Hotel Booking Analysis

**By: Aniket Ojha**

**CAPSTONE PROJECT**

# Abstract:

Ever since the pandemic hit worldwide, almost all industry sectors have observed a negative business impact with ‘Hotel’ sector under ‘Travel and Tourism’ industry taking a major blow. With near end of the global pandemic the business is thriving more than ever but it’s not the same as before. With advent of technology, people have become more aware of multiple mediums to book reservations at luxurious Hotels at affordable prices. This has made many entrepreneurs and business leaders realize how important it is to be strategic and leverage data analysis to penetrate the market more and gain more market shares. By applying Exploratory Data Analysis principles and feature engineering, this study aims to explore merits of analysis on Hotels dataset that governs bookings and can help in increasing customer base. The decision support and benefits reported in this study advocate significance of Data Analysis.The dataset contains booking information for two different hotels. One being City Hotel and another Resort Hotel along with information on various booking criteria such as booking season, time of booking, length of stay, number of adults, children and babies, parking spaces, pricing data, market segment and many more.

Primary objective is to explore and inspect the dataset; and discover important features using Exploratory Data Analysis that can govern bookings and help hotels penetrate deep into the market, thereby attracting more customers. Analyze booking and pricing trends to draw out insights to make intelligent business decisions.

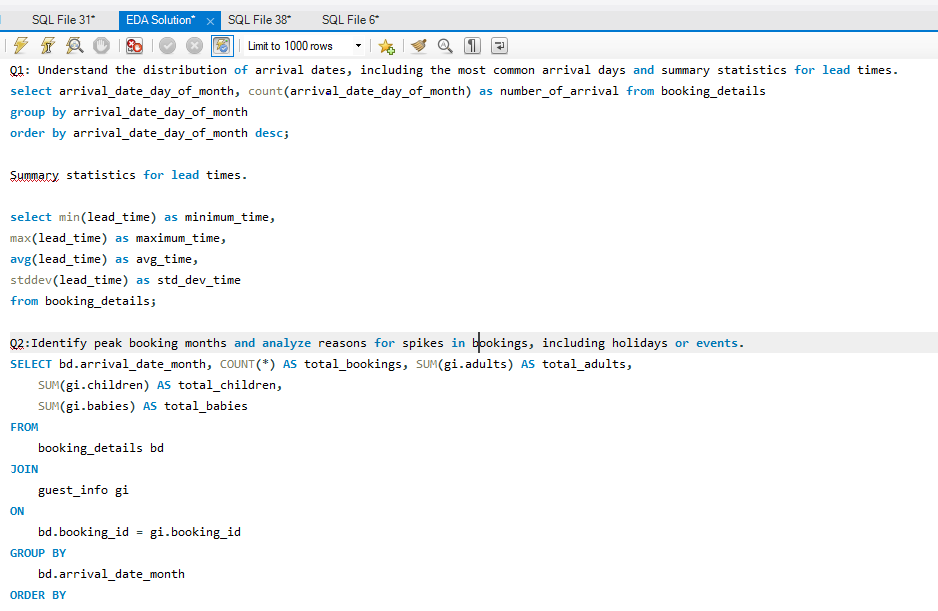
Secondary objective is to help the customers in deciding the best period to visit places while availing low accommodation cost benefits.

**Problem Statement**

* Data Exploration and Inspection:
  + Libraries imported:
    - Data manipulation: numpy and pandas
    - Data visualization : matplotlib and seaborn
  + Initial dataset size: 119390 rows and 32 columns.
  + Data contains following features:
* hotel: Resort Hotel or City Hotel
* is\_canceled: Value indicating if the booking was canceled (1) or not (0)
* lead\_time: Number of days that elapsed between the entering date of the booking and the arrival date
* arrival\_date\_year: Year of arrival date
* arrival\_date\_month: Month of arrival date
* arrival\_date\_week\_number: Week number of year for arrival date

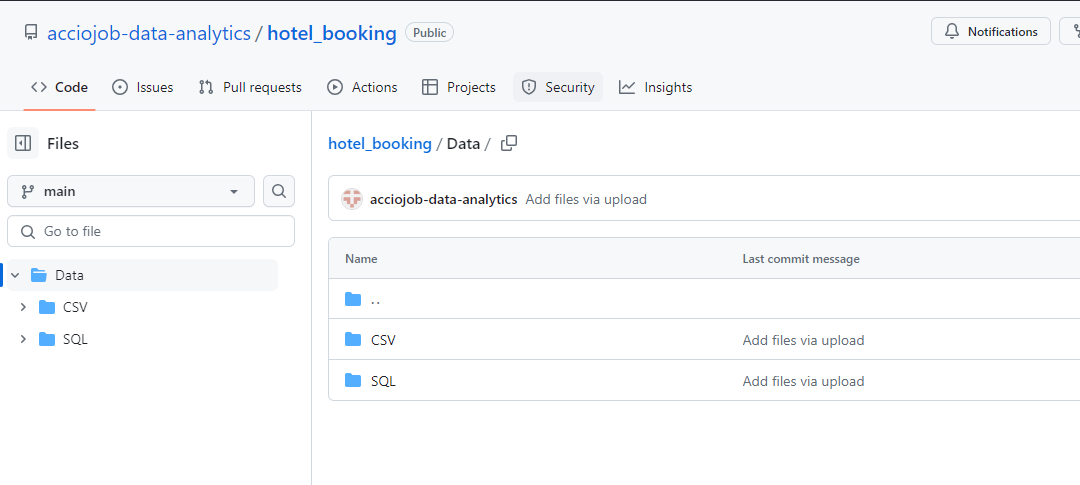
arrival\_date\_day\_of\_month: Day of arrival date

* stays\_in\_weekend\_nights: Number of weekend nights
* stays\_in\_week\_nights: Number of week nights.
* adults: Number of adults
* children: Number of children
* babies: Number of babies
* meal: Type of meal booked
* country: Country of origin.
* market\_segment: Market segment designation (TA/TO)
* distribution\_channel: Booking distribution channel.(T/A/TO)
* is\_repeated\_guest: is a repeated guest (1) or not (0)
* previous\_cancellations: Number of previous booking
* **previous\_bookings\_not\_canceled**: Number of previous bookings not cancelled by the customer prior to the current booking
* **reserved\_room\_type**: Code of room type reserved.
* **assigned\_room\_type**: Code for the type of room assigned to the booking.
* **booking\_changes**: Number of changes made to the booking
* **deposit\_type** : No Deposit, Non Refund , Refundable.
* **agent**: ID of the travel agency that made the booking
* **company**: ID of the company/entity that made the booking .
* **days\_in\_waiting\_list** : Number of days the booking was in the waiting list before it was confirmed to the customer
* **customer\_type**: type of customer. Contract, Group, Transient, Transient party.
* **adr**: Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights
* **required\_car\_parking\_spaces**: Number of car parking spaces required by the customer
* **total\_of\_special\_requests**: Number of special requests made by the customer (e.g. twin bed or high floor)
* **reservation\_status**: Reservation last status.



**3. Extraction: SQL queries**

* 1. SQL queries were used to extract the necessary data from the database. This involved writing specific queries to retrieve tables such as booking details, guest information, room details, and reservation statuses.

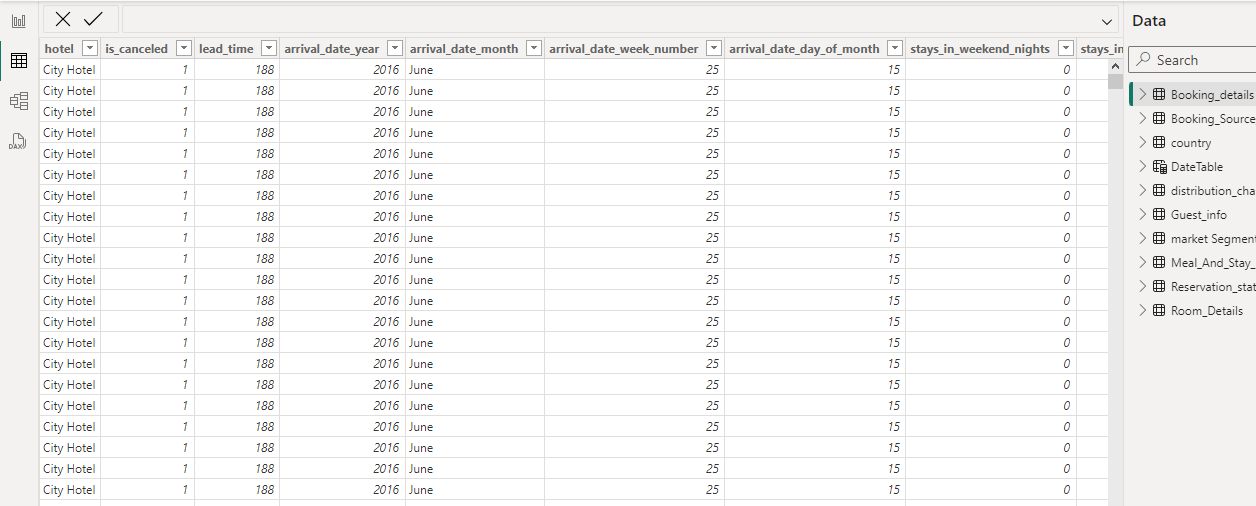


1. Data Collection:
2. **Source: Accio Github** 
   1. The data was sourced from the Accio github database, which includes comprehensive information on bookings, guests, and related details.

**4. Data Preparation:**

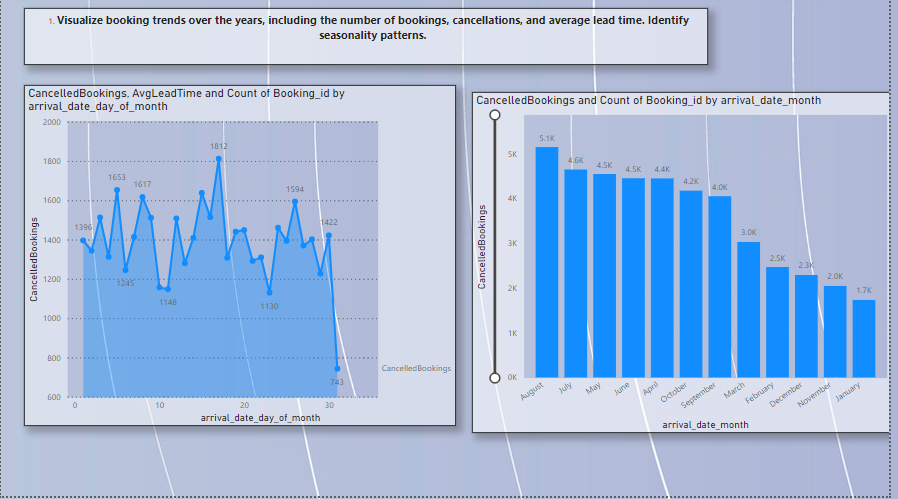
**Cleaning: Handling missing values, removing duplicates**

* 1. **Handling Missing Values:** Identified and filled missing values where appropriate, or removed entries that were incomplete and could not be corrected.
  2. **Removing Duplicates:** Ensured that duplicate records were identified and removed to maintain data integrity.



* Exploratory Data Analysis (EDA):

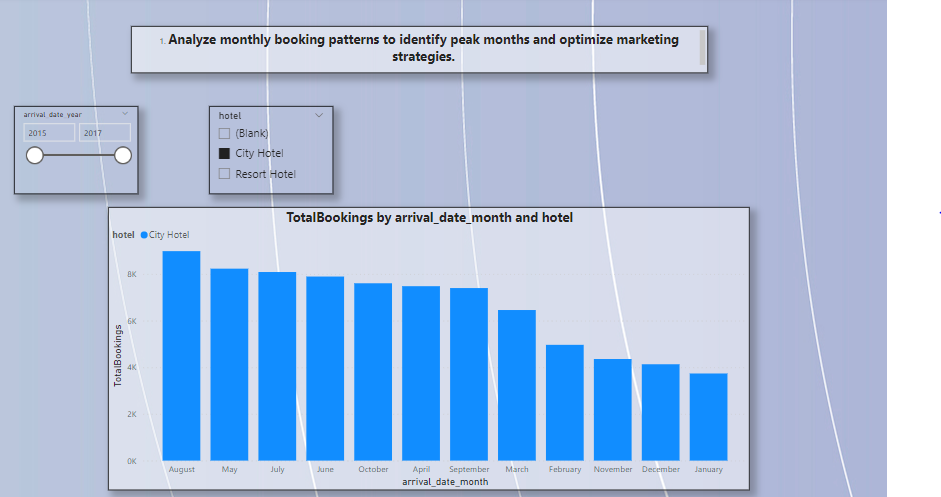
1. **Visualize booking trends over the years, including the number of bookings, cancellations, and average lead time. Identify seasonality patterns.**



* Exploratory Data Analysis (EDA):

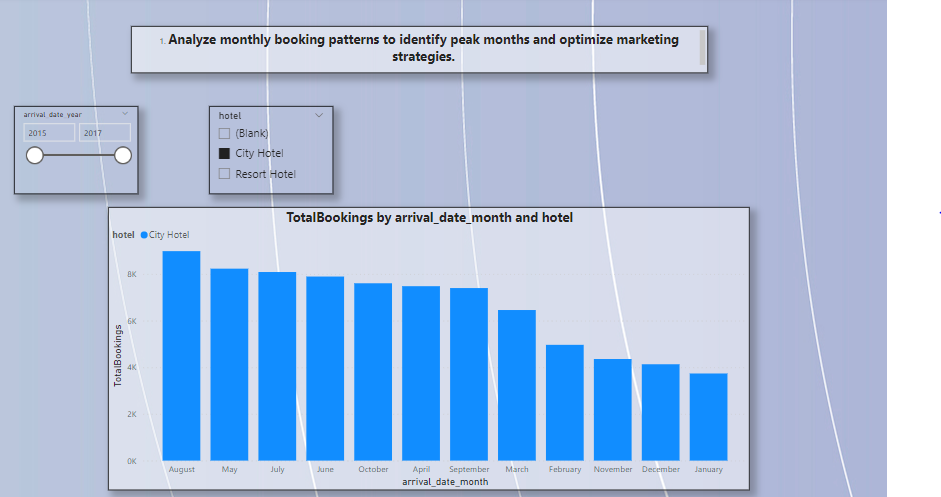
**2. Analyze monthly booking patterns to identify peak months and optimize marketing strategies.**

Insight: Visualizing booking trends over the years helps identify significant patterns and changes. By analyzing the number of bookings and cancellations alongside the average lead time, we can determine peak seasons and periods of low activity. Seasonality patterns, such as increased bookings during holidays or specific months, can be identified, allowing the hotel to optimize marketing strategies and resource allocation. Understanding these trends helps in forecasting and planning, ensuring better customer satisfaction and efficient use of resources



Insight: Monthly booking patterns reveal the hotel's busiest and slowest months. Identifying peak months allows the hotel to allocate resources efficiently, ensuring adequate staffing and inventory. Conversely, recognizing off-peak periods enables targeted marketing campaigns to boost occupancy. By understanding these patterns, the hotel can implement promotional strategies during slower months and optimize pricing and offers during peak periods, ultimately maximizing revenue and enhancing guest experience.

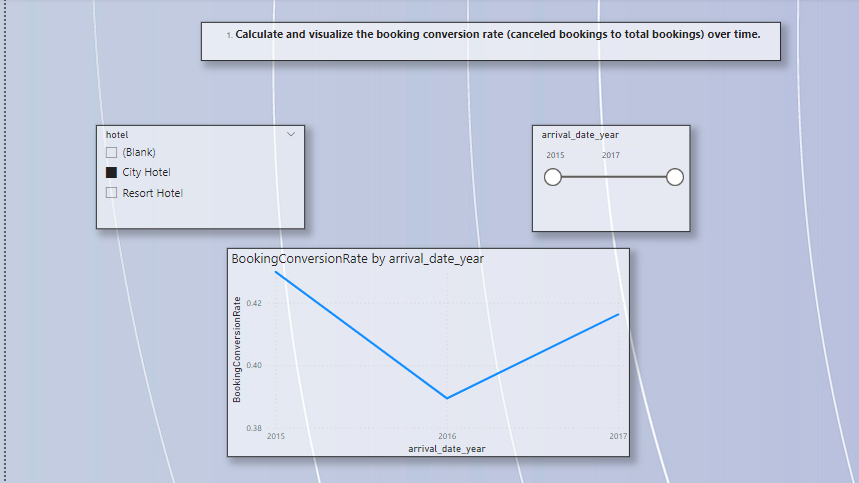
* Exploratory Data Analysis (EDA):
* **Compare stays in weekend nights and weekday nights to determine preferences and variations by hotel type.**



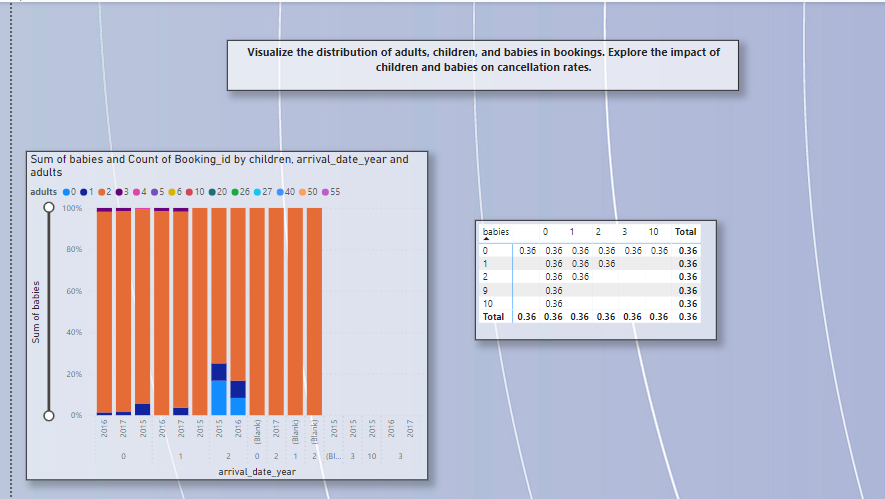
Insight: Analyzing stays in weekend versus weekday nights provides valuable insights into guest preferences. Typically, weekend stays might indicate leisure travelers, while weekday stays could suggest business travelers. Comparing these patterns across different hotel types (Resort vs. City Hotel) helps tailor services and amenities to meet specific guest needs. For instance, resorts may see higher weekend occupancy, prompting offers for extended stays, while city hotels might focus on corporate packages during weekdays.

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

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



Insight: The booking conversion rate, representing the ratio of canceled bookings to total bookings, is crucial for understanding customer behavior and potential revenue loss. By visualizing this metric over time, the hotel can identify trends and patterns in cancellations. High cancellation rates may indicate issues with booking policies, guest dissatisfaction, or external factors. Addressing these issues through improved policies, customer engagement, or flexible booking options can reduce cancellations and enhance overall booking efficiency.



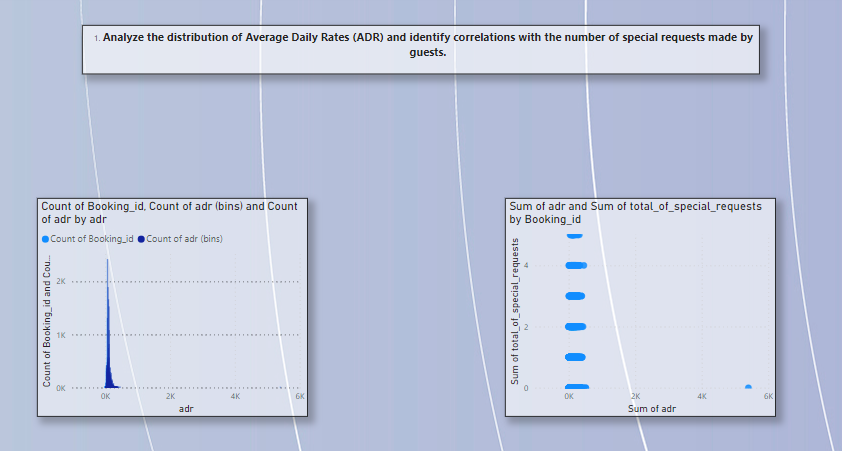
* **Exploratory Data Analysis (EDA):**
* **Visualize the distribution of adults, children, and babies in bookings. Explore the impact of children and babies on cancellation rates.**

Insight: Visualizing the distribution of adults, children, and babies in bookings provides insights into the guest demographic and helps tailor services accordingly. Understanding the impact of children and babies on cancellation rates is essential, as families with young children might have higher cancellation tendencies due to unforeseen circumstances. By identifying this trend, the hotel can develop family-friendly policies, flexible booking options, and targeted marketing campaigns to enhance guest experience and reduce cancellations.

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

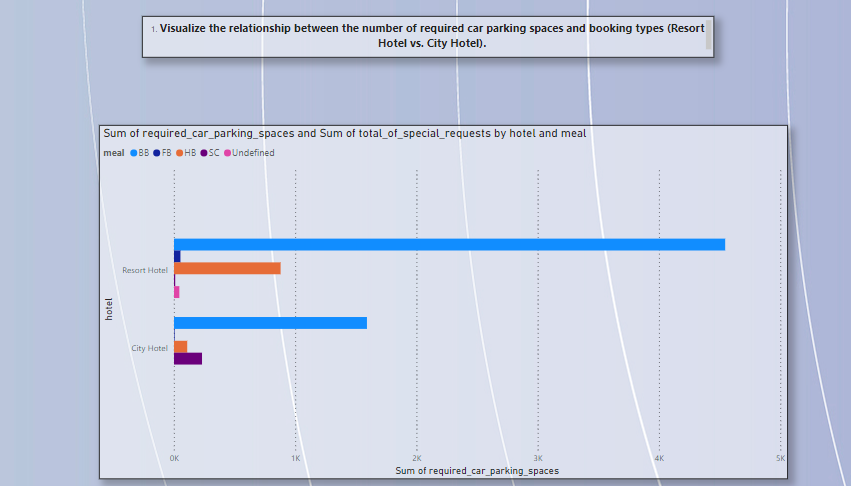
**Analyze the distribution of Average Daily Rates (ADR) and identify correlations with the number of special requests made by guests.**

**Insight**: Analyzing the distribution of ADR alongside the number of special requests helps understand the relationship between pricing and guest expectations. Higher ADRs might correlate with an increase in special requests, indicating that guests paying premium rates expect additional services. Identifying these correlations enables the hotel to enhance its service offerings for higher-paying guests, ensuring that their expectations are met, which can lead to increased satisfaction and repeat bookings.



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

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



Insight: Visualizing the relationship between car parking space requirements and booking types (Resort vs. City Hotel) helps identify patterns and guest needs. Resorts might have higher parking demands due to longer stays and more leisure travelers arriving by car, while city hotels may have varied requirements depending on business travelers and short stays. Understanding these patterns aids in optimizing parking space allocation, improving guest experience, and planning future infrastructure investments.

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

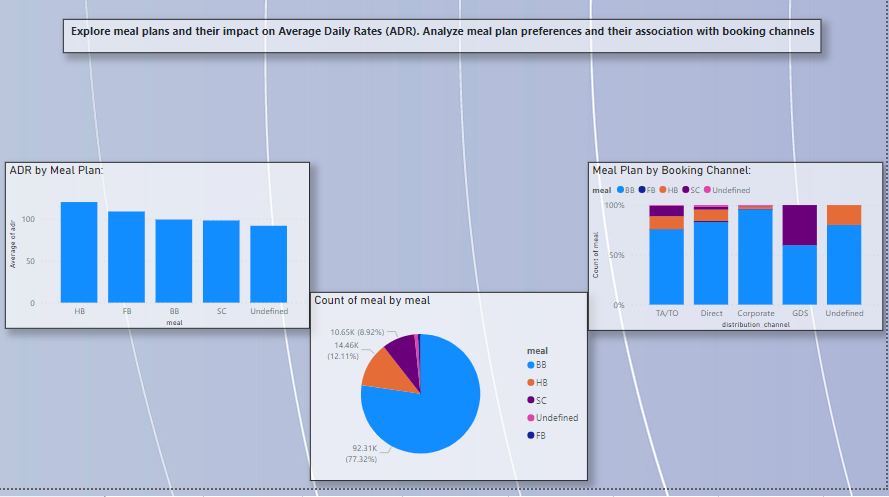
**Use Power BI to explore how the total number of special requests made by guests varies by hotel type and customer type (e.g., Transient, Group).**



**Insight**: Exploring the variation in special requests by hotel type and customer type reveals critical insights into guest needs and preferences. Transient guests might have different requirements compared to groups, and these needs can vary between resort and city hotels. By identifying these patterns, the hotel can tailor its services and amenities to meet the specific demands of each customer segment, enhancing guest satisfaction and potentially increasing loyalty and repeat bookings.

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

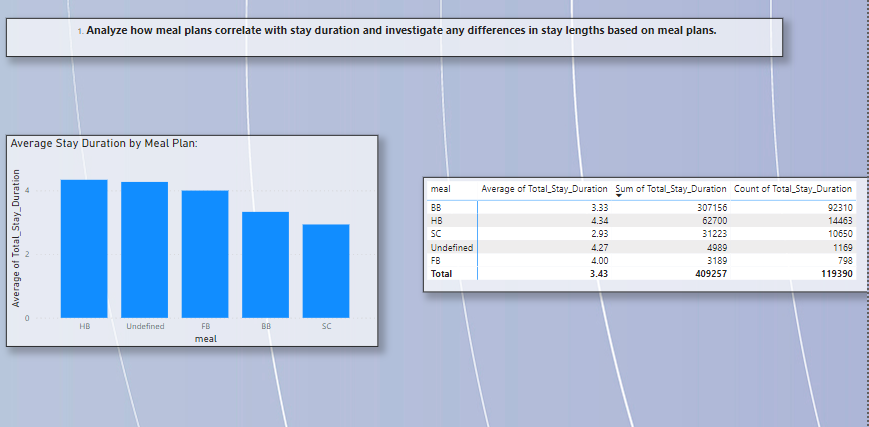
**Explore meal plans and their impact on Average Daily Rates (ADR). Analyze meal plan preferences and their association with booking channels**



**Insight**: Examining meal plans and their impact on ADR provides insights into how different meal offerings influence guest spending. Analyzing preferences across various booking channels helps identify which channels attract guests opting for specific meal plans. For instance, guests booking through travel agents might prefer all-inclusive packages, while direct bookings might favor bed-and-breakfast options. Understanding these associations aids in optimizing meal plan offerings and marketing strategies to maximize revenue.

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

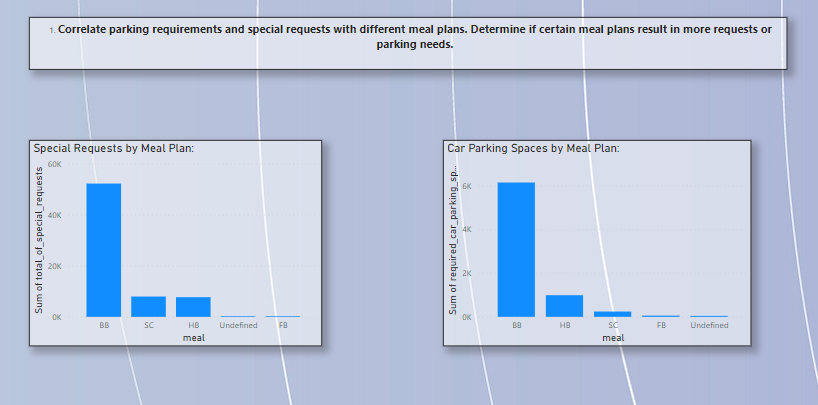
**Analyze how meal plans correlate with stay duration and investigate any differences in stay lengths based on meal plans.**

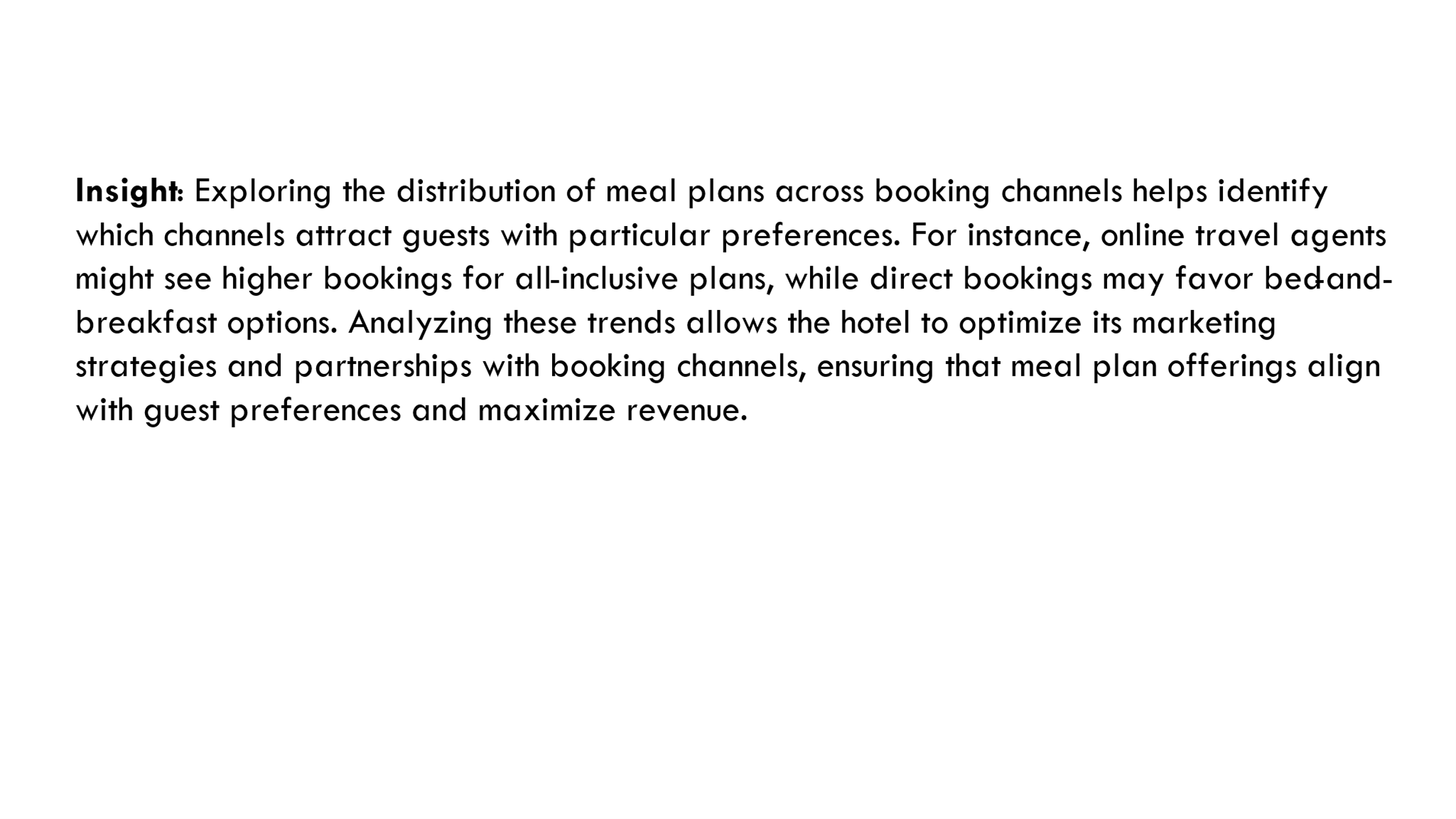


**Insight**: Analyzing the correlation between meal plans and stay duration helps identify if certain meal options attract longer stays. Guests opting for full-board or all-inclusive plans may tend to stay longer, benefiting from the convenience and value of bundled meals. Understanding these differences allows the hotel to design targeted packages and promotions that cater to guests' preferences, potentially increasing average stay lengths and overall revenue.

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

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





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

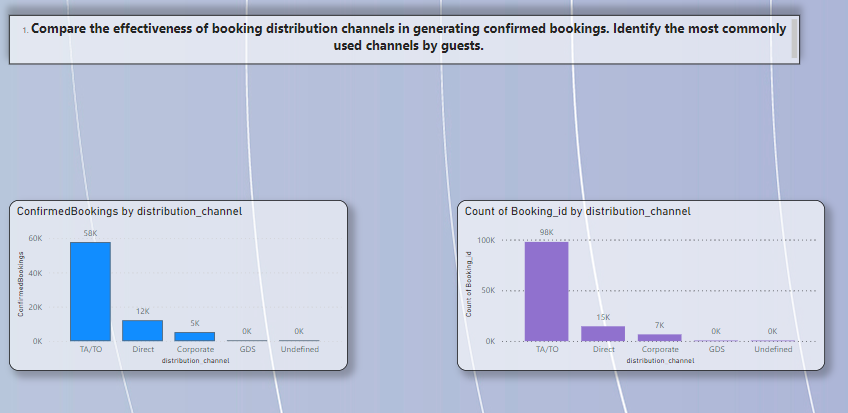
**Visualize booking distribution across different market segments and analyze cancellation rates within each segment.**



**Insight**: Visualizing booking distribution across market segments and analyzing cancellation rates helps identify which segments are most valuable and which are prone to cancellations. For example, business travelers might have lower cancellation rates compared to leisure travelers. Understanding these dynamics allows the hotel to tailor its marketing and retention strategies for each segment, improving overall booking stability and customer satisfaction.

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

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



**Insight**: Comparing the effectiveness of booking distribution channels helps identify which channels generate the most confirmed bookings. Channels such as direct bookings, online travel agents, and corporate bookings can be analyzed for their performance. Identifying the most commonly used channels allows the hotel to focus its marketing efforts and partnerships on these channels, optimizing booking conversion rates and maximizing revenue.

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

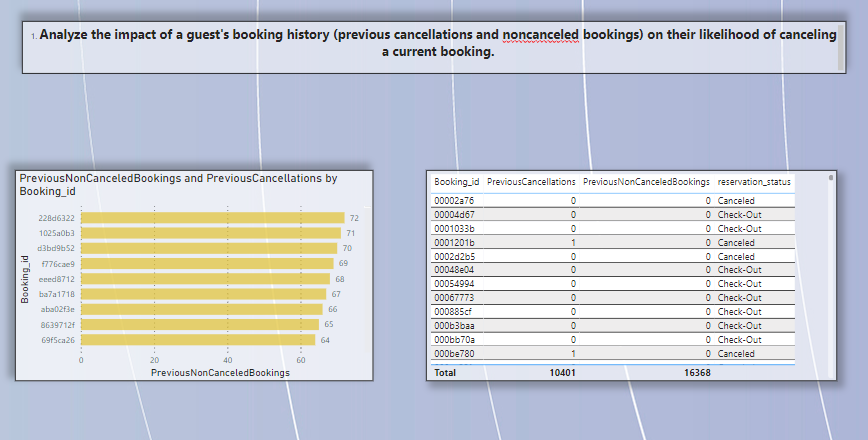
**Visualize the percentage of repeated guests for each hotel type (Resort Hotel vs. City Hotel) over time. Explore factors influencing guest retention.**



**Insight**: Visualizing the percentage of repeated guests for resort and city hotels over time helps understand guest loyalty and retention patterns. Analyzing factors influencing retention, such as service quality, amenities, and guest experience, provides insights into what drives repeat bookings. This information enables the hotel to enhance its offerings and create targeted loyalty programs to increase guest retention and build a loyal customer base.

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

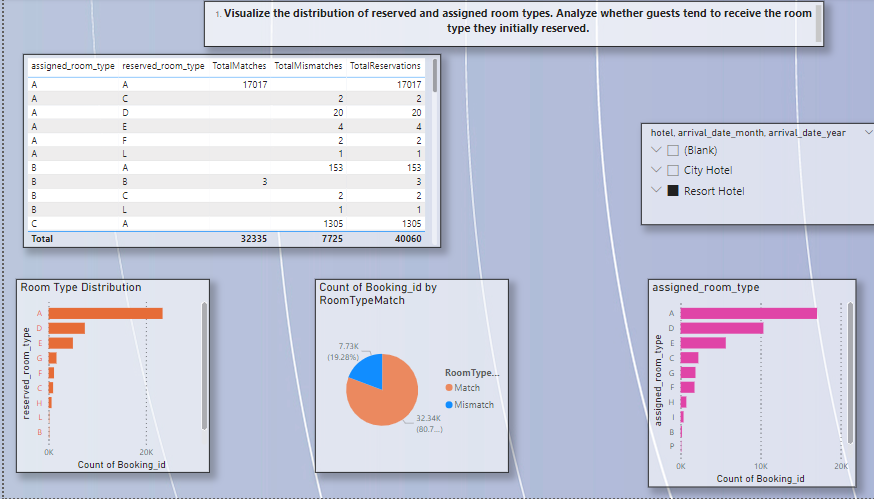
**Analyze the impact of a guest's booking history (previous cancellations and non canceled bookings) on their likelihood of canceling a current booking.**



**Insight**: Analyzing a guest's booking history provides insights into their likelihood of canceling current bookings. Guests with a history of cancellations might have a higher probability of future cancellations, while those with consistent non-canceled bookings may be more reliable. Understanding these patterns allows the hotel to implement targeted retention strategies, such as flexible booking policies or personalized offers, to reduce cancellations and improve booking stability.

* Exploratory Data Analysis (EDA):

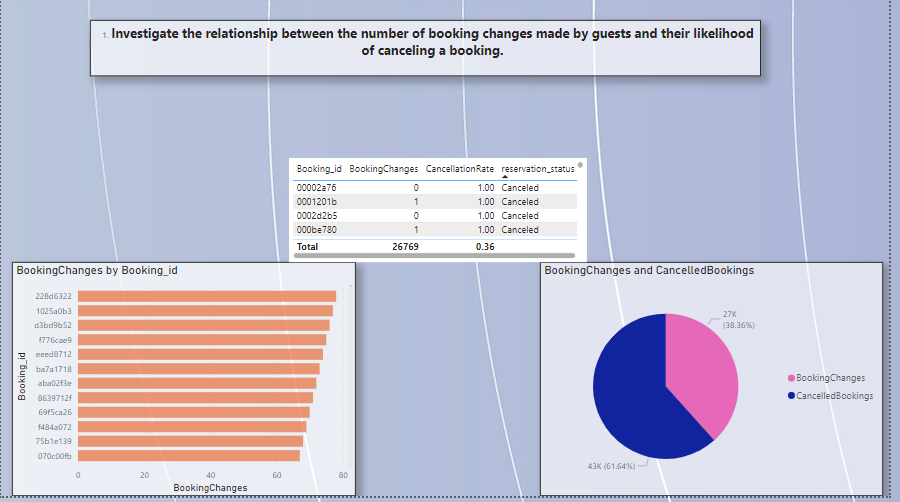
**Visualize the distribution of reserved and assigned room types. Analyze whether guests tend to receive the room type they initially reserved.**



**Insight**: Visualizing the distribution of reserved and assigned room types helps understand if guests are receiving the rooms they initially reserved. Analyzing discrepancies between reserved and assigned room types can identify potential issues in the booking process or inventory management. Ensuring guests receive their preferred room type enhances satisfaction and trust, reducing the likelihood of complaints and cancellations.

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

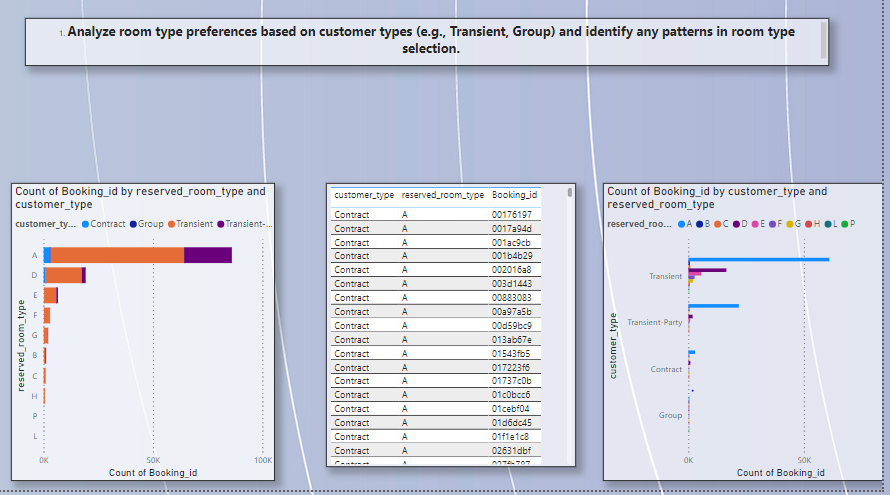
**Investigate the relationship between the number of booking changes made by guests and their likelihood of canceling a booking.**



**Insight**: Investigating the relationship between booking changes and cancellations helps identify if frequent changes indicate a higher likelihood of cancellations. Guests making multiple changes may have uncertain plans, increasing the risk of cancellation. Understanding this relationship allows the hotel to offer flexible booking options and targeted communication to such guests, reducing cancellation rates and improving booking stability.

* Exploratory Data Analysis (EDA):

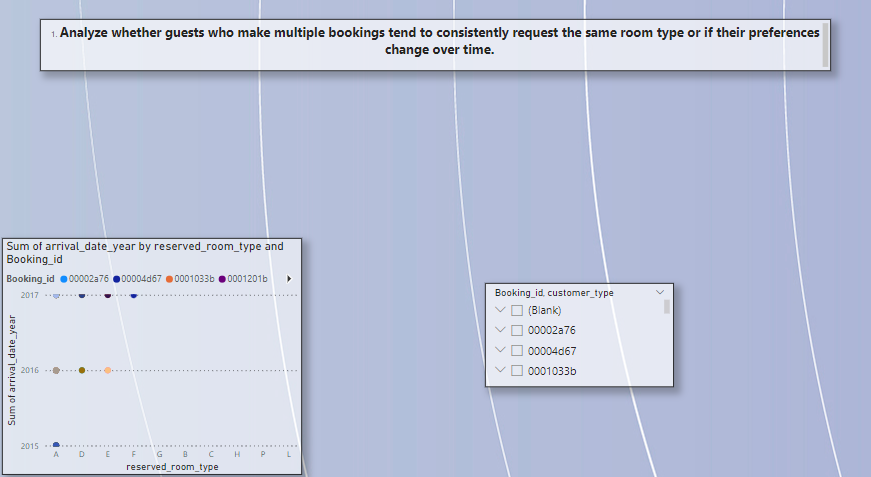
**Analyze room type preferences based on customer types (e.g., Transient, Group) and identify any patterns in room type selection.**



**Insight**: Analyzing room type preferences based on customer types provides insights into what different segments prefer. Transient guests might favor standard rooms, while groups may prefer suites or connecting rooms. Identifying these patterns allows the hotel to tailor its room offerings and marketing strategies to meet the specific needs of each customer type, enhancing guest satisfaction and optimizing room occupancy.

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

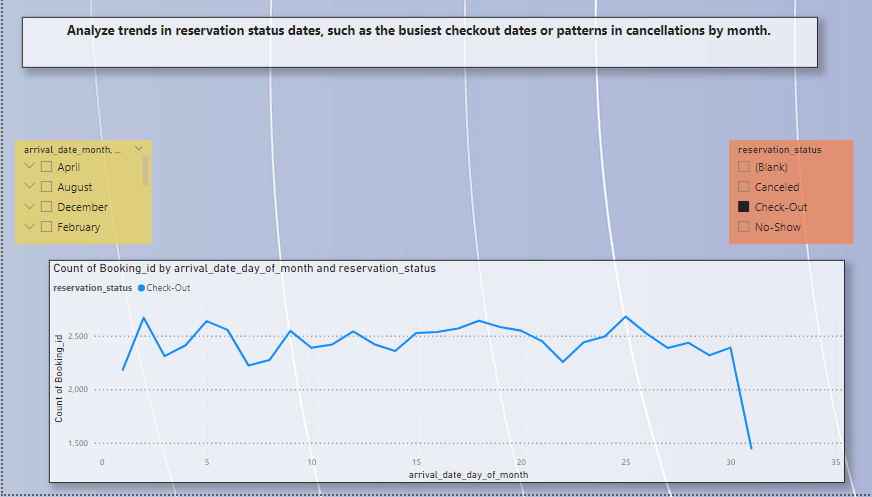
**Analyze whether guests who make multiple bookings tend to consistently request the same room type or if their preferences change over time.**



**Insight**: Analyzing guests' room type preferences over multiple bookings helps identify if they consistently request the same type or change preferences over time. Consistent preferences may indicate strong loyalty to specific room features, while changing preferences might suggest evolving needs or tastes. Understanding these trends allows the hotel to personalize guest experiences and anticipate their needs, fostering loyalty and repeat bookings.

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

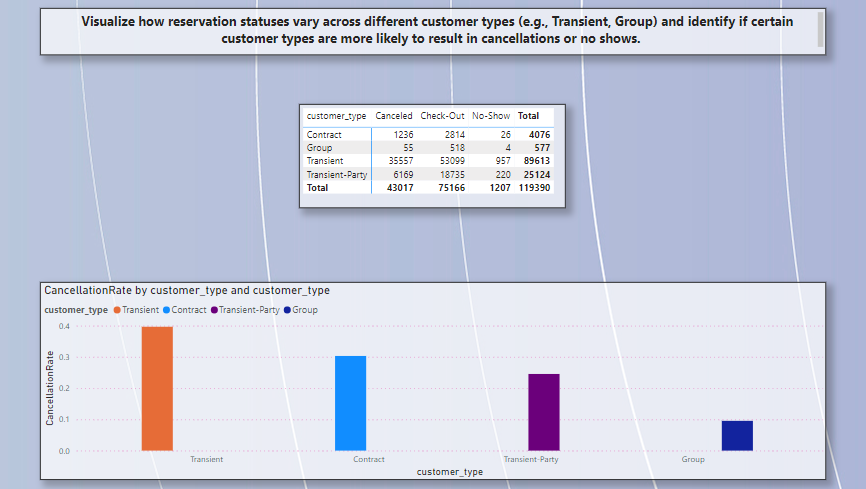
**Analyze trends in reservation status dates, such as the busiest checkout dates or patterns in cancellations by month.**



**Insight**: Analyzing trends in reservation status dates helps identify the busiest checkout dates and patterns in cancellations by month. This information is crucial for operational planning, resource allocation, and marketing strategies. Understanding these trends allows the hotel to optimize staffing, enhance guest services during peak periods, and address factors leading to cancellations, ensuring a smoother operation and better guest experience.

* Exploratory Data Analysis (EDA):

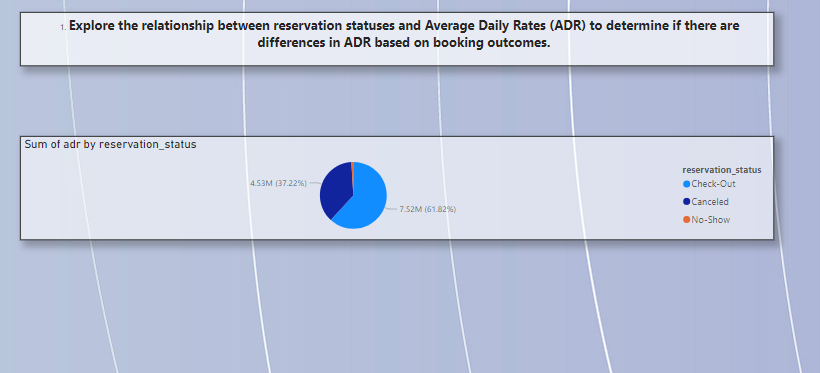
**Visualize how reservation statuses vary across different customer types (e.g., Transient, Group) and identify if certain customer types are more likely to result in cancellations or no shows.**



**Insight**: Visualizing reservation statuses across different customer types reveals which segments are more prone to cancellations or no-shows. For instance, transient guests might have higher cancellation rates due to unpredictable travel plans. Identifying these patterns helps the hotel tailor its booking policies and customer communication, reducing the risk of cancellations and no-shows, and ensuring a more stable booking environment.

Exploratory Data Analysis (EDA):

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



**Insight**: Exploring the relationship between reservation statuses and ADR helps determine if booking outcomes affect revenue. Higher ADRs for checked-out bookings compared to canceled or no-show bookings indicate that confirmed stays contribute significantly to revenue. Understanding these differences allows the hotel to focus on strategies that maximize ADR and reduce cancellations, such as targeted promotions and flexible booking options, enhancing overall financial performance.

### Project Summary

This project involves a comprehensive analysis of hotel booking data to extract actionable insights and improve operational efficiency. Data was sourced from the hotel's booking system database and prepared through cleaning, transformation, and loading into Power BI for visualization. Key areas of analysis include booking trends, guest preferences, cancellation rates, and ADR correlations.

Booking trends were visualized to identify seasonality patterns, peak booking months, and average lead times, helping to optimize marketing strategies and resource allocation. Guest preferences were analyzed to understand room type and meal plan choices across different customer segments, enabling tailored service offerings. Cancellation rates were examined to identify influencing factors, aiding in developing retention strategies.

Further analysis of ADR correlations with booking behaviors assisted in refining pricing strategies. Parking needs and special requests were explored to facilitate better resource planning. The effectiveness of booking distribution channels was analyzed to guide marketing and partnership efforts. Market segment analysis provided insights into booking distribution and cancellation rates, aiding in creating targeted promotional strategies.

Reservation status trends, including common checkout dates and cancellation patterns, were analyzed to manage peak periods efficiently. Customer type behavior and historical booking impact were evaluated to enhance loyalty programs and guest retention efforts. Overall, this project provides a robust framework for leveraging data-driven insights to improve hotel operations and guest satisfaction.