



Exploratory Data Analysis

Topic - Hotel Booking Analysis

By:-

Ajit kumar Toppo

Columns present in dataset

1. Hotel
2. is_canceled
3. Lead_time
4. arrival_date_year
5. Arrival_date_month
6. arrival_date_week_number
7. Arrival_date_day_of_month
8. stays_in_weekend_nights
9. stays_in_week_nights
10. Adults
11. Children
12. Babies
13. meal
14. Country
15. Market_segment
16. distribution_channel
17. is_repeated_guest
18. previous_cancellations
19. Previous_bookings_not_canceled
20. Reserved_room_type
21. Assigned_room_type
22. booking_changes
23. Deposit_type
24. Agent
25. Company
26. Days_in_waiting_list
27. customer_type
28. adr
29. required_car_parking_spaces
30. total_of_special_requests
31. Reservation_status
32. reservation_status_date



dataset (119390 , 32)



#	Column	Non-Null Count	Dtype
0	hotel	119390 non-null	object
1	is_canceled	119390 non-null	int64
2	lead_time	119390 non-null	int64
3	arrival_date_year	119390 non-null	int64
4	arrival_date_month	119390 non-null	object
5	arrival_date_week_number	119390 non-null	int64
6	arrival_date_day_of_month	119390 non-null	int64
7	stays_in_weekend_nights	119390 non-null	int64
8	stays_in_week_nights	119390 non-null	int64
9	adults	119390 non-null	int64
10	children	119386 non-null	float64
11	babies	119390 non-null	int64
12	meal	119390 non-null	object
13	country	118902 non-null	object
14	market_segment	119390 non-null	object
15	distribution_channel	119390 non-null	object
16	is_repeated_guest	119390 non-null	int64
17	previous_cancellations	119390 non-null	int64
18	previous_bookings_not_canceled	119390 non-null	int64
19	reserved_room_type	119390 non-null	object
20	assigned_room_type	119390 non-null	object
21	booking_changes	119390 non-null	int64
22	deposit_type	119390 non-null	object
23	agent	103050 non-null	float64
24	company	6797 non-null	float64
25	days_in_waiting_list	119390 non-null	int64
26	customer_type	119390 non-null	object
27	adr	119390 non-null	float64
28	required_car_parking_spaces	119390 non-null	int64
29	total_of_special_requests	119390 non-null	int64
30	reservation_status	119390 non-null	object
31	reservation_status_date	119390 non-null	object

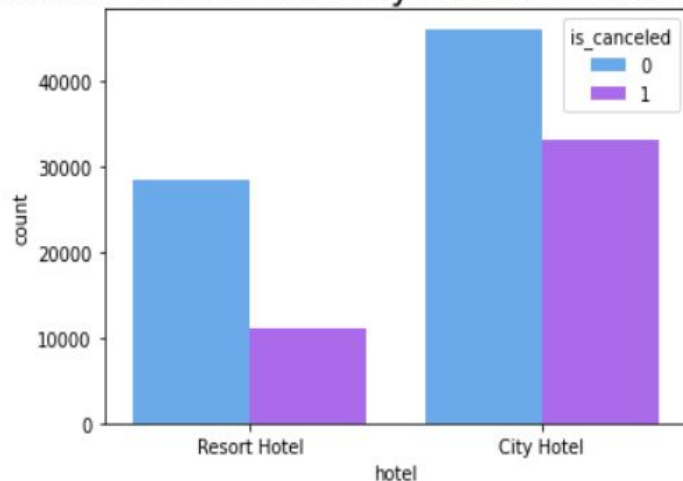
Q1. How much of the bookings done get cancelled?



118727- total number of records available after data cleaning

0	74590
1	44137

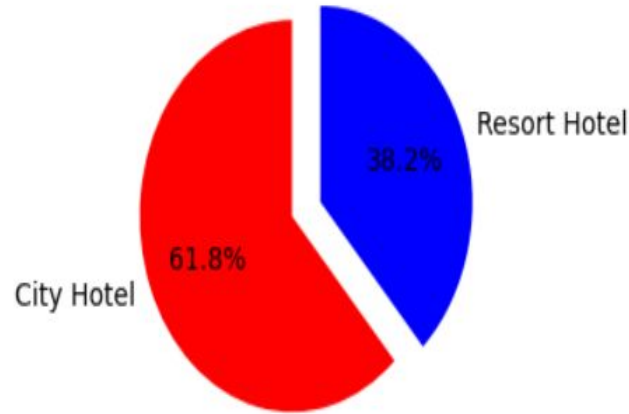
Cancelation rates in City hotel and Resort hotel



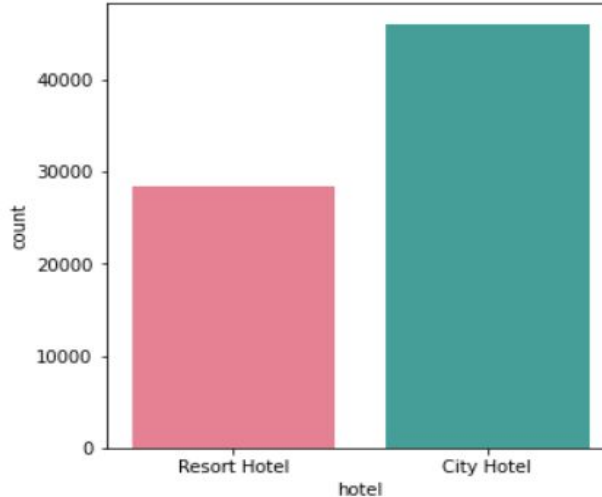
Q2. which type of hotel has higher number of bookings City Hotel or Resort Hotel ?



Percentage of booking in City Hotel vs Resort Hotel



Number of bookings in City Hotel vs Resort Hotel



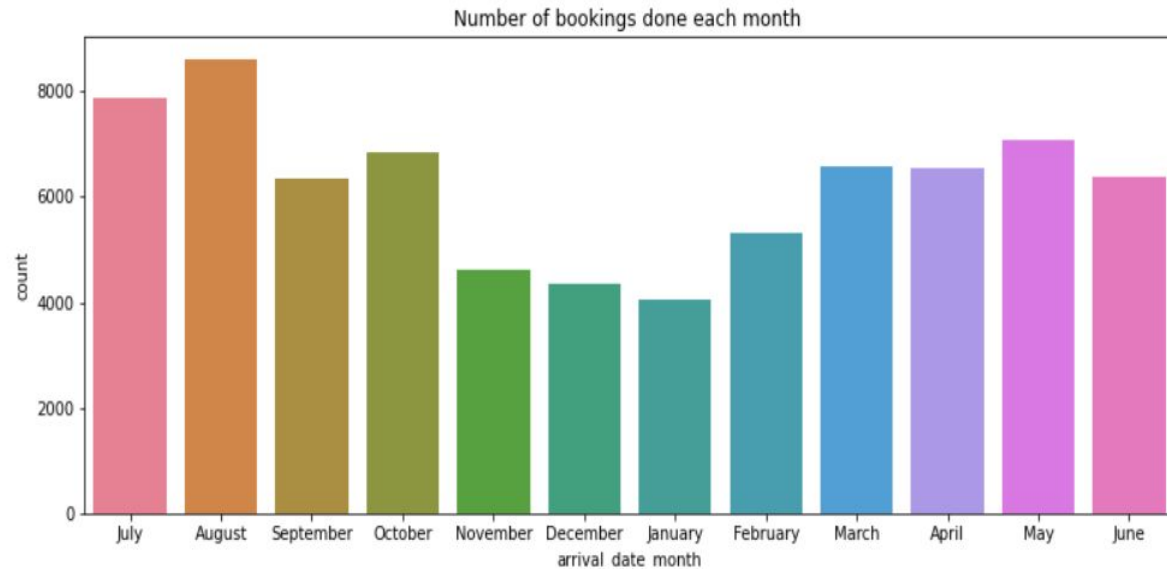
City Hotel - 46082

Resort Hotel - 28508

Q3. In which month maximum number of bookings were done?

Number of bookings done each month-

August	8604
July	7880
May	7091
October	6854
March	6566
April	6528
June	6384
September	6360
February	5304
November	4611
December	4347
January	4061



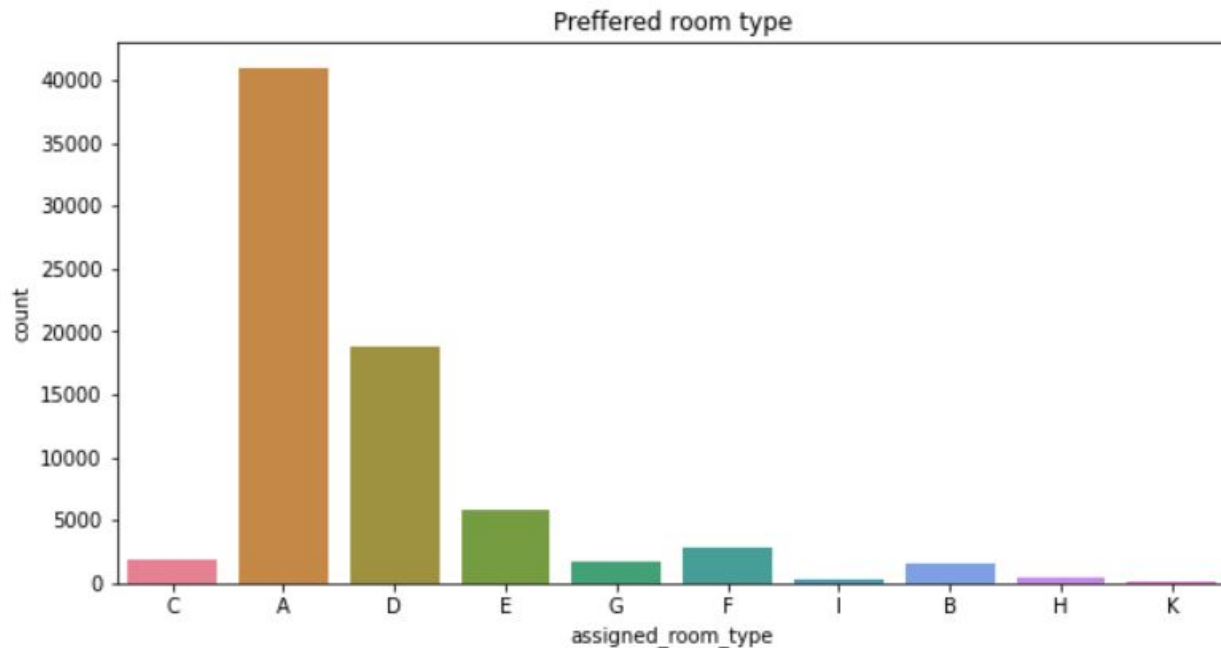
Q4. From which country maximum number of bookings are done?

- 1.Portugal-20977
- 2.United Kingdom-9668
- 3.France-8468
- 4.Spain-6383
- 5.Germany-6067



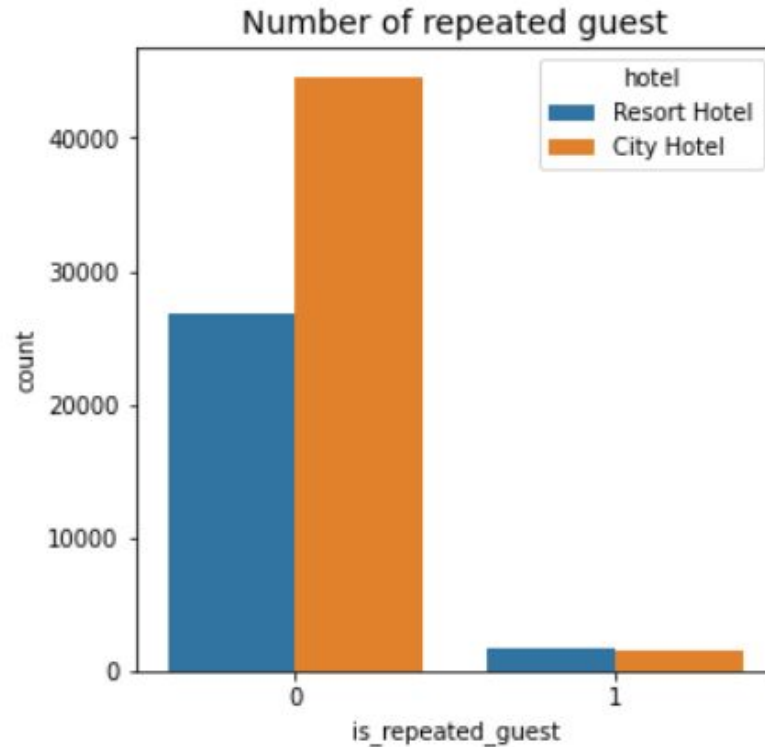
Q5. Which is the most preferred type of room for booking?

A	40930
D	18796
E	5774
F	2806
C	1905
G	1755
B	1643
H	457
I	348
K	176



Q6. How many guest are repeated customers?

0	71389
1	3201



Q7. Which type of meal package is most preferred by the customers?

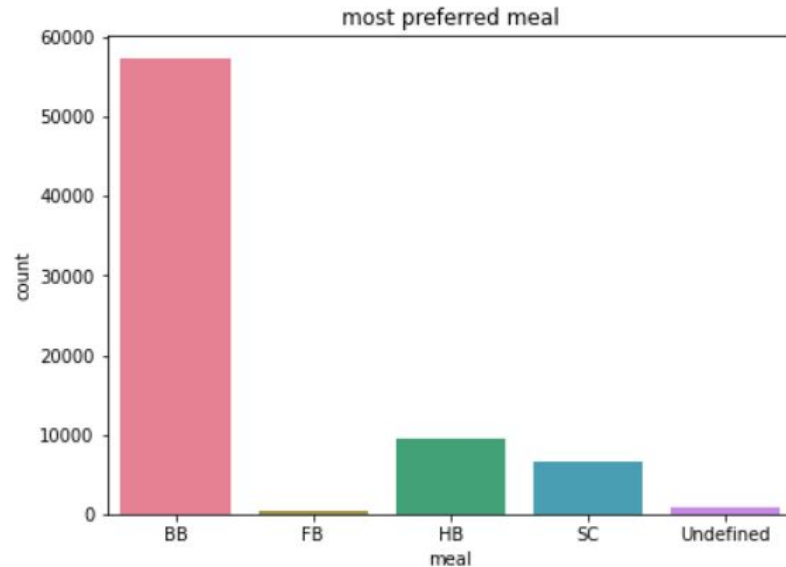
BB	57341
HB	9449
SC	6601
Undefined	879
FB	320

Undefined/SC – no meal package;

BB – Bed & Breakfast;

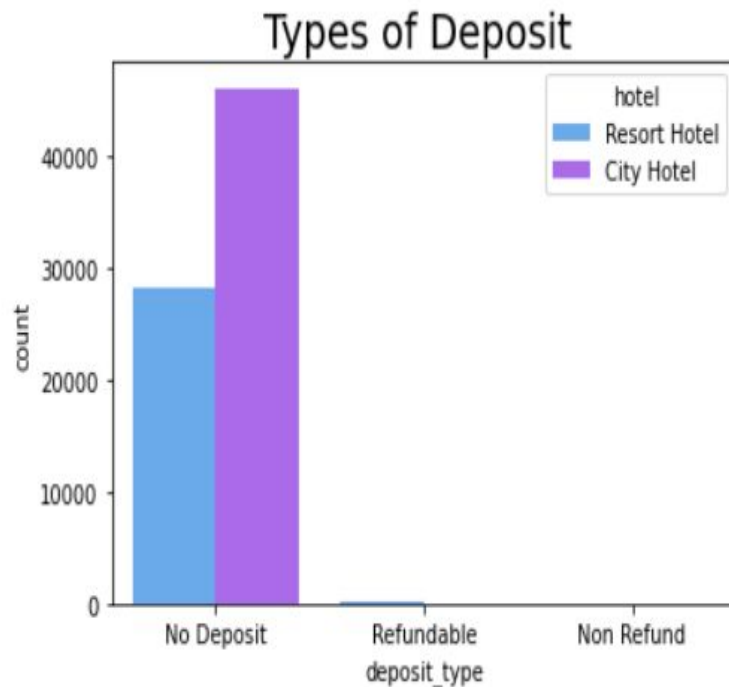
HB – Half board (breakfast and one other meal
– usually dinner);

FB – Full board (breakfast, lunch and dinner)



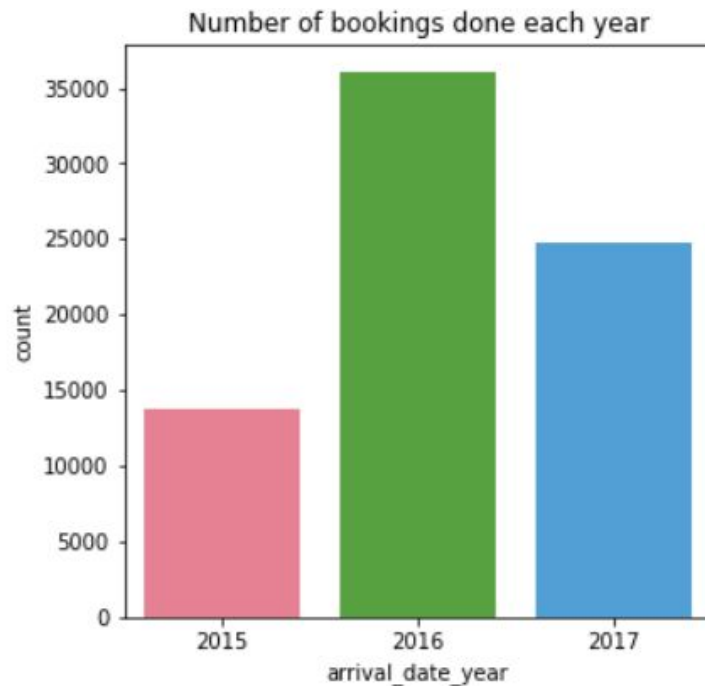
Q8. which type of deposit customers prefer?

No Deposit	74371
Refundable	126
Non Refund	93



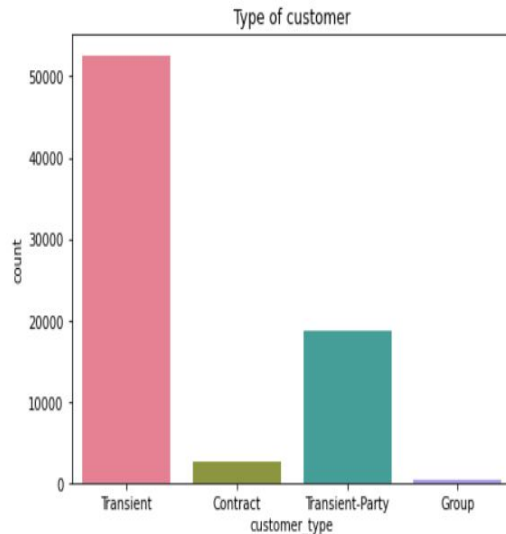
Q9. In which year most of the bookings were done?

2016	36051
2017	24814
2015	13725



Q10. Which type of customers are most likely to visit?

Transient	52597
Transient-Party	18673
Contract	2810
Group	510



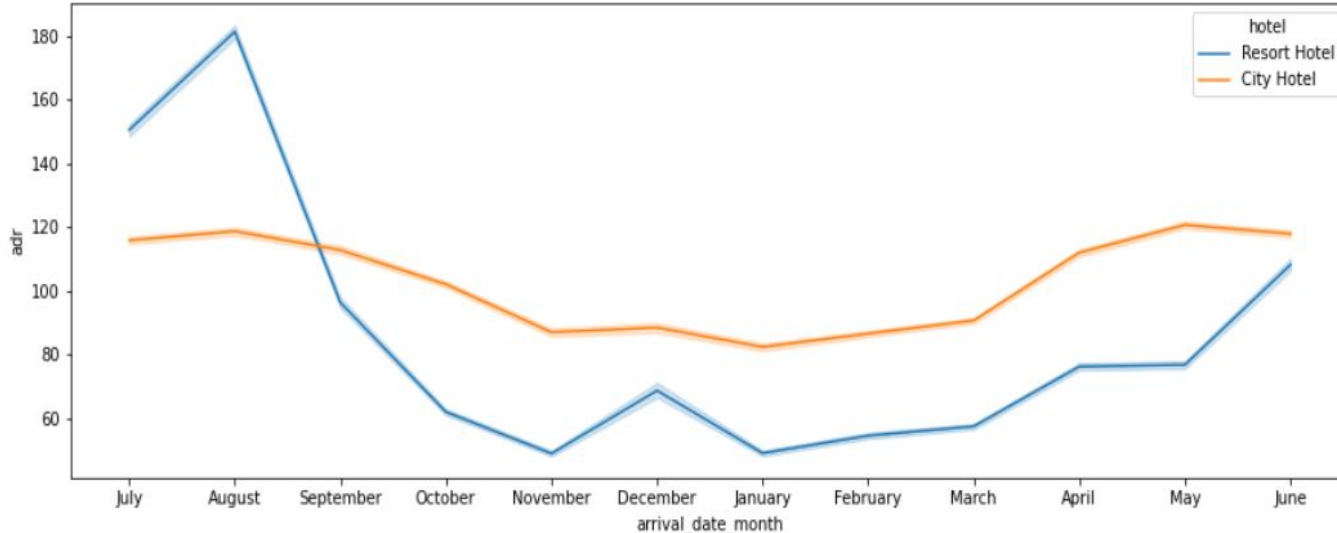
Contract - when the booking has an allotment or other type of contract associated to it;

Group – when the booking is associated to a group;

Transient – when the booking is not part of a group or contract, and is not associated to other transient booking;

Transient-party – when the booking is transient, but is associated to at least other transient booking

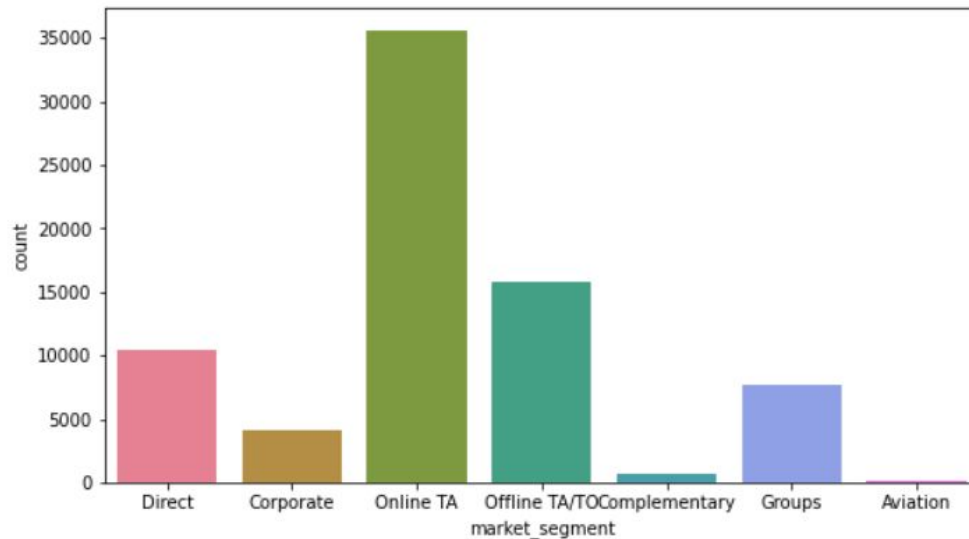
Q11. How the ADR varies over the year for both the hotels?



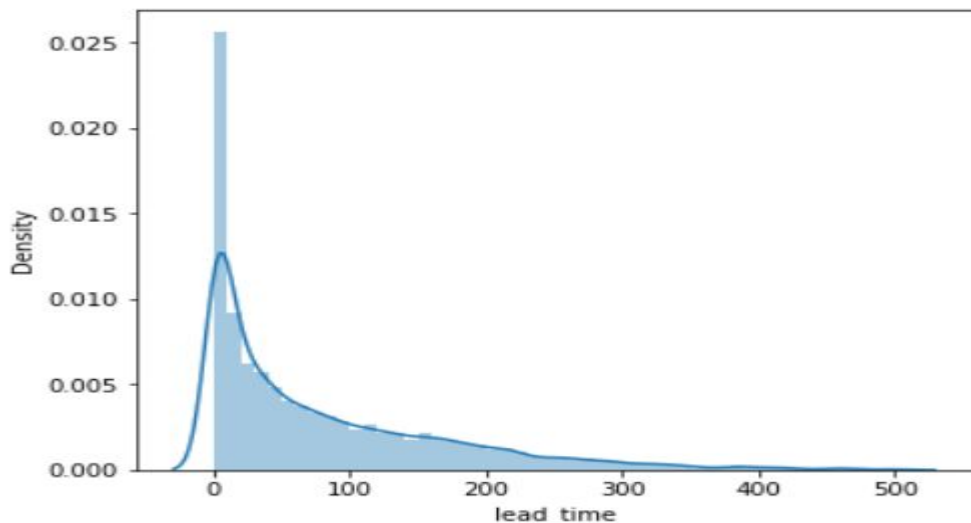
ADR(**average daily rate**) is used to measure the average revenue that a hotel receives for each occupied guest room per day.

Q12. Which is the most preferred method for the booking?

Online TA	35599
Offline TA/TO	15854
Direct	10504
Groups	7692
Corporate	4121
Complementary	637
Aviation	183



Q13. Before how many days most of the bookings were done?



We see that there is a positive skewness in the lead time. Most of booking done are planned for a close time but still we can see few percentage of people who do planning a long before their arrival.

