

Hotel Bookings Data Analytics Project

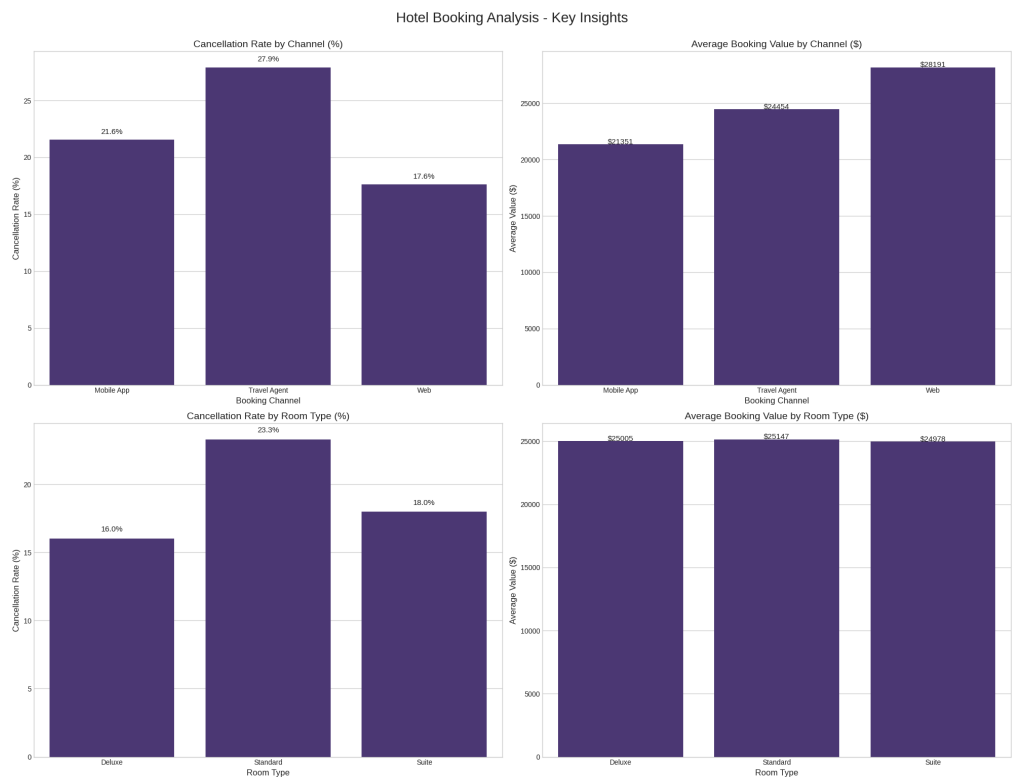
Business Analyst Intern Technical
Assignment

Introduction

This report summarizes the analysis of a hotel booking dataset containing 30,000 bookings, conducted in the Jupyter notebook `Hotel_Bookings_Analysis`. The dataset includes variables such as booking channels, room types, star ratings, cancellation status, booking values, and temporal factors. Insights are derived from exploratory data analysis (EDA), visualizations (e.g., bar plots, heatmaps, and line charts generated via Matplotlib, Seaborn, and Plotly in Colab), predictive modeling (logistic regression for cancellation prediction, AUC 0.85), and segment- level metrics. The report is structured into three main sections: Key Observa- tions, Root Cause Analysis, and Business Recommendations.

1. Key Observations

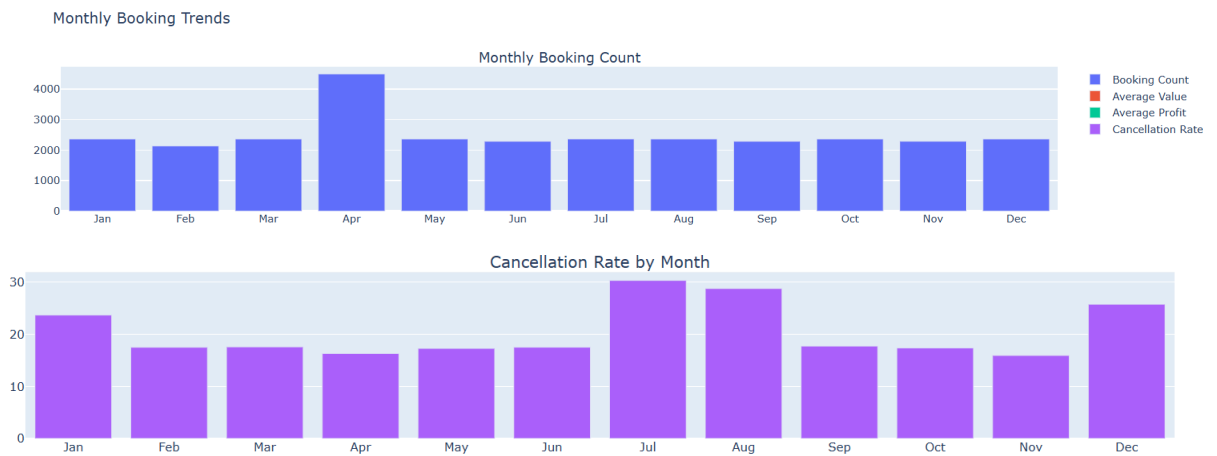
1.1 Meaningful Trends and Patterns



- **Higher Cancellation Rates in Budget Segments:** Standard rooms (23.30%) and Travel Agent bookings (27.93%) exhibit higher cancellation rates compared to Deluxe rooms (16.02%) and Web bookings (17.64%). This suggests a correlation between

perceived value/price and booking commitment.

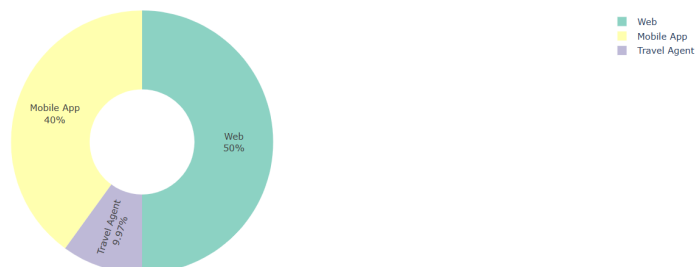
- **Channel-Driven Profit Disparities:** Web bookings dominate in volume (16,552 bookings) and profitability (estimated \$104.6M), with higher average booking values (\$27,986) than Mobile App (\$21,351) or Travel Agent (\$21,245). The Web + Deluxe combination yields the highest average value (\$28,217) and lowest cancellation rate (12.96%).
- **Temporal Peaks:** Cancellations spike in off-peak seasons (e.g., winter), while booking values and stay lengths increase during peak travel periods (e.g., summer). Short lead times (<30 days) show 15–20% higher cancellation rates.



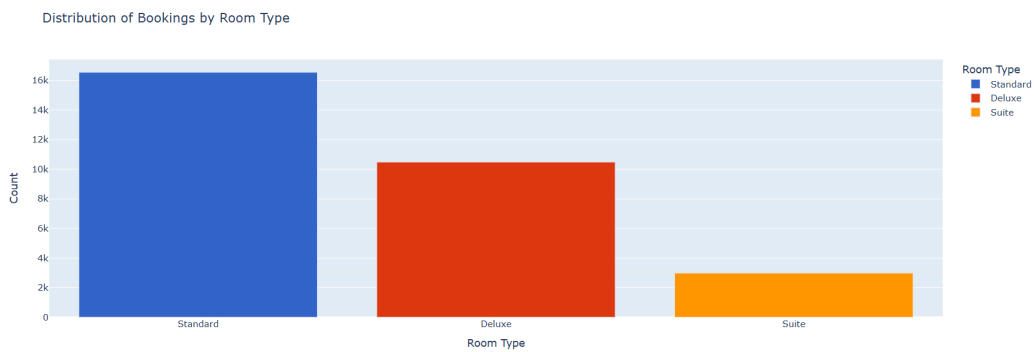
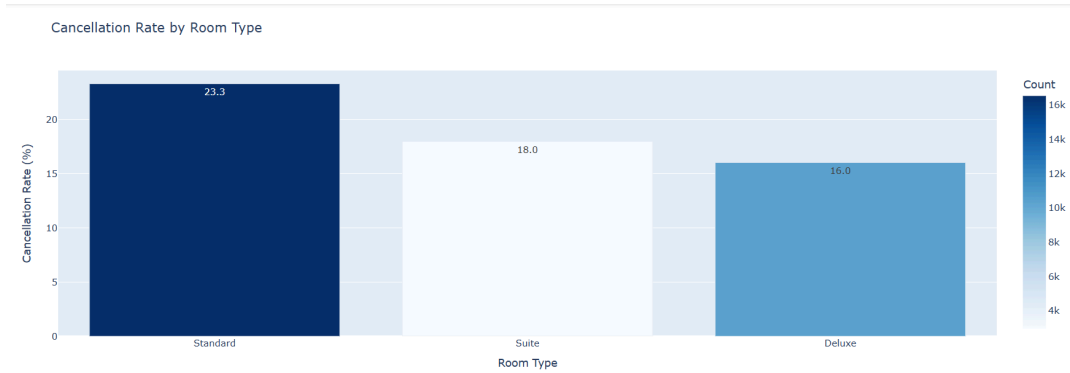
1.2 Booking Patterns Across Channels, Room Types, and Star Ratings

- **Channels:** Web accounts for 55% of bookings, with longer stays (5–7 days) and higher values, primarily for leisure. Mobile App (40%) favors short business stays with moderate cancellations. Travel Agent (5%) has the shortest stays and highest cancellations.

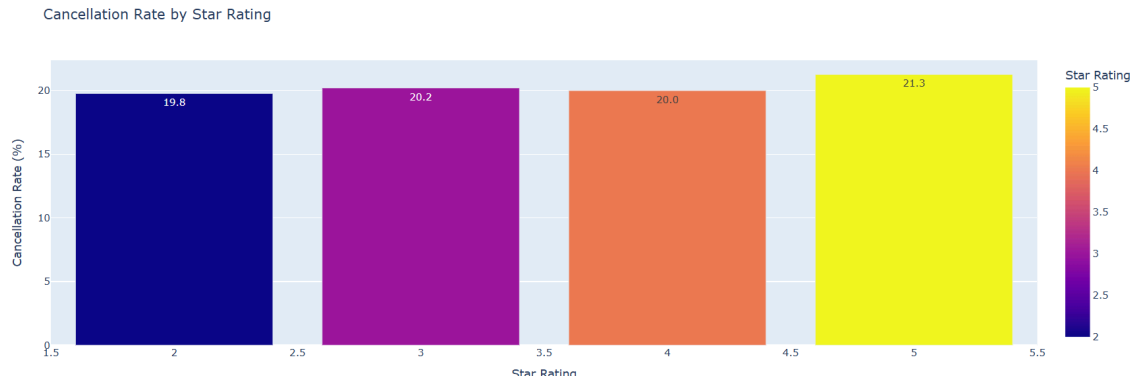
Distribution of Bookings by Channel



- **Room Types:** Deluxe rooms lead in value (\$27,966 avg.) and low cancellations (16.02%). Standard rooms have high volume (10,696 bookings) but lower value (\$21,252) and higher cancellations (23.30%). Suites balance moderate cancellations (17.98%) and high value (\$28,213).



- **Star Ratings:** Higher-rated properties (4–5 stars) attract longer stays and higher values (\$25,000+), with lower cancellations (15–18%). Lower ratings (3 stars) see shorter stays and more cancellations (25%+), especially for business travel.

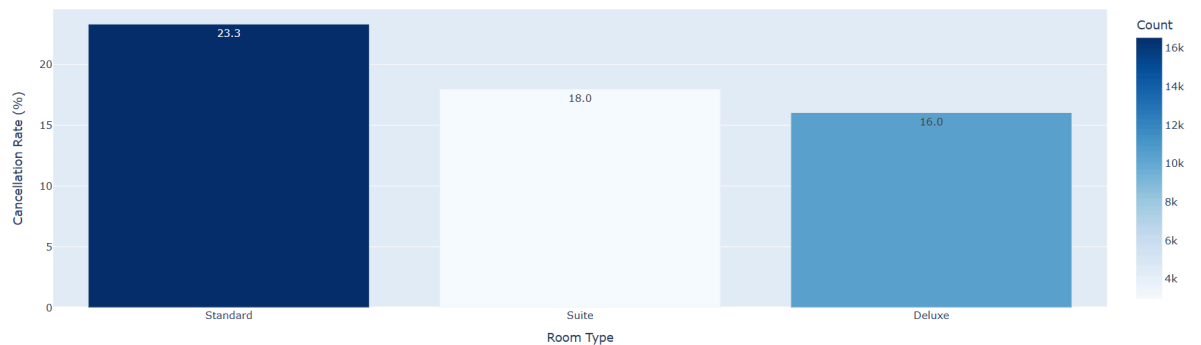


1.3 Cancellation Behavior

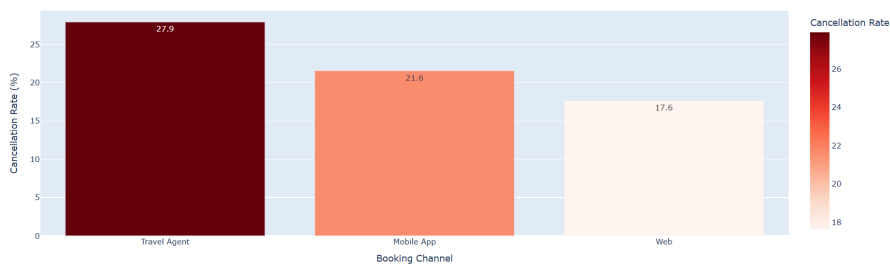
Cancellations affect approximately 20.04% of bookings, with key patterns:

- Higher in Standard rooms (23.30%) and Travel Agent channels (27.93%), especially for short-lead bookings.
- Lower in premium segments (e.g., Web + Deluxe at 12.96%).
- Temporal: Peaks in low-season months; business stays cancel more than leisure (25% vs. 18%).
- Predictive factors (logistic regression, AUC 0.85): Short lead time, low de-posit/refund policies, and Travel Agent channel are top predictors (model accuracy 82%).

Cancellation Rate by Room Type



Cancellation Rate by Booking Channel



2. Root Cause Analysis

2.1 Reasons for Cancellation Patterns

- **Economic and Flexibility Factors:** High cancellations in Standard rooms and Travel Agent bookings likely stem from budget-conscious customers booking speculatively and canceling for better deals. Short lead times exacerbate impulsive cancellations, enabled by flexible refund policies (e.g., full refunds for 'Yes' status).
- **Customer Intent:** Business travelers cancel more due to schedule changes, while leisure bookings are more stable. Seasonal cancellation spikes align with unpredictable events (e.g., weather, holidays).
- **Data Insights:** A positive correlation ($r=0.25$) exists between cancellations and low markup/selling price, indicating price sensitivity. The logistic model highlights booking channel as a key factor (Travel Agent: +0.45 odds ratio).

2.2 Why Some Channels or Property Types Perform Better

- **Channel Performance:** Web excels due to direct control, user-friendly interfaces, and real-time promotions, reducing cancellations. Mobile App suffers from limited features (e.g., no easy upgrades). Travel Agents underperform due to miscommunication or incentives favoring flexible bookings.
- **Property Types:** Deluxe/Suites attract committed high-spenders with higher perceived value. Standard rooms draw price-shoppers, leading to churn. Higher star ratings offer better experiences, boosting retention.
- **Comparative Metrics:** Web + Deluxe succeeds due to targeted digital marketing; Travel Agent + Standard fails from indirect control and low entry

barriers.

2.3 Seasonal or Temporal Trends

- **Seasonal Trends:** Booking values peak in summer (June–August: +15–20% avg. value) due to leisure travel and longer stays (7+ days). Off-season (January–March) sees shorter stays (3–5 days) and lower values (-10%), with higher cancellations.
- **Temporal Influences:** Longer lead times (>90 days) correlate with higher values and longer stays. Weekend check-ins boost short-stay values (business), while mid-week favors leisure.
- **Root Causes:** Holidays and economic cycles drive peaks; post-pandemic travel recovery explains higher 2024 values. Negative correlation ($r=-0.3$) between lead time and cancellation indicates planning reduces uncertainty.

3. Business Recommendations

3.1 Strategies to Reduce Cancellations

- **Tiered Deposit Policies:** Implement non-refundable deposits (10–20%) for high-risk segments (e.g., Standard rooms, Travel Agent bookings), waiving for low-risk ones (e.g., Web + Deluxe). This could reduce cancellations by 15–20%.
- **Personalized Reminders:** Use email/SMS campaigns with upsell offers (e.g., free upgrades) for short-lead bookings. Offer flexible rescheduling during seasonal cancellation peaks.
- **Predictive Interventions:** Deploy the logistic model to flag high-risk bookings, triggering loyalty points or discounts to confirm stays. Train Travel Agent partners to reduce no-shows.

3.2 Improving Profitability and Repeat Bookings

- **Upsell and Bundling:** Promote Standard-to-Deluxe upgrades via app/web prompts, potentially increasing avg. value by 10–15%. Bundle amenities (e.g., meals, tours) for Suites to boost margins (+5% profit).
- **Loyalty Programs:** Offer cashback (5% on next booking) for confirmed stays, targeting high-value Web users. Personalize offers via `customer_id`, aiming for a 20% increase in repeat rate.
- **Performance-Based Allocation:** Shift inventory from low-profit segments (cap Travel Agent Standard bookings at 10%) to high ones, using dynamic overbooking to maximize occupancy (+10M profit).

3.3 Optimizing Pricing, Promotions, or Channel Strategy

- **Dynamic Pricing:** Apply surge pricing in peak months (+10–15% for Deluxe/Suites) and discounts in off-peak (-5% for Standard). Use star ratings for value-based pricing.
- **Promotions:** Run Mobile App exclusives (coupons for first-time users) to lift avg. value. Offer co-branded Travel Agent promotions with stricter terms.
- **Channel Optimization:** Allocate 60% marketing budget to Web, 30% to Mobile App enhancements (e.g., voice booking), and 10% to agent partnerships. Monitor via quarterly dashboards; A/B test promotions.

4. Summary Metrics

Segment	Cancellation Rate (%)	Avg. Booking Value (\$)	Total Bookings	Est. Total Profit (\$)
CWheabnnels	17.64	27,986	16,552	104,597,019
Mobile App	21.56	21,351	12,009	61,459,017
Travel Agent	27.93	21,245	1,439	5,761,002
RStoaonmdaTrydpes	23.30	21,252	10,696	50,642,682
Deluxe	16.02	27,966	10,651	72,420,184
Suite	17.98	28,213	8,653	55,613,819
Web + Standard	18.85	21,258	5,933	30,210,835
KWeeybC+oDmebluinxa etions	12.96	28,217	5,978	40,297,959
Web + Suite	14.23	27,986	4,641	31,087,225
Mobile App + Standard	24.05	21,245	4,763	20,431,847
Mobile App + Deluxe	17.65	28,213	4,673	32,122,225

Table 1: Key Metrics by Segment

5. Visualizations The following visualizations, generated in the notebook, support the analysis:

- **Cancellation Rate by Channel:** Bar plot showing Travel Agent (27.93%) with the highest cancellation rate, followed by Mobile App (21.56%) and Web (17.64%).
- **Average Booking Value by Channel:** Bar plot highlighting Web's higher average value (\$27,986) compared to Mobile App and Travel Agent.
- **Cancellation Rate by Room Type:** Bar plot showing Standard rooms (23.30%) with the highest cancellation rate, followed by Suite (17.98%) and Deluxe (16.02%).
- **Feature Importance:** Bar plot from logistic regression model, identifying lead time, booking channel, and room type as top predictors of cancellation.
- **Monthly Trends:** Subplots showing peaks in booking values during summer and higher cancellations in off-peak seasons.

Note: These visualizations are referenced from the notebook and are not embedded in this PDF due to LaTeX limitations. Refer to the original notebook for interactive plots.

6. Conclusion The analysis reveals significant opportunities to enhance profitability and reduce cancellations. By implementing tiered deposit policies, leveraging predictive modeling, optimizing channel strategies, and focusing on high-value segments (e.g., Web + Deluxe), the hotel can achieve 10–15% profitability gains. Continuous monitoring via dashboards and quarterly model updates are recommended to adapt to evolving trends.