Capstone Project
Hotel Booking Analysis

Team members
Mohammad Jibran
Siddhi H.Thakur

Points for Discussion

- ❏ **Data Summary**
- ❏ **Ratio of different types of hotels**
- ❏ **Which year hotel businesses were most booked**
- ❏ **Busiest months for hotels**
- ❏ **Hotel cancellation**
- ❏ **Special request by guest**
- ❏ **Market segment and distribution channel**
- ❏ **Pre deposit**
- ❏ **Hotel Booking rates**
- ❏ **Guest preferences**

Data Summary

Our topic for EDA Project is "Hotel booking demand" database.
This dataset contains the complete details of two type of hotels namely the City hotel and the Resort hotel, and includes details such as duration of stay, when bookings are made, number of adults, children among other things.

Some important columns description

hotel
- The category of hotels, which are two Resort hotel and City hotel.

is_cancelled
- The value of column show the cancellation type. If the booking was cancelled or not. Values[0,1], where 0 indicates not cancelled.

lead_time
- No of days that elapsed between entering date of booking into property management system and arrival date.

stays_in_weekend_nights
- No of weekend nights (Sat/Sun) the guest stayed or booked to stay at the hotel

stays_in_week_nights
- No of week nights (Mon - Fri) the guest stayed or booked to stay at the hotel

market_segment (A group of people who share one or more common characteristics, lumped together for marketing purposes)
- TA: Travel agents
- TO: Tour operators

distribution_channel (A distribution channel is a chain of businesses or intermediaries through which a good or service passes until it reaches the final buyer or the end consumer)

- TA: Travel agents
- TO: Tour operators

deposit_type

- No Deposit - no deposit was made
- Non Refund – deposit was made in the value of the total stay Cost
- Refundable – a deposit was made with a value under the total cost of stay

adr(average daily rate)

- Average daily rate = Sum of all loging transaction/Total no. of staying night

required_car_parking_spaces

- Number of car parking spaces required by the customer

Exploratory Data Analysis (EDA)

Importing Libraries :

We have to import all needed libraries those are required in project for data manipulation , cleaning and visualization.

Data Preparation and Cleaning :

Apply some command to clean and get information about dataset like data summary using describe(), shape, columns and info(), drop unwanted or irrelevant columns, treatment of NULL values. And many other commands for cleaning like finding and removing duplicates values.

Example – Column agents and column companies has many missing values making the columns irrelevant for our analysis, hence we removed them. Finding all Null and Duplicate values and dropping them . After doing the data cleaning we are left with 86913 rows and 29 columns.

Plotting different features against one another:

We use variety of graphs for comparing categories and Performance of model.
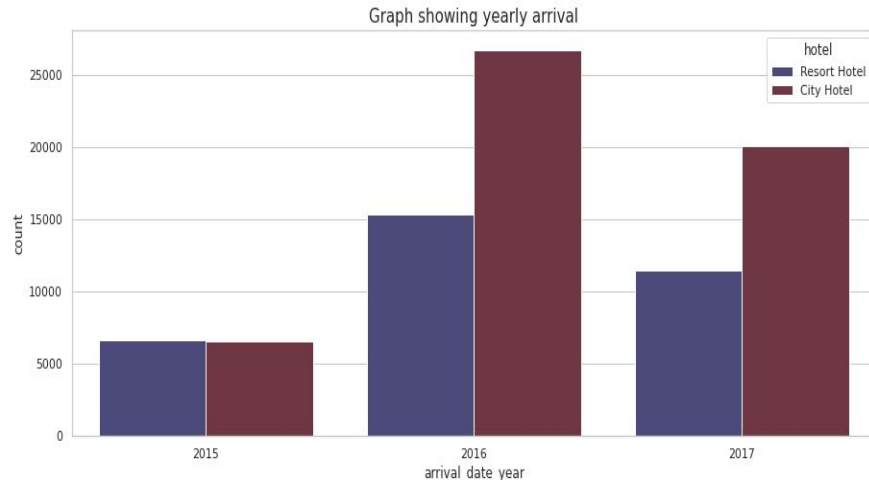
Ratio of different types of hotels

**We can see that the majority of the people preferred city hotel over resort hotel.**
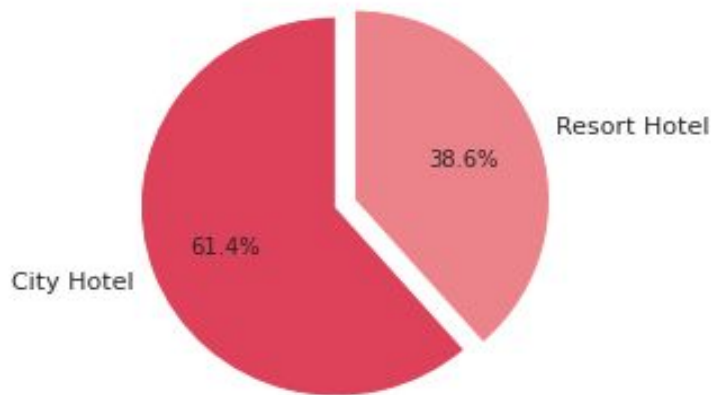
How many bookings made by each hotel yearly?

From this graph, we can conclude that maximum number of customers preferred City hotel. In 2016, both hotels got maximum business as compared to their their other business years.
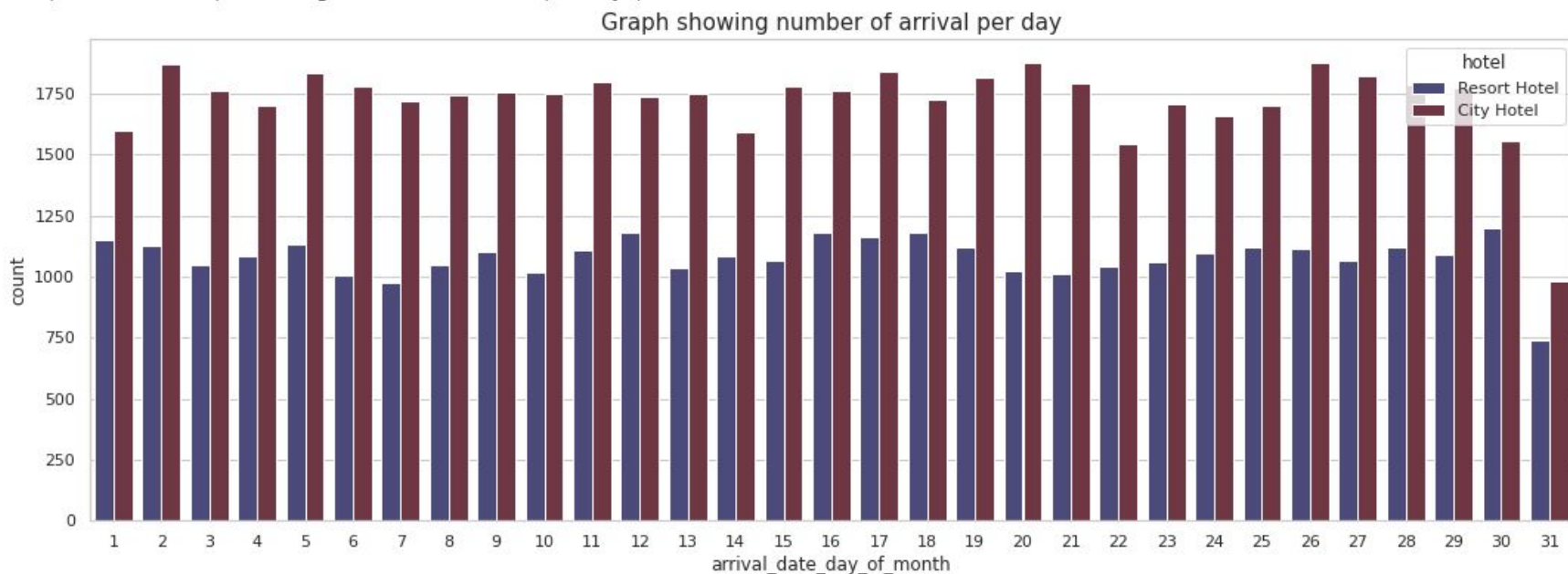
How many bookings made by each hotel every month and which is the Busiest month?

**This graph conclude that August month is the highest arrival month of customers.**

Graph showing number of arrival per month
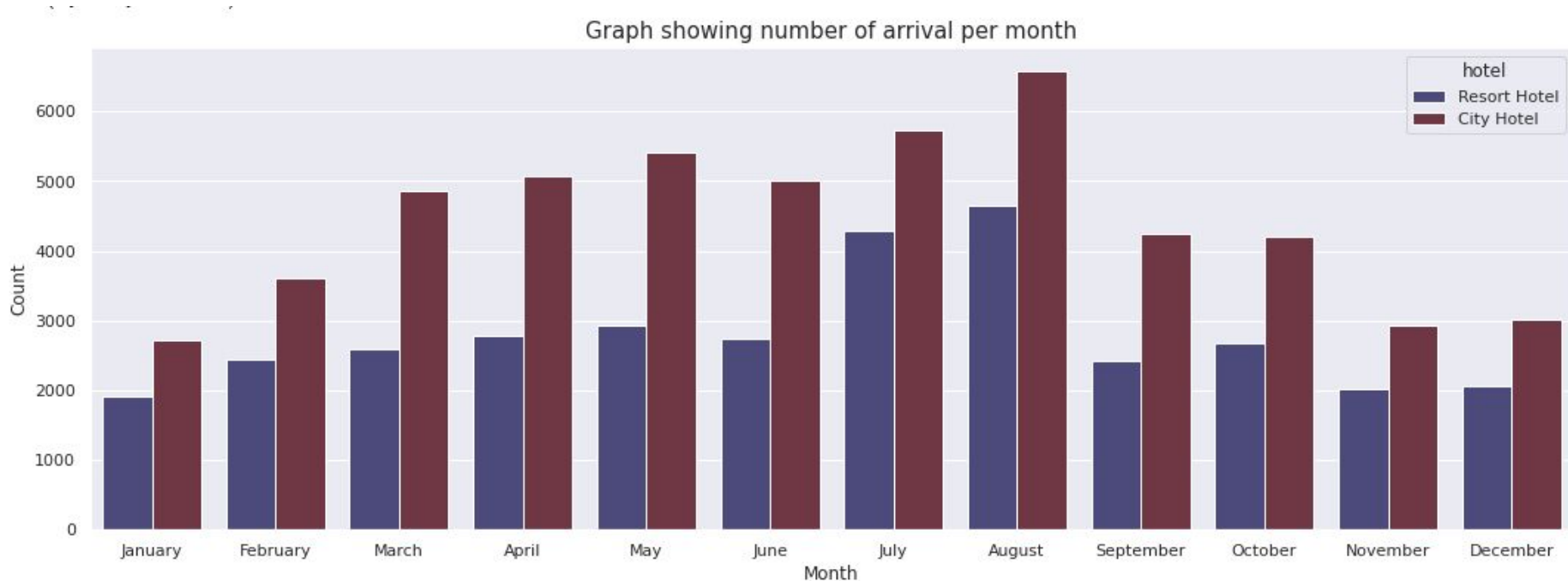
At which period of the month hotels demands are lesser

We can see that month ending has the least number of arrivals

Graph showing number of arrival per day

How often people cancel their booking.

From this graph it can be concluded, that majority of bookings are not cancelled of both hotels.
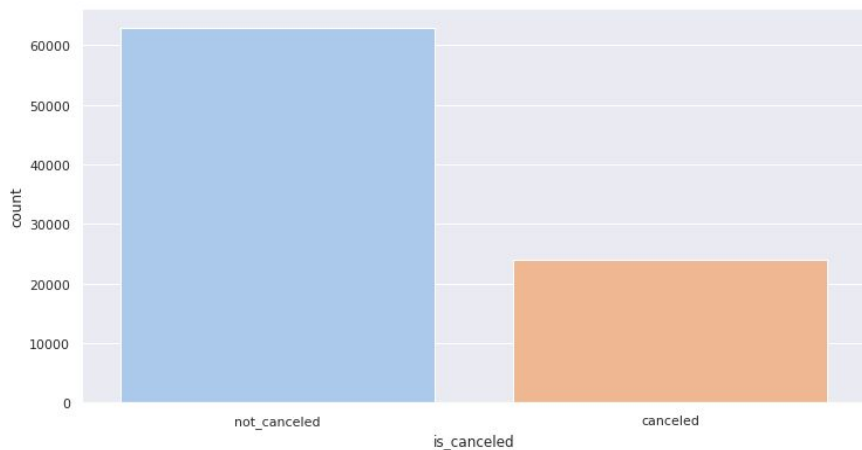
Which type of hotel usually face more cancellation than other.

**From this graph it can be concluded, that City hotel has highest cancellation rate.**

Which country's visitors didn't cancelled their bookings.

**From, this graph we conclude that Visitors from Portugal, Spain, French and UK didn't cancel their bookings**

How many guests requested for car parking?

From the bar chart It is clear that the majority of guests do not require car parking in both types of hotel.


Graph showing guests require Car Parking
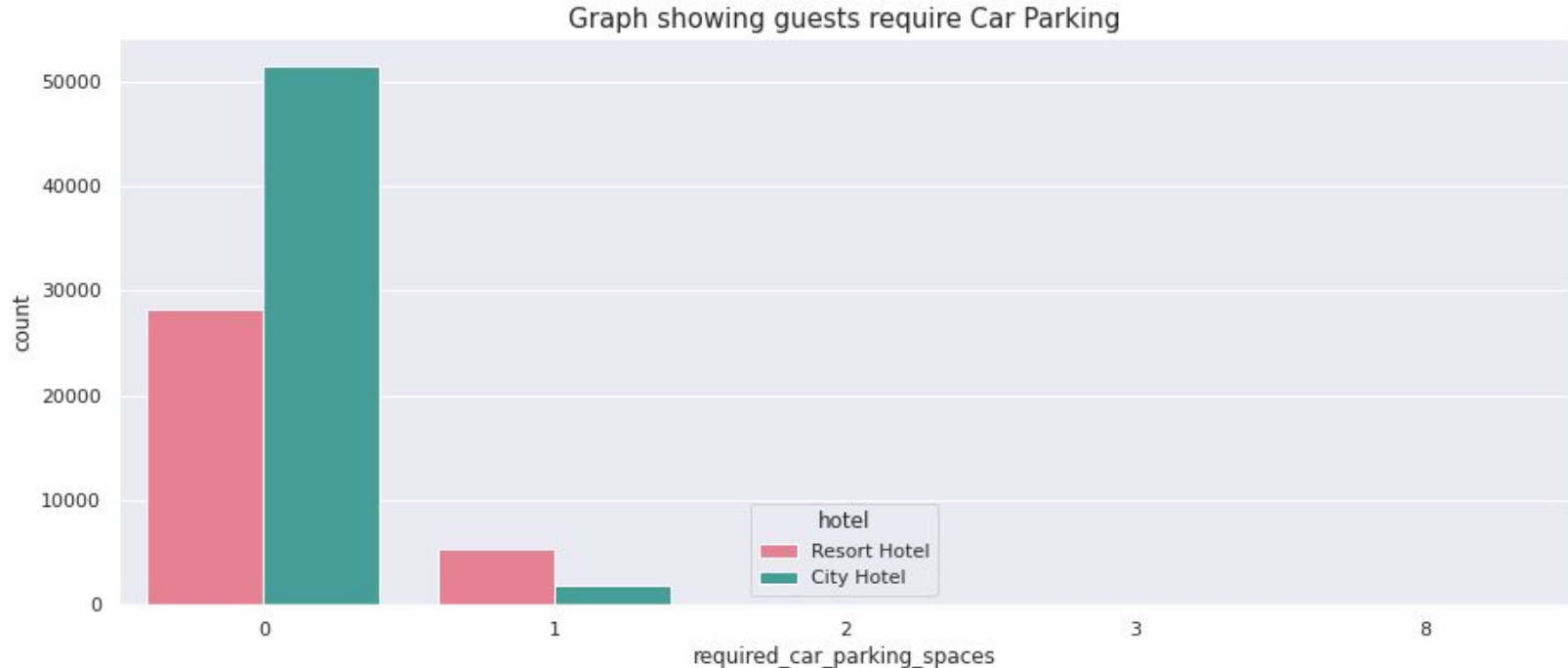
Which distribution channel gave most of the bookings?

As seen in bar graph majority of booking done by 'online TA' and TA whether online/offline have the highest distribution rates.
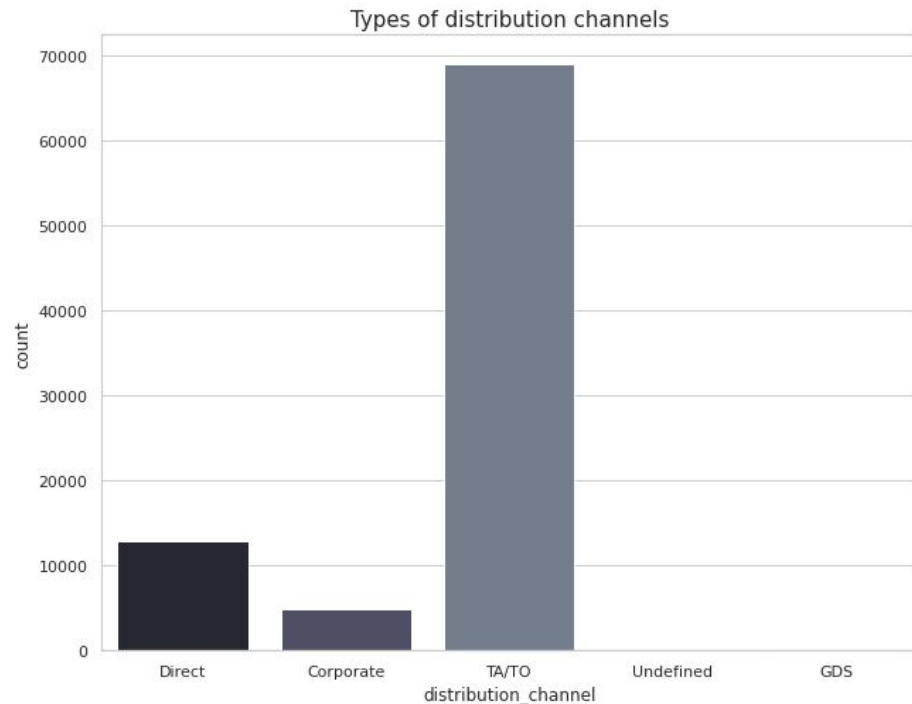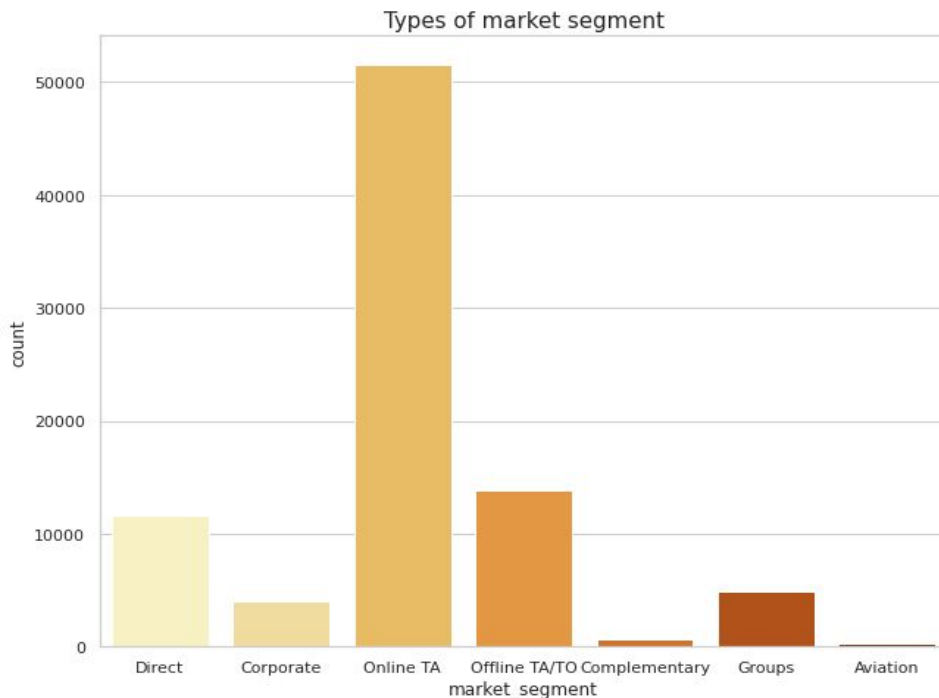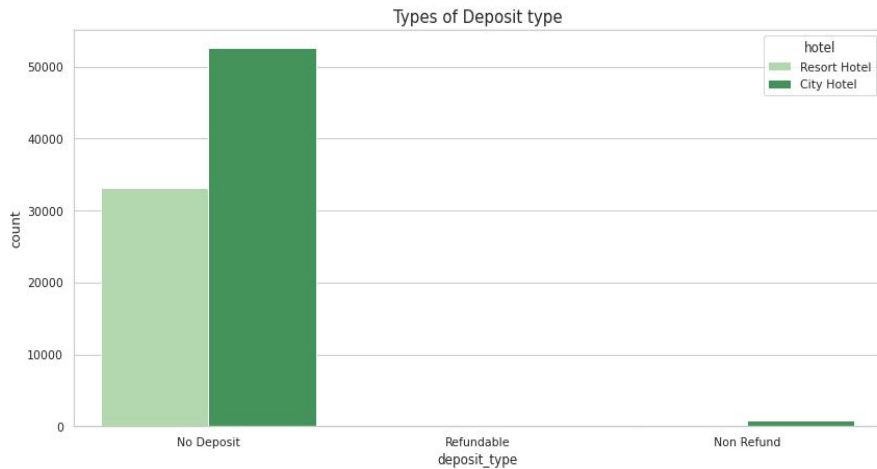
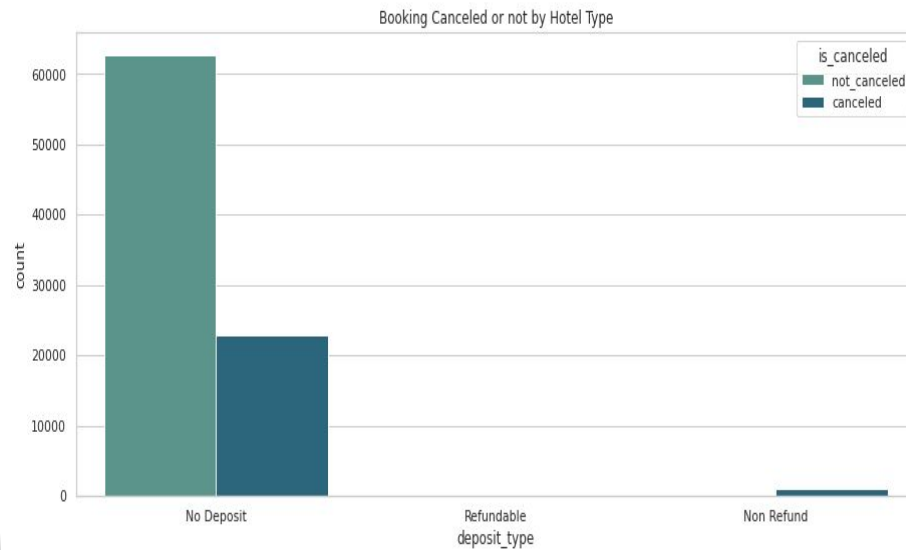# Which is the most preferred hotels types as per business segment?

This graph show that,both type of hotels have highest preference on 'online TA'.

Mostly prefered Hotel with respect to market Segment
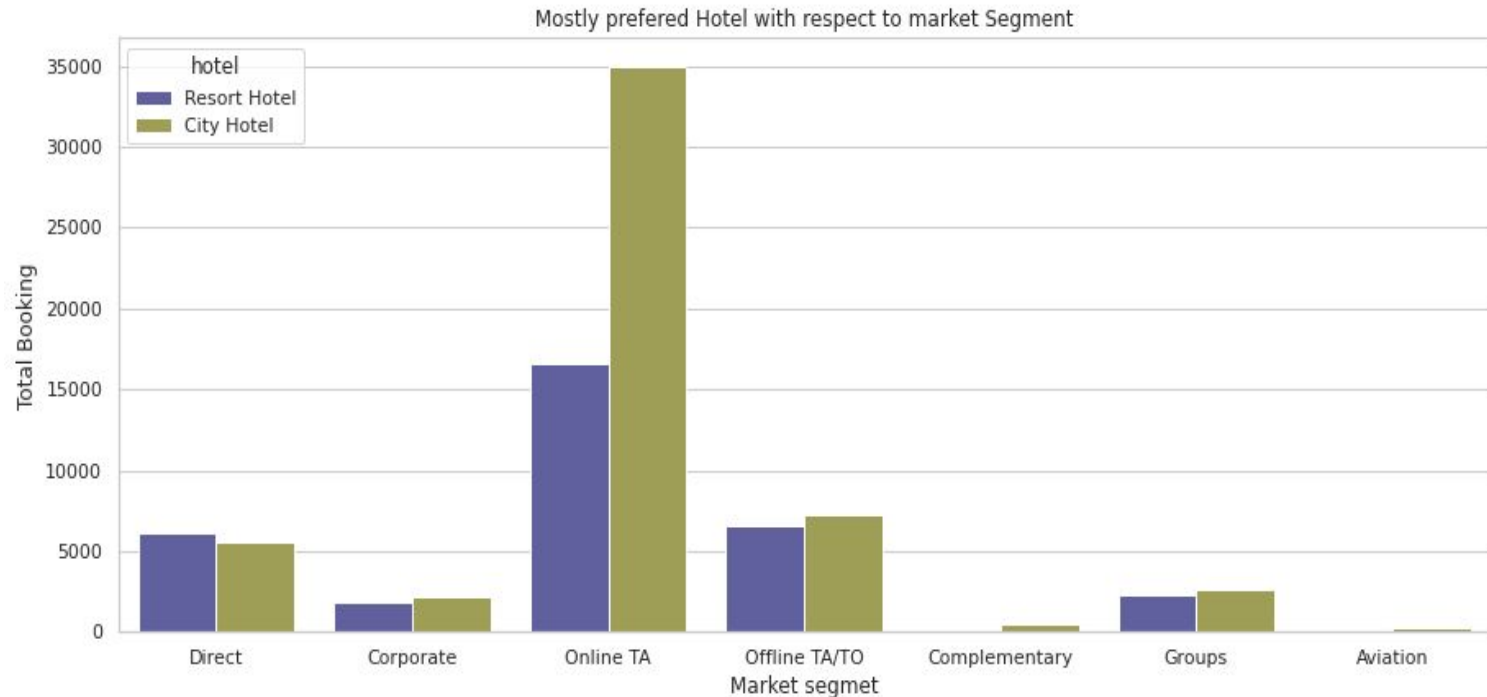
Did hotels asked for deposits for bookings.

Types of Deposit type

Is deposit type and cancellation have any pattern?

Booking Canceled or not by Hotel Type

# Which country's guest have high number of booking?

Western Europe has the most visitors, with Portugal, France, the United Kingdom, and Spain topping the list. These top 10 countries should be targeted for promotion.

TOP 10 countries from where most guest had arrived
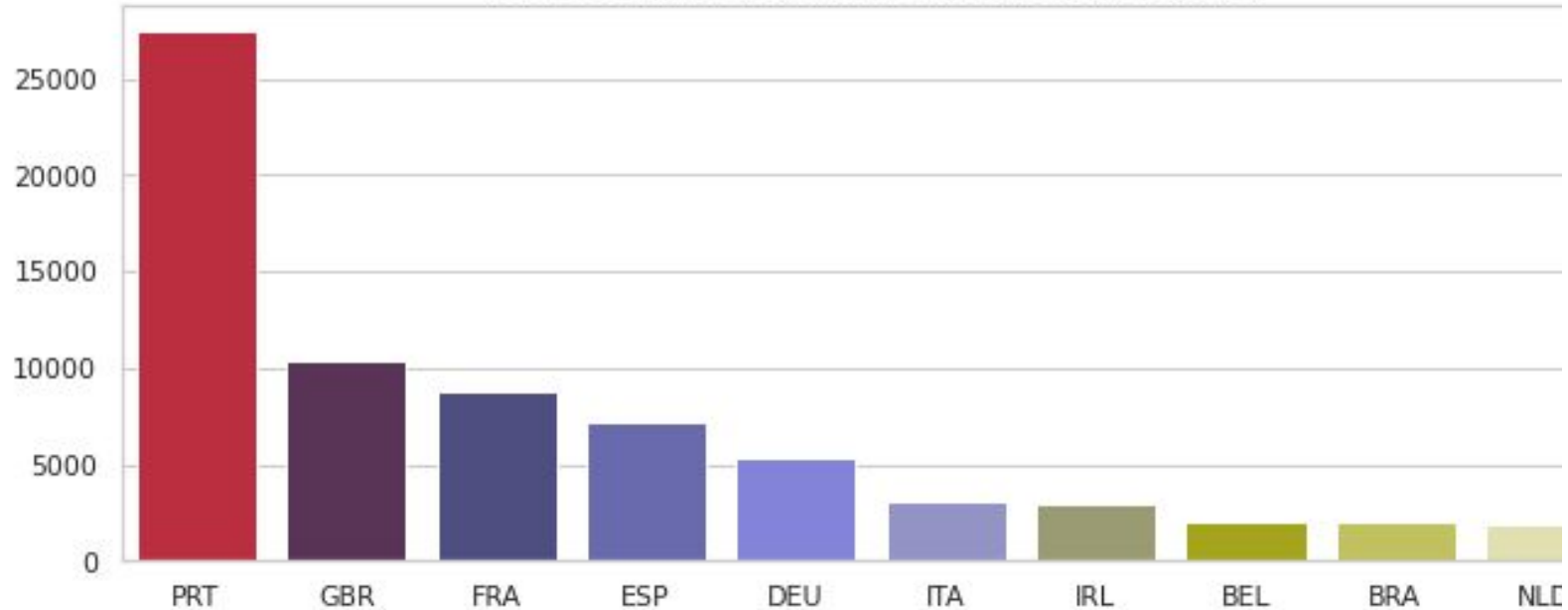
Which months have highest and lowest booking rates?

From below graph, it can be concluded, that both categories of hotels , the month of November & January have lowest Average Daily Rate (adr) and month of August has the highest Average Daily Rate (adr).
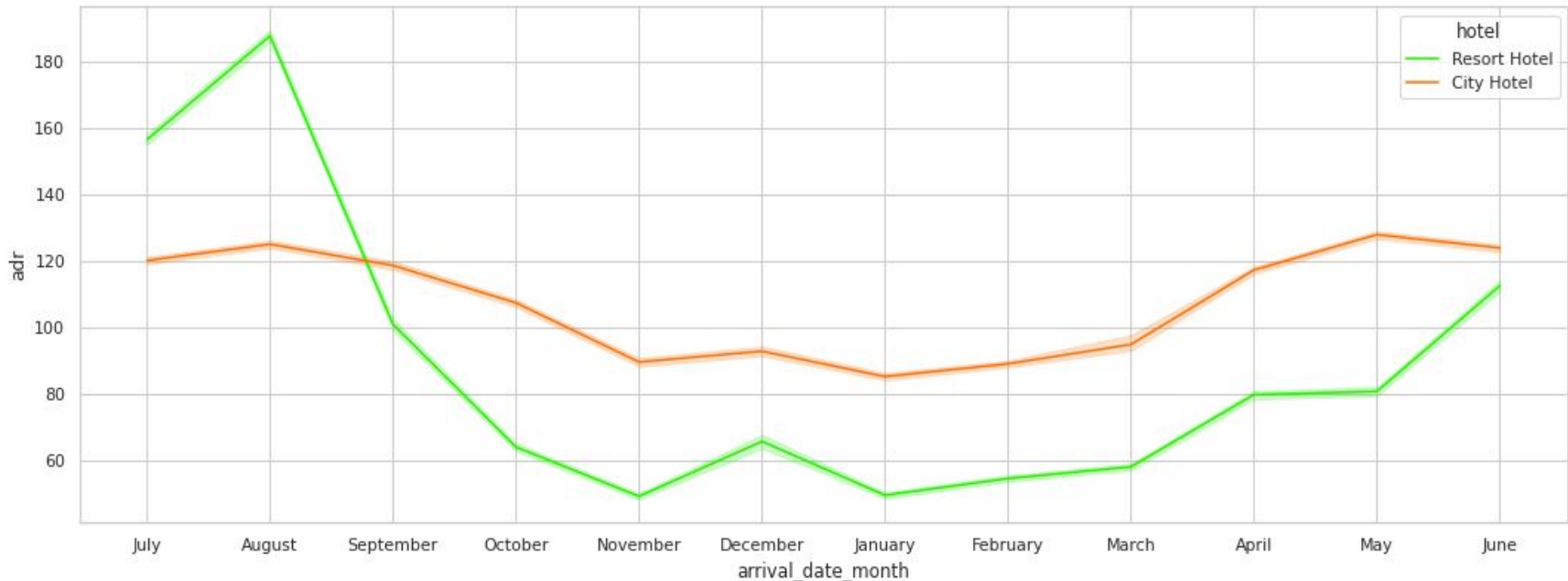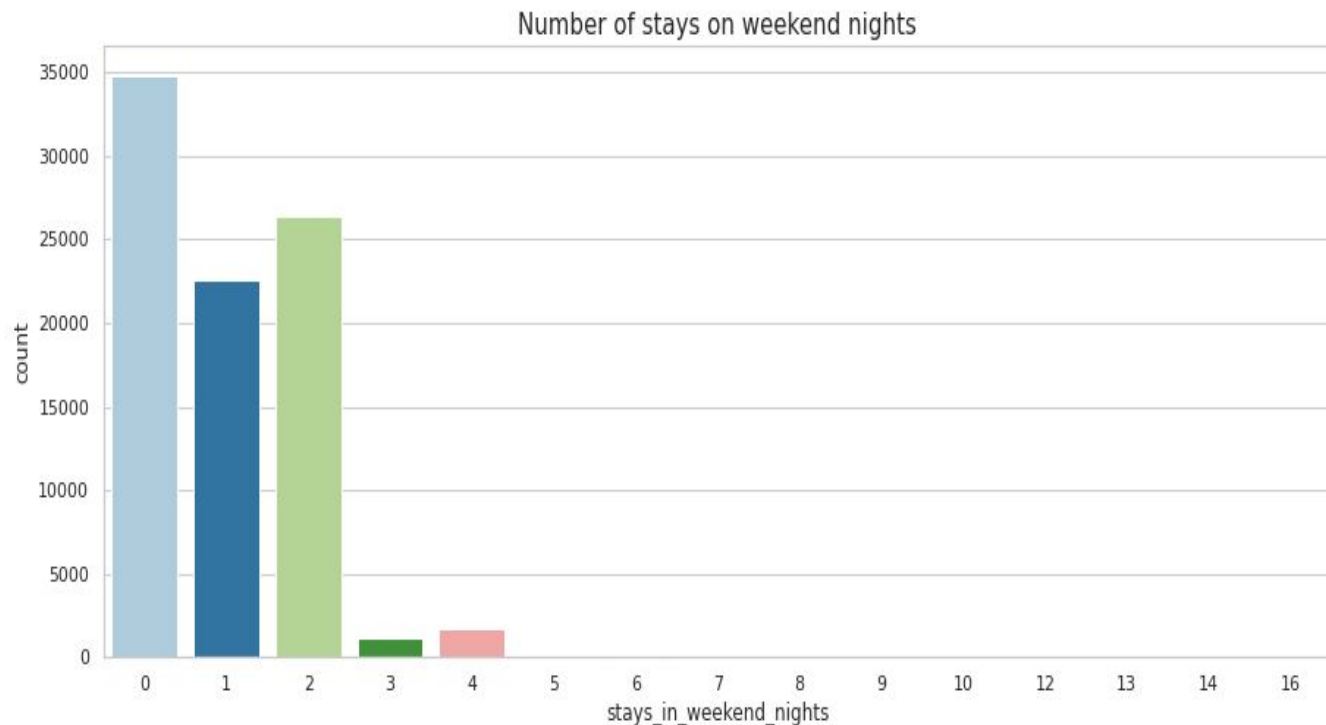
What is the average cost of a person staying in each type of hotel?

Resort hotel rates are significantly higher than City Hotel. The prices of city hotels do not fluctuate significantly.

Average Daily rate per person
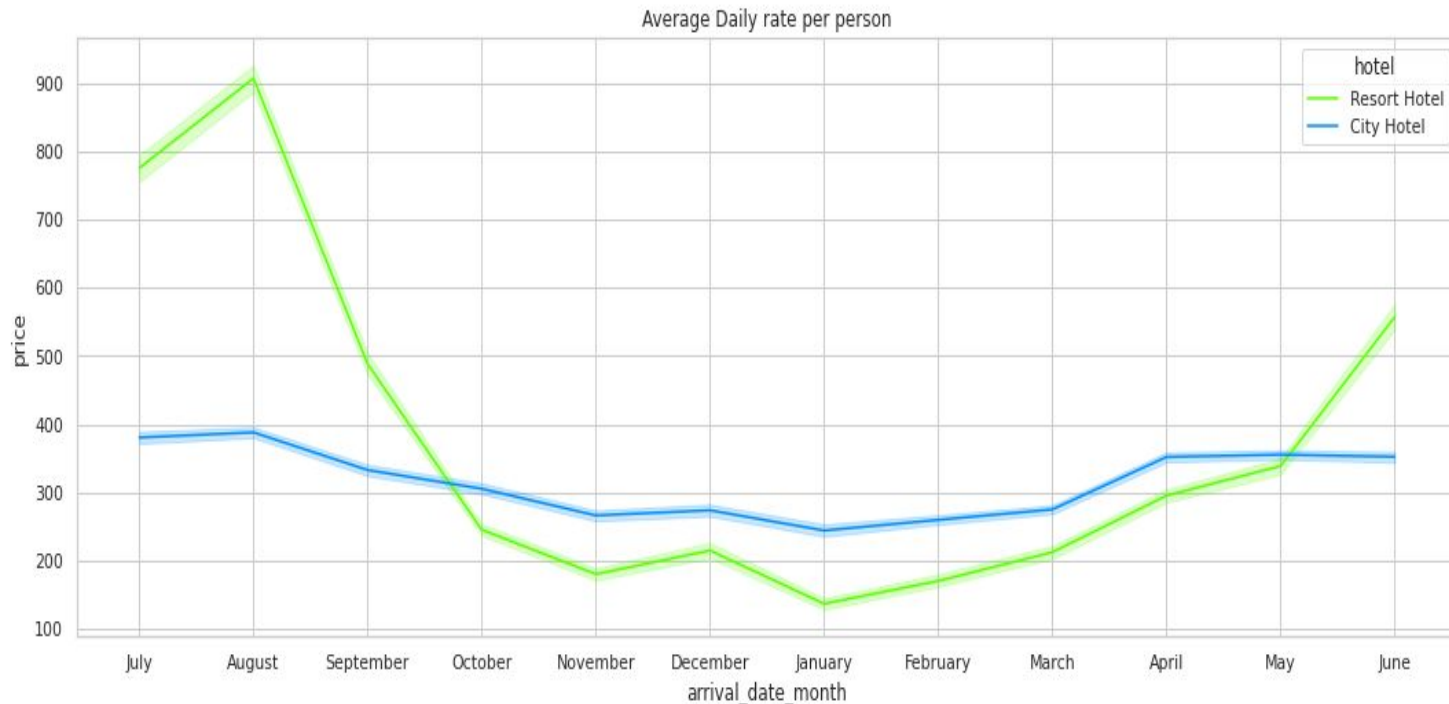
Whether guest prefers to stay over a weekend or weekday?
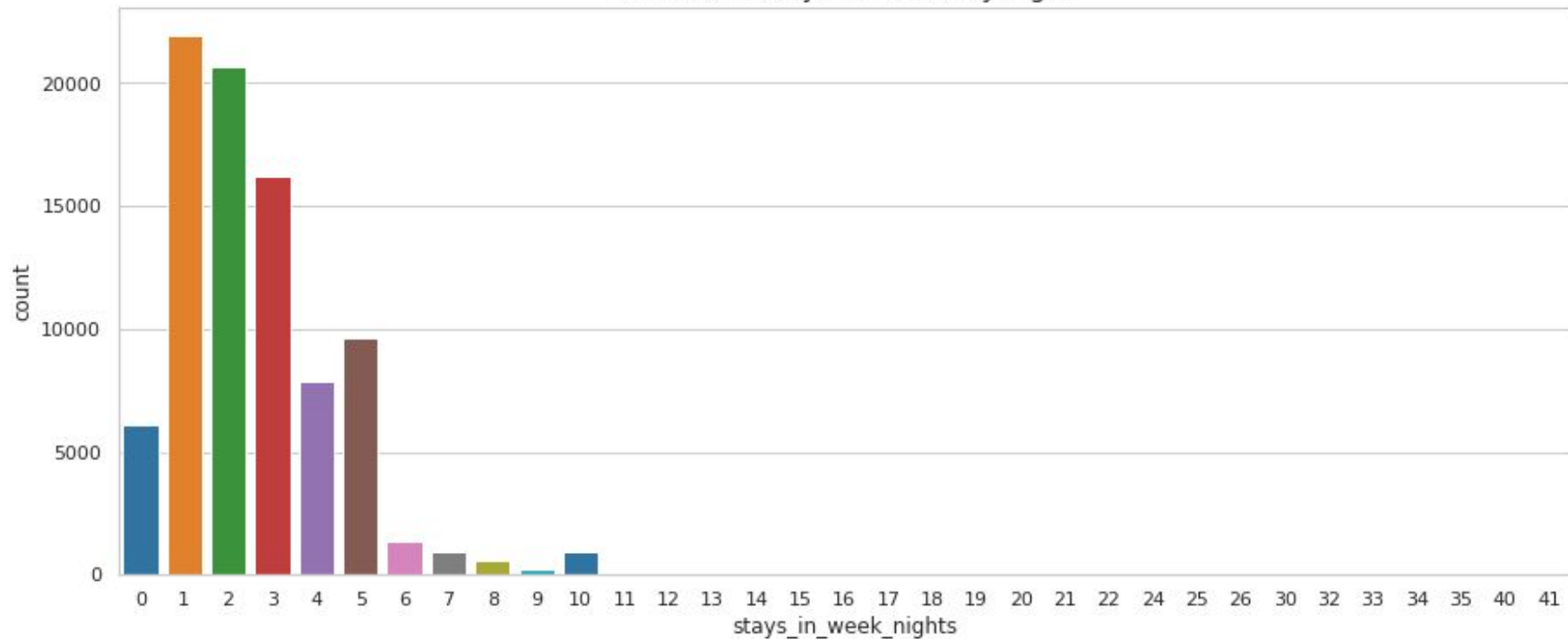
Stay on weekend nights -

Number of stays on weekend nights

From this graph, we got to know, that the majority of the stays appear to be on weekday nights as compared to weekend nights.
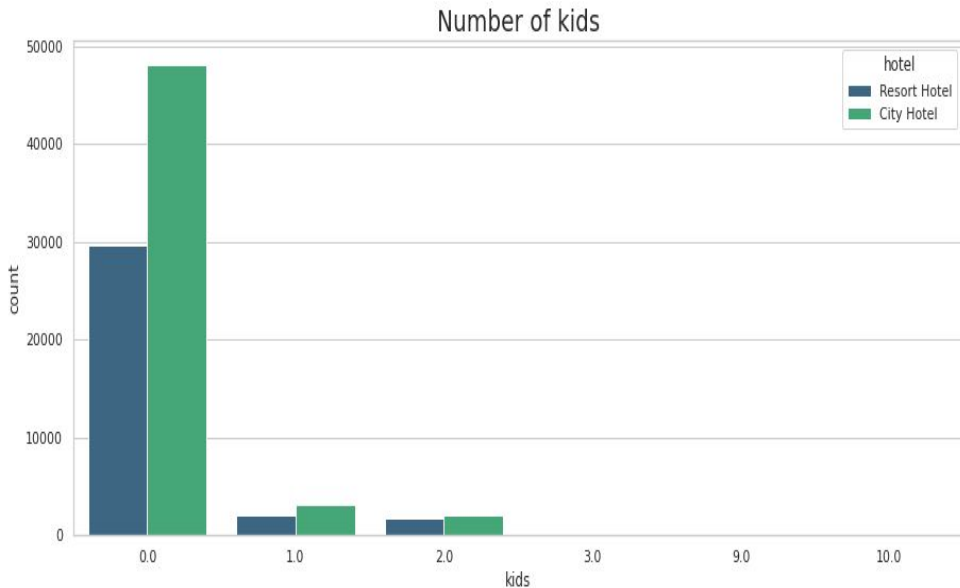
Stay on weekday nights -
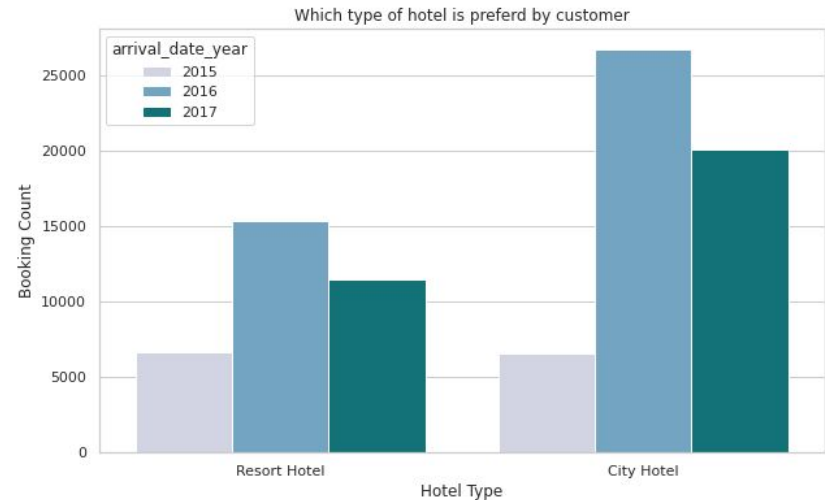

Number of stays on weekday night

Do people with children have a particular preference for the type of hotel?

As shown in figure Families with Kids have no particular preference for their Kids in any types of hotel.

Which type of hotel is most preferred by the customers?

From the graph we can conclude that guest highly prefered city hotel every year.
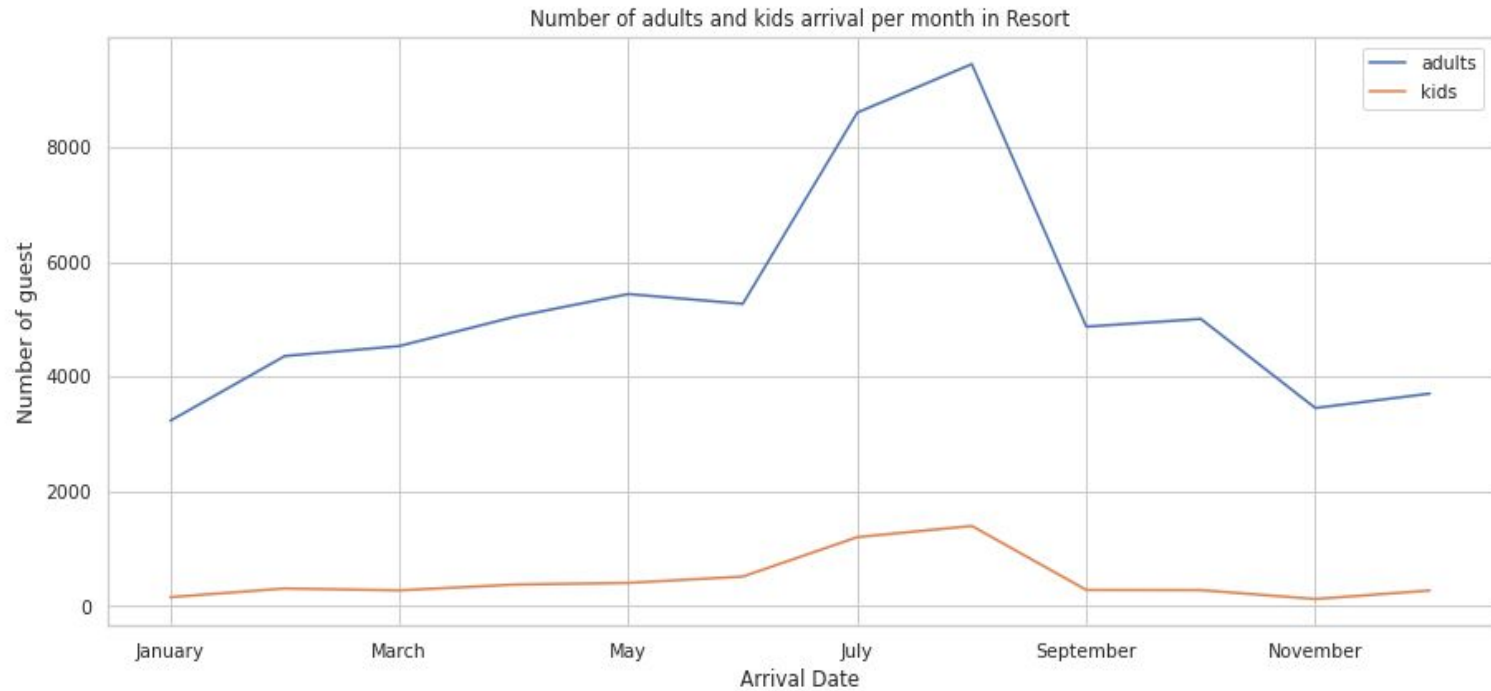
How many adults and kids guests arrived per month in each hotels?

As we can see in both city and resort hotel the arrival of adult guests per month is much higher than kids.

For City hotels

Number of adults and kids arrival per month in City hotel

For Resort hotels

Number of adults and kids arrival per month in Resort

Conclusion

- **Most bookings are done in the city hotel. Resort hotel tend to be on the costlier side and most people with family tend to go with city hotel.**

- **Analysis shows that August is the busiest month followed by May, maybe because of the vacations or summer season.**

- **Most of the visitors are from Western Europe; Portugal being the highest followed by Great Britain and France.**

- **Mostly people don't cancel their bookings but if compared with the resort hotel, city hotels have high cancellation rate because they also have high booking rates. Most Western Europe visitors didn't cancel their booking**