Hotel Performance Business Intelligence Report

By Ellie Crossley-Fells

Date of Publication: 15/07/2025

Table of Contents

1. Executive Summary

This report evaluates the operational performance of two hotels managed by Sleep; City Hotel and Resort Hotel, focusing on booking trends, pricing strategies, guest behavior and demographics, and agent performance.

- **Hotel Comparison:** The City Hotel consistently hosts more guests, has a higher average daily rate (ADR) compared to the Resort Hotel, and Resort Hotel recieves a higher proportion of special requests.

- **Seasonality of Bookings:** August has the largest volume of bookings overall. January is the quietest month (likely driven by a reduced number of guests going on holiday, after increased spending during the Christmas period). February has the lowest Average ADR, offering potentially better booking opportunities for guests. For visitors wanting the best rates for a summer holiday, June offers the lowest ADR out of the summer months.

- **Stay Duration Savings:** Lowest daily rates for visitors are typically for stays between 20 to 25 nights, with all ADRs in this duration being under £100. For shorter stay, 7 days offers the lowest median ADR. Average stay length however is 3-4 days - consider deals in this length of stay too.

- **Guest Demographics:** Most guests come from Portugal, Great Britain, and France with no visitors from Canada or Greenland within the dataset period. City Hotel appears to hosts slightly more adults (suggesting a business/solo travelers), while Resort Hotel serves more families.

- **Agent Performance:** Agent 9 is the top-performing agent, with the largest number of bookings. Agent 240 is runner-up.

- **Cancellation Forecast**: 37% of bookings cancel. A Random Forest model is being developed to determine what factors increase likelihood of cancellation. Deposit type and Country have the strongest correlations with a cancelled booking.

- **Recommendations**: include targeted promotions during off-peak months, offer packages for extended stays and target marketing to fit each hotel's guest profile.

2. Introduction
Sleep operates both a City Hotel and a Resort Hotel. The primary business aims is to produce a model to classify the likelihood of cancellations at these hotels, and also to better understand booking behaviour, optimal pricing strategies, and guest demographics to improve occupancy, which in turn will improve revenue. This report addresses several specific management questions, including peak booking periods, optimal stay length for best rates, special request patterns, and agent performance, using data collected across all bookings for the two hotels, from between July 2015 to August 2017.

3. Methods and Findings

Data Cleaning

- Checked for null values and converted column data types.

- Created a unique identifier column for rows to aid in potential tracking and outlier analysis.

- Created a derived arrival date column and standardized month, day, and year data types.

- Ensured variables in text columns (e.g. reservation status column) were consistent; addressed spelling variations.

- Replaced null values in Agent and Company fields with "No Agent" and "No Company."

- Created a total stay duration column by summing weekday and weekend night stays columns.

- Removed outliers: some ADRs were 0 and these were filtered out during investigation of optimal length of stay for best average daily rate.
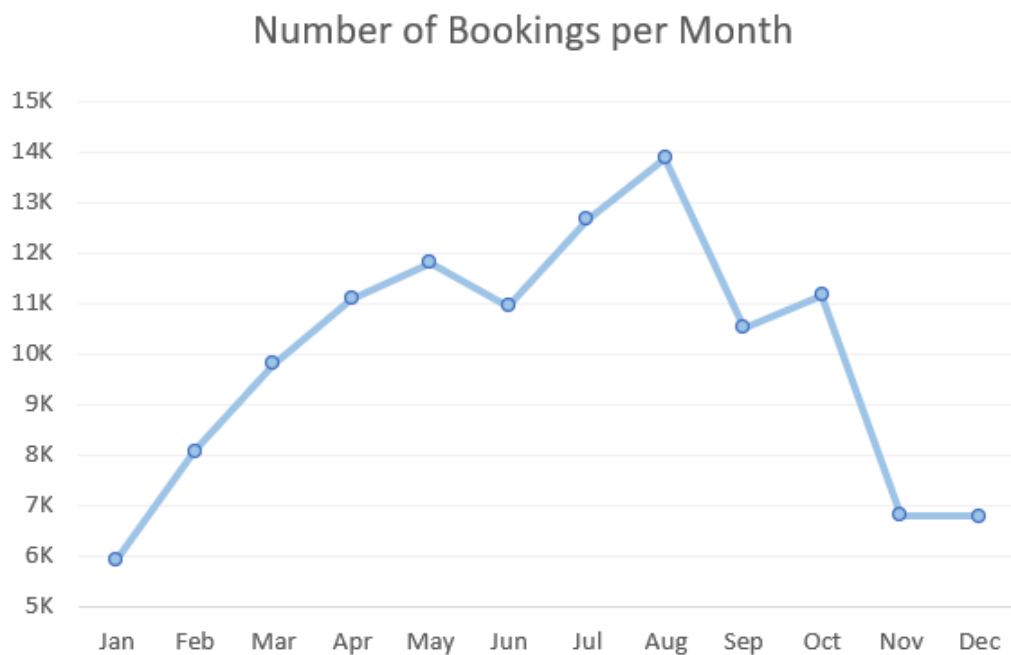
Booking Trends & Best Time to Book

Figure 1: Line chart showing total number of bookings per month

- Busiest Month: August (c.13.8k bookings)

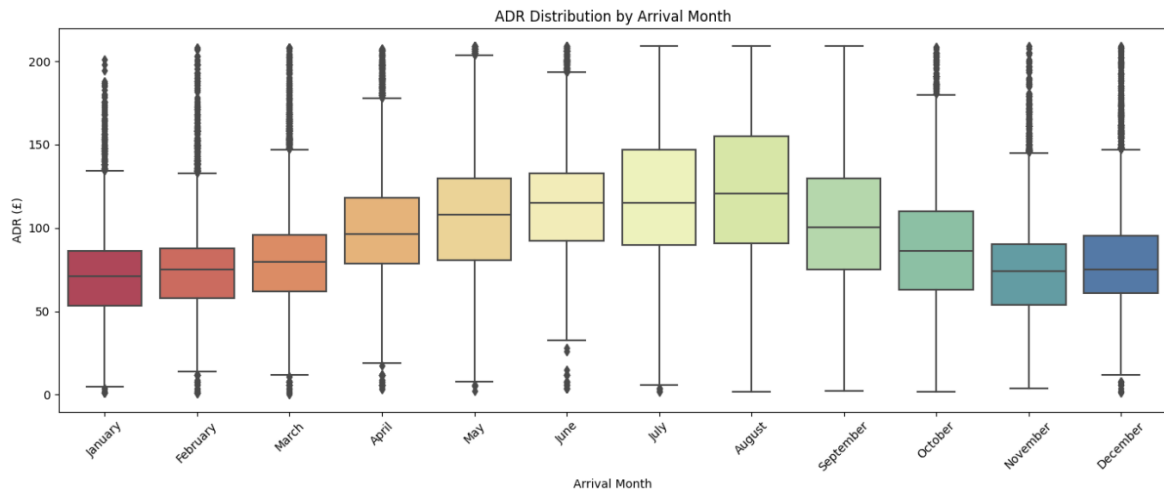- Quietest Month: January (c.5.9k bookings)

Seasonal Insights

Figure 2: Boxplot showing the distribution of the ADR per month

Low Season

- January, February, March, November, and December have the lowest median ADR (center lines in the boxes), between £60 and £80.

- These months also have a narrower interquartile range (IQR, the hight of the boxes), indicating consistently lower prices and less rate variability.

High Season

- May, June, July, and August show the highest median ADRs, with values reaching well above £100.

- The boxes are taller for these months, indicating more rate variability, caused by higher demand due to school holidays.

Month	Volume	Average ADR	Max ADR
Jan	5929	£70.36	£284.20
Feb	8068	£73.58	£284.20
Mar	9793	£80.14	£285.00
Apr	11089	£100.38	£336.00
May	11791	£108.70	£510.00
Jun	10939	£116.67	£333.67
Jul	12661	£126.79	£508.00
Aug	13877	£140.11	£450.00
Sep	10508	£105.05	£345.00
Oct	11160	£87.91	£300.00
Nov	6794	£73.79	£260.00
Dec	6780	£81.08	£451.50

Figure 3: Table showing the volume of bookings, average ADR and maximum ADR per month.

- Rooms are easiest to book in winter due to lower numbers of bookings; January and February offer the lowest average rates (£70.40 and £73.60 respectively). February has less variation in the ADR compared to Jan and November has the lowest maximum ADR out of all months.

- June is the cheapest month for summer stays, with a Maximum ADR of £333 compared to £510 for May, £508 for July, and £450 for August.

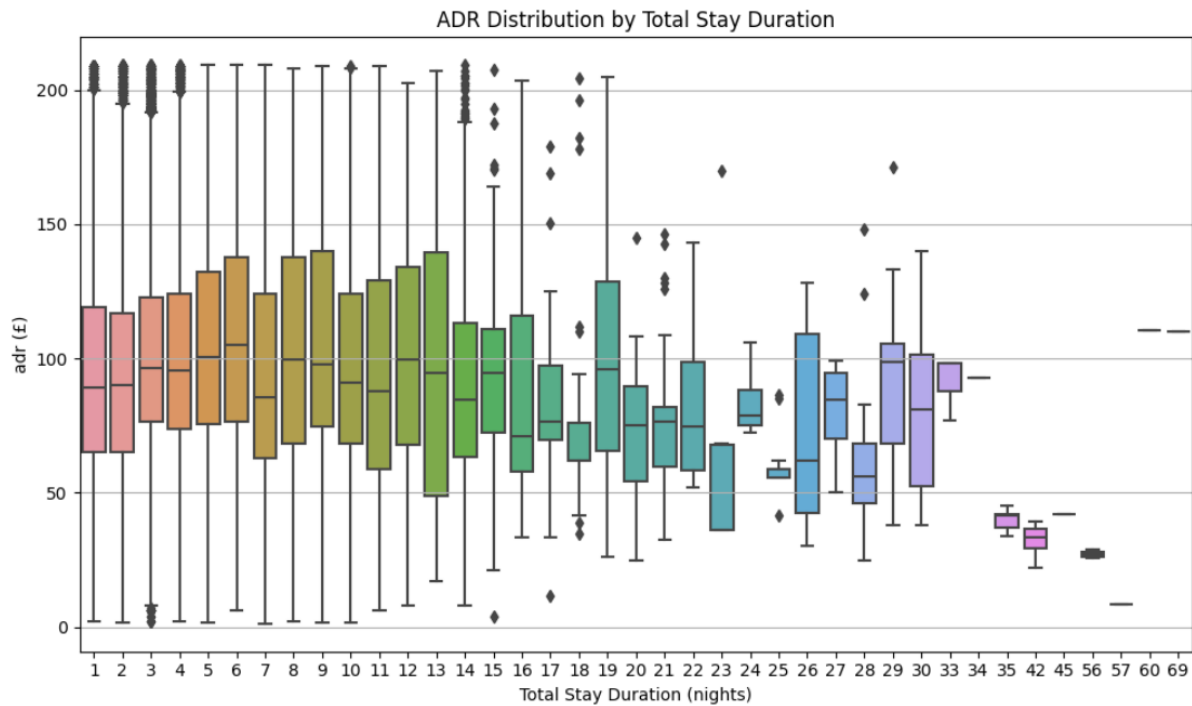**Optimal Stay Duration and ADR**

Figure 4: Boxplot showing ADR vs length of stay. ADRs of £0 have been excluded from this plot (often complimentary stays and cancellations)

Boxplot analysis indicated:

- Short Stays (1–13 nights): Higher, more variable ADR.

- Long Stays (17–24+ nights): Lower, more stable ADR.

- Stays of 20–25 nights offer the lowest, most consistent nightly rates.

Special Requests Analysis

Figure 5: Bart chart showing total number of special requests per hotel

- City Hotel: 43,000 special requests.

- Resort Hotel: 25,000 special requests.

- However, 62% of bookings have special requests at Resort Hotel, and 54.7% at City Hotel - meaning **Resort Hotel has a disproportionately higher number of special requests.**
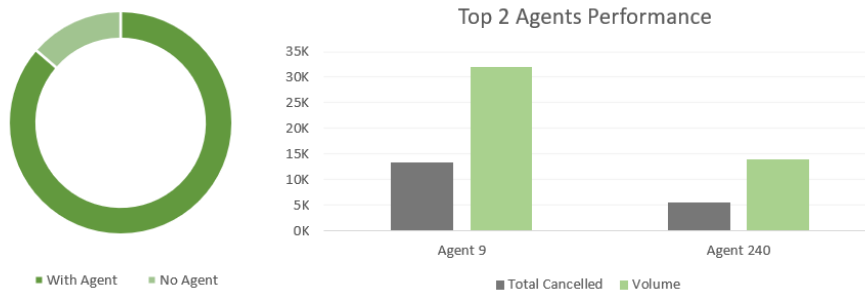
Agent Performance

Figure 6: Pie chart of bookings made with an agent vs no agent and bart chart of total bookings for top 2 agents

- Majority of Bookings (86%) are made via agents.

- Top Agent: Agent 9 with 32,000 bookings, 147 repeat guests, zero wait time.

- Runner-up: Agent 240 with 13,900 bookings, 255 repeat guests, zero wait time.
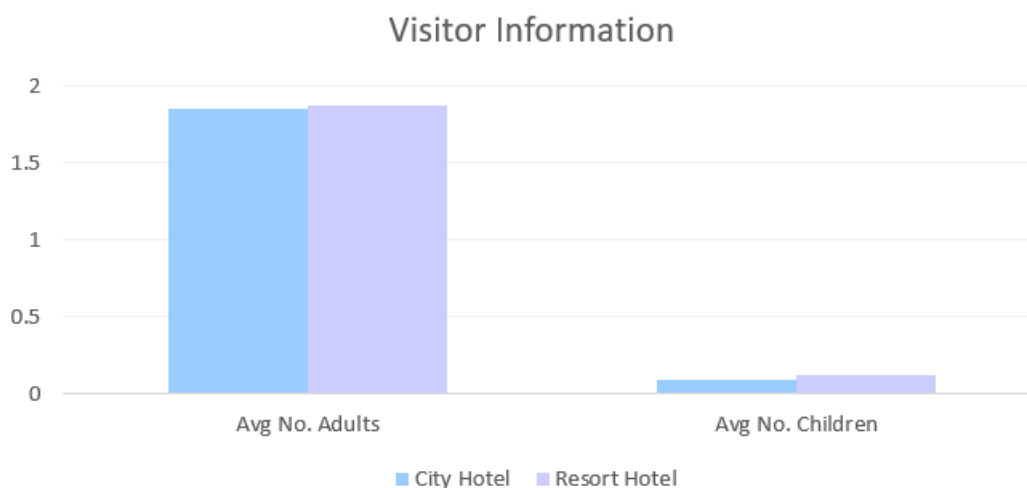
Further Summary Statistics and Insight

Figure 7: Guest Profile: Both hotels average a similar number of adults and children per booking. City Hotel typically hosts slightly more adults (suggesting a business or solo traveler focus), while Resort Hotel serves more families.
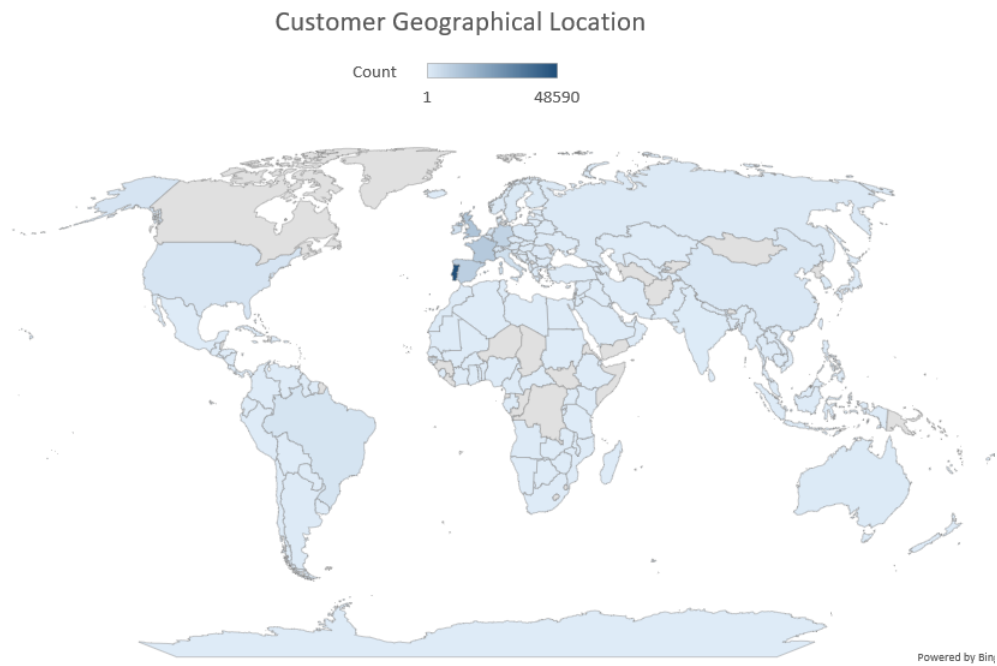
Customer Geographical Location

Count
1   48590

Figure 8: Top 3 countries that guests visit from are Portugal, Great Britain, and France

Avg Stay Duration

Figure 9: Average Stay: City Hotel is 3 nights; Resort Hotel is 4 nights.

Total Bookings   Avg Daily Rate

Figure 10: Rates: City Hotel's total number of bookings and ADR is higher than Resort Hotel

- Outliers: Maximum stays (City: 57 nights, Resort: 69 nights) are rare, potentially atypical.

Random Forest Forecasting Model to Predict Cancellations

- Analysis showed 37% of bookings are cancelled.

- A random forest is a machine learning method that builds and combines many individual decision trees to make a final prediction.

- It can captures complex relationships in data and process datasets with numerous numerical and categorical features without requiring extensive feature selection. It is also resilient to noise and outliers: ideal for real-world data.

```
print(X_train_tree_model_final.corrwith(y_train).abs().sort_values(ascending=False).head(10))
```

```
deposit_type                0.4810
agent                       0.3802
country                     0.3517
lead_time                   0.2921
arrival_date                0.2726
market_segment              0.2669
booking_id                  0.2417
total_of_special_requests   0.2339
assigned_room_type          0.2032
required_car_parking_spaces 0.1985
```

Figure 11: A look at some variables that are correlated with cancellations: deposit type and country are the 2 highest.
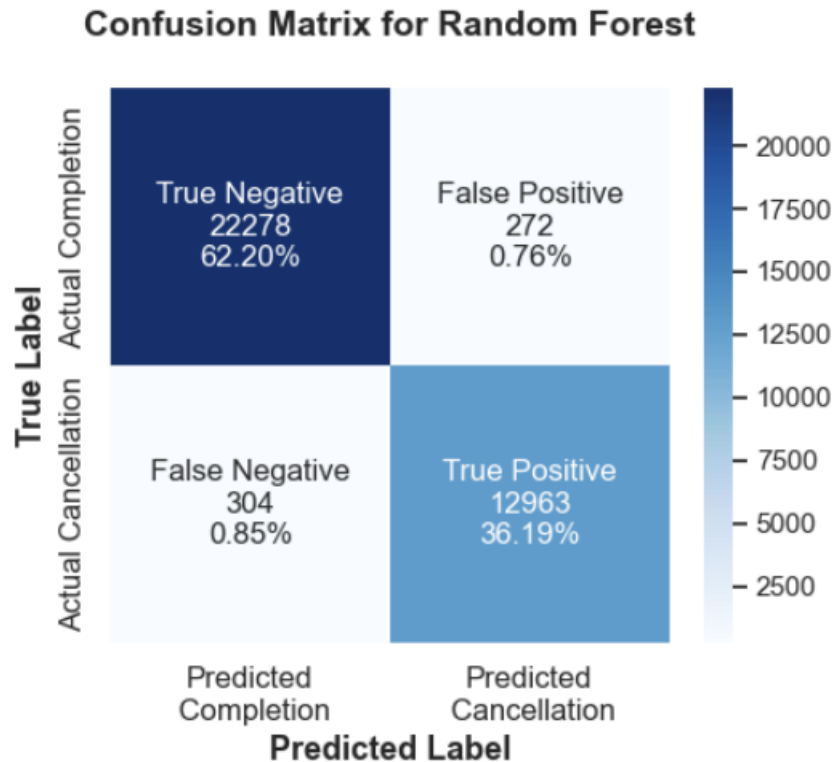
Figure 12: Confusion matrix to evaluate the Random Forest Model.

Confusion Matrix Results

- **True Negative):** 22,278 (62.2%)
  The model correctly predicted 62.2% bookings would be completed. This is the largest portion, meaning the model is good at detecting completed bookings.

- **False Positive:** 272 (0.76%)
  In 272 cases, the model wrongly predicted a booking would be canceled, but it actually was completed.

- **False Negative):** 304 (0.85%)
  In 304 instances, the model predicted the booking would be completed, but it actually was canceled.

- **True Positive:** 12,963 (36.2%)
  The model correctly identified 36.2% bookings that were canceled.

Criticisms of the model
Much more development, testing and evaluation of this model is required before it can be used for business decision making. Parameters need to be refined, and further investigation is required, it however provides a basis of how to predict cancellations going forward.

4. Conclusions and Recommendations
Conclusions

- Seasonal demand peaks in August, making winter the best time for easy bookings and better rates, or June for a summer holiday.

- Extended stays (20-25 nights) yield the best value for nightly rates for customers, or 7 days for a shorter stay (lowest median ADR for bookings under 10 days)

- City Hotel handles more volume and and travelers, while Resort Hotel is a preferred choice for families and has more special requests.

- Agent partnerships contribute the majority of bookings.

- Top source markets are Portugal, Great Britain, and France.

Recommendations

- **Promote Off-Peak Winter Bookings:** Offer targeted packages and deals for winter months to boost bookings.

- **Stay Duration Packages** Offer special packages and deals for 20-25 night bookings and 7 day bookings. Average stay length however is 3-4 days - consider deals in this length of stay too.

- **Tailored Marketing:** Focus City Hotel campaigns and deals on solo/business travelers, and families for the Resort Hotel.

- **Strengthen Agent Relationships:** Recognize and reward top-performing agents like Agent 9 and Agent 240, investigate poorer performing agents to see what can be done to improve performance and booking volumes.

5. References
This report is based on the dataset "Del-HotelBookings.csv" and its metadata "Del-HotelBookingsMetadata - V2.0.xlxs", and was processed as described in the Methods section of this report.