**CAPSTONE PROJECT ON HOTEL BOOKING ANALYSIS**

**OVERVIEW**

This hotel booking analysis provides a clear snapshot of guest booking patterns, stay behaviors, and reservation trends. It highlights key factors such as preferred hotel types, lead times, peak arrival periods, and cancellation rates, helping to understand how guests plan and experience their stays. By analyzing when and why bookings happen — and how they vary across weekdays, weekends, and seasons — the insights support smarter decisions for pricing, promotions, and resource planning. Overall, this analysis helps hotel managers optimize occupancy, boost revenue, and deliver better guest experiences.

**THE PROCESS**

1. **Dataset Acquisition from GitHub:**

The dataset for this project was obtained from a GitHub repository, which contained Hotel-related data in CSV and SQL formats. These files included key guest booking, cancellation, meal and stay duration records, providing a foundation for analyzing factors influencing Hotel Analysis.

1. **Data Loading:**

The CSV file was imported into Excel for preliminary exploration, while the SQL file was loaded into MySQL for structured querying. This helped in understanding the data structure, identifying missing values, and preparing for further processing.

1. **Data Cleaning (ETL Process):**

To ensure data consistency and accuracy, an ETL (Extract, Transform, Load) process was performed. This involved removing duplicates, handling missing values, correcting data types, and normalizing records in Excel. The cleaned dataset was then structured for integration with analytical tools.

1. **Tool Integration:**

After data preparation, the refined dataset was connected to Excel, Power BI, and MySQL for deeper analysis and visualization. Excel was used for data manipulation, Power BI for interactive dashboards, and MySQL for advanced querying and data retrieval.

1. **Problem Definition & Insights Generation:**

Using Power BI, key problem statements were formulated to explore factors affecting Hotel Analysis. Various charts, graphs, and visual representations were created to analyze trends in total bookings, cancellation rates, customer meal preferance and room reserved by them. These insights provided a clear understanding of how different variables influence Hotel Analysis.

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

An in-depth Exploratory Data Analysis (EDA) was conducted using Excel and SQL to uncover patterns and correlations. SQL queries helped in filtering data, performing statistical analysis, and drawing relationships between customer behaviour, guest satisfaction and Hotel preferance. Excel’s pivot tables and statistical tools further refined these insights.

1. **Presentation & Reporting:**

Finally, the key findings, insights, and recommendations were compiled into a PowerPoint presentation. This report effectively summarized the analysis with graphs, tables, and actionable suggestions for improving Hotel Analysis.

1. **Detail Documentation:** Compile a detailed report that meticulously documents the entire project lifecycle.Include sections on data collection,transformation,problem statement formulation,toll integration,Power BI solution,EDA insights and Powerpoint visualizations.

**OBJECTIVE**

The primary objective of the Hotel Booking Analysis is to convert raw booking data into clear, actionable insights that empower hotel managers, analysts, and decision-makers to make informed, data-driven choices. In today’s competitive hospitality industry, understanding guest booking behavior is essential for maximizing occupancy, revenue, and guest satisfaction.

This analysis focuses on uncovering patterns in how and when guests book their stays, whether they choose city hotels or resort properties, how far in advance they plan (*lead time*), and how their bookings fluctuate across different days, weeks, and months. It also closely examines cancellation rates, weekend versus weekday stays, and other factors that directly impact daily operations and long-term profitability.

By exploring these trends, the goal is to help hotels better forecast demand, plan staffing and resources more efficiently, and set dynamic pricing strategies that respond to peak seasons and low-demand periods. Understanding cancellation behavior further supports smarter overbooking policies and reduces lost revenue.

Additionally, the insights from this booking analysis enable hotels to design targeted marketing campaigns and promotional offers that align with guest needs — for example, weekend getaway packages for leisure travelers or weekday deals for business guests. It also creates opportunities to strengthen guest loyalty by identifying booking behaviors of repeat customers.

Overall, the objective is to equip stakeholders with a powerful, data-backed understanding of the entire booking lifecycle — from reservation to check-out — so they can make strategic decisions that drive growth, improve guest experiences, and keep the hotel competitive in a dynamic market. This analysis ultimately turns complex booking data into a roadmap for sustainable success.

**SIGNIFICANCE**

The significance of conducting a Hotel Booking Analysis lies in its ability to bridge the gap between raw reservation data and strategic decision-making. In the hospitality industry, booking trends are a powerful indicator of guest behavior, market demand, and the hotel’s operational efficiency. Understanding these trends is critical for hotels to remain competitive, profitable, and responsive to changing guest expectations.

Through detailed booking analysis, hotel managers can identify when guests prefer to stay — such as peak seasons, weekends, or weekdays — and how far in advance they make reservations. This information is vital for effective demand forecasting, allowing hotels to plan resources, staff schedules, and inventory to match expected occupancy levels.

Moreover, analyzing lead times and cancellation patterns helps hotels develop flexible pricing and overbooking strategies, reducing lost revenue from last-minute cancellations or no-shows. Insights into booking behavior also highlight which customer segments, such as business travelers or families, are driving occupancy. This knowledge enables the hotel to design targeted marketing campaigns and special offers that resonate with these groups, increasing bookings during off-peak periods.

A robust booking analysis also sheds light on the impact of external factors like holidays, local events, or seasonality, which directly influence booking surges or drops. By recognizing these patterns, hotel management can respond proactively — for example, by adjusting room rates dynamically or launching promotional deals to attract guests during low-demand periods.

In addition, the significance extends to enhancing the guest experience. By understanding what guests want — longer stays, weekend getaways, or short business trips — hotels can tailor their services, amenities, and packages to meet diverse needs. This personalization not only boosts guest satisfaction but also strengthens loyalty and encourages repeat bookings.

In summary, the hotel booking analysis is a vital tool that transforms complex data into meaningful insights, helping hotels make smarter decisions, maximize revenue, optimize resources, and deliver exceptional guest experiences in an increasingly competitive market.

**DATA DICTIONARY**

1. **Booking\_Details**

The Booking\_Details table contains essential details related to hotel reservations. It includes a unique booking identifier and information about the type of hotel (Resort Hotel or City Hotel). Additionally, it records the booking's cancellation status (0 for not canceled, 1 for canceled), lead time (number of days between booking and arrival), year, month, week number, and day of the month of arrival. The table also captures the number of weekend and weekday nights stayed

1. **Room\_Details**

The Room\_Details table provides information related to room reservations and changes made to them. It is associated with the Booking\_Details table via the booking identifier. This table includes details about the type of room initially reserved, the type of room eventually assigned, and the number of changes made to the booking. It offers insights into room allocation dynamics and booking modifications.

1. **Reservation\_Status**

The Reservation\_Status table records the status of reservations over time. It is connected to the Booking\_Details table through the booking identifier. This table captures the reservation's last status (e.g., Canceled, Check-Out) and the date on which this status was recorded. It is valuable for tracking the progression of reservations and understanding their final outcomes.

1. **Guest\_Info**

The Guest\_Info table provides insights into the guests associated with each booking. It is linked to the Booking\_Details table via the booking identifier. This table records the number of adults, children, and babies accompanying the booking, offering an understanding of the composition of guests for each reservation.

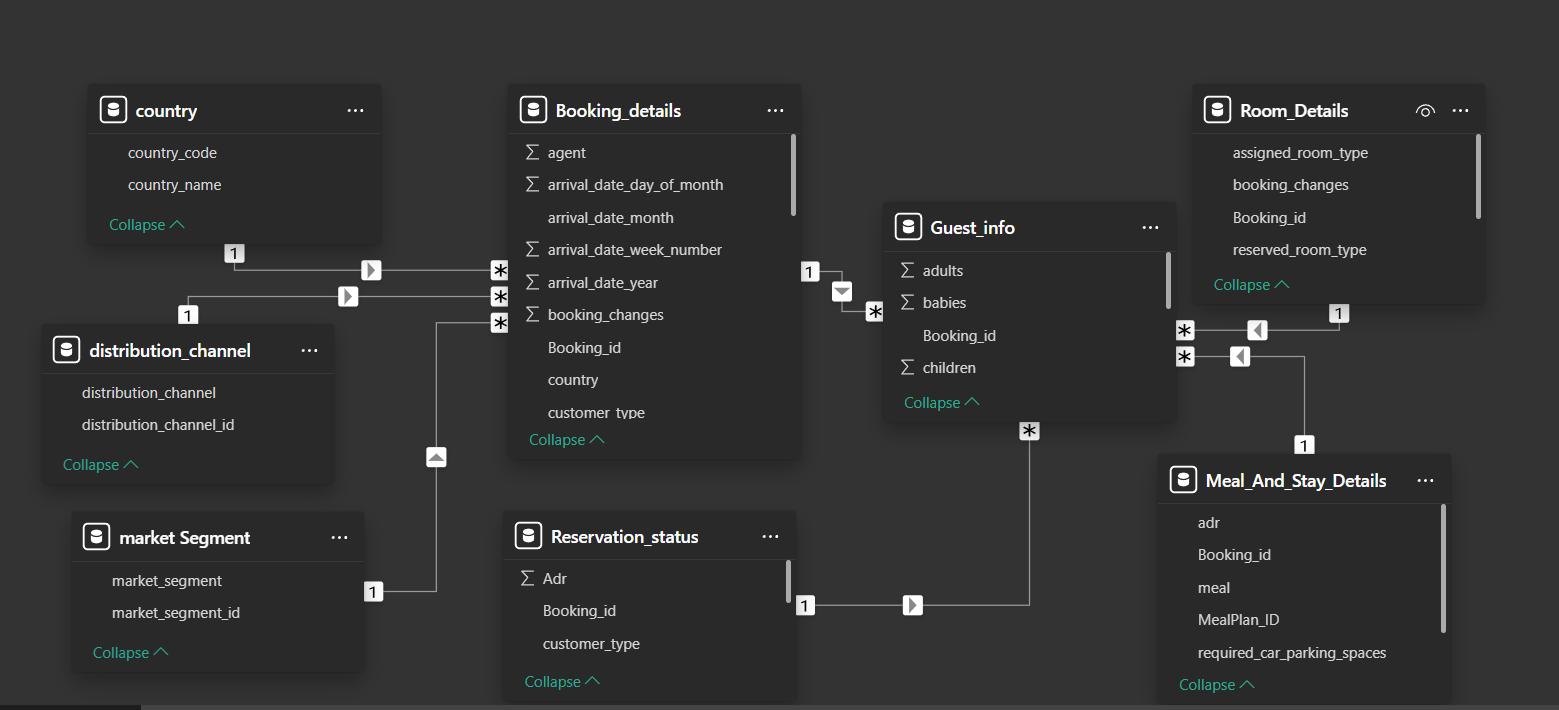
1. **Meal\_And\_Stay\_Details**

The Meal\_And\_Stay\_Details table complements the booking information by specifying meal-related and stay-related attributes. It connects to the Booking\_Details table via the booking identifier. This table includes the type of meal booked (e.g., Bed & Breakfast, Half Board), the Average Daily Rate (ADR) for the stay, the number of required car parking spaces, and the total count of special requests made by the guest.

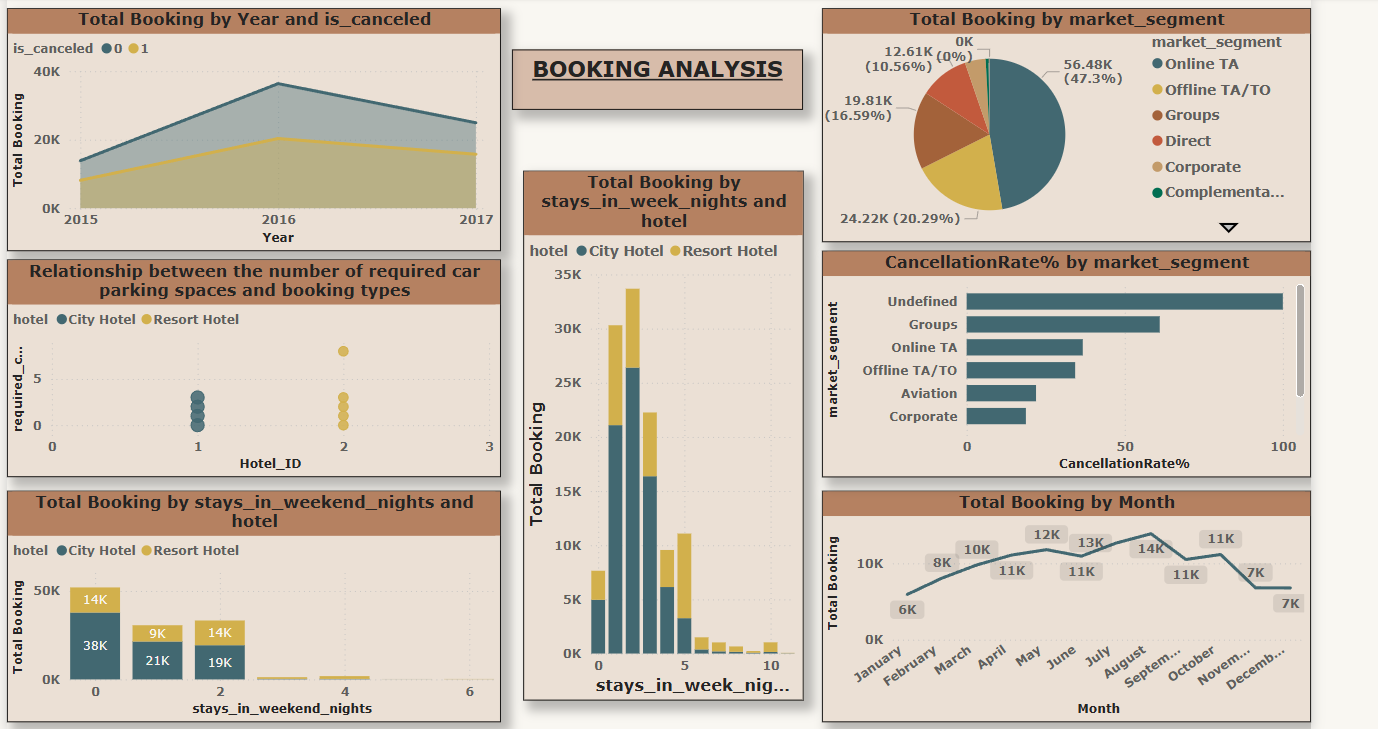
1. **Booking\_Source\_and\_History**

The Booking\_Source\_and\_History table is crucial for understanding the source of bookings and the historical behavior of guests. It is connected to the Booking\_Details table via the booking identifier. This table encompasses information such as the market segment (e.g., Online Travel Agents, Direct Booking), distribution channel (e.g., Online Travel Agents, Direct Booking), and whether the guest is a repeated visitor (0 for not repeated, 1 for repeated). Additionally, it records the number of previous booking cancellations, previous bookings that were not canceled, the deposit type (e.g., No Deposit, Non-Refund), the booking agent's ID, the company's ID, the number of days a booking spent on the waiting list, and the customer type (e.g., Transient, Group).

ER DIAGRAM



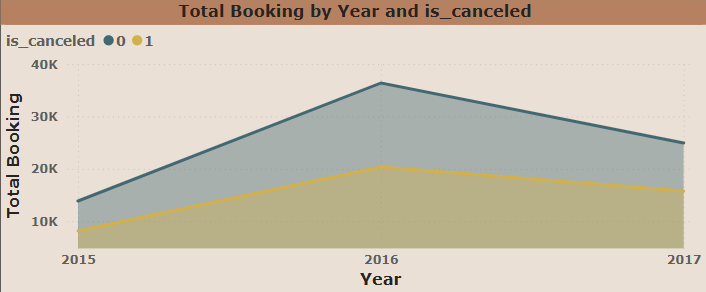
**POWER BI PROBLEM STATEMENT**



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

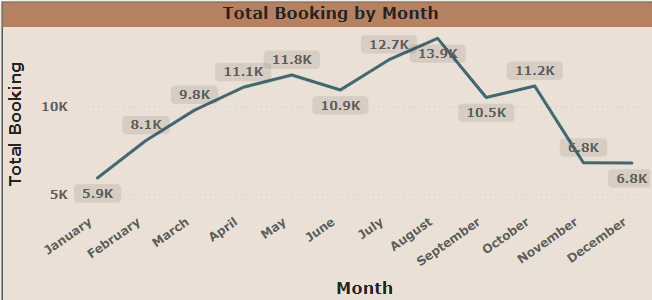
The provided area chart displays "Total Booking by Year and is\_canceled," illustrating booking trends and cancellation rates from 2015 to 2017. The blue/grey area represents bookings that were not canceled ("is\_canceled" = 0), while the yellow/brown area indicates canceled bookings ("is\_canceled" = 1).

In 2015, total bookings were relatively low, with a small proportion of cancellations. Both total bookings and cancellations saw a significant increase in 2016, reaching their peak for the observed period. In 2016, the number of non-canceled bookings was approximately 35,000, and canceled bookings were around 20,000. By 2017, there was a noticeable decline in both total bookings and cancellations. While the overall number of bookings decreased from 2016 to 2017, the proportion of canceled bookings appears to have remained substantial relative to non-canceled bookings. This suggests that while the booking volume fluctuated, cancellations continued to be a significant factor throughout the years, peaking in 2016.



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

The line chart illustrates the "Total Booking by Month," showing a clear seasonality in booking patterns throughout the year. Starting from a relatively low point of 5.9K in January, total bookings steadily increase, reaching 8.1K in February and 9.8K in March. April and May continue this upward trend with 11.1K and 11.8K bookings respectively, indicating a strong spring booking season. There's a slight dip in June to 10.9K, but bookings rebound sharply in July and August, reaching their peak at 12.7K and 13.9K respectively, suggesting a robust summer booking period. Following this peak, bookings begin to decline, dropping to 10.5K in September and 11.2K in October. The decline becomes more significant towards the end of the year, with 6.8K bookings in both November and December, indicating a slower winter booking period. Overall, the chart highlights that the peak booking months are during the summer (July and August), while the beginning and end of the year (January, February, November, and December) experience the lowest booking volumes.



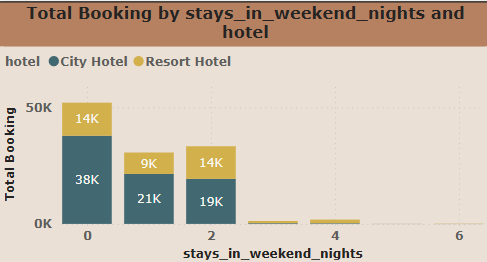
**Compare stays in weekend nights and weekday nights to determine preferences and variations by hotel type.**

The two stacked bar charts illustrate booking patterns based on the type of hotel (City Hotel or Resort Hotel) and the number of weekend nights (stays\_in\_weekend\_nights) and week nights (stays\_in\_week\_nights).

For weekend stays, the majority of bookings are for 0 weekend nights, meaning guests primarily book stays that do not include a weekend. Within this category, City Hotels account for a significantly higher proportion of bookings (38K) compared to Resort Hotels (14K). As the number of weekend nights increases to 1 and 2, the total booking volume decreases, and while City Hotels still lead in bookings for 1 weekend night (21K vs 9K), Resort Hotels see a slight increase in their share for 2 weekend nights (14K vs 19K). Bookings for 3 or more weekend nights are negligible for both hotel types.

Conversely, for weekday stays, the highest booking volume occurs for 1 and 2 week nights. City Hotels consistently dominate bookings across almost all weekday night categories. For 1 week night, City Hotels have 26K bookings compared to 9K for Resort Hotels, and a similar trend is observed for 2 week nights. Bookings for 0 week nights are also present, with City Hotels having 5K bookings. As the number of week nights increases beyond 2, total bookings decline significantly for both hotel types, though City Hotels maintain a larger share up to around 5 week nights. Bookings for very long week stays (e.g., 6 or more nights) are minimal, similar to long weekend stays.

In summary, City Hotels are overwhelmingly more popular for both weekend and weekday stays, particularly for shorter durations. The data suggests that guests frequently book City Hotels for non-weekend stays and single weekend night stays, while Resort Hotels see a comparatively smaller, though notable, share, especially for 2 weekend nights. Generally, longer stays, whether on weekends or weekdays, are less common for both hotel types.



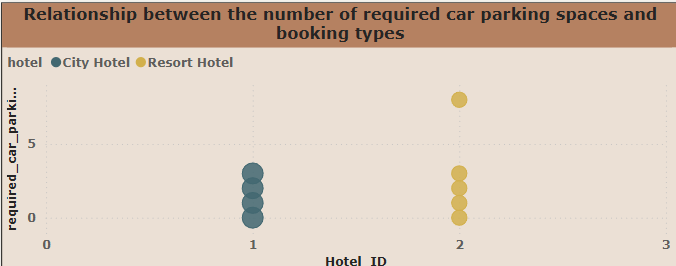


**Visualize the relationship between the number of required car parking spaces and booking types (Resort Hotel vs. City Hotel).**

The scatter plot illustrates the "Relationship between the number of required car parking spaces and booking types," distinguishing between City Hotels and Resort Hotels (though the x-axis is labeled "Hotel\_ID" with values 1 and 2, implying these correspond to the two hotel types).

For City Hotels (Hotel\_ID 1, represented by blue/grey dots), all observed bookings indicate that 0 car parking spaces were required. This suggests that guests booking City Hotels generally do not require parking, perhaps due to urban locations with good public transport access or different transportation preferences.

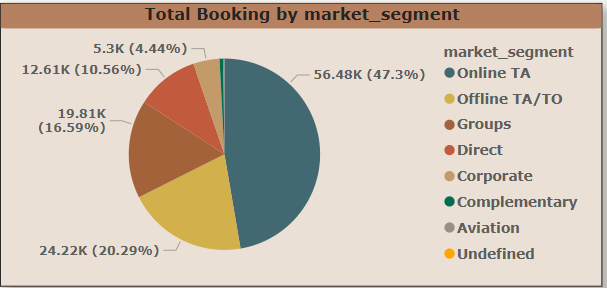
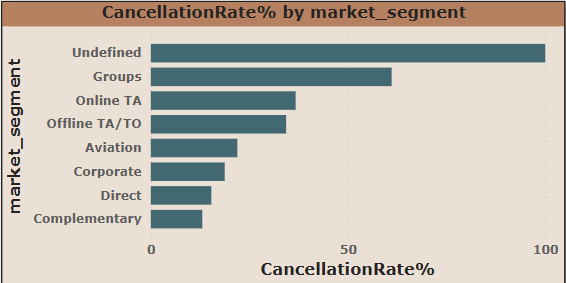
In contrast, for Resort Hotels (Hotel\_ID 2, represented by yellow/brown dots), while some bookings also show 0 required car parking spaces, there are instances where 1 car parking space was required. There is also one outlier instance where 7 car parking spaces were required for a Resort Hotel booking. This indicates that guests at Resort Hotels are more likely to require parking, which is consistent with resort locations often being less accessible by public transport and encouraging private vehicle use. The single instance of 7 required parking spaces could point to a group booking or an event at a Resort Hotel. Overall, the chart highlights a clear difference in parking requirements between the two hotel types, with Resort Hotels having a greater demand for parking.

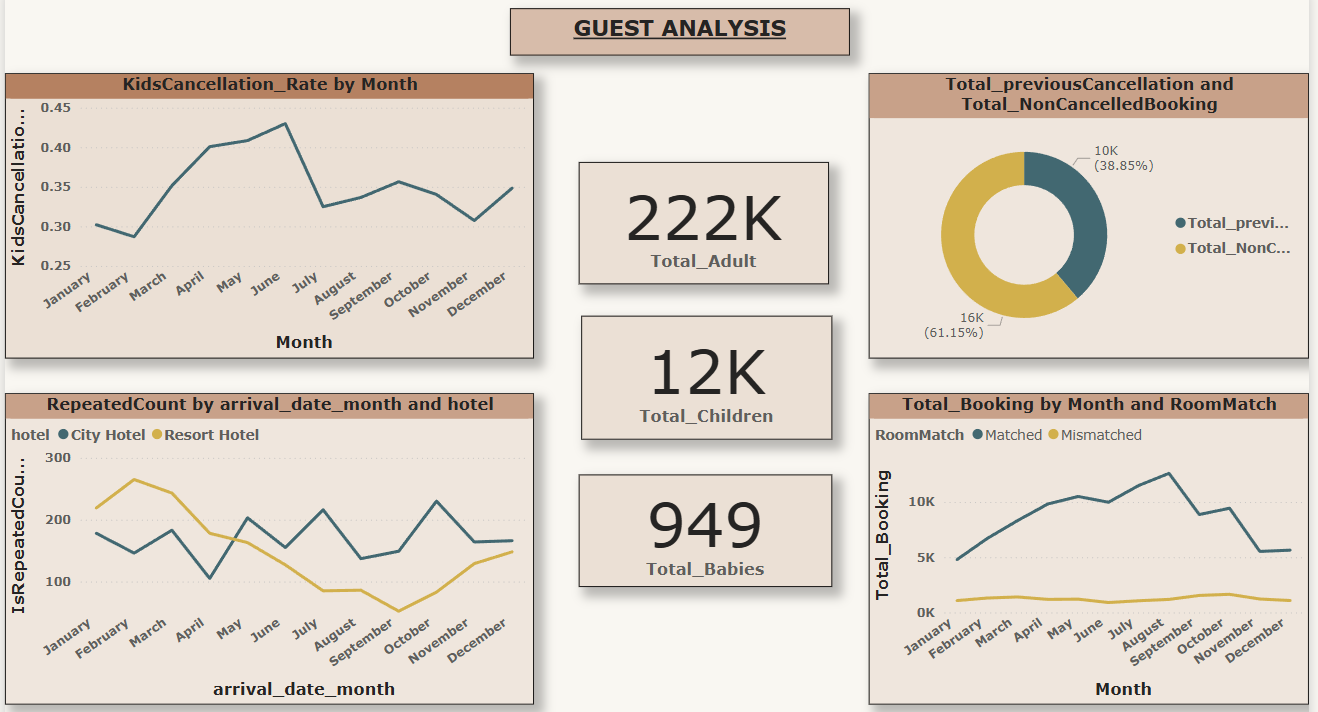


**Visualize booking distribution across different market segments and analyze cancellation rates within each segment.**

The provided charts offer insights into booking distribution and cancellation rates across different market segments. The pie chart, "Total Booking by market\_segment," reveals that "Online TA" (Online Travel Agencies) is the dominant booking channel, accounting for a substantial 47.3% (56.48K) of total bookings. This is followed by "Offline TA/TO" (Offline Travel Agencies/Tour Operators) at 20.29% (24.22K) and "Groups" at 16.59% (19.81K). "Direct" bookings represent 10.56% (12.61K), while other segments like "Corporate," "Complementary," "Aviation," and "Undefined" contribute smaller percentages.

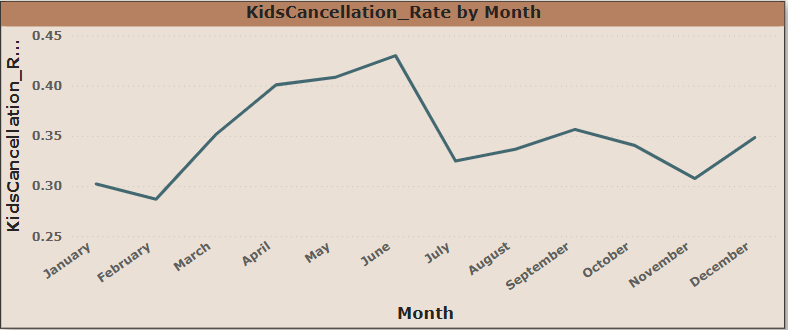
Complementing this, the "CancellationRate% by market\_segment" bar chart highlights the varying cancellation behaviors across these segments. Notably, the "Undefined" market segment has an extremely high cancellation rate, approaching 100%. "Groups" also exhibit a high cancellation rate, exceeding 50%. While "Online TA" is the largest booking segment, its cancellation rate is moderate, around 30-40%. Similarly, "Offline TA/TO" also shows a moderate cancellation rate, slightly lower than "Online TA." "Aviation," "Corporate," and "Direct" segments generally have lower cancellation rates, typically below 25%, with "Complementary" having the lowest cancellation rate among all. This combined view suggests that while online and offline travel agencies drive the majority of bookings, significant cancellation risks are associated with "Undefined" and "Groups" segments, necessitating closer monitoring and potentially different strategies for these high-risk areas.





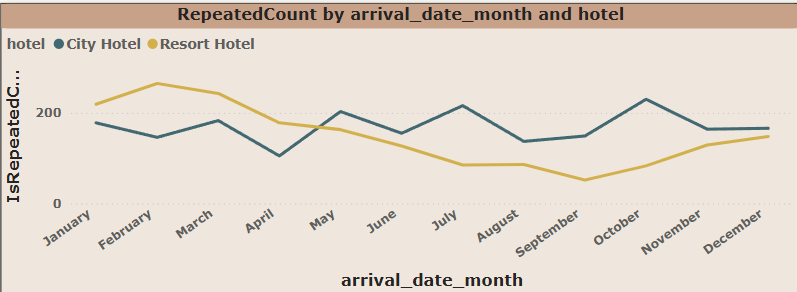
**Visualize the distribution of adults, children, and babies in bookings. Explore the impact of children and babies on cancellation rates.**

The line chart "KidsCancellation\_Rate by Month" illustrates the fluctuating cancellation rates for bookings involving children throughout the year. Starting in January with a cancellation rate of approximately 0.30, there's a slight dip in February before a noticeable increase through March, April, and May, reaching a peak of around 0.43 in June. This suggests that cancellations for bookings with children are highest during the early summer months. Following this peak, there's a sharp decrease in July, bringing the rate down to around 0.32. The cancellation rate then experiences a moderate rise in August and September, before gradually declining through October and November, reaching its lowest point in November at roughly 0.30. Finally, there's a slight uptick again in December. Overall, the chart highlights that bookings with children are most prone to cancellation in late spring and early summer (April to June), with lower cancellation rates observed at the beginning and end of the year, as well as a significant drop in July.



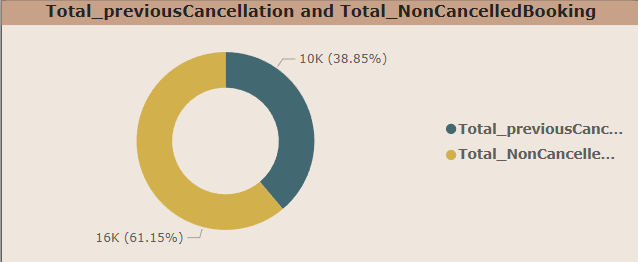
**Visualize the percentage of repeated guests for each hotel type (Resort Hotel vs. City Hotel) over time. Explore factors influencing guest retention.**

The line chart titled "RepeatedCount by arrival\_date\_month and hotel" illustrates the trend of repeated bookings for City Hotels and Resort Hotels across different months of arrival. For Resort Hotels, the repeated booking count starts relatively high in January and February, reaching a peak in March, and then generally declines through the summer and early autumn months, hitting its lowest point around September. It shows a slight recovery towards the end of the year. In contrast, City Hotels show a more fluctuating pattern. Their repeated booking count is moderate in January and February, dips in April, but then sees an increase in May. While Resort Hotels decline, City Hotels experience an increase in repeated bookings in July and October. Generally, Resort Hotels tend to have a higher "RepeatedCount" in the first quarter of the year, while City Hotels show more consistent or even increasing repeated bookings during mid-year and towards the autumn, especially in July and October, where they surpass Resort Hotels. This suggests that repeat guests might favor resort hotels more in the early part of the year, potentially for winter getaways, while city hotels maintain a more consistent or even growing appeal for repeat visitors through the warmer months and into the fall.



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

The doughnut chart illustrates the proportion of "Total\_previousCancellation" versus "Total\_NonCancelledBooking." The chart clearly shows that "Total\_NonCancelledBooking" accounts for the larger share, representing 16K bookings or 61.15% of the total. In contrast, "Total\_previousCancellation" makes up 10K bookings, which is 38.85% of the total. This indicates that while a significant portion of previous interactions involved cancellations, a greater majority ultimately resulted in non-canceled bookings. This suggests a relatively healthy booking success rate, although the 38.85% previous cancellation rate highlights an area that could potentially be analyzed further to understand underlying reasons and reduce future cancellations.

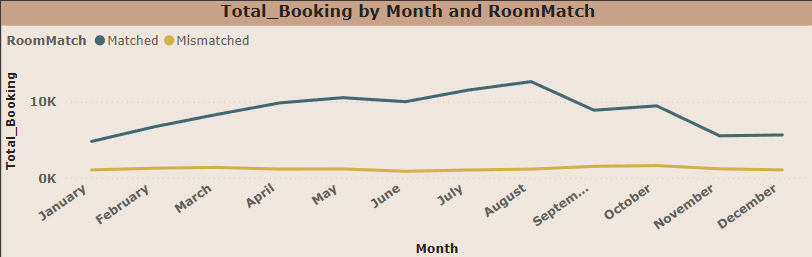


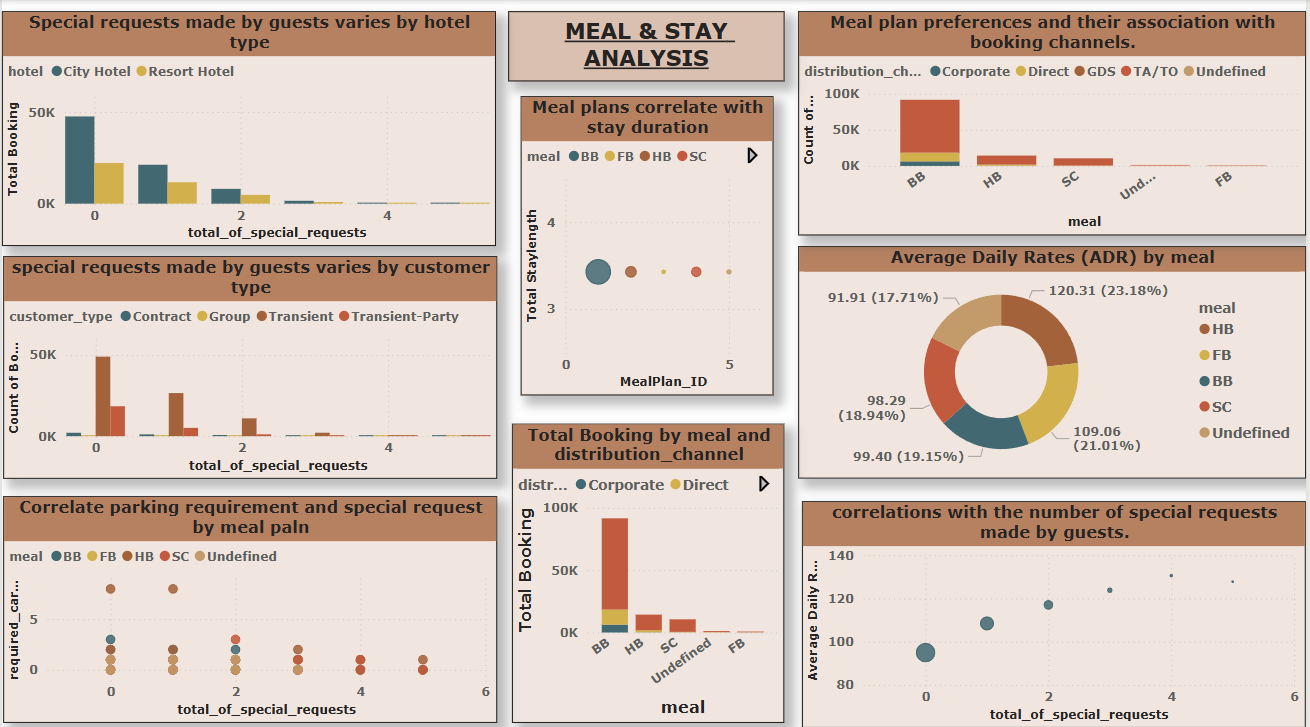
**Analyze whether guests who make multiple bookings tend to consistently request the same room type or if their preferences change over time.**

The line chart titled "Total\_Booking by Month and RoomMatch" illustrates the total number of bookings each month, categorized by whether the assigned room matched the reserved room type ("Matched") or not ("Mismatched").

A clear and consistent trend emerges: the vast majority of bookings consistently have a "Matched" room. The "Matched" bookings (blue line) show a general upward trend from January, peaking in August with over 10,000 bookings, before gradually declining towards the end of the year, with a noticeable drop in November and December. This pattern suggests a strong seasonal influence on overall booking volume.

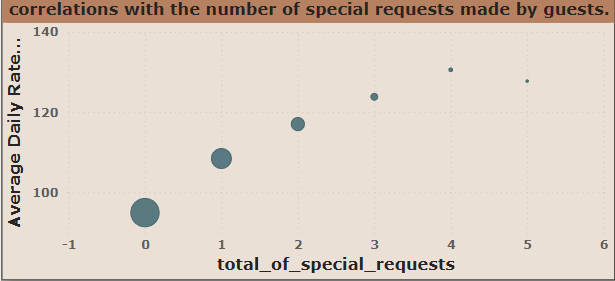
In stark contrast, "Mismatched" bookings (yellow line) remain consistently low throughout the entire year, fluctuating only slightly and always staying well below 2,000 bookings per month. This indicates that instances of room type mismatches are a rare occurrence, suggesting effective room assignment processes. The volume of mismatched rooms does not appear to follow the same strong seasonal pattern as matched bookings, remaining relatively flat. Overall, the chart demonstrates a high success rate in providing guests with their reserved room types, with mismatches being a minor issue that does not significantly contribute to the overall booking volume trends.





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

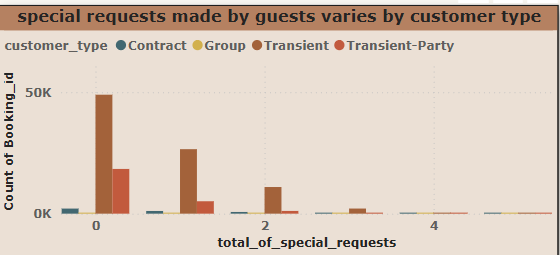
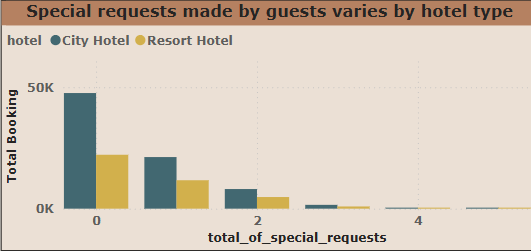
The bubble chart illustrates the correlation between the "Average Daily Rate" and the "total\_of\_special\_requests" made by guests. There appears to be a positive correlation between the two variables. As the number of special requests increases from 0 to 4, the average daily rate also generally increases. For example, bookings with 0 special requests have an average daily rate below 100. This rate steadily rises as the number of special requests goes up, reaching over 130 for 4 special requests. The size of the bubbles indicates the volume of bookings at each point; for instance, the largest bubble is at 0 special requests, suggesting that most bookings do not involve special requests. While the average daily rate continues to rise with more special requests up to 4, there's a slight decrease in the average daily rate when 5 special requests are made, and the bubble size also appears smaller for higher numbers of special requests (4 and 5), indicating fewer bookings with a very high number of special requests. Overall, guests who make more special requests tend to pay a higher average daily rate, with the bulk of bookings having few to no special requests.



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

The two bar charts provide insights into how special requests vary by hotel type and customer type. The "Special requests made by guests varies by hotel type" chart shows that for both City Hotels and Resort Hotels, the vast majority of bookings involve zero special requests. City Hotels consistently have a higher number of bookings across all special request categories compared to Resort Hotels. As the number of special requests increases, the total number of bookings for both hotel types sharply decreases, indicating that guests rarely make more than two special requests.

The "Special requests made by guests varies by customer type" chart reveals that "Transient" customers make the highest number of bookings, and they are also the most frequent makers of special requests, particularly for 0, 1, and 2 requests. "Transient-Party" customers follow, but with significantly fewer bookings and special requests than "Transient" guests. "Group" and "Contract" customers make very few special requests, with most of their bookings having zero special requests. This suggests that individual, non-group travelers ("Transient") are more likely to have specific needs or preferences they communicate through special requests, regardless of whether they are staying at a City or Resort Hotel. Across both charts, the overarching trend is that a high volume of bookings have no special requests, and as the number of requests increases, the booking volume drastically drops, with City Hotels and Transient customer types being the primary sources of bookings with special requests.

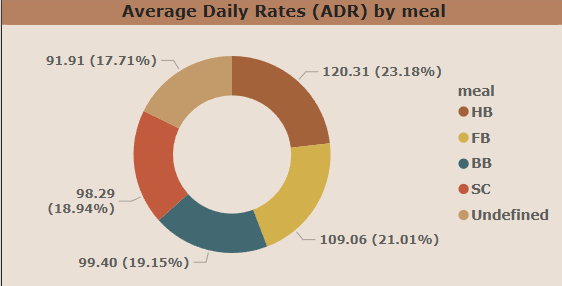
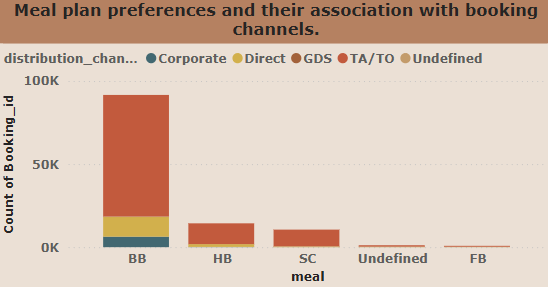


**Explore meal plans and their impact on Average Daily Rates (ADR). Analyze meal plan preferences and their association with booking channels.**

The provided charts offer insights into meal plan preferences, their association with booking channels, and their impact on Average Daily Rates (ADR). The bar chart "Meal plan preferences and their association with booking channels" clearly shows that "BB" (Bed & Breakfast) is by far the most popular meal plan, accounting for the highest count of bookings. Within BB bookings, "TA/TO" (Travel Agent/Tour Operator) is the dominant distribution channel, followed by "Direct" bookings, and then "Corporate" and "GDS" (Global Distribution System). "HB" (Half Board) and "SC" (Self-Catering) are the next most popular meal plans, but with significantly fewer bookings compared to BB, and TA/TO continues to be a major distribution channel for these as well. "Undefined" and "FB" (Full Board) meal plans have very minimal bookings.

The doughnut chart "Average Daily Rates (ADR) by meal" indicates the average daily rates associated with each meal plan. "HB" has the highest average daily rate at 120.31, representing 23.18% of the total ADR distribution. "FB" follows with an ADR of 109.06 (21.01%), while "BB" has an ADR of 99.40 (19.15%). "SC" and "Undefined" meal plans have lower ADRs at 98.29 (18.94%) and 91.91 (17.71%) respectively.

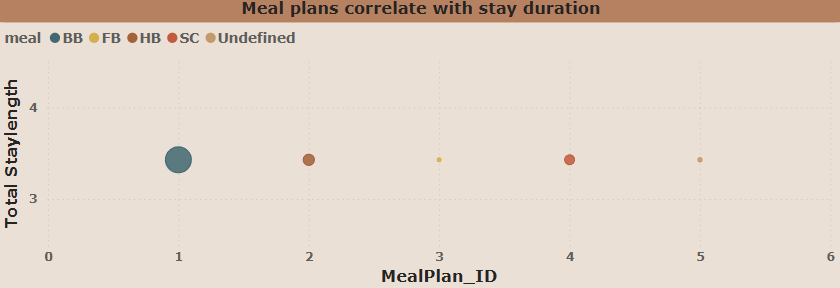
In summary, while Bed & Breakfast (BB) is the most frequently booked meal plan, particularly through travel agents, Half Board (HB) commands the highest Average Daily Rate. This suggests that even though fewer guests opt for HB, hotels generate more revenue per booking from them. The dominance of TA/TO as a distribution channel across most meal plans highlights their crucial role in securing bookings.



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

The bubble chart titled "Meal plans correlate with stay duration" illustrates the relationship between different meal plans (represented by MealPlan\_ID) and the average total stay length. It appears that while the average total stay length hovers around 3 to 3.5 nights across various meal plans, there isn't a strong, clear linear correlation where a specific meal plan consistently leads to significantly longer or shorter stays.

The largest bubble, representing "BB" (Bed & Breakfast) with MealPlan\_ID 1, indicates that it's associated with a stay length of approximately 3.3 nights. "HB" (Half Board), "SC" (Self-Catering), and "FB" (Full Board) all seem to cluster around a similar average stay length of just over 3 nights. The "Undefined" meal plan also falls within this narrow range. The chart suggests that regardless of the meal plan chosen, guests tend to book for similar durations, with "BB" being the most popular choice based on the size of its bubble (implying a higher volume of bookings), despite not showing a notably different average stay duration compared to other meal plans. This implies that the choice of meal plan may not be a primary driver for the length of stay.



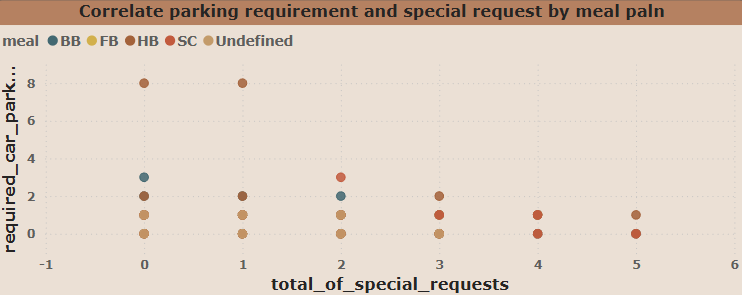
**Correlate parking requirements and special requests with different meal plans. Determine if certain meal plans result in more requests or parking needs.**

The scatter plot titled "Correlate parking requirement and special request by meal plan" explores the relationship between the number of required car parking spaces, the total number of special requests, and different meal plans.

The most prominent observation is that the majority of data points, regardless of meal plan, cluster around 0 required car parking spaces, especially when the total number of special requests is low (0, 1, or 2). This suggests that most guests do not require car parking, or at least not more than one space, and this holds true even if they have some special requests.

However, there are notable exceptions. For both 0 and 1 special request, bookings with the "Undefined" meal plan show a requirement for 8 car parking spaces. This is a significant outlier compared to other meal plans and suggests that "Undefined" meal plan bookings, when they occur, can be associated with a very high parking demand, possibly indicating group bookings or events.

For other meal plans like "BB" (Bed & Breakfast), "FB" (Full Board), and "SC" (Self-Catering), the required car parking spaces generally range from 0 to 3, even as the number of special requests increases. There's no clear linear trend indicating that more special requests lead to a higher parking requirement. The data points become sparser as the total number of special requests increases beyond 2, implying fewer bookings fall into these categories.

In summary, while most bookings, irrespective of meal plan or a few special requests, do not require car parking, the "Undefined" meal plan stands out with a significant demand for parking, potentially for large groups. For other meal plans, parking requirements remain low even with some special requests.

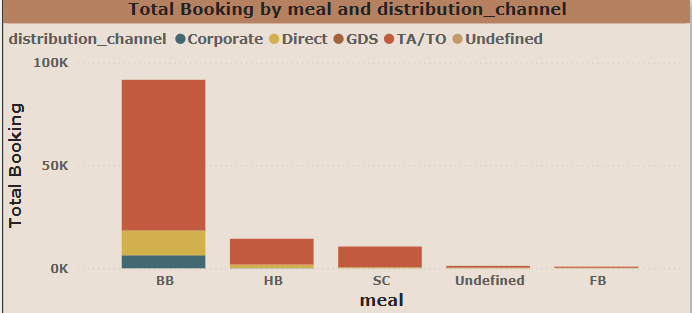
**Explore how meal plans are distributed across various booking channels. Analyze if certain channels are associated with specific meal plans.**

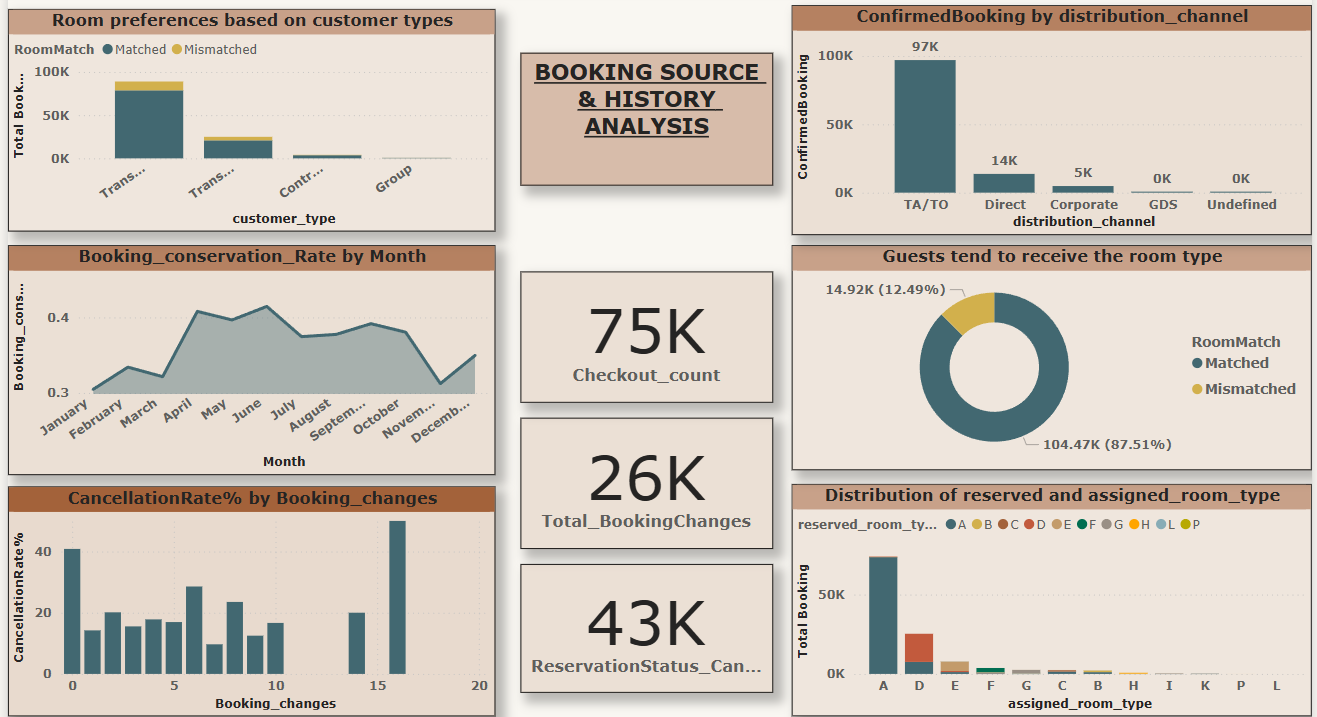
The stacked bar chart titled "Total Booking by meal and distribution\_channel" illustrates the volume of bookings for different meal plans and the distribution channels through which these bookings are made.

The chart clearly shows that "BB" (Bed & Breakfast) is overwhelmingly the most popular meal plan, accounting for the highest total bookings, significantly surpassing all other meal types. Within the "BB" category, "TA/TO" (Travel Agent/Tour Operator) is the dominant distribution channel, responsible for the largest portion of bookings, followed by "Direct" bookings and then "Corporate" bookings.

"HB" (Half Board) and "SC" (Self-Catering) are the next most booked meal plans, but their total booking volumes are substantially lower than "BB". For both "HB" and "SC," "TA/TO" remains the primary distribution channel, though the relative contribution of "Direct" bookings appears to be proportionally higher for HB compared to BB. "GDS" (Global Distribution System) contributes a very small amount across all meal plans where it is visible.

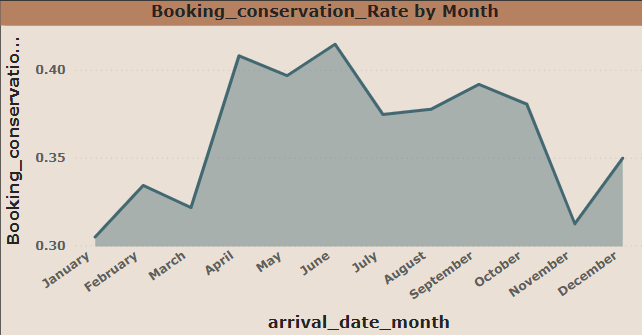
Finally, "Undefined" and "FB" (Full Board) meal plans have a very minimal number of total bookings, indicating low demand for these options.

In summary, the hotel relies heavily on the "TA/TO" channel for securing bookings, particularly for the highly popular "BB" meal plan. While "HB" and "SC" also contribute to bookings, their volume is much lower, and "Undefined" and "FB" meal plans are practically negligible in terms of total bookings.



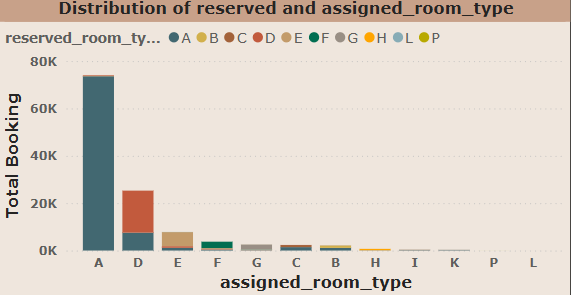
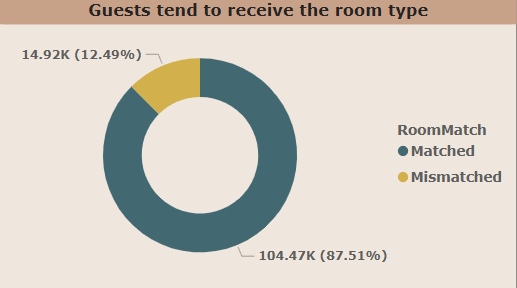
**Calculate and visualize the booking conversion rate (canceled bookings to total bookings) over time.**

The provided chart illustrates the "Booking\_conservation\_Rate" across different months of the year. The rate starts relatively low in January, hovering around 0.30, and sees a slight increase in February before dipping in March. A significant surge occurs in April, where the rate peaks around 0.41, maintaining a high level through May and June, with June showing the highest point just above 0.41. Following this peak, there's a noticeable decline in July, and the rate stabilizes at a slightly lower level through August and September, before a further drop in October and November. The year concludes with a noticeable recovery in December, with the rate climbing back towards 0.35. Overall, the booking conservation rate demonstrates a strong seasonality, peaking in the spring/early summer months (April-June) and experiencing a dip towards the end of the year, with a slight recovery in December.



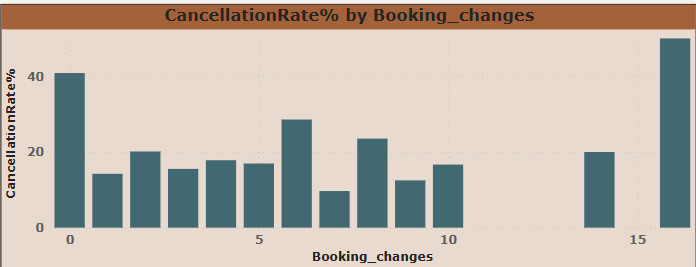
**Visualize the distribution of reserved and assigned room types. Analyze whether guests tend to receive the room type they initially reserved.**

The two charts combined offer insights into room allocation and guest satisfaction regarding room types. The first chart, "Distribution of reserved and assigned\_room\_type," reveals that room type 'A' is overwhelmingly the most frequently reserved and assigned room, accounting for a significant majority of total bookings (over 70,000). Room type 'D' is the second most common, with around 25,000 bookings, followed by 'E', 'F', and 'G' in decreasing order of frequency, with other room types having minimal bookings. Importantly, within each assigned room type, the dominant color (likely representing the reserved room type 'A' or 'D') indicates that guests generally receive the room type they reserved. This observation is strongly supported by the second chart, "Guests tend to receive the room type," which shows that a vast majority of guests, 87.51% (104.47K), receive a "Matched" room type, meaning their assigned room matches their reserved room. Only a small percentage, 12.49% (14.92K), experience a "Mismatched" room, where the assigned room differs from the reserved one. This suggests a highly efficient and accurate room assignment process, particularly for the most popular room types like 'A' and 'D'.



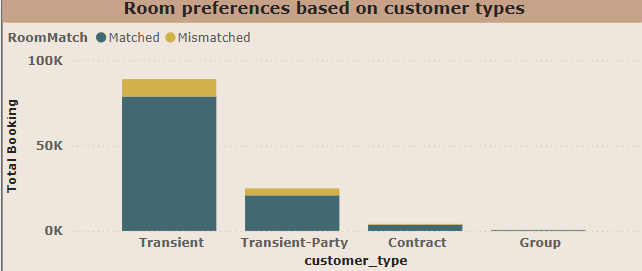
**Investigate the relationship between the number of booking changes made by guests and their likelihood of canceling a booking.**

The chart displays the Cancellation Rate % based on the number of booking changes. It reveals a strong correlation between the number of changes made to a booking and the likelihood of cancellation. Notably, bookings with zero changes have a high cancellation rate of approximately 40%. As the number of booking changes increases, the cancellation rate generally fluctuates but tends to decrease significantly compared to bookings with no changes, though there's a slight peak around 6 changes (nearly 30% cancellation rate) and another around 8 changes (around 23%). Interestingly, for bookings with a very high number of changes (e.g., around 16), the cancellation rate spikes dramatically, reaching almost 50%. This suggests that while minor adjustments might reduce cancellation risk compared to rigid initial bookings, an excessive number of changes could indicate indecision or evolving circumstances that ultimately lead to cancellation.



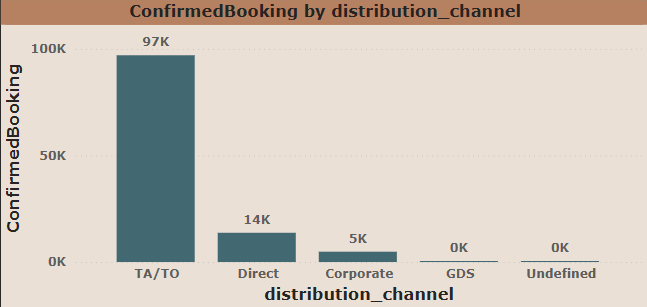
**Analyze room type preferences based on customer types (e.g., Transient, Group) and identify any patterns in room type selection.**

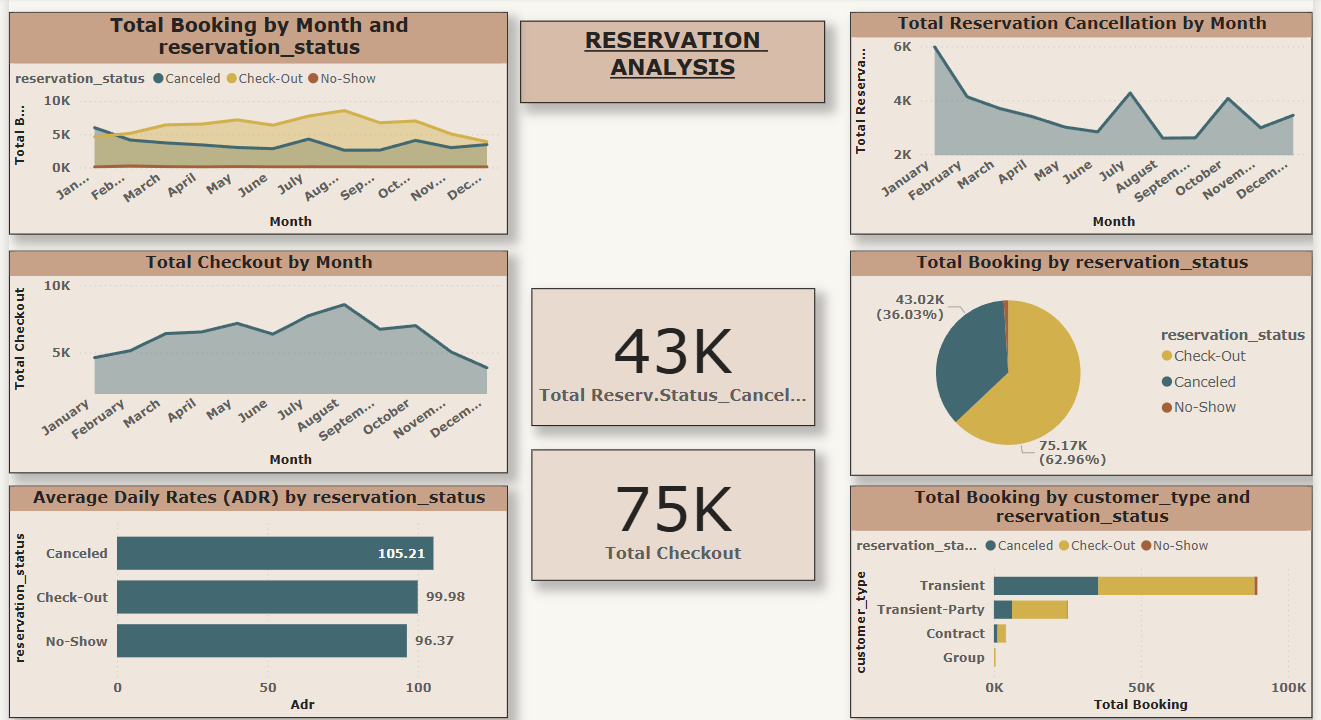
The chart illustrates room preferences, specifically the "RoomMatch" (matched vs. mismatched rooms), across different customer types, measured by "Total Booking." The "Transient" customer type accounts for the overwhelming majority of bookings, with well over 80,000 total bookings, and within this group, the vast majority of rooms are matched to their preferences, indicating a high level of satisfaction or accurate assignment. "Transient-Party" customers represent the second largest group, with total bookings around 20,000, and they also show a high proportion of matched rooms, though there appears to be a slightly larger absolute number of mismatches compared to "Transient" customers. "Contract" and "Group" customer types have significantly fewer bookings, both below 5,000. While their overall booking numbers are low, the chart suggests that even for these smaller segments, a high percentage of their room preferences are met, as indicated by the dominant "Matched" portion of their respective bars. In summary, across all customer types, the system appears to be highly effective in assigning guests to their preferred room types, with "Transient" customers being the largest and most consistently satisfied segment in terms of room matching.



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

The chart illustrates the volume of "ConfirmedBooking" across various "distribution\_channel" categories. It clearly shows that the TA/TO (Travel Agent/Tour Operator) channel is by far the most dominant source of confirmed bookings, accounting for a massive 97,000 bookings. This figure is significantly higher than any other channel. The Direct channel comes in second, with a modest 14,000 confirmed bookings, followed by the Corporate channel, which generated 5,000 bookings. Both GDS (Global Distribution System) and Undefined channels contribute negligible amounts to confirmed bookings, both registering 0K. This indicates a strong reliance on travel agents and tour operators for generating confirmed bookings, highlighting their crucial role in the booking ecosystem.

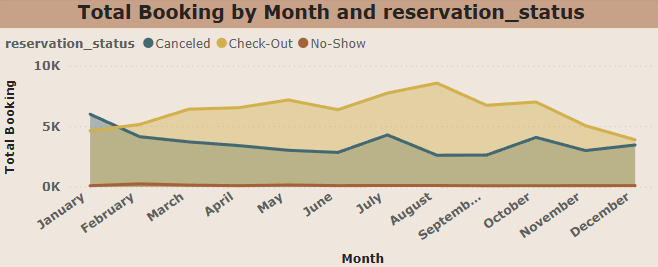
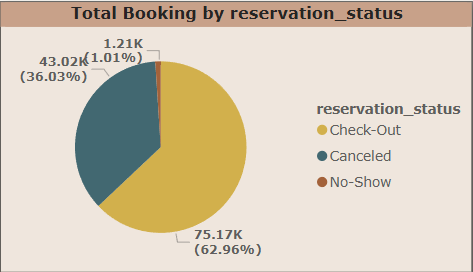




**Provide an overview of reservation statuses over time, including the percentage of canceled, checked-out, and No show bookings.**

The two charts together provide a comprehensive view of booking outcomes and their monthly distribution. The first pie chart, "Total Booking by reservation\_status," clearly indicates that the majority of bookings, specifically 62.96% (75.17K), result in a "Check-Out," meaning the guest completed their stay. "Canceled" bookings represent a significant portion at 36.03% (43.02K), while "No-Show" bookings are a small minority at 1.01% (1.21K).

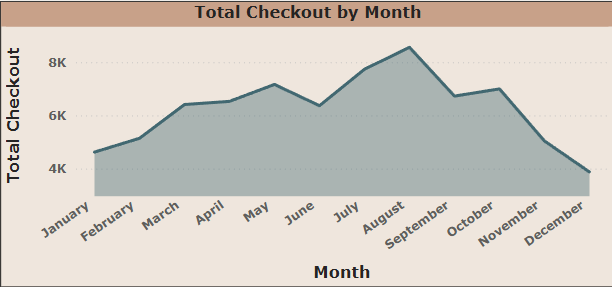
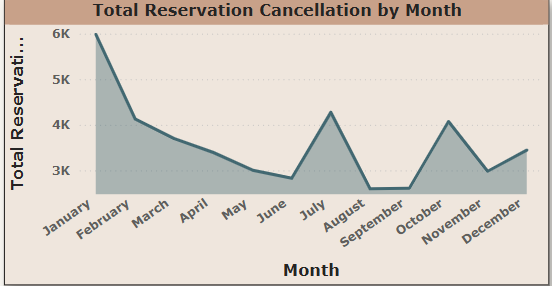
The second chart, "Total Booking by Month and reservation\_status," breaks down these reservation statuses by month. It shows a consistent pattern where "Check-Out" bookings generally outnumber "Canceled" bookings throughout the year, peaking in the summer months, particularly July and August. "Canceled" bookings, while substantial, remain below "Check-Out" numbers in most months, with a notable peak in January and another smaller peak in October. "No-Show" bookings remain consistently very low across all months. This combined view highlights that despite a significant cancellation rate overall, the business successfully converts a majority of its bookings into completed stays, with the summer months being particularly strong for check-outs.



**Analyze trends in reservation status dates, such as the busiest checkout dates or patterns in cancellations by month.**  
The two charts, "Total Reservation Cancellation by Month" and "Total Checkout by Month," provide a contrasting view of booking outcomes throughout the year. The cancellation chart shows a very high number of cancellations in January, exceeding 6,000, which then steadily decreases through the spring and early summer, reaching a low point around 2,800 in June. There's a notable spike in cancellations in July and another in October, both exceeding 4,000, before declining towards the end of the year.

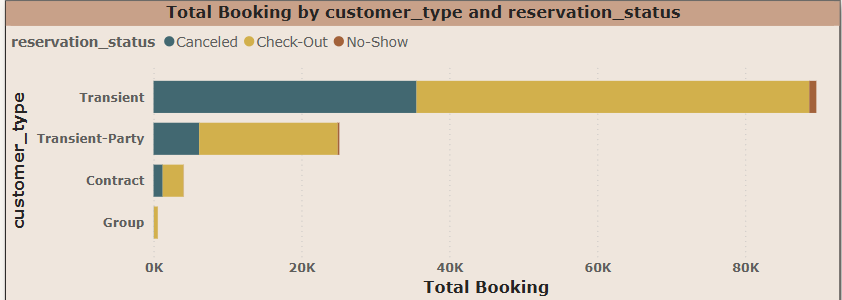
Conversely, the "Total Checkout by Month" chart indicates a different trend for successful stays. Checkouts start around 4,500 in January and generally show an increasing trend, peaking significantly in August with over 8,000 checkouts. While there's a dip in September, checkouts remain relatively strong through October and November before decreasing towards the year-end.

In summary, the combined view suggests a period of high cancellations early in the year (January-February) that gradually subsides as the year progresses, indicating that many bookings made for later in the year are less likely to be cancelled. Concurrently, the number of successful checkouts steadily increases, reaching its zenith in the summer months (July-August), which likely corresponds to peak travel seasons. The spikes in cancellations in July and October might be related to changes in travel plans around summer holidays or autumn events, despite high checkout volumes during these periods. This dual trend highlights distinct seasonal patterns for both booking cancellations and completed stays.



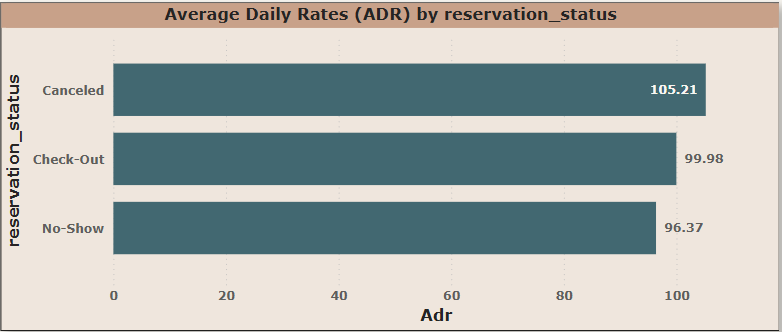
**Visualize how reservation statuses vary across different customer types (e.g., Transient, Group) and identify if certain customer types are more likely to result in cancellations or no-shows.**

The chart displays "Total Booking" segmented by "customer\_type" and "reservation\_status" (Canceled, Check-Out, No-Show). It clearly shows that "Transient" customers account for the overwhelming majority of total bookings, with a large proportion of these bookings resulting in "Check-Outs" (successful stays) and a significant, though smaller, portion being "Canceled." "No-Show" bookings for transient customers are negligible. "Transient-Party" customers represent the second largest booking group, also showing a higher proportion of "Check-Outs" compared to "Canceled" bookings, with minimal "No-Shows." "Contract" and "Group" customer types contribute very little to the total bookings in comparison, with "Contract" bookings having a small number of cancellations but mostly check-outs, and "Group" bookings being almost entirely "Check-Outs" with extremely few cancellations or no-shows. Overall, the data highlights the dominance of transient customers in total bookings and a general trend across all customer types where successful check-outs outweigh cancellations, though cancellations remain a notable factor, particularly for "Transient" bookings.



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

The chart displays the Average Daily Rates (ADR) categorized by reservation status. It reveals a clear pattern: bookings that are ultimately canceled have the highest average daily rate, at 105.21. This suggests that higher-priced bookings are more prone to cancellation. In contrast, bookings that result in a check-out (completed stays) have a slightly lower ADR of 99.98. The lowest average daily rate, at 96.37, is observed for no-show reservations. This implies that guests who do not show up for their bookings tend to have initially booked at a comparatively lower rate.

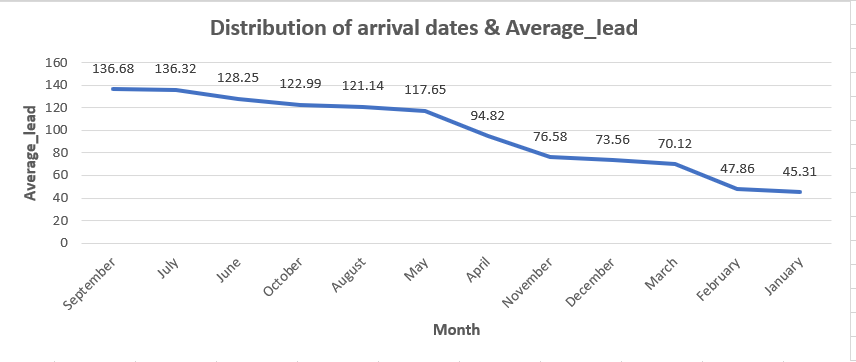


**EDA PROBLEM STATEMENT**

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

Peak Lead Times: September (136.68 days) and July (136.32 days) have the highest average lead times — meaning guests book their stays well in advance for these months.Summer & Early Fall: June, July, August, and September all have high lead times (120–137 days). This suggests these are peak or high-demand seasons, likely due to summer holidays and favorable weather.Moderate Lead Times: October and May also see relatively long lead times (118–123 days), indicating these shoulder months are planned in advance too.Short Lead Times: January (45.31 days) and February (47.86 days) have the lowest average lead times, showing bookings are often more last-minute after the holiday season.Off-Peak Months: March, November, and December have moderate to low lead times (70–76 days), possibly indicating off-peak or shoulder seasons with more spontaneous bookings.

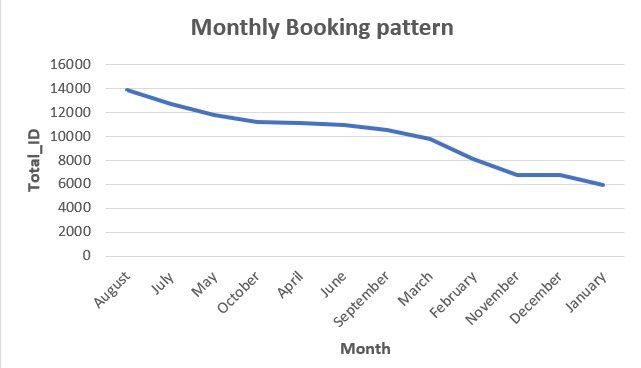
This pattern helps hotels forecast demand and manage pricing. They can expect longer booking windows for peak summer and early fall, and shorter windows in winter and early spring.Marketing campaigns and promotional offers can be planned accordingly — more aggressive last-minute deals for winter months, and early-bird offers for summer.



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

Peak Booking Months: August (13,877) and July (12,661) have the highest number of bookings, followed by May (11,791) and October (11,160). This suggests that summer and early autumn are the busiest months for guest arrivals.Strong Spring Performance: April and June also see high booking volumes (around 11,000), showing that spring and early summer are strong periods.Moderate Demand: September and March have moderate bookings (around 10,000), indicating stable activity during late summer and early spring.Lowest Booking Months: January (5,929), December (6,780), and November (6,794) have the lowest booking volumes, likely due to post-holiday slowdown and off-season travel trends. February is slightly higher but still low at 8,068.The pattern indicates clear seasonality — with peak demand in summer and low demand in winter.

Hotels can use this to plan staffing, inventory, promotions, and pricing — for example, higher rates and stricter cancellation policies during summer, and discounts or packages to boost occupancy in winter.Marketing efforts can focus on early booking promotions for summer months and last-minute deals or special events for off-peak months.

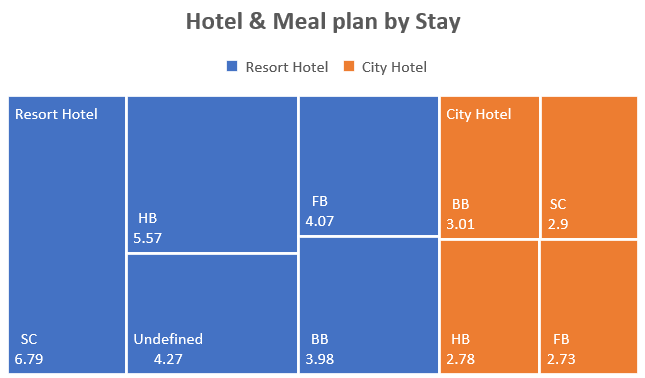


**Calculate the average length of stays for different hotel types and explore variations by meal plans.**

Resort Hotels:Guests staying at Resort Hotels tend to have higher average values (likely stays per guest, or spend per stay — depending on what the number represents — here it’s 6.79 for SC).City Hotels:City Hotel guests show lower average values across all meal types compared to Resort Hotels.

Resort Hotels attract longer-staying or higher-spending guests who prefer self-catering or meal plans — so focusing on diverse dining options, on-site kitchens, and flexible packages would align with guest preferences.

City Hotels guests appear to prefer basic inclusions (BB) and eat out more, so offering convenient breakfast options and local dining recommendations can enhance guest experience.This insight can help tailor marketing strategies, pricing, and service design for each hotel type.



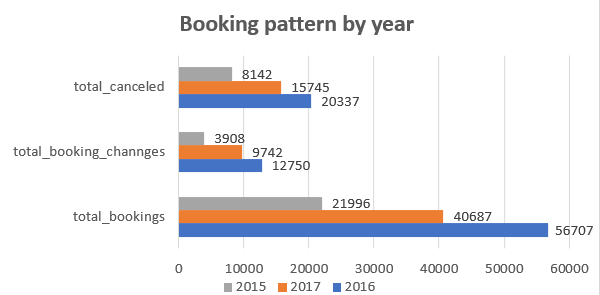
**Analyze how booking patterns have evolved over the years, including yearoveryear changes in bookings and cancellations.**

Growth Trend: Bookings increased significantly from 21,996 in 2015 to 56,707 in 2016 — a 158% increase, showing major growth.

Booking Changes:Total booking changes also rose with total bookings:3,908 in 2015 → 12,750 in 2016 → 9,742 in 2017.The number of changes remains high, showing that guests frequently modify their reservations — flexibility may be an important factor for customers.

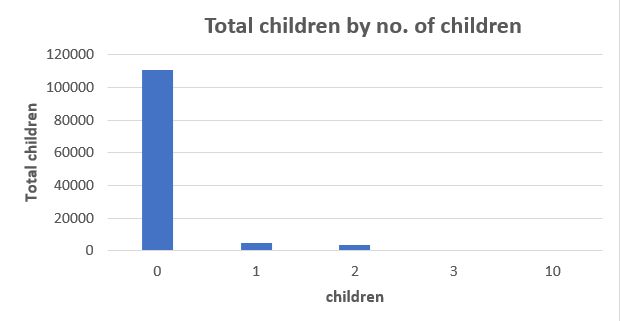
Cancellations:Cancellation volume is high across all years: 8,142 in 2015 (37% of bookings),20,337 in 2016 (36% of bookings),15,745 in 2017 (39% of bookings).The cancellation rate stays high (~36–39%), which could impact revenue predictability and operational planning.

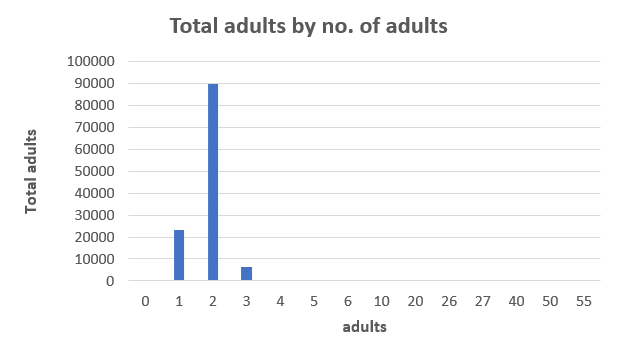
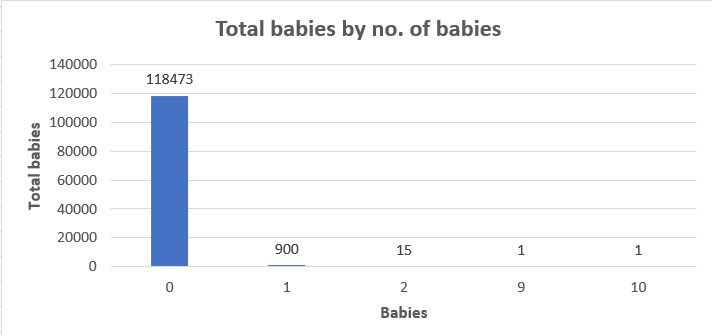
The large jump in bookings from 2015 to 2016 shows successful marketing or expansion but the drop in 2017 signals a need to analyze why bookings declined — maybe guest satisfaction, competition, or market shifts.High cancellation rates highlight the importance of clear cancellation policies, flexible booking options, or incentives for non-refundable bookings to protect revenue.Frequent changes suggest that flexibility and easy modification options are valued by guests — maintaining good customer support for changes could boost satisfaction.



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

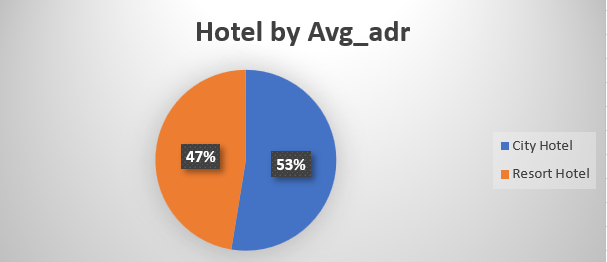
The majority of bookings are for 1–2 adults with no children or babies, indicating the hotel’s core market is solo travelers, couples, or business guests.Family stays (with children or babies) form a small segment of total bookings.There are some clear data outliers (e.g., 55 adults, 10 babies) — these should be cleaned or checked during data preprocessing, as they are unrealistic.Marketing and services can be tailored mainly towards couples and solo travelers, with family-friendly options as a smaller but important secondary focus.





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

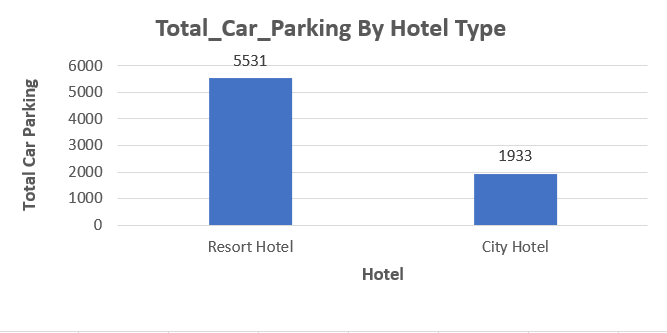
City Hotels have a higher average daily rate (ADR) at 105.3, compared to Resort Hotels at 94.95.This suggests City Hotels generally charge more per night, likely due to urban location advantages, business travelers, or higher demand in cities.Resort Hotels have slightly lower ADRs, possibly to attract longer leisure stays or family guests.City Hotels generate higher nightly revenue per room on average.



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

Resort Hotels have a much higher total car parking count (5,531) compared to City Hotels (1,933).This indicates that guests at Resort Hotels are more likely to arrive by car, probably due to remote or leisure destinations that require private transport.City Hotels have fewer parking slots, likely because guests use public transport, taxis, or walk due to urban location convenience.

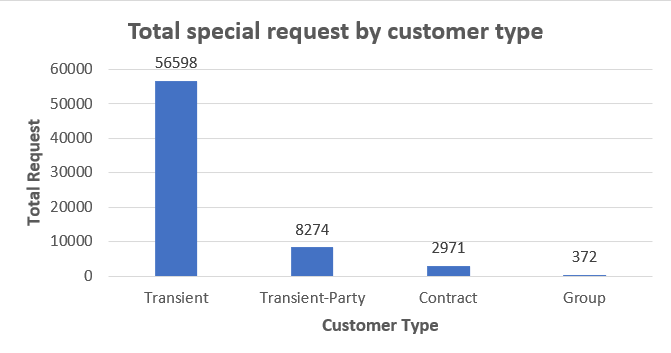
Parking is a more critical amenity for Resort Hotels than for City Hotels.



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

Transient guests have by far the most special requests (56,598) — these are individual travelers, so they likely have more personalized needs.Transient-Party guests have the next highest (8,274), reflecting that small leisure groups or families also make frequent requests.Contract customers (corporate or long-term) and Groups have far fewer special requests (2,971 and 372, respectively), likely because their stays are more standardized or managed through agreements.

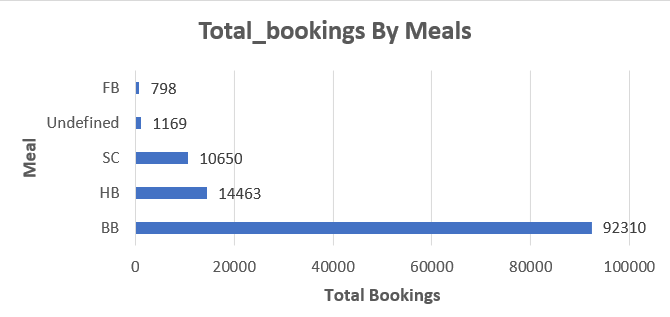
Individual travelers generate the majority of special requests, so hotels should focus on personalized service for them.



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

Resource Allocation: The establishment should prioritize resources (staffing, food purchasing, kitchen capacity) towards the BB option, as it generates the most demand.Menu and Service Customization: For HB and SC, understanding the specific needs of these guests can help tailor services or offerings (e.g., specific dinner menus for HB, well-equipped kitchenettes for SC).Pricing Strategy: The pricing for different meal plans should reflect their popularity and perceived value. There might be an opportunity to review the pricing and offerings for FB to make it more appealing, or conversely, to de-emphasize it if the demand is consistently low.

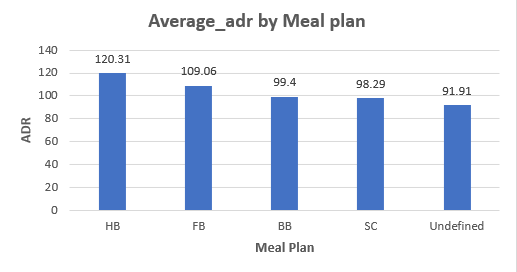
Data Integrity: Investigating and resolving the 'Undefined' category will lead to more reliable data and better decision-making.Promotional Activities: Marketing and promotional activities should be aligned with the most popular meal plans, particularly BB, while also subtly promoting HB and SC to their respective target segments.



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

The highest average daily rate is associated with Half Board (HB) at 120.31, followed by Full Board (FB) at 109.06. Bed & Breakfast (BB) and Self-Catering (SC) have similar ADRs of 99.4 and 98.29 respectively, both notably lower than HB and FB. The lowest ADR is for Undefined meal types at 91.91.

While BB is the most booked meal plan (from the previous data), HB and FB command significantly higher average daily rates, indicating that guests are willing to pay more for more inclusive meal packages, particularly for Half Board. This suggests that the perceived value or convenience of having more meals included directly translates to a higher price point for the accommodation.



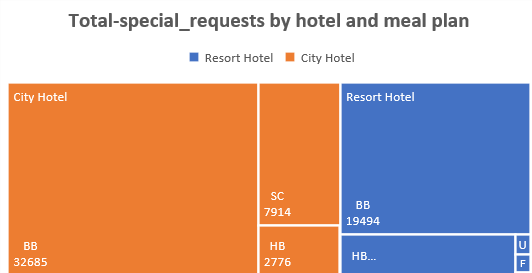
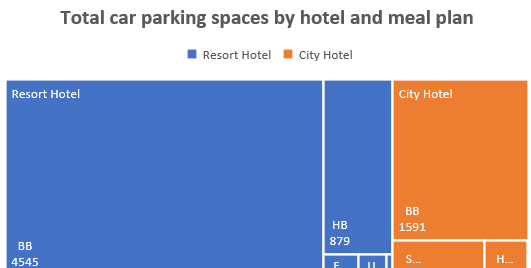
**Investigate the distribution of required car parking spaces and special requests by hotel type and meal plan.**

The data shows the breakdown of car parking spaces requested and special requests made per meal type for both Resort Hotels and City Hotels.

Resort Hotels:BB (Bed & Breakfast) leads significantly in both requested car parking spaces (4545) and special requests (19494).HB (Half Board) also shows substantial numbers for both car parking (879) and special requests (4916).FB (Full Board), SC (Self-Catering), and Undefined meal types have very low numbers for both categories.

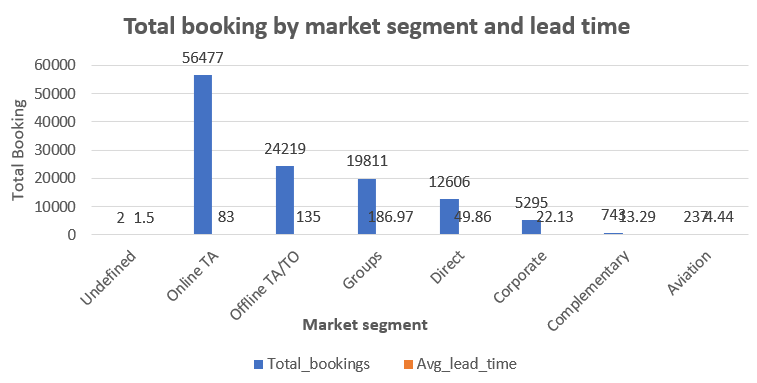
City Hotels:BB (Bed & Breakfast) again dominates in special requests (32685) but has fewer car parking spaces requested (1591) compared to Resort Hotel BB.SC (Self-Catering) has a surprisingly high number of special requests (7914) relative to its car parking requests (231) in City Hotels.HB (Half Board) shows moderate numbers.FB (Full Board) has extremely low numbers.

Across both hotel types, Bed & Breakfast (BB) meal plans are consistently associated with the highest volume of special requests, indicating that guests booking BB often have specific needs or preferences. Additionally, there's a strong correlation between BB and car parking space requests at Resort Hotels, suggesting these guests frequently arrive by car. Conversely, Full Board (FB) guests generally have the fewest car parking needs and special requests in both hotel types. The high special request volume for Self-Catering (SC) at City Hotels is a notable difference, suggesting guests choosing SC in urban settings may have more specific needs.



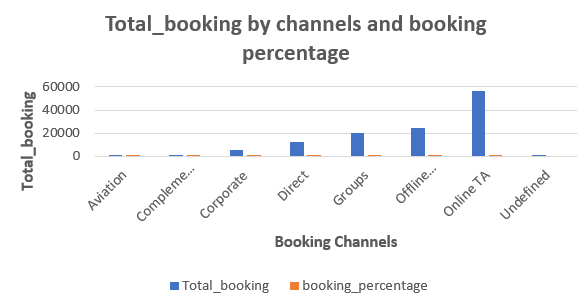
**Understand the distribution of bookings across different market segments and calculate summary statistics for lead times within each segment.**

The business heavily relies on Online TAs for the highest volume of bookings, suggesting a strong digital presence is crucial. However, Offline TAs/TOs and Groups contribute significant bookings with much longer lead times, indicating these segments plan their stays well in advance. This insight is critical for forecasting, inventory management, and developing targeted marketing strategies for segments with different booking behaviors. For instance, campaigns for Groups can be planned far ahead, while Corporate and Direct bookings might require more immediate attention and last-minute offers.



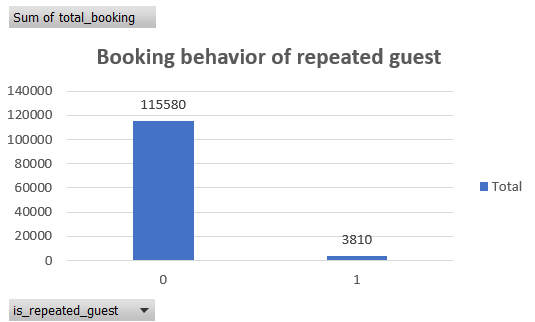
**Analyze the distribution of bookings through different booking channels (e.g., online travel agents, direct bookings) and calculate the percentage of bookings through each channel.**

The business has a heavy reliance on Online Travel Agents (nearly half of all bookings). While Online TAs are crucial for volume, the combined contribution of Offline TAs/TOs and Groups (over 36%) highlights the importance of traditional travel channels and group sales. This means a balanced strategy is needed: optimizing presence and conversion on online platforms while also nurturing relationships with traditional agents and managing group sales effectively. Diversification beyond just online channels is evident and important for overall booking stability.



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

The business has an extremely low rate of repeat customers. This indicates a significant opportunity for improvement in customer retention strategies. While acquiring new customers is important, focusing on initiatives to encourage repeat business could lead to more stable revenue streams and potentially lower marketing costs in the long run.



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

The most crucial insight is that guests who have canceled once before (previous\_cancellation = 1) are highly likely to cancel again, with a near 95% cancellation rate. While overall cancellation rates are high (around 33-34% for new bookers), identifying and potentially flagging or requiring pre-payment from guests with a single prior cancellation could significantly reduce future cancellations. Bookings from guests with very high numbers of prior cancellations (e.g., 14 or more) are almost guaranteed to be canceled, suggesting these might be problematic bookings.



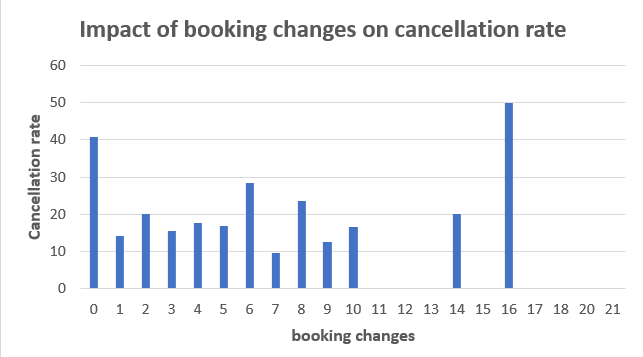
**Understand the distribution of reserved and assigned room types. Calculate summary statistics for the consistency between reserved and assigned room types.**

A significant majority of bookings (nearly 87.5%) are "matched," indicating a generally good alignment or successful processing of bookings. However, a mismatch rate of almost 12.5% is notable and suggests that a substantial number of bookings (around 1 in 8) encounter some form of discrepancy or issue. Addressing these mismatches could improve operational efficiency, customer satisfaction, and potentially reduce revenue leakage or data inaccuracies. Understanding the root causes of these mismatches would be the next crucial step.



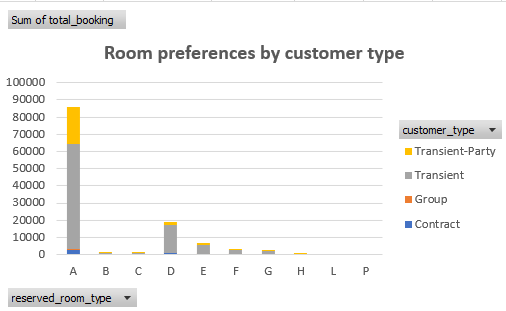
**Analyze the impact of booking changes on cancellation rates. Calculate cancellation rates for bookings with different numbers of changes.**

Counterintuitively, bookings that undergo no changes (0 changes) have the highest cancellation rate, while bookings with a small number of changes (1 to 5 changes) are significantly less likely to be canceled. This suggests that guests who actively manage or modify their bookings are more committed to their stay. The high cancellation rate for unchanged bookings might indicate impulsive bookings or those made without firm plans, leading to a higher likelihood of cancellation. Encouraging guests to confirm details or make minor adjustments after initial booking could potentially lead to lower cancellation rates.



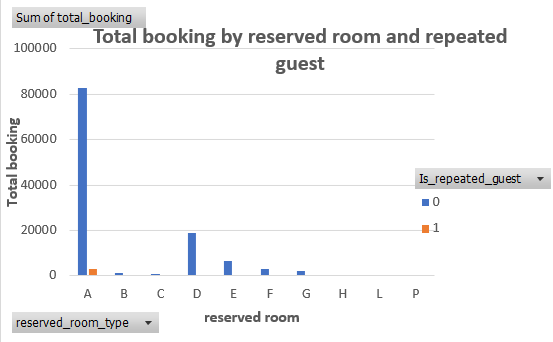
**Explore how room type preferences vary across different customer types (e.g., Transient, Group). Identify if certain customer types have specific room preferences.**

The business is overwhelmingly dominated by "Transient" bookings, particularly for Room Type A. This indicates that individual travelers are the primary customer base, and Room Type A is their preferred choice. While other customer types (Transient-Party, Contract, Group) and room types exist, their booking volumes are significantly lower. Therefore, optimizing the offering and marketing for Transient customers and Room Type A should be the top priority, while still managing the smaller, distinct needs of other segments. The low booking numbers for certain room types (e.g., L, P, H) might warrant a review of their market demand or operational efficiency.



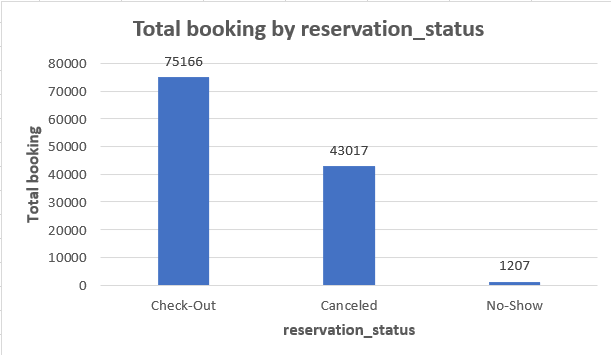
**Examine whether guests who make multiple bookings have consistent room type preferences or if their preferences change over time.**

The business is overwhelmingly reliant on new guests, who primarily book Room Type A. While repeat guests also show a preference for Room Type A, their overall contribution to bookings is minimal. This reinforces the need for strong new customer acquisition strategies, but it also highlights a significant untapped opportunity to encourage repeat bookings, especially for popular room types like A. Understanding why new guests don't become repeat guests could lead to substantial growth.



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

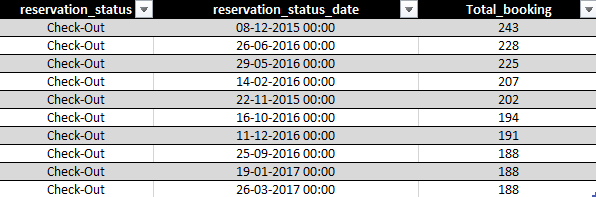
The most significant insight is the very high volume of canceled bookings (43,017), which represents a substantial portion of the total booking activity compared to actual check-outs. This indicates a major challenge with cancellations that could be impacting revenue and resource planning. While a large number of guests successfully check out, the high cancellation rate suggests a need for strategies to reduce cancellations, such as stricter cancellation policies, reconfirmation calls, or personalized offers. The relatively low number of no-shows is positive, but cancellations remain the primary concern.



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

The provided data shows a list of "Check-Out" reservation statuses along with their corresponding dates and the "Total\_booking" count for each date. The dates range from late 2015 to early 2017. The "Total\_booking" column appears to represent the number of bookings associated with that particular "Check-Out" date, with values generally in the range of 188 to 243.

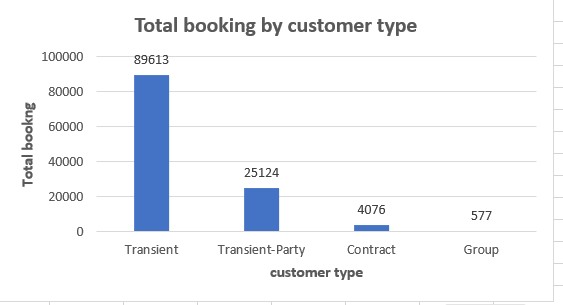
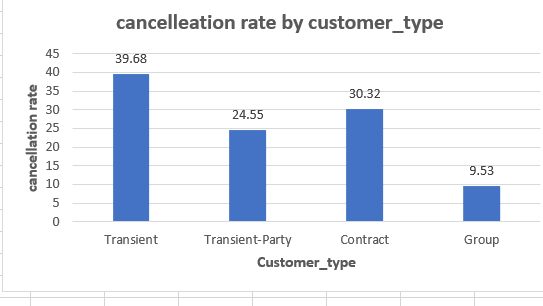
The data primarily highlights the volume of hotel check-outs on specific dates. It indicates that the booking numbers for these check-out dates were consistently high, generally above 180, suggesting a healthy occupancy or booking rate during the displayed period (late 2015 to early 2017).



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

The data presents booking and cancellation rate information categorized by "customer\_type." "Transient" customers account for the vast majority of bookings (89,613) but also have the highest cancellation rate (39.68%). "Transient-Party" customers have a significant number of bookings (25,124) with a moderate cancellation rate (24.55%). "Contract" customers have fewer bookings (4,076) and a cancellation rate of 30.32%. "Group" bookings are the lowest in volume (577) but exhibit the lowest cancellation rate (9.53%).

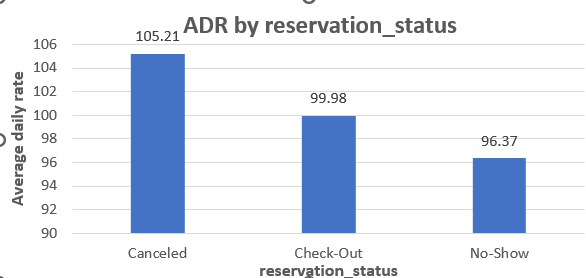
"Transient" customers are the most prevalent booking type but also represent the highest risk for cancellations. Conversely, "Group" bookings, though fewer in number, are the most reliable in terms of low cancellation rates. This suggests that while individual travelers ("Transient") drive volume, strategic focus on retaining "Group" and potentially "Transient-Party" customers could lead to more stable revenue due to lower cancellation risks.

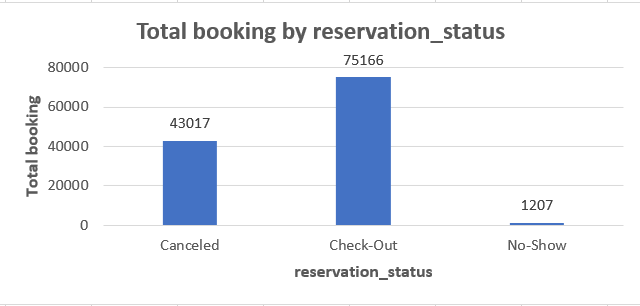


**Investigate whether there are differences in Average Daily Rates (ADR) based on reservation status (e.g., canceled vs. checkedout).**

The data categorizes hotel reservation statuses by "Canceled," "Check-Out," and "No-Show," showing the total bookings and average daily rate for each. "Check-Out" bookings represent the largest volume (75,166) with an average daily rate of 99.98. "Canceled" bookings are also significant in number (43,017) and surprisingly have the highest average daily rate (105.21). "No-Show" bookings are the lowest in volume (1,207) with the lowest average daily rate (96.37).

A substantial number of bookings are being canceled, and these canceled bookings actually had a higher average daily rate than completed "Check-Out" stays. This indicates a significant potential for revenue loss from cancellations, particularly from higher-priced bookings. While "Check-Outs" are the primary revenue driver, managing and reducing cancellations, especially those with higher average daily rates, should be a key focus to maximize profitability.

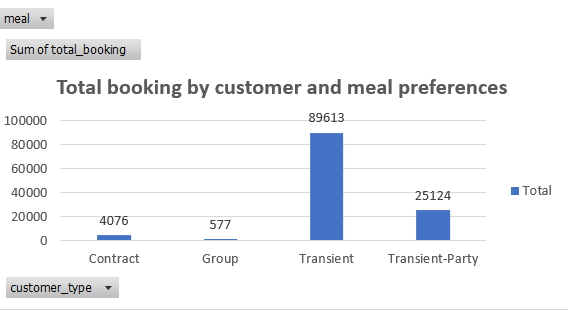




**Compare the distribution of meal plans among different customer types (e.g., Transient, Group) to identify preferences.**

This data breaks down the "total\_booking" count by "customer\_type" and the "meal" plan selected.Transient customers form the largest group by far, with the majority opting for "BB" (Bed & Breakfast) meal plans (70,692 bookings), followed by "SC" (Self-Catering) and "HB" (Half Board).Transient-Party customers also heavily favor "BB" (17,859 bookings), with "HB" as the next popular choice.Contract customers primarily book "BB" (3,260 bookings), followed by "HB."Group bookings are much lower in volume across all meal types, with "BB" being the most common choice (499 bookings)."Undefined" and "FB" (Full Board) meal plans have significantly fewer bookings across all customer types.

"Bed & Breakfast" (BB) is overwhelmingly the most preferred meal plan across all customer types, indicating its widespread popularity. "Transient" customers are the dominant segment, and their strong preference for BB should be a key consideration for hotel offerings and marketing. While other meal plans exist, their demand is considerably lower, especially for "Full Board" and "Undefined" options.



**CONCLUSION**

In conclusion, the **Hotel Booking Analysis** serves as a crucial foundation for understanding the complex dynamics of guest behavior, market demand, and operational performance in the hospitality sector. By systematically analyzing booking trends, lead times, arrival patterns, stay durations, and cancellation rates, hotels gain a clear picture of how their rooms are being booked, when demand peaks and drops occur, and what factors influence guest decisions.

This analysis has highlighted how different guest segments — such as business travelers, leisure guests, families, or repeat customers — contribute uniquely to overall occupancy and revenue. For example, city hotels may benefit more from weekday business bookings, while resort properties often see higher demand for weekend or holiday stays. Such insights enable hotels to strategically design pricing models that capture maximum revenue during high-demand periods while creating attractive offers to boost bookings during low seasons.

Another key takeaway is the importance of understanding cancellation trends and lead times. High cancellation rates can directly impact occupancy forecasts and revenue targets if not anticipated and managed proactively. By using these insights, hotels can refine their cancellation policies, improve overbooking strategies, and ensure that last-minute cancellations do not result in empty rooms and lost income.

Furthermore, this analysis emphasizes the critical role of aligning marketing and sales efforts with real booking data. Knowing when guests typically book and what influences their choices helps hotels develop targeted promotions, loyalty programs, and value-added packages that meet specific guest needs. This not only drives new bookings but also strengthens relationships with existing customers, encouraging repeat visits and positive reviews.

Beyond revenue and occupancy optimization, the broader impact of booking analysis lies in operational efficiency and guest experience. With better demand forecasting, hotels can plan staffing, housekeeping, and inventory needs more accurately, reducing waste and controlling costs. They can also tailor services and amenities to match guest expectations — whether that means more family-friendly offerings, business services, or personalized touches for loyal customers.

In an increasingly competitive and dynamic market, hotels that leverage booking data effectively gain a significant advantage. They can respond faster to market changes, make confident decisions backed by real trends, and stay ahead of shifting guest preferences. Ultimately, this **Hotel Booking Analysis** provides the actionable insights needed to drive sustainable growth, enhance profitability, and deliver memorable guest experiences that build long-term loyalty and brand reputation.

**THANK YOU**