Investigate Business Hotel using Data Visualization

By:

Irham Maulana Ahmadi

Full Repository:

Irham Maulana Ahmadi

irhammahmadi@gmail.com
www.linkedin.com/in/irhamahmadi
www.linktr.ee/IrhamMAhmadio7

"I am a graduate of the Faculty of Petroleum Engineering UPN "Veteran" Yogyakarta, seeking an opportunity to apply skills in data analysis and data science. I am skilled in operating Python for data processing, visualization, and machine learning. Tableau for data reporting, and SQL. highly enthusiastic individual, eager to learn new things, and can work in a team or individually, "

OVERVIEW

It is essential for a company to always analyze its business performance. On this occasion, we will delve deeper into business in the hospitality sector. Our focus is to find out how our customers behave in making hotel reservations and its relationship to the rate of cancellation of hotel reservations. We will present the results of the insights we find in data visualization to make it easier to understand and more persuasive.

DATA PROCESSING

- Duplicate Checking

    this process is carried out to see whether there are duplicate rows. If duplicate rows are found, action will be taken which is deemed most appropriate to the condition of the existing data.

    After a meticulous examination of duplicates, numerous redundant rows were pinpointed. The absence of unit identification in each row posed challenges in validating data accuracy. In the end, it was decided to keep one entry from each set of duplicates.

- Null Values Handling

    After doing some analysis, we found that several columns have Null values. Columns with a relatively small number of Null values, such as the `children` and `city` columns, will have their rows deleted, while the `agent` and `company` columns, which have a relatively high percentage of Null values, will be seen first to see whether the whole row makes sense to be Null or not.

- Data Correction

    At this stage, values that were found to have incorrect values will be corrected or deleted. some of them are:
    - Incorrect Day of Month (Correction)
        - February = dates exceed 28, and not leap years.
        - September = dates exceed 30
    - Incorrect Week Number (Correction)
        `arrival_date_week_number` calculation starts at the beginning of March

- adr values (Delete)
  - adr values <= 0
- Stay Duration (Delete)
  - Stay duration = 0
- Undefined Value
  - `meal` (Correction)
  - `distribution_channel` (Delete)
  - `market_segment` (Delete)
- Unnecessary Rows (Delete)
  With several criteria taken into consideration for rows removal:
  - Rows without Customer
  - Rows without Adult
  - Rows with too many babies but accompanied by relatively few adults

- Data Type Conversion
  Columns containing data types considered unsuitable will be changed.
  - `children`: float -> int
  - `total_guests`: float -> int
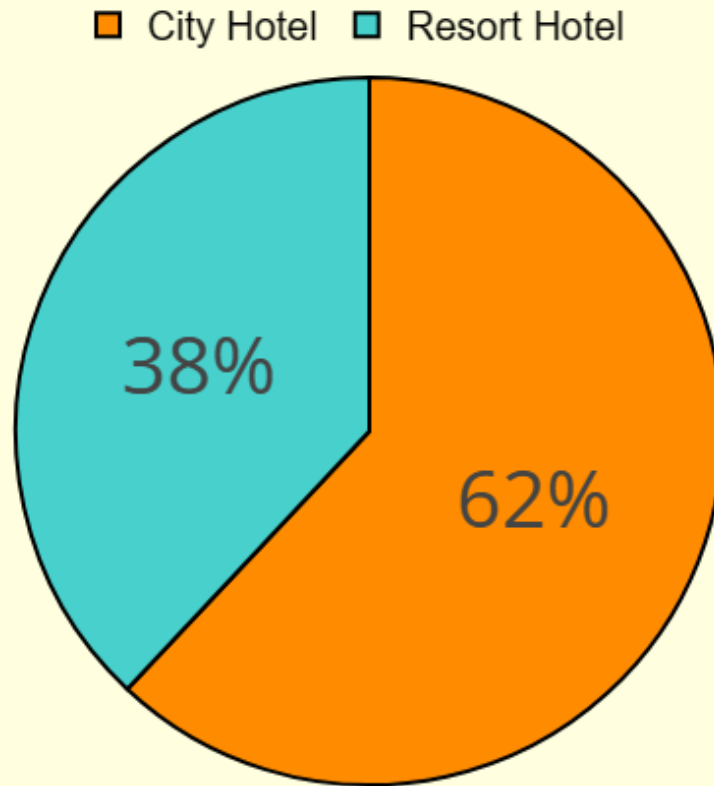
INSIGHT AND VISUALIZATION

1. Transaction

A. Pie Chart for Comparing Total Transactions
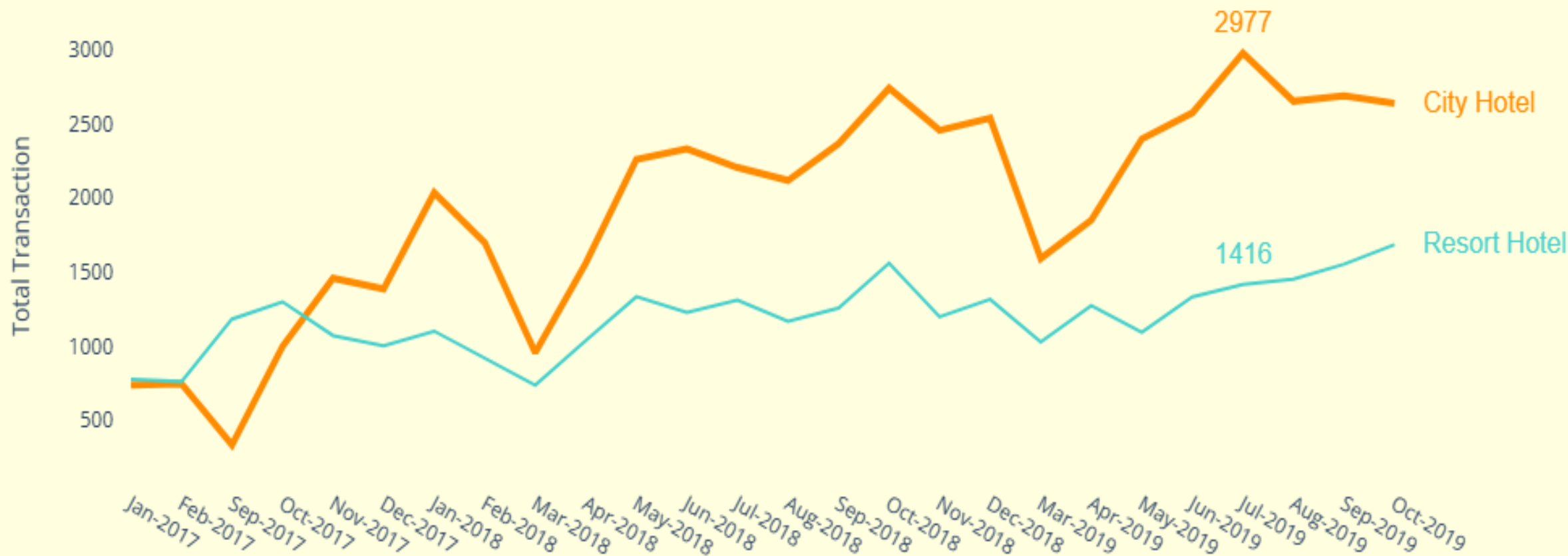
The City Hotel dominates transactions each month.

☐ City Hotel ☐ Resort Hotel

38%

62%

**62%** of the total transactions are conducted for **City Hotels**.

- Upon further elaboration of the transaction history, it becomes increasingly evident that City Hotels dominate the transactions. It is observed that starting around November 2017, the number of transactions for City Hotels consistently surpasses those for Resort Hotels, with the largest difference recorded in July 2019, reaching approximately 1561 transactions.

- Upon a more detailed examination of the transaction history, it becomes increasingly clear that City Hotels are predominant in transactions. From around November 2017 onward, the number of transactions for City Hotels consistently exceeds those for Resort Hotels, with the most significant difference recorded in July 2019, amounting to approximately 1561 transactions.

- The graph also highlights a distinct pattern, consistently showing that March records the lowest transactions each year.

City Hotel consistently leads in monthly average transactions

Both hotel types exhibit a consistent pattern, with Avg. transaction peaks in June to August during the mid-year holiday season and the lowest transaction in February.
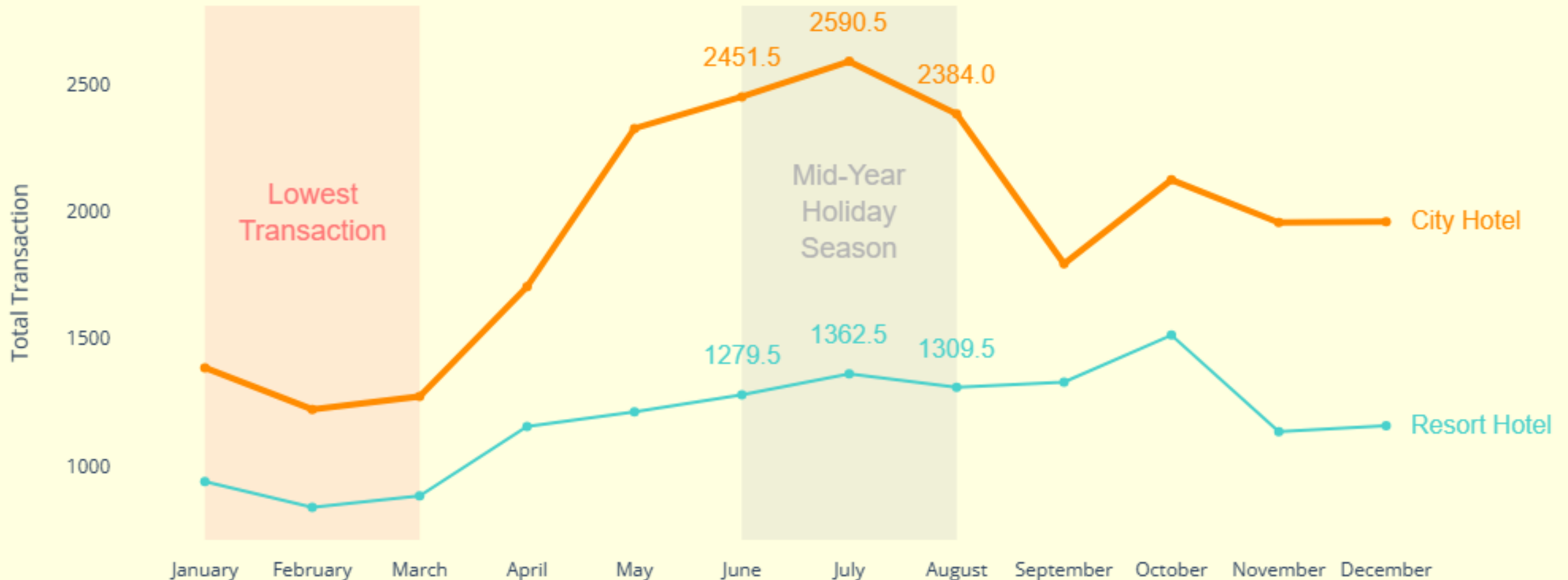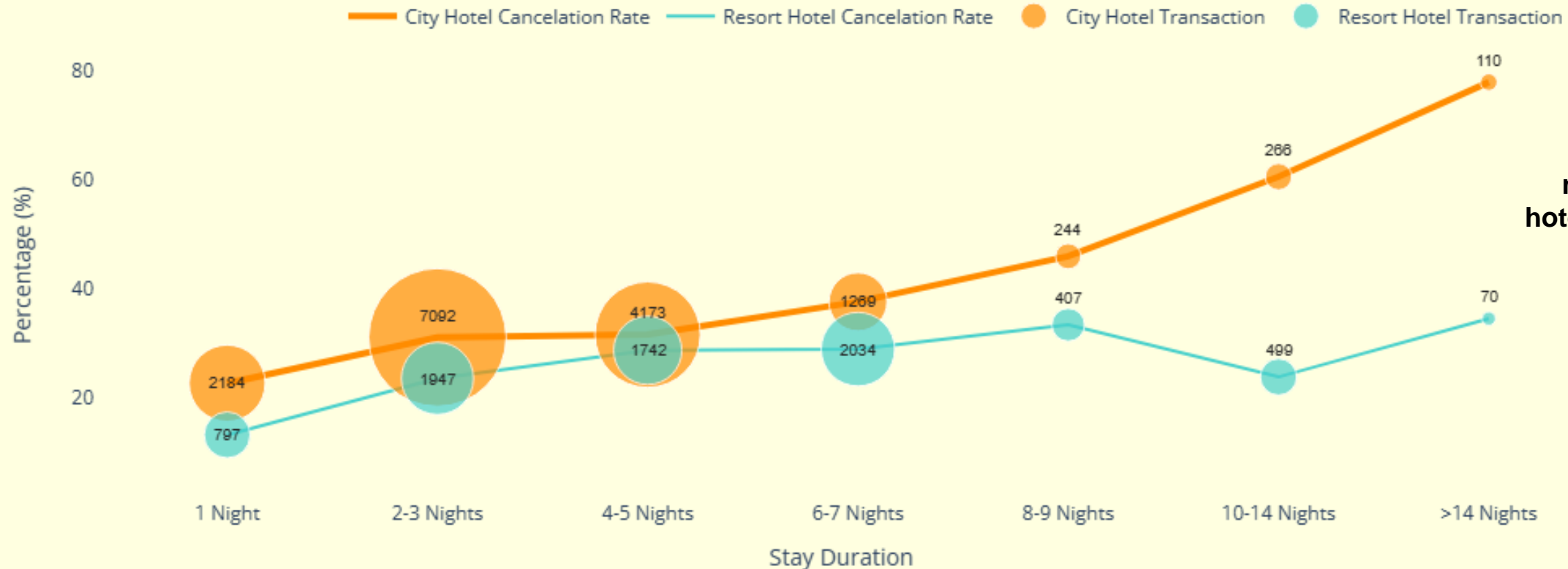
- Having previously determined the consistent upward trend in transaction volume each month, a plot illustrating the average monthly transaction volume was generated, as depicted in the above graph. The graph clearly indicates a noteworthy impact of the mid-year holiday season on hotel room reservations. Specifically, the months of June to August stand out for both hotel types, registering high levels of room bookings, with city hotels experiencing their highest reservation rates during these months.

- As noted in the earlier analysis, March emerges as a month with a relatively low transaction contribution, trailing behind February, which holds the record for the lowest average transactions. This pattern is attributed to the lack of transaction data in February 2019, as revealed in the historical monthly transaction graph. This absence results in a smaller average compared to March, which, based on the historical monthly transaction graph, would likely have been a month with the lowest average transactions.

Cancellation rates rise with longer stays

Guests prefer booking at least one month in advance, but longer lead times increase the likelihood of cancellations.

Legend: City Hotel Cancelation Rate · Resort Hotel Cancelation Rate · City Hotel Transaction · Resort Hotel Transaction

City hotel guests mostly stay for 2-3 nights, while resort hotel guests generally opt for a 6-7 night duration.

City Hotel Transaction values: 2184, 7092, 4173, 1269, 244, 266, 110

Resort Hotel Transaction values: 797, 1947, 1742, 2034, 407, 499, 70

Y-axis: Percentage (%)

X-axis (Stay Duration): 1 Night, 2-3 Nights, 4-5 Nights, 6-7 Nights, 8-9 Nights, 10-14 Nights, >14 Nights

Cancellation Rate Increases as Lead Time Increases

Guests prefer booking at least one month in advance, but longer lead times increase the likelihood of cancellations.
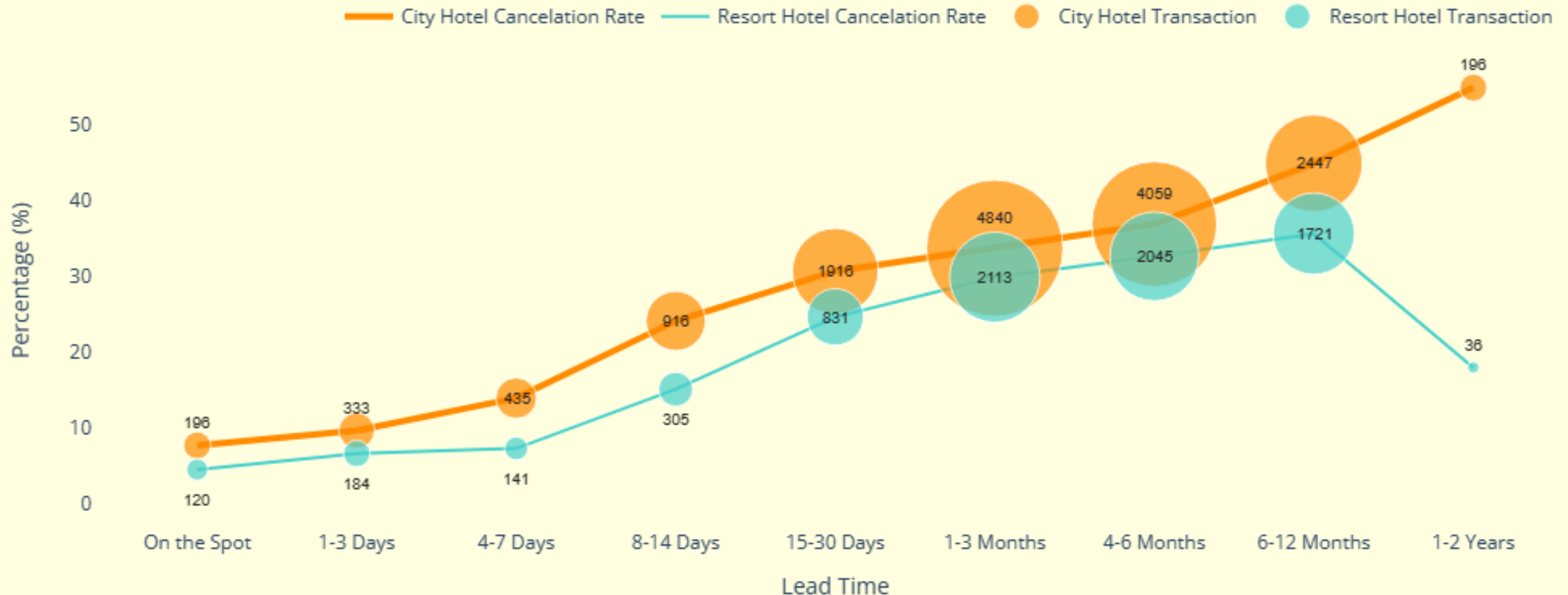
SUMMARY

1. The City Hotel stands out as the most popular type, with room bookings comprising 62% of the total transactions.
2. The mid-year holiday season (June, July, August) substantially contributes to elevated room booking levels for both hotel types.
3. The months in the first quarter (January, February, March) consistently represent the period with the lowest transaction contributions each year.
4. As the duration of the stay increases, the probability of reservation cancellations also rises.
5. City hotel guests usually stay for 2-3 nights, while resort hotel guests generally opt for a 6-7 night duration.
6. Guests prefer booking at least one month in advance, but longer lead times increase the likelihood of cancellations.