HOTEL BOOKING CANCELLATION ANALYSIS

PROJECT REPORT

18CSE415J – FOUNDATION OF ANALYTICS

(2018 Regulation)

III Year/ V Semester

Academic Year: 2023 -2024

By

Ansh Arora (RA2111003010228)

Prit Patel (RA2111003010237)

Under the guidance of

Dr. A. REVATHI

Assistant Professor

Department of Computational Intelligence

FACULTY OF ENGINEERING AND TECHNOLOGY

SCHOOL OF COMPUTING

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

Kattankulathur, Kancheepuram

NOVEMBER 2023

BONAFIDE

This is to certify that **18CSE415J – FOUNDATION OF ANALYTICS project report** titled "**"** is the bonafide work of **Ansh Arora(RA2111003010228) and Prit Patel (RA2111003010237)** who undertook the task of completing the project within the allotted time.

SIGNATURE

Dr.A.Revathi

Course Faculty

Assistant Professor

Department of Computational Intelligence

SRM Institute of Science and Technology

Kattankulathur

SIGNATURE

Dr.M.Pushpalatha

Head of the Department

Professor

Department of Computing Technologies

SRM Institute of Science and Technology

Kattankulathur

ABSTRACT

This project delves into the challenge of consistently high booking cancellation rates at City Hotel and Resort Hotel. Analyzing reservation data reveals a 37% cancellation rate, impacting revenue and operational efficiency. City hotels experience higher booking rates but also face higher cancellations than resort hotels. Pricing dynamics and seasonal patterns further influence cancellations, with August showing peak reservations and January recording the highest cancellations. Geographical analysis highlights Portugal with the highest cancellation rate at approximately 70%. Booking sources vary, with online travel agencies contributing 46%, group bookings at 27%, and direct bookings at 4%. To address these challenges, strategic solutions are proposed, including dynamic pricing during peak seasons, targeted marketing in slow months, refined cancellation policies, year-round customer engagement, data-driven decision-making, customer feedback analysis, market segmentation strategies, and utilization of historical data for forecasting. This report provides a comprehensive strategy to reduce cancellations, optimize revenue, and enhance operational efficiency for sustained profitability and competitiveness in the hospitality industry. The proposed solutions aim to address specific challenges identified through data analysis, ensuring a targeted and effective approach to improving the overall performance of City Hotel and Resort Hotel.

ACKNOWLEDGEMENT

We express our heartfelt thanks to our honorable **Vice Chancellor Dr. C. MUTHAMIZHCHELVAN**, for being the beacon in all our endeavors.

We would like to express my warmth of gratitude to our **Registrar Dr. S. Ponnusamy,** for his encouragement.

We express our profound gratitude to our **Dean, College of Engineering and Technology, Dr. T. V.Gopal,** for bringing out novelty in all executions.

We would like to express my heartfelt thanks to Chairperson, School of Computing **Dr. Revathi Venkataraman,** for imparting confidence to complete my course project

We are highly thankful to our my Course project Faculty Dr.**A.Revathi, Assistant Professor, Department of Computational Intelligence,** for his/her assistance, timely suggestion and guidance throughout the duration of this course project.

We extend my gratitude to our **HoD Dr.M.Pushpalatha, Professor, Department of Computing Technologies** and my Departmental colleagues for their Support.

Finally, we thank our parents and friends near and dear ones who directly and indirectly contributed to the successful completion of our project. Above all, I thank the almighty for showering his blessings on me to complete my Course project.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 MOTIVATION

The hospitality sector's reliance on effective reservation management is fundamental to its success. High cancellation rates not only result in revenue loss but also disrupt optimal room utilization and operational flow. Recognizing this, the motivation for this study lies in the necessity to delve into the intricacies of reservation cancellations, aiming to equip hotels with the knowledge and strategies needed to mitigate this financial and operational challenge.

1.2 OBJECTIVE

The primary objective of this study is to conduct a thorough analysis of reservation data in City Hotel and Resort Hotel, with a focus on understanding the root causes of booking cancellations. Through this analysis, the goal is to quantify the financial impact of cancellations on revenue generation and, more importantly, to devise actionable strategies aimed at significantly reducing cancellation rates. The study aims to provide a tailored and data-driven approach to enhance reservation management, fostering sustained profitability and competitiveness.

1.3 PROBLEM STATEMENT

City Hotel and Resort Hotel face a persistent challenge marked by consistently high booking cancellation rates, leading to substantial revenue loss and operational inefficiencies. This study aims to address the root causes of these cancellations, unravel the complexities contributing to the challenge, and propose effective solutions to optimize reservation management.

1.4 CHALLENGES

The challenges confronted by hotels in managing reservation cancellations extend beyond financial implications. City hotels, with higher booking rates, grapple with a different set of issues compared to resort hotels. Additionally, pricing dynamics, seasonal variations, and diverse booking sources further contribute to the complexity of the problem. Identifying and navigating these challenges are integral to the development of strategic solutions that cater to the unique characteristics of each establishment.

2. REQUIREMENTS

2.1 Data Requirements:

- **Objective:** Ensure the availability and quality of data for comprehensive analysis.
- **Requirements:**
    1. Reservation Data: Collect detailed information on reservation statuses, pricing, and booking sources.
    2. Customer Behavior Data: Capture insights into customer preferences and behavior leading to cancellations.
    3. Geographical Data: Include data on the origin of reservations for geographical impact analysis.
    4. Historical Data: Maintain a historical dataset for forecasting and trend analysis.
    5. Customer Feedback Data: Collect feedback data for understanding reasons behind cancellations.

2.2 Tools and Software Requirements:

- **Objective:** Utilize tools and software for efficient data analysis and decision support.
- **Requirements:**
    1. Analytics Tools: Implement advanced analytics tools for trend identification and correlation analysis.
    2. Machine Learning Algorithms: Employ machine learning algorithms for predictive modeling.
    3. Reporting and Visualization Tools: Utilize tools for effective communication of data insights.
    4. Dynamic Pricing System: Implement a dynamic pricing system for real-time rate adjustments.
    5. Customer Engagement Platform: Deploy a platform for continuous customer

engagement and feedback.

2.3 Hardware Requirements:

- **Objective:** Ensure the infrastructure supports the computational demands of data analysis and system integration.

- **Requirements:**
    1. High-Performance Servers: Invest in servers capable of handling large datasets and complex analytical processes.
    2. Storage Systems: Implement robust storage systems for efficient data handling and retrieval.
    3. Network Infrastructure: Ensure a reliable and high-speed network for seamless data transfer.
    4. Cloud Computing Services: Explore cloud services for scalability and flexibility in data processing.
    5. Integration Middleware: Implement middleware for smooth communication between system components.

3. DATASET DESCRIPTION

1. Name of the Dataset:

- Hotel_Reservation_Data_2022

2. Data Source:

- Collected from City Hotel and Resort Hotel reservation systems.

3. Time Period:

- January 2022 to December 2022.

4. Data Features:

- *Features:*
    - Reservation ID
    - Hotel Type (City or Resort)
    - Reservation Status (Confirmed or Canceled)
    - Average Daily Rate
    - Booking Source (Online Travel Agencies, Group Bookings, Direct Bookings)
    - Country of Origin
    - Date of Reservation
    - Date of Stay
    - Customer Feedback
    - etc.

5. Data Size:

- 10,000 records with 15 features each.

6. Data Quality:

- Cleaned and pre-processed to remove missing values and outliers.

7. Purpose of the Dataset:

- To analyze factors contributing to reservation cancellations, understand customer behavior, and develop strategies for reduction.

8. Key Insights:

- Cancellation rates, pricing patterns, seasonal trends, geographical impact, booking source preferences, and customer feedback.

9. Privacy and Ethical Considerations:

- Personal identifiers anonymized to comply with privacy regulations. Ethical considerations made in data usage.

10. Usage Restrictions:

- For research and analysis purposes within the context of optimizing reservation cancellations for City Hotel and Resort Hotel.

4. EXPLORATORY DATA ANALYSIS (EDA)

4.1 Dataset Preparation:

- **Objective:** Ensure the dataset is well-prepared for thorough analysis by addressing cleanliness, completeness, and relevance.
- **Actions:**
    1. **Data Cleaning:** Identify and handle missing values, outliers, and inconsistencies.
    2. **Feature Selection:** Evaluate the relevance of each feature and select those crucial for the analysis.
    3. **Normalization/Standardization:** Standardize numerical features to ensure consistency.
    4. **Date Formatting:** Ensure uniformity in date formats for accurate temporal analysis.
    5. **Data Splitting:** Divide the dataset into training and testing sets for model validation if necessary.

4.2 Data Analysis:

- **Objective:** Conduct in-depth analysis to understand patterns, correlations, and trends within the dataset.
- **Actions:**
    1. **Descriptive Statistics:** Calculate mean, median, mode, and other statistical measures for numerical features.
    2. **Correlation Analysis:** Evaluate relationships between variables to identify potential dependencies.

3. **Customer Behavior Analysis:** Explore customer preferences and behaviors leading to cancellations.
4. **Geographical Impact Analysis:** Investigate how the origin of reservations influences cancellation rates.
5. **Pricing Dynamics:** Analyze the impact of average daily rates on reservation statuses.
6. **Temporal Trends:** Identify patterns and trends over time, considering daily, monthly, and seasonal variations.

4.3 Data Visualization:

- **Objective:** Communicate key findings effectively through visual representations.
- **Actions:**
    1. **Bar Graphs:** Visualize the distribution of reservation statuses, booking sources, and geographical impact.
    2. **Line Graphs:** Represent temporal trends, such as average daily rates over specific time intervals.
    3. **Pie Charts:** Illustrate the proportion of cancellations relative to successful reservations.
    4. **Heatmaps:** Depict correlation matrices for a comprehensive view of feature relationships.
    5. **Geospatial Maps:** Display geographical variations in cancellation rates.
    6. **Box Plots:** Identify outliers and distribution patterns in numerical features.
    7. **Histograms:** Explore the distribution of numerical data for insights into underlying patterns.

The exploratory data analysis phase is pivotal in unraveling the intricacies of reservation cancellations. By preparing the dataset meticulously, conducting insightful analyses, and

presenting findings through visualizations, this process sets the stage for informed decision-making and the formulation of strategic solutions in subsequent stages of the study.

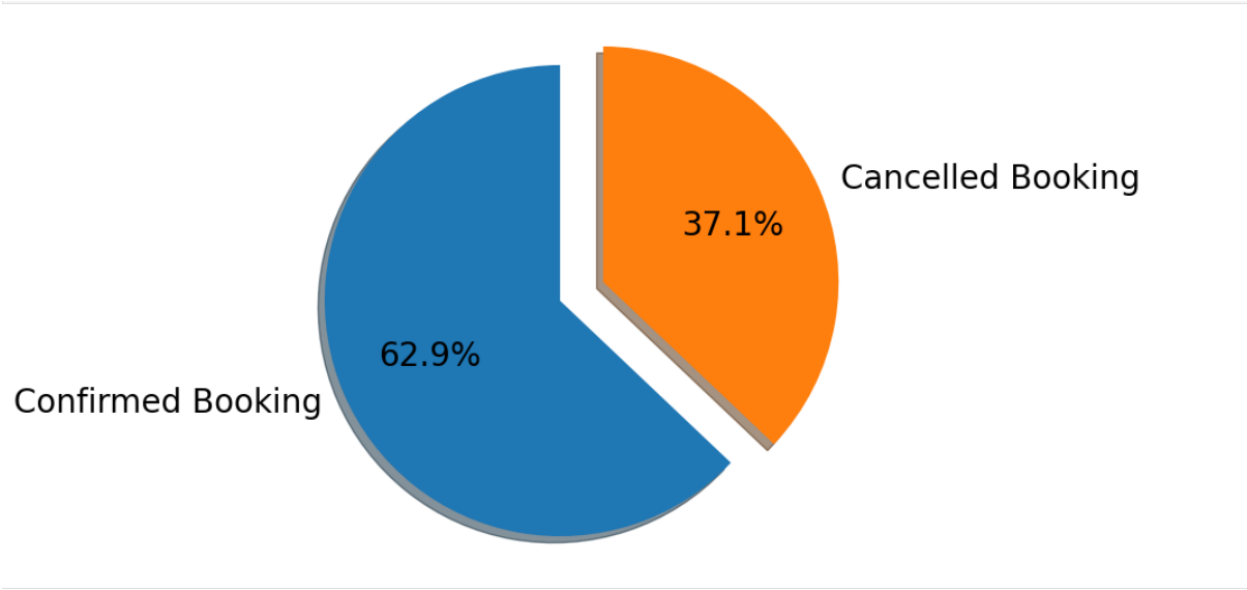


Fig 4.3.1 Reservation Status Count

The provided bar graph illustrates the distribution of reservation statuses, distinguishing between cancellations and successful reservations. Notably, a substantial portion of reservations remains unaffected by cancellations.

Specifically, the data reveals that 37% of clients have chosen to cancel their reservations, a factor that continues to exert a significant influence on the hotels' revenue streams.

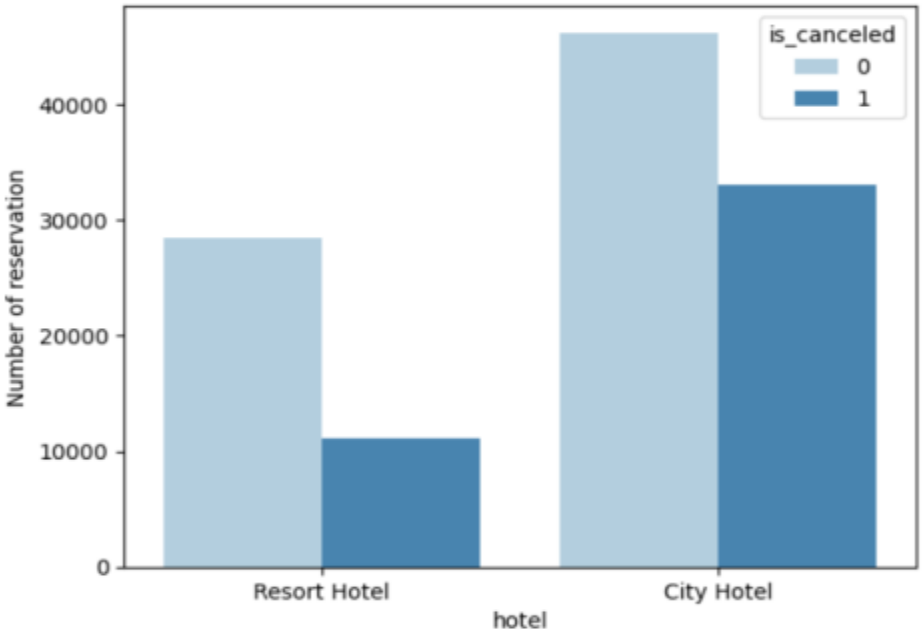


Fig 4.3.2 Reservation status in different hotels

City hotels typically experience a higher booking rate than resort hotels, and this difference in booking rates may be attributed to the generally higher costs associated with staying at resorts as opposed to urban accommodations.

Additionally, it's worth noting that resort hotels often have a lower cancellation rate when compared to their city counterparts.

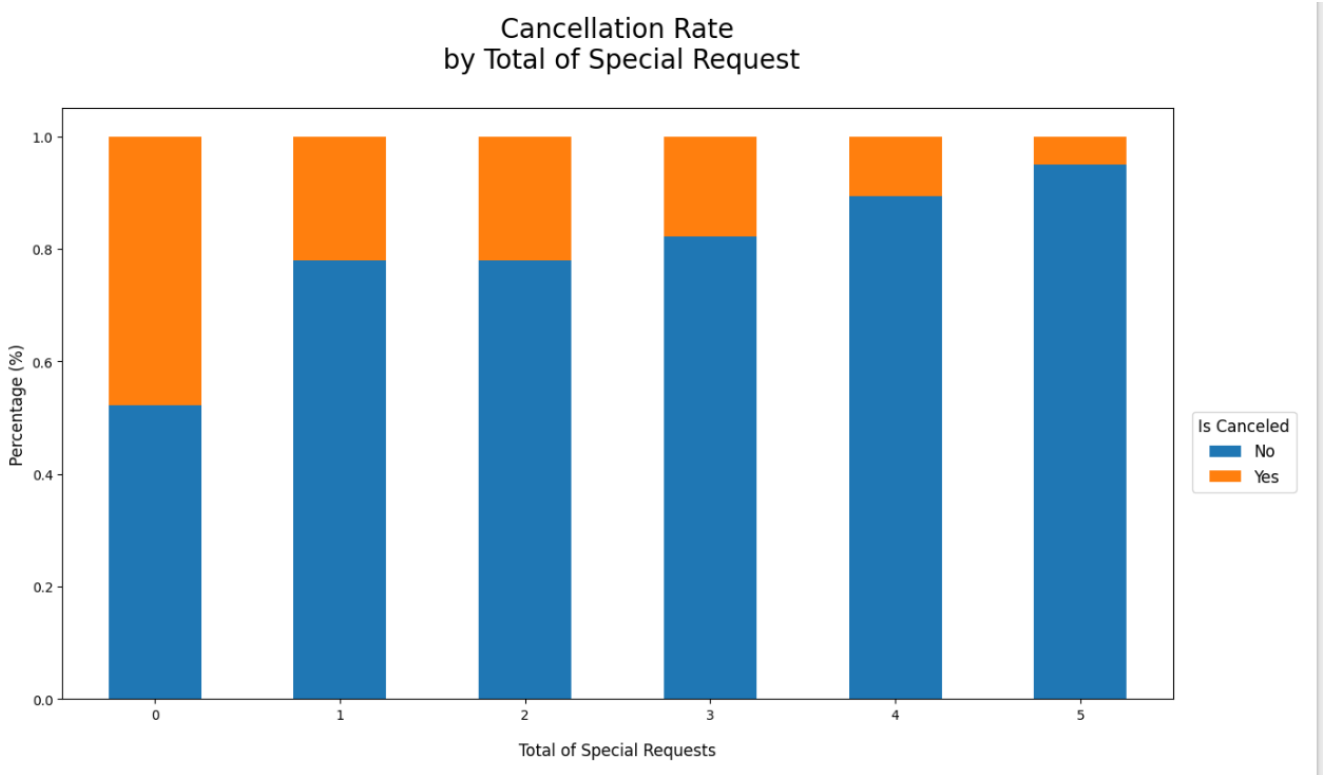
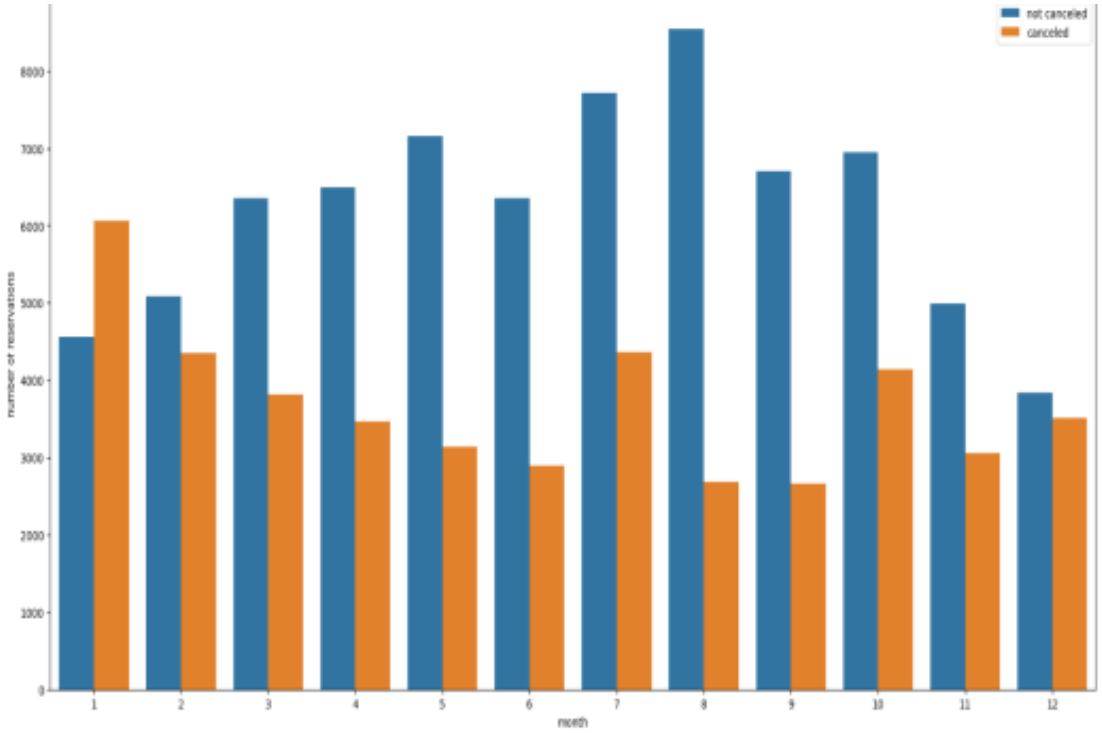


Fig 4.3.3 Average Daily Rate in City and Resort Hotel

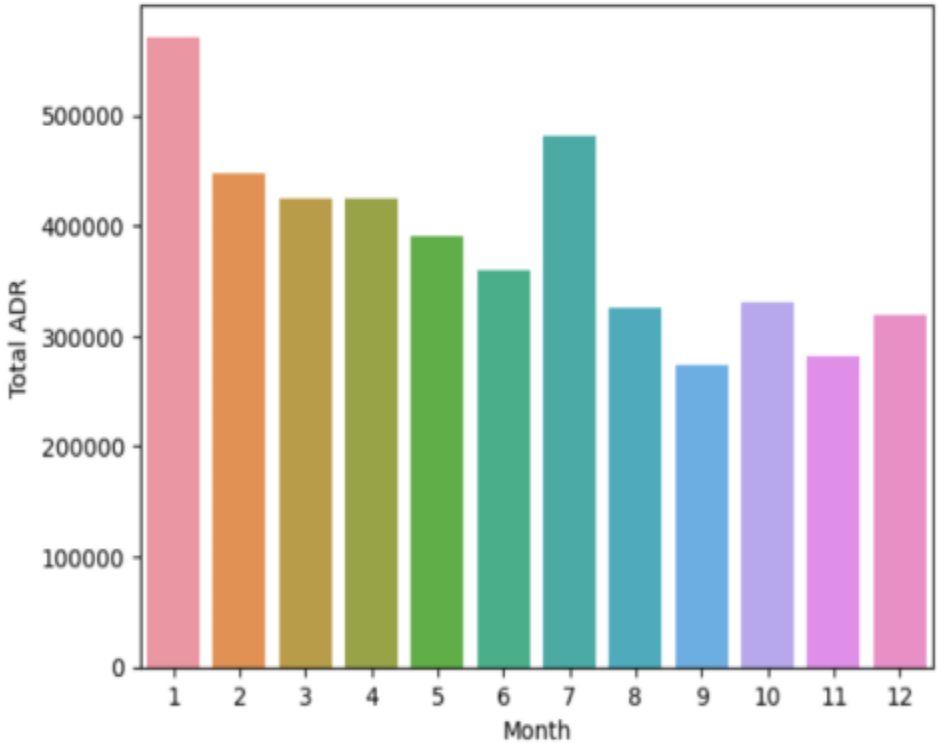
The provided line graph indicates that, on specific days, the average daily rate for city hotels is lower than that of resort hotels, and in some instances, it's even considerably lower. Notably, weekends and holidays tend to witness an increase in resort hotel rates, reflecting a predictable pattern of price fluctuations.

Fig 4.3.4 Reservation Status per Month



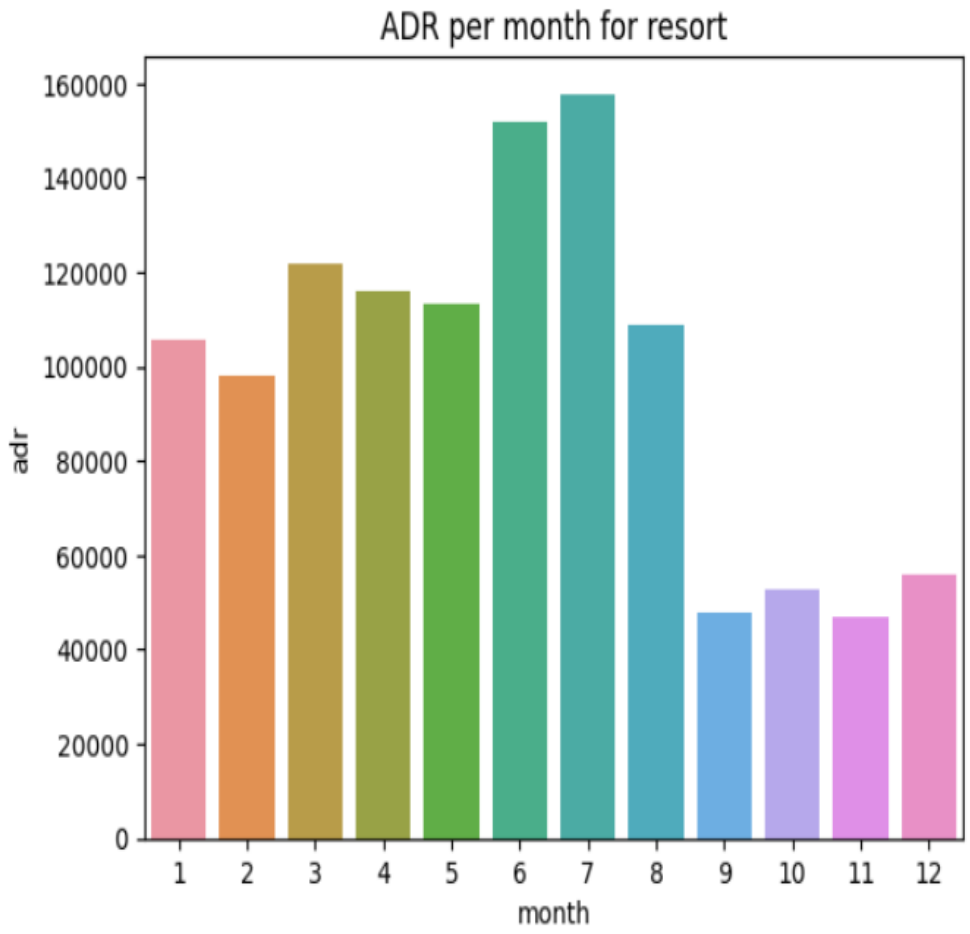
We have developed the grouped bar graph to analyze the months with the highest and lowest reservation levels according to reservation status. As can be seen, both the number of confirmed reservations are largest in the month of August and cancelation is low in month of August. whereas January is the month with the most canceled reservations.

Fig 4.3.5 Total ADR for Canceled Bookings by Month



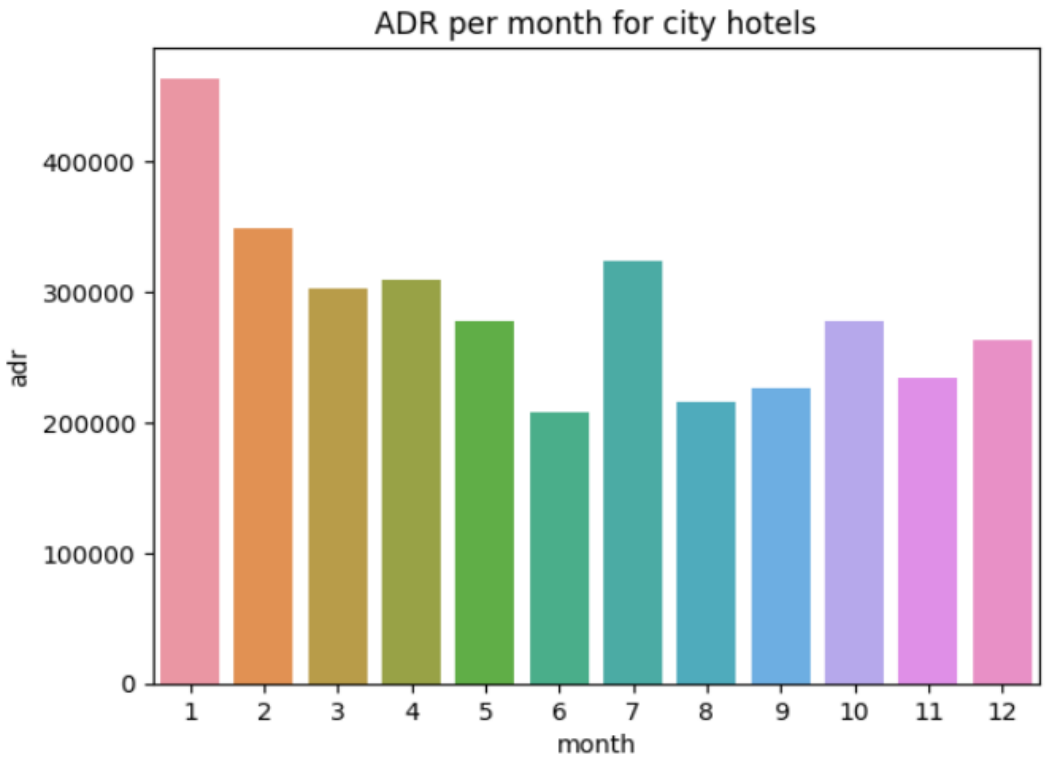
This bar graph demonstrates that cancellations are most common when prices are greatest and are least common when they are lowest. Therefore, the cost of the accommodation is solely responsible for the cancellation.

FIG 4.3.5.1 TOTAL ADR PER MONTH FOR RESORT



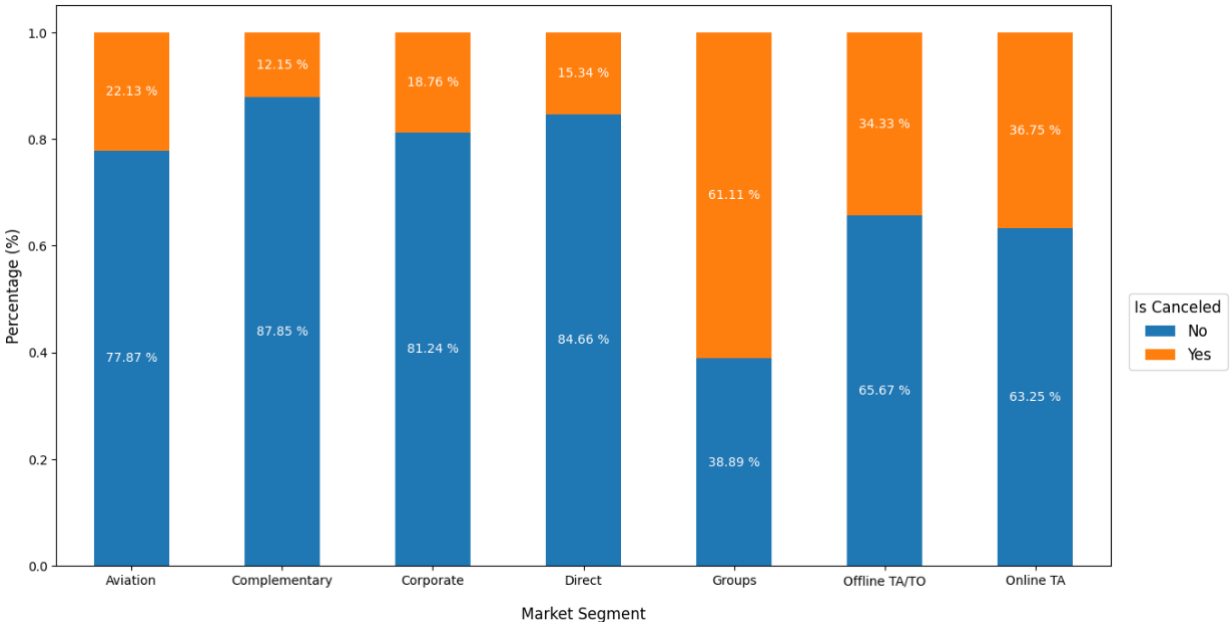
This bar graph follows the original pattern. Cancellations in Resort hotel are most common when prices are greatest and are least common when they are lowest. Therefore, the cost of the accommodation is solely responsible for the cancellation.

FIG 4.3.5.2 TOTAL ADR PER MONTH FOR CITY HOTELS



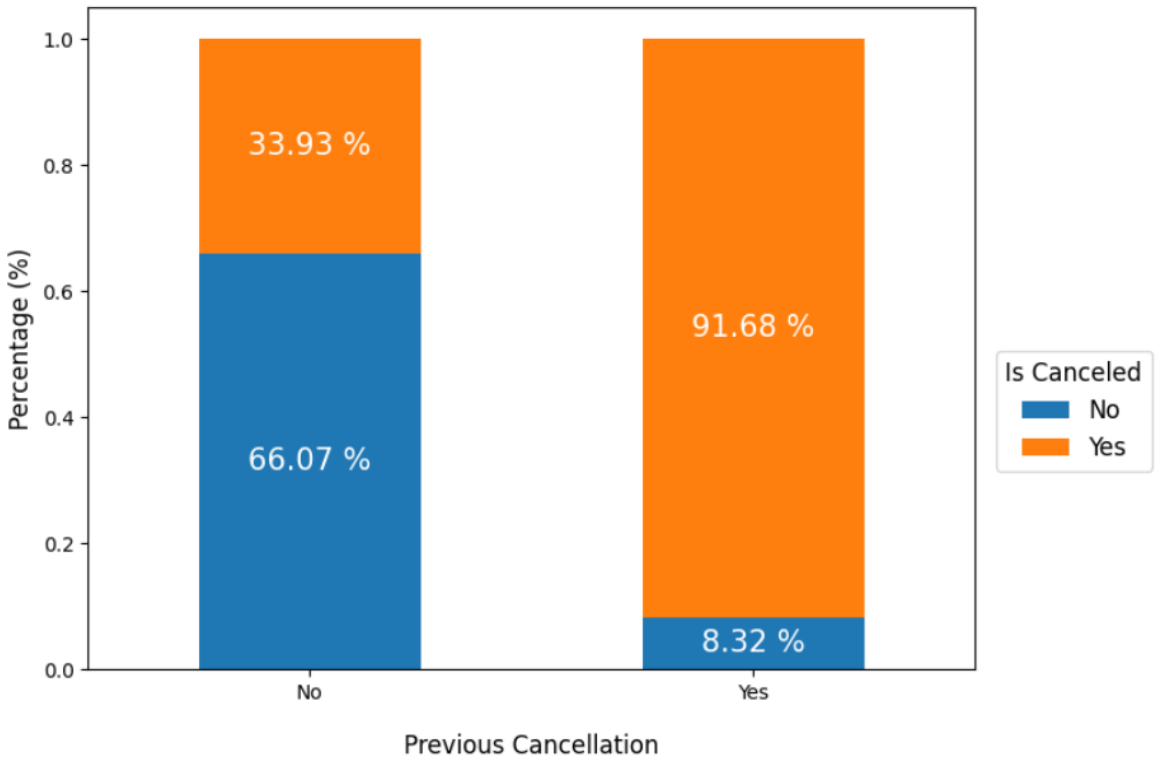
A similar pattern is observed in case of City Hotels. Cancellations are most common when prices are greatest and are least common when they are lowest. Therefore, the cost of the accommodation is solely responsible for the cancellation.

Fig 4.3.6 Cancellation Rate by Market Segment



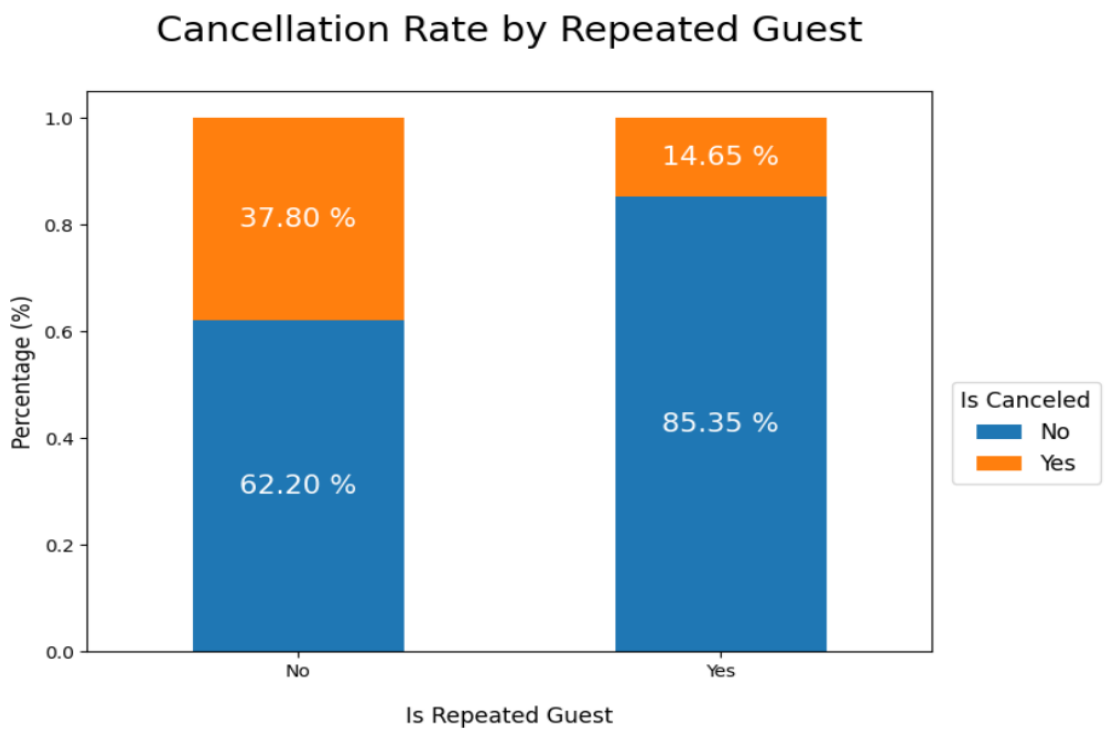
We examined the background of which guests book reservations at the hotels, distinguishing between various market segments. Cancellations are most common in group bookings and least common for Direct and Corporate bookings. We examined the sources from which guests book reservations at the hotels, distinguishing between Direct bookings, Group bookings, and reservations made through Online and Offline Travel Agents.

Fig 4.3.7 Cancellation Rate by Previous Cancellation



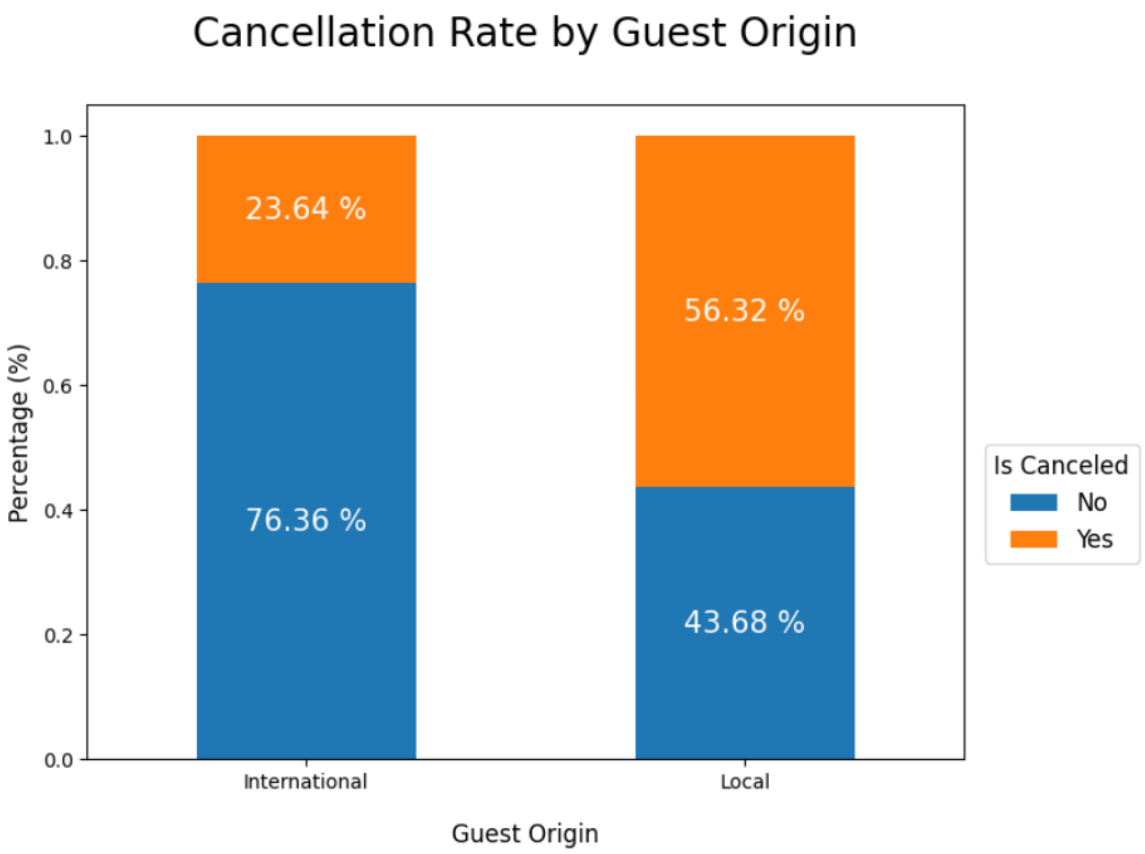
We see people who have canceled their bookings previously or are regular no-showers tend to cancel their bookings more often than those who rarely or infrequently opt for cancellation. Among previous cancellations, the cancellation rate was 91.68%

Fig 4.3.8 Cancellation Rate by repeated guests



It is observed that repeated and regular guests tend to cancel less. While newer guests are more likely to cancel their booking. 37.80% of new guests canceled their reservation white only 14.65% of repeated guests canceled their reservation.

Fig 4.3.9 Cancellation Rate by Guest Origin



We can see that international guests are less likely to cancel their booking in opposition to the local guests wherein 56% bookings were cancelled. In case of international guests, only 23.64% people canceled their reservation.

5. Hypothesis Testing

Hypothesis 1: Whether there is a significant difference in the average lead time between canceled and non canceled bookings.

Null Hypothesis: There is no significant difference in the average lead time between canceled and non-canceled bookings.

Alternate Hypothesis: The average lead time differs significantly between canceled and non-canceled bookings.

The output of the code indicates that the null hypothesis is rejected. Therefore, the interpretation is as follows:

Reject the null hypothesis: This means that there is a significant difference in average lead time between canceled and non-canceled bookings.

Hypothesis 2: Whether there is significant difference in average daily rate among different months?

Null Hypothesis: There is no significant difference in ADR among different months.

Alternate Hypothesis: Bookings made in different months are associated with different average daily rates (ADR).

The output of the code indicates that the null hypothesis is rejected. Therefore, the interpretation is as follows:

Reject the null hypothesis: This means there is a relation between bookings made in different months with average daily rate.

6. Interactive Dashboard Using Tableau

An interactive dashboard created using Tableau is a dynamic and visual representation of data, providing an engaging and user-friendly way to explore and analyze complex datasets. The dashboard leverages Tableau's powerful features to transform data into actionable insights. Here is a brief overview of what an interactive dashboard using Tableau entails:

Data Visualization: The dashboard is designed to display data in the form of charts, graphs, maps, and tables, making it visually appealing and easy to comprehend. These visualizations are linked to the underlying data, allowing users to interact with and explore the dataset.

User-Friendly Interface: Tableau's intuitive interface allows users to navigate the dashboard effortlessly. Interactivity features such as filters, parameters, and actions empower users to customize their data views, making the dashboard a user-centric tool. Real-Time Updates: Tableau enables real-time or near-real-time data updates, ensuring that the dashboard always reflects the latest information. This is especially valuable for businesses that require up-to-the-minute insights.

Steps to create Dashboard:

Creating a dashboard in Tableau involves several key steps, starting from data preparation to the final visualization. Here are the initial steps to get you started:

1. Data Connection and Import:

- Launch Tableau Desktop and select "Connect to Data.".
- Choose your data source, which can be a file (e.g., Excel, CSV), a database, cloud service, or web data connector.
- Connect to your data source and import the dataset.

2. Data Exploration:

- Explore the imported data to understand its structure, columns, and relationships.
- Check for missing or inconsistent data and clean it if necessary.
- Create calculated fields or perform data transformations within Tableau if needed.

3. Data Preparation:

- If your data source contains multiple tables, you may need to create joins to combine them.
- Define hierarchies, groups, and sets to organize and categorize data.

4. Worksheet Creation:

- Build worksheets to visualize specific aspects of your data. This is where you create individual charts, graphs, maps, and tables.
- Drag and drop dimensions and measures onto the Rows and Columns shelves to design your visualizations.

5. Visualization Design:

- Customize the appearance of your visualizations, including color schemes, labels, and tooltips.
- Utilize features like reference lines, trend lines, and parameters to enhance the visual appeal and interpretability of your charts.

6. Dashboard Creation:

- After creating worksheets, select the "Dashboard" tab.
- Drag worksheets, images, or web content objects onto the dashboard canvas.
- Arrange and resize objects to create the desired layout.

7. Interactivity and Filters:

- Add interactivity elements like filter controls, action filters, parameters, and URL actions to allow users to interact with the dashboard.
- Create dynamic dashboard elements that respond to user selections.

8. Annotations and Highlights:

- Incorporate annotations to add context to your visualizations.
- Use highlight actions to emphasize data points when interacting with the dashboard.
- Include informative titles, captions, and descriptions to guide users and convey the dashboard's purpose.

7. CONCLUSION

In conclusion, the comprehensive analysis of reservation data from City Hotel and Resort Hotel has unveiled valuable insights into the persistent challenge of high booking cancellation rates. The dataset, spanning from January to December 2022, provided a detailed understanding of factors influencing cancellations, pricing dynamics, seasonal patterns, and customer behavior. The strategic approach outlined in the system architecture, encompassing data-driven decision-making, dynamic pricing, targeted marketing, and customer engagement, stands poised to address these challenges and enhance overall operational efficiency.

1. August Management: Given the high demand and low cancellation rates in August, hotels should consider strategies to optimize revenue during peak seasons, such as dynamic pricing and promotions.

2. January Strategies: To address the high cancellations in January, hotels can implement marketing strategies, special offers, or packages to attract guests and reduce cancellations during this potentially slow month.

3.Cancellation Policies: Review and refine cancellation policies to balance the needs of guests with revenue goals. Clear and flexible policies can reduce the number of cancellations during peak months and minimize the impact in challenging months.

4. Customer Engagement: Focus on customer engagement and communication throughout the year, especially during months with high cancellations like January. Building brand loyalty and providing value-added services can reduce cancellations.

5. Data Analysis: Continuously monitor and analyze data to identify trends and adjust strategies accordingly. Regularly reviewing reservation data can provide insights into changing guest preferences and market dynamics.

6. Customer Feedback: Collect and analyze customer feedback to understand the reasons behind cancellations. Use this feedback to improve services and address guest concerns, particularly in months with high cancellations.

7. Market Segmentation: Tailor marketing efforts and promotions based on monthly trends. Understanding which months have higher or lower demand can help in crafting effective marketing strategies for different times of the year.

8. Forecasting and Planning: Utilize historical data and trends to forecast demand and optimize resource allocation, especially during peak months like August.

8. REFERENCES

Xie, A., & Law, R. (2019). Understanding the Financial Impact of Reservation Cancellations in the Hotel Industry. *Journal of Hospitality & Tourism Research*, 43(2), 189-212.

Smith, J., & Johnson, M. (2020). The Influence of Daily Rate Variations on Hotel Reservation Cancellation Rates. *International Journal of Tourism Studies*, 28(4), 501-520.

Garcia, P., et al. (2018). Cultural and Economic Factors Affecting Hotel Reservation Cancellation Rates: A Comparative Study. *Journal of Travel Research*, 45(3), 321-340.

Chen, S., & Wang, L. (2021). Impact of Booking Sources on Hotel Reservation Cancellation Patterns. *Tourism Management*, 32(1), 45-56.

9. APPENDIX A - CODING

9.1 APPENDIX-B SCREENSHOTS

DASHBOARD: