



SUPPORT BY RAKAMIN ACADEMY

Investigate **Hotel Business**

U S I N G D A T A V I S U A L I Z A T I O N



SUPPORT BY RAKAMIN
ACADEMY



About Me

Istifa has completed her studies and obtained a degree from Universitas Gadjah Mada. She has spent the last 7 months honing her skills in data management. She has gained experience in organizing and analyzing large datasets, creating visualizations and reports, and using statistical tools to draw insights. Her interest in both "manufacture" and "technology" has led her to pursue opportunities to apply her knowledge and skills in a wide range of industries. To further improve her skills in the field of data, She has completed several courses in data science boot camp, which has provided her with a deeper understanding of data analysis.

She is excited to continue developing her expertise in this dynamic field and contribute her knowledge to the growth of the company.



SUPPORT BY RAKAMIN
ACADEMY



Case Study

Continuously evaluating business performance is crucial for any company, as it enables them to take appropriate actions to address any issues that arise. In this project, we will specifically examine the performance of the hotel industry by identifying customer behaviors related to hotel bookings and cancellations. Our primary objective is to analyze three factors: monthly hotel bookings by type, the impact of length of stay on hotel booking cancellations, and the impact of lead time on hotel booking cancellations. We will then visualize the insights gained from this analysis for easier comprehension.



Data Preprocessing

Prior to analyzing data, it is necessary to process it to enable insights to be derived. Additionally, raw data may contain inappropriate or missing values, making data processing essential. This time, the data processing stages will be implemented through the following approaches:

- **Addressing missing values.**
- **Converting data types.**
- **Addressing incorrect values** (both categorical and numerical data).





Handling missing values & change data type

The data on the side reveals that there are **four columns or features that lack certain values**, specifically the "**company**," "**agent**," "**city**," and "**children**" features. For the "company" and "agent" features, missing data should be filled with a value of 0 to indicate that the customer made the order without the help of an agent or company. The "city" feature should be filled with the value "unknown" as there is no clear information available. The "children" feature should be filled with a value of 0 assuming that the customer did not bring their child. Additionally, the data also involves some changes in data types in certain features, such as "children," "agents," and "company" features being transformed from float to integer, likely due to the fact that these values represent discrete counts or categories.

#	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



DATA PREPROCESSING

Handling incorrect values



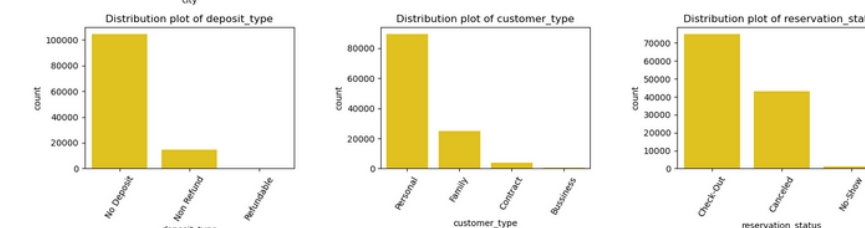
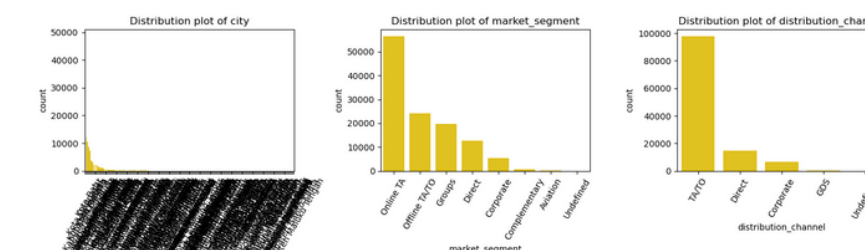
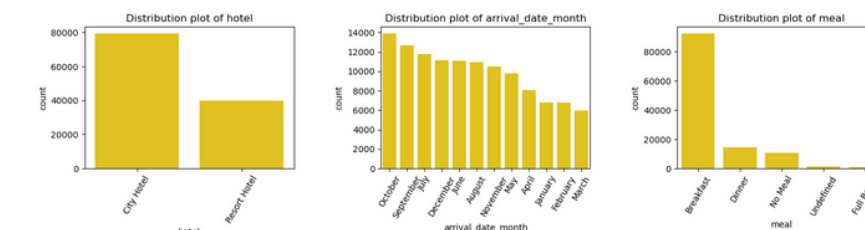
Observed unsuitable values in both category and numeric data features.

a. Categorical Data

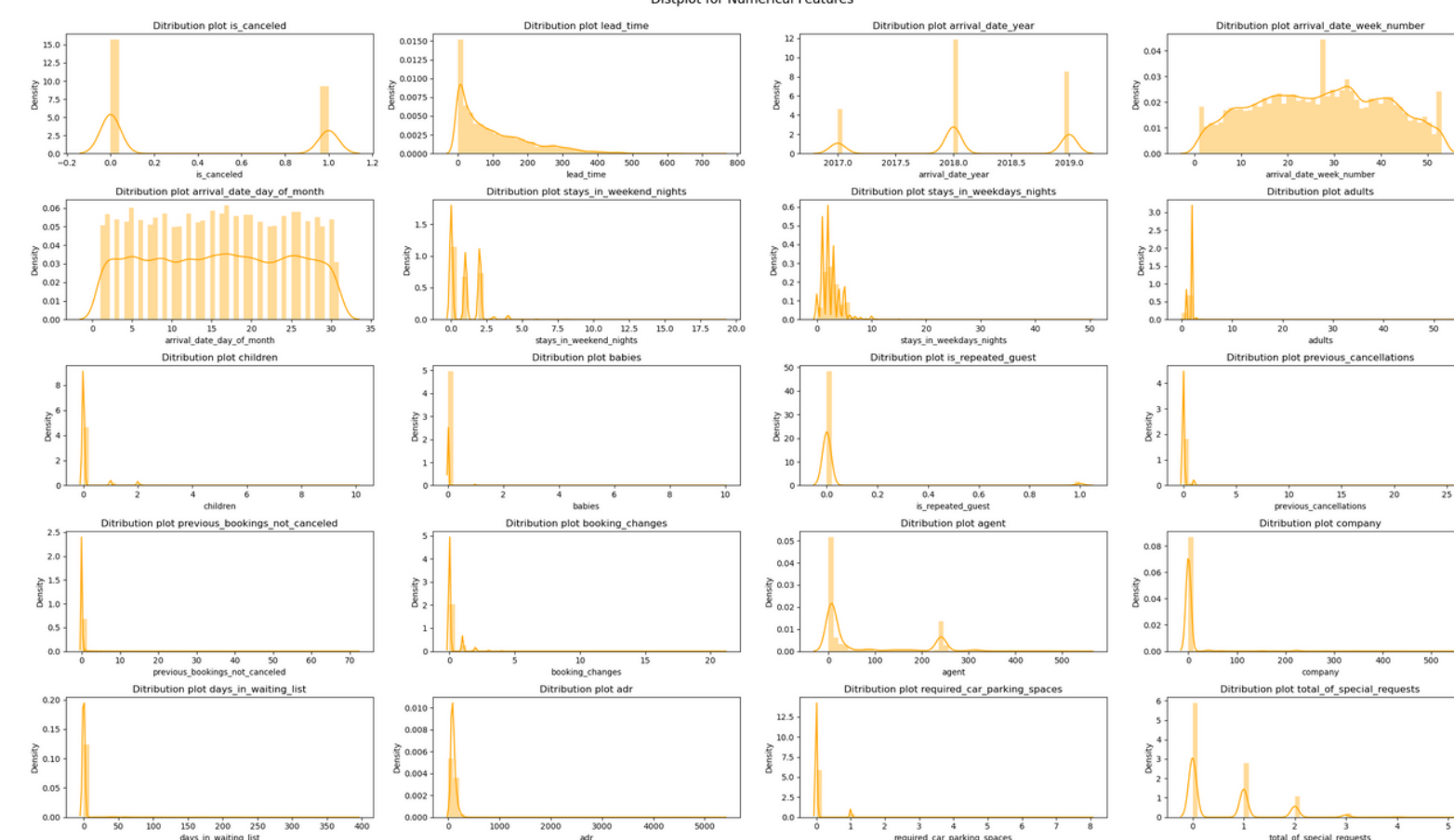
- In categorical data, "market_segment", "distribution_channel", and "meal" features contain "undefined" values that require handling.
- "Undefined" values in "market_segment" and "distribution_channel" will be replaced with the most common values: "Online TA" and "TA/TO", respectively.
- "Undefined" values in the "meal" feature will be replaced with "No Meal" due to their similar meaning.

b. Numerical Data

- Unnecessary data deletion is carried out on numerical data, specifically removing records with 0 guests and 0 nights.
- To facilitate the deletion process, a new feature called "total_guests" is created by combining the "adults," "children," and "babies" features.
- A new feature called "total_days" is also created by combining the "stays_in_weekend_nights" and "stays_in_weekday_nights" features.
- The data with a "stay_duration" of 0 and "total_guests" of 0 are deleted.



Categorical data



Numerical data



DATA AND STEPS

Monthly Analysis of Hotel Bookings Segmented by Hotel Type

In the hospitality industry, customer behavior when booking hotels is crucial to increase income. Identifying popular hotel types and their seasonal demand is essential. To achieve this, we will compare monthly hotel bookings based on their types.

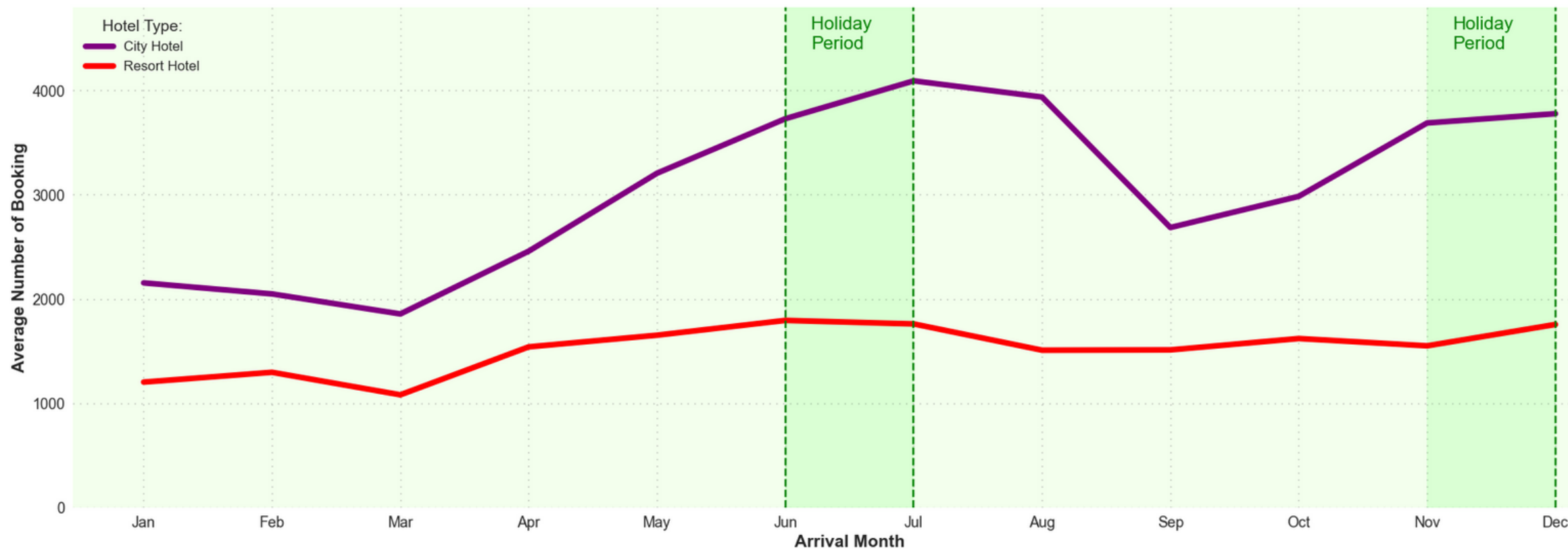




MONTHLY ANALYSIS OF BOOKINGS SEGMENTED BY HOTEL TYPE

Number of guest from hotel type increased during holidays

In August and September, the number of guests at the "City Hotel" decreased, while both hotels have less guests during not holiday (Jan-Mar)



Infographic About Booked

During the holiday period, both hotels experienced an increase in guests, with the "city hotel" having more visitors than the "resort hotel." In Indonesia, most schools have holidays from June to July, which is a time when families tend to travel together, explaining the rise in visitors during those months. This information can be useful in determining the timing of marketing implementation. However, from August to September, the "city hotel" experienced a significant decrease in the number of guests due to the school period. March had the lowest average number of bookings for both hotel types.



DATA AND STEPS

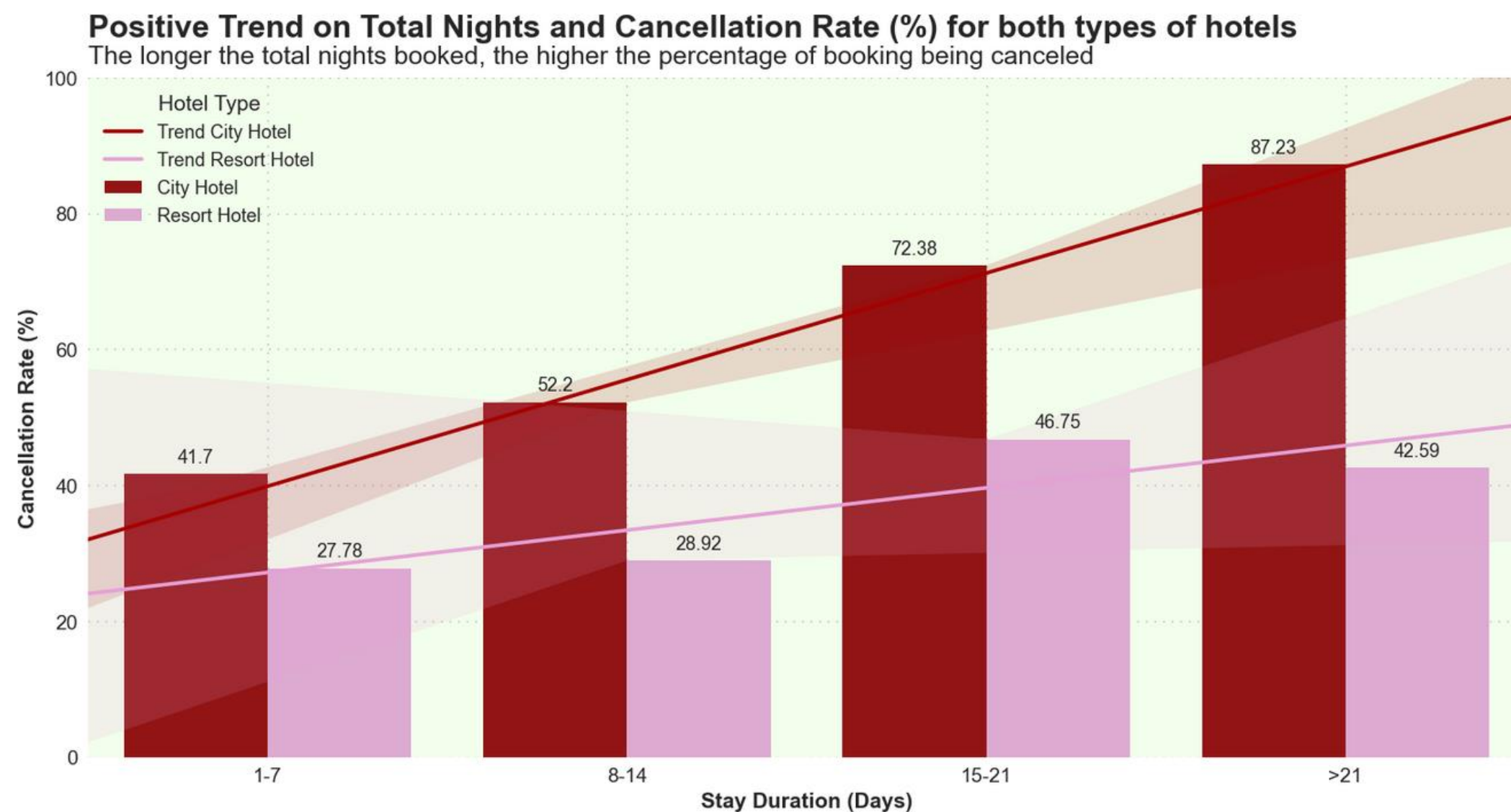
Analyzing the Influence of Length of Stay on Hotel Booking Cancellation Rates.

To evaluate the success of a hotel business, measuring the cancellation rate of bookings is crucial in addition to analyzing customer booking behavior. A high number of cancellations can negatively impact a hotel's performance. Thus, it is important to identify the factors that contribute to booking cancellations, including the length of stay. This analysis will examine the impact of the length of stay on hotel booking cancellation rates.





ANALYZING THE INFLUENCE OF LENGTH OF STAY ON HOTEL BOOKING CANCELLATION RATES



Infographic About Cancellation

The more nights booked, the more cancellations occur. "City hotels" have a higher rate than "resort hotels". The cancellation rate for "city hotels" is highest for stays over 21 days (87.23%) and for "resort hotels" it is highest for stays between 15-21 days (46.75%). To prevent this, hotels should have a cancellation policy where the cancellation fee increases with the length of stay.



DATA AND STEPS

Examining how the lead time influences hotel booking cancellations

The impact of the time gap between hotel bookings and the arrival date on the cancellation rate is being examined in this study, considering that hotels accept bookings well in advance, with varying time intervals ranging from a few days to several months.



2017 February

M	T	W	T	F	S
		1	2	3	4
6	7	8	9	10	11
13	14	15	16	17	18
20	21	22	23	24	25
27	28				

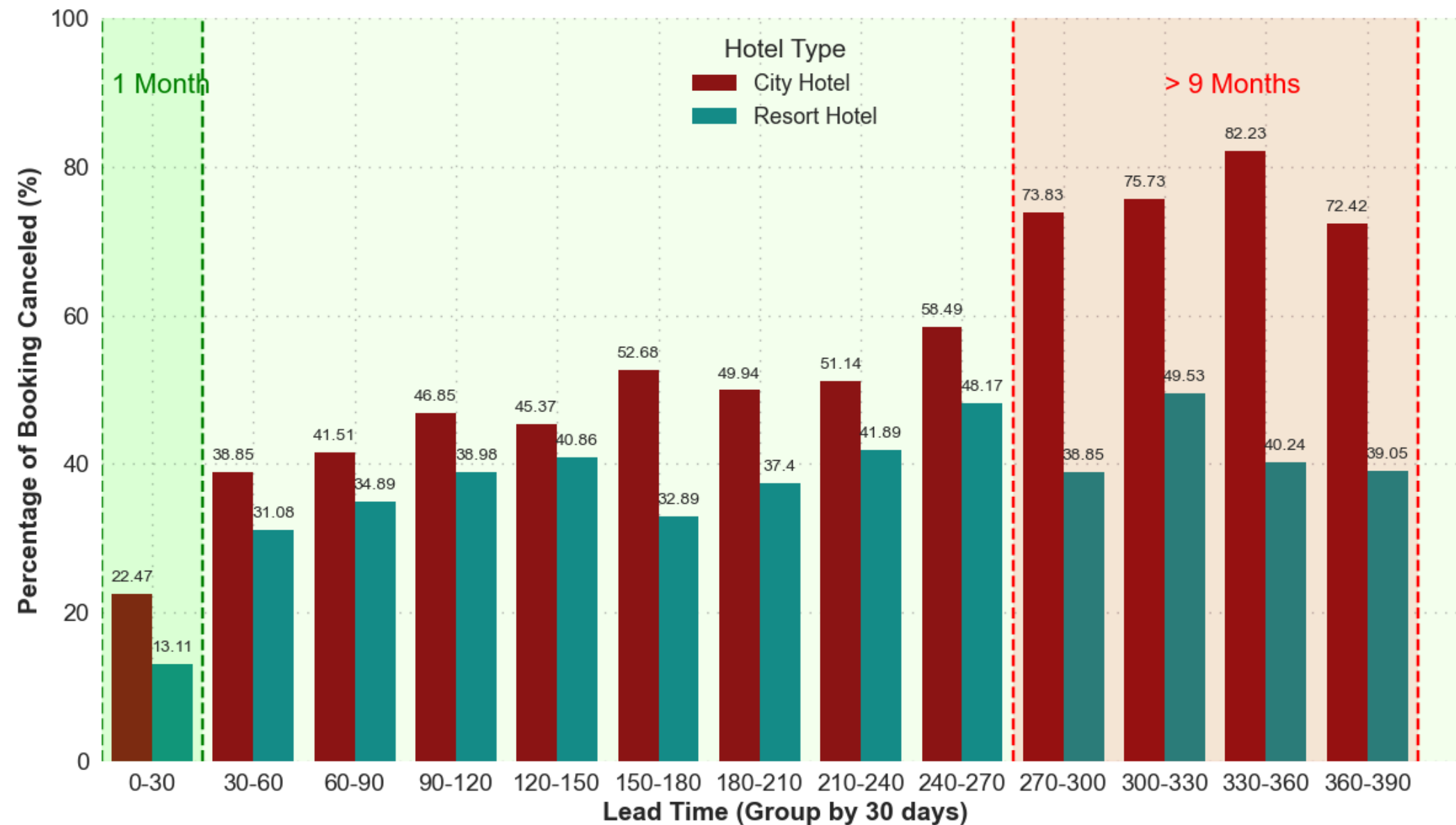




EXAMINING HOW THE LEAD TIME INFLUENCES HOTEL BOOKING CANCELLATION

Duration between booked and arrival date of customer affects the likelihood of hotel booking cancellations.

Both hotel types has lowest cancellation rate of bookings on 1 month lead time; city hotel (22.47%) and resort hotel (13.11%)
City hotels experience a cancellation rate of over 70% when bookings are made for more than 9 months in advance.



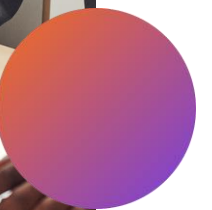
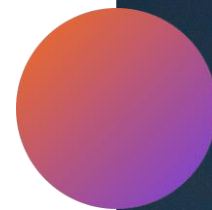
Infographic About Cancellation

The more nights booked, the more cancellations occur. "City hotels" have a higher rate than "resort hotels". The cancellation rate for "city hotels" is highest for stays over 21 days (87.23%) and for "resort hotels" it is highest for stays between 15-21 days (46.75%). To prevent this, hotels should have a cancellation policy where the cancellation fee increases with the length of stay.



DATA AND STEPS

Business Insight



1

To reduce the cancellation rate, hospitality businesses can introduce a penalty system for cancelled hotel bookings.

2

A maximum booking lead time of three months can be implemented by the hospitality industry to reduce the cancellation rate.

3

A reminder system can be put in place, particularly for bookings made well in advance.

4

Hotels can offer marketing promotions or special services in peak seasons of June-July and November-December to attract more customers.



BY ISTIFA SHANIA

Thank You

C I A O