

**HOTEL BOOKING ANALYSIS**

Sujata Kumari

**PROJECT OVERVIEW**

**Problem Statement:**

Analyze booking patterns, guest preferences, and factors influencing cancellations. Use SQL and Excel to identify trends in booking sources and revenue. Develop a Power BI dashboard for tracking booking trends and optimizing hotel operations.

**Dataset Description**

The given dataset is a comprehensive collection of hotel booking information that can be used for analysis and insights in various aspects of hotel industry.

It contains a wide range of data related to hotel reservations, including booking details, guest information, booking source, room details, and reservation status.

It aims to explore and visualize trends, patterns, and key factors affecting hotel bookings and cancellations.

**PROCESS**

**1. Data Acquisition:**

Obtain the relevant dataset for hotel bookings, including details on bookings, guests, room preferences, meal plans, special requests, parking, and cancellations. Ensure comprehensive coverage of key attributes such as booking sources, market segments, and stay durations.

**2. Data Transformation and Enhancement:**

* **Data Cleaning:** Address inconsistencies, missing values, and ambiguous fields to ensure data quality.
* **Enhancement:** Introduce new calculated fields, such as cancellation rates, booking lead time, average daily rates (ADR), and guest retention metrics.
* **Segmentation:** Categorize data into meaningful segments, e.g., guest types (repeated or new), hotel types (City vs. Resort), and booking channels.

**3. Connecting with Analytical Tools:**

* Integrate the dataset with **Power BI**, **Excel**, and **MySQL Workbench** for comprehensive analysis and visualization.
* Establish seamless connections for automated data updates where necessary.

**4. Problem Statement Solution in Power BI:**

* Utilize Power BI to address specific problem statements, such as:
  + Identifying cancellation trends and factors.
  + Analyzing room preferences and guest behaviour over time.
  + Understanding seasonality and peak booking periods.
* Employ dynamic visualizations (e.g., bar charts, scatter plots, heat maps) for clear insights.

**5. Exploratory Data Analysis (EDA):**

Perform EDA in **Excel** and **SQL Workbench** to:

* Investigate the distribution of key metrics such as ADR, parking needs, and stay durations.
* Identify patterns like high-cancellation channels or long-stay guests.
* Examine guest behaviour and retention metrics to refine operational strategies.

**6. Creation of Visual and Insightful PowerPoint:**

* Develop a PowerPoint presentation encapsulating:
  + **Objectives:** Key goals like improving booking efficiency or reducing cancellations.
  + **Analysis:** Visuals and insights for problem statements such as meal preferences or booking channel performance.
  + **Conclusions:** Summarized actionable recommendations based on analysis.
* Include a **comparative analysis** section highlighting the differences between City Hotels and Resort Hotels.

**7. Detailed Documentation:**

* Compile a report covering:
  + **Data Collection & Cleaning:** Steps taken to prepare the dataset.
  + **Tools Integration:** Explanation of connections with Power BI, Excel, and SQL Workbench.
  + **Problem Statements & Insights:** Address each analysis question with visualizations and interpretations.
  + **Future Recommendations:** Suggestions for improving data tracking and operational decisions.

**OBJECTIVE**

The hospitality industry thrives on understanding and catering to diverse guest preferences while optimizing operational efficiencies. Hotel bookings, being a critical component of this industry, are influenced by various factors such as guest demographics, seasonal trends, market segmentation, and operational challenges like cancellations and special requests. Analyzing these dynamics is essential for hotels to remain competitive and enhance their services.

This project focuses on a comprehensive examination of hotel booking data to uncover trends, patterns, and key drivers of guest behaviour and operational performance. By delving into aspects like booking sources, reservation statuses, room preferences, meal plans, and guest characteristics, the analysis aims to offer actionable insights for improving hotel strategies.

The primary goal is to identify opportunities for enhancing customer satisfaction, reducing cancellations, and optimizing revenue streams. Additionally, the project seeks to explore the interplay between guest preferences, seasonal demand, and operational requirements, providing a well-rounded understanding of the factors influencing hotel bookings.

Through this analysis, hotels can refine their marketing strategies, tailor their services to meet diverse guest needs, and implement data-driven solutions to increase efficiency and profitability. The insights derived will be instrumental in shaping a more responsive and guest-centric approach to hospitality management.

**Significance of Hotel Booking Analysis**

Hotel booking analysis is a critical tool in the hospitality industry, offering valuable insights to hoteliers, marketers, and policymakers. It aids in understanding guest behaviours, optimizing resources, and creating strategies to enhance guest satisfaction and profitability. Here's an exploration of its significance:

**1. Enhancing Guest Experience:**  
By analyzing booking patterns, preferences, and special requests, hotels can tailor their offerings to meet the specific needs of guests. For example, understanding meal plan preferences or room type demands helps in providing personalized services, leading to higher guest satisfaction and retention.

**2. Revenue Optimization:**  
Hotel booking analysis allows businesses to evaluate key metrics such as Average Daily Rate (ADR), occupancy rates, and cancellation patterns. This data informs dynamic pricing strategies, ensuring competitive pricing during peak seasons and maximizing revenue during low-demand periods.

**3. Operational Efficiency:**  
Understanding trends like lead times, booking channels, and special requests enables hotels to manage their resources more effectively. For instance, forecasting parking space demand or special request frequency ensures that hotels allocate staff and amenities efficiently.

**4. Identifying Market Segment Opportunities:**  
Insights into booking channels and market segments highlight where the majority of bookings originate. By identifying the most effective channels, such as Online Travel Agents or direct bookings, hotels can focus their marketing efforts and allocate budgets strategically.

**5. Mitigating Cancellations:**  
Analyzing cancellation trends provides actionable insights into the factors driving booking withdrawals, such as lead times, guest demographics, or external conditions. Hotels can then implement strategies like flexible policies or targeted retention efforts to reduce cancellations and improve booking stability.

**6. Strategic Decision-Making:**  
Hotel booking analysis helps decision-makers understand long-term trends, including seasonal patterns and demographic shifts. This information supports strategic planning, from launching promotional campaigns during off-peak periods to investing in amenities that attract high-value guests.

**7. Competitive Benchmarking:**  
By comparing performance across distribution channels, room types, and guest preferences, hotels can benchmark themselves against competitors. Understanding what drives higher bookings or guest loyalty for competitors helps in refining their service offerings.

**8. Sustainability and Resource Management:**  
Data-driven insights into guest behaviour enable hotels to make environmentally conscious decisions, such as optimizing meal plans or energy usage based on occupancy patterns. This contributes to cost savings and sustainable operations.

In conclusion, hotel booking analysis empowers stakeholders in the hospitality sector to make informed decisions that enhance guest satisfaction, streamline operations, and boost profitability. By leveraging insights from this analysis, hotels can remain agile and competitive in a dynamic market, ultimately driving long-term success.

**Data Dictionary**

Table – booking\_details

Fields:

***hotel :*** the type of hotel (categorical),

(categories: Resort Hotel, city Hotel)

***is\_canceled :*** Whether the booking was cancelled or not (binary),

{Values: 1 (cancelled) 0(not cancelled)}

***Lead\_time :*** The number of days between booking and arrival (numerical)

***arrived\_date\_year :*** The year of arrival (numerical)

***arrival\_date\_ month :*** The month of arrival (Categorical),

(Categories: January, February, ..., Dec.)

***arrival\_date\_week\_number :*** The week number of the arrival date (numerical)

***arrival\_date\_day\_of\_month :*** The day of the month arrival (numerical)

***Stay\_in\_weekend\_nights*** ***:*** The number of weekend nights (numerical)

***Stay\_in\_week\_nights :*** The number of weekdays nights (numerical)

***country :*** The country of origin (categorical)

***Booking\_id :*** unique id of guest (numerical)

Table – guest\_info

Fields:

***adults :*** Number of adults (numerical)

***children :*** Number of children (numerical)

***babies :*** Num. of babies (numerical)

Table – meal\_and\_stay\_details

Fields:

***meal :*** Type of meal booked (Categorical),

(Categories: BB (Bed & Breakfast), HB (Half Board), FB(Full Board),

SC (self catering))

***adr :*** Average Daily Rate (numerical)

***required\_car\_Parking\_spaces :*** No. of car parking Spaces required (numerical)

***total\_of\_special\_requests :*** Total no. Special request made by guest(numerical)

Table – Booking\_source\_and\_History

Fields:

***country :*** The country of origin (categorical)

***market segment :*** Market segment designation (categorical)

***distribution\_channel :*** Booking distribution channel (categorical)

***is\_repeated\_guest :*** whether the guest is repeated guest (binary) ,

{Values : O (not), 1 (repeated guest) }

***previous\_cancellations :*** No. of previous booking cancellation by the guest (num)

***Previous\_booking\_not\_canceled :*** No. of previous booking not cancelled by the

guest (numerical)

***deposit\_type :*** Type of deposit made (categorical),

(Categories: No deposit, Non-Refund, Refundable)

***agent :*** ID of the booking agent (numerical)

***Company :*** ID of the company booking (numerical)

***days\_in\_waiting\_list :*** Number of days the booking was on the waiting list

(numerical)

***customer\_type :*** Type of customer (categorical)

(categories: Transient, Transient\_Party, contract, Group.)

Table – Room\_Details

Fields:

***reserved\_room\_type :*** Type of room reserved (categorical)

***assigned\_room\_type :*** Type of room assigned (categorical)

***booking\_changes :*** Number of booking changes made by the guest (numerical)

Table – Reservation\_Status

Fields:

***reservation\_Status :*** Reservation Last Status (categorical),

(Categories: Cancelled, check-out, No-show)

***reservation\_status\_date :*** Date at which the reservation Status was

recorded (date time)

Table – Country

Fields:

***country\_code :*** code of country (numerical),

***country\_name :*** Name of country (categorical)

Table – Distribution\_Channel

Fields:

***distribution\_channel :*** Distribution Channel name (categorical) (e.g. Online, Direct, Corporate)

***distribution\_channel\_id :*** corresponding distribution channel id (numerical)

Table – Market Segment

Fields:

***market\_segment :***  Market segment name (categorical),

(Categories: Direct, corporate, groups etc)

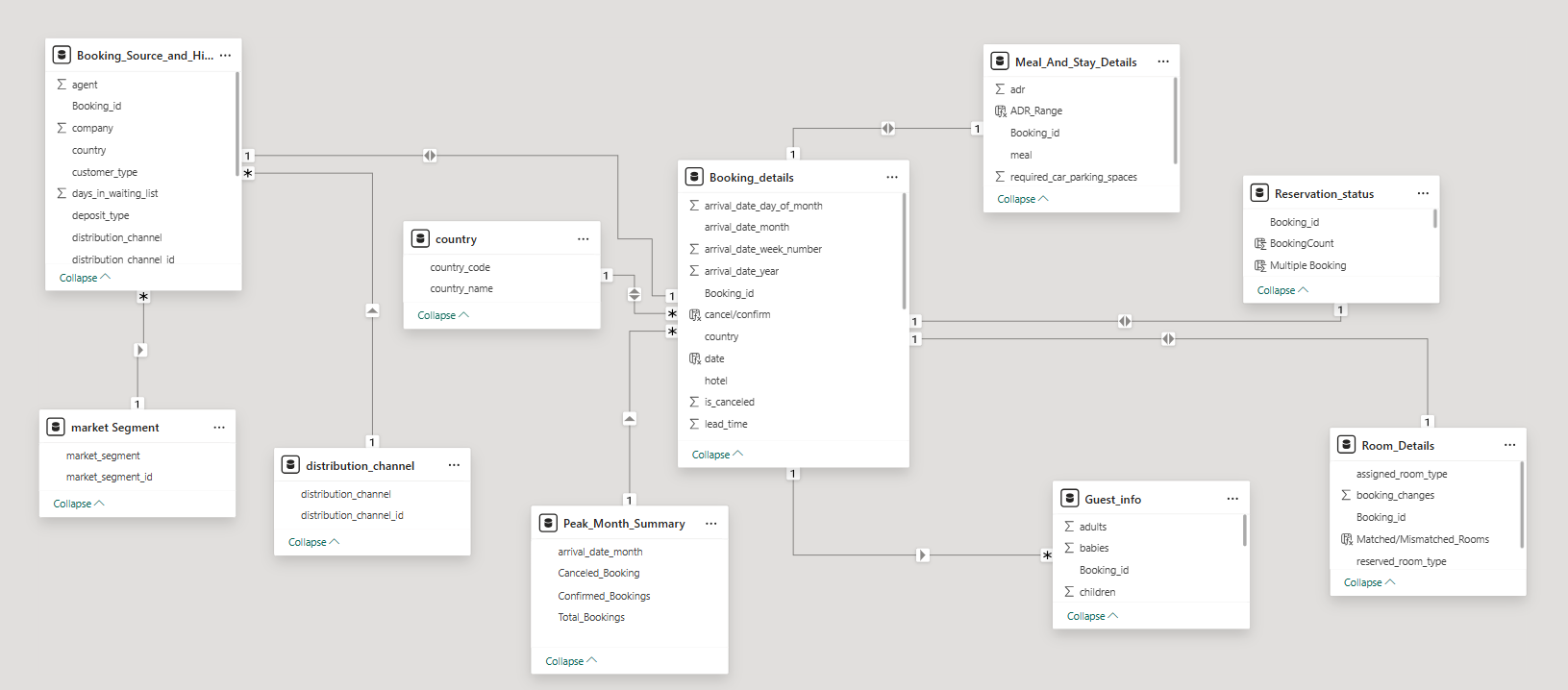
***market\_segment\_id :*** Market segment id (numerical)

Table – Peak Month Summary

Fields:

***arrival\_date\_month , Total\_Bookings, Confirmed\_Booking, Canceled\_Booking***

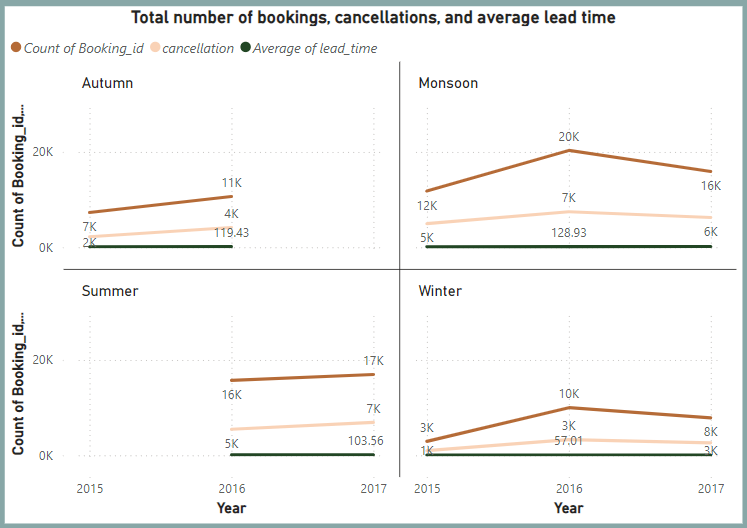
**ER DIAGRAM**



**POWER BI**

**PROBLEM STATEMENTS**

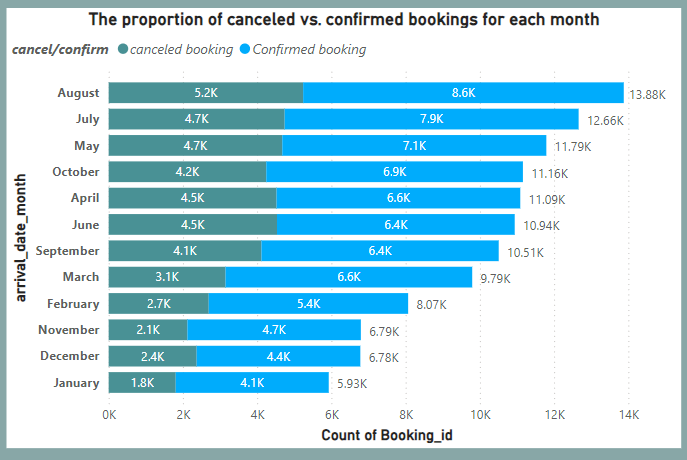
**Que. 1. Visualize booking trends over the years, including the number of bookings, cancellations, and average lead time. Identify seasonality patterns.**

****

Insights on Seasons and Years:

1. Best Year and Season:
   * 2016 was the most successful year with peak bookings in all seasons, especially Monsoon (20k bookings) and strong growth in Autumn and Winter.
   * Monsoon was the best season overall, consistently leading in bookings across years.
2. Worst Year and Season:
   * 2015 had the lowest bookings, particularly in Winter (2.9k) and Autumn (7k).
   * Winter was the weakest season, with low demand and high cancellation rates every year.
3. Overall Insight:
   * Focus on Monsoon and Summer for high-demand strategies, while targeted promotions and discounts are needed for Winter and low-performing years like 2015.

**Que. 2. Analyze monthly booking patterns to identify peak months and optimize marketing strategies.**

****

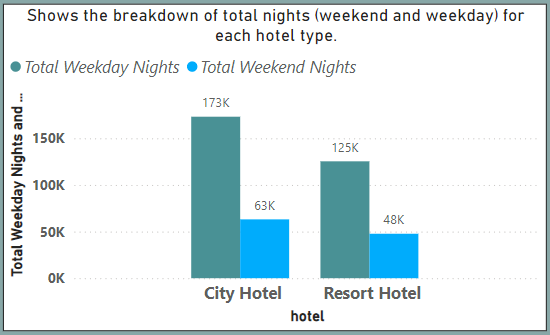
Certain months, such as August, July show significantly high confirmed bookings, indicating peak travel demand but also high cancellations, likely due to flexible policies or changing plans. In contrast, off-peak months such as December and January see lower bookings and cancellations, presenting opportunities for targeted promotions.

By analyzing monthly booking patterns, we identify that:

1. Peak Months (e.g., August, July) should focus on retention strategies to minimize cancellations. Offering discounts for early confirmations or more stringent cancellation policies could help reduce lost revenue.
2. Off-Peak Months (e.g., February and March) require marketing campaigns to attract more guests. Promotional offers, packages, or collaboration with local events could be effective.

In conclusion, peak months offer opportunities to maximize revenue through efficient booking management, while off-peak months require creative marketing efforts to drive demand.

**Que 3. Compare stays in weekend nights and weekday nights to determine preferences and variations by hotel type.**

****

The bar chart highlights the differences in guest preferences for weekday and weekend stays across City Hotels and Resort Hotels. City Hotels exhibit a strong preference for weekday stays, with 173K nights compared to just 63K nights on weekends. This indicates that City Hotels primarily cater to business travellers or work-related stays during the week.

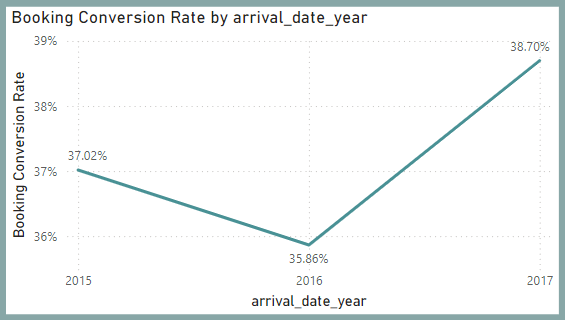
On the other hand, Resort Hotels show a smaller but steadier demand, with 125K weekday nights and 48K weekend nights. This suggests that Resort Hotels are more leisure-oriented, attracting guests seeking relaxation throughout the week, though weekday stays are still more common.

City Hotel shows a sharp drop in stays from weekdays to weekends, suggesting a dominant business-oriented clientele.

Resort Hotel shows a less drastic drop, indicating a relatively stable demand throughout the week, likely from leisure travellers.

Overall, City Hotels outperform Resort Hotels in total stays, particularly during weekdays. However, the smaller variation in Resort Hotel stays between weekdays and weekends reflects a more balanced demand pattern. These trends emphasize the distinct customer bases and purposes of stay for the two hotel types.

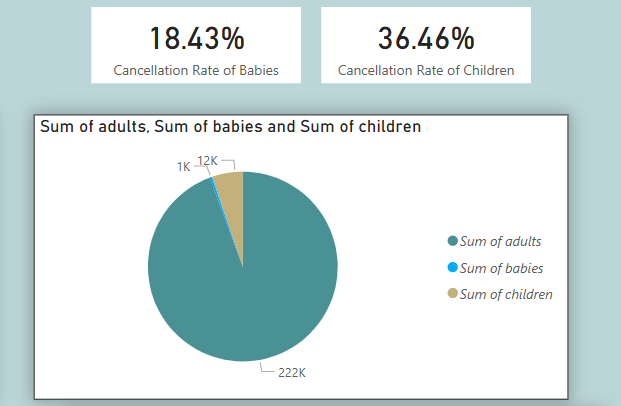
**Que 4. Calculate and visualize the booking conversion rate (canceled bookings to total bookings) over time.**

****

Insight: Booking Conversion Rate by Year

1. Trend Overview: The booking conversion rate shows changes over the years:
   * It dropped from 2015 to 2016, meaning fewer bookings were confirmed compared to inquiries or availability.
   * It improved in 2017, showing a recovery and better booking performance.
2. Possible Reasons for Changes:
   * The drop in 2016 could be due to more cancellations, competition, or issues with pricing or marketing.
   * The recovery in 2017 might mean the business improved its strategies, like offering better deals or improving the booking process.
3. What It Means:
   * Businesses can use this data to understand what worked well and what didn’t, helping them plan better strategies to keep the booking rate steady or growing.

**Que 5. Visualize the distribution of adults, children, and babies in bookings. Explore the impact of children and babies on cancellation rates.**

****

1️.Booking Composition (Pie Chart)

* Adults dominate bookings, accounting for the vast majority.
* Children make up a small but notable segment, while babies represent a tiny fraction of overall bookings.

2️. Cancellation Rates

* The cancellation rate for children is nearly double that of babies.
* Despite their lower numbers, children’s bookings exhibit a significant cancellation risk.

Impact and Insights

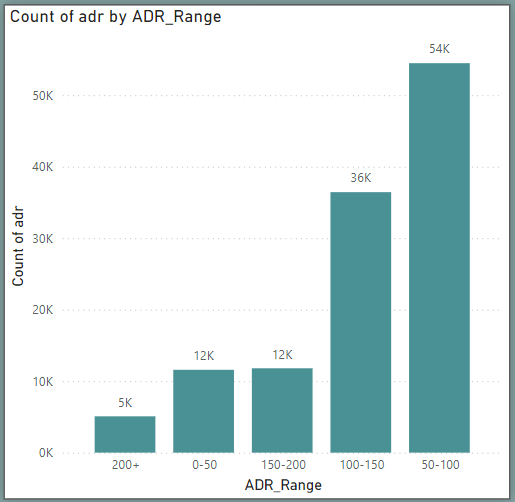
Children’s Impact on Cancellations

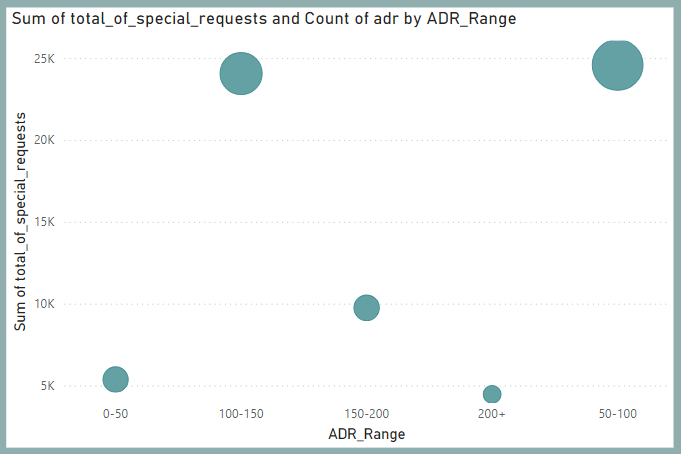
* With a high cancellation rate (36.46%), even though children represent fewer bookings than adults, their cancellations can disproportionately affect planning (e.g. room allocations, meal plans, activities)

Babies’ Impact on Cancellations

* Babies have a lower cancellation rate (18.43%). Since their bookings are quite rare (~1K), their overall impact on cancellations is relatively minor.

**Que 6. Analyze the distribution of Average Daily Rate (ADR) and identify correlations with the number of special requests made by guests.**



****

The chart shows how the ADR Range (Average Daily Rate) is related to the Total Special Requests made by guests. Guests in the 50–100 ADR Range made the most bookings and also the highest number of special requests, making it the most popular price category. This means that guests in this range are more likely to request extra services, which could reflect higher engagement and specific needs.

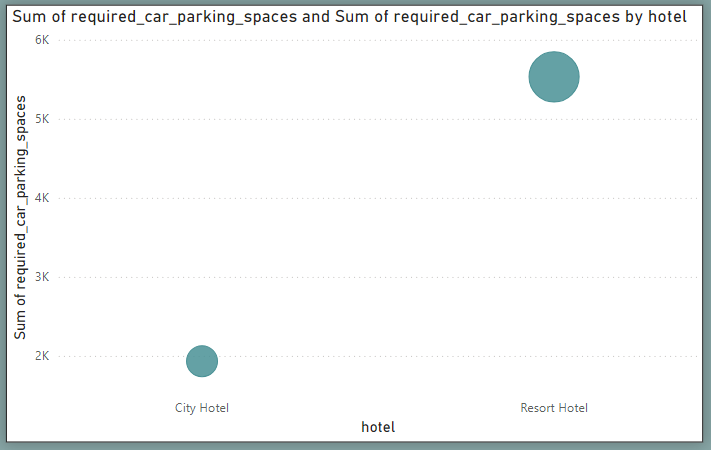
The 0–50 ADR Range had fewer special requests and bookings, suggesting guests in this budget-friendly range might have simpler needs or are less interested in extra services. The 100–150 ADR Range showed moderate levels of both bookings and special requests, indicating balanced guest activity. On the other hand, the 200+ ADR Range had the fewest bookings and special requests, suggesting it caters to a small, luxury-focused group of guests who don’t typically ask for extras.

Recommendations

1. Focus on the 50–100 ADR Range:
   * Improve processes to handle the high volume of special requests efficiently and ensure guest satisfaction in this key segment.
2. Boost the 200+ ADR Range:
   * Create special offers, luxury experiences, or personalized packages to attract more high-paying guests.
3. Explore the 0–50 ADR Range:
   * Look for opportunities to upsell additional services, even for budget-friendly guests.
4. Enhance the 100–150 ADR Range:
   * Focus on understanding what these guests value most to provide tailored services and ensure a good experience.

This analysis helps identify which price ranges are performing well and where there is room for improvement to maximize revenue and guest satisfaction.

**Que 7. Visualize the relationship between the number of required car parking spaces and booking type (Resort Hotel vs City Hotel).**

****

Here's a breakdown of the insights:

1. Bubble Size: The size of the bubbles indicates the total number of required car parking spaces for each hotel type. Larger bubbles represent a higher number of parking spaces.

- The Resort Hotel has a much larger bubble compared to the City Hotel, indicating that Resort Hotels require significantly more car parking spaces.

2. Vertical Axis (Sum of required car parking spaces): This axis represents the total number of required parking spaces.

- The City Hotel is closer to the 2,000–3,000 range for required parking spaces.

- The Resort Hotel is closer to the 5,000–6,000 range, implying a higher demand for parking at Resort Hotels.

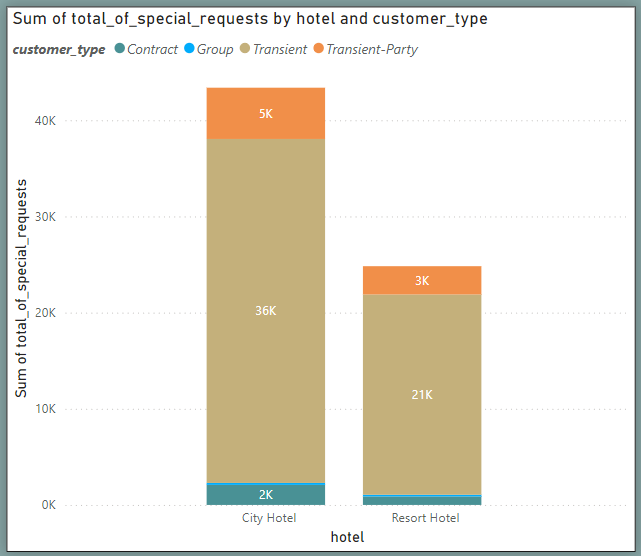
3. Hotel Type: The horizontal axis indicates the hotel types, with City Hotel on the left and Resort Hotel on the right. The chart clearly separates the two categories and shows that Resort Hotels consistently require more parking spaces than City Hotels.

Conclusion:

- \*Resort Hotels\* require a significantly higher number of car parking spaces compared to City Hotels.

- This difference might be due to factors like location (resorts being in more isolated areas), the need for visitors to bring cars, and the longer stays often associated with resorts.

**Que 8. Use power BI to explore how the total number of special requests made by guests varies by hotel type and customer type (Transient, Group).**

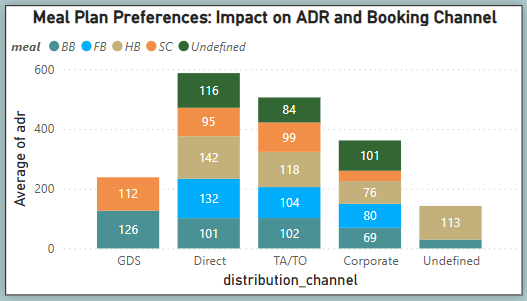
****

The visual reveals clear differences in the total special requests made by guests across hotel and customer types. City Hotels receive significantly more requests (43,000) than Resort Hotels (26,000), suggesting urban travellers have more specific needs, likely due to the diverse guest profiles and fast-paced city environments.

Transient guests dominate special requests in both hotel types, with City Hotels receiving 36,000 and Resort Hotels 21,000 requests. This indicates individual travellers often seek more personalized services. Group bookings contribute fewer requests—5,000 in City Hotels and 3,000 in Resort Hotels—likely due to standardized arrangements. Contract customers make the fewest requests, with minimal numbers in City Hotels (2,000) and negligible in Resort Hotels.

Overall, City Hotels see higher demand for customization, especially from Transient guests, while Resort Hotels face fewer requests. These insights can guide hotels in tailoring services and optimizing resources to meet guest preferences effectively.

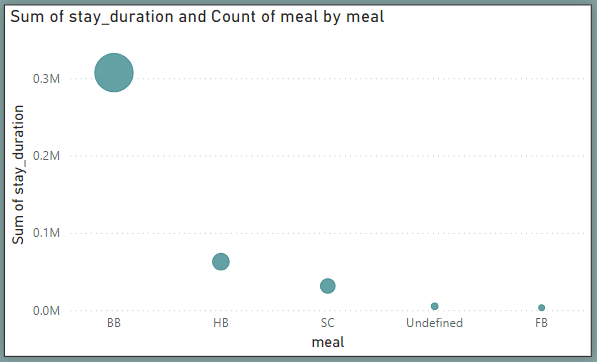
**Que 9. Explore meal plans and their impact on Average Daily Rates (ADR). Analyze meal plan preferences and their association with booking channels.**



Insight:  
The impact of meal plan preferences on the Average Daily Rate (ADR) varies across booking channels. The Direct and TA/TO (Travel Agent/Tour Operator) channels have the highest ADRs overall, with a noticeable preference for the HB (Half Board) meal plan, as indicated by higher segments in the stacked bars. GDS (Global Distribution System) bookings tend to have lower ADRs, with a relatively higher proportion of BB (Bed & Breakfast). The Corporate channel shows a more balanced ADR distribution, but with a higher proportion of FB (Full Board) and BB meal plans. Undefined channels have moderate ADRs, with no clear meal plan dominance.

Recommendation:  
To maximize ADR, hotels might prioritize marketing the HB meal plan through Direct and TA/TO channels, as these show the highest potential for revenue generation. Additionally, encouraging meal plan upgrades in channels like Corporate could boost ADR performance.

**Que 10. Analyze how meal plans correlate with stay duration and investigate any differences in stay lengths based on meal plans.**

****

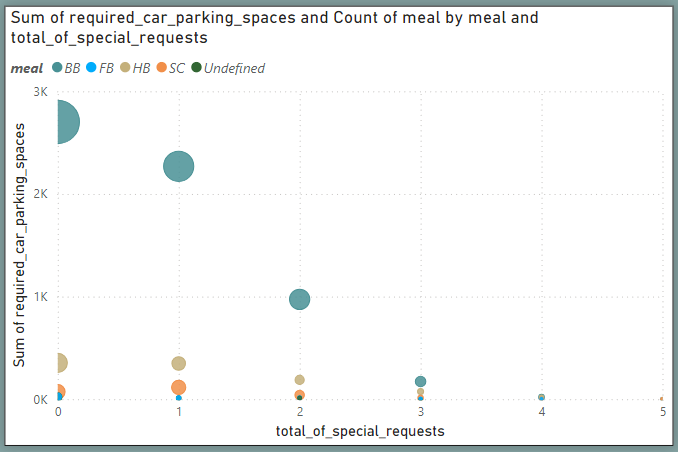
The analysis reveals that the BB (Bed & Breakfast) meal plan is the most popular among guests, with the highest cumulative stay duration (0.3M) and the largest number of bookings. This suggests that BB strikes a balance between cost and convenience, making it the preferred choice, especially for extended stays.

In contrast, HB (Half Board) and SC (Self-Catering) plans show significantly lower stay durations and booking counts, with HB performing slightly better.

Meal plans like Undefined and FB (Full Board) have minimal contributions, indicating they are rarely selected by guests.

Overall, guests on the BB meal plan tend to stay longer, highlighting its appeal for extended visits. Hotels should focus on optimizing their BB offerings while exploring strategies to make HB, SC, and other less popular plans more attractive, particularly for guests seeking longer stays.

**Que 11. Correlate parking requirements and special requests with different meal plans. Determine if certain meal plans result in more requests or parking needs.**

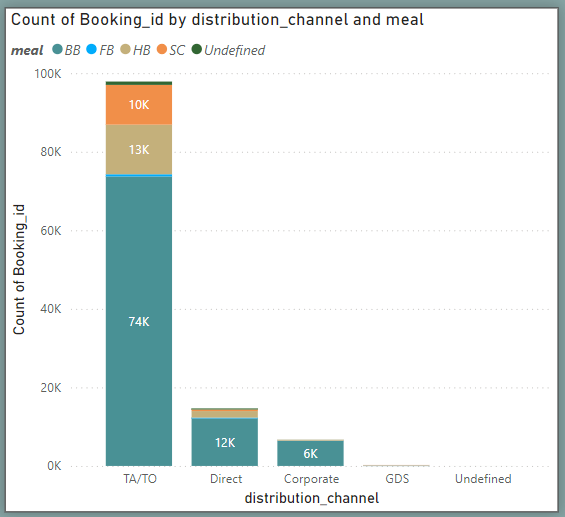
****

Insights:

1. BB Meal Plan Leads in Parking and Requests: Guests choosing the BB (Bed & Breakfast) meal plan have the highest demand for car parking spaces, reaching up to 3,000, and also tend to make more special requests, typically ranging between 1 and 3. This suggests BB guests are more likely to request personalized services and require additional facilities.
2. Lower Demand for Other Meal Plans: Guests opting for FB (Full Board), HB (Half Board), SC (Self-Catering), and Undefined meal plans show significantly fewer parking needs and special requests. These meal plans cater to a smaller or less service-intensive group of guests, indicating different preferences or requirements.
3. Correlation Between Parking and Requests: There is a clear trend where higher parking needs often align with 1-2 special requests, especially for BB guests. This reflects the broader popularity of the BB meal plan and the expectations of those guests for a more tailored hotel experience.

To enhance guest satisfaction, expand parking facilities or implement reservation systems for high-demand BB guests. Streamline special request handling through digital tools and staff training, and promote underutilized meal plans like FB, HB, and SC with tailored offers.

**Que 12. Explore how meal plans are distributed across various booking channels. Analyze if certain channels are associated with specific meal plans.**

****

The chart shows that TA/TO (Travel Agent/Tour Operator) is the largest distribution channel with 74K bookings, predominantly offering the BB (Bed & Breakfast) meal plan.

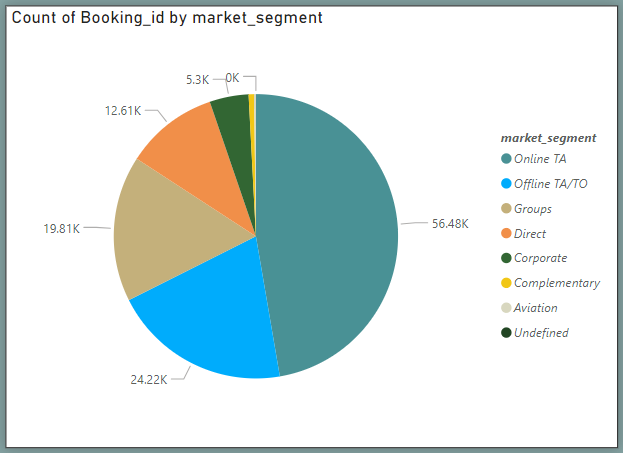
Other meal plans like FB (Full Board), HB (Half Board), and SC (Self-Catering) are present but in much smaller proportions.

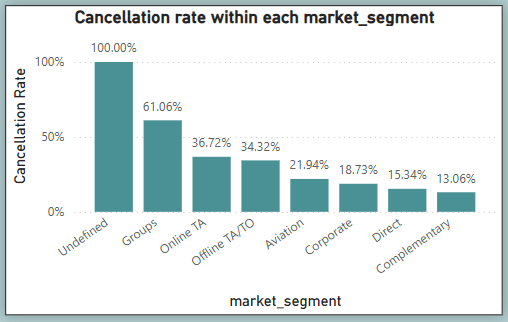
Direct and Corporate channels have fewer bookings (12K and 6K, respectively) and also favour BB, though with slightly more variety in meal plan choices compared to TA/TO. The GDS channel and Undefined category have negligible or no representation.

Overall, BB is the most preferred meal plan across all channels, especially in the TA/TO segment, while Direct and Corporate bookings show some diversity in meal plans.

These insights suggest that hotels and booking platforms could focus on offering more customized meal plans for direct and corporate bookings, while keeping the BB option prominent for TA/TO customers.

**Que 13. Visualize booking distribution across different market segments and analyze cancellation rates within each segment.**

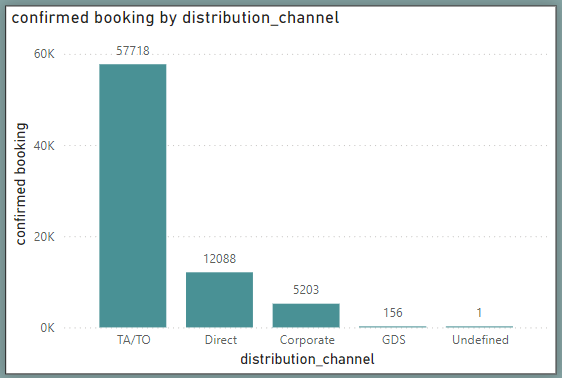
****

****

Overall Analysis:  
Online TA and Offline TA/TO have the highest booking volumes, but Groups show the highest cancellation rate (61.06%). The Undefined segment has a 100% cancellation rate but a negligible share of bookings. Focusing on reducing cancellations in Groups and Online TA segments could improve booking stability and revenue.

Recommendation:  
To reduce cancellations, hotels should focus on understanding why Groups and Online TA segments have high cancellation rates. Strategies like flexible policies, clearer communication, and targeted incentives could improve booking reliability.

**Que 14. Compare the effectiveness of booking distribution channels in generating confirmed bookings. Identify the most commonly used channels by guests.**

****

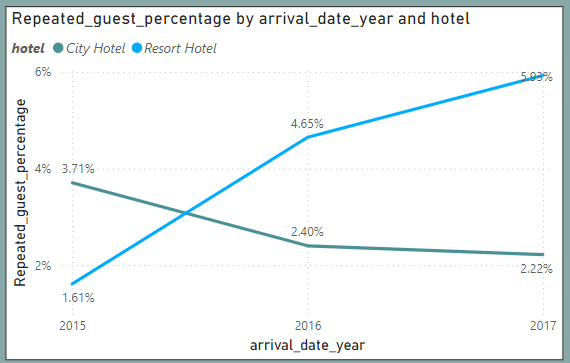
Insights from the Visual:

1. Dominant Channel (TA/TO): The Travel Agents/Tour Operators (TA/TO) channel is the most effective, generating the highest number of confirmed bookings (57,718). This suggests that a majority of guests rely on intermediaries for their bookings, making it the most commonly used channel.
2. Direct Bookings: The Direct channel is the second most popular, with 12,088 confirmed bookings. This indicates that a significant portion of guests prefer to book directly, either through the hotel's website or in-person interactions.
3. Corporate Channel: The Corporate channel accounts for 5,203 confirmed bookings, catering mainly to business travelers. This reflects a niche but steady demand.
4. Minimal Use of GDS: The Global Distribution System (GDS) has a negligible share of confirmed bookings (156), indicating it is not a primary channel for this hotel or guest segment.
5. Undefined Bookings: There is only 1 undefined booking, which suggests data integrity is well-maintained, with minimal inconsistencies.

Suggestions:

* Strengthen TA/TO partnerships.
* Enhance direct booking platforms to boost profitability.
* Expand corporate packages to attract more business clients.
* Investigate GDS for potential growth opportunities.

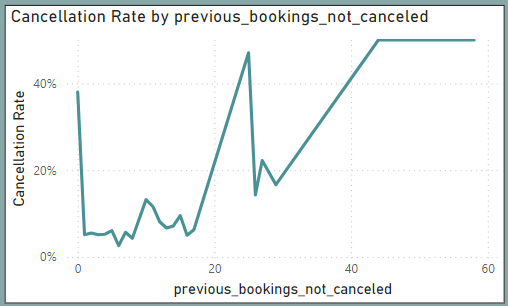
**Que 15. Visualize the percentage of repeated guests for each hotel type (Resort Hotel vs. City Hotel) over time. Explore factors influencing guest retention.**

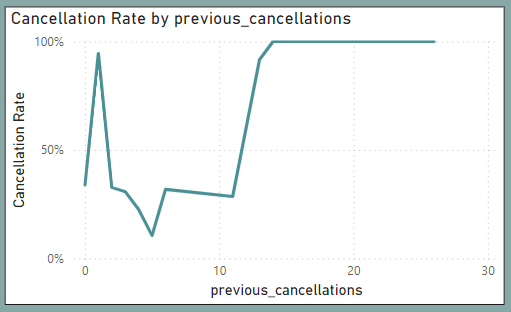


1. City Hotel:
   * The percentage of repeated guests consistently declined over the years. It started at 3.71% in 2015, dropped to 2.40% in 2016, and further decreased to 2.22% in 2017.
   * This indicates potential challenges in retaining guests, possibly due to increased competition, service quality issues, or a focus on one-time business travellers rather than repeat customers.
2. Resort Hotel:
   * The percentage of repeated guests steadily increased. It started at 1.61% in 2015, rose to 4.65% in 2016, and peaked at 5.93% in 2017.
   * This suggests effective strategies in guest retention, possibly through superior leisure experiences, loyalty programs, or enhanced amenities catering to vacationers.
3. Recommendations:
   * City Hotels should consider implementing loyalty programs, improving guest experiences, or targeting leisure travellers to increase repeat bookings.
   * Resort Hotels can further capitalize on their upward trend by promoting exclusive offers and enhancing personalized services to retain their loyal customer base.

This analysis emphasizes the importance of tailored strategies for each hotel type to improve guest retention.

**Que 16. Analyze the impact of a guest's booking history (previous cancellations and noncanceled bookings) on their likelihood of canceling a current booking.**

****

****

Cancellation Rate by previous\_bookings\_not\_canceled

Insight:

* Guests with 0 past successful bookings show a high cancellation rate (~40%+).
* As the number of previous non-canceled bookings increases (1–20), the cancellation rate remains low (mostly under 20%), indicating trust built with repeated successful stays.
* However, beyond 20, the cancellation rate rises sharply again, especially beyond 30–40 previous successful bookings, reaching 45–50%.

Interpretation:

* First-time or new guests are more likely to cancel.
* Repeat guests with a moderate history of completed bookings (1–20) are the least likely to cancel.
* Very frequent guests (perhaps corporate or group travellers with many prior stays) show higher cancellation rates, possibly due to bulk bookings or flexible plans.

Cancellation Rate by previous\_cancellations

Insight:

* Guests with even 1 previous cancellation show a dramatically increased cancellation rate, nearing 100%.
* The rate dips slightly with 2–10 past cancellations but remains high (~30–60%).
* From 11 previous cancellations onward, the cancellation rate stabilizes near 100%.

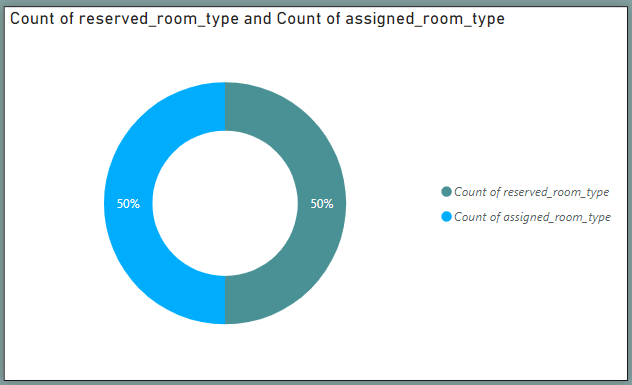
Interpretation:

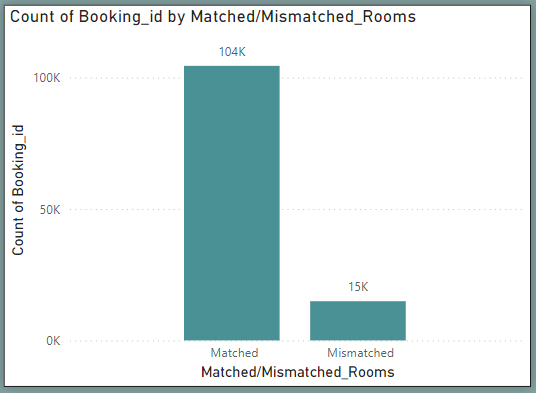
* A guest’s history of cancellations is a strong predictor of future cancellations.
* If someone has canceled bookings in the past, especially multiple times, there is an extremely high chance they will cancel again.

Overall Conclusion:

* Past behaviour strongly influences future booking reliability.
* Guests with no or few cancellations and several completed bookings are the most trustworthy.
* Conversely, guests with a history of cancellations or no prior positive booking history are high-risk for cancellation

**Que 17. Visualize the distribution of reserved and assigned room types. Analyze whether guests tend to receive the room type they initially reserved.**

****

****

The analysis shows that 104K bookings had matched room types, indicating that most guests received the room they reserved, reflecting effective room allocation and high guest satisfaction.

However, 15K bookings had mismatched room types, likely due to issues like overbooking or operational inefficiencies, which could lead to dissatisfaction.

Recommendations:

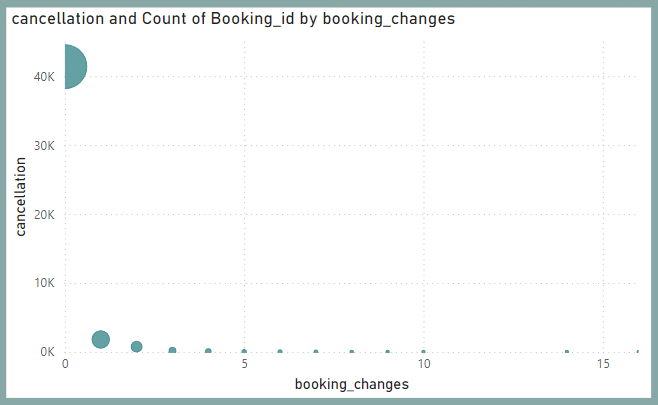
1. Implement better inventory management systems to reduce mismatches. For example, block reserved room types in advance to avoid accidental reallocation.

2. If a mismatch is unavoidable, inform guests in advance and offer upgrades or other compensations to maintain satisfaction.

3. Actively gather feedback from guests about their experience with assigned rooms to further improve room allocation and operational efficiency.

Overall, the hotel's performance in matching reserved room types is strong, but addressing mismatched cases can enhance guest satisfaction and operational excellence.

**Que 18. Investigate the relationship between the number of booking changes made by guests and their likelihood of canceling a booking.**

****

**** Correlation Between Booking Changes and Cancellations:

* The scatter plot suggests a clear relationship between the number of changes made to a booking and the likelihood of cancellation.
* Bookings with fewer or no changes have a lower cancellation count, while bookings with more changes tend to have higher cancellations.

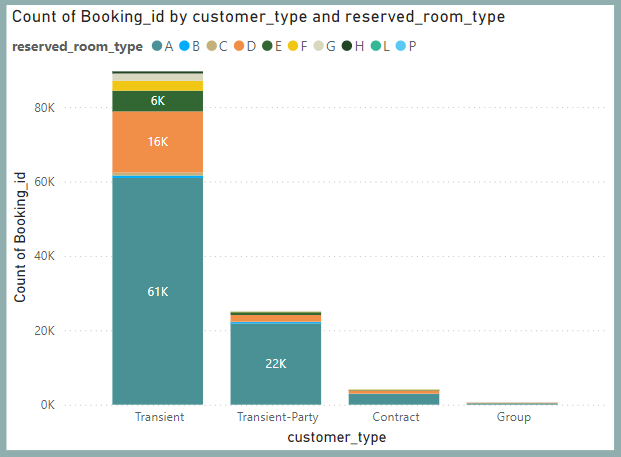
 Significance of Booking Changes:

* Frequent modifications (e.g., changes in dates, room type, or guest details) might indicate uncertainty from the guest, which increases the chances of cancellation.
* Bookings with many changes may require special attention from the hotel to reduce the risk of cancellations.

 What This Means for Hotels:

* Monitoring bookings with multiple changes could allow hotels to proactively reach out to guests, clarify their needs, and possibly prevent cancellations.
* Hotels can also implement policies or incentives for guests to finalize their plans early to reduce booking changes.

**Que 19. Analyze room type preferences based on customer types (e.g., Transient, Group) and identify any patterns in room type selection.**

****

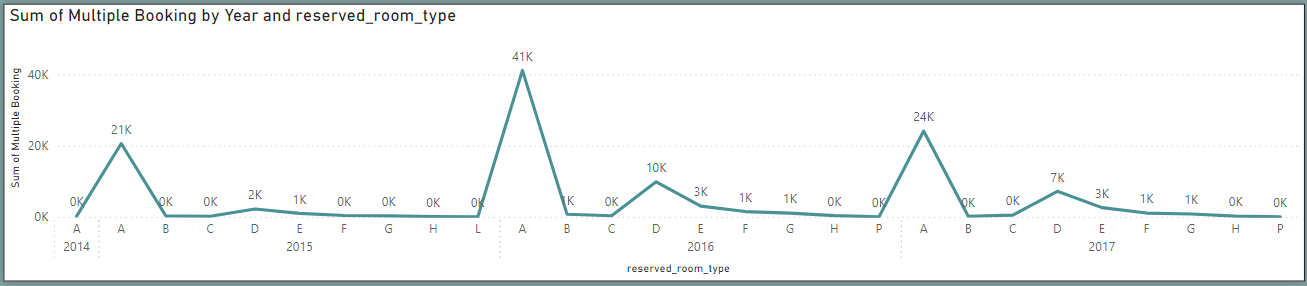
Insights :

1. Transient Customers:  
   Transient customers dominate bookings, with 61K preferring Room Type A, followed by D (16K) and E (6K). This highlights Room Type A’s appeal for short-term stays.
2. Transient-Party Customers:  
   Transient-party customers also favor Room Type A (22K bookings), showing consistency in preferences across transient categories.
3. Contract and Group Customers:  
   Contract and group customers have minimal bookings, with Room Type A still preferred. Their lower volume suggests situational or niche bookings.

Suggestions:

* Focus on optimizing and promoting Room Type A, as it is the most popular.
* Explore why certain room types (e.g., B, C, P) are underutilized and consider strategies like discounts or targeted marketing to boost their bookings.

**Que 20. Analyze whether guests who make multiple bookings tend to consistently request the same room type or if their preferences change over time.**



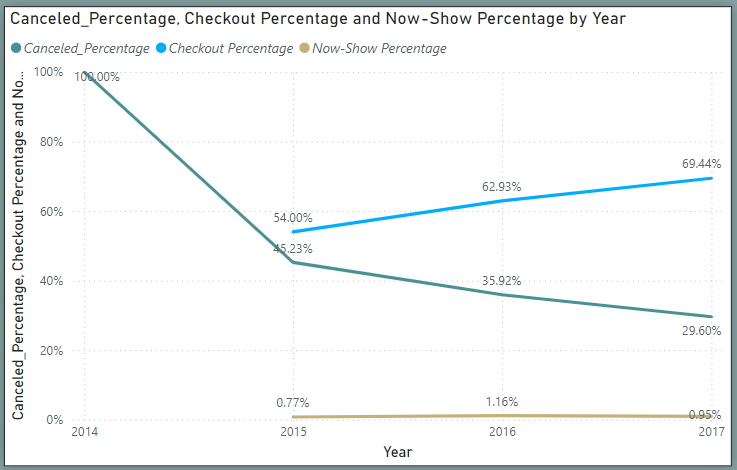
Insights:

1. Room Type "A" Dominates: Guests making multiple bookings consistently prefer room type "A" across all years (2014–2017), with significant peaks in 2015 (41K bookings) and 2016 (24K bookings).
2. Other Room Types Are Rarely Chosen: Room types "B," "C," "D," and "E" show occasional use, but their booking counts are much lower, suggesting minimal preference changes among repeat bookers.
3. Stable Preferences: The data indicates that guests who make multiple bookings tend to maintain consistent preferences, with little evidence of changing room types over time.

Suggestions:

1. Segment multiple-booking guests based on demographics (e.g., age, nationality) to understand why room type "A" is so popular and if specific groups drive this preference.
2. Offer targeted promotions or packages for room types "B," "C," "D," and "E" to diversify guest preferences and optimize room allocation.
3. Use guest reviews or survey data to determine if satisfaction with room type "A" drives repeat bookings, or if the trend is influenced by availability and pricing.

**Que 21. Provide an overview of reservation statuses over time, including the percentage of canceled, checkedout, and noshow bookings.**



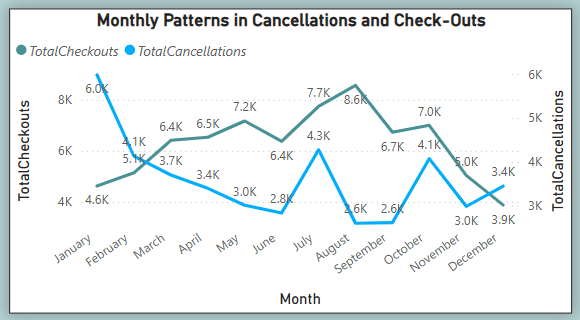
This chart analyses trends in cancellation rate, checkout rate, and no-show rate from 2014 to 2017:

1. Cancellation Rate (Blue Line):
   * The cancellation rate decreased significantly from 100% in 2014 to 29.60% in 2017, indicating improved booking stability over time. This could be attributed to stricter policies or better customer awareness.
2. Checkout Rate (Green Line):
   * The checkout percentage has steadily increased from 0% in 2014 to 69.44% in 2017, reflecting an increase in completed stays. This trend suggests better guest retention and reduced cancellations.
3. No-Show Rate (Yellow Line):
   * The no-show rate remains minimal across all years, fluctuating slightly between 0.77% (2015) and 0.95% (2017), indicating a consistent low impact on hotel operations.

Key Takeaways:

* Positive Trend: The growing checkout percentage and declining cancellation rate demonstrate improved guest commitment and operational efficiency.
* Actionable Insight: Maintain efforts to further reduce cancellations by enhancing communication, offering flexible policies, or introducing loyalty programs.
* Low No-Show Impact: The consistently low no-show rate highlights that this is not a major operational concern.

**Que 22. Analyze trends in reservation status dates, such as the busiest checkout dates or patterns in cancellations by month.**



Checkouts Trend:

* The hotel experiences peak checkouts between May and September, with September (8.6K) and August (7.7K) being the highest.
* February (4.6K) and December (3.9K) show the lowest checkouts, indicating off-season periods.

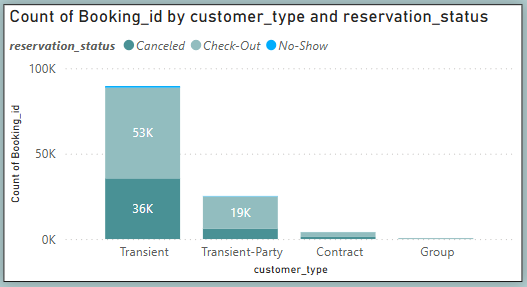
Cancellations Trend:

* January (6.0K) records the highest cancellations, followed by August (6.0K) and November (5.0K).
* September (2.6K) and June (2.8K) have minimal cancellations, showing high booking reliability.

Checkout vs Cancellation Pattern:

* Months with high checkouts generally have low cancellations, especially September, indicating strong customer intent.
* Early-year months (Jan–Feb) show a reverse pattern, with more cancellations than checkouts, possibly due to post-holiday booking changes or weather issues.

**Que 23. Visualize how reservation statuses vary across different customer types (e.g., Transient, Group) and identify if certain customer types are more likely to result in cancellations or noshows.**

****

Transient customers account for the majority of bookings, but also show the highest number of cancellations (36K) and check-outs (53K).

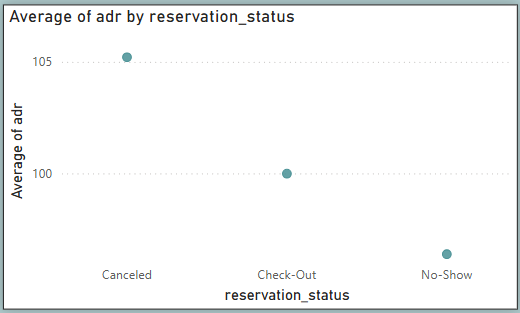
Transient-Party has fewer bookings but a high cancellation rate relative to its total.

Contract and Group customer types have minimal bookings with negligible cancellations or no-shows.

Recommendation:

* Prioritize retention strategies for Transient customers—e.g., flexible policies or incentives to reduce cancellations.
* Monitor Transient-Party behavior closely to understand and mitigate potential booking risks.
* Maintain Contract and Group booking relationships, which show high reliability.

**Que 24. Explore the relationship between reservation statuses and Average Daily Rates (ADR) to determine if there are differences in ADR based on booking outcomes.**



Canceled bookings have the highest ADR, followed by Check-Out, while No-Shows have the lowest ADR. This suggests high-value bookings are more likely to be cancelled, potentially impacting revenue projections.

Recommendation:

* Implement stricter cancellation policies or require deposits for high-ADR bookings to reduce lost revenue.
* Investigate no-show reasons for low-ADR bookings and consider overbooking strategies to mitigate the impact.

**EDA**

**PROBLEM STATEMENTS**

**1.Understand the distribution of arrival dates, including the most common arrival days and summary statistics for lead times.**

with cte as (

SELECT

concat(arrival\_date\_year,'-',arrival\_date\_month,'-',arrival\_date\_day\_of\_month) as arrival\_dates

,STR\_TO\_DATE(concat(arrival\_date\_year,'-',arrival\_date\_month,'-',arrival\_date\_day\_of\_month),'%Y-%M-%d') as arrival\_

,count(\*) as Booking\_count

,max(lead\_time) as max\_lead

,min(lead\_time) as min\_lead

, round(avg(lead\_time),2) as avg\_lead

, round(stddev(lead\_time),2) as stddev\_lead

FROM hotel\_booking.booking\_details

Group by 1,2

order by 3 desc

limit 10

)

select arrival\_dates,Booking\_count, max\_lead,min\_lead,avg\_lead,stddev\_lead

,dayname(arrival\_) as 'Days'

from cte

***Most bookings occur on Thursdays and Fridays with highly variable lead times, indicating both weekend travel trends and diverse booking behaviors.***

**2.** **Identify peak booking months and analyze reasons for spikes in bookings, including holidays or events.**

SELECT

arrival\_date\_year AS year,

arrival\_date\_month AS month,

COUNT(\*) AS total\_bookings

FROM booking\_details

GROUP BY arrival\_date\_year, arrival\_date\_month

ORDER BY total\_bookings DESC;

***August and July month dominate booking numbers.March to May month show steady growth. November and December months experience a significant decline in bookings.***

**3. Calculate the average** **of stays for different hotel types and explore variations by meal plans.**

SELECT   
 hotel as hotel\_type,  
 meal as meal\_plan,  
 ROUND(AVG(stays\_in\_weekend\_nights + stays\_in\_week\_nights), 2) as avg\_length\_of\_stay  
 FROM   
 booking\_details as bs  
 join meal\_and\_stay\_details as ms  
 on   
 bs.Booking\_id=ms.Booking\_id  
GROUP BY   
 hotel, meal  
ORDER BY   
 hotel\_type, avg\_length\_of\_stay DESC;

***Resort Hotels attract longer stays, with SC as the preferred meal plan, while City Hotels cater to shorter stays, favoring BB as the top meal plan.***

**4.** **Analyze how booking patterns have evolved over the years, including yearoveryear changes in bookings and cancellations.**

SELECT

arrival\_date\_year as booking\_year,

COUNT(\*) as total\_bookings,

SUM(CASE WHEN is\_canceled = 1 THEN 1 ELSE 0 END) as total\_cancellations,

COUNT(\*) - LAG(COUNT(\*)) OVER (ORDER BY arrival\_date\_year) as bookings\_change,

SUM(CASE WHEN is\_canceled = 1 THEN 1 ELSE 0 END) - LAG(SUM(CASE WHEN is\_canceled = 1 THEN 1 ELSE 0 END)) OVER (ORDER BY arrival\_date\_year) as cancellations\_change

FROM

booking\_details

GROUP BY

arrival\_date\_year

ORDER BY

booking\_year;

***Bookings peaked in 2016 with a big rise from 2015, but dropped in 2017, while cancellations followed the same up-and-down pattern.***

**5.** **Understand the distribution of the number of adults, children, and babies and identify any outliers.**

WITH ctc AS (

SELECT

b.booking\_id, b.hotel, g.adults, g.children g.babies,

(g.adults + g.children + g.babies) AS total\_guests,

CASE

WHEN (g.adults + g.children + g.babies) > 10 THEN 'Outlier'

ELSE 'Normal'

END AS outlier\_flag

FROM

Guest\_Info AS g

JOIN booking\_details AS b

ON g.booking\_id = b.booking\_id

)

SELECT

ctc.hotel,

SUM(ctc.adults) AS Total\_Adults,

SUM(ctc.children) AS Total\_Children,

SUM(ctc.babies) AS Total\_Babies

FROM

ctc

WHERE

ctc.outlier\_flag='Normal'

GROUP BY

ctc.hotel;

***City Hotel had more guests overall, while Resort Hotel hosted more families with babies, and bookings with more than 10 guests were rare outliers.***

**6.** **Calculate summary statistics for ADR and explore differences between Resort Hotel and City Hotel bookings.**

SELECT

hotel,

ROUND(AVG(adr), 2) as average\_adr,

ROUND(MIN(adr), 2) as min\_adr,

ROUND(MAX(adr), 2) as max\_adr,

ROUND(STDDEV(adr), 2) as std\_dev\_adr

FROM

booking\_details as bd

join meal\_and\_stay\_details as m

on

bd.Booking\_id=m.Booking\_id

GROUP BY

hotel

ORDER BY

hotel;

***City Hotel had a higher average ADR, but Resort Hotel showed more price variation and included some unusual low rates.***

**7.** **Analyze the distribution of required car parking spaces for each hotel type and determine if one type attracts more guests with cars.**

SELECT

hotel,

SUM(required\_car\_parking\_spaces > 0) AS bookings\_with\_parking,

SUM(required\_car\_parking\_spaces) AS total\_parking\_spaces,

COUNT(\*) AS total\_bookings

FROM

booking\_details as bd

join meal\_and\_stay\_details as m

on

bd.Booking\_id=m.Booking\_id

GROUP BY

hotel

ORDER BY

bookings\_with\_parking DESC;

***Resort Hotels attract more guests with cars compared to City Hotels and offer significantly more parking spaces to accommodate this demand.***

**8.** **Compare the total number of special requests made by different customer types (e.g., Transient, Group) and identify which customer type makes more requests.**

SELECT

customer\_type,

SUM(total\_of\_special\_requests) AS total\_special\_requests, COUNT(\*) AS total\_bookings,

FROM

booking\_source\_and\_history as b

join meal\_and\_stay\_details as m

on

b.Booking\_id=m.Booking\_id

GROUP BY

customer\_type

ORDER BY

total\_special\_requests DESC;

***Transient customers make the highest number of special requests, followed by Transient-Party customers, while Groups and Contract customers have significantly fewer special requests.***

**9.** **Understand the distribution of meal plans (e.g., BB, HB, FB, SC) and identify any patterns or preferences.**

SELECT

meal AS meal\_plan,

COUNT(\*) AS total\_bookings,

ROUND(100 \* COUNT(\*) / (SELECT COUNT(\*) FROM Meal\_And\_Stay\_Details), 2) AS percentage\_of\_bookings

FROM

meal\_and\_stay\_details

GROUP BY

meal

ORDER BY

total\_bookings DESC;

***Guests overwhelmingly prefer BB meal plans, suggesting its popularity and suitability for a wide range of travelers, while FB and Undefined options have limited appeal.***

**10.** **Analyze Average Daily Rates (ADR) by meal plan type to identify variations in pricing.**

SELECT

meal AS meal\_plan,

ROUND(AVG(adr), 2) AS average\_adr

FROM

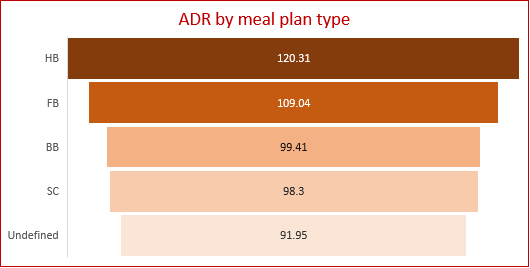
Meal\_And\_Stay\_Details

GROUP BY

meal

ORDER BY

average\_adr DESC;



***Half Board leads in ADR, reflecting its premium value, followed by Full Board and Bed & Breakfast. Self-Catering appeals to budget-conscious guests, while Undefined plans show minimal preferences.***

**11.** **Investigate the distribution of required car parking spaces and special requests by hotel type and meal plan. 12 .Compare the distribution of meal plans among different customer types (e.g., Transient, Group) to identify preferences.**

SELECT

hotel as hotel\_type,

sum(required\_car\_parking\_spaces) as total\_parking\_spaces,

sum(total\_of\_special\_requests) as total\_special\_requests,

meal as meal\_plan,

customer\_type

FROM

meal\_and\_stay\_details as m

join booking\_details as h

on m.Booking\_id = h.Booking\_id

join booking\_source\_and\_history as b

on m.Booking\_id = b.Booking\_id

GROUP BY

hotel,customer\_type,meal

ORDER BY

customer\_type, hotel\_type , total\_special\_requests desc;

***City Hotels attract guests with special requests, while Resort Hotels cater to leisure travelers with high parking demand.***

***Bed & Breakfast is the most popular meal plan, appealing to transient guests, while Half Board suits groups, Full Board caters to hotel-focused stays, Self-Catering attracts budget-conscious travelers, and Undefined plans see minimal use.***

**12.** **Understand the distribution of bookings across different market segments and calculate summary statistics for lead times within each segment.**

SELECT

m.market\_segment,

COUNT(\*) AS total\_bookings,

ROUND(AVG(bd.lead\_time), 2) AS avg\_lead\_time,

MIN(bd.lead\_time) AS min\_lead\_time,

MAX(bd.lead\_time) AS max\_lead\_time,

ROUND(STDDEV(bd.lead\_time), 2) AS stddev\_lead\_time

FROM

booking\_details as bd

JOIN

booking\_source\_and\_history bs

ON bd.Booking\_id = bs.Booking\_id

JOIN

market\_segment as m

ON bs.market\_segment\_id = m.market\_segment\_id

GROUP BY

m.market\_segment

ORDER BY

total\_bookings DESC;

***Online TA leads with 56,477 bookings and an average lead time of 83 days, highlighting a mix of early and last-minute bookings.***

***Offline TA/TO and Groups show the longest lead times , reflecting advanced planning.***

***Direct and Corporate bookings have shorter lead times, driven by spontaneous travel and business needs.***

***Aviation bookings have the shortest lead time, catering to predictable requirements like crew stays.***

**13.** **Analyze the distribution of bookings through different booking channels (e.g., online travel agents, direct bookings) and calculate the percentage of bookings through each channel.**

SELECT

distribution\_channel AS booking\_channel,

COUNT(\*) AS total\_bookings,

ROUND((COUNT(\*) \* 100.0 / (SELECT COUNT(\*) FROM Booking\_Source\_and\_History)), 2) AS percentage\_bookings

FROM

booking\_source\_and\_history as b

join distribution\_channel as d

on b.distribution\_channel\_id= d.distribution\_channel\_id

GROUP BY

distribution\_channel

ORDER BY

total\_bookings DESC;

***TA/TO dominates with 81.98% bookings, while direct and corporate channels contribute 12.27% and 5.59%, highlighting areas for diversification.***

**14.** **Calculate the proportion of repeated guests and investigate their booking behavior. Identify any patterns or differences in preferences compared to firsttime guests.**

SELECT

CASE

WHEN is\_repeated\_guest = 1

THEN "Repeated Guest"

ELSE "First Time Guest"

END AS Guest,

ROUND(100.0 \* COUNT(\*) / SUM(COUNT(\*)) OVER (), 2) AS percentage\_of\_bookings,

ROUND(AVG(lead\_time), 2) AS avg\_lead\_time,

ROUND(AVG(total\_of\_special\_requests), 2) AS avg\_special\_requests,

ROUND(AVG(stays\_in\_weekend\_nights + stays\_in\_week\_nights), 2) AS avg\_stay\_length,

ROUND(AVG(adr), 2) AS avg\_adr

FROM

meal\_and\_stay\_details m

JOIN

booking\_source\_and\_history b ON b.Booking\_id = m.Booking\_id

JOIN

booking\_details bd ON b.Booking\_id = bd.Booking\_id

GROUP BY

1

ORDER BY

total\_bookings DESC;

***Repeated guests (3.19%) prefer shorter stays (1.93 nights), lower lead times (30.79 days), and lower ADR (64.45), but make slightly more special requests, indicating loyalty-driven but cost-conscious behavior.***

**15.** **Explore the impact of a guest's booking history on their likelihood of cancelling a current booking. Calculate cancellation rates based on previous cancellations and noncancelled bookings.**

SELECT

CASE

WHEN previous\_cancellations = 0 AND previous\_bookings\_not\_canceled = 0 THEN 'No Booking History'

WHEN previous\_cancellations > 0 THEN 'previous\_Cancellations'

WHEN previous\_bookings\_not\_canceled > 0 THEN 'Non\_canceled booking'

ELSE 'N/A'

END AS booking\_history\_category,

SUM(CASE WHEN is\_canceled = 1 THEN 1 ELSE 0 END) AS total\_cancellations,

ROUND(100.0 \* SUM(CASE WHEN is\_canceled = 1 THEN 1 ELSE 0 END) / COUNT(\*), 2) AS cancellation\_rate

FROM

booking\_source\_and\_history as bs

JOIN

booking\_details as b

ON b.Booking\_id = bs.Booking\_id

GROUP BY

1

ORDER BY

total\_bookings DESC;

***Guests with a history of previous cancellations show a significantly higher cancellation rate of 91.64%, contributing to the majority of total cancellations, whereas guests with no prior cancellations exhibit a very low cancellation rate of 2.35%, indicating a much more reliable booking behaviour.***

**16.** **Understand the distribution of reserved and assigned room types. Calculate summary statistics for the consistency between reserved and assigned room types.**

SELECT

reserved\_room\_type,

COUNT(\*) AS total\_bookings,

SUM(CASE WHEN reserved\_room\_type = assigned\_room\_type THEN 1 ELSE 0 END) AS consistent\_room\_type,

SUM(CASE WHEN reserved\_room\_type != assigned\_room\_type THEN 1 ELSE 0 END) AS inconsistent\_room\_type,

ROUND(SUM(CASE WHEN reserved\_room\_type = assigned\_room\_type THEN 1 ELSE 0 END) \* 100.0 / COUNT(\*), 2) AS Consistency\_Percentage

FROM

room\_details

GROUP BY

reserved\_room\_type

ORDER BY 1;

***Most guests received the room type they originally reserved, with overall consistency above 85% for common room types like A, D, E, and G. However, rare types like L had very low consistency (only 16.67%).***

**17.** **Analyze the impact of booking changes on cancellation rates. Calculate cancellation rates for bookings with different numbers of changes.**

SELECT

booking\_changes,

ROUND(100.0 \* SUM(CASE WHEN is\_canceled = 1 THEN 1 ELSE 0 END) / COUNT(\*), 2) AS cancellation\_rate

FROM

room\_details r

JOIN

booking\_details b ON r.Booking\_id = b.Booking\_id

GROUP BY

booking\_changes

ORDER BY

booking\_changes ASC

limit 11;

***Bookings with no changes have the highest cancellation rate, while minor changes reduce cancellations, but excessive changes can increase the likelihood of cancellations.***

**18.** **Explore how room type preferences vary across different customer types (e.g., Transient, Group). Identify if certain customer types have specific room preferences.**

SELECT

customer\_type,

reserved\_room\_type,

COUNT(\*) AS total\_bookings

FROM

booking\_source\_and\_history bsh

JOIN

room\_details r ON bsh.Booking\_id = r.Booking\_id

GROUP BY

customer\_type, reserved\_room\_type

ORDER BY

customer\_type, percentage\_of\_bookings DESC;

***Transient and Transient-Party customers prefer Room Type A, while Contract customers favor Room Type D. Room Types B, C, G, and H are less favored, particularly by Transient-Party and Group customers***

**19.** **Examine whether guests who make multiple bookings have consistent room type preferences or if their preferences change over time.**

SELECT

distinct rd.reserved\_room\_type,

CASE

WHEN COUNT(r.booking\_id) > 1 THEN "Multiple\_Booking"

ELSE "Single\_Booking"

END AS multiple\_booking,

DATE\_FORMAT(r.reservation\_status\_date, '%Y') AS year,

COUNT(r.booking\_id) AS booking\_count

FROM

reservation\_status r

JOIN

room\_details rd

ON

r.booking\_id = rd.booking\_id

GROUP BY

1,3

ORDER BY

1;

***Room type preferences remain consistent over the years, with Room Type "A" being the most popular across all years, accounting for over 70% of total bookings. Preferences for other room types like "B," "C," and "D" show steady but much smaller proportions, indicating specific but less common demand.***

**20.** **Understand the distribution of reservation statuses and calculate summary statistics for reservation status dates.**

SELECT

reservation\_status,

COUNT(\*) AS total\_bookings,

MIN(reservation\_status\_date) AS earliest\_status\_date,

MAX(reservation\_status\_date) AS latest\_status\_date,

DATEDIFF(MAX(reservation\_status\_date), MIN(reservation\_status\_date)) AS date\_range

FROM

reservation\_status

GROUP BY

reservation\_status

ORDER BY

total\_bookings DESC;

***Most bookings result in check-outs, indicating high fulfilment rates. Cancellations are significant, suggesting potential areas for policy or process improvement. No-shows are minimal but still worth addressing to optimize room utilization.***

**21.** **Analyze trends in reservation status dates, including the most common checkout dates and any seasonality patterns.**

SELECT

DATE\_FORMAT(reservation\_status\_date, '%Y') AS Total\_year,

DATE\_FORMAT(reservation\_status\_date, '%M') AS month\_name,

CASE

WHEN MONTH(reservation\_status\_date) IN (12, 1, 2) THEN 'Winter'

WHEN MONTH(reservation\_status\_date) IN (3, 4, 5) THEN 'Summer'

WHEN MONTH(reservation\_status\_date) IN (6, 7, 8, 9) THEN 'Monsoon'

WHEN MONTH(reservation\_status\_date) IN (10, 11) THEN 'Autumn'

ELSE 'Unknown'

END AS seasons,

reservation\_status,

COUNT(\*) AS total\_reservations

FROM

reservation\_status

WHERE

reservation\_status = 'Check-Out'

GROUP BY

1,2,3

ORDER BY

1;

***Monsoon and Summer seasons are the most popular for bookings, with Summer showing a growing trend.***

***Winter remains the least popular, while Autumn shows moderate and steady growth.***

***Total reservations have grown over the years, highlighting an increasing demand across all seasons.***

**22.** **Explore how reservation statuses vary across different customer types (e.g., Transient, Group) using Excel or SQL. Calculate cancellation rates by customer type.**

SELECT

reservation\_status,

customer\_type,

ROUND(SUM(CASE WHEN reservation\_status = 'Canceled' THEN 1 ELSE 0 END) \* 100.0 / COUNT(b.Booking\_id), 2) AS cancellation\_rate,

COUNT(b.Booking\_id) AS total\_bookings,

SUM(CASE WHEN reservation\_status = 'Canceled' THEN 1 ELSE 0 END) AS total\_canceled

FROM

booking\_source\_and\_history AS b

JOIN

reservation\_status AS r ON b.Booking\_id = r.Booking\_id

GROUP BY

customer\_type, reservation\_status

ORDER BY

customer\_type, reservation\_status;

***Transient customers are more likely to cancel or not show up but also account for a majority of check-outs.***

***Contract and Group customers have fewer cancellations and no-shows, indicating more stable booking behaviour.***

**23.** **Investigate whether there are differences in Average Daily Rates (ADR) based on reservation status (e.g., canceled vs. checkedout).**

SELECT

r.reservation\_status,

ROUND(AVG(adr), 2) AS avg\_adr,

COUNT(\*) AS total\_bookings

FROM

reservation\_status r

JOIN

meal\_and\_stay\_details as m ON r.Booking\_id = m.Booking\_id

GROUP BY

r.reservation\_status

ORDER BY

avg\_adr DESC;

***Canceled bookings have the highest ADR, while checked-out bookings dominate in volume with moderate rates.***

**Conclusion**

This hotel booking analysis revealed valuable insights into guest behaviour, booking patterns, and revenue drivers. Seasonal trends showed that Monsoon and Summer had the highest booking volumes, while Winter remained the weakest season, indicating areas for targeted marketing.

Guests preferred the BB (Bed & Breakfast) meal plan and Room Type A, especially for longer stays, making them key offerings for customer satisfaction. A strong link was found between frequent booking changes and cancellations, suggesting the need for proactive guest communication.

In terms of booking sources, Travel Agents/Tour Operators (TA/TO) generated the highest number of bookings but also faced the most cancellations. In contrast, Direct and Corporate bookings were more stable and reliable. Focusing on these channels can help increase revenue and reduce booking risk.

Revenue trends highlighted that the ₹50–100 ADR range attracted the highest number of bookings and special requests, making it the most profitable segment. Additionally, City Hotels mostly catered to short-term business guests, while Resort Hotels attracted longer-stay leisure travellers and had higher repeat booking rates.

Overall, these insights can help hotels optimize pricing, improve service quality, reduce cancellations, and boost profitability through data-driven decision-making.

**REPORT ENDED**