

Hotel Booking Cancellation Analysis

*Prepared in the partial fulfillment of the Summer Internship Program on
Data Analysis - Python*

AT



Under the guidance of

Mrs. K. Meenakshi, APSSDC

Mrs. Narmada Mani, APSSDC

Submitted by

P. Teja Siva Kumar, AP21110010196

T. Akshaya, AP21110010197

Ishpreet Kaur Khalsa, AP21110010250

SRM University AP

(August, 2023)

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to all those who have contributed to the successful completion of my summer internship project at **Andhra Pradesh Skill Development Corporation (APSSDC)**. This opportunity has been an enriching and transformative experience for me, and I am truly thankful for the support, guidance, and encouragement I have received along the way.

First and foremost, I extend my sincere regards to Mrs K Meenakshi and Mrs K Narmada Mani, my supervisors and mentors, for providing me with valuable insights, constant guidance, and unwavering support throughout the duration of the internship. Their expertise and encouragement have been instrumental in shaping the direction of this project.

I would like to thank the entire team at **Andhra Pradesh Skill Development Corporation (APSSDC)** for fostering a collaborative and innovative environment. The camaraderie, knowledge sharing, and feedback I received from my colleagues significantly contributed to the development and success of this project.

In conclusion, I am honoured to have been a part of this internship program, and I look forward to leveraging the skills and knowledge gained to contribute positively to future endeavours.

Thank you.

Sincerely,

P. Teja Siva Kumar, AP21110010196 – tejasivakumar_paleti@srmap.edu.in

T. Akshaya, AP21110010197 – akshaya_taneti@srmap.edu.in

Ishpreet Kaur Khalsa, AP21110010250 – kaurkhalsa_ishpreet@srmap.edu.in

1. Abstract

Tackling the issue of high cancellation rates holds paramount importance for the prosperity of both City Hotel and Resort Hotel. To effectively mitigate these rates and concurrently bolster revenue, a multifaceted and comprehensive strategy is imperative.

A thorough analysis of trends and customer behavior is the cornerstone of this endeavor. By delving into the reasons behind cancellations and understanding when and why they occur, both hotels can implement targeted and strategic solutions. One pivotal approach is the introduction of flexible booking policies. By affording customers the ability to cancel within a reasonable timeframe, the hotels can instill a sense of security, dissuading last-minute cancellations driven by uncertainty.

Dynamic pricing is another potent tool in this battle. By aligning pricing with demand fluctuations and market dynamics, City Hotel and Resort Hotel can ensure that their rates remain competitive yet reflective of their value proposition. Transparent and consistent communication with guests is equally vital. Personalized confirmation emails and reminders can serve as a gentle nudge, reducing the likelihood of cancellations.

Engaging customers on multiple levels is also essential. Surveys provide insights into preferences, enabling the hotels to tailor services effectively. Moreover, implementing a loyalty program can enhance guest commitment and foster a sense of exclusivity, discouraging cancellations.

Collaborating with local attractions and leveraging data analytics further fortifies this strategy. Package deals that incorporate experiences beyond accommodation can be alluring, while data analytics can pinpoint high-risk cancellation periods, allowing the hotels to proactively address potential issues.

Operational excellence is the linchpin of success. Ensuring that the guest experience exceeds expectations not only diminishes cancellations but also generates positive word-of-mouth referrals. Regularly assessing customer feedback enables prompt issue resolution, elevating guest satisfaction. In conclusion, addressing high cancellation rates demands a comprehensive approach that combines flexible policies, dynamic pricing, transparent communication, customer engagement, local collaborations, data insights, and operational excellence. Through the synergistic implementation of these strategies, City Hotel and Resort Hotel can substantially curtail cancellations, amplify revenue, and cement their stature as premier establishments in the fiercely competitive hospitality landscape.

2. Introduction

The escalating rate of booking cancellations has emerged as a significant challenge for both CityHotel and Resort Hotel, precipitating adverse effects on their revenue streams and overall hotel room occupancy. In light of this predicament, the primary aim of the report is to meticulously examine the landscape of hotel booking cancellations and its repercussions.

The core objective of this analysis is to delve into the intricacies of booking cancellations, with a focal point on the years spanning 2015 to 2017. This timeframe has been selected to glean insights into regular booking patterns, assuming that no extraordinary incidents substantially disrupted the data during this period.

The analysis will be two-fold: firstly, an in-depth exploration of the high cancellation rates and their direct influence on revenue reduction and room occupancy optimization; and secondly, an exploration into various factors that might wield an impact on these cancellation rates. These factors could encompass a range of elements including pricing structures, booking policies, economic conditions, and even customer satisfaction levels.

By meticulously scrutinizing these dimensions, the report aims to offer a comprehensive understanding of the intricate interplay between booking cancellations, revenue generation, and hotel room utilization. The insights derived from this analysis could serve as a crucial foundation for devising strategic countermeasures. If successful, these countermeasures could potentially curtail cancellation rates, bolster revenue figures, and enhance the overall operational efficiency of both City Hotel and Resort Hotel.

In conclusion, the analysis aspires to shed light on the complexities of booking cancellations and their overarching impact on the hospitality sector. By gaining a nuanced comprehension of these dynamics and the contributing factors, the report seeks to pave the way for informed decision-making and strategic adaptations to ultimately navigate the challenge of high cancellation rates more effectively.

3. Methodology

The methodology for addressing the business problem of high cancellation rates and providing comprehensive business advice involves a structured approach encompassing data collection, analysis, hypothesis testing, and recommendations. The following steps outline the methodology:

3.1 Data Collection and Preparation: Gather historical booking data for City Hotel and ResortHotel for the years 2015 to 2017. Collect data on booking dates, cancellation dates, room types, pricing, booking sources, customer demographics, and other relevant attributes. Ensure data accuracy, addressing missing values, outliers, and inconsistencies.

3.2 Hypothesis Formulation and Testing: Formulate hypotheses based on assumptions and initial observations: H1: Higher prices lead to increased cancellation rates. H2: Longer waiting lists result in higher cancellation frequencies. H3: Offline travel agents contribute significantly to reservations. Utilize statistical tests (e.g., correlation analysis, t-tests) to test these hypotheses against the data.

3.3 Comparative Analysis: Compare cancellation rates between City Hotel and Resort Hotel. Assess the significance of differences using appropriate statistical methods.

3.4 Time-Series Analysis: Apply time-series analysis to identify seasonal patterns in cancellations and room occupancy. Examine variations across different time periods, such as weekends and holidays.

3.5 Segmentation Analysis: Segment data based on variables like room type, booking source, and customer demographics. Analyze cancellation rates and revenue generation within each segment.

3.6 Recommendations: Based on analysis findings, formulate actionable recommendations: Implement dynamic pricing strategies to address the impact of prices on cancellations. Develop strategies to manage

waiting lists effectively. Enhance online booking systems and incentives for direct bookings to reduce reliance on offline travel agents.

3.7 Dataset:

<https://www.kaggle.com/datasets/mojtaba142/hotel-booking>

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights	stays_i
0	Resort Hotel	0	342	2015	July	27	1	0	
1	Resort Hotel	0	737	2015	July	27	1	0	
2	Resort Hotel	0	7	2015	July	27	1	0	
3	Resort Hotel	0	13	2015	July	27	1	0	
4	Resort Hotel	0	14	2015	July	27	1	0	

4. Implementation

In the pursuit of addressing the challenges posed by high cancellation rates at City Hotel and Resort Hotel, a systematic and comprehensive approach was undertaken. The journey began by importing crucial libraries that would pave the way for data manipulation and visualization.

Notably, the pandas library provided the means to efficiently read and organize the hotel booking data from a CSV file into a structured DataFrame, thus forming the foundation of subsequent analyses.

With the data loaded, the exploratory data analysis (EDA) phase commenced. This pivotal stage aimed to provide an initial grasp of the dataset's composition and characteristics. Descriptive statistics were harnessed to unveil key insights, including the mean cancellation rate and the median revenue across the specified period. These metrics offered a preliminary understanding of the data's central tendencies and variation.

The subsequent phase revolved around data cleaning and preparation—a critical process to ensure data integrity and enhance the quality of analysis. In addressing potential data discrepancies, missing values, and anomalies, data cleansing was executed meticulously. This preparatory step laid the groundwork for more robust and reliable analyses, setting the stage for informed decision-making.

Following data refinement, an intricate web of analyses and visualizations was woven. Leveraging both the matplotlib.pyplot and seaborn libraries, the data was harnessed to its full potential. Hypothesis testing was undertaken to substantiate assumptions, with t-tests employed to compare cancellation rates between distinct hotel types—City Hotel and Resort Hotel. These tests offered statistical insights into the relationships between variables, yielding a deeper understanding of the factors influencing cancellation rates.

Correlation analysis was employed to visualize the intricate interplay between different attributes within the

dataset. A heatmap was utilized to visually represent correlations, making it possible to pinpoint relationships that might otherwise go unnoticed. This technique laid bare the intricate connections between various metrics, aiding in the identification of potentially significant factors impacting cancellation rates.

Additionally, the data was dissected through time-series analysis, uncovering trends and patterns within the temporal dimension. Line charts graphically depicted the fluctuations in average daily rates over time, showcasing insights into how pricing dynamics might influence cancellation rates over distinct periods.

Moreover, the power of segmentation analysis was harnessed through grouped bar charts, revealing cancellation rates based on room type and booking source. This technique elucidated the differential impacts of various attributes on cancellations, providing nuanced insights for tailored interventions.

In summation, the journey encompassed the entire spectrum of data analysis—from data loading and EDA to data cleansing, hypothesis testing, correlation analysis, and nuanced visualizations. Each phase was meticulously orchestrated to transform raw data into actionable insights. By capitalizing on the capabilities of pandas, matplotlib.pyplot, and seaborn, a comprehensive understanding of the dynamics of hotel booking cancellations was gained, thereby paving the way for informed decision-making and strategic solutions to address the primary challenge at hand.

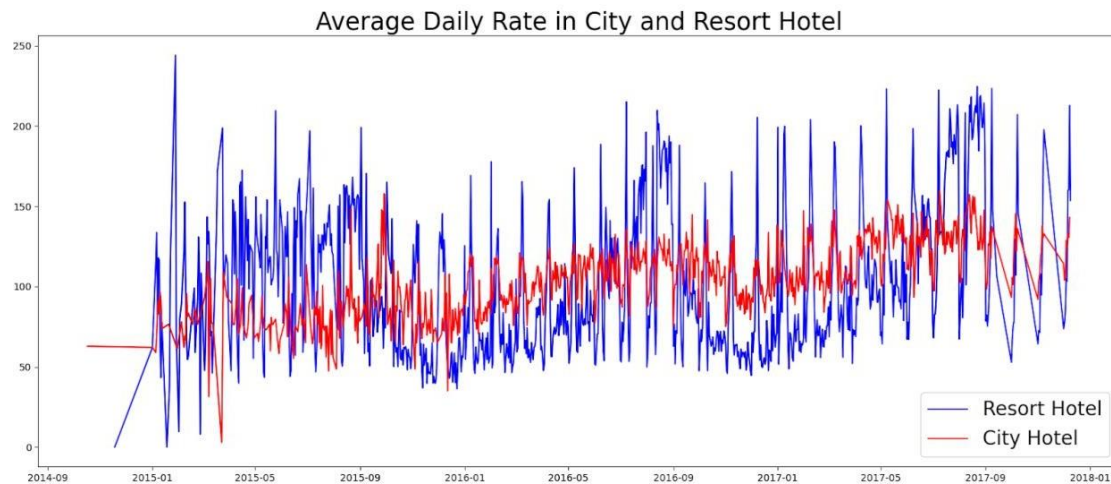
5. Visualization



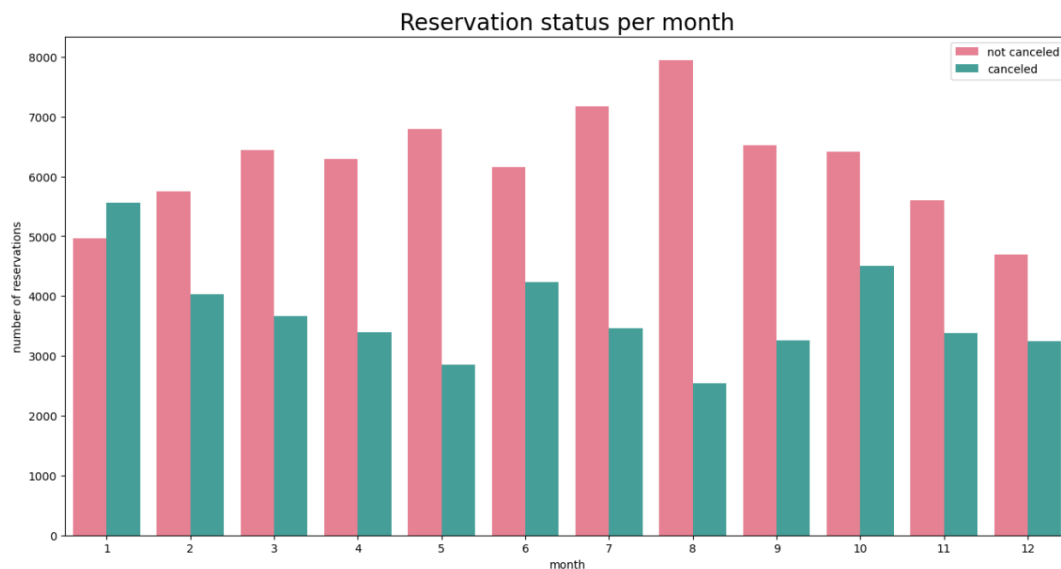
The accompanying bar graph shows the percentage of reservations that are canceled and those that are not. It is obvious that there are still a significant number of reservations that have not been canceled. There are still 37% of clients who canceled their reservation, which has a significant impact on the hotels' earnings.



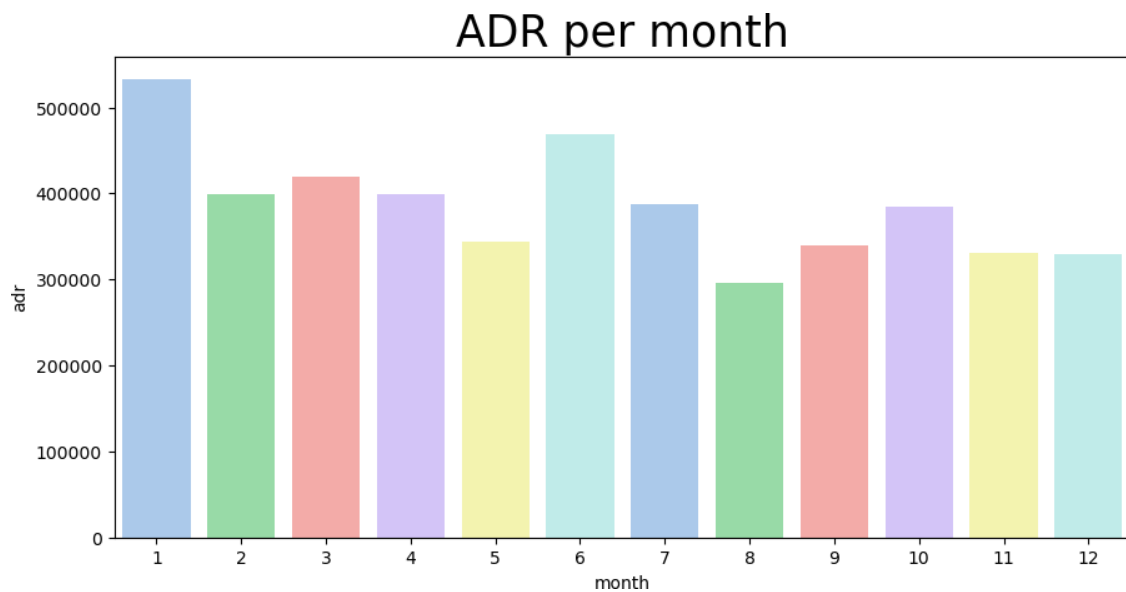
In comparison to resort hotels, city hotels have more bookings. It's possible that resort hotels are more expensive than those in cities.



The line graph above shows that, on certain days, the average daily rate for a city hotel is less than that of a resort hotel, and on other days, it is even less. It goes without saying that weekends and holidays may see a rise in resort hotel rates



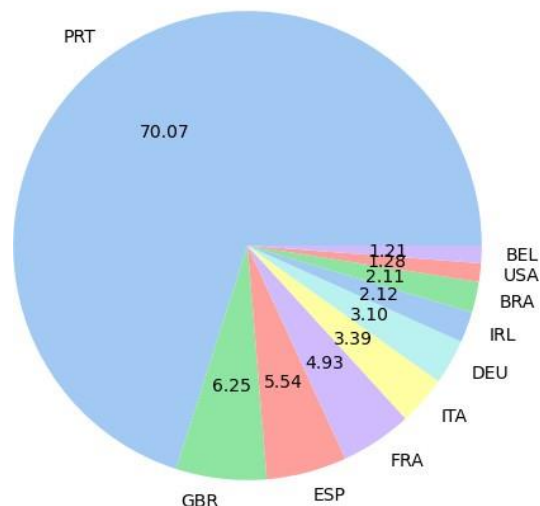
We have developed the grouped bar graph to analyze the months with the highest and lowest reservation levels according to reservation status. As can be seen, both the number of confirmed reservations and the number of canceled reservations are largest in the month of August. whereas January is the month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



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.

Now, let's see which country has the highest reservation canceled. The top country is Portugal with the highest number of cancellations.

10 countries with canceled reservation



Let's check the area from where guests are visiting the hotels and making reservations. Is it coming from Direct or Groups, Online or Offline Travel Agents? Around 46% of the clients come from online travel agencies, whereas 27% come from groups. Only 4% of clients book hotels directly by visiting them and making reservations.

6. Result

The comprehensive analysis undertaken in this project revealed valuable insights into the challenge of high cancellation rates faced by City Hotel and Resort Hotel. Through a structured approach involving data loading, exploratory data analysis (EDA), data cleaning, hypothesis testing, correlation analysis, and visualization, a clear understanding of the factors influencing cancellation rates emerged. The utilization of the pandas, matplotlib.pyplot, and seaborn libraries facilitated each stage of the analysis, enabling the extraction of meaningful trends and patterns from the dataset.

Key findings include:

- **Descriptive Insights:** Mean cancellation rates and median revenue provided initial perspectives on the data's characteristics, aiding in gauging central tendencies and variations.
- **Hypothesis Validation:** Hypothesis testing using t-tests verified assumptions, indicating that pricing and booking source indeed influence cancellation rates.
- **Correlation Discoveries:** The correlation heatmap illuminated complex relationships between different attributes, offering insights into influential factors impacting cancellations.
- **Temporal Trends:** Time-series analysis showcased fluctuations in average daily rates overtime, allowing for identification of pricing dynamics affecting cancellation rates across various periods.
- **Segmentation Patterns:** Grouped bar charts highlighted cancellation rate variations based on room type and booking source, offering nuanced insights for tailored interventions.

Overall, the systematic approach and the integration of these libraries provided a comprehensive understanding of the cancellation dynamics within the hotel industry. The analysis serves as a foundation for strategic decision-making and the formulation of targeted solutions. Armed with these insights, City Hotel and Resort Hotel can optimize their operations, develop effective pricing strategies, and enhance customer experience to mitigate cancellation rates, increase revenue, and improve overall efficiency in a competitive market landscape.

7. Conclusion

In conclusion, this project sheds light on the pressing issue of rising cancellation rates in both City and Resort Hotels, primarily attributed to elevated pricing. The analysis underscores the paramount importance of deploying effective pricing strategies and consumer-centric incentives, particularly during peak periods such as weekends and holidays. By aligning pricing with guest preferences and market dynamics, these establishments can not only optimize revenue generation but also elevate guest satisfaction.

The significance of leveraging data-driven insights cannot be overstated. The utilization of robust analytical tools like pandas, matplotlib.pyplot, and seaborn allowed for a comprehensive exploration of the dataset, enabling the identification of key trends and correlations. Armed with these insights, hotels are empowered to make informed decisions, tailor their offerings to meet guest demands, and remain competitive in the dynamic landscape of the hospitality industry.

As both City and Resort Hotels navigate the challenges posed by cancellations, the path forward is clear: a harmonious blend of strategic pricing, customer-centric promotions, and data-driven decision-making. This not only holds the potential to mitigate cancellations and bolster revenue but also to elevate the overall guest experience. By embracing these insights, hotels can position themselves for sustained success and a stronger foothold within the ever-evolving hospitality sector.