

Capstone Project

Hotel Booking Data Analysis

TEAM MEMBER:

ANKIT PATEL

SNEHA H V

SHRUNGA M



Point for Discussion



- Data Summary
- Step involved in data analysis
- Most of the booking arrived from which country
- Top 10 Countries visited by the guests
- What is average daily rate(ADR)?
- Why ADR?
- Distribution of ADR (Average price as per room type)
- Distribution of ADR As per Month
- Analyze Special requests done by customers?
- Which month has the highest number of guests?
- Customers most preferred type hotel
- Booking by Market Segment
- Demand for different room types
- Which month has the highest number of cancellation
- Which hotel has the highest number of cancellation
- Which Year has the highest number of cancellation
- Deposit Type
- Which was the most booked accommodation type(Single, Couple, Family)
- Meal Analysis
- Car parking space analysis
- What is the relationship between special requests and cancellation?
- How long do people stay at the hotels?
- What is the relationship between lead time and cancellation?
- Conclusion

Data Summary :

Based on the initial assessment we found that the data was pretty much clean except for some missing values in a few columns. Upon using the `info()` method we draw out the following key insights about the data:-

- The dataset has a shape of (119390, 32) which means that it contains approximately 1.2 lakh rows and 32 columns.
- Our Dataset has 4 columns with float64 dtype, 16 columns with int64 dtype, and 12 columns with object dtype.

In our Dataset, we observed null values in the following columns:

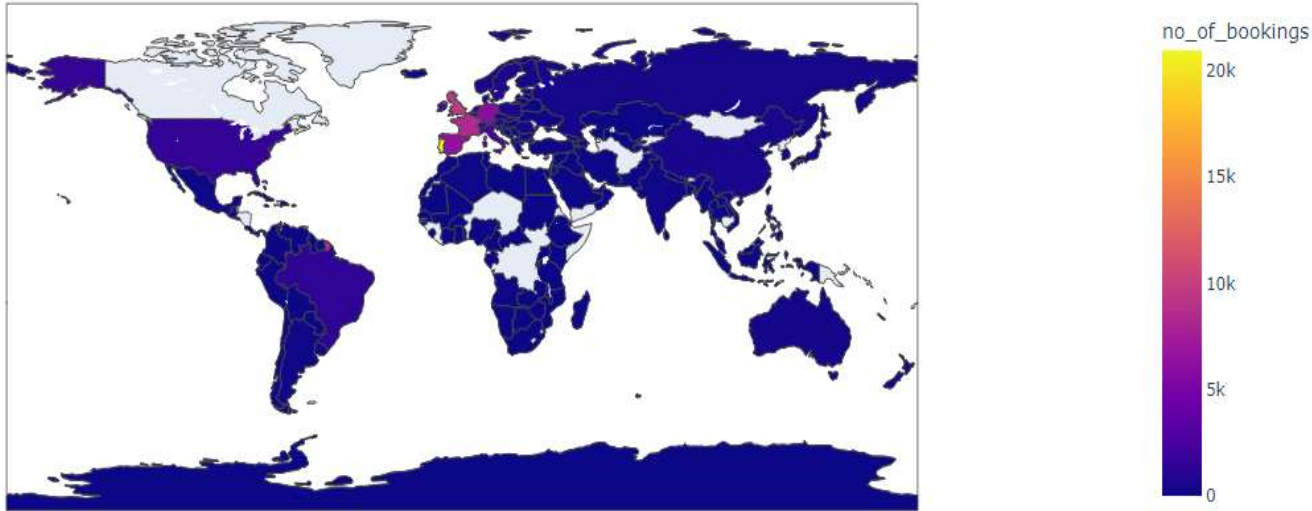
- 4 null values in the children column
- 488 null values in the country column
- 16,340 null values in the agent column
- 112,593 null values in the company column

Steps involved in the Data Analysis:

- **Framing the questions:** Before any form of analysis, it is important to frame the questions that we want to know from the data.
- **Filtering out the ideas:** After getting a long list of questions and assumptions that we want to solve from the dataset.
- **Cleaning the Data:** As mentioned before, the dataset provided to us by the Alma Better faculty was pretty much clean from the start. However, it had four variables with null values, so we had to take them into account before proceeding further with our analysis.
- **EDA analysis:** By EDA we mean exploratory data analysis. In this, we looked at the data frame and decided our target variables (Important Columns) based upon which we were going to conduct further analysis.
- **Visualization of Insights:** After we completed the analysis of our data, we used matplotlib and seaborn libraries of python to present our analysis graphically.
- **Drawing Conclusions and Finding Answers:** Finally, we wrapped up each analysis by drawing out conclusions from them

Most of the booking arrived from which country

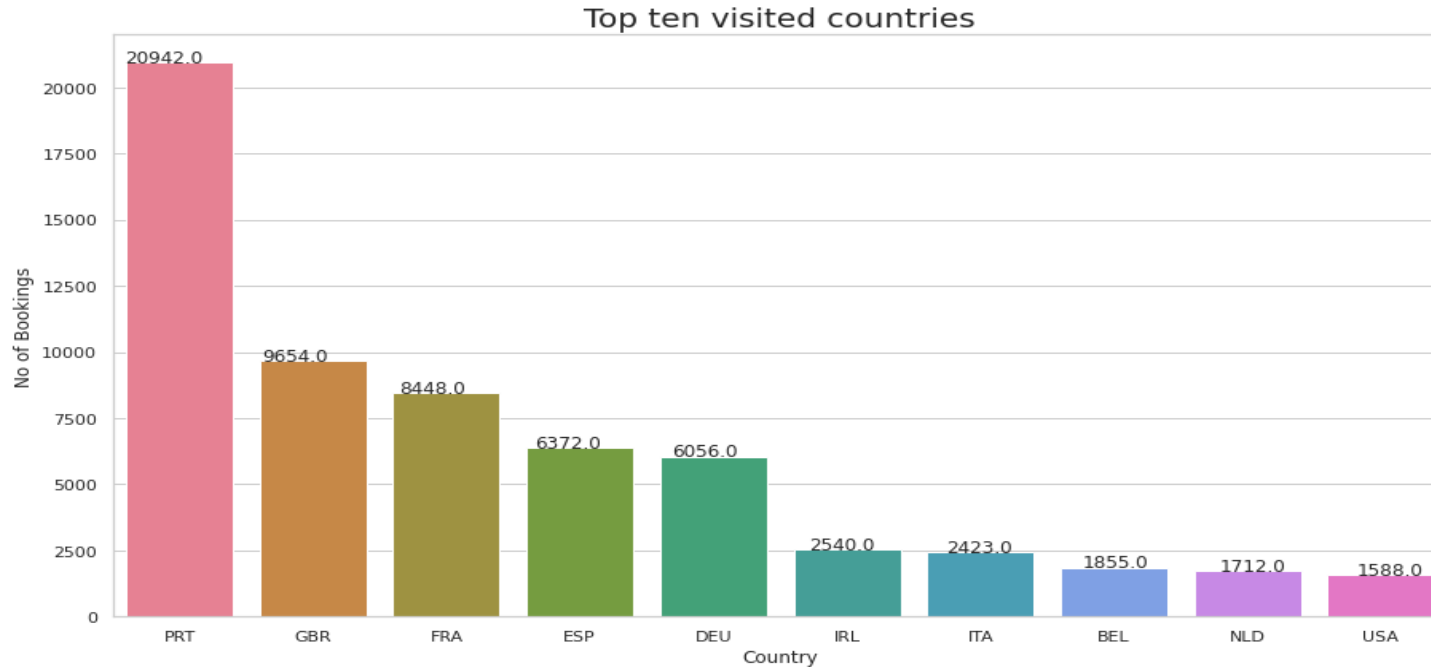
Number of Bookings per Country



Insight:

- Most of the bookings come from European countries.

Top 10 Countries from where the Guests arrived



Insight:

➤ Top arrived guests Countries are Portugal, Great Britain, France, Spain, Germany, etc.

What is Average Daily Rate (ADR)?

The average daily rate (ADR) is a performance indicator used in the hospitality sector to measure the strength of revenues generated. It is measured as the total revenues generated by all the occupied rooms in a hotel or lodge divided by the total number of occupied rooms over a given time period. It is a simple average that shows the revenues generated per occupied room.


$$\text{ADR} = \frac{\text{Room Revenue}}{\text{\# Rooms Sold}}$$

Summary:

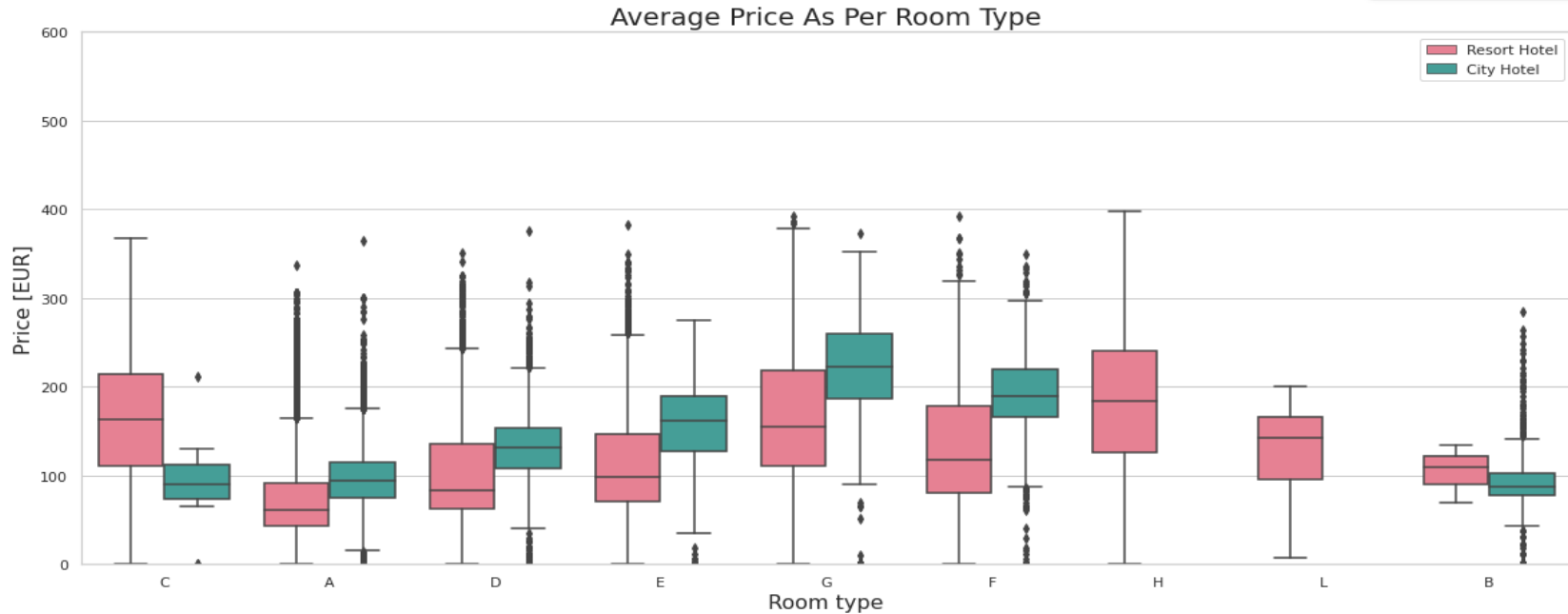
- The average daily rate (ADR) is a useful tool to maximize revenues in the hospitality sector.
- The ADR is measured as the total revenues generated by all the occupied rooms in the hotel or lodge divided by the total number of occupied rooms over a given time period.
- The average daily rate includes only the occupied rooms and not the total available stock.



Why ADR?

- The average daily rate (ADR) measures the average rental revenue earned for an occupied room per day.
- The operating performance of a hotel or other lodging business can be determined by using the ADR.
- Multiplying the ADR by the occupancy rate equals the revenue per available room.
- Hotels or motels can increase the ADR through price management and promotions.

Distribution of ADR (Average Price As Per Room Type) :



Insights:

- Room type 'G' in City Hotel have more average price as compared to other room type.
- In Resort Hotel type 'H' has shown the highest average price among other room type.
- Though there not much difference in between them.

Distribution of ADR As Per Months



Insights:

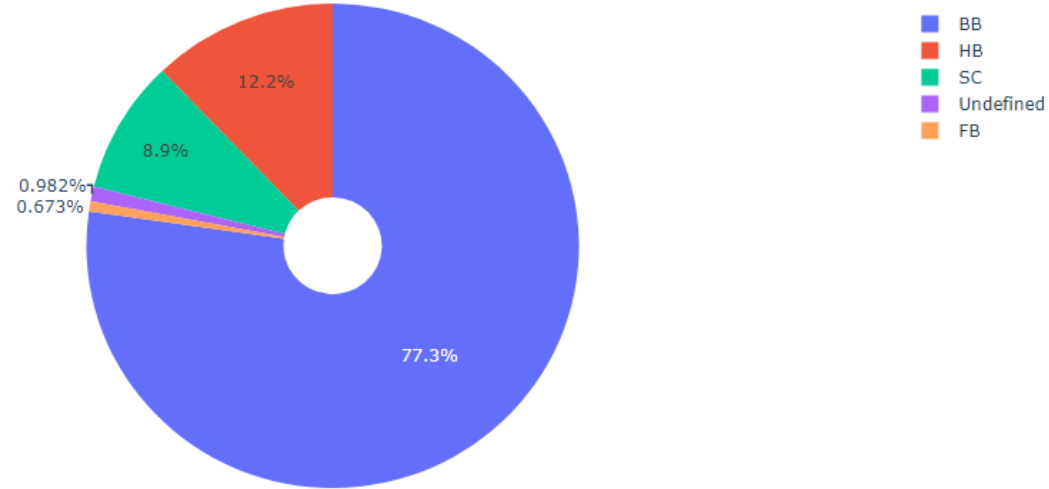
- August is the best month for both the hotels business because average adr is the highest in August.
- Average ADR is good in May, June, July, August as compared to other months.

Conclusion:

- Hotels are getting the best business during May-August.
- Hotels can make strategies to get more business during May-August.
- Average ADR is increasing May - August that shows hotel business is growing every year with a decent number.

Meal Analysis

Meal Analysis

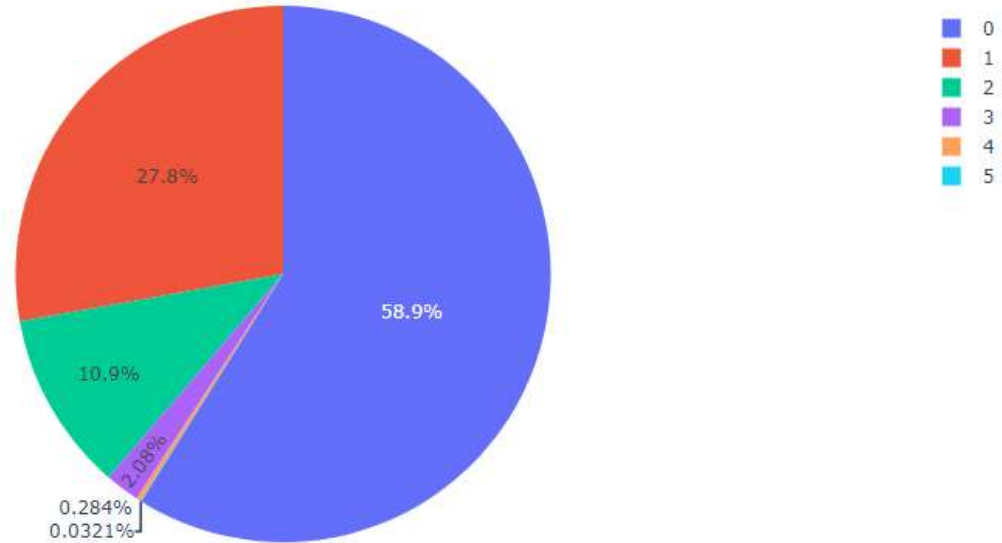


Insight:

- Almost 77.3% of people preferred **Bed and Breakfast** at the hotel where they stay.
- Most people prefer to eat lunch and dinner outside the hotel.
- People love to explore city local food at different different places.

Analyze Special Requests done by Customers?

total_of_special_requests



Conclusion :

- More than 50% of bookings don't have any special requests.

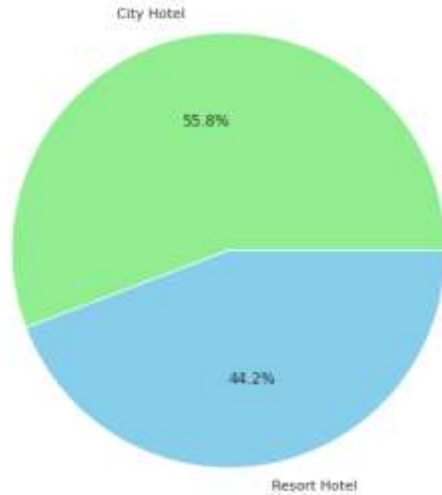
Which Month Has the Highest Number Of Guests



Insight:

- August have the highest bookings for City and Resort hotels, and January has the least bookings of both the hotel.
- Most people are coming to city hotels from May to August and fewer bookings from November to March.
- In Resort hotel most booking between May to August whereas in month June have less booking.

In which hotel do guest prefer to stay more?



Insight:

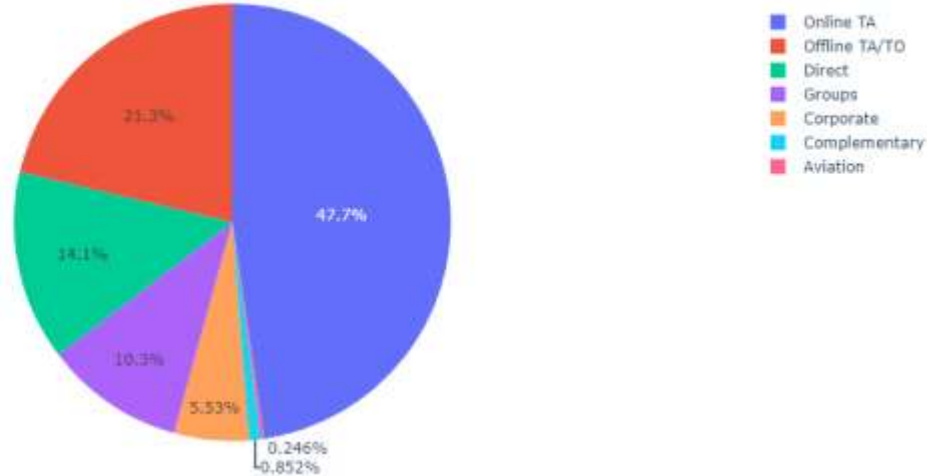
- Stay duration is almost equal for both the types.

Conclusion:

- City hotels are slightly more preferred.

Bookings by Market Segment

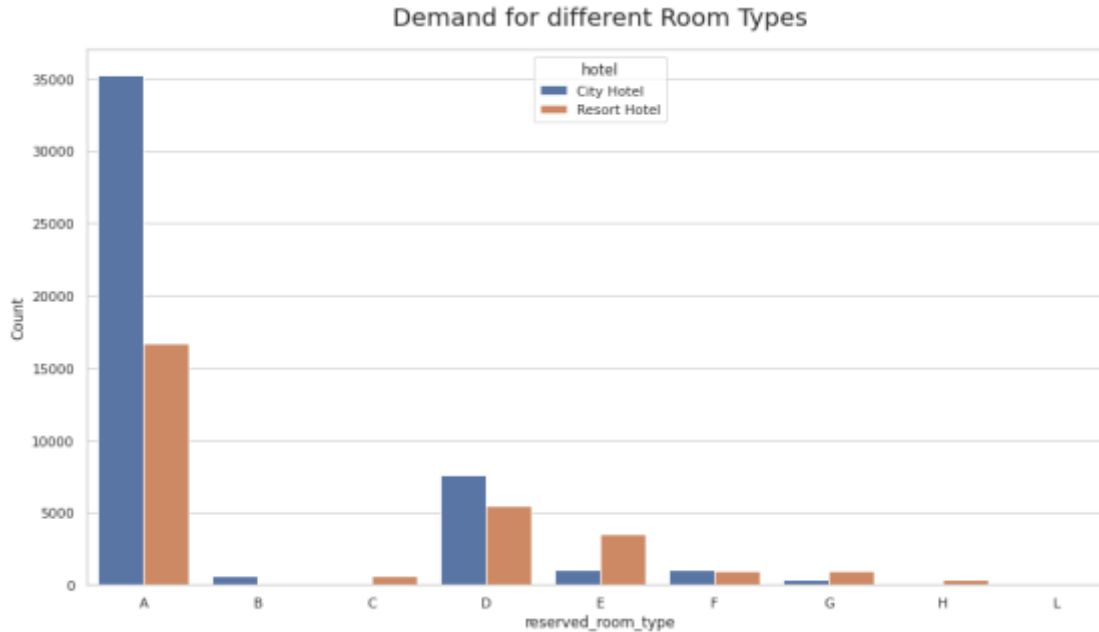
Booking Per Market Segment



Conclusion:

- Most Of The Booking Is Through Travel Agencies, Preferably Online Mode

Demand for different Room types



Conclusion:

- Demand for room type A in both hotels is highest.

Which Month Has The Highest No Of Cancellation



Insights:

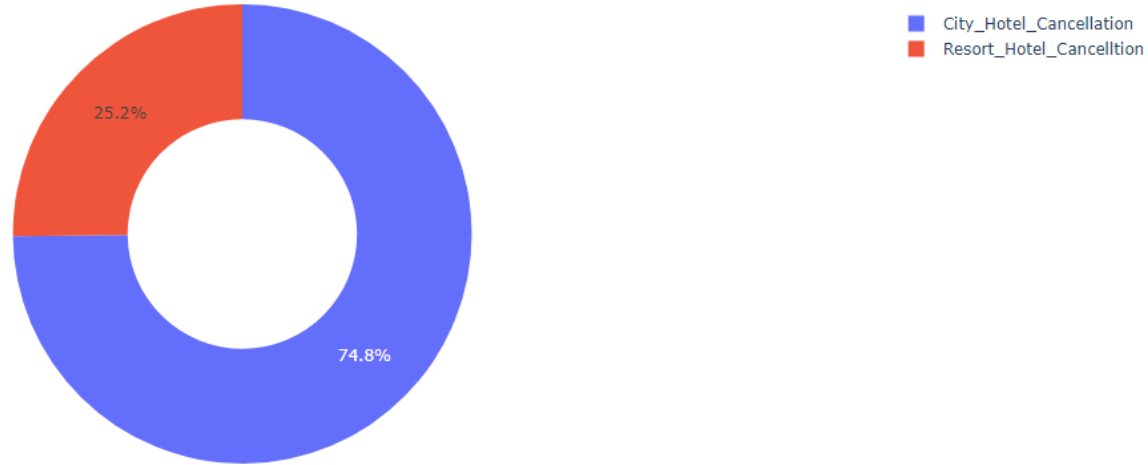
- The highest number of booking cancellations Months is between May to August.

Conclusion:

- Both City and Resort Hotel has the highest booking cancellation in August Month.

Which Hotel Has the Highest Number Of Cancellation?

Booking Cancelled



Conclusion:

- Highest Number of cancellations done in City Hotel, compared to the Resort Hotel.

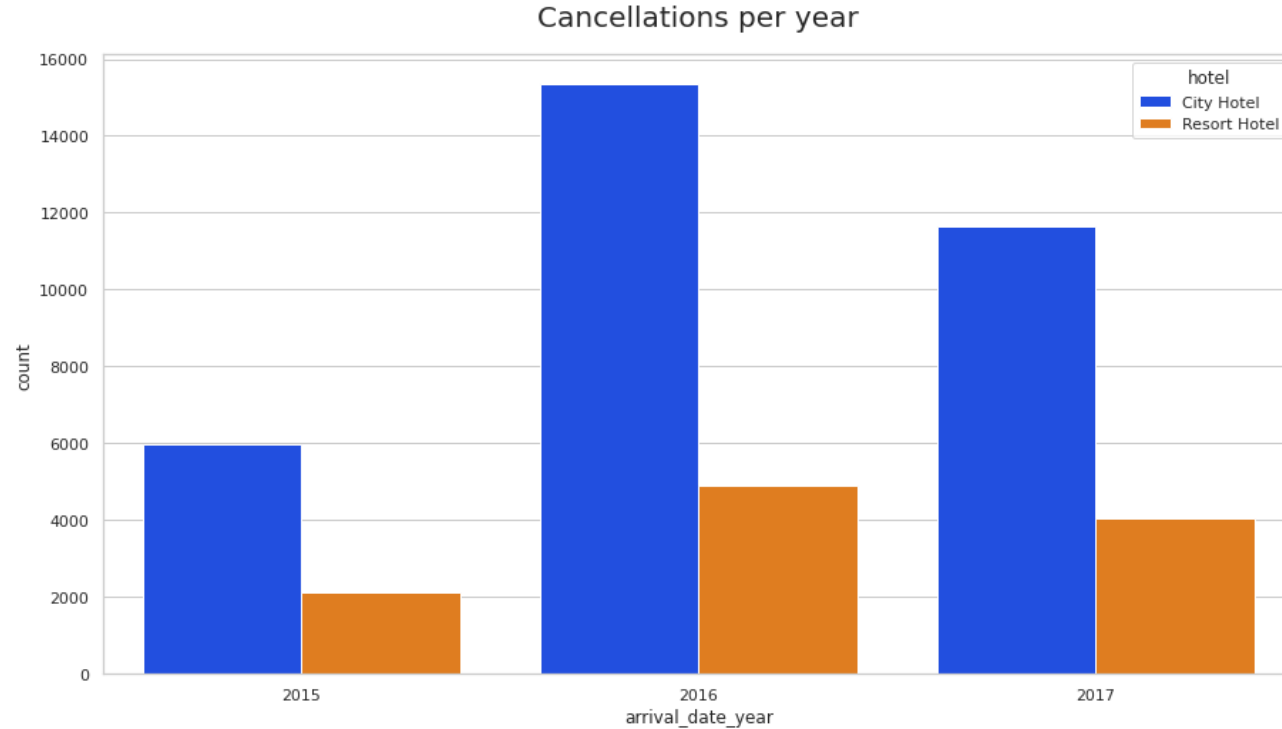
Which Month Has The Highest No Of Booking



Conclusion

- Both City and Resort Hotel has the highest booking in August.

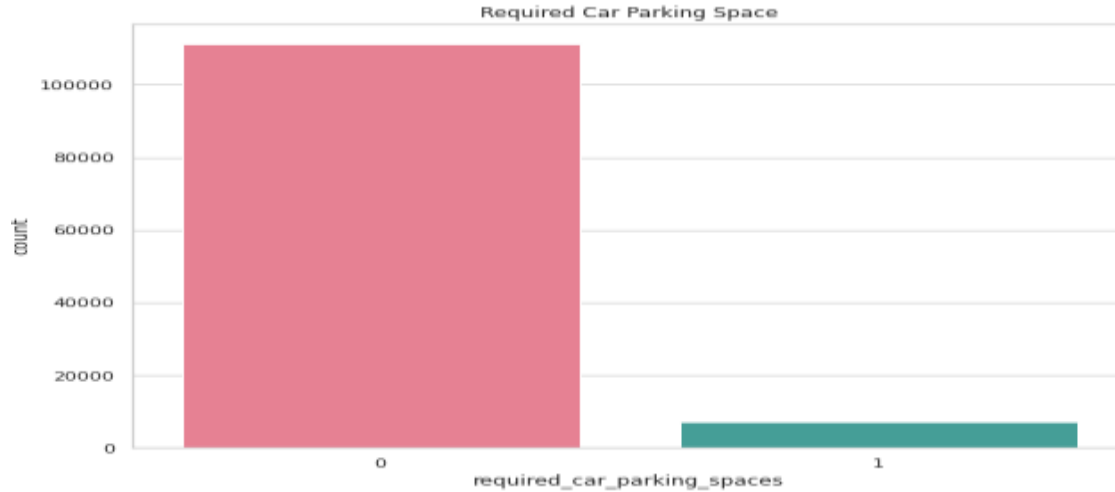
Which Year Has The Highest No Of Cancellation



Conclusion:

- Highest Number of cancellations done in City Hotel, compared to the Resort Hotel in 2016.

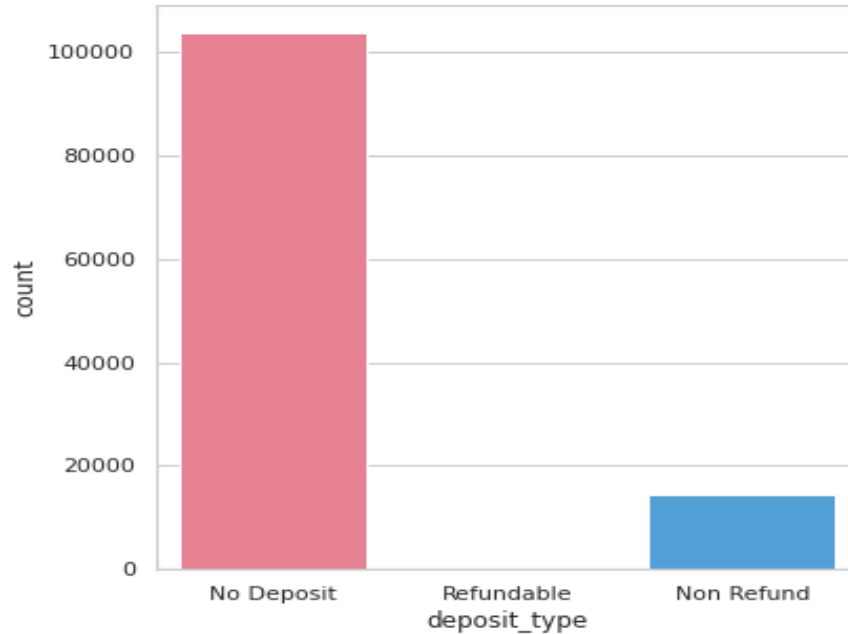
Required Car Parking Space?



Insight:

- Maximum hotels have 0-1 car parking space.
- The hotels are having international tourism so people prefer Travel with taxis and their local transport service.
- Hotel parking space does not affect the business.

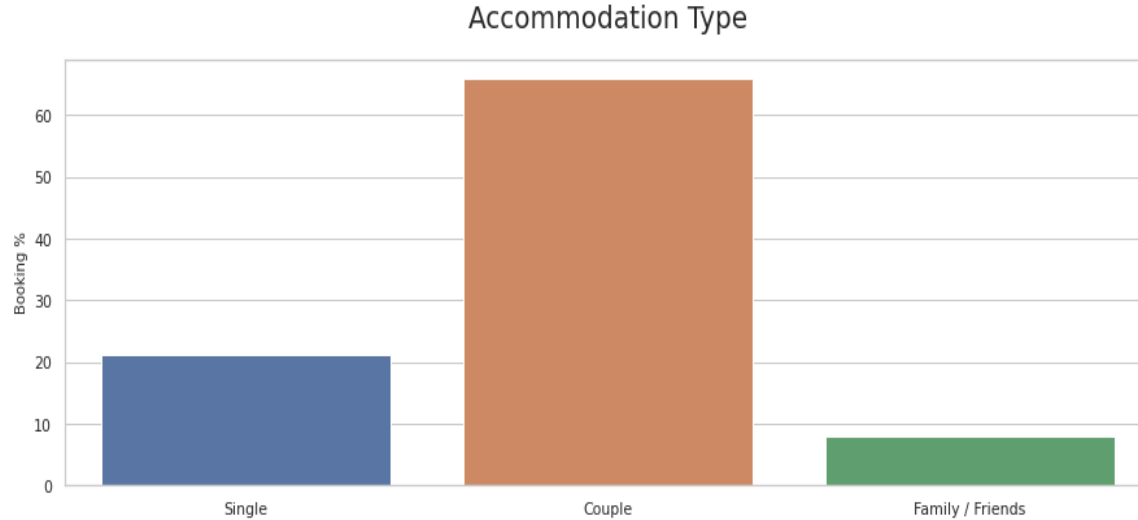
Deposit Type?



Conclusion:

- The Visitors Tends To Prefer Booking Hotel Which Doesn't require Any Deposit.

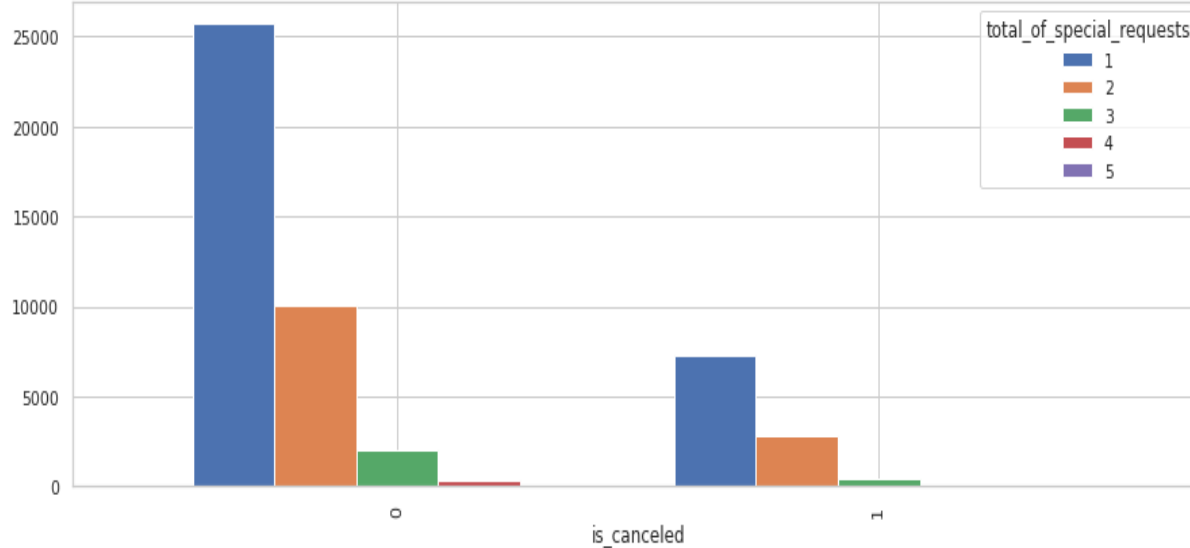
Which was the most booked accommodation type (Single, Couple, Family)?



Conclusion:

- A couple (or 2 adults) is the most popular accommodation type.
- So hotels can make plans accordingly.

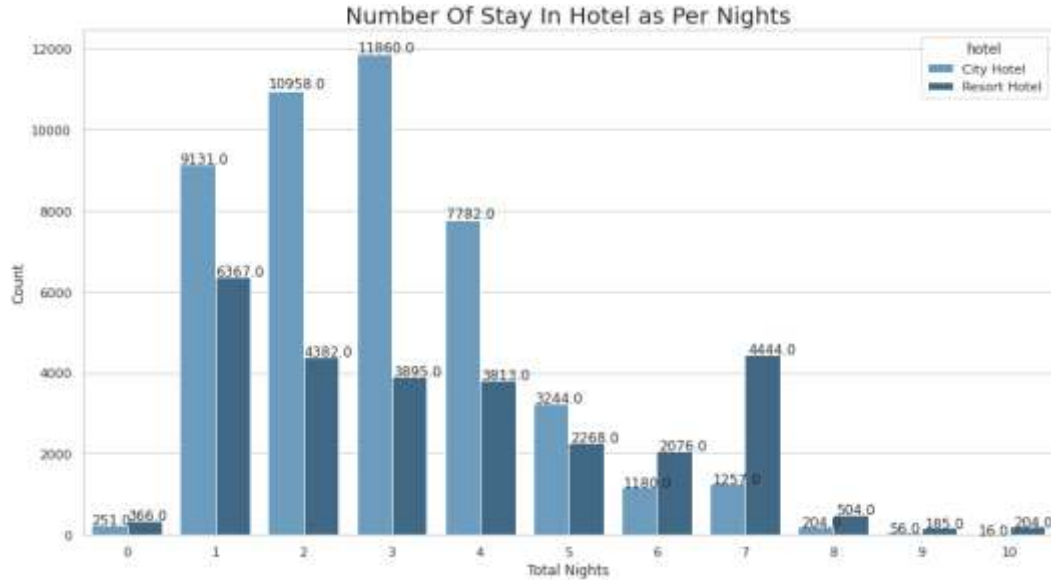
What is the relationship between Special Requests and cancellation?



Conclusion:

- If we have any special request then the chances of cancellation of booking are less.

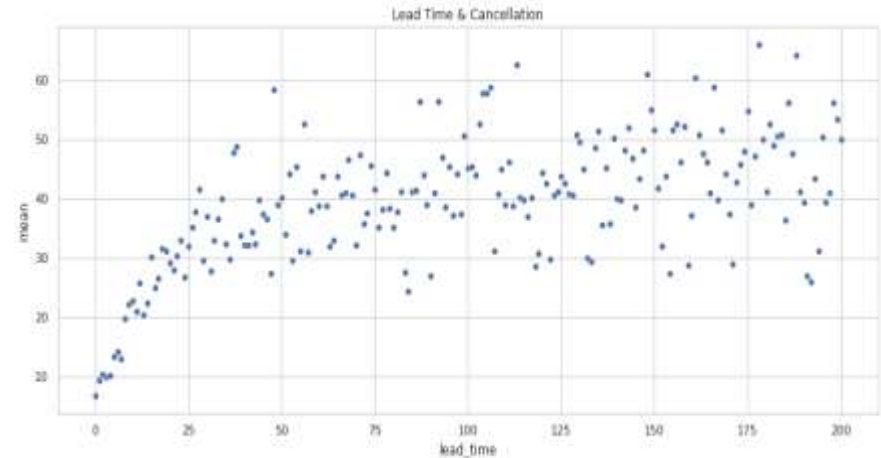
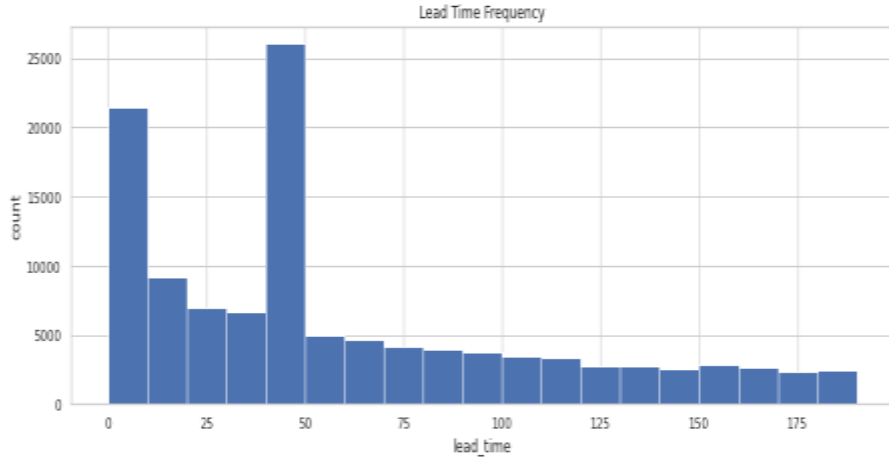
How Long Do People Stay At The Hotels?



Conclusion:

- For Both The hotels, guests preferred to stay less than one week.

What is the relationship between lead time and cancellation?



Insights:

- The highest number of lead times is 50 days.

Conclusion:

- Lead time has a positive correlation with cancellation.

Conclusion:

- Maximum number of booking has been observed in Month of August and lowest in January.
- Customer most preferred hotel is “city hotel.
- Portugal is the country which has maximum number of hotel bookings.
- Highest number of booking have seen for both hotels is 2016.
- Cancellation is more in city hotels as compare to resort hotel.
- Online TA brings maximum bookings and also maximum cancellations .
- The average ADR for hotels is in between 90-120 which is approximate 100.
- The average “ADR” of a city hotel is more than a resort hotel.
- Average ADR value is highest of the city hotels when 5 member stays at time in the hotel and 4 is for resort hotel.
- Anguila has best average adr value among all country.
- August is the best month for the hotel business in all years(2015-2017) because average adr is the highest of August for all years.
- Average ADR is increasing every year that shows hotel business is growing every year with a decent increment.
- Almost 77.3% of people preferred **Bed and Breakfast** at the hotel where they stay
- Maximum hotels have 0-1 car parking space.
- Hotel parking space does not affect the business.

Thank
you