Title Page

Long Island University

Optimizing Hotel Booking Strategies: *An Analytical Approach to Cancellations, Pricing, and Offers*

Submitted by: Roger Christy

Course: Data-Driven Decision Making
Instructor:
Gurpreet Singh

Submitted on:

Contents

Title Page 1
Background 2
Project Impact 3
Problem Scenario/Business Issue 3
Objective/Goals 4
Framework for Analysis 4
Data Exploration/ Data Visualization and Manipulation 5
Average Daily Rate (ADR) Trends in City and Resort Hotels 8
Strategic Implications 9
Strategic Implications 10
Impact of ADR on Reservation Cancellations by Month 10
Strategic Recommendations 11
Reservation Cancellation Rates by Country 11
Strategic Recommendations 12
Reservation Channels Breakdown 12
Strategic Recommendations 13
Predictive Modelling on Research Questions 13
Research Question 1: Why Do People Cancel Their Hotel Bookings? 14
Research Question 2: How Can We Help Hotels Have Fewer Cancellations? 15
Research Question 3: How Can We Help Hotels Decide on Their Prices and Special Offers? 17
Methodology/Model Building/Model Selection 19
Conclusions and Strategic Recommendations 21

Background

The hospitality industry plays a crucial role in the global economy but faces ongoing challenges, especially with high booking cancellation rates. On average, around 40% of hotel reservations worldwide are cancelled, with some online travel agencies reporting even higher rates of up to 50% (Hospitality Tech, Hotel Management). These cancellations significantly impact hotel revenue, disrupt operational planning, and create inefficiencies.

Traditional methods, such as overbooking or imposing strict cancellation policies, have been used to manage cancellations. However, these strategies often fall short, leading to dissatisfied customers and potential damage to the hotel's reputation.

With advancements in data analytics and machine learning, hotels now have the opportunity to better understand booking behaviors and implement smarter strategies. These tools allow hotels to make data-driven decisions, optimize operations, and address the challenges of high cancellation rates effectively.

Project Impact

This project focuses on using data-driven approaches to analyze and predict booking cancellations, optimize pricing strategies, and create targeted offers. By applying these solutions, hotels can improve revenue management, minimize operational inefficiencies, and enhance customer satisfaction.

The goal is to provide practical recommendations that address the challenges faced by the hospitality industry while aligning with its dynamic nature. Through the use of advanced analytical techniques, this project aims to transform traditional hotel management practices, offering a strategic framework for hotels to navigate challenges and achieve long-term profitability and resilience.

SOURCES

https://www.siteminder.com/r/hotel-industry-statistics/

https://www.oliandalex.com/the-impact-of-hotel-cancellation-fees-on-travelers/

https://www.hotelminder.com/everything-you-need-to-know-about-hotel-cancellations

https://hoteltips101.com/how-to-handle-cancellations-at-hotels/

https://shrgroup.com/2023/06/21/we-need-to-talk-about-cancellations/

https://www.mirai.com/blog/cancellations-shooting-up-implications-costs-and-how-to-reduce-them/

Problem Scenario/Business Issue

City Hotel and Resort Hotel are struggling with an increase in reservation cancellations, which has negatively affected their revenue and daily operations. The key challenges include:

1. **Revenue Loss**:
   - Cancelled bookings leave rooms vacant, directly reducing income.
   - Overbooking strategies fail to fully cover last-minute cancellations, resulting in unsold rooms.
2. **Higher Operational Costs**:
   - Even cancelled bookings require maintenance and staffing, increasing fixed costs.
   - Unpredictable booking patterns make it difficult to efficiently manage staff and resources.
3. **Pricing Challenges**:
   - Current pricing strategies don’t adapt to changes in demand, leading to under-priced rooms during busy periods and overpriced rooms during low-demand times.
   - Generic discounts fail to appeal to specific customer needs, reducing their effectiveness.

Gaps in Current Solutions

- Stricter cancellation policies and broad discounts often upset customers or fail to address the real causes of cancellations.
- These traditional approaches lack precision and do not optimize pricing or customer targeting.

Project Focus

This project aims to use advanced data analysis and machine learning to:

1. **Predict Cancellations**: Accurately forecast cancellations to take proactive steps.
2. **Improve Pricing**: Create dynamic pricing strategies to boost revenue and customer satisfaction.
3. **Tailor Offers**: Recommend customized packages and promotions to attract specific customer segments.

Objective/Goals

This project aims to provide practical, data-driven solutions to help City Hotel and Resort Hotel tackle challenges related to reservation cancellations and pricing issues. The key objectives are:

1. **Understand Why Guests Cancel**:
    - Analyse past booking data to uncover patterns and behaviors that lead to cancellations.
2. **Improve Pricing Strategies**:
    - Evaluate how current pricing models affect bookings and revenue.
3. **Predict Cancellations**:
    - Use machine learning to forecast cancellations, allowing the hotels to act in advance.
4. **Optimize Pricing**:
    - Create smart pricing strategies to maximize revenue during busy and slow seasons.
5. **Design Targeted Offers**:
    - Develop special deals and packages to reduce cancellations and attract valuable customers.

Framework for Analysis

This section explains the assumptions and research questions that guide our analysis. These form the foundation for understanding booking cancellations and identifying ways to improve hotel operations.

Assumptions

1. **Data is Reliable**:

- The data from 2015 to 2017 reflects accurate booking trends, with no major external events disrupting its reliability.

2. **Strategies are Relevant**:
   - The data and insights are applicable to the hotel's current operations.
3. **Ideas are Feasible**:
   - Implementing the proposed strategies is practical and won't cause unexpected challenges for the hotels.
4. **New Approaches**:
   - The strategies and recommendations are innovative and haven't been tried before by the hotels.
5. **Cancellations Impact Revenue**:
   - Cancelled bookings are a significant cause of revenue loss for the hotels.
6. **Rooms Stay Empty**:
   - When bookings are cancelled, the rooms typically remain unoccupied during the reserved period.
7. **Booking Trends are Year-Specific**:
   - Most guests book and cancel within the same calendar year, reflecting predictable seasonal patterns.

Research Questions

1. **Why Do Guests Cancel?**
   - What factors drive guests to cancel their bookings?
2. **How Can Hotels Reduce Cancellations?**
   - What strategies can help hotels minimize cancellations?
3. **How Can Hotels Improve Pricing and Offers?**
   - How can hotels optimize their prices and create better promotions to attract guests?

Main Focus

This project focuses on understanding why cancellations happen and providing solutions to address them. The goal is to help hotels:

- Minimize cancellations.
- Keep rooms occupied.
- Improve revenue and operational efficiency.

Data Exploration/ Data Visualization and Manipulation

The data shows that **37% of hotel reservations are cancelled**, which has a significant impact on the hotel's financial and operational performance. While this cancellation rate is slightly better than the industry average of **40%**, it still represents a major challenge, even for well-managed hotels.

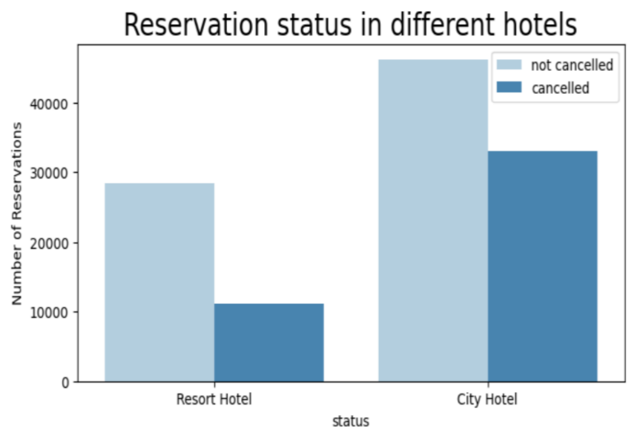
Financial Impact

- With an **average daily rate (ADR) of $100**, a **37% cancellation rate** could result in substantial revenue losses, amounting to millions of dollars annually depending on the number of bookings.
- Addressing cancellations is critical to improving the hotel's revenue management and financial stability.

Operational Challenges

1. **Room Inventory Issues**:
   - Frequent cancellations make it hard to manage room availability effectively.
2. **Staff Scheduling Problems**:
   - Unpredictable booking patterns lead to overstaffing or understaffing, increasing costs or reducing service quality.
3. **Difficulty in Planning**:
   - Cancellations make it challenging to forecast demand and set optimal pricing strategies.

Comparative Analysis: Resort vs. City Hotel Bookings and Cancellations



The data highlights key differences between City Hotels and Resort Hotels, including booking trends, cancellation rates, and pricing strategies:

1. Booking Trends

- **City Hotels** attract more bookings, likely due to their appeal for short-term stays and business travellers.
- **Resort Hotels** have fewer bookings but cater to guests seeking leisure-focused, longer stays.

2. Cancellation Rates

- **City Hotels** experience higher cancellation rates (**45%**) compared to Resort Hotels (**25%**).
- **Resort Hotels** benefit from lower cancellation rates, as guests are often more committed to their vacation plans.

3. Pricing Differences

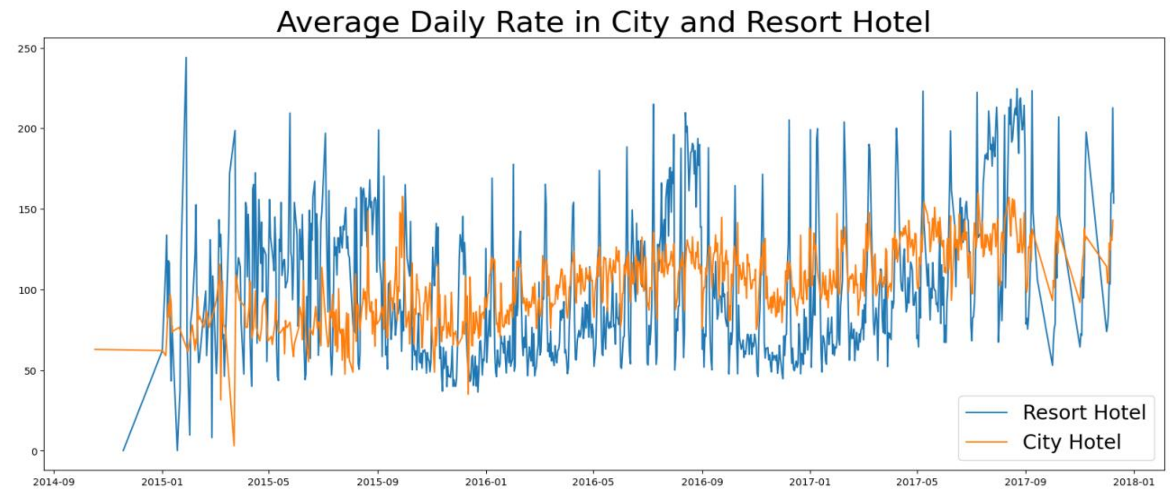
- **Resort Hotels** typically charge a higher **Average Daily Rate (ADR)**, attracting guests who are less likely to cancel.
- **City Hotels**, with generally lower ADRs, attract more price-sensitive guests who are more prone to cancellations.

4. Operational Needs

- **City Hotels** need to improve their facilities and guest experiences to reduce cancellations.
- Strategies like personalized offers and better room maintenance can help retain guests.

5. Strategic Focus

- For both hotel types, understanding key factors such as lead time, room type, and guest preferences is crucial for reducing cancellations.
- **City Hotels**, in particular, should analyze guest expectations more deeply to address their higher cancellation rates effectively.



Average Daily Rate (ADR) Trends in City and Resort Hotels

The analysis of ADR trends from 2014 to 2018 highlights key patterns and opportunities for revenue optimization:

1. Rate Fluctuations

- ADR changes are influenced by seasonality, special events, and travel demand.
- **Resort Hotels** show larger price swings compared to City Hotels, reflecting their reliance on peak vacation periods.

2. City vs. Resort Pricing

- **Resort Hotels** generally have higher ADRs, with holiday peaks reaching up to **$250**.
- **City Hotels** sometimes match or exceed Resort ADRs during high-demand periods, such as major urban events.

3. Pricing Stability

- Between 2016 and 2017, **City Hotels** demonstrated stable pricing, which could attract repeat customers by providing predictable rates.

4. Peak Period Premiums

- Both hotel types increase ADRs by **20-30%** on weekends and holidays to capitalize on higher demand.

5. Competitive Pricing

- Occasional price overlaps between City and Resort Hotels reflect targeted efforts to attract customers during special events or off-peak periods.

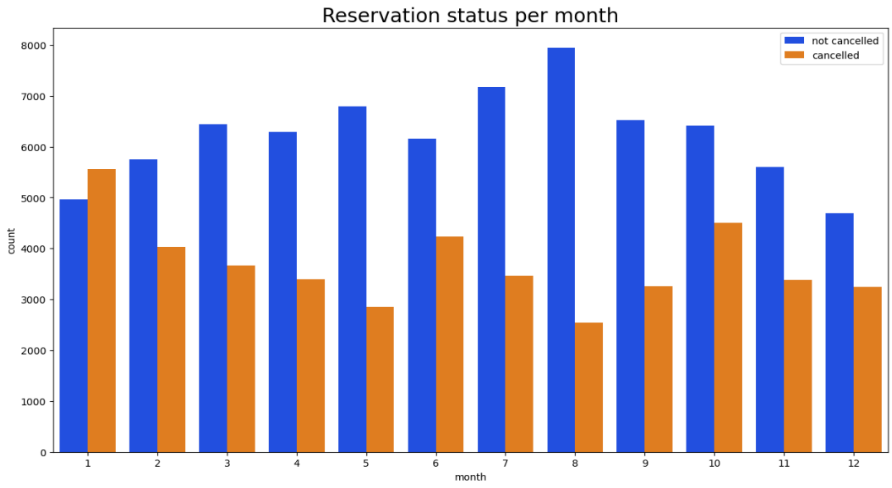
6. Weekend Trends

- Weekend ADRs are **15-25% higher** than weekday rates for both hotel types, taking advantage of increased travel activity.

Strategic Implications

1. **For Resort Hotels**:
   - Refine dynamic pricing models to maximize revenue during peak periods and holidays.
2. **For City Hotels**:
   - Maintain consistent pricing to build loyalty while introducing premium rates during holidays and events.
3. **For Both Hotel Types**:
   - Use these insights to design promotional offers and optimize pricing strategies during high-demand periods.

Monthly Reservation Trends and Cancellation Rates



The data reveals important patterns in monthly reservation behaviors, providing insights to help improve hotel performance:

1. Peak Reservation Period:

- **August** has the highest number of reservations (around 8,000), including both confirmed and cancelled bookings.
- This is likely due to increased travel during the summer season.

2. High Cancellation Month:

- **January** has the highest cancellation rate, despite fewer overall bookings.
- This could be due to post-holiday financial strain, bad weather, or shifting travel plans.

3. Seasonal Trends:

- Spring and fall see lower cancellation rates, reflecting more stable travel demand during these periods.

Strategic Implications

1. **Targeted Marketing**:
   - Launch promotional campaigns in **January** to address high cancellations, such as flexible offers or special discounts.
2. **Demand Management**:
   - Prepare for the **August surge** by optimizing staffing and resources to handle the high demand efficiently.
3. **Cancellation Mitigation**:
   - Investigate why cancellations are high in January and offer solutions like refundable bookings or off-season packages to attract guests.

Impact of ADR on Reservation Cancellations by Month

The analysis shows a clear relationship between pricing (ADR) and cancellation rates, providing valuable insights for improving booking stability:

1. How ADR Affects Cancellations:

- Higher ADRs are linked to more cancellations, as price sensitivity influences guest decisions.

2. High-ADR Cancellations:

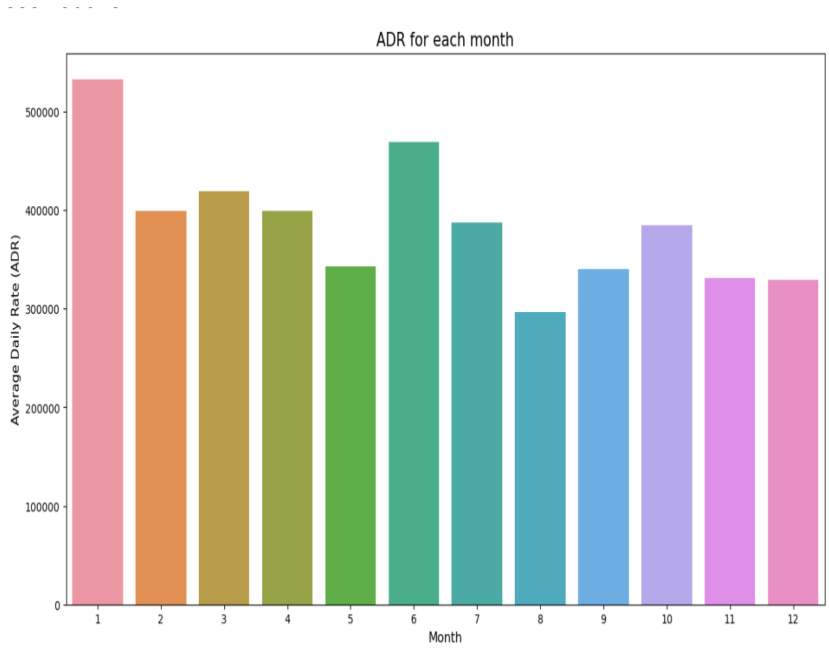
- **January** has the highest ADR (around $500) and the highest cancellation rate, likely due to cost concerns after the holidays.

3. Low-Cancellation Periods:

- **August**, despite moderate-to-high ADRs, has the lowest cancellation rates, showing that guests are willing to pay more during peak summer travel.

4. Cancellation Extremes:

- **January**: Requires strategies like discounted packages or flexible payment plans to reduce cancellations.
- **August**: Demonstrates the importance of aligning pricing with guest value perception, which can be replicated during other high-demand periods.



Strategic Recommendations

1. **Flexible Pricing**:
   - Offer tiered pricing or early-bird discounts in high-ADR months like **January** to retain bookings.
2. **Value-Based Offers**:
   - Use August as a model by pairing premium pricing with added benefits, such as inclusive packages or enhanced guest experiences.
3. **Guest Behavior Analysis**:
   - Study seasonal patterns and preferences to refine pricing strategies for both high-demand and low-demand periods.

Reservation Cancellation Rates by Country

The data highlights the top countries contributing to hotel reservation cancellations, providing key insights into regional patterns:

1. Global Cancellation Overview:

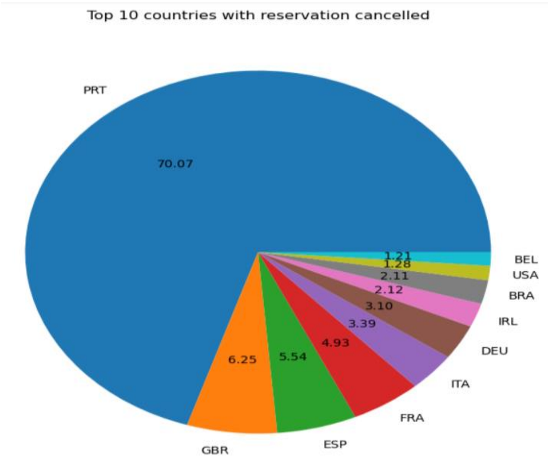
- **Portugal** accounts for **70.07%** of cancellations, making it the most critical market for addressing this issue.
- **The United Kingdom (6.25%)** and **Spain (5.54%)** also contribute notably to cancellations.

2. Why Portugal Dominates:

- Portugal's high cancellation rate could be due to factors like:
  - Economic uncertainty.
  - Flexible booking policies that encourage cancellations.
  - Specific behaviors or preferences of local guests.

3. Other Key Insights:

- While the **UK** and **Spain** have lower shares, their cancellation rates are significant enough to warrant attention.
- Countries like the **USA** and **Belgium** contribute smaller shares but reflect diverse guest behaviors that may present new opportunities.



Strategic Recommendations

1. **For Portugal**:
   - Introduce stricter cancellation policies.
   - Launch targeted loyalty programs and promotional offers to retain bookings.
2. **For the UK and Spain**:
   - Offer flexible pricing and package deals to address price sensitivity during high-cancellation periods.
3. **For Global Markets**:
   - Analyze trends in smaller contributing countries to identify emerging challenges or opportunities.

Reservation Channels Breakdown

The data highlights the key sources of hotel reservations and their contributions, offering insights into revenue opportunities and challenges:

1. Online Travel Agencies (OTAs):

- **46% of bookings (56,402)** come from OTAs, making them the largest source.
- However, OTAs charge high commission fees, which can reduce hotel profitability.

2. Offline Travel Agents/Tour Operators (TA/TO):

- Account for **20.3% of bookings (24,159)**, showing their continued importance, especially for guests who prefer traditional booking methods.

3. Group Bookings:

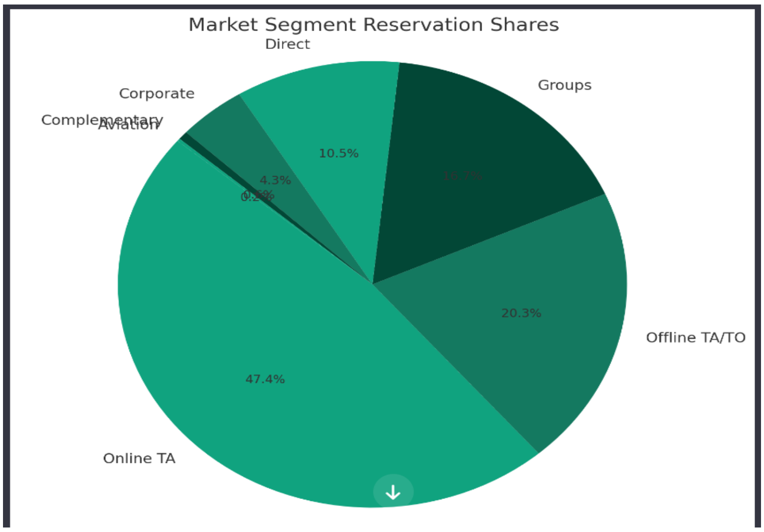
- Represent **27% of reservations (19,806)**, a significant revenue driver.
- Tailored packages for group travelers can increase this segment's profitability.

4. Direct Bookings:

- Only **4% of bookings (12,448)** come directly from guests, presenting a major growth opportunity.
- Increasing direct bookings can reduce dependency on OTAs and improve profit margins.

5. Other Sources:

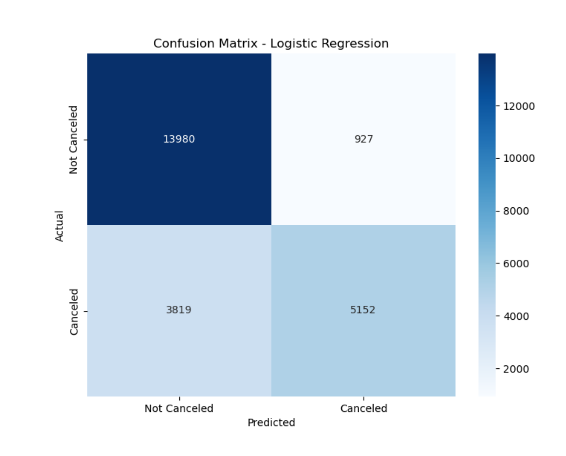
- Smaller contributions come from **corporate bookings (5,111)**, **complementary stays (734)**, and **aviation-related bookings (237)**, which are important for diversification.

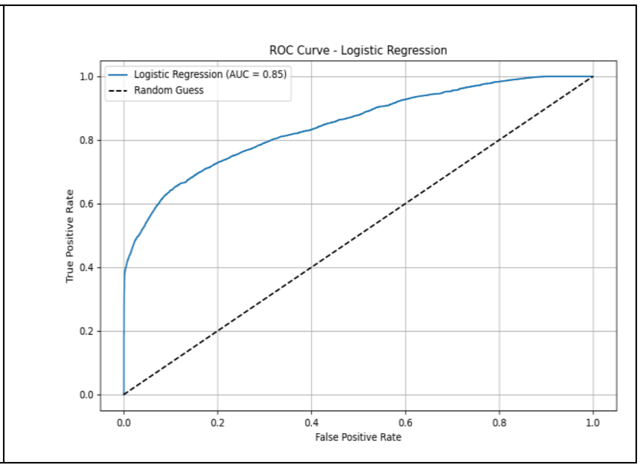


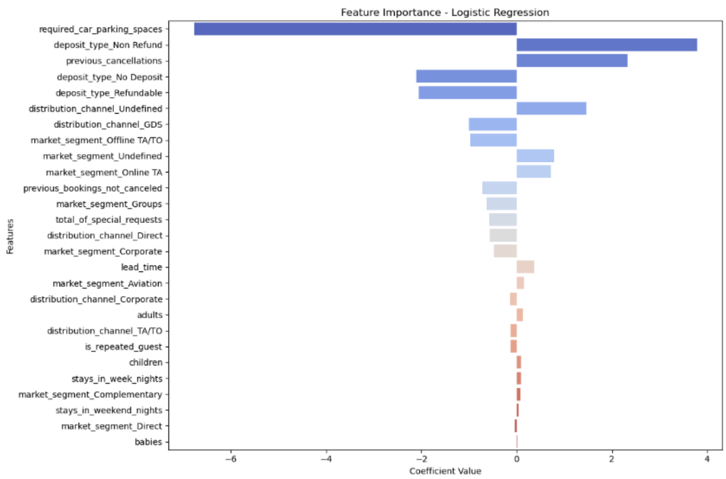
Strategic Recommendations

1. **Increase Direct Bookings**:
   - Offer loyalty programs, personalized discounts, and direct marketing to attract more direct reservations.
2. **Optimize OTA Strategy**:
   - Use OTAs to reach a wider audience while negotiating lower commission rates and targeting high-value customers.
3. **Enhance Group Packages**:
   - Develop competitive group travel packages with bundled amenities to boost group bookings.
4. **Diversify Revenue Channels**:
   - Strengthen partnerships with Offline TA/TOs and corporate clients to reduce over-reliance on OTAs and balance revenue streams.

Predictive Modelling on Research Questions







Research Question 1: Why Do People Cancel Their Hotel Bookings?

To understand why guests cancel their bookings, a **Logistic Regression model** was used to analyze booking characteristics and customer behavior.

Key Results

1. **Model Performance**:
   - **Accuracy**: The model correctly predicted cancellations 80.1% of the time, showing strong reliability.
   - **ROC AUC Score**: Achieved 84.6%, indicating good ability to distinguish between cancelled and non-cancelled bookings.
2. **Important Factors**:
   - **Lead Time**: Longer booking lead times increased the chances of cancellations.
   - **Deposit Type**: Non-refundable deposits reduced the likelihood of cancellations.
   - **Special Requests**: Guests making more special requests were less likely to cancel, indicating stronger commitment.
3. **Confusion Matrix Insights**:

    - **Correct Predictions (True Positives)**: 5,152 bookings were accurately identified as cancellations.
    - **Missed Cancellations (False Negatives)**: 3,819 bookings were incorrectly classified as non-canceled, suggesting room for improvement.

Actionable Insights

1. **Guest Behavior**:
    - Guests with long lead times or refundable deposits are more likely to cancel, highlighting the need for tailored strategies to retain these bookings.
2. **Channel Analysis**:
    - Higher cancellation rates from Online Travel Agencies (OTAs) suggest the need for closer monitoring and specific retention efforts for this segment.

Recommendations

1. **Flexible Booking Policies**:
    - Offer tiered pricing options that vary by cancellation terms to balance guest flexibility and protect revenue.
2. **Target High-Risk Guests**:
    - Focus retention campaigns on guests with long lead times or those who frequently cancel, using personalized offers or engagement strategies.
3. **Increase Guest Commitment**:
    - Encourage guests to make special requests or choose non-refundable deposits, as these behaviors are linked to lower cancellation rates.

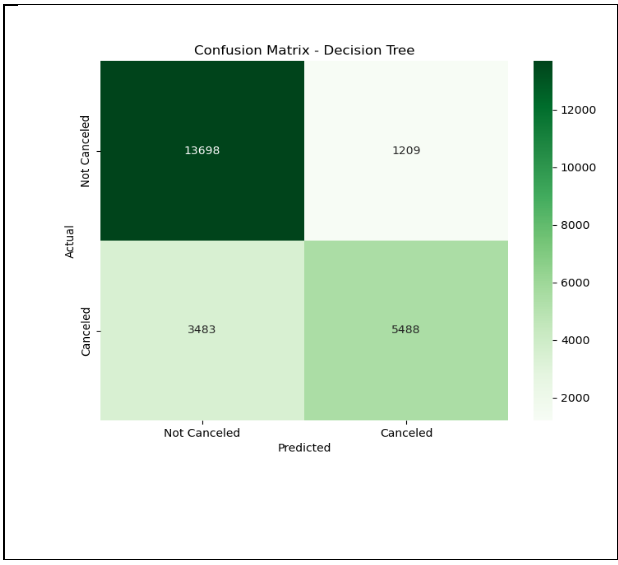
Research Question 2: How Can We Help Hotels Have Fewer Cancellations?

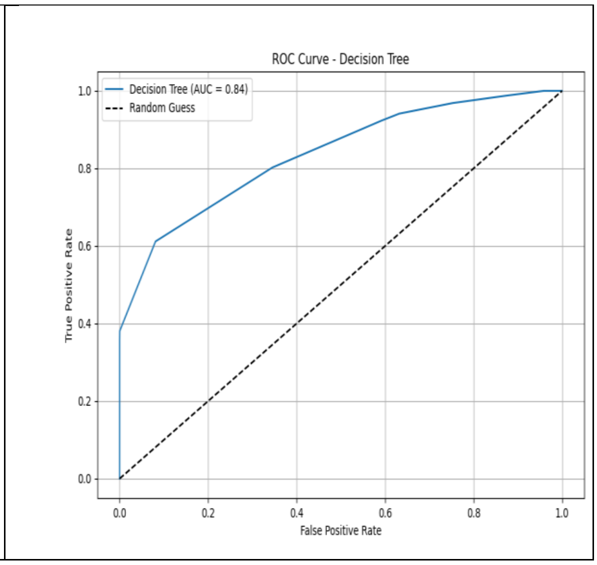
The Decision Tree model reveals key factors influencing cancellations and offers actionable strategies to minimize them.

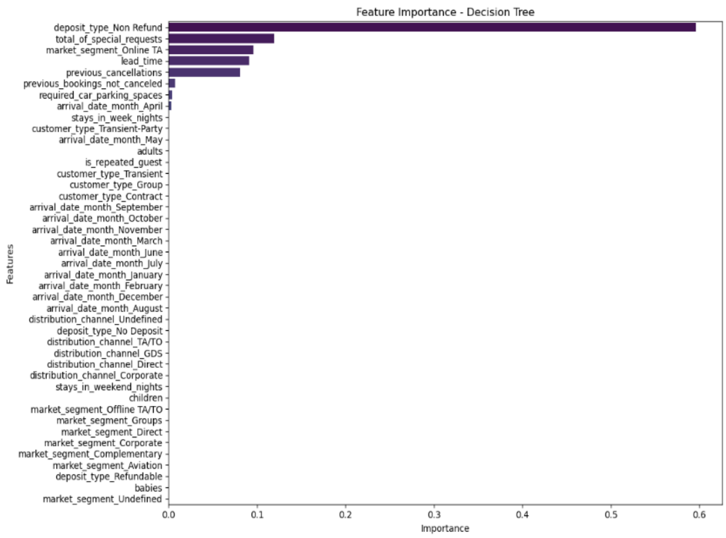
Key Findings

1. **Most Important Factors**:
    - **Deposit Type**: Non-refundable deposits significantly reduce cancellations.
    - **Special Requests**: Guests with more requests are less likely to cancel, showing stronger commitment.
    - **Market Segment**: Bookings via Online Travel Agencies (OTAs) are more prone to cancellations.
    - **Lead Time**: Longer lead times increase the chances of cancellations.
    - **Seasonality**: Certain months (e.g., January) have higher cancellation rates.
2. **Model Performance**:
    - **Accuracy**: 80.3%, indicating strong reliability in predicting cancellations.
    - **ROC AUC Score**: 84%, reflecting excellent ability to differentiate canceled and non-canceled bookings.
    - **Confusion Matrix**:
        - Correctly predicted cancellations: 5,488.
        - Missed cancellations: 3,483, showing room for improvement.

3. **Customer Behavior Insights**:
   - **Transient Customers**: They cancel more often compared to group bookings.
   - **High-Risk Channels**: OTAs and specific months (e.g., January) have higher cancellation tendencies.







Recommendations

1. **Promote Commitment**:
   - Encourage non-refundable deposits by offering added perks like discounts or free services.
2. **Personalized Guest Engagement**:
   - Increase opportunities for guests to make special requests, enhancing their commitment to bookings.
3. **Focus on High-Risk Segments**:
   - Target guests with long lead times or those booking through OTAs with personalized retention efforts.

4. **Seasonal Promotions**:
    - Offer discounts and flexible packages during high-cancellation months, such as January, to retain bookings.
5. **Optimize OTA Collaboration**:
    - Partner with OTAs to identify high-risk customers and create retention strategies tailored to their needs.

Strategies

1. **Flexible Booking Policies**:
    - Introduce options like partial refunds or flexible rescheduling to meet guest needs without losing revenue.
2. **Incentives for Shorter Lead Times**:
    - Provide last-minute deals or exclusive packages to reduce cancellations associated with long lead times.
3. **Loyalty Programs**:
    - Develop rewards for repeat customers and group travelers to encourage long-term relationships.
4. **Data-Driven Campaigns**:
    - Use the model to identify high-risk bookings and proactively engage guests with personalized offers and follow-ups.
5. **Operational Adjustments**:
    - Prepare for peak cancellation periods by adjusting staffing, marketing efforts, and promotional strategies.

Research Question 3: How Can We Help Hotels Decide on Their Prices and Special Offers?

The analysis reveals key factors influencing the **Average Daily Rate (ADR)** and provides actionable strategies to optimize pricing and revenue.

Key Findings

1. **Factors Influencing ADR**:
    - **Lead Time**: Guests booking earlier tend to pay more, as longer lead times positively impact ADR.
    - **Market Segments**:
        - **Corporate and Direct bookings** have the highest ADRs.
        - **OTA and group bookings** show lower ADRs, requiring improvement strategies.
    - **Seasonality**:
        - ADRs peak during **August** and **December** due to high demand.
        - Lower ADRs in **January** and **February** indicate the need for targeted offers.
    - **Special Requests**: Guests making more special requests are willing to pay higher rates, supporting premium pricing strategies.
2. **Model Performance**:
    - **MSE**: Indicates the model effectively captures ADR variations with minimal errors.

    - **R^2 Score**: Confirms a strong link between features and ADR, validating the model's accuracy.
3. **Seasonal Trends**:
    - Peak pricing opportunities in **August** and **December**.
    - Low ADRs in off-peak months like **January** and **February**.
4. **Market Segment Insights**:
    - **Corporate and Direct bookings** are the most profitable.
    - **OTA and group bookings** need targeted efforts to boost profitability.

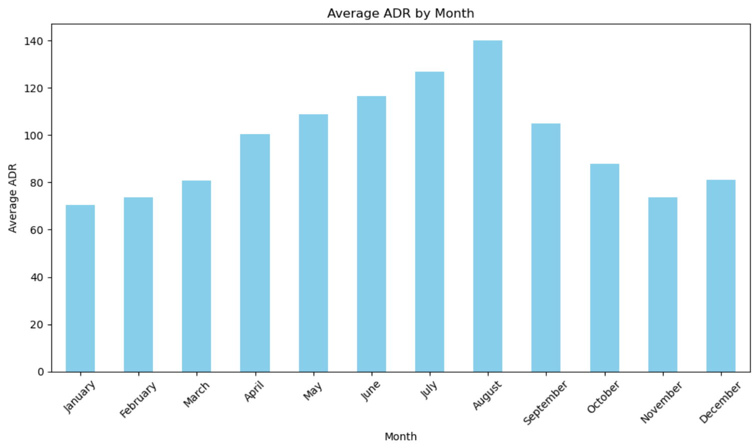
Recommendations

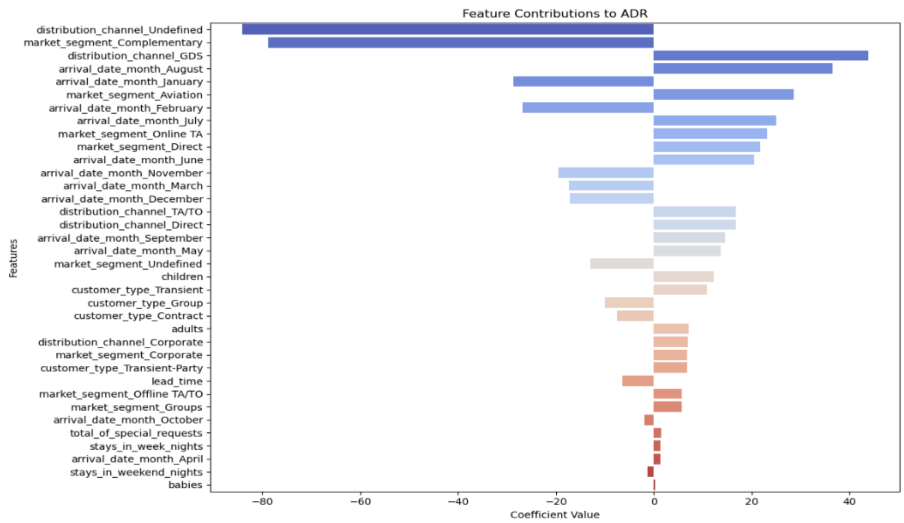
1. **Dynamic Pricing**:
    - Adjust pricing in real-time to capitalize on peak seasons like **August** and **December**.
    - Offer **early-bird discounts** for advanced bookings to encourage higher ADRs.
2. **Specialized Packages**:
    - Create tailored packages for high-ADR segments like **Corporate and Direct bookings**.
    - Include bundled services such as premium amenities or exclusive deals.
3. **Seasonal Promotions**:
    - Offer promotions during low-demand months (**January** and **February**) to maintain occupancy.
    - Add value to bookings with offers like **free breakfast** or extended stays.
4. **Premium Services**:
    - Upsell tailored options (e.g., spa packages, room upgrades) to guests making special requests.
5. **Market Optimization**:
    - Work with OTAs to promote higher-tier rooms and attract higher-paying customers.
    - Develop loyalty programs to encourage **Direct bookings** and reduce reliance on OTAs.

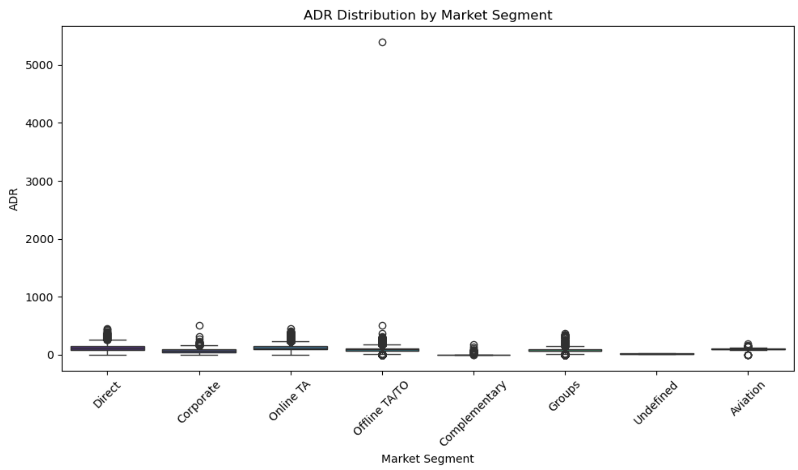
Actionable Strategies

1. **Early-Bird Promotions**:
    - Offer discounts for bookings made **3+ months in advance** to leverage lead time effects.
2. **Seasonal Adjustments**:
    - Maximize profits during peak months (**August**, **December**) with higher ADRs.
    - Provide incentives for bookings during off-peak periods.
3. **Focus on Profitable Segments**:
    - Target **Corporate and Direct bookings** with exclusive offers like business-class rooms.
4. **Data-Driven Pricing**:
    - Monitor ADR trends across segments and seasons to adjust pricing dynamically.
5. **Upselling Opportunities**:
    - Use guest data to upsell premium amenities based on preferences and special requests.

.







Methodology/Model Building/Model Selection

This section outlines the methods and models used to analyze the data, predict cancellations, and optimize pricing strategies, with Python as the primary tool for its versatile libraries in data analysis and machine learning.

Data Preparation

1. **Cleaning**:
   - Missing values were handled: `children` set to 0, `agent` and `company` set to 0, and `country` filled as 'Unknown.'
   - Outliers were managed using statistical techniques to ensure accurate analysis.
2. **Feature Engineering**:
   - Key features selected: `lead_time`, `previous_cancellations`, `market_segment` (for cancellations), and `adr`, `arrival_date_month` (for pricing).
   - Categorical data (e.g., `market_segment`) was one-hot encoded, and numerical data was standardized for consistency.

Model Selection

1. **Logistic Regression**:
   - Used to predict cancellations, focusing on features like `lead_time` and `deposit_type` to identify cancellation behavior.
2. **Decision Tree Classifier**:
   - Highlighted actionable factors such as `total_of_special_requests` and `deposit_type` to suggest strategies for reducing cancellations.
3. **Multiple Linear Regression**:
   - Analyzed pricing trends and ADR, examining variables like `lead_time` and `arrival_date_month` to provide recommendations on pricing and offers.

Model Training and Evaluation

- **Data Split**: Training (80%) and testing (20%) sets.
- **Metrics**:
   - Classification models: Accuracy, ROC-AUC, precision, and recall.
   - Regression models: Mean Squared Error (MSE) and R^2 Score.
- Cross-validation ensured reliable and generalizable results.

Data Visualization

- **EDA Techniques**:
   - Bar graphs, pie charts, and heatmaps analyzed cancellation rates, market segments, and ADR patterns.
- **Model Insights**:
   - Feature importance plots (Decision Trees) and regression coefficients (Linear Regression) identified key drivers.
   - Confusion matrices and ROC curves demonstrated model performance visually.

Challenges and Solutions

1. **Data Imbalance**:

    - Addressed using balanced class weights and oversampling techniques.
2. **Seasonality**:
    - Incorporated `arrival_date_month` to account for seasonal effects on cancellations and pricing.
3. **Multicollinearity**:
    - Managed by selecting independent variables after correlation analysis, ensuring robust regression models.

Tools and Techniques

- **Data Handling**: pandas for cleaning and preprocessing.
- **Visualization**: seaborn and matplotlib for advanced graphs.
- **Machine Learning**: scikit-learn for building and evaluating models.
- **Approach**:
    - Combined EDA to explore data patterns with machine learning to predict outcomes and derive actionable insights.

Conclusions and Strategic Recommendations

The analysis highlights the importance of dynamic pricing strategies to address high cancellation rates during off-peak months like January, where elevated ADRs correlate with increased cancellations. Conversely, peak periods such as August and December provide an opportunity for hotels to maximize revenue through premium pricing. Market segmentation reveals that while Online Travel Agencies (OTAs) dominate bookings, they also exhibit the highest cancellation rates. In contrast, Corporate and Direct bookings generate higher ADRs and fewer cancellations, presenting a profitable focus area. Furthermore, long lead times are strongly associated with cancellation risks, emphasizing the need for incentives to encourage shorter booking windows. The role of special requests in increasing ADRs underscores the potential of upselling personalized services. Seasonality remains a critical factor, with higher bookings during summer and holiday seasons and elevated cancellations in early months like January and February.

- To address these challenges, hotels should adopt dynamic pricing models that adjust rates in real-time based on market conditions, seasonality, and demand patterns. Early-bird discounts can secure advanced bookings, while premium packages during high-demand months can optimize revenue. Flexible booking policies, such as tiered cancellation options or non-refundable deposits with added perks, cater to guest needs while minimizing financial risks. Portugal-specific campaigns, including enhanced service quality and infrastructure, can reduce its disproportionately high cancellation rates. Corporate and Direct bookings should be incentivized with loyalty rewards and tailored packages to maintain their profitability. Personalized marketing campaigns targeting high-risk segments, such as guests with long lead times, and upselling premium services like room upgrades or luxury amenities can further enhance revenue.
- Operationally, hotels should prepare for peak seasons by optimizing staffing and resources while using predictive models to anticipate high-risk bookings and engage at-risk guests proactively. Collaborations with OTAs can help target cancellation-prone customers with tailored retention strategies. Leveraging AI-driven tools for dynamic pricing adjustments and machine learning models to predict cancellations

can significantly enhance operational efficiency and decision-making, helping hotels maintain steady occupancy and profitability.

Bibliography/References/Works Cited

1. Hotel Booking Dataset, retrieved from Kaggle.
2. https://www.siteminder.com/r/hotel-industry-statistics/
3. https://www.oliandalex.com/the-impact-of-hotel-cancellation-fees-on-travelers/
4. https://www.hotelminder.com/everything-you-need-to-know-about-hotel-cancellations
5. https://hoteltips101.com/how-to-handle-cancellations-at-hotels/
6. https://shrgroup.com/2023/06/21/we-need-to-talk-about-cancellations/
7. https://www.mirai.com/blog/cancellations-shooting-up-implications-costs-and-how-to-reduce-them/
8. Python Libraries Documentation: pandas, matplotlib, seaborn, and scikit-learn.

Data Source and Description

- **Source**: The dataset was sourced from Kaggle and contains anonymized records of hotel bookings.
- **Scope**: Covers the years 2015 to 2017, including data on City and Resort Hotels.
- **Records**: The dataset contains 119,390 rows and 32 columns.
- **Key Variables**:
  - `lead_time`: Days between booking and arrival.
  - `adr`: Average Daily Rate (price per night).
  - `is_canceled`: Whether the booking was canceled (1) or not (0).
  - `market_segment`: Booking source such as Direct, Corporate, or OTA.
  - `arrival_date_month`: Month of the booking.