**EDA**

* 62.9% of the total bookings were cancelled.
* Months of July and August receive the highest number of bookings. January has the least number of bookings.
* Months of July and August have the highest number of booking cancellations. Thus the number of booking cancellations follows the volume of bookings for the months.
* On categorising the customers based on their country of origin, we see that Portugal makes up for 22.2 percent of bookings that are not cancelled(highest for a region/country). The next in the list is Great Britain and then france. On looking at the bookings that were cancelled, Portugal makes up for 59.1% of the total booking cancellations. Next, comes Great Britain and Spain. Therefore Portugal and Great Britain have the highest percentages of cancellation and non-cancellation.
* On comparing the deposit types for cancellations, we see that the customers that have not made a deposit are more likely to cancel a booking than deposits made with no refund or refundable policies.
* Comparing the customer types for cancellations,we see that transient customers have a higher probability of cancelling a booking when compared to bookings on a contract or group.
* Most of the transient customers make bookings with no deposit policy.Group customers are a minority.
* Average Daily Rate is a statistical unit that is often used in the lodging industry. The number represents the average rental income per paid occupied room in a given time period. City hotel has a higher average ADR than the resort hotel.
* Average ADR is lowest in the month of January. Hence in January when the hotel  receives the least number of bookings, the price is kept to the minimum. This month would ideally be more profitable for the customer.
* 14.4% of the repeated guests have cancelled their bookings. 37.7% of the first time guests have canceled their bookings. This shows that repeated guests have a lower probability of cancelling the bookings.
* Couples( no children) form the majority of the customers. As expected in accordance to the popularity of the hotels among people types, couples show a higher probability of cancellation.
* Longer the Lead Time, the higher is the probability of cancellation
* The correlation between lead time and cancellation is found to be 0.29
* The spearman correlation between deposit type and cancellation - both of which are categorical is found to be 0.48.