

Investigate Business Hotel using Data Visualization



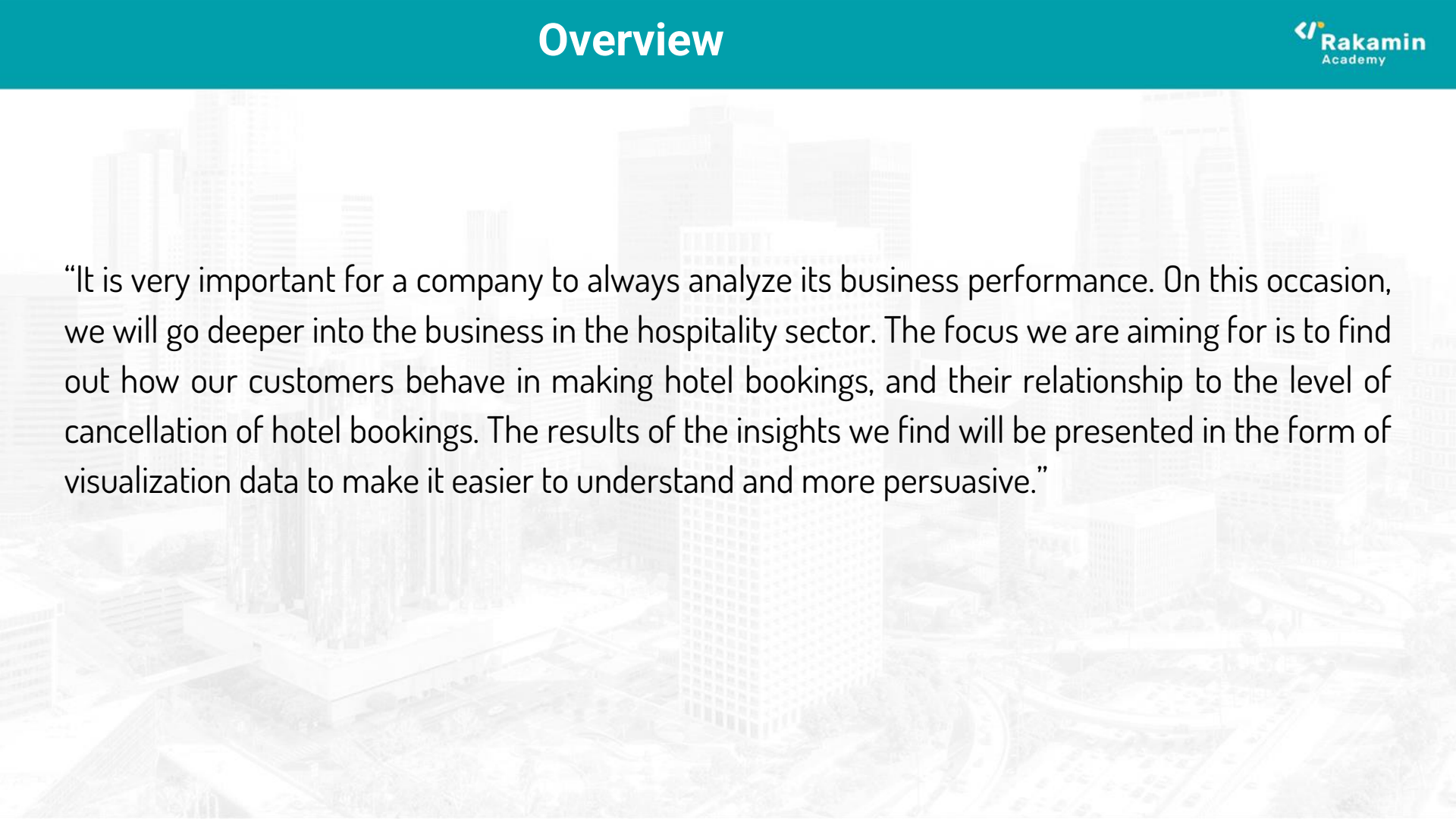
Created by:

Moh. Harwin Prayoga

mh.prayoga11@gmail.com

<https://www.linkedin.com/in/mhprayoga/>

“Yoga is an Engineering Physics graduate at Institut Teknologi Sepuluh Nopember that passionate about Data Science and Data Analysis. He joined a data science bootcamp to forge his programming skill in Python and RDBMS querying with SQL and gratefully his final project team got awarded as The Winner of Best Final Project Team. With Data Science skills and experience, he believes that he is able to deliver strategic insights and recommendations through data to achieve company goals.”

A faded, light-colored background image of a city skyline with various skyscrapers and buildings, serving as a backdrop for the text.

“It is very important for a company to always analyze its business performance. On this occasion, we will go deeper into the business in the hospitality sector. The focus we are aiming for is to find out how our customers behave in making hotel bookings, and their relationship to the level of cancellation of hotel bookings. The results of the insights we find will be presented in the form of visualization data to make it easier to understand and more persuasive.”

DATA PREPROCESSING

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

➤ Data Exploration

- There is "Undefined" in the meal, market_segment, and distribution_channel columns.
- The company and agent columns have numeric values, possibly referring to the id of each company and agent.
- Column children, agent, company uses the data type "float64" but does not have a decimal value.

➤ Handling Missing Value

- Features that have a value of 'Null' : children, city, agent, company.
- The children feature contains very few 'Null' values. Can be filled with '0'.
- The enterprise feature contains many 'Null' values. Can be filled with '0'. The 'Null' value of the agent feature can be filled with '0' (not from the agent).
- The 'Null' value of the city feature was decided to be deleted.

➤ Checking Duplicated Data

- Found 33223 duplicate data but will not be deleted because the dataset does not have unique features such as id / customer_id.

➤ **Fixing Data Type**

- Convert column children, agent, company to “int64”

➤ **Handling Unnecessary Feature & Data**

- Change the value "Undefined" in the meal column to "No Meal"
- Create a "total_guest" column from the addition of the columns adults, children, babies. Then drop the column adults, children, babies.
- Create a “stay_duration” column from the addition of the stays_in_weekend_nights and stays_in_weekdays_nights columns
- Omits rows that have 0 guests and 0 stay_duration

➤ **Statistical Analysis**

- The feature 'adr' (Average Daily Rate) has a negative value which is an anomaly, so a drop on that line is required.

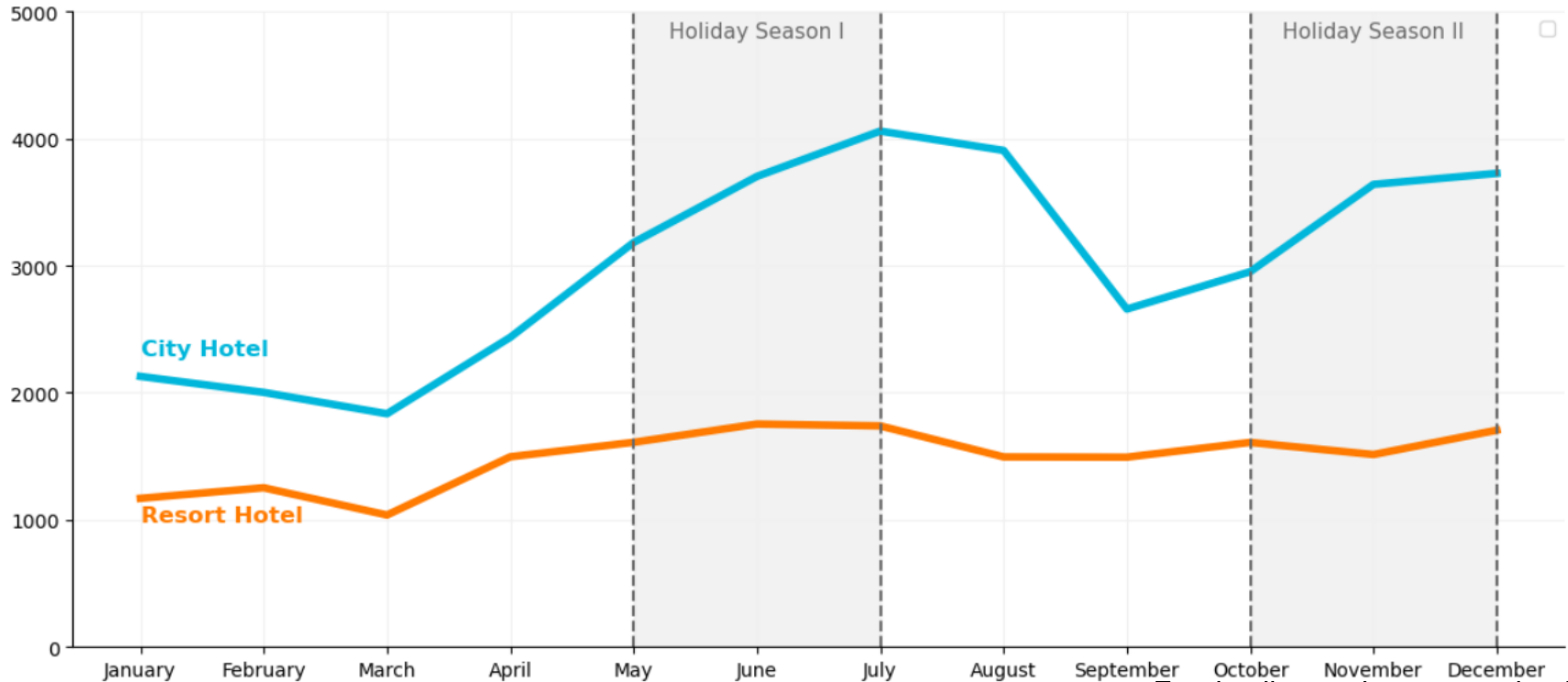
DATA VISUALIZATION 1

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

Monthly Hotel Booking Analysis Based on Hotel Type

The Average Amounts of Bookings are Increasing During the Holiday Season

Both types of hotels tend to decrease other than high season (holiday).
The lowest average number of bookings for both hotel types happened in March.



For details, see jupyter notebook [here](#)

Observations:

- Both types of hotels were increasing during the holiday season (Holiday Season I: Ramadhan & Eid Fitr Season, Holiday Season II: Christmas & New Year Season).
- City Hotel have more average amount of bookings than Resort Hotel.
- The lowest number of bookings for both types were happened in March and a significant decrease of bookings for City Hotel happened in September, most possibly due to towards the end of the semester for students.

DATA VISUALIZATION 2

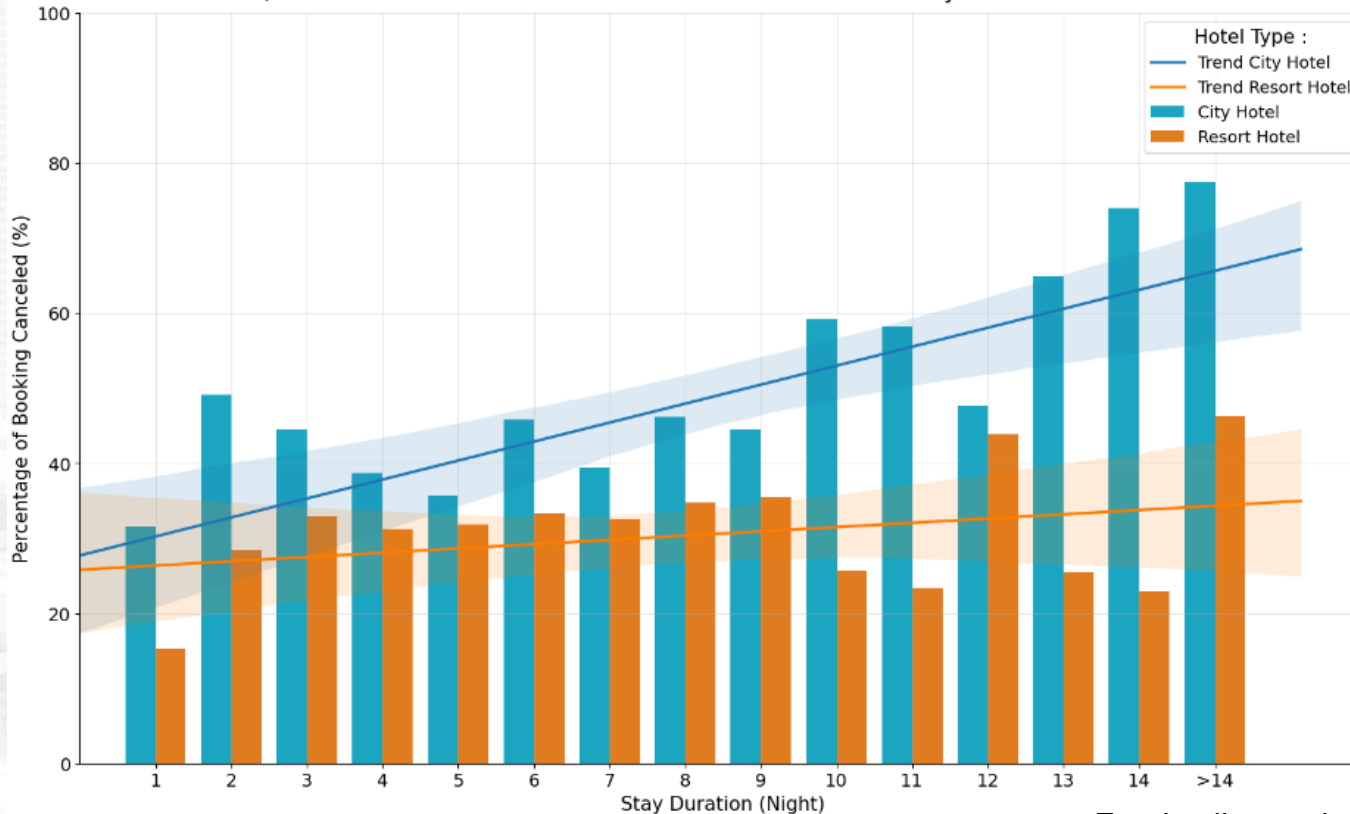
Supported by:
Rakamin Academy
Career Acceleration School
www.rakamin.com

Impact Analysis of Stay Duration on Hotel Bookings Cancellation Rates

Both Types have Positive Trends to Cancellation Percentage

Overall, the more stay duration, the higher chance to do the cancellation.

Also, Resort Hotels have a lesser effect on cancellation than City Hotels



For details, see jupyter notebook [here](#)

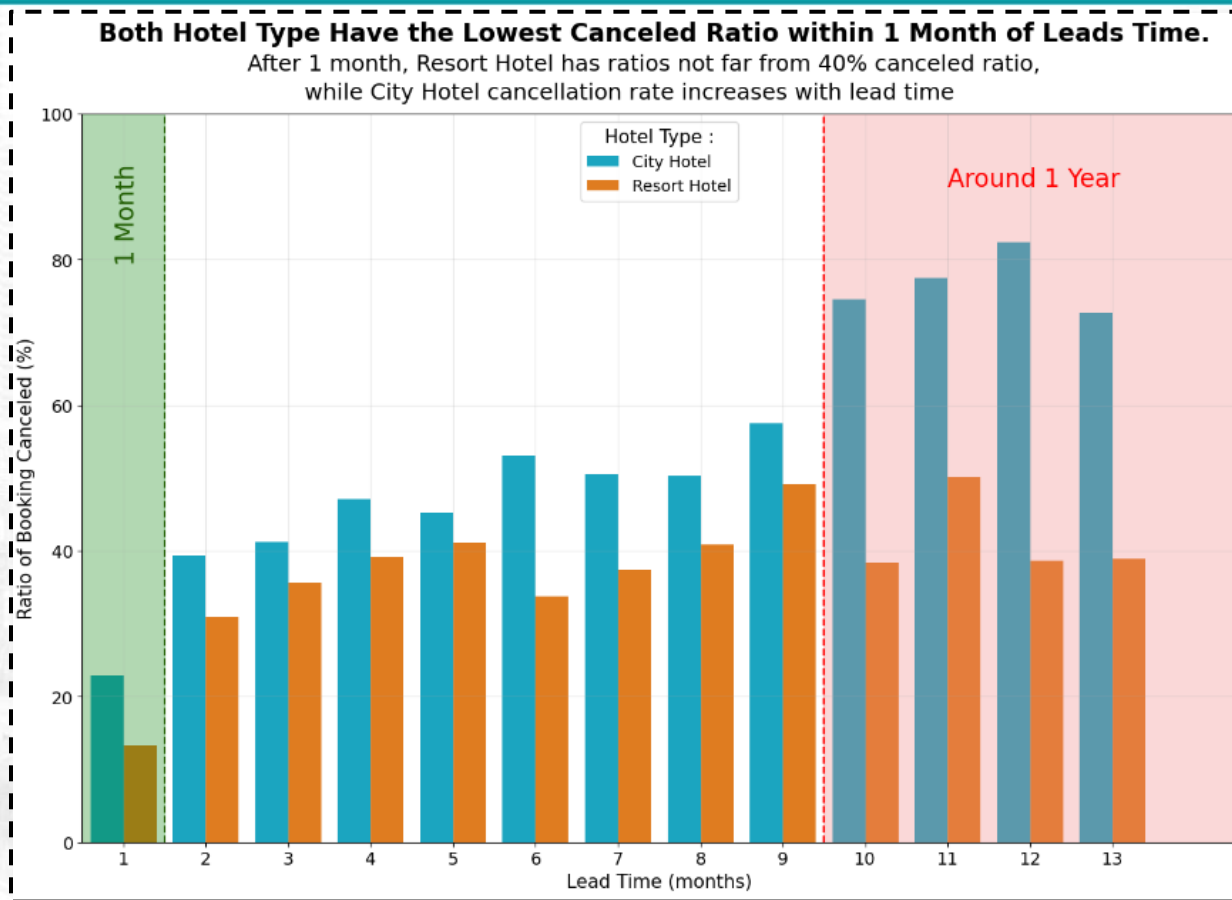
Observations:

- The positive trends on both types indicate that The longer the stay duration that customers ordered, the higher chance the customers do the cancellation for both hotel types.
- The Resort Hotel has a lesser chance of cancellation against stay duration than the City Hotel.

DATA VISUALIZATION 3

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

Impact Analysis of Lead Time on Hotel Bookings Cancellation Rate



For details, see jupyter notebook [here](#)

Observations:

- Overall, the cancellation rate for both types of hotels is affected by customer waiting time.
- The lowest cancellation rate based on lead time for both types occurred in the first month.
- City Hotels have experienced a significant increase in cancellation rates in just over 10 months.
- Resort Hotel cancellation rate based on lead time is more stable than City Hotel.