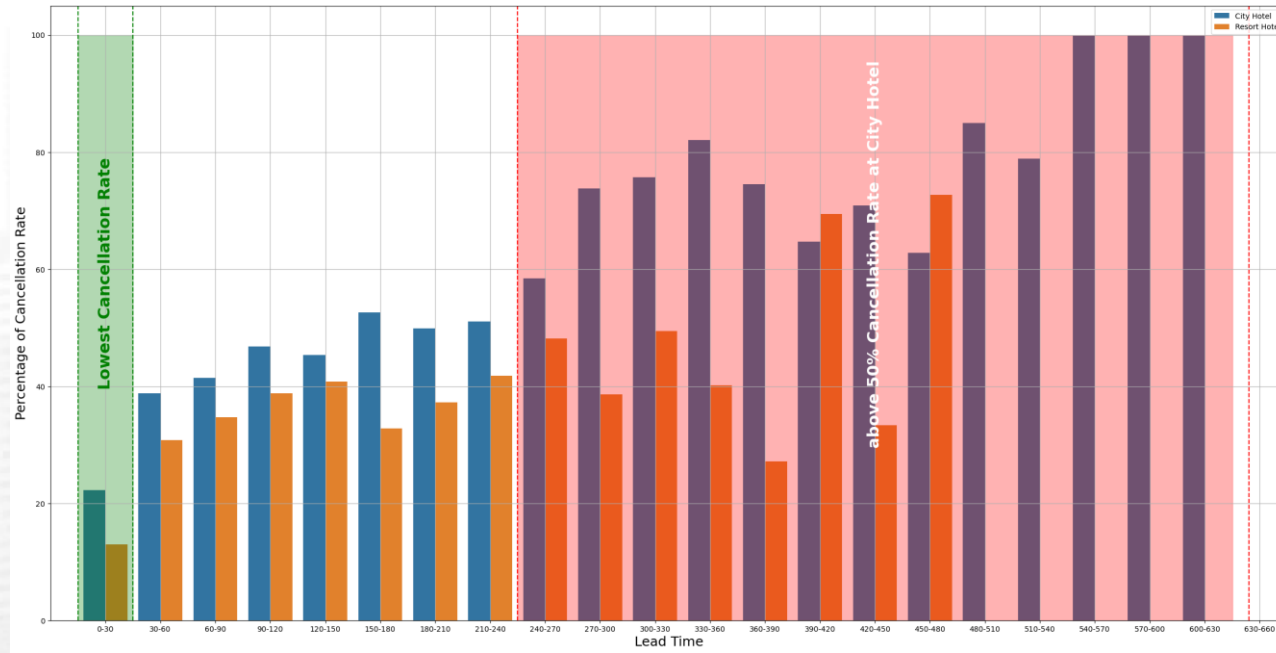
Mengubah metode pembayaran yang dimana customer membayar saat tiba di hotel menjadi pembayaran saat pemesanan, jika terjadi pembatalan dibuat kebijakan pengembalian dana customers (refund) dengan mencantumkan ketentuan fees for canceling orders (biaya pembatalan) berdasarkan hari sebelum kedatangan.

Untuk selengkapnya, dapat melihat repositorynya [disini](#)

hotel_type	total_nights_stays	Total Cancel	percentage_cancel
City Hotel	0	23	6.95
City Hotel	1	4103	30.91
City Hotel	2	10434	48.70
City Hotel	3	9486	44.37
City Hotel	4	4853	38.65
City Hotel	5	1776	35.54
City Hotel	6	943	45.80
City Hotel	7	808	39.24
City Hotel	8	176	45.71
City Hotel	9	93	43.66
City Hotel	10	116	58.29
City Hotel	11	291	64.52
Resort Hotel	0	12	3.12
Resort Hotel	1	1168	15.07
Resort Hotel	2	1729	27.81
Resort Hotel	3	1865	32.75
Resort Hotel	4	1505	31.19
Resort Hotel	5	887	31.83
Resort Hotel	6	592	32.93
Resort Hotel	7	2161	32.76
Resort Hotel	8	265	34.15
Resort Hotel	9	220	35.03
Resort Hotel	10	240	25.53
Resort Hotel	11	477	28.67

Lead Time on Hotel Bookings Cancellation Rate

Dampak Lead Time terhadap Tingkat Pembatalan Pemesanan Hotel Berdasarkan Jenis Hotel



- Persentase cancellation rate paling sedikit terdapat pada Lead Time kurang dari 1 bulan
- Persentase cancellation rate pada leadtime 240 hingga 630 sudah menyentuh angka diatas 50%, yang artinya kemungkinan akan pesanan dilanjutkan menjadi kedatangan adalah 50:50, customer hanya melakukan pemesanan untuk mendapatkan slot tempat , sedangkan jadi tidaknya menjadi kedatangan masih melihat perkembangan situasi dan kondisi costumers itu sendiri.
- Terlebih pada rentang lead time 540 hingga 630 tingkat pembatalan adalah 100%, yang artinya customers hanya sampai proses pemesanan saja.

Business Recommendations

Mengubah metode pembayaran yang dimana customer membayar saat tiba di hotel menjadi pembayaran saat pemesanan, jika terjadi pembatalan dibuat kebijakan pengembalian dana customers (refund) dengan mencantumkan ketentuan fees for canceling orders (biaya pembatalan) berdasarkan hari sebelum kedatangan.