Hotel Booking Analysis

- Preetham Krishnamurthy

Table of Contents

Contents

Dataset Link: https://www.kaggle.com/datasets/mojtaba142/hotel-booking/data

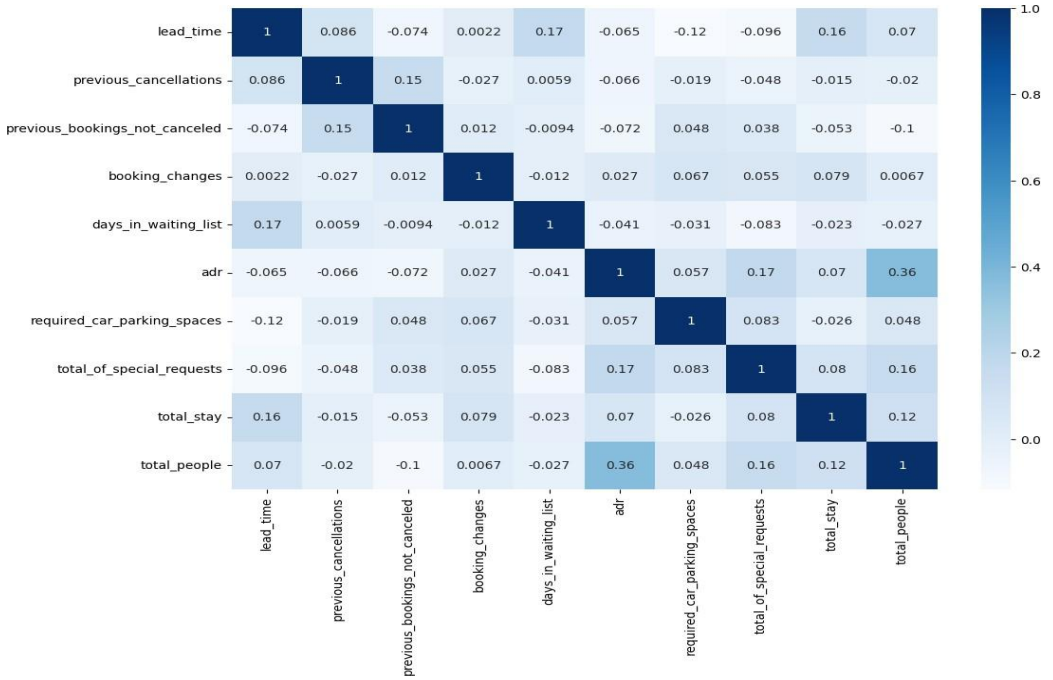
1. **Analytical Question**

Primary Question: The fundamental question is which reasons drive customers to cancel hotel reservations, as well as which aspects assist hotels in understanding customer requirements for a better booking rate.

And also from the hotel dataset, we can extract booking information, such as the peak hotel booking period. What are the reasons behind the customer's decision to cancel their hotel reservation? Which type of hotel—a city hotel or a resort hotel, in my case—do they prefer most? What kind of food do they like best, and which company or agencies get the most reservations? We are able to find out what kind of deposit the majority of our clients choose, as well as whether they are returning visitors. We can also examine the average number of adults, kids, and infants that check into the hotel, as well as the nation from which the most and least booking happen. And we are able to contrast the hotel rates in each period.

2. Exploratory Data Analysis and Data Visualization

The above code represents different columns that are considered for the Correlation. we have not considered few columns like "is_cancelled," "arrival_date_year," "arrival_date_week_number," "arrival_date_of_month," "is_repeated_guest," "company," and "agent" because they are categorical data with numerical types. And also I have excluded the columns for stays in weekend nights, stays in week nights, adults, babies, and children because the total_stay and total_people columns have been introduced.

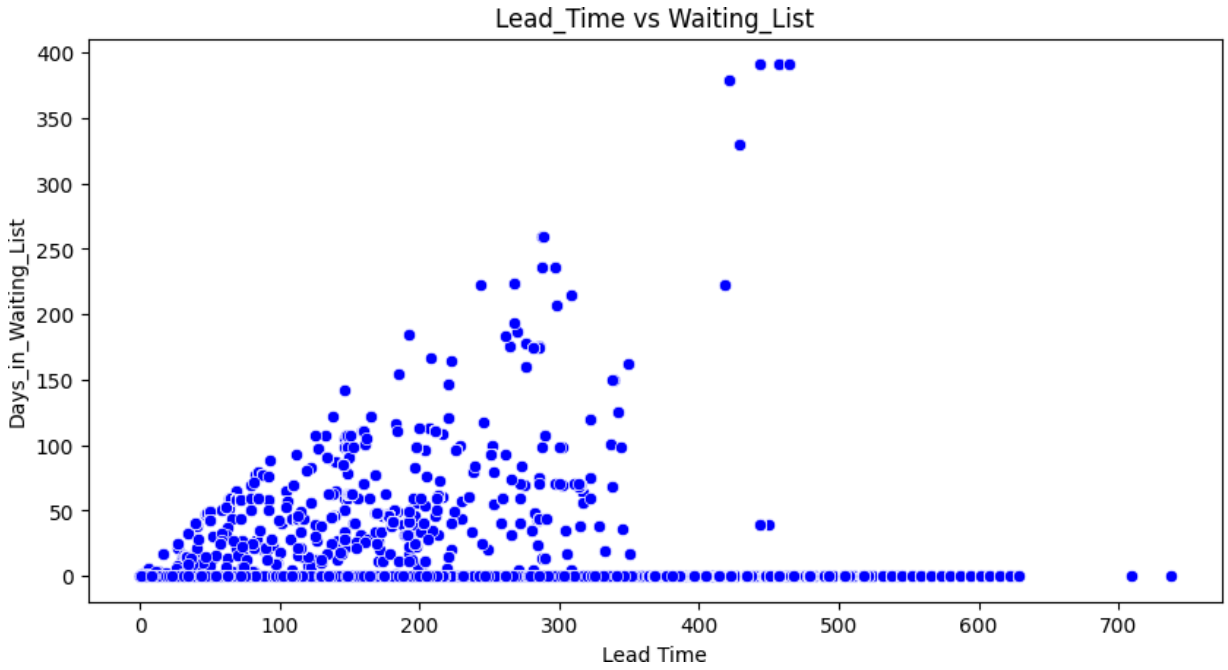


The above heat map visualization represents different correlation between the columns. From these we can find that

- Days on the waiting list have a small correlation with lead time. This means that the longer the lead time, the shorter the time on the waiting list.
- The total length of stay and lead time are slightly correlated. This could imply that for lengthier hotel stays, guests often plan little before their arrival.
- Adr is marginally associated with total_people, which makes sense given that a bigger number of people equals more revenue and hence more adr.

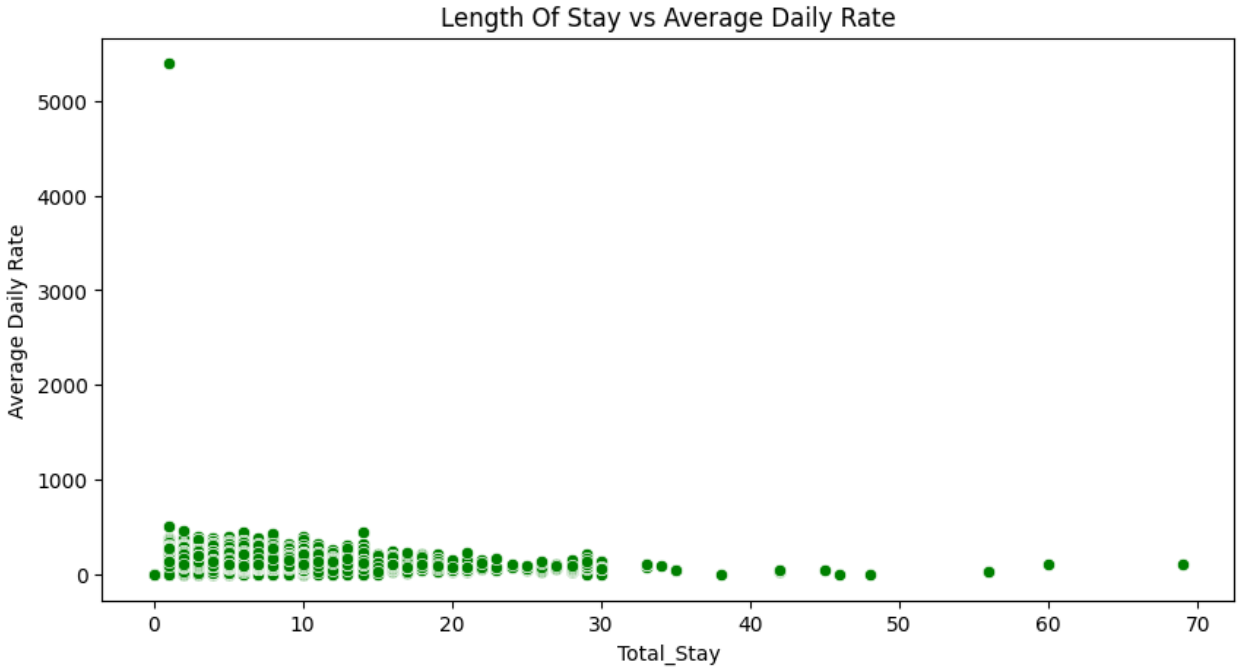
For the above output we would formulate few questions

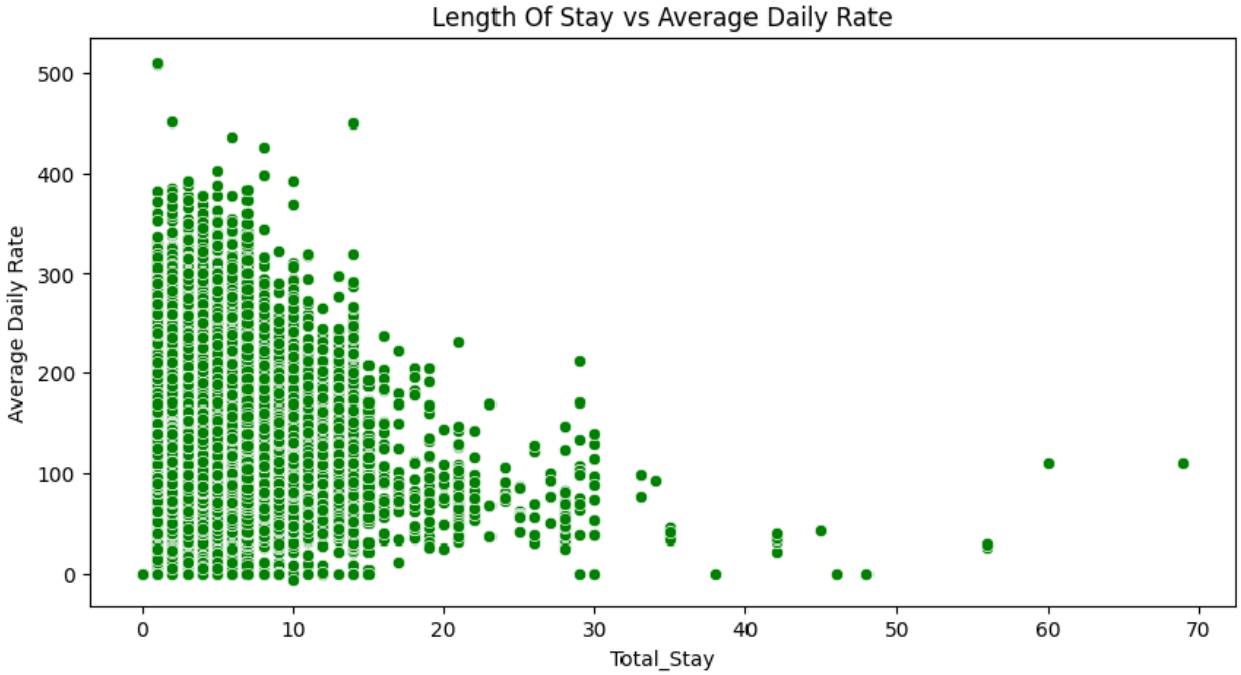
1. **Is the number of days on the customer's waiting list affected by lead time?**



From the above graph, we may conclude that with a shorter lead time, the customer may have to wait for a longer period of time.For example, if a guest makes a reservation with very little notice (less lead time), the hotel may have limited availability or may need to hustle to prepare the room, perhaps causing check-in or room ready delays.

2. Does the length of stay affect the average daily rate..?

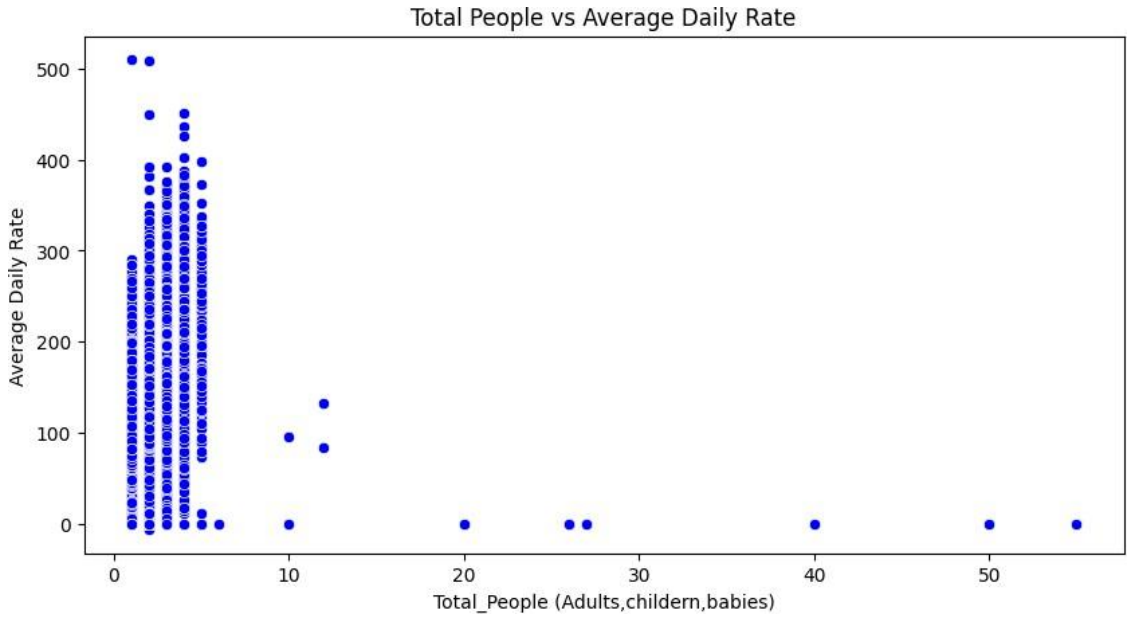


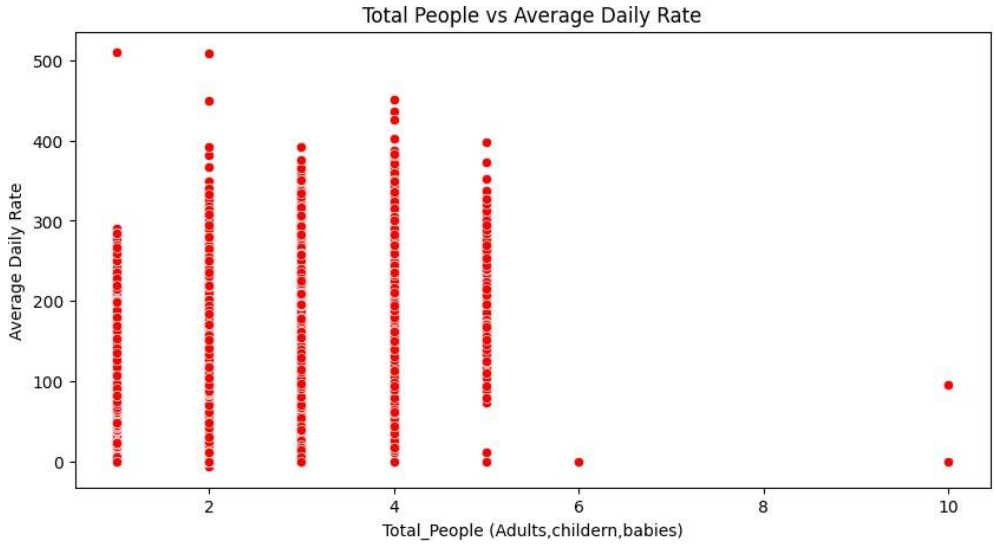


From the above two graphs the first scatter plot has a higher adr rate which is outlier so we will remove it for better analysis. After removing outliers The Second scatter plot

shows that as the length of the total stay grows, the average daily rate drops. This suggests that the customer can obtain a better deal if he or she stays longer.

3. Does the total number of people affect the average daily rate

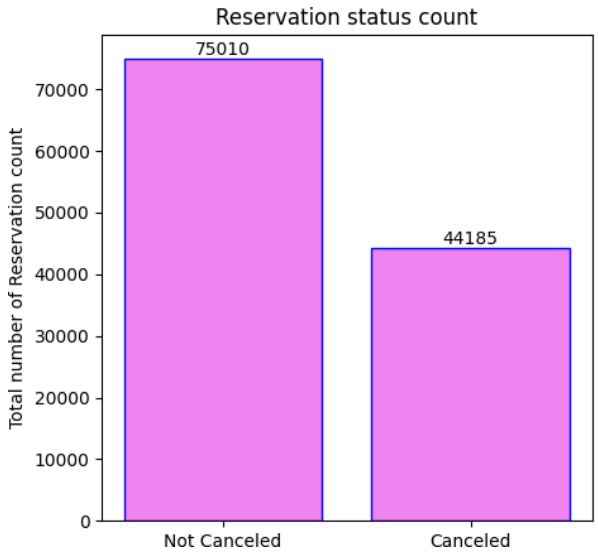




The scatterplots above show how the overall number of persons influences the average daily rate. In the first scatter plot, we can notice a few outliers, such as total_people being greater than 10, which is the least likely scenario. So we're

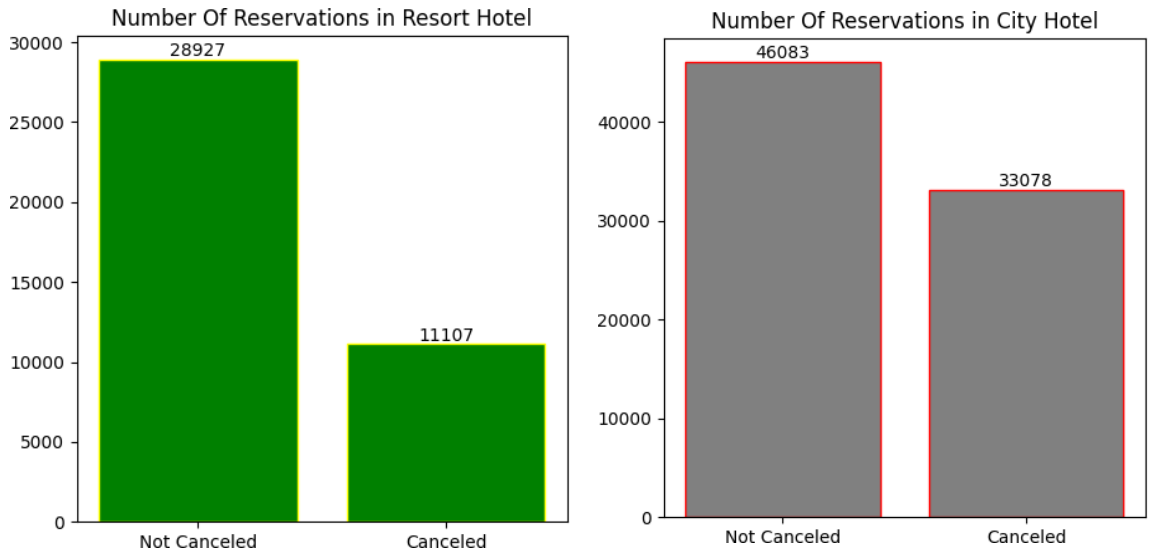
eliminating total_people>10. Except for a few outliers, the average daily rate increases as the number of people increases.

Q1: What is the total number of reservations at both hotels?



The resulting graph shows that over 37% i.e 44185 of the reservations are canceled. But still we cannot determine which hotel is having high cancellation and booking rate. So, we would do further analysis.

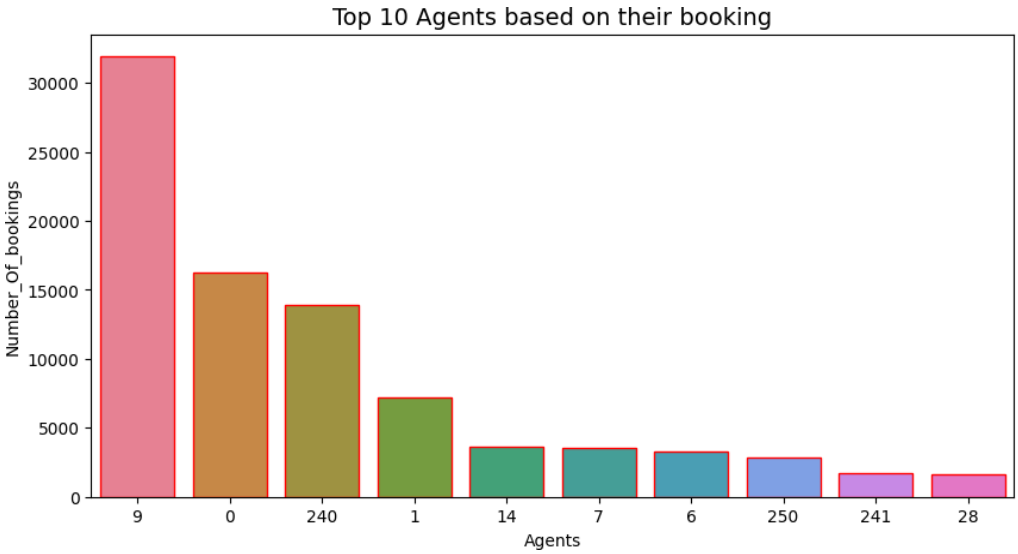
Q2: What is the number of reservations at resort hotels and City Hotels ?



The following graph shows that 72% of resort hotels have not canceled their booking. And 58% of city hotels have not canceled their booking.
Based on these reservations we can conclude that cancelation of bookings are higher in City hotels when compared with Resort Hotels.

Based on this data, we may conclude that City Hotel should address consumer concerns about why their reservations are being canceled.

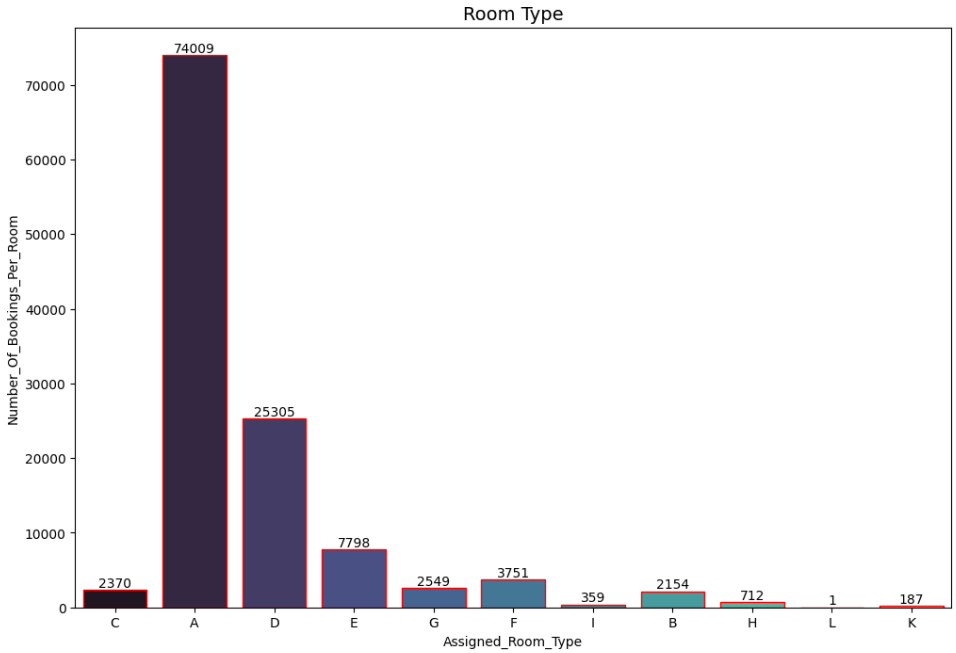
Q3: Which agents are making the most bookings?



The graph above demonstrates that agent 9 makes most of the bookings, and that the majority of bookings are made by individuals who avoid using an agent or company to book i.e., direct booking.

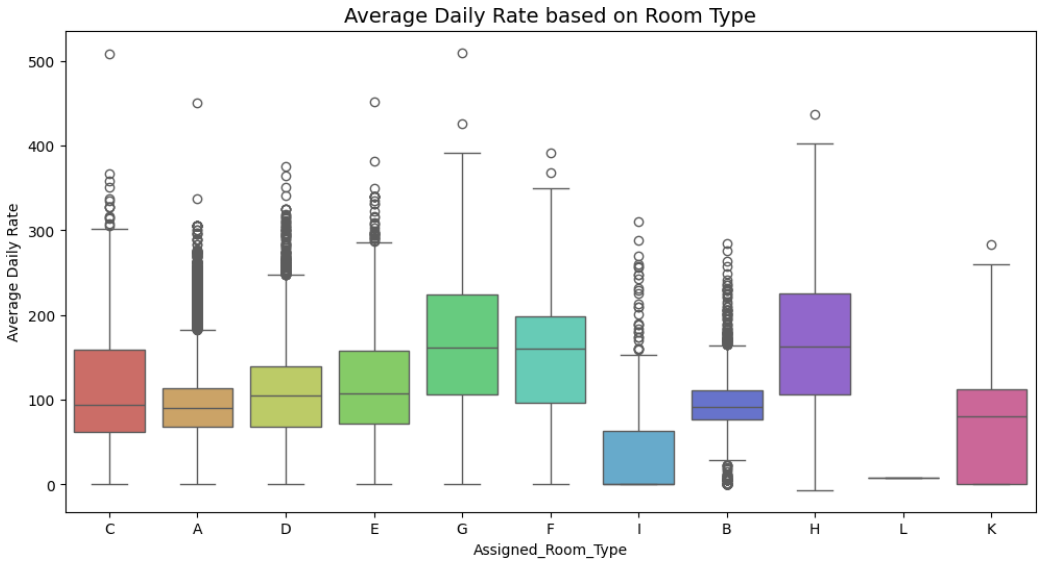
Conclusion: Hotels can provide offers or discounts to customers who book directly rather than through an agent or the company. So that hotels can increase their revenue by reducing commissions paid to agents or companies. Customers would be satisfied, and the hotel would benefit from free marketing.

Q4: Which Room Types are mostly preferred by customers?



According to the above box plot, the most wanted room types are Rooms A ,D and E hence these room types should be increased. So, that hotels have better chances of making higher revenue.

Q5: How Room type is affecting the Average Daily Rate?

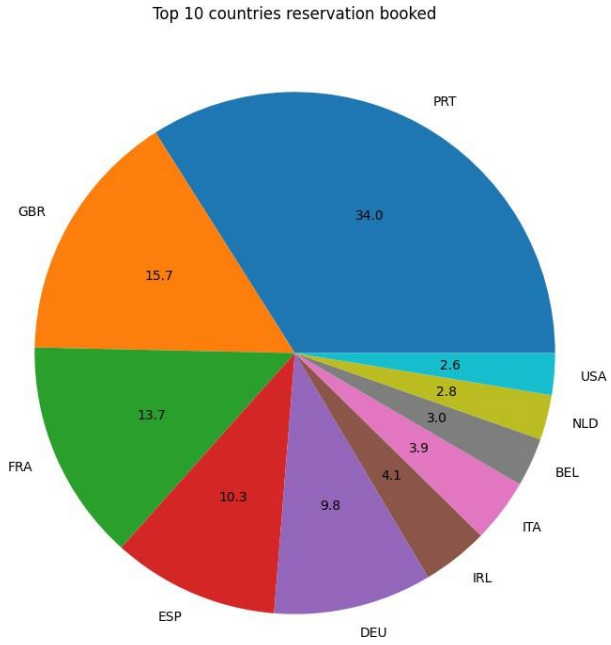


The resulting box plot provides insight, as the costs for Room A and Room D are relatively reasonable. As a result, this room type attracts the most bookings. Also, we can observe

that the ADR for room types H, G, and C is high, hence these rooms will be booked in lower numbers.

Conclusion: To optimize revenue, hotels can add room types A, G, and H. Customers can also choose Rooms A and D for a more affordable price.

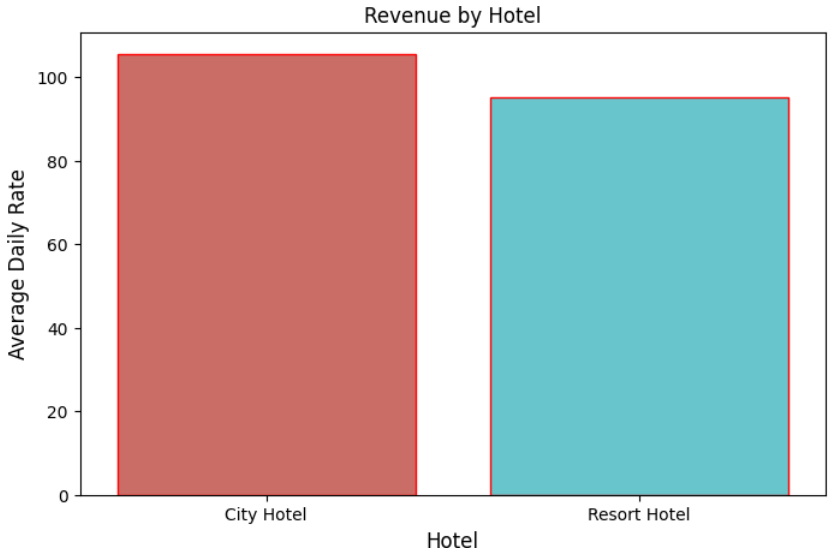
Q6: Which country accounts for the majority of customers?



The above pie chart shows that the majority of the customers are from Portugal, the United Kingdom, and France. The United States, Netherlands, and Belgium have the fewest customer bookings.

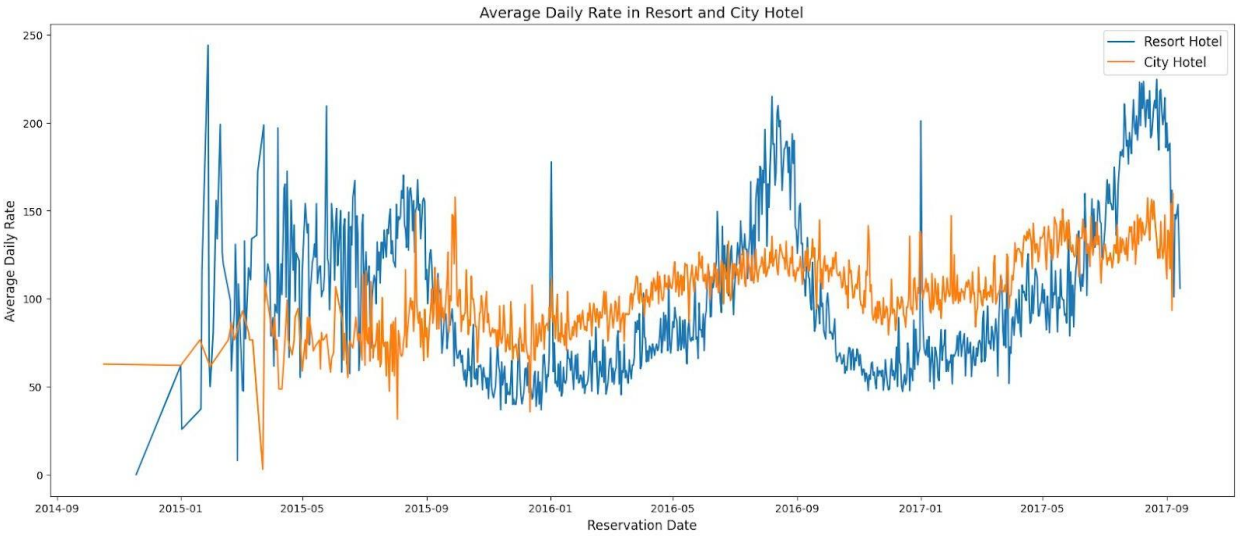
Conclusion: Because some nations have a smaller percentage of bookings, hotels can offer unique discounts and marketing in these countries to promote bookings from these regions of the world.

Q7: Which hotel has the highest revenue?



The above bar plot depicts income by hotel. The average daily cost at a resort hotel is slightly less than that of a city hotel. As a result, city hotels will generate greater revenue. Conclusion: Because the resort hotel receives the majority of its customer bookings, the resort hotel's average revenue rate will be significantly greater than that of the city hotel.

Q8: In which time frame the average daily rate for hotels is high?

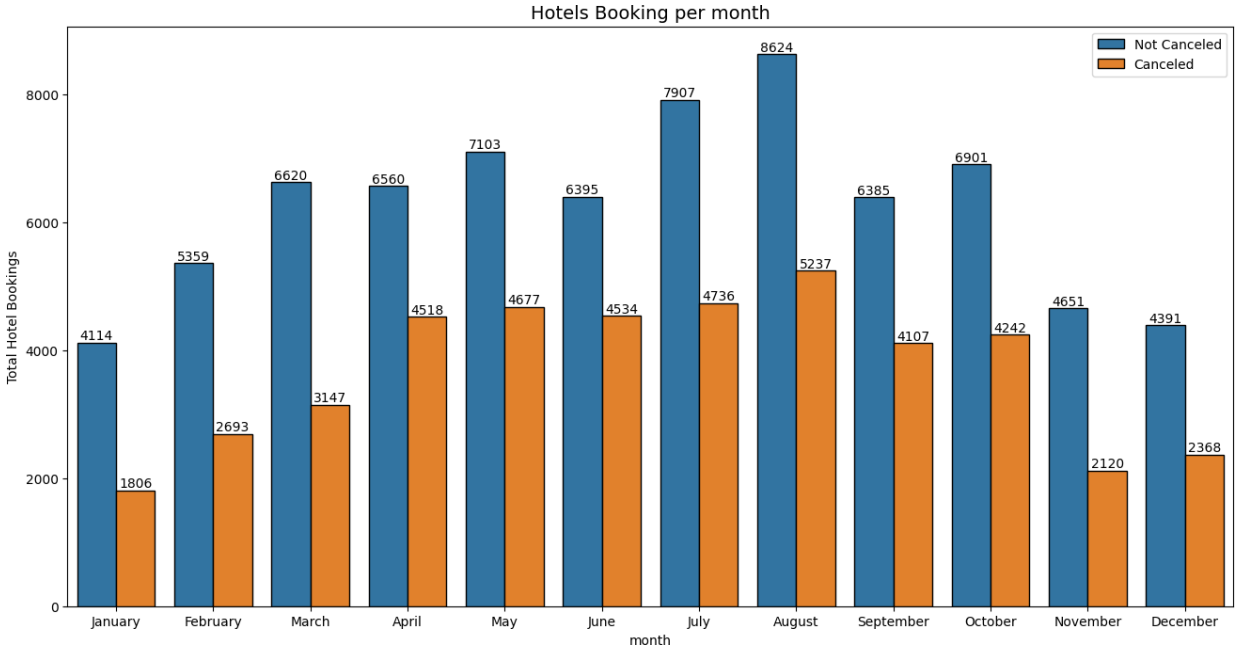


The line graph above displays the average daily rate for city and resort hotels from 2014 to 2017. According to the data presented above, the price of city hotels has stayed consistent over time, whereas the price of resort hotels has risen over time.

We can conclude from this that, because resort hotels have higher costs, the number of bookings will be lower than for city hotels. Additionally, we can see that resort hotels have greater costs. Customers may be charged a hefty cancellation fee, therefore the frequency of cancellations at resort hotels will be smaller than at city hotels.

Conclusion: From this perspective I recommend that resort hotels should evaluate their prices so that the number of bookings can progressively increase.

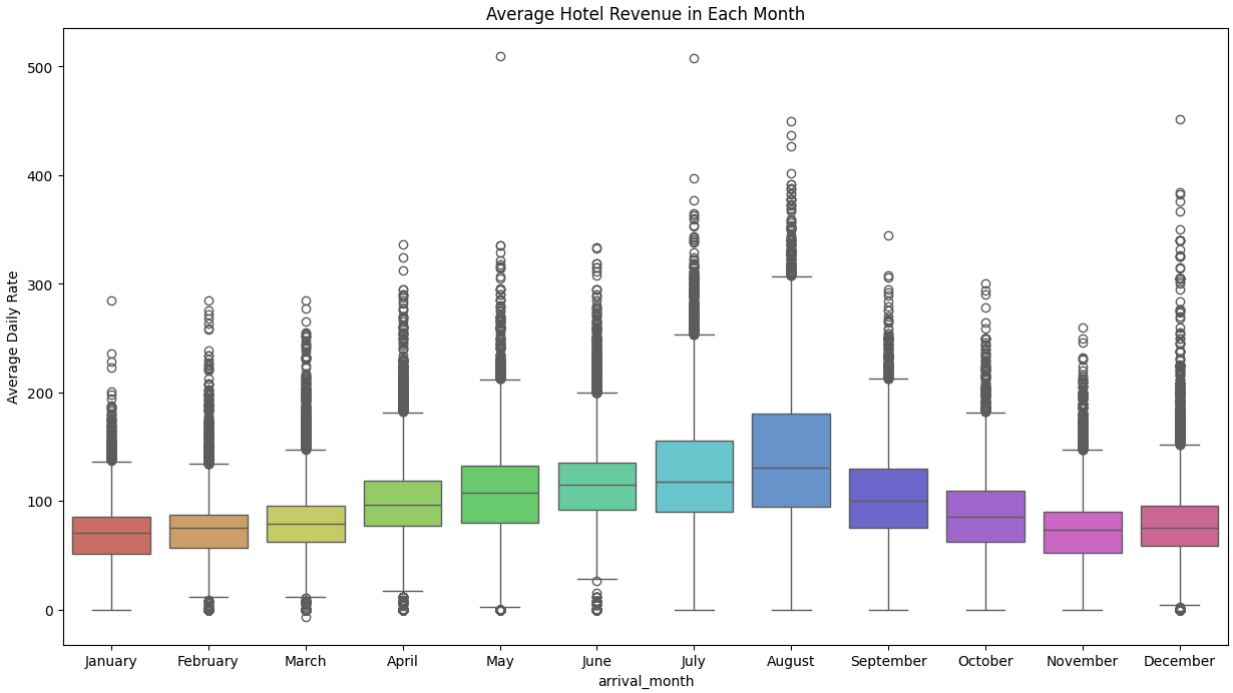
Q9: Which month has a higher amount of hotel bookings?



The accompanying bar graph shows an increase in client arrivals in May, July, and August due to summer vacation.

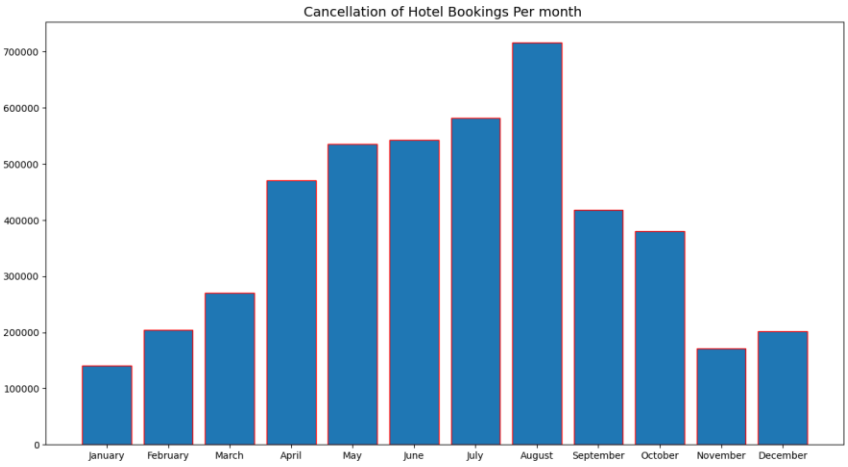
Also, we can see that as the number of bookings increases in these months, the average daily fee of hotels should steadily rise. To confirm this, we would plot adr by month.

Q10: Which month sees the highest hotel revenue?



The above box plot represents the average hotel revenue in each month. In a previous question I have addressed that in may, july and august months have the highest bookings. To verify it as we can see ADR grows from the beginning until the middle of the year, reaches a peak in August, and then falls till the end of the year. However, hotels generate significant revenue, particularly at the end of the year.

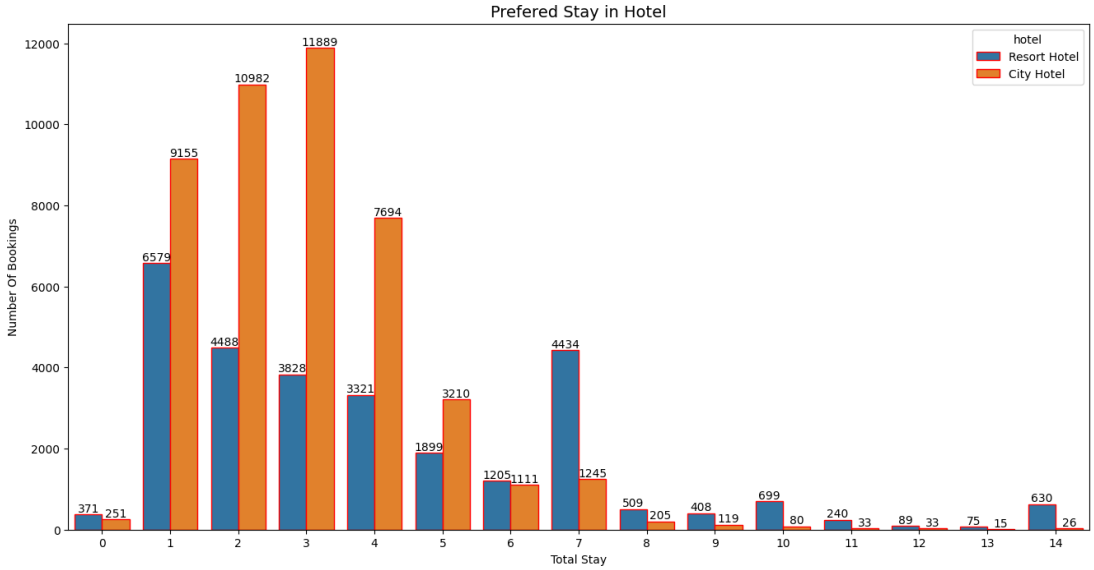
Q11: Which month has seen the highest hotel booking cancellation?



The bar plot above indicates monthly hotel booking cancellations. It shows that booking cancellations were high in May, July, and August. We can also conclude that increased

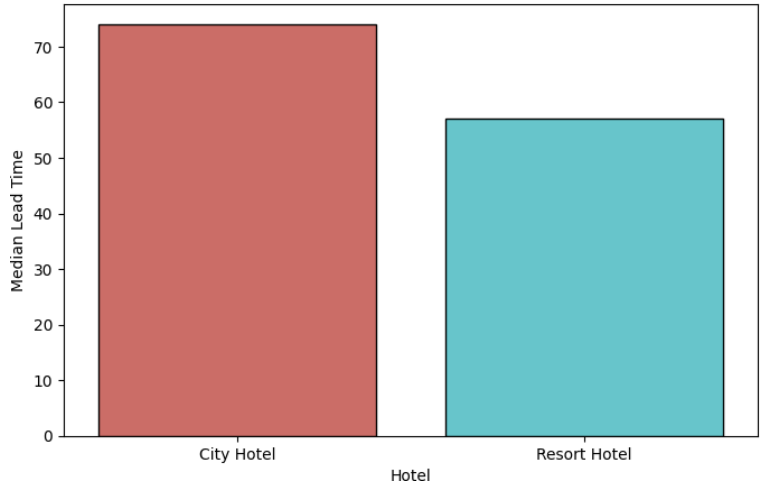
bookings in these months required customers to be on a longer waiting list, resulting in a higher cancellation rate.

Q12: Which hotel does the customer think is the best for a longer stay?



The bar graph above shows the hotel's preferred stay. City hotels are favored for stays of four days or less, whereas resort hotels are preferred for longer stays. Resort hotels are more popular for vacation plans, therefore stays are typically longer, whereas city hotels are not as popular.

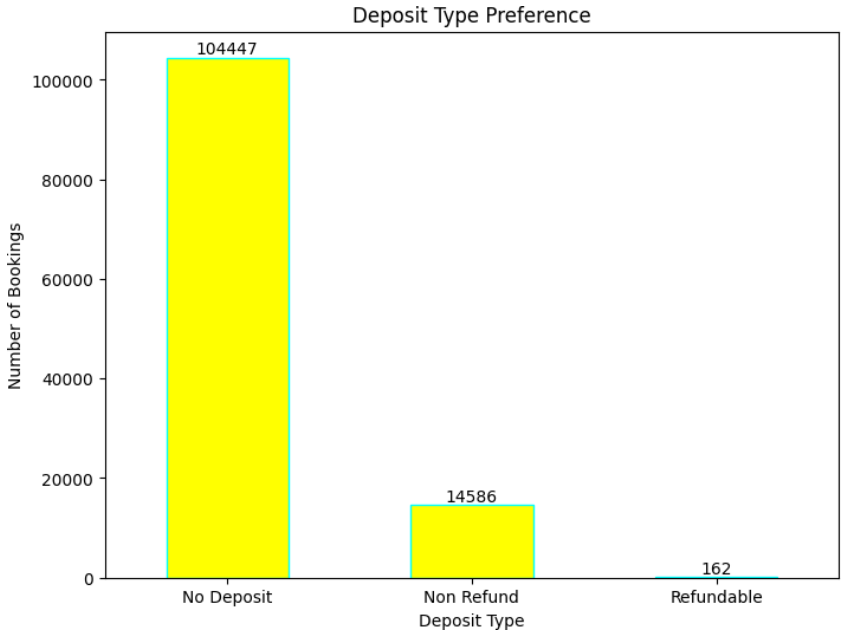
Q13: Which hotel has the longest lead time?



The above bar plot indicates which hotel has the highest lead time. The graph above indicates that the City Hotel's median lead time is somewhat longer. Additionally, the

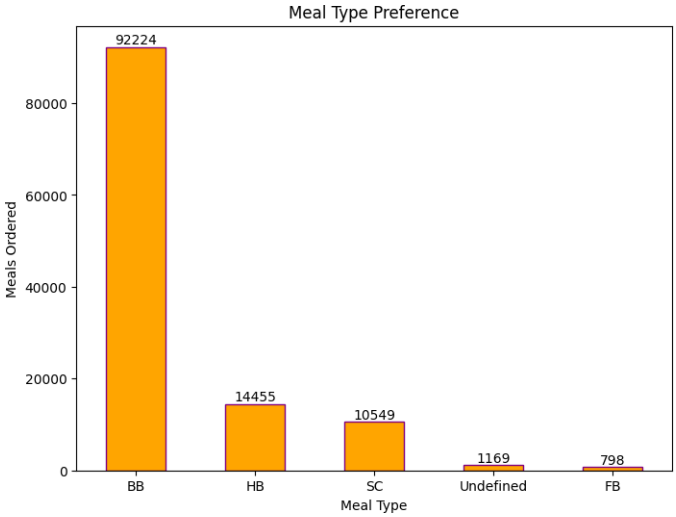
median lead time is noticeably longer in city hotels, indicating that guests typically schedule their hotel stays far in advance.

Q14: Which deposit type do most customers prefer?



The above bar plot depicts the deposit types preferred by customers. According to the above plot, the majority of consumers chose 'No deposit' as their deposit type since everyone wants the same level of certainty and no one wants to make a deposit or pay in advance.

Q15: Which particular type of meal does the customer prefer the most?



The above graph indicates the meal type preference chosen by most customers. According to the graph above, bed and breakfast is the most popular meal type among customers, followed by half-board meals.

3. Conclusion and Recommendation:

According to our overall data, resort hotels receive fewer bookings due to their higher price range, while city hotels have a higher cancellation rate due to their long waiting period. Furthermore, hotels do not receive many bookings at the end of the year, which is cause for concern.

To overcome all of these issues and meet consumer expectations, I would recommend that resort hotels give lower prices to repeat customers and adjust their pricing structure so that they can make a large number of bookings. City hotels must reduce the waiting time for clients to book by carefully organizing room reservations. So that clients do not have to wait any longer to confirm their booking. Furthermore, both hotels might present deals and packages at the end of the year to encourage bookings during certain periods.