

# Investigate Hotel Business



**Created by:**

**Fikri Diva Sambasri**

[fikri.sambasri@gmail.com](mailto:fikri.sambasri@gmail.com)

<http://www.linkedin.com/in/fikridivasambasri>

I'm final year informatics Engineering student with experience in working on several projects related to Data Science using Python and several data science tools during bootcamp Data Science at Rakamin Academy to deliver insights from machine learning model. I also have ability in Data Analyst position with experience using Google BigQuery and Google Data Studio.

**Supported by:**  
**Rakamin Academy**  
Career Acceleration School  
[www.rakamin.com](http://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.”

## 1. Mengatasi data null / data kosong

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

Pada persoalan pertama terdapat temuan missing value:

1. Terdapat 4 kolom yang memiliki missing value yaitu kolom city, agent, company, dan children.
2. Kolom company memiliki nilai missing value yang besar sehingga bisa diisi menggunakan nilai 0. Begitu juga untuk kolom agent dan children
3. Isi nilai kosong pada kolom city menggunakan nilai 'unknown'.

```
# fill missing value

# isi nilai missing value pada kolom company dengan nilai 0
df['company'] = df['company'].fillna(0)
# isi nilai missing value pada kolom agent dengan nilai 0
df['agent'] = df['agent'].fillna(0)
# isi nilai missing value pada kolom city dengan nilai unknown
df['city'] = df['city'].fillna('unknown')
# isi nilai missing value pada kolom children dengan nilai 0
df['children'] = df['children'].fillna(0)
```

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

## 2. Mengatasi data yang tidak sesuai

Pada persoalan kedua terdapat 3 temuan:

1. Terdapat nilai **Undefined** pada kolom meal sehingga nilai tersebut dapat diganti dengan nilai 'No Meal'.

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

```
# drop nilai "undefined" ganti dengan No Meal pada kolom meal  
df["meal"].replace("Undefined", "No Meal", inplace=True)
```

## 3. Mengidentifikasi data-data yang tidak berhubungan

Pada persoalan ketiga terdapat temuan:

1. Terdapat nilai 0 untuk kolom adults, kolom children, dan kolom babies yang mana jika nilainya 0 pada ketiga kolom logikanya tidak ada yang memesan hotel. Sehingga kita dapat mengeclude data tersebut dari analisis

arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights	stays_in_weekdays_nights	adults	children	babies	meal	city
41	6	0	3	0	0.0	0	No Meal	Kota Denpasar
42	12	0	0	0	0.0	0	No Meal	Kota Denpasar
47	20	1	2	0	0.0	0	No Meal	Kota Batu
53	30	1	4	0	0.0	0	No Meal	Kota Denpasar
53	30	2	4	0	0.0	0	No Meal	Kota Denpasar

```
# drop baris yang memiliki nilai 0 pada kolom adults, children, and babies
df.drop(df[(df['adults'] == 0) & (df['children'] == 0) & (df['babies'] == 0)].index, inplace=True)
```

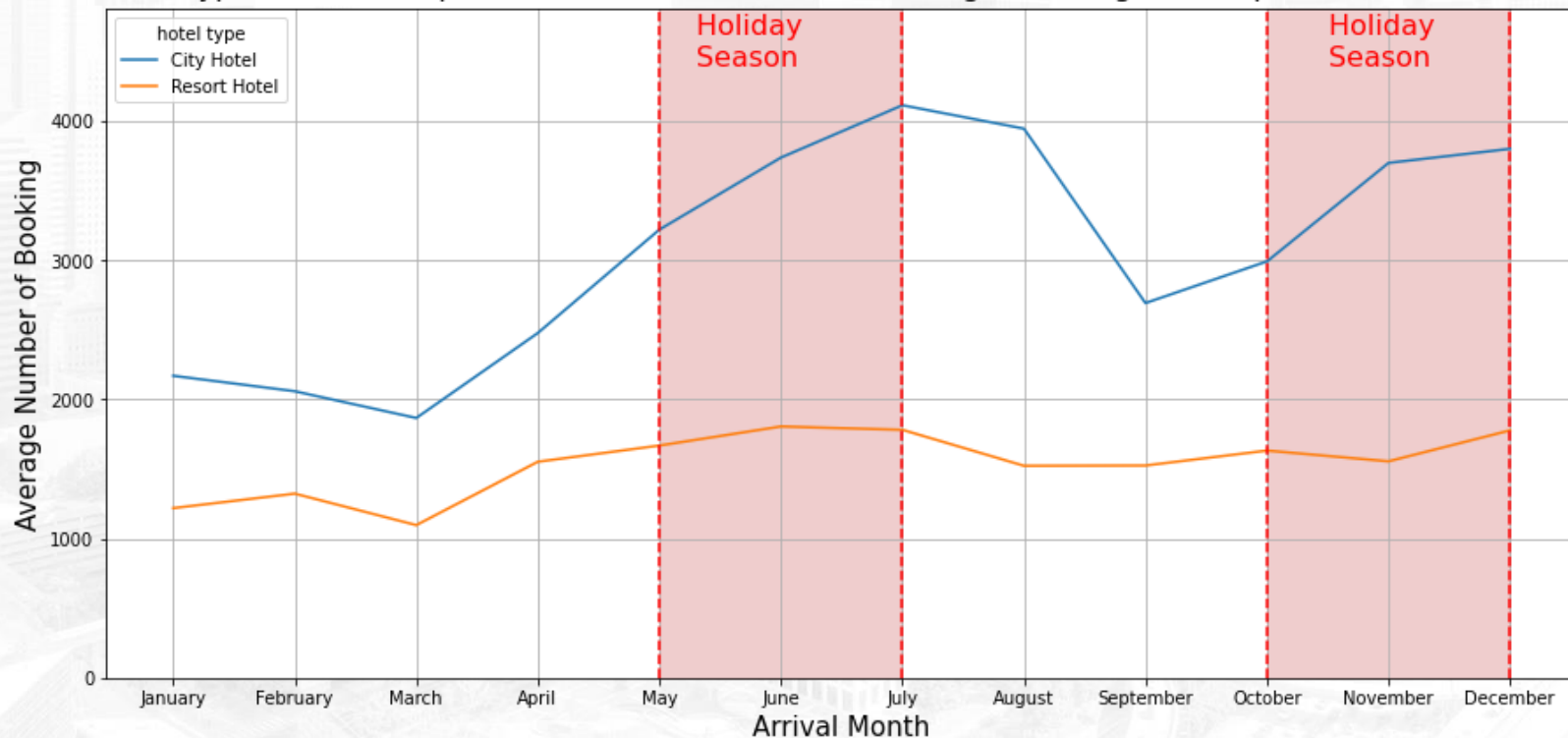
# Monthly Hotel Booking Analysis Based on Hotel Type

Supported by:  
**Rakamin Academy**  
Career Acceleration School  
[www.rakamin.com](http://www.rakamin.com)

# Monthly Hotel Booking Analysis Based on Hotel Type

## Comparison of the number of bookings for each type of hotel by month

Both types of hotels experienced a decrease in hotel bookings from august to september



Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



## Insight/Analisis:

Analisis ini berfokus untuk melihat *trend* dari pemesanan hotel untuk setiap jenis hotel per bulannya.

Kedua jenis hotel mendapatkan jumlah pemesanan terbanyak di sekitar bulan Juni dan Juli. Jika dipikirkan lebih dalam, waktu tersebut bertepatan dengan liburan Hari Raya dan liburan anak sekolah.

Selain itu, di bulan lainnya yang tidak berbarengan dengan suatu libur apapun terlihat pemesanan jauh lebih rendah. Sehingga untuk Hotel Perkotaan mengalami penurunan signifikan pada bulan Agustus dan September.

Kemudian di bulan Oktober dan Desember pemesanan hotel mengalami kenaikan seperti bulan Juni dan Juli, mungkin bulan tersebut adalah bulan disaat liburan hari tahun baru / hari natal.



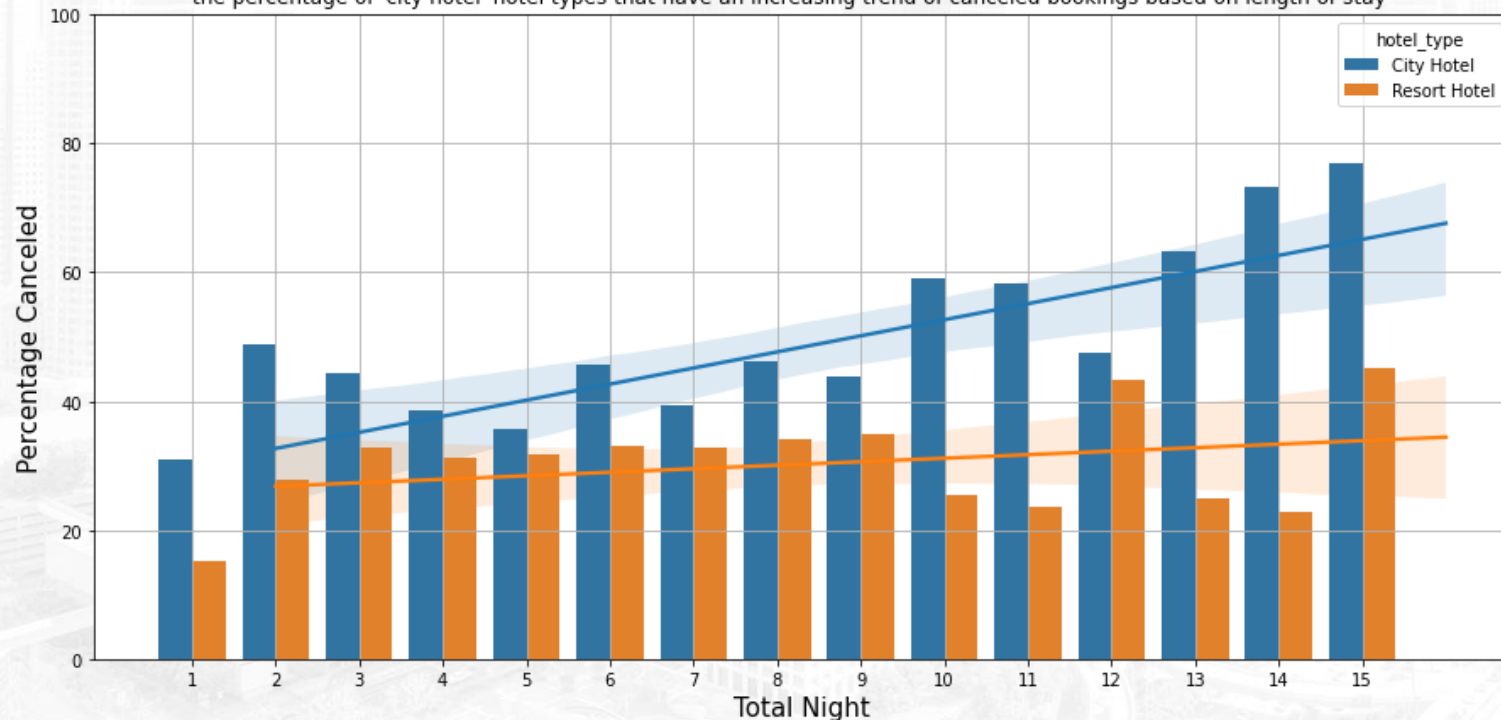
# Impact Analysis of Stay Duration on Hotel Bookings Cancel

Supported by:  
**Rakamin Academy**  
Career Acceleration School  
[www.rakamin.com](http://www.rakamin.com)

# Impact Analysis of Stay Duration on Hotel Bookings Cancel

## Trend of total bookings against percentage of canceled bookings

the percentage of 'city hotel' hotel types that have an increasing trend of canceled bookings based on length of stay



## Insight/Analisis:

Analisis ini berfokus untuk melihat *trend* dari durasi menginap dengan tingkat pembatalan pemesanan hotel.

Kedua jenis hotel memiliki trend positif, dimana semakin lama durasi menginap, semakin tinggi kemungkinan pemesanan tersebut dibatalkan.

Selain itu, terlihat untuk hotel di perkotaan mengalami trend positif yang signifikan untuk pemesanan yang dibatalkan dibandingkan dengan hotel resort.

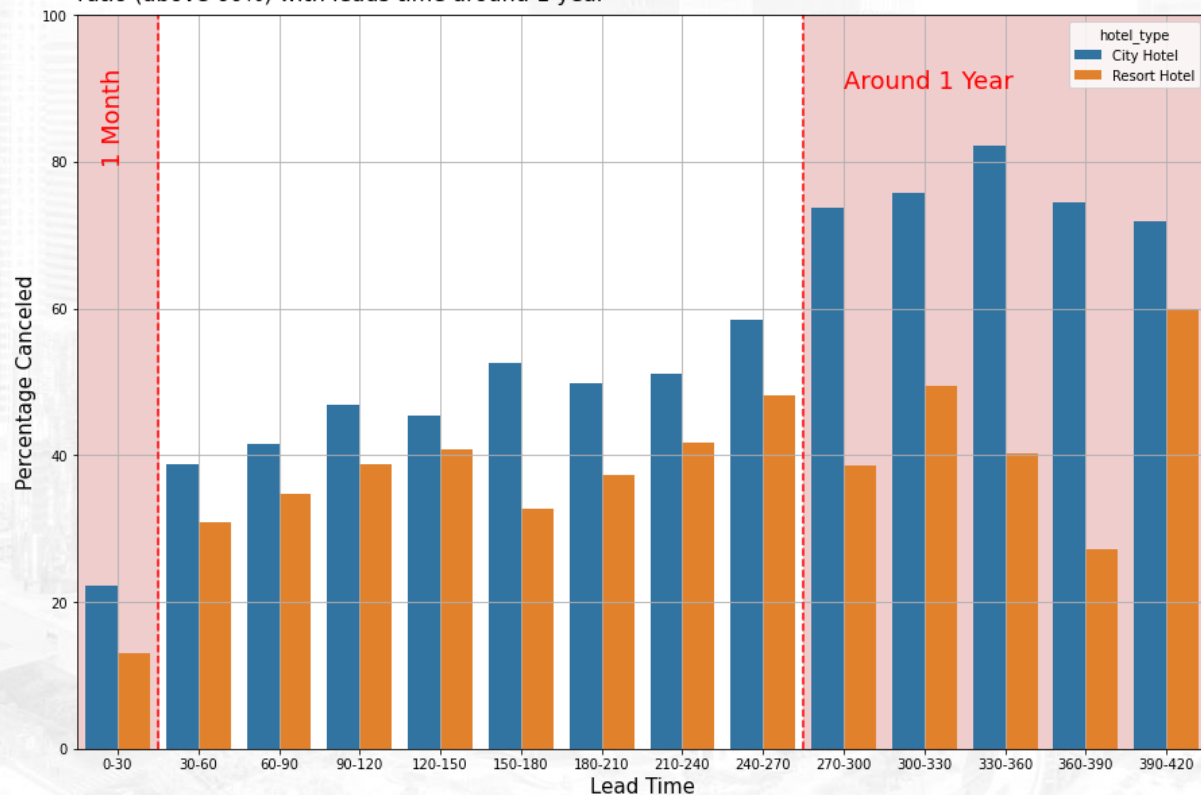
# Impact Analysis of Lead Time on Hotel Bookings Cancel

Supported by:  
**Rakamin Academy**  
Career Acceleration School  
[www.rakamin.com](http://www.rakamin.com)

# Impact Analysis of Lead Time on Hotel Bookings Cancel

## Comparison of order ratios to the percentage of cancellation rates for hotel booking

Resort Hotel has quite stagnant (in around 40%) canceled ratio, while City Hotel has quite high canceled ratio (above 60%) with leads time around 1 year



Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

## Insight/Analisis:

Analisis ini berfokus untuk melihat *trend* antara waktu tunggu (jarak waktu pemesanan hotel hingga waktu kedatangan) dan tingkat pembatalan pemesanan hotel.

Tingkat pembatalan pemesanan paling rendah ada pada pemesanan yang memiliki waktu tunggu kurang dari 30 hari dan berlaku untuk kedua jenis hotel. Hotel resort cukup stagnan dengan nilai 40% rasio pembatalan, sedangkan hotel jenis perkotaan memiliki rasio yang cukup tinggi 60% Ketika waktu tunggu sekitar 1 tahun.