EDA Case Study - Business Analytics in Hospitality Domain

Background:

AtliQ Grands owns multiple five-star hotels across India. They have been in the hospitality industry for the past 20 years. Due to strategic moves from other competitors and ineffective decision-making in management, AtliQ Grands are losing its market share and revenue in the luxury/business hotels category. As a strategic move, the managing director of AtliQ Grands wanted to incorporate "Business and Data Intelligence" to regain their market share and revenue.

1. Import the data file in R

```
df<-read.csv("Hospitality_Data.csv",header=TRUE)
head(df)
```

```
         booking_id property_id booking_date Month.Name  Day.Name
weekday
1 May012216558RT11       16558     27-04-22       April Wednesday
weeke day
2 May012216558RT12       16558     30-04-22       April  Saturday we
ekend day
3 May012216558RT13       16558     28-04-22       April  Thursday
weeke day
4 May012216558RT14       16558     28-04-22       April  Thursday
weeke day
5 May012216558RT15       16558     27-04-22       April Wednesday
weeke day
6 May012216558RT16       16558     01-05-22         May     Sunday we
ekend day
  no_guests room_category booking_platform ratings_given booking_st
atus
1         3          RT1    direct online             1     Checked
Out
2         2          RT1           others          null      Cance
lled
3         2          RT1         logtrip             5     Checked
Out
4         2          RT1           others          null      Cance
lled
5         4          RT1    direct online             5     Checked
Out
6         2          RT1           others             4     Checked
Out
  Booking.status revenue_generated revenue_realized Revenue_lost We
ek.of.Year
1              1             10010            10010            0
18
2              0              9100             3640         5460
18
3              1              9100             9100            0
18
4              0              9100             3640         5460
18
5              1             10920            10920            0
18
6              1              9100             9100            0
19
  No.of.Days dim_rooms property_name category  city successful_book
ings
1          1  Standard  Atliq Grands    Luxury Delhi
18
2          1  Standard  Atliq Grands    Luxury Delhi
18
3          3  Standard  Atliq Grands    Luxury Delhi
18
4          1  Standard  Atliq Grands    Luxury Delhi
18
5          1  Standard  Atliq Grands    Luxury Delhi
18
6          2  Standard  Atliq Grands    Luxury Delhi
18
  capacity Unsuccessful_bookings
1       19                    1
2       19                    1
3       19                    1
4       19                    1
5       19                    1
6       19                    1
```

2. Obtain table of booking status using showing count

Using table function

```
Booking_status<-table(df$booking_status)
Booking_status
```

```
  Cancelled Checked Out     No Show
        232          720           48
```

Using dplyr

```
library(dplyr)
Booking_status<-df %>% group_by(booking_status) %>%
  summarise(Count=length(booking_id)) %>% as.data.frame()
Booking_status
```

```
  booking_status Count
1      Cancelled   232
2    Checked Out   720
3       No Show    48
```

3. Show the table of total bookings (count) by each booking platform with percentages

Using table function

```
Platform<-table(df$booking_platform)
Platform
```

```
direct offline  direct online         journey         logtrip   makey
ourtrip
            57             82              57              94
194
        others       tripster
           438             78
```

```
Platform<-round(prop.table(Platform)*100,2)
Platform
```

```
direct offline  direct online         journey         logtrip   makey
ourtrip
           5.7            8.2             5.7             9.4
19.4
        others       tripster
          43.8            7.8
```

Using dplyr

```
Platform<-df %>%
  group_by(booking_platform) %>%
  summarise(Count = n()) %>%
  mutate(Percentage = (Count / sum(Count)) * 100) %>% as.data.frame
()
Platform
```

```
  booking_platform Count Percentage
1   direct offline    57        5.7
2    direct online    82        8.2
3          journey    57        5.7
4          logtrip    94        9.4
5     makeyourtrip   194       19.4
6           others   438       43.8
7         tripster    78        7.8
```

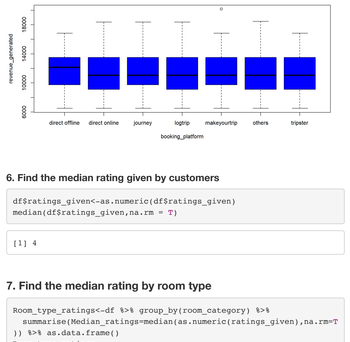
4. Show table of booking volume by day of week and arrange in descending order

```
Booking_Day<-df %>% group_by(Day.Name) %>% summarise(Total_Bookings
=n_distinct(booking_id)) %>% arrange(desc(Total_Bookings))  %>% as.
data.frame()
Booking_Day
```

```
   Day.Name Total_Bookings
1    Sunday            206
2    Friday            166
3  Thursday            164
4 Wednesday            139
5  Saturday            135
6   Tuesday            100
7    Monday             90
```

5. Visualize the spread of revenue generated by each booking platform

```
boxplot(revenue_generated~booking_platform,data=df,col="blue",main
= "Revenue Generated by Booking Platform")
```



6. Find the median rating given by customers

```
df$ratings_given<-as.numeric(df$ratings_given)
median(df$ratings_given,na.rm = T)
```

```
[1] 4
```

7. Find the median rating by room type

```
Room_type_ratings<-df %>% group_by(room_category) %>%
  summarise(Median_ratings=median(as.numeric(ratings_given),na.rm=T
)) %>% as.data.frame()
Room_type_ratings
```

```
  room_category Median_ratings
1           RT1              4
2           RT2              4
3           RT3              4
```