Capstone project: Exploratory data analysis

HOTEL BOOKING EDA

-LAVLEEN PRABHAKAR DEKATEY

Content:-

1. Data Visualization-

1) Maximum cancellation considering country, hotel type and customer type

2) Customer type which are repeated guest through different channels

3) Cancellation related to Lead time.

4) Cancellation percentage related to lead time

5) Deposit type which have cancellation

6) Is assigned room type for reserved room type a reason for cancellation

7) Customer with different number of booking changes doing reservation cancellation

8) Customer with different wait duration having reservation cancelled

9) Cancellation in different year.

2. Solution to business objective

Problem statement:-

There is approximately 37% cancellation during 3 years period whose reason can become reason for falling ratings of the hotels in the long run. so we need to find the reason for cancellation of booking.

Data summary:-

Dataset name: Hotel Bookings.csv

Dataset Shape: (119390, 32)

Dataset column: ['hotel', 'is_canceled', 'lead_time', 'arrival_date_year', 'arrival_date_month', 'arrival_date_week_number', 'arrival_date_day_of_month', 'stays_in_weekend_nights', 'stays_in_week_nights', 'adults', 'children', 'babies', 'meal', 'country', 'market_segment', 'distribution_channel', 'is_repeated_guest', 'previous_cancellations', 'previous_bookings_not_canceled', 'reserved_room_type', 'assigned_room_type', 'booking_changes', 'deposit_type', 'agent','company', 'days_in_waiting_list', 'customer_type', 'adr', 'required_car_parking_spaces', 'total_of_special_requests', 'reservation_status', 'reservation_status_date']

Cleaning dataset:-

As we are not considering agent and company named column and as it has more null values. So to avoid mess we are removing this two column.

Data Visualization:-

1)Maximum cancellation considering country, hotel type and customer type

- We used simple line plot chart as it will be able to show wide variety of country name on single chart and different customer type having different number of reservation cancellation.

maximum cancellation considering country, hotel type and customer type
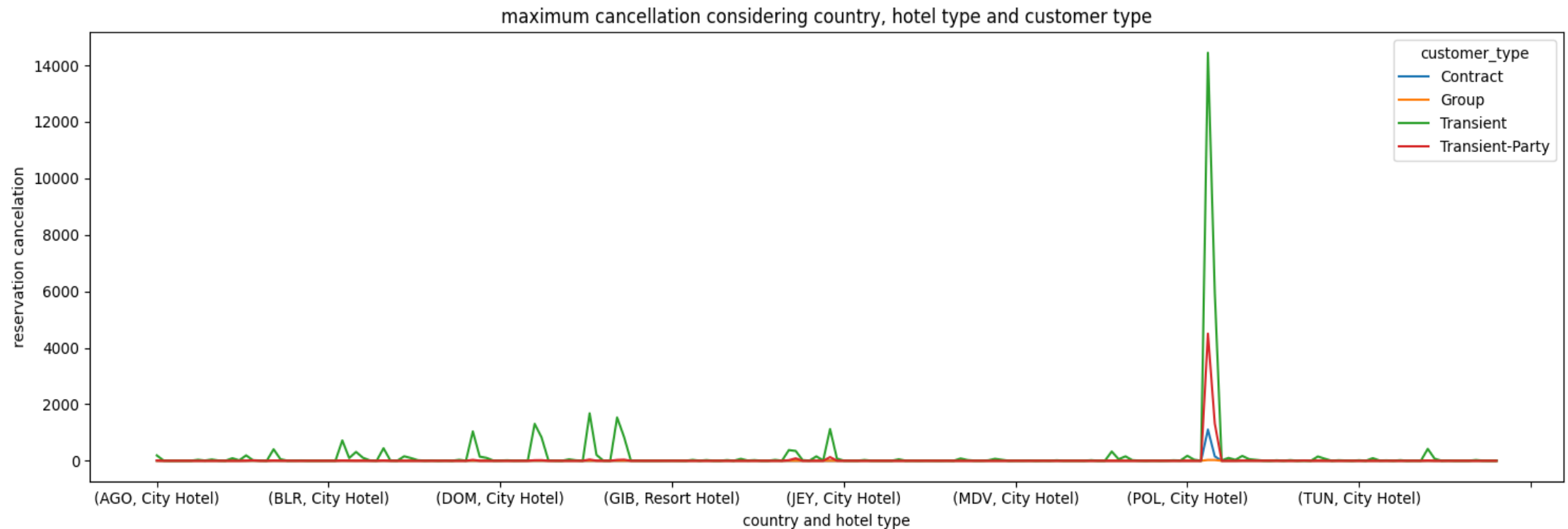
2) Customer type which are repeated guest through different channels

- We used a bar chart as it is easy and suitable to show the customer type and market segment which the customer belongs to that is suitable for showing categories.

customer type which are repeated guest through different channels
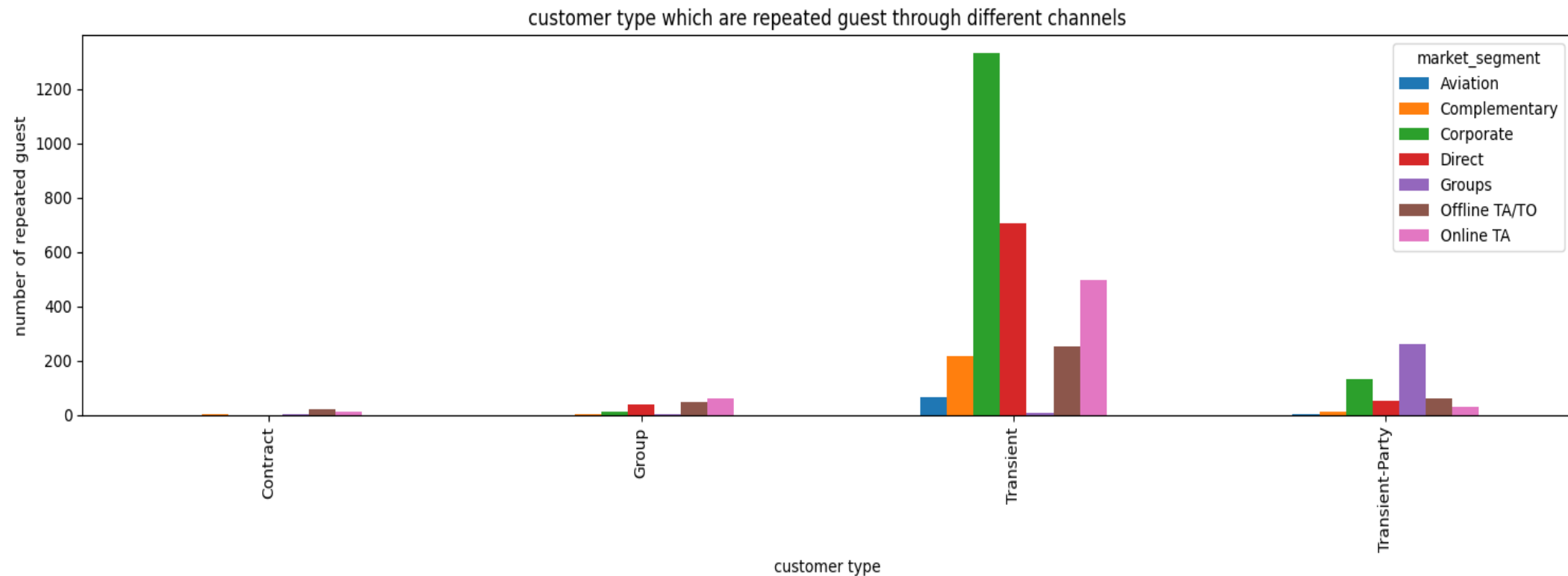
3) Cancellation related to Lead time

- We used correlation heat map to show how much lead time is related to customer who cancel the reservation.

- We used a bar chart to show different category of lead time according to days of lead time v/s number of reservation cancellation because it is category v/s number.


cancelation related to lead time
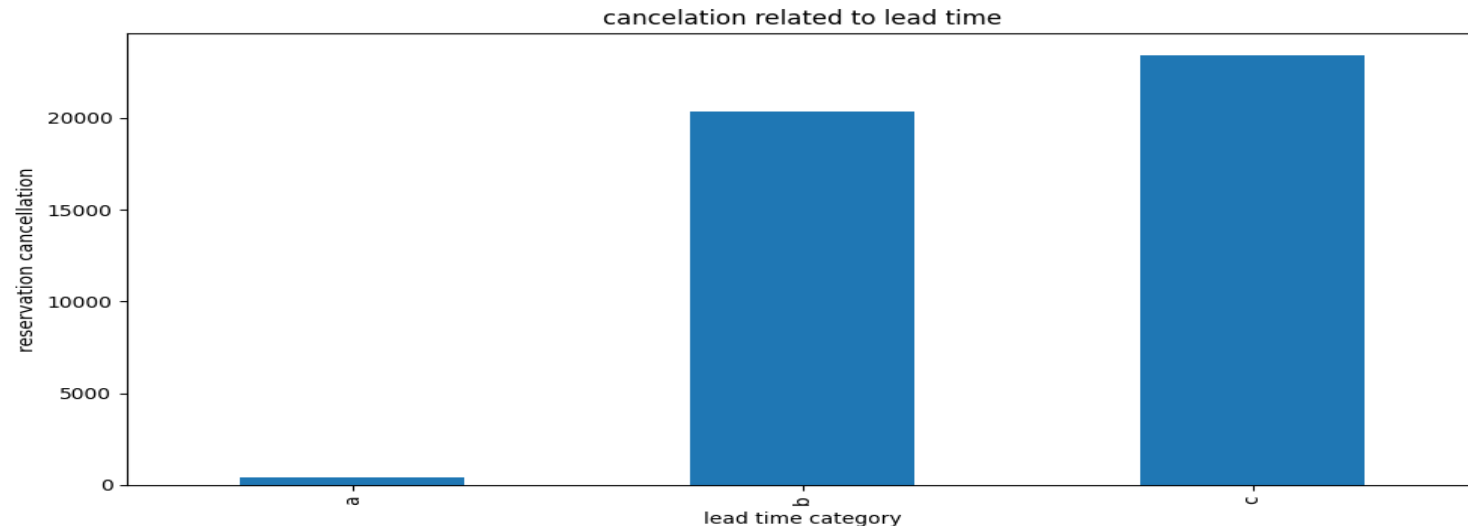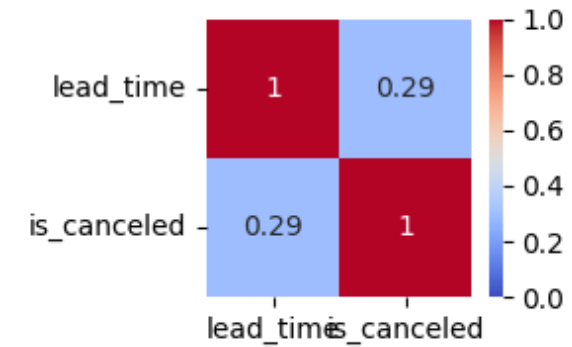
4) Cancellation percentage related to lead time

- We used line plot to see the behavior of reservation cancellation percentage value according to lead time category.
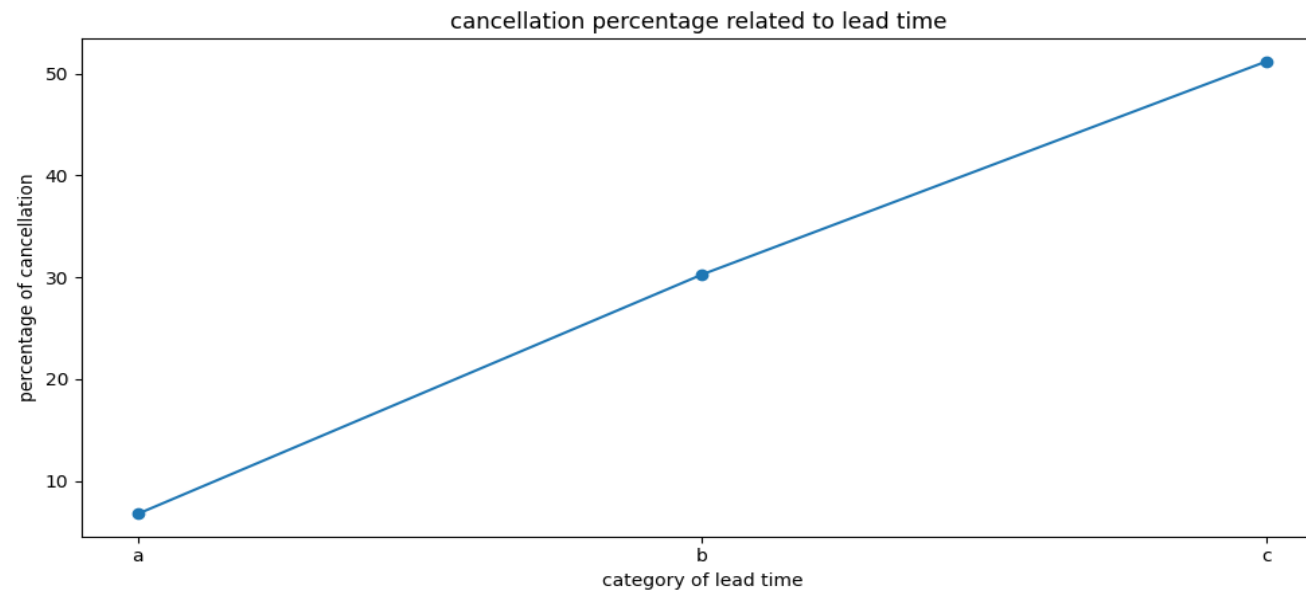
5) Deposit type which have cancellation

- We have used a line plot so that we can know how the graph falls down or goes up.

- We used bar chart to show the percentage of cancellation belonging to different deposit type.
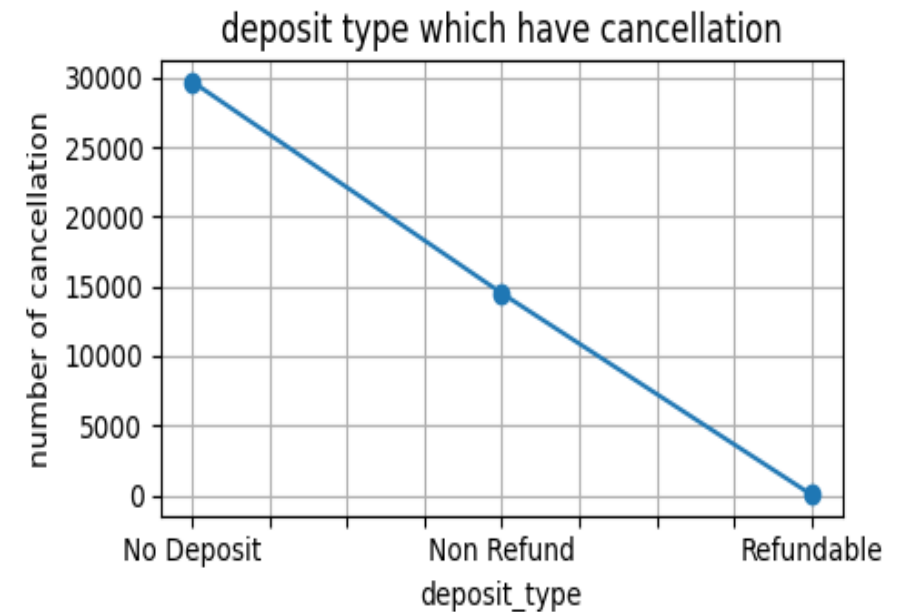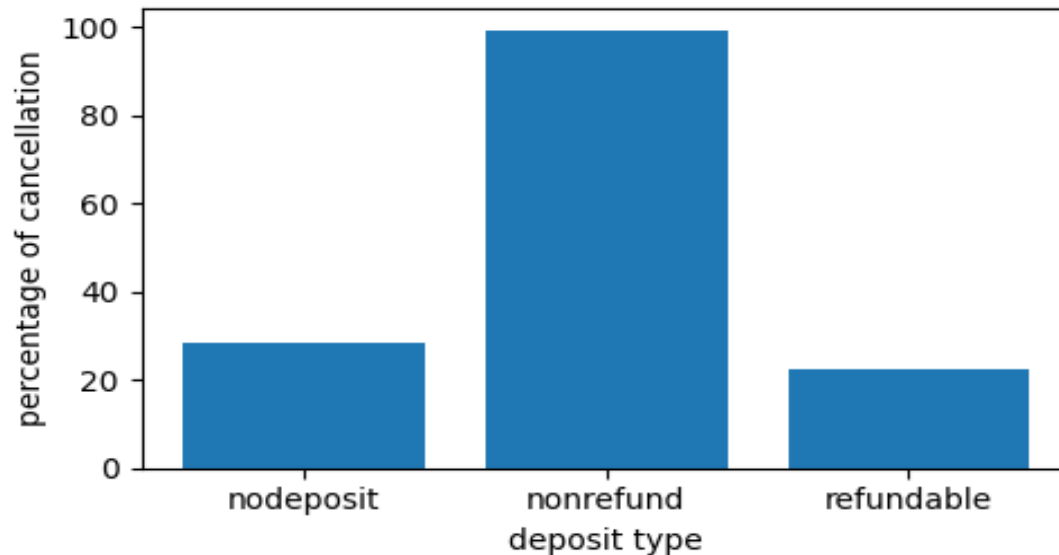
6) Is assigned room type for reserved room type a reason for cancellation

- Here we used bar chart to show the customer who have reserved room type but assigned with different room type and there corresponding reservation cancellation.
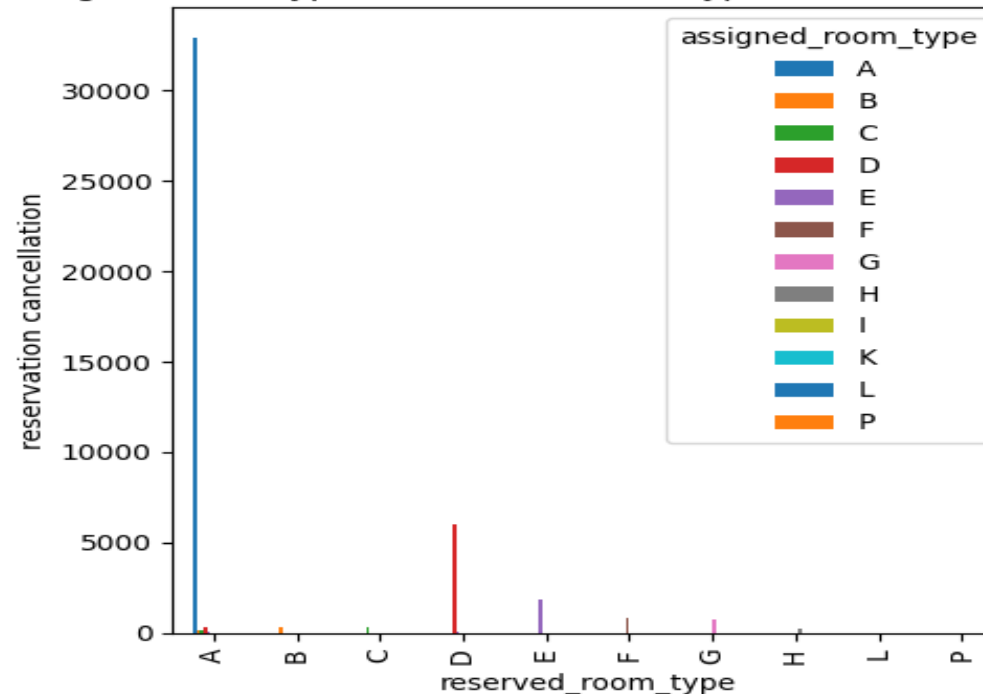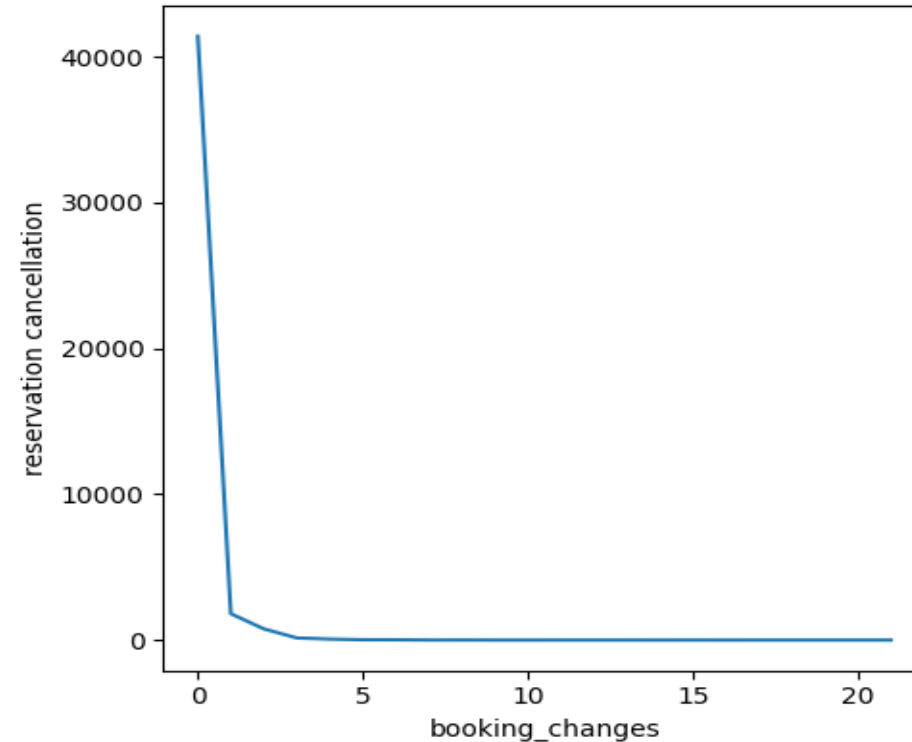
7) Customer with different number of booking changes doing reservation cancellation

- We used line chart to show how the number of booking changes play a role in cancellation of reservation.

customer with different number of booking changes doing reservation cancellation
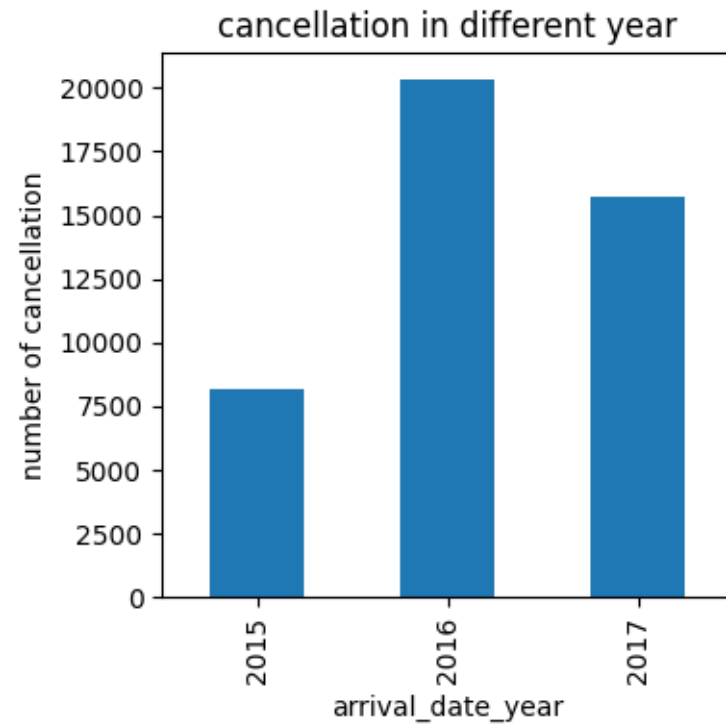
8) Customer with different wait duration having reservation cancelled

- We used line plot chart to show how days in waiting list is involved in reservation cancellation.

customer with different wait duration having reservation cancelled
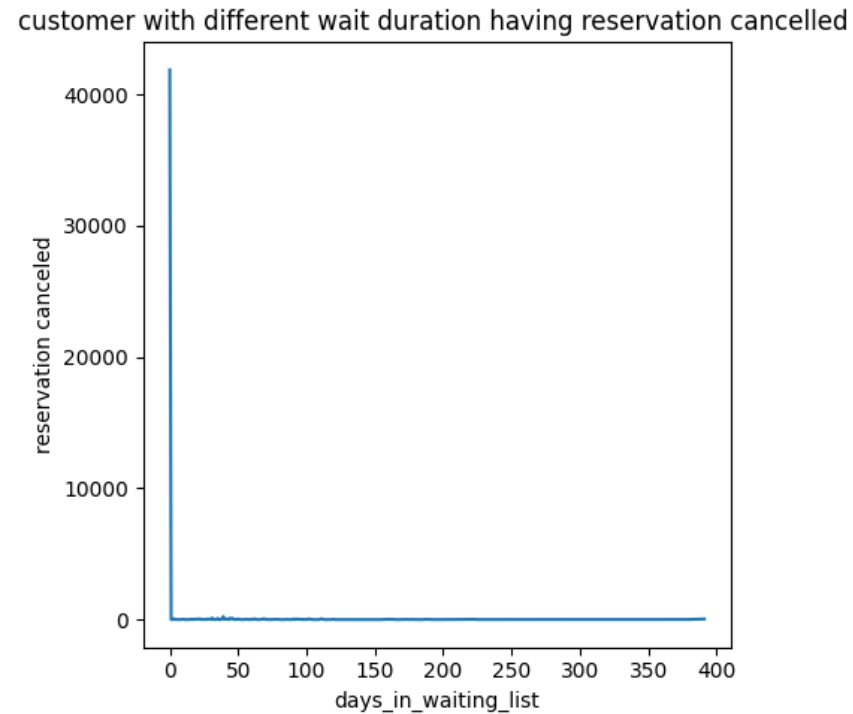
9) Cancellation in different year.

- We used the bar chart to find out which year has maximum number of cancellation.

Solution to business objective:-

1) Its too risky to sustain a business in portugal country. So you should structure the business in such a way that losses are bearable and services of the hotel should be upto the level of satisfaction of customer.

2) There are more customer of the transient customer type and they are from corporate market segment, so we should try to grab more customer from other market segment by creating offers like discounts for group of customers.

3) We should keep the maximum limit of lead time so as to reduce the possible number of reservation cancellation.

4) While setting up a business you should structure the system such that security deposit are high near to the total accommodation cost. So that the customer will think twice before cancelling reservation.

5) We should look after the condition and services provided to 'A' and 'D' type room. And make our rooms competent and satisfactory related to others hotel.

6) Customers making booking changes more than thrice are normally not cancelling the reservation.

7) We should be prepared to sustain losses as much as happened in year 2016.

Conclusion:-

We should shape our business in such a way that it should be able to sustain losses due to reservation cancellation in coming year such as it was in 2016, have competent and satisfactory rooms to attract more customer by increasing our reviews positively. Manage lead time and booking changes, so as to keep control of the number of cancellation. Losses should not be only seen from business side it should also be seen from customer side like the condition of the room and services we provide so as to bring loyal customers in the long run.