Capstone Project Submission

Team Member's Name, Email and Contribution:
Name -Deepak Kumar Dubey Email -dkd31994@gmail.com This project was done by me individually.
Please paste the GitHub Repo link.
Github Link:-https://github.com/dkd99/hotel-booking-analysis
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)
So I had a hotel booking data at my hand through which i had to analyze some KPI's which are average daily rate,average length of stay,meals ordered per booking,total bookings. After looking at the data some questions popped in my head. Those were- **1)What months of the year get the most bookings?Does the average daily rate get affected by this??** **Approach**-groupby the data with months and then use aggregate count function to get monthly bookings and mean function for adr column to get average adr for each month .I also found out average adr over the whole time period and drawn it on the same chart where average monthly adr was plotted to get better visualization. **Conclusion**- may to september months were getting the most amount of bookings and average adr for these months was also high as compared to other months. **2)Is there any relation between number of bookings happening and lead time?** **Approach-**Average monthly bookings and avg monthly lead time were organized in a data frame.Using that data frame i found out correlation coefficient between two variables.

Conclusion- Correlation coefficient came out around 0.82 which represents strong positive correlation.

3)How many bookings got canceled?

Approach- Group by data based on canceled(1) or not canceled(0) and then use aggregate count function.
conclusion- Around 30% bookings were getting canceled.

4)Is there any relation between the number of bookings and cancellations?
Approach-Arranged both the data in a data frame and then found out correlation coefficient between them.
Conclusion-correlation coefficient of 0.98 was found out which represents almost perfect correlation.

5)What type of hotel gets more bookings?
Approach-Grouped by data based on hotel type and then used aggregate count function.
Conclusion-City type hotels were getting twice the number of bookings than Resort type.

6)People who ought to stay for more nights opt for which type of hotel?
Approach-Grouped by data based on hotel and then found average nights of stays using aggregate function.
Conclusion-People who ought to stay longer preferred Resort type hotels.

7)Does the size of the family affect which type of hotel the family opts for?
Approach-Divided the data and arranged it into 3 categories-single adult,couples having no babies and children and families having more than 3 members.
Conclusion-Found out families having children and babies preferred Resort type hotels over City types a little more.

8)How many bookings required parking space and which type of hotel did they opt for?
Approach-Counted number of bookings which required parking and then those who required parking, grouped them based on hotel type.
Conclusion-Just around 6-7% of people required parking spaces and most of them opted for Resort type hotels.

9)What is the average size of families against all the bookings?
Approach-Divided the data and arranged it into 3 categories-single adult,couples having no babies and children and families having more than 3 members.
Conclusion-Couples(no babies and children) booked around 60% of the bookings followed by single adults.

10)Through what type of distribution channels do customers get their bookings done?
Conclusion-Most of them are done through online traveling agents.

12)How many customers had to wait for their bookings to get confirmed?Which type of hotel had more waiting time?
conclusion-Around 3 % people had to wait .Since most people opt for city type hotels ,city type has more avg waiting time.

13)How many customers opted for a meal and which type of meal was most preferred?

Conclusion-Almost 98 % of customers opted for meals and most opted for bed and breakfast time i.e BB type.

14)What was the average number of nights stay when averaged over all bookings?
Conclusion-3.42 nights .Distribution plot was drawn for clear visualization.