AlmaBetter

Capstone Project

HOTEL BOOKING ANALYSIS

By – Lavanya Shinde

CONTENTS

i. Agenda and Problem Statement
ii. Work Flow
iii. Data Review
iv. Types of Data in Dataset
v. Insights to find out
vi. Exploratory Data Analysis
vii. Insights to be found

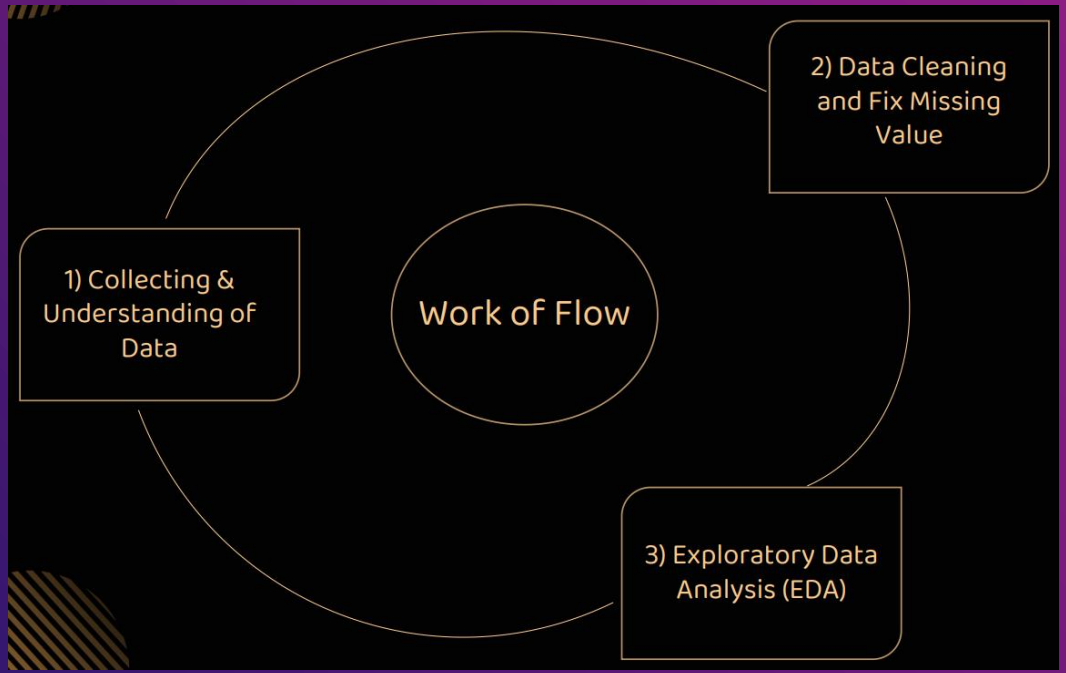
01

AGENDA AND PROBLEM STATEMENT

- We Have a Dataset of Hotel Booking, and we discuss about this dataset as well as cover some query with analyzing.
- In Analyzing the data, we work with Hotel Booking Dataset in this dataset containing include the information such as hotel resort vs city , booking cancellation , types of customers , meal of customers , agent information & many more.
- Hotel Industry is Highly Traffic Industry with Lakh of People with Lakh of Data and with collecting and maintain the information for analyzing , give big benefit to Hotel Industry.
- The main target behind this project is to explore and then analyze the data for discover important factors and find out the insights to solve out problem in Hotel Management , So they easily observe advantages & weakness and perform campaigns to boost their business as well as Performance.

02 WORKFLOW

- Here, is a simple workflow used in our project.



03 DATA REVIEW

In the Given Hotel Booking Dataset there are 119390 number of rows and 32 number of columns. So, let's understand every columns which is contain in data.

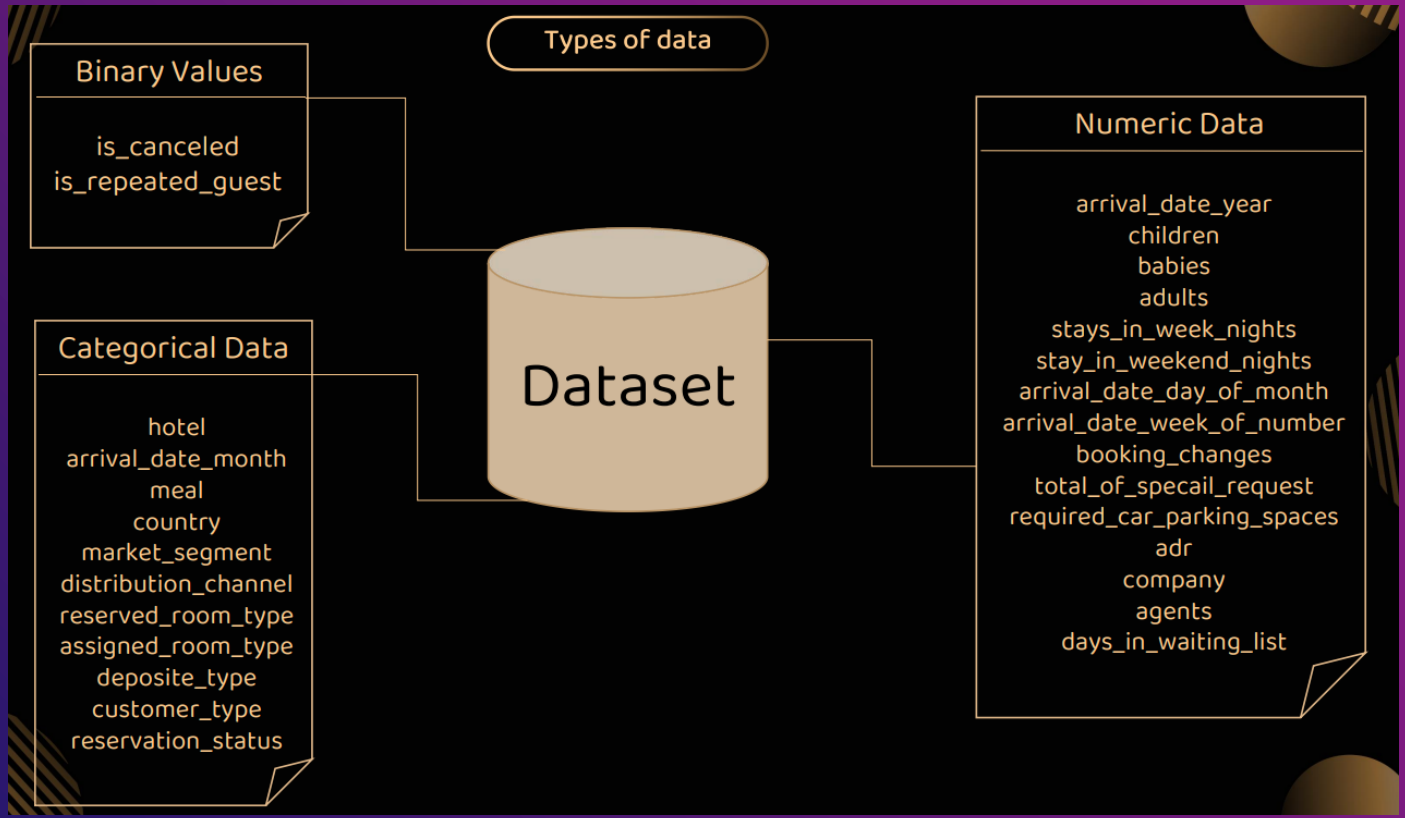
1. hotel :- In this column contain two categorical data which is Resort Hotel & City Hotel .
2. is_cancelled :- In this column value show the cancellation type. If the booking was cancelled Value indicate 1 & 0 indicates not cancelled.
3. lead_time :- Data show the time between reservation and actual arrival .
4. arrival_date_year :- Data show Year of arrival date .
5. arrival_date_month :- Data show Month of arrival date .
6. arrival_date_week_number :- Data of Week number of year for arrival date .
7. arrival_date_day_of_month :- Day of arrival date .
8. stays_in_weekend_nights :- Total Number of weekend nights .
9. stays_in_week_nights :- Total Number of week nights .
10. adults : - Total Number of adults .

11. children :- Total Number of children in Hotel .
12. babies :- Total Number of babies in Hotel .
13. meal :- Type of meal booked by a Customers .
14. country :- Country of origin of Customers .
15. market_segment :-Designation of market segment . (TA/TO) .
16. distribution_channel :- Booking distribution channel.(T/A/TO) .
17. is_repeated_guest :- is a repeated guest or not. (1) means yes & (0) means not.
18. previous_cancellations :- Number of previous bookings that were cancelled by the customer prior to the current booking .
19. previous_bookings_not_canceled :- Number of previous bookings not cancelled by the customer prior to the current booking
20. reserved_room_type :- Code of room type reserved.
21. assigned_room_type : Code for the type of room assigned to the booking.
22. booking_changes : Number of changes made to the booking from the moment the booking

23. deposit_type :- Values of No Deposit, Non Refund , Refundable.
24. agent :- Values of ID of the travel agent that made the booking.
25. company :- Contain ID of the company or entity that made the booking .
26. days_in_waiting_list :- Number of days the booking was in the waiting list before it was confirmed to the customer.
27. customer_type :- Data Contain types of customer & ontract,Group,transient,Transient party.
28. adr :- Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights.
29. required_car_parking_spaces :- Data of Number of car parking spaces required by the customer.
30. total_of_special_requests :- Data of Number of special requests made by the customer
31. reservation_status :- Reservation status.
32. reservation_status_data :- Date of reservation status.

04

TYPES OF DATA IN DATASET


Types of data
Binary Values
is_canceled
is_repeated_guest
Categorical Data
hotel
arrival_date_month
meal
country
market_segment
distribution_channel
reserved_room_type
assigned_room_type
deposite_type
customer_type
reservation_status
Dataset
Numeric Data
arrival_date_year
children
babies
adults
stays_in_week_nights
stay_in_weekend_nights
arrival_date_day_of_month
arrival_date_week_of_number
booking_changes
total_of_specail_request
required_car_parking_spaces
adr
company
agents
days_in_waiting_list

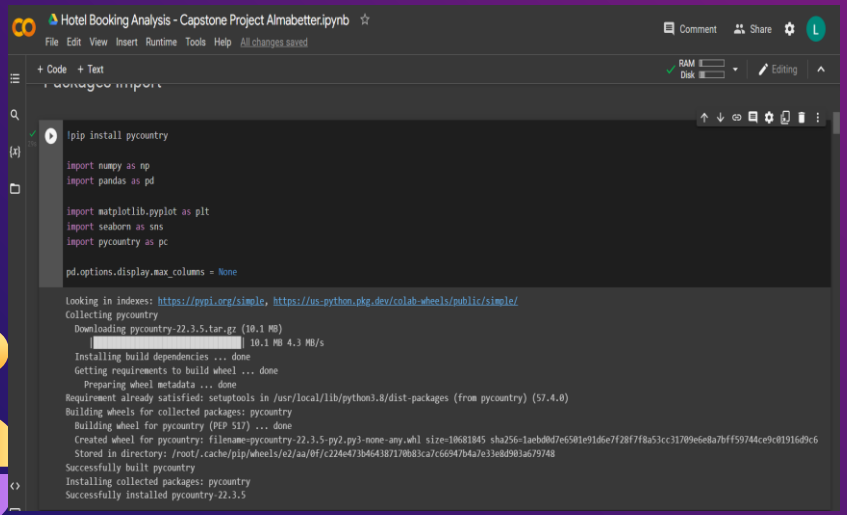
05 INSIGHTS TO BE FOUND

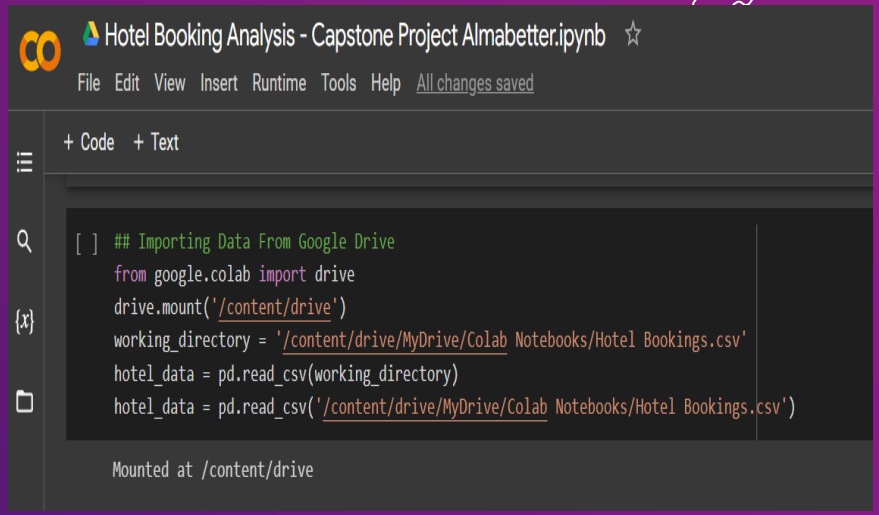
1. How Many Booking Were Cancelled?
2. What is the booking ratio between Resort Hotel and City Hotel?
3. What is the percentage of booking for each year?
4. Which is the busiest month for hotel?
5. From which country most guest come?
6. How Long People Stay in the hotel?
7. Which was the most booked accommodation type (Single, Couple, Family)?
8. How many guests repeated in hotel ?
9. Types of Customers ?
10. Meal Plan Of Customers.
11. Which agent makes highest no. of bookings ?
12. Which is the most preferred room type by the customers ?
13. Which Hotel type has the highest ADR ?

06

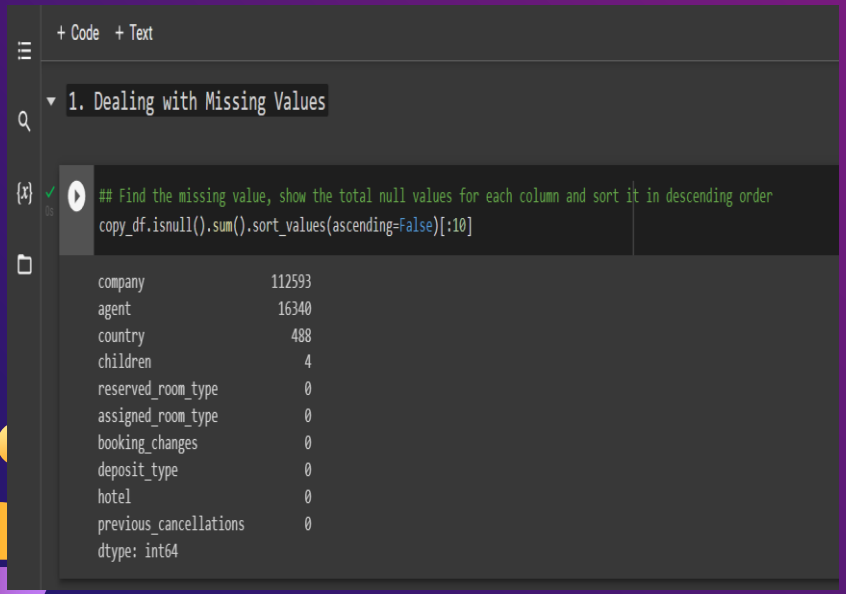
EXPLORATORY DATA ANALYSIS (EDA)

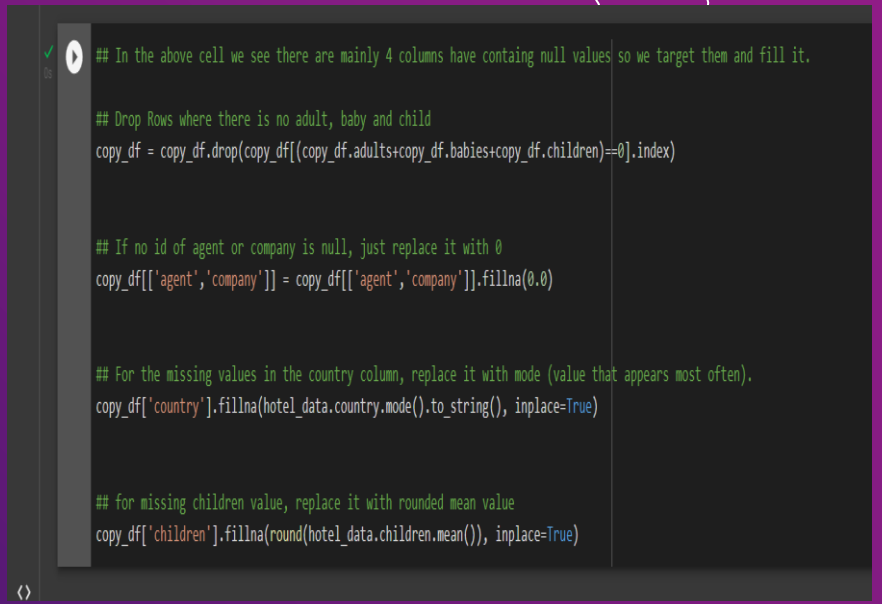
1. We Import all required library in code, so we take advantages ofl ibrary to solve out our problem. If in future we need more library, so we import in this collab. Currently we add numpy , pandas , matplotlib , seaborn, pycountry etc.
2. Then we add our Data-Setfile ie., excel file in it . Our Data-Setfile is in google drive so we import google drive to link with that file & we import google drive then we give location of ourfile then callfile with pandas library with the function of pd.read_csv() .This function read file excelfile.





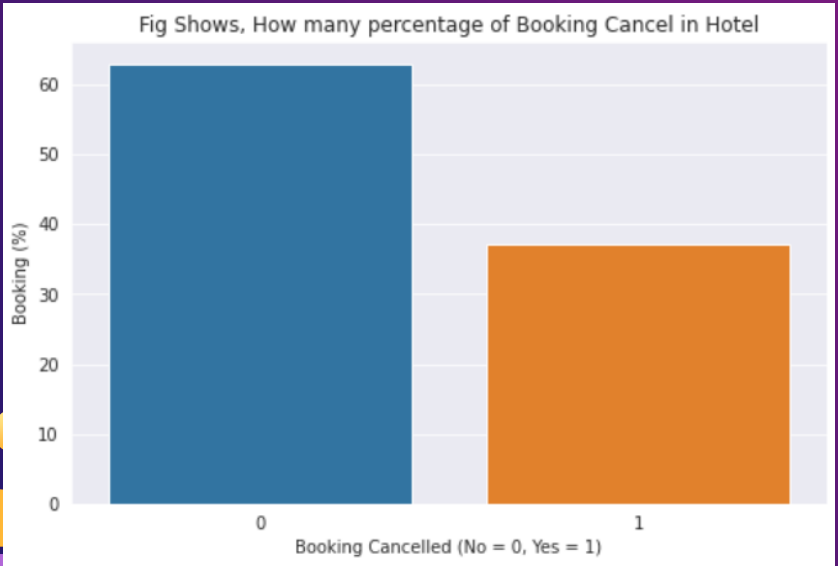
3. Lets Check Missing Value in Dataset & Then Target Every missing value to fill & make data c





Questions to be answered:

1. How Many Booking Were Cancelled ?

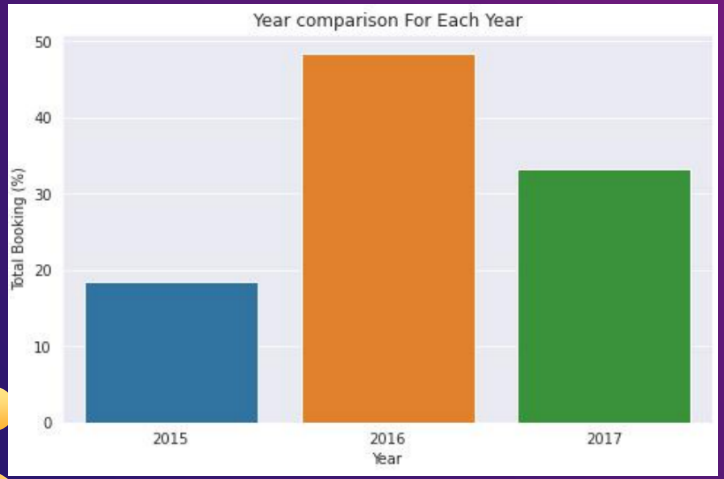


Key Insights:
According to visualization, There are near 37% Booking Cancelled By Customers & remaining 63% of Booking is Safe.

2. What is the booking ratio between Resort Hotel and City Hotel

Key Insights:
ccording to visualization , There are 61% of Booking done
in City Hotel & remaining 39% of Booking done in Resort Hotel. So, City Hotel capture more customers than the Resort Hotel.

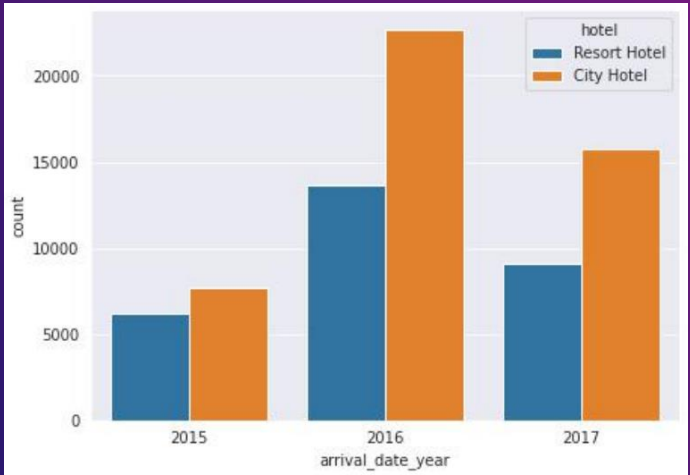
3. What is the percentage of booking for each year ?



Key Insights:

Bar Chart Show the Total % of Booking done in Each Year. In Visualiztion we see that in 2015 Booking done by nearly 19%
and in 2016 48% Booking & in 2017 Booking done 33% .
So, 2016 is Successful Year For Hotel Booking.

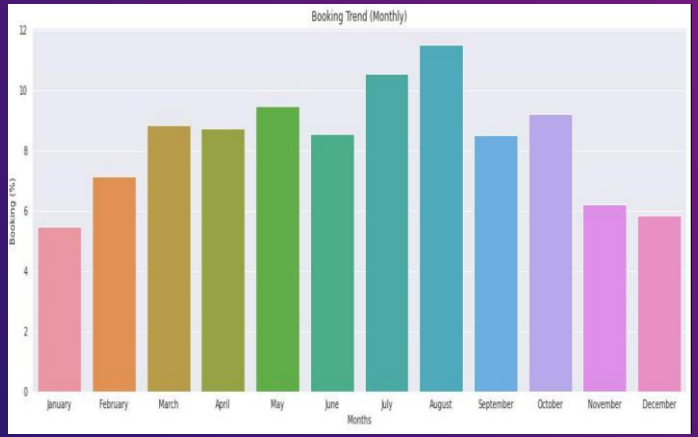
Let's separte it by hotel



Key Insights:

After Comparing the Booking with City vs Resort Hotel. We see City Hotel is most successful for booking in 2015 , 2016 & 2017 with the comparision of Resort Hotel.

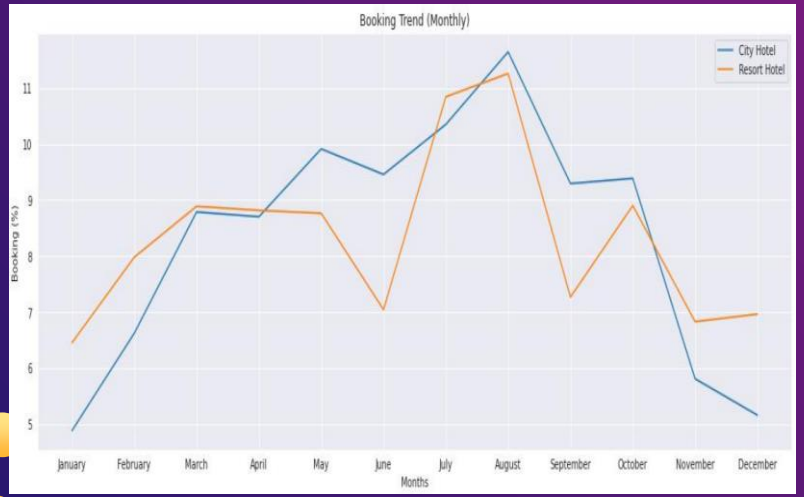
4) Which is the most busy month for hotel ?



Key Insights:

Here is , According to visualization , There are August Month is most busy month for Hotel & July is 2nd most busy month for hotel & then May, October & so on..

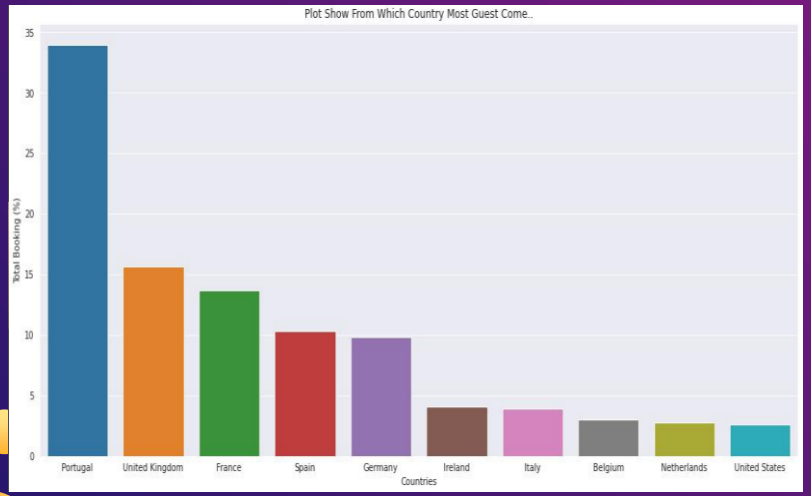
Compare Busy Month for Hotel Between Resort vs City



Key Insights:

Here city Hotel is busy month than Resort in May, June, Aug, Sep, oct month and in Jan, Feb, Mar, april, july, nov & dec month busy for resort hotel more than city hotel.

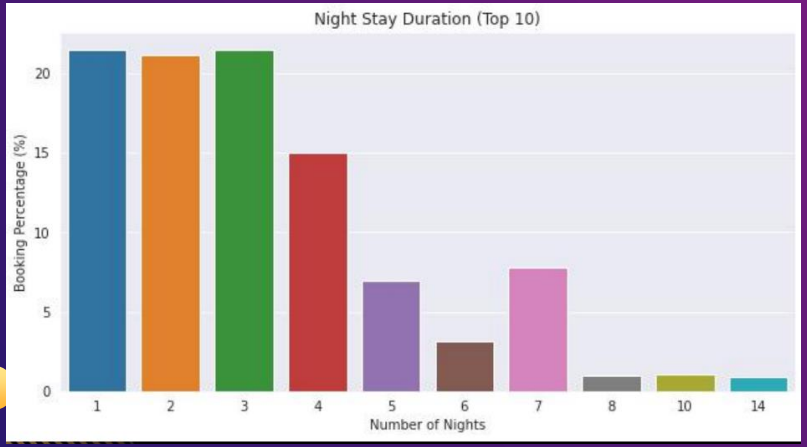
5) From which country most guest come ?



Key Insights:

Most Guest come from Portugal. Nearly 34% Guest come from Portugal and united kingdom 2nd highest country with near 16% & France is in 3rd position with near 14% of guest.

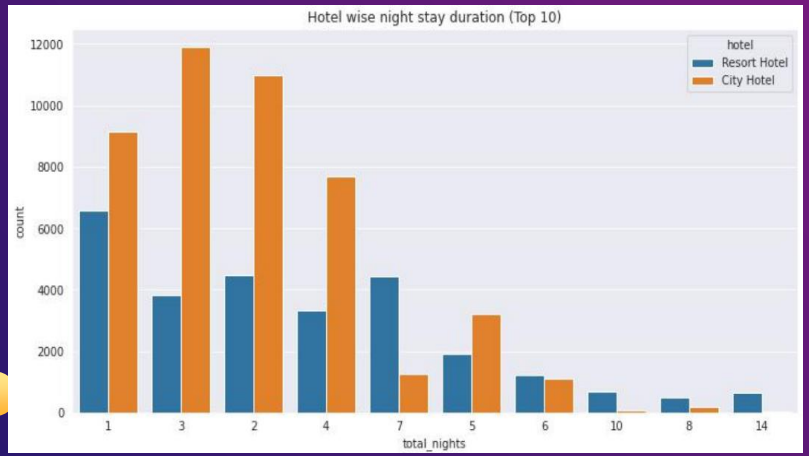
6) How Long People Stay in the hotel ?



Key Insights:

In visualization we see most of people choose 3 Night in hotel. Near 22.5% of people choose 3 night & 22% of people choose 1 night & 21%
people choose 2 nights

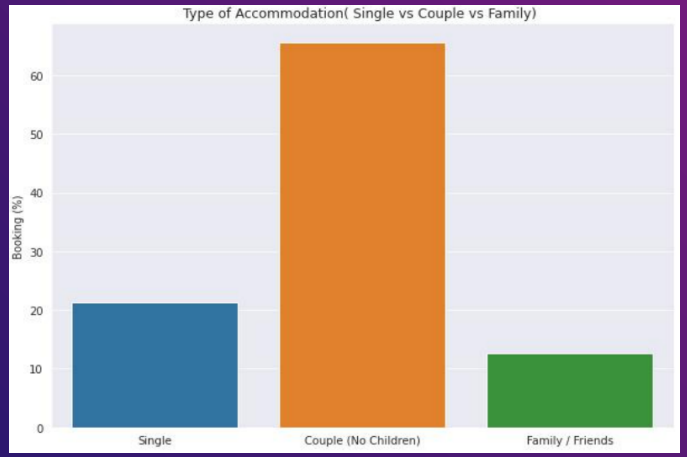
Compare Hotel wise night stay duration resort vs city



Key Insights:

Here is City hotel is more guest for 1, 2, 3, 4 & 5 nights than the resort hotel. In resort hotel people love to stay with 6, 7, 8, 10 & 14 nights.

7) Which is the most booked type of accommodation (Single, Couple, Family) ?



Key Insights:

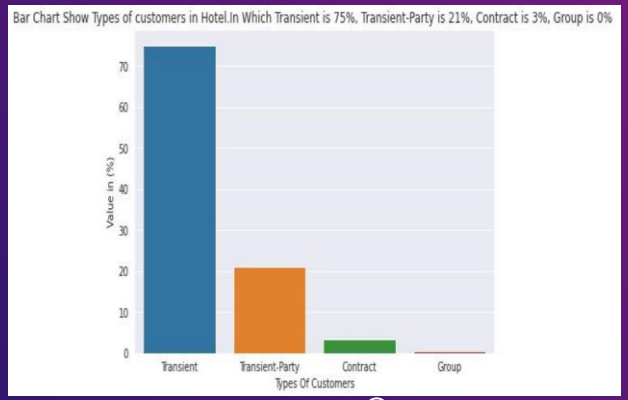
Here is mostly Couple book hotel more than single & Family/Friends members. Ratio of couple for booking hotel is near 66% and for single is 22% and Family/Friends with 12% .

8) How many guests repeated in hotel ?

Key Insights:

Here is only 3% of guest repeated in hotel and remaining 97% of guest not repeated.

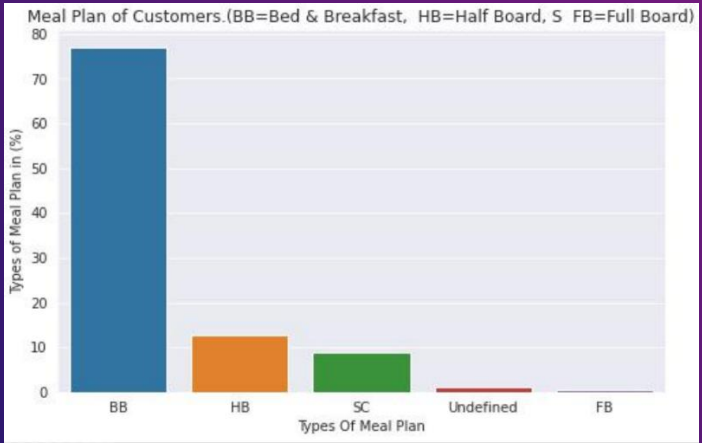
9) Types of Customers ?



Key Insights:

Transient Customer is mostly visited customer more than Transient-party , contract & Group.

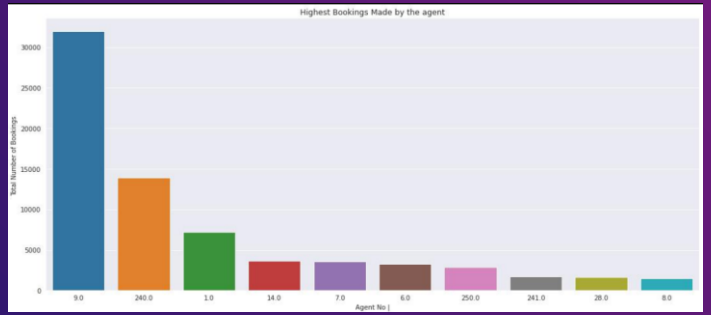
10) Meal Plan of Customers ?



Key Insights:

BB is mostly demanded type of meal according to guest data . Here is BB means Bed & Breakfast. The ratio of BB is nearly 77% & HB is 12.5% & SC is 9.5%. Remaining is Undefined type of meal.

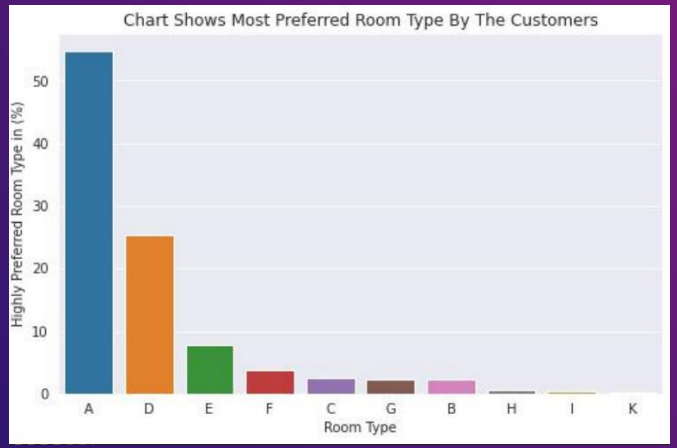
11) Which agent makes highest no. of bookings ?



Key Insights:

Agent no 9 is make most number of booking with above 32000. Agent no 240 is in the 2nd position who has make booking more than 14000.

12) Which is the most preferred room type by the customers ?



Key Insights:

‘A’ type of room most preferred by customers compare of others room types & ‘D’ type of room 2nd most choice by customer & ‘ E ’ is in 3rd position.

13) Which Hotel type has the highest ADR ?

Key Insights:

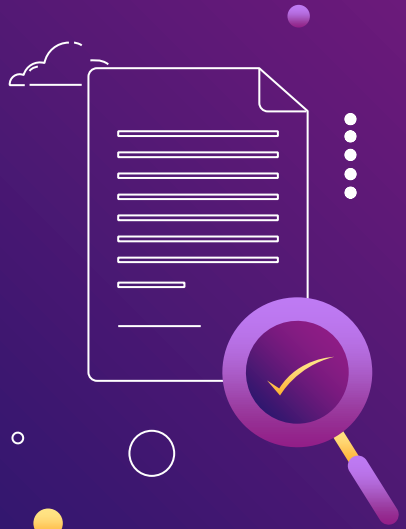
Here is we see in visualization City Hotel is more ADR Than Resort Hotel. City hotel is with nearly 105 of adr & Resort is 85 of adr. More adr(average daily rate) means more revenue. so it means city hotel s are generating more revenues than the resort hotels.

07

INSIGHTS TO BE FOUND

1. There are near 37% Booking Cancelled By Customers & remaining 63% of Booking is Safe.
2. There are 61% of Booking done in City Hotel & remaining 39% of Booking done in Resort Hotel.
3. There are in 2015 Booking done by nearly 19% and in 2016 48% Booking & in 2017 Booking done 33% . So, 2016 is Successful Year For Hotel Booking. We see City Hotel is most successful for booking in 2015 , 2016 & 2017 with the comparison of Resort Hotel.
4. There are August Month is most busy month for Hotel & July is 2nd most busy month for hotel . If we compare , city Hotel is busy month than Resort in May, June, Aug, Sep, oct month and in Jan, Feb, Mar, April, July, Nov & dec month busy for resort hotel more than city hotel.
5. Most Guest come from Portugal. Nearly 34% Guest come from Portugal and united kingdom 2nd highest country with near 16% & France is in 3rd position with near 14% of guest.

6. In visualization we see most people choose 3 Night in hotel. Near 22.5% of people choose 3 night & 22% of people choose 1 night & 21% people choose 2 nights. With comparison City hotel is more guest for 1, 2, 3, 4 & 5 nights than the resort hotel. In resort hotel people love to stay with 6, 7, 8, 10 & 14 nights.
7. Couple book hotel more than single & Family/Friends members. Ratio of couple for booking hotel is near 66% and for single is 22% and Family/Friends with 12% .
8. only 3% of guest repeated in hotel and remaining 97% of guest not repeated.
9. Transient Customer is mostly visited customer more than Transient-party , contract & Group.
10. BB is mostly demanded type of meal according to guest data . Here is BB means Bed & Breakfast. The ratio of BB is nearly 77% & HB is 12.5% & SC is 9.5%. Remaining is Undefined type of meal.

11. Agent no 9 is make most number of booking with above 32000. Agent no 240 is in the 2nd position who has make booking more than 14000.
12. ‘A’ type of room most preferred by customers compare of others room types & ‘D’ type of room 2nd most choice by customer & ‘ E ’ is in 3rd position.
13. In visualization City Hotel is more ADR Than Resort Hotel. City hotel is with nearly 105 of adr & Resort is 85 of adr. More adr(average daily rate) means more revenue. so it means city hotels are generating more revenues than the resort hotels.

THANK YOU!