

# Investigate Hotel Business using Data Visualization

By : Wanda Dwi Putri

# Background



Hotel merupakan salah satu akomodasi industri pariwisata yang sangat penting. Untuk meningkatkan performa bisnis hotel kita perlu mendalami dan melakukan analisis terhadap data yang dimiliki. **Fokus kita adalah melakukan eksplorasi data untuk mengetahui perilaku customer dalam melakukan pemesanan hotel dan hubungannya terhadap tingkat pembatalan pemesanan hotel.** Insight yang ditemukan akan disajikan dalam bentuk data visualisasi agar lebih mudah untuk dipahami.

# Dataset

## **Dataset : hotel\_bookings\_data.csv**

Dataset yang digunakan merupakan dataset yang telah dimodifikasi oleh Tim Rakamin. Tiap baris dalam dataset berisi informasi mengenai reservasi City Hotel dan Resort Hotel, termasuk reservasi yang dibatalkan maupun tidak dibatalkan.

Dari hasil pengamatan diketahui bahwa :

- Dataset terdiri dari 119390 baris dan 29 kolom.
- Beberapa kolom masih memiliki tipe data yang belum sesuai.
- Kolom company, agent, city, dan children masih memiliki nilai null(missing values).

# Tools

**Programming Language : Python**



**Python Library : pandas & Numpy**



**Visualization : Matplotlib & Seaborn**



## Data Preprocessing

### → Handle Missing Value

- Missing value terbanyak terdapat dalam feature company hingga 94%, sehingga feature ini harus di drop dari dataframe.
- Missing values pada feature agent akan ditangani dengan impute features menggunakan modus.
- Missing values pada feature country dan children yang jumlahnya dibawah 1% akan dihapus dari dataset.

### → Handle Incorrect Data Type

Penanganan dilakukan pada feature children dan agen yang memiliki tipe data float. Untuk feature children tipe data float akan diubah menjadi integer. Sedangkan tipe data agent akan diubah menjadi object.

### → Handle Unsuitable Values

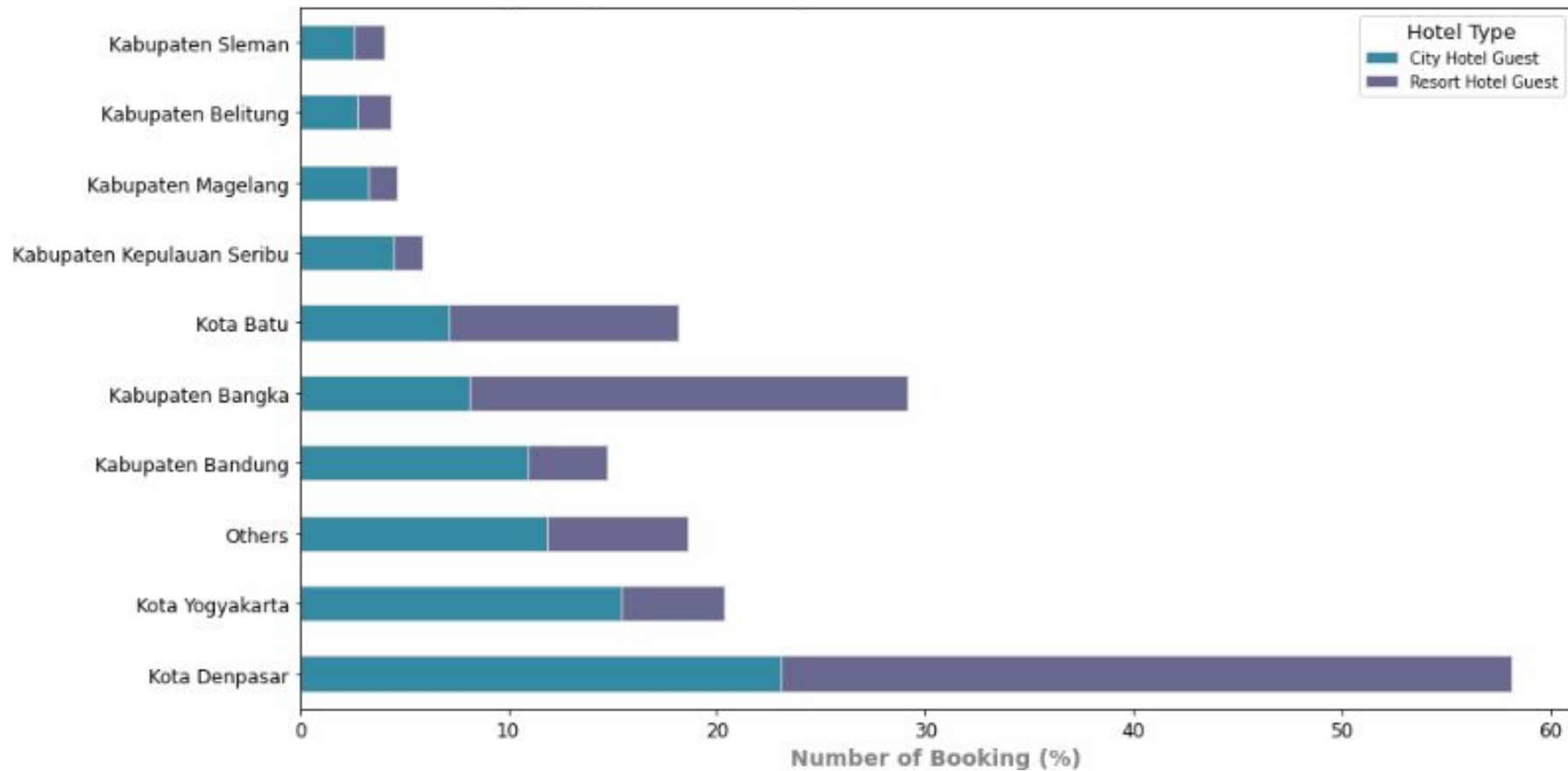
Feature meal dan distribution\_channel memiliki value yang tidak sesuai, sehingga perlu ditangani dengan mengganti nilai 'Undefined' pada feature meal dengan nilai 'No Meal' dan nilai 'TA/TO' pada feature distribution\_channel.

### → Remove Unnecessary Data

Actual visitor merupakan customer yang memiliki nilai total\_guests dan stay\_duration lebih dari 0. Nilai total\_guests diperoleh dari kombinasi feature adults, children, dan babies. Nilai stay\_duration diperoleh dari kombinasi feature stay\_in\_weekend\_nights dan stay\_in\_weekdays\_nights. Tahap selanjutnya menggunakan df\_final yang merupakan data frame dengan actual visitor.

## Data Visualization

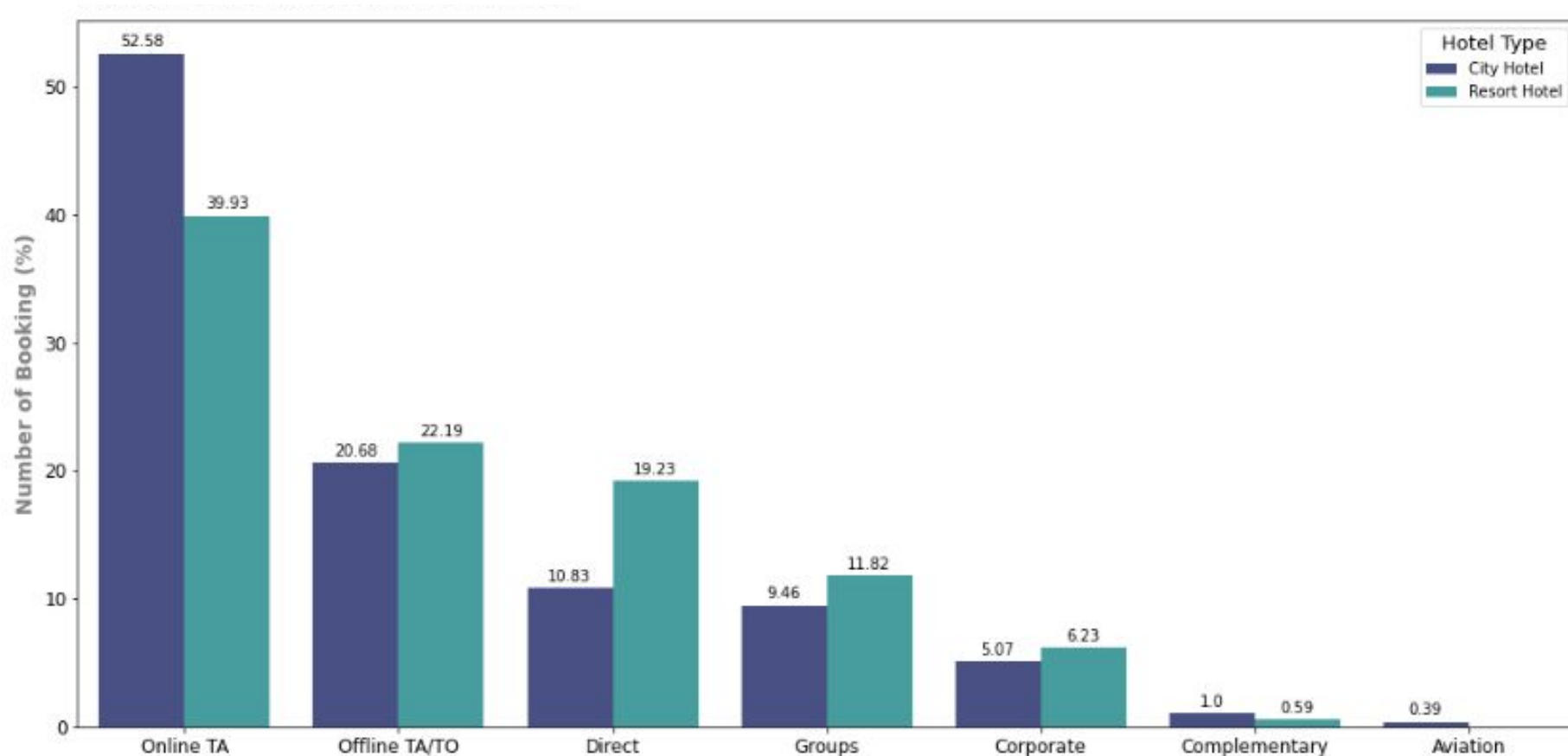
## Top 10 Origin Cities of Hotel Guests



	City	City Hotel Guest	Resort Hotel Guest
0	Kota Denpasar	23.07	35.09
1	Kota Yogyakarta	15.42	4.95
2	Others	11.83	6.73
3	Kabupaten Bandung	10.93	3.76
4	Kabupaten Bangka	8.16	20.98
5	Kota Batu	7.15	10.99
6	Kabupaten Kepulauan Seribu	4.47	1.34
7	Kabupaten Magelang	3.23	1.37
8	Kabupaten Belitung	2.74	1.62
9	Kabupaten Sleman	2.58	1.44

Dari Top 10 Origin Cities of Hotel Guest, Kota Denpasar merupakan kota asal tamu hotel yang paling banyak melakukan pemesanan di City Hotel dan Resort Hotel. Sedangkan Kabupaten Sleman merupakan kabupaten dengan pemesanan yang paling sedikit. Untuk kota - kota dengan total tamu di bawah 1% dikelompokkan ke dalam kategori Others.

## Number of Booking by Distribution Channel



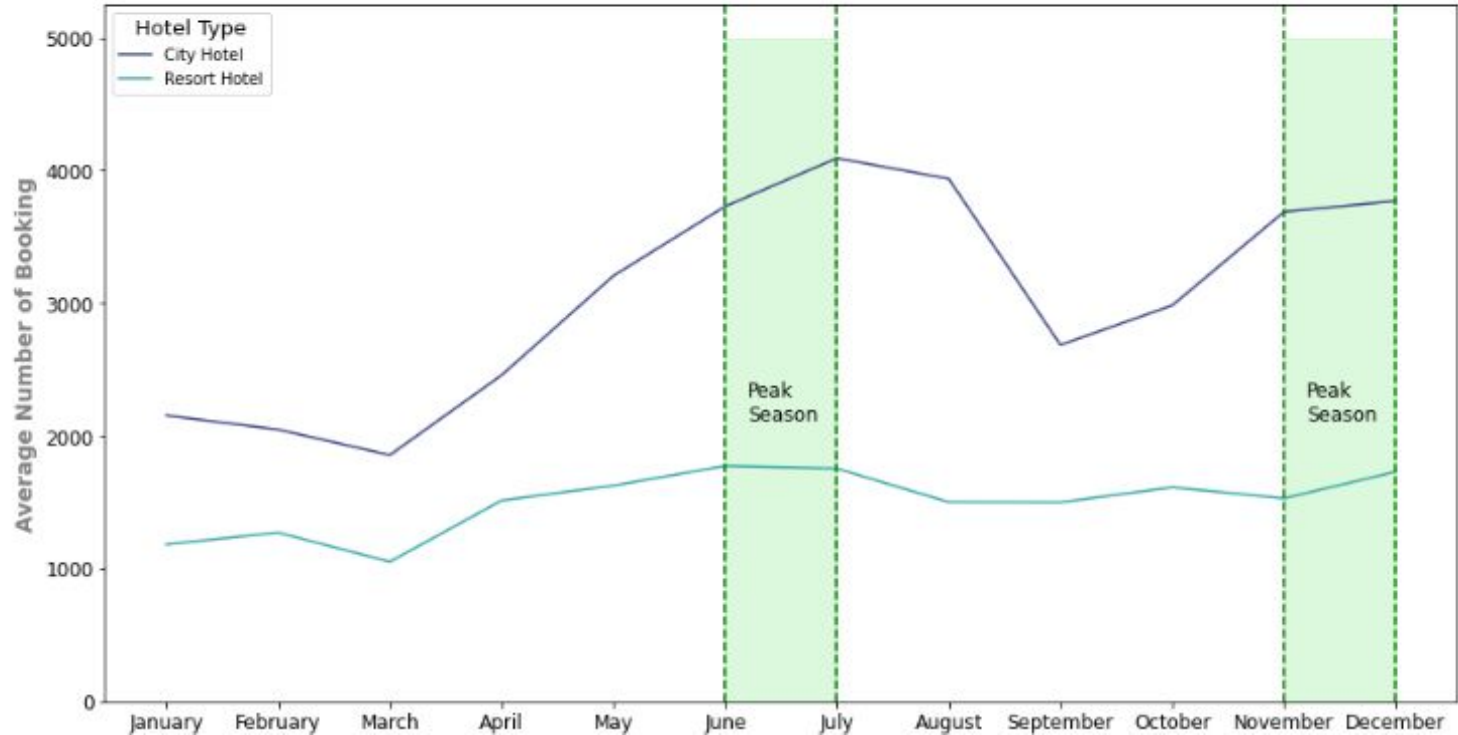
Pemesanan City Hotel dan Resort Hotel paling banyak dilakukan secara online menggunakan Travel Agent (Online TA). Pemesanan yang dilakukan secara Offline melalui Tour Operators (Offline TO) berada di posisi kedua paling banyak dilakukan oleh customers. Selebihnya customers melakukan pemesanan secara langsung, groups, corporate, complementary, dan Aviation.

Lebih dari setengah pemesanan **City Hotel** sebesar 52.58% dilakukan secara online. 20.68% pemesanan dilakukan secara offline, 10.83% pemesanan dilakukan secara langsung, dan pemesanan paling sedikit dilakukan melalui aviation sebesar 0.39%. Dari data tersebut terlihat hanya City Hotel yang menerima pemesanan melalui Aviation.

Pemesanan **Resort Hotel** paling banyak dilakukan secara online sebesar 39.93% dan paling sedikit dilakukan melalui Complementary sebanyak 0.59%.

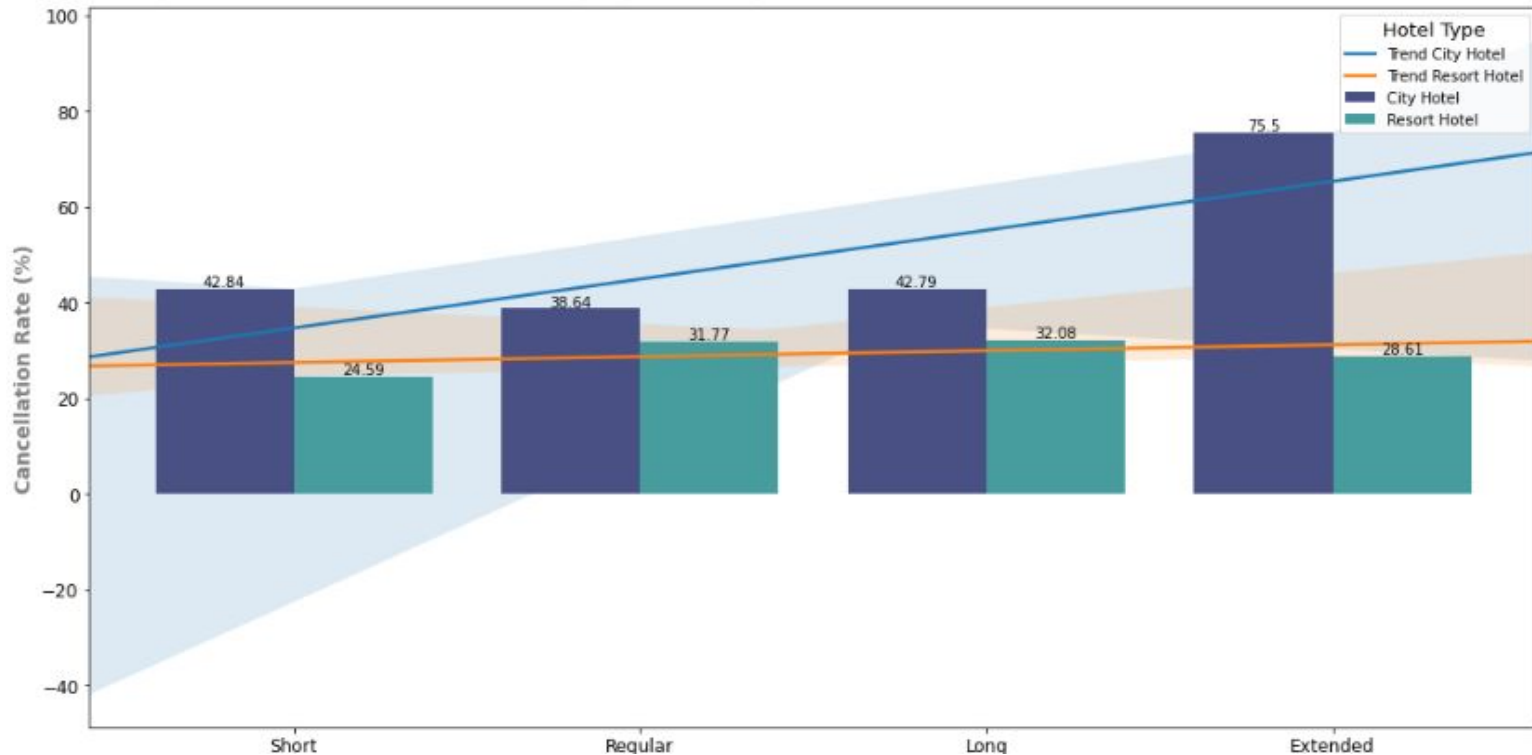


## Average Number of Bookings per Month



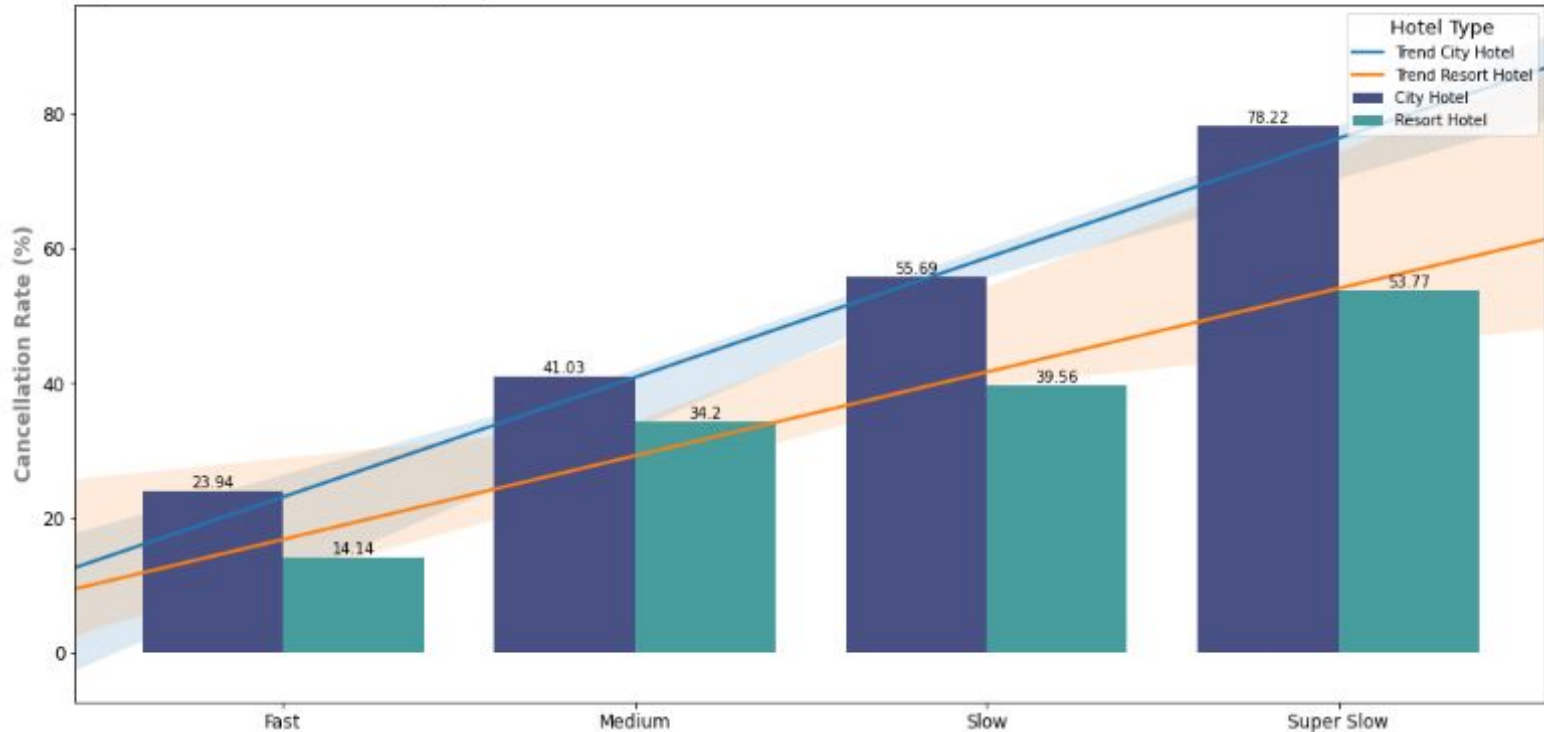
Dilihat dari rata-rata jumlah pemesanan hotel per bulan, City Hotel maupun Resort Hotel mengalami lonjakan pemesanan selama peak seasons. Peak Seasons adalah puncak musim liburan yang biasanya terjadi saat akhir tahun hingga awal tahun. Di Indonesia, peak seasons terjadi pada bulan Juni - Juli & November- Desember.

## Impact Stay Duration on Cancellation Bookings



Berdasarkan bar chart di atas terdapat tren positif antara cancellation rate dan stay duration, yang berarti bahwa durasi menginap mempengaruhi pembatalan pemesanan hotel. Semakin lama durasi menginap, maka semakin tinggi kemungkinan customer melakukan pembatalan pemesanan hotel dan sebaliknya, durasi menginap yang singkat memiliki kecenderungan pembatalan kecil.

## Impact Lead Time on Cancellation Bookings



Terdapat tren positif antara Cancellation Rate dan Lead Time, yang berarti bahwa lead time pemesanan hotel mempengaruhi pembatalan pemesanan. Semakin lama lead time pemesanan hotel maka semakin besar kecenderungan customers untuk melakukan pembatalan pemesanan.

# Summary

- Kota asal tamu yang paling banyak melakukan reservasi City Hotel & Resort Hotel adalah Kota Denpasar.
- Reservasi Hotel paling banyak dilakukan secara online menggunakan Travel Agent (Online TA).
- Berdasarkan rata - rata jumlah reservasi hotel mengalami lonjakan pada saat peak seasons (Juni - Juli & November - Desember).
- Durasi menginap (stay duration) dan Lead Time memiliki pengaruh terhadap pembatalan pemesanan hotel.

# Recommendation

City Hotel maupun Resort Hotel harus memberlakukan kebijakan pembatalan (*cancellation policy*) yang tidak merugikan hotel, seperti tidak memberikan *fully refund* untuk reservasi yang dilakukan saat *peak seasons*.