**HOTEL BOOKING ANALYSIS REPORT**

**BUSINESS PROBLEM**

Hotels often face operational and financial challenges due to booking cancellations, fluctuating occupancy rates, inconsistent guest behavior, and unpredictable revenue trends. These challenges affect revenue forecasting, resource planning, and customer satisfaction.

The business lacked a clear, data-driven understanding of key performance metrics, guest behavior, seasonal trends, and cancellation drivers — leading to inefficiencies in pricing strategy, customer targeting, and operational decision-making.

**DATA DESCRIPTION**

This data contains observations for a City Hotel and a Resort Hotel. Each observation represents a hotel booking between the 1st of July 2015 and 31st of August 2017, including booking that effectively arrived and booking that were cancelled.

The data comprises hotel booking records from both **City Hotels** and **Resort Hotels**, capturing a wide range of guest and reservation attributes. Key booking details include whether the reservation was cancelled (is\_canceled), how far in advance it was made (lead\_time), and the arrival date details like year, month, and day. Guest composition is recorded through the number of adults, children, and babies, while the duration of stay is broken down into weekend and weekday nights. Booking sources (market\_segment, distribution\_channel), meal preferences, and the guest’s country of origin are also included, providing insights into customer behavior and travel patterns.

The dataset further includes indicators of customer loyalty (is\_repeated\_guest), past booking behavior, room preferences (reserved\_room\_type vs. assigned\_room\_type), and financial aspects like deposit type and average daily rate (adr). Operational data such as booking changes, waiting list duration, car parking needs, and special requests are also available. The final booking outcome is recorded via reservation\_status and its update date. Additionally, personal identifiers like name, email, phone number, and credit card information are present, offering deeper guest profiling while highlighting the importance of data privacy in analysis.

**PROJECT OBJECTIVE**

1. Booking & Guest Metrics
2. Cancellation & No-Show Analysis
3. Revenue & Pricing KPIs
4. Seasonal & Time-Based Trends
5. Operational KPIs
6. Geographical Insights
7. Customer Behavior & Segmentation

**ASSUMPTIONS AND HYPOTHESIS**

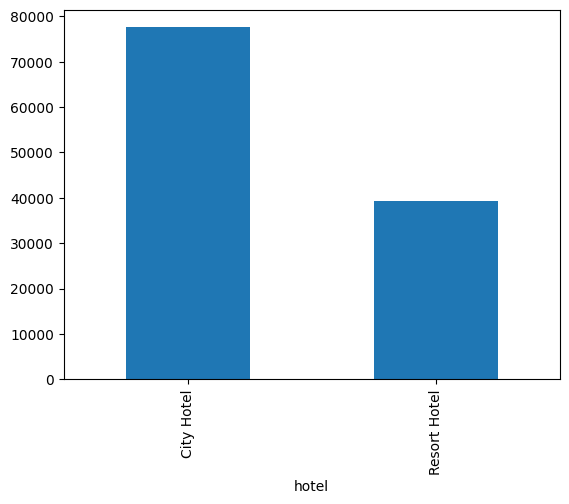
1. We assume that we have data up to 18 months (547 days) of advance booking; data beyond 18 months of advance booking are not taken into consideration.
2. We assume that a person can remain at a hotel for a maximum of 10 days on weekends and 30 days on weeknights; any further data is capped at 10 and 30, respectively.
3. We assume that each room can accommodate a maximum of six adults, five children, and five infants (babies). Anything over that is limited(capped) to those upper limits, respectively.
4. We assume that the waiting list maximum is 300 days. Greater than that is limited to that number.

**BUSINESS QUESTIONS (KPIs)**

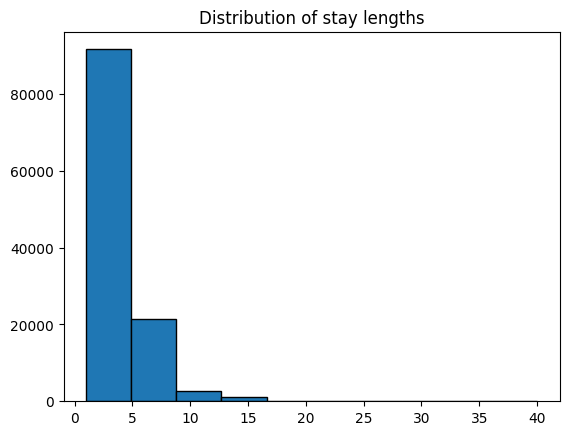
1. Total number of booking ( ie. From 1st of July 2015 and 31st of August 2017)
2. Distribution by hotel type
3. Distribution of stay lengths (total no.of days stayed in hotel)
4. How many bookings are from repeated guests?
5. What is the average number of adults, children, and babies per booking?
6. what is reservation status of each booking
7. What is the overall cancellation rate?
8. What is the cancellation rate by hotel type?
9. Which market segments have the highest cancellation rates?
10. How does lead time affect cancellation likelihood?
11. What is the average ADR (Average Daily Rate) overall and What is the total revenue generated?
12. Compare ADR by room type and hotel type.
13. What are the peak months for arrivals?
14. What are the top revenue-generating months?
15. How does ADR change over months and hotel types?
16. Which agents generate the most bookings?
17. What are the top 10 countries by number of bookings?
18. Which countries have the highest cancellation rates?
19. How do repeated guests behave vs. new ones in terms of: ADR, Lead time and Special requests
20. What are the most common room types reserved vs. assigned?
21. Weekend vs Weekday Stays

**KPI TRACKING AND INSIGHTS GENERATION**

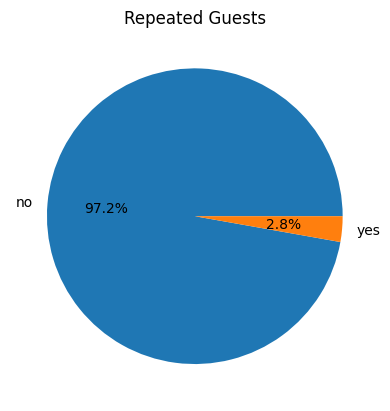
1. Total of **116,916 hotel bookings** recorded between **1st July 2015 and 31st August 2017**, covering over two years of reservation activity. This includes both city and resort hotel stays across multiple seasons and customer segments.
2. Majority of bookings were made at City Hotels, accounting for **77,610** reservations compared to **39,306** at Resort Hotels. This indicates a significantly higher demand for urban accommodations, likely driven by business travelers, short stays, and better connectivity. Resort Hotels, while fewer, may cater more to seasonal or leisure travelers.



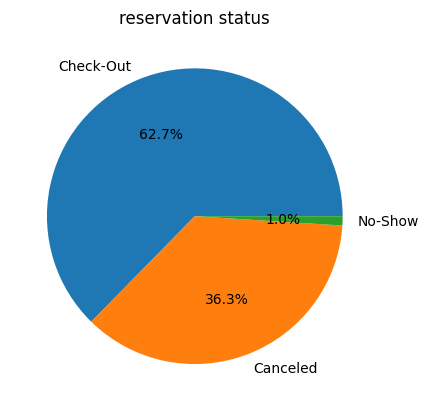
1. The distribution of stay lengths shows that most guests stayed between 2 to 4 nights, with a **median of 3 nights**. The **average stay duration** was approximately 3.46 nights, indicating short leisure or business trips. While a few bookings extended up to **40 nights**, such long stays were rare.



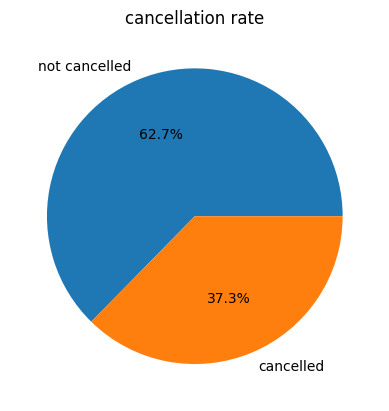
1. Approximately **2.8% of all bookings** were made by **repeated guests**, while the remaining **97.2%** were by **new guests**. This indicates that the majority of customers are first-time bookers, highlighting an opportunity to build a stronger loyalty program. Encouraging repeat bookings could lead to more consistent revenue and improved guest retention.



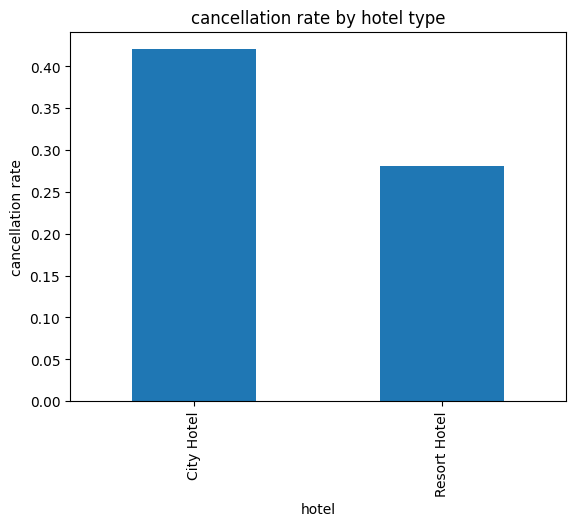
1. On average, each booking consisted of approximately **1.86 adults**, indicating that most bookings were made by either solo travelers or couples. The number of **children per booking was around 0.10**, and **babies averaged just 0.008**, suggesting that families with young children made up a relatively small portion of the guests. This highlights a predominantly adult guest profile, useful for tailoring marketing and services.
2. The **reservation status** indicates the final outcome of each hotel booking. A majority of bookings (≈62.65%) ended in a successful **Check-Out**, showing completion of stay. However, a significant portion (≈36.34%) were **Canceled**, and a small fraction (≈1%) resulted in **No-Show**, indicating guests did not arrive.



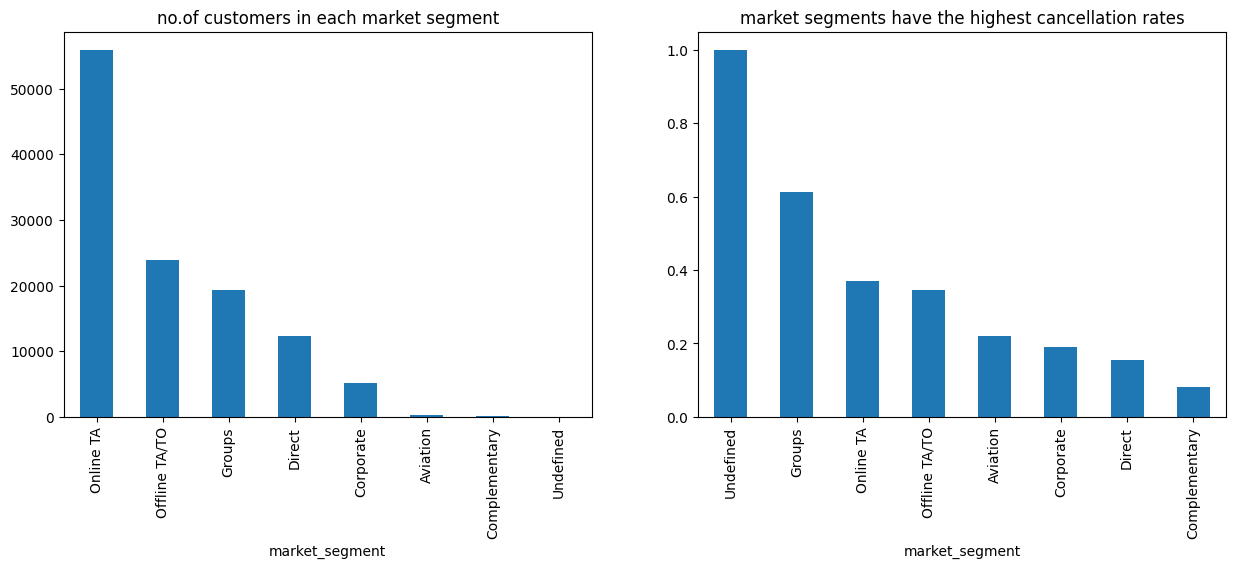
1. The overall cancellation rate in the hotel booking data is approximately **37.35%**, indicating that more than one-third of all reservations were eventually canceled. Conversely, around **62.65%** of the bookings were honored and completed as planned. This high cancellation rate highlights a potential area of concern for hotel management, particularly regarding revenue forecasting and resource allocation. (from this and 6 we can say that the one who didn’t showed up is considered as cancelled)



1. The cancellation rate for **City Hotel** is significantly higher at **42.04%**, compared to **28.08%** for **Resort Hotel**. This indicates that bookings at city locations are more likely to be canceled, possibly due to more flexible booking conditions or frequent business-related travel changes. This insight suggests the need for better cancellation policies or incentives to reduce churn in city hotels.



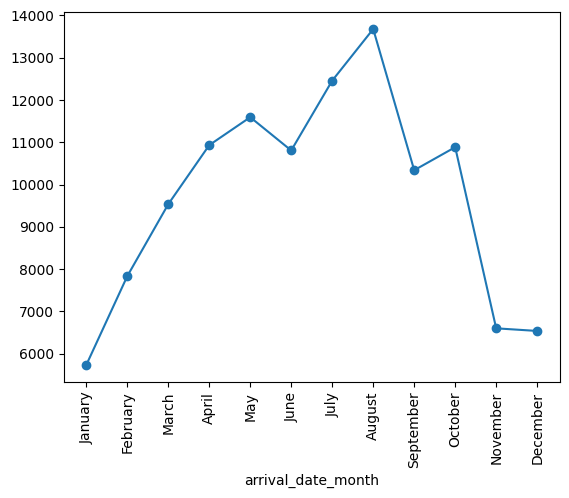
1. The majority of customers booked through the **Online TA** and **Offline TA/TO** market segments, indicating these are the most common booking sources. However, the **Groups** segment shows the highest cancellation rate at over **61%**, followed by **Online TA (Not considering that Undefined segment)**. This suggests that group and online bookings are more likely to cancel, which could impact revenue forecasting and inventory planning.



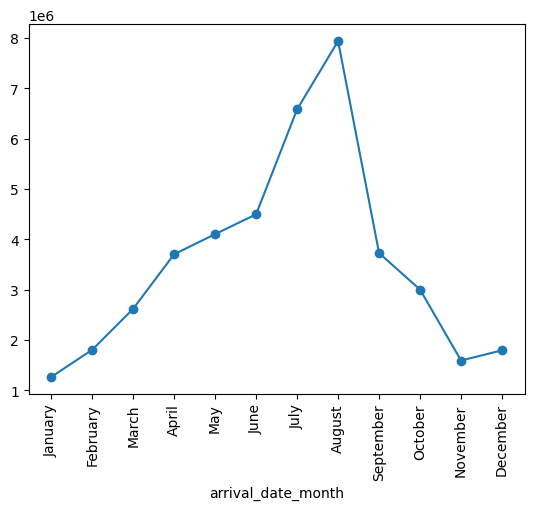
1. Guests who canceled had a significantly higher average lead time (≈142 days) compared to those who did not cancel (≈81 days). This suggests that bookings made far in advance are more prone to cancellations, possibly due to changing plans or lack of commitment.
2. The average ADR (Average Daily Rate) across all bookings was approximately **$103.63**, indicating the typical earning per occupied room per day. Based on this rate and the total number of nights stayed, the **total revenue generated** was around **$42.6 million**. This highlights the significant earning potential of the hotel during the analysis period.
3. The Average Daily Rate (ADR) varies significantly across room types and hotel categories. Resort Hotels generally have higher ADRs for premium room types (e.g., H, G, F), indicating luxury pricing, while City Hotels show higher ADRs for standard room types like A and B due to urban demand. Room types like H, I, and L are exclusive to one hotel type, suggesting differentiated offerings across hotel categories.



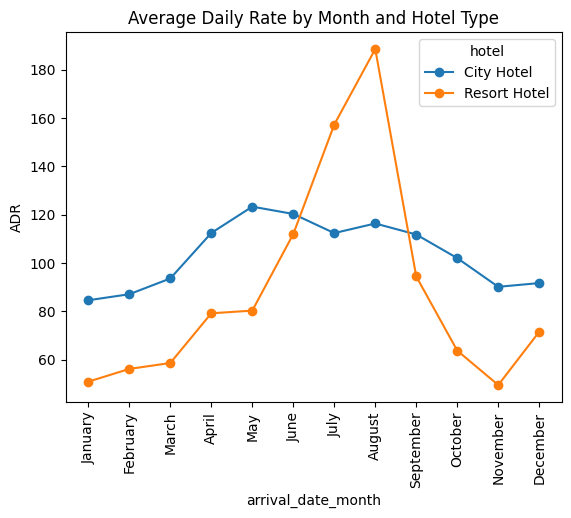
1. The peak months for hotel arrivals are **August**, **July**, and **May**, with August having the highest number of arrivals at 13,677. These months indicate strong demand likely due to summer vacations and holidays. In contrast, the winter months such as **January** and **December** show significantly lower booking volumes, suggesting clear seasonal trends in customer behavior.



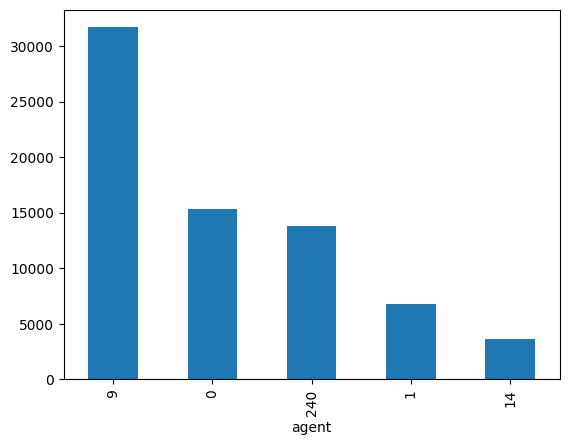
1. The top revenue-generating months were **August, July, and June**, with August alone contributing over **7.9 million** in revenue. These summer months consistently outperformed others, indicating a peak travel season with higher occupancy and room rates. This seasonal trend highlights the importance of optimizing pricing and marketing strategies during summer to maximize profitability.



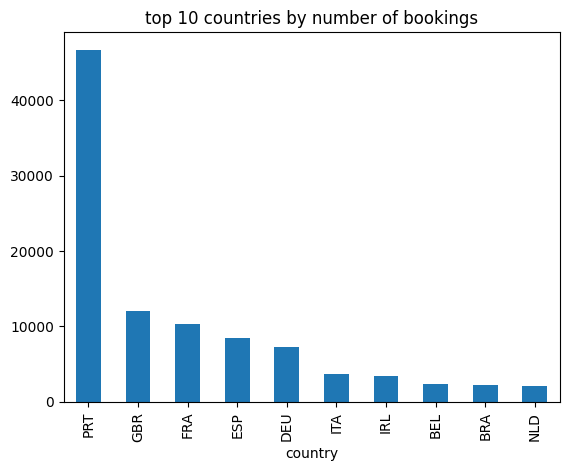
1. The Average Daily Rate (ADR) shows clear seasonal and hotel-type patterns. Resort Hotels peak sharply in August (₹188.5) and July (₹157.3), indicating high summer demand, while City Hotels maintain more stable ADRs year-round, peaking in May (₹123.3) and June (₹120.3). Overall, Resort Hotels experience greater ADR fluctuations across months compared to City Hotels.



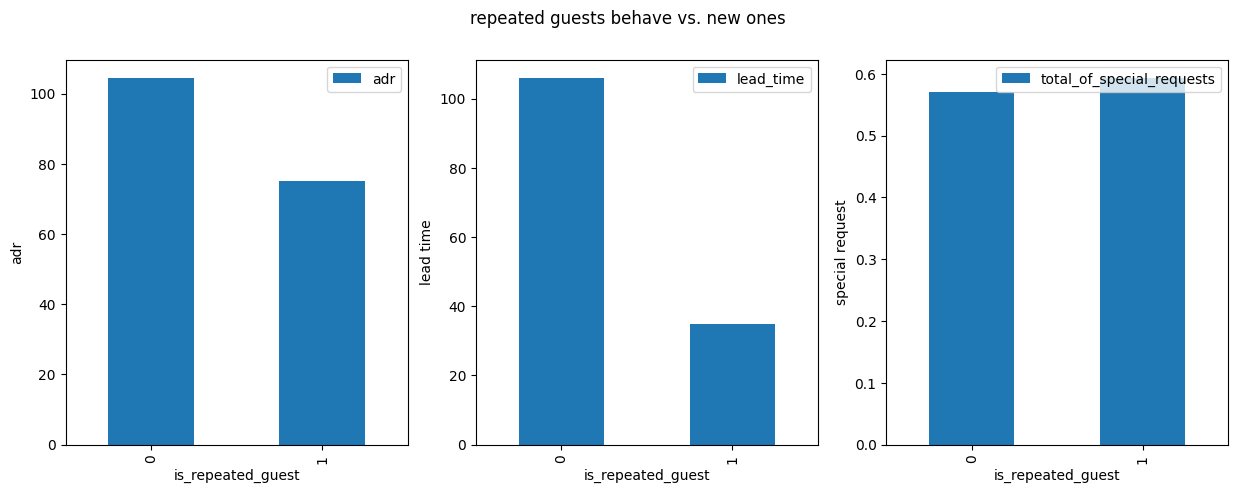
1. Agent **ID 9** generated the highest number of bookings with **31,666 reservations**, indicating strong performance or strategic partnerships. Interestingly, a large number of bookings (**15,373**) were made without a specified agent (ID 0), possibly via direct or corporate channels. Agents 240, 1, and 14 also contributed significantly, highlighting a concentration of bookings among a few key agents.



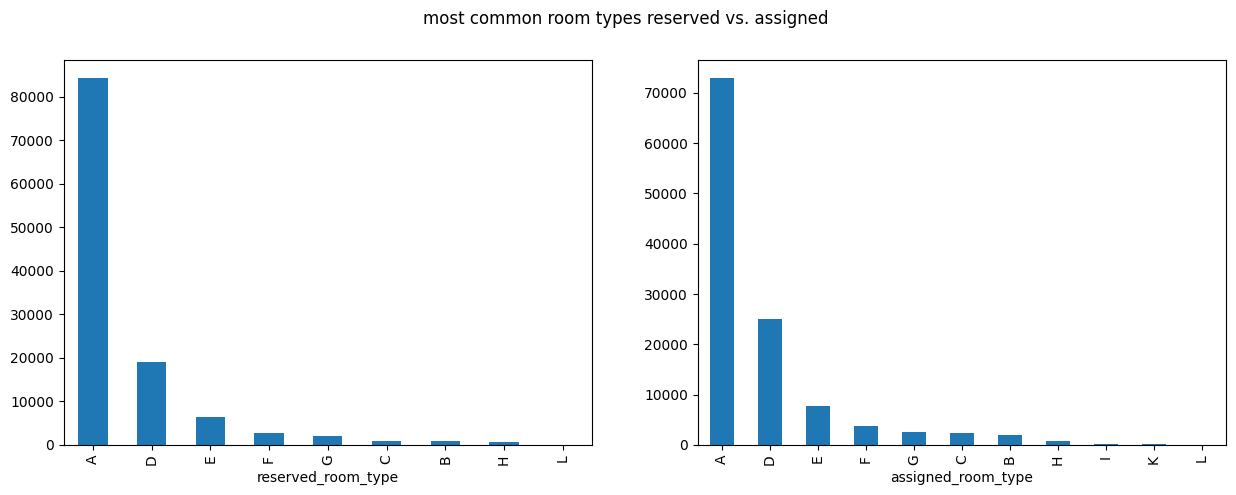
1. The top 10 countries by number of bookings reveal that **Portugal (PRT)** leads significantly with over 46,000 bookings, followed by the **UK (GBR)**, **France (FRA)**, and **Spain (ESP)**. These countries represent the hotel’s major customer base, likely due to geographic proximity, tourism trends, and ease of travel. Understanding this distribution helps in designing targeted marketing and service strategies for key international markets.



1. The analysis revealed that several countries, including **Benin (BEN), Fiji (FJI), and Hong Kong (HKG)**, had the highest cancellation rates, with many at or near **100%**. These high rates are likely due to low booking volumes from these regions, making them more sensitive to cancellations. In contrast, countries with a larger number of bookings like **Portugal (PRT), Brazil (BRA), and the UK (GBR)** had moderate cancellation rates, reflecting more stable guest behavior.
2. Repeated guests tend to book with a significantly shorter lead time (35 days) compared to new guests (106 days), indicating more spontaneous bookings. Their average daily rate (ADR) is also lower (₹75 vs ₹104), possibly due to loyalty benefits or negotiated rates. Interestingly, repeated guests make slightly more special requests, reflecting familiarity with hotel services and personalized preferences.



1. The most commonly reserved room type was **A**, followed by **D** and **E**, while the most frequently assigned room types were also **A**, **D**, and **E**, but with notable variation in counts. A total of **12.02%** of bookings had a mismatch between the reserved and assigned room types, indicating that guests were not always given the room they initially selected. This could be due to overbooking, upgrades, or operational constraints in room availability.



1. Most guests stayed **0 or 1 night during weekends**, indicating that weekend stays are generally shorter or skipped altogether. In contrast, **weekday stays are longer**, with the majority staying **2 to 3 nights**, suggesting that the hotel is often used for business or midweek travel. This pattern can guide targeted promotions for weekend stays to boost occupancy.

