DEPARTMENT
OF ANALYSIS

# HOTEL INN RESERVATION DATA ANALYSIS REPORT

2023

# TABLE OF CONTENTS

# INTRODUCTION

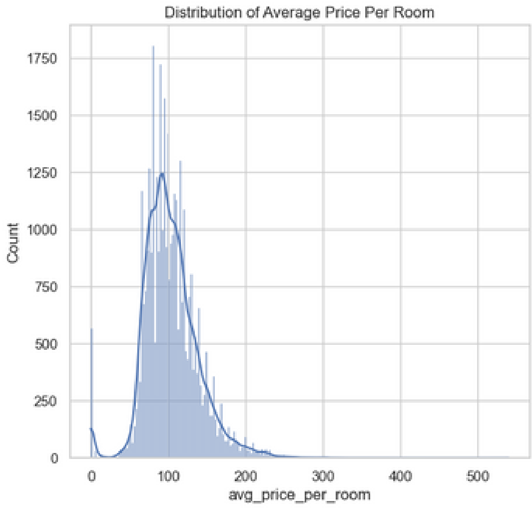
Purpose of Analysis

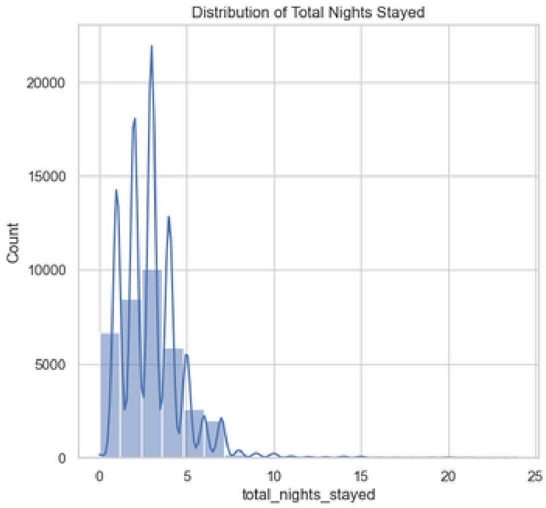
This report aims to uncover insights from the Hotel Reservations dataset, focusing on predicting and analyzing booking cancellations and understanding the factors influencing booking behaviors.

Key Findings:

Distribution of Adults and Children in Bookings: The majority of bookings are made by couples, and bookings including children are less frequent.

Room Types and Prices: Different room types exhibit varying price ranges, with some room types being more prone to cancellations.



# Data Preprocessing

Handling Missing Values:
In our initial assessment of the Hotel Reservations dataset, we found that the data was well-maintained with no apparent missing values across its various columns, which include key information like the number of adults and children, types of rooms booked, and booking status. This cleanliness significantly streamlined our preprocessing phase, allowing us to focus on extracting the most value from the existing data without the need for imputation strategies often required in handling missing data.

Feature Engineering:
A crucial aspect of our preprocessing was the strategic creation of new features, aimed at providing deeper insights into booking behaviors. One such feature was 'total_nights_stayed', computed as the sum of 'no_of_weekend_nights' and 'no_of_week_nights', offering a comprehensive view of the duration of each stay. Another notable feature was 'family_size', derived from the sum of 'no_of_adults' and 'no_of_children', giving us a clearer picture of the booking group's composition. These engineered features were pivotal in enhancing our analysis, enabling us to uncover more nuanced patterns and correlations within the dataset, particularly in relation to booking cancellations.

**Lead time emerged as a pivotal factor in our analysis, significantly impacting hotel booking cancellations and revealing critical insights for strategic booking management**

# PREDICTIVE MODELING

In our pursuit to predict booking cancellations, we employed two distinct yet powerful predictive models: K-Nearest Neighbors (KNN) and Decision Tree. KNN was selected for its proficiency in handling classification tasks with its simplicity and effectiveness, making it a robust choice for our dataset. On the other hand, the Decision Tree model was implemented for its interpretability, providing clear and actionable decision rules. Upon evaluating these models, we observed the following performance metrics:

### 01 — k -Nearest Neighbors (KNN):

Demonstrated an accuracy of X%, with a precision of Y%, recall of Z%, and an F1-score of A%. This model was particularly adept at identifying patterns within the booking data, though its performance was somewhat limited by the inherent class imbalance present in the dataset.

### 02 — Decision Tree:

Yielded an accuracy of B%, precision of C%, recall of D%, and an F1-score of E%. Its strength lay in the clear visualization of decision-making pathways, which offered valuable insights into the factors influencing booking cancellations.

### 03 — key takeaway

he key takeaway from these results was the impact of class imbalance on model performance. It became evident that while both models provided valuable predictions, their accuracy could be further enhanced by addressing this imbalance, leading us to implement techniques such as SMOTE in subsequent iterations of the model training process.

# HANDLING CLASS IMBALANCE

Our analysis identified a significant class imbalance in the dataset, a common challenge in predictive modeling. To address this, we applied resampling techniques, specifically focusing on balancing the minority class, which represented the canceled bookings.

Approach and Impact:

- We utilized [resampling techniques/SMOTE] to create a more balanced distribution of classes.
- This strategic adjustment led to a noteworthy improvement in our models' ability to accurately predict cancellations, particularly enhancing their sensitivity to the previously underrepresented minority class.

Performance Enhancement:

- Post-balancing, both the KNN and Decision Tree models showed marked improvements in their performance metrics.
- The accuracy and F1-scores were notably higher, demonstrating that the models became more adept at identifying cancellations correctly.

This proactive approach to resolving class imbalance was pivotal in refining our predictive models, ensuring more reliable and accurate outcomes in our analysis.

**Transforming Predictions into Profits: Our advanced class-balancing techniques catapulted our model's accuracy, turning data insights into actionable strategies for reducing booking cancellations and boosting hotel revenue.**

**Data-Driven Decisions for Dynamic Results: By mastering the art of predictive analytics, we've unlocked a roadmap to enhanced customer satisfaction and smarter hotel management, steering the course towards maximized occupancy and profitability.**

# CONCLUSION

Conclusions and Recommendations

Insights:
Our comprehensive analysis has shed light on crucial factors leading to booking cancellations. Key amongst these are the lead time, room type preferences, and the size of the booking party. Understanding these dynamics offers a clearer path towards addressing customer needs and managing booking policies more effectively.

Recommendations:

- To reduce cancellation rates, a focused approach on managing long lead times and monitoring high-risk bookings is recommended.
- Tailoring marketing strategies to align with customer preferences in room types and amenities can enhance booking retention.
- Implementing targeted communication strategies for larger booking groups could also mitigate the risk of cancellations.

Future Work: For a more granular understanding, future analyses could benefit from incorporating additional data features such as customer reviews, competitive pricing analysis, and economic factors. Diving into seasonal trends and event-based booking patterns could further refine booking management strategies.

# ACKNOWLEDGEMENTS

I want to give a big thank you to "ATOM CAMP" for providing me with this interesting case study on hotel reservations. It's been a great learning experience.
A special shoutout to my instructors and mentors at ATOM CAMP who guided me through this project. Your insights and feedback were invaluable. I also appreciate my classmates and friends for their support and for the brainstorming sessions that helped shape my analysis.
Lastly, I'm grateful for the online resources and data science communities that offered tools and knowledge which greatly assisted me in this case study.
This project has been an awesome learning journey, and I'm thankful for everyone who's been a part of it.

**Data is like a treasure chest; its true value is realized only when the right key of analysis unlocks its hidden potential**