Welcome

Problem Statement:

understand the movement of the price as the day approaches the check-in date. Attached is a randomly sampled booking data from five different cities with check-in between 10/10/2016 – 12/31/2016.

Approach:

Merged the raw data from multiple cities, then started creating the analysis, going through the data cleaning process then creating new variables to dig deeper in the funnel before building the hypothesis.
Packages used are: Pandas - Numpy - Statistics - Scipy - Seaborn - matplotlib - Scikit learn

At which part of the funnel should this message be implemented? On a homepage, search results page, at the property level, or booking form?

Answer: Homepage, results page, and personalized notification on the user's email. As it says it needs to start with an action verb and urgent sense to act. Using my analysis can highlight the percentage the user can save money while booking the desired destination.

What assumptions are you making about the data?

Answer: Main focus was:

- Select the most revenue hotels and resorts in the different cities.
- Create a hypothesis regarding time period between booking and the check-in.
- Trying to use the model to predict the hotel prices values using linear regression and gradient boosting regressor, followed by evaluation of the accuracy of models.
- Check the booking days that have high cumulative total revenue for all cities.
- Check the profits of the hotels that the users booked in less than 20 days, since this data is related to the prime months. That means this category pay more on specific hotels and I wanted to check this factor as it will definitely add more value when contributing the discounts and offers.

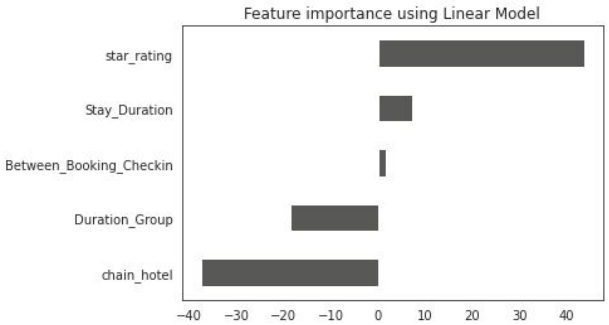
What factors play a role on price in this analysis?

Answer:

Most important factors in this analysis that positively impact the prices are the hotel rating - Stay duration (The time period between the check-in and check-out from the hotel) - the duration between the booking and check-in date.

Then other factors that have strong negative impact are: Hotel chain - Duration group.

Kindly note that the strong negative correlation means the more of something makes the less of the other thing. like the more time prior between the booking and check-in dates the less money spend on booking, and so forth.



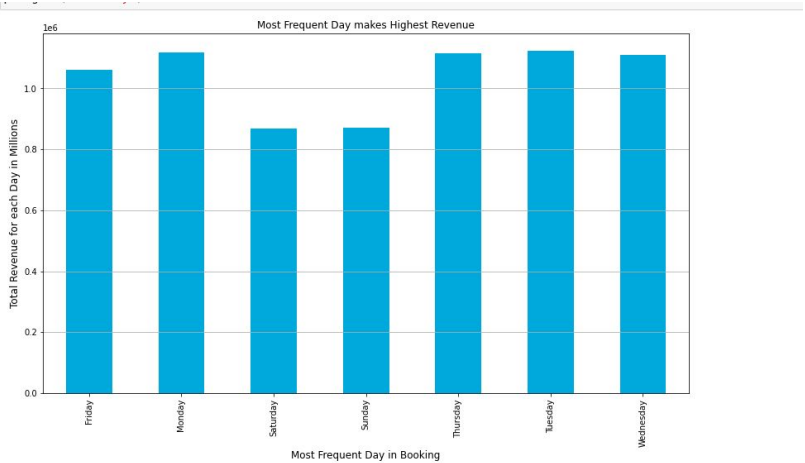
What other factors may influence the analysis that isn’t available in the dataset? What would improve the analysis?

Answer:
(Practical metrics insights I was able to obtain if):

- The number of occupied rooms in each hotel: so I can generate the ADR "Average Daily Rate" metric for the hotels. “I have only the total revenue for less than 3 months”
- Average length of stay (ALOS) metric to help me personalise my analysis and categorize the users using clustering analysis. So we can promote "email-campaigns" where to send personalised discounts to the right users and increase the overall conversion rates.

(General features could play a vital role in this analysis are):

- Hotel Reviews. “Users select the hotel mainly on the rating, reviews and safety”
- How close the hotel is to the city centers and malls.
- Raw data for at least 1 year to check the seasonality/trends in the analysis.



What conclusions can you draw from the data?

Answer:

The popularity of the hotel depends on the hotel rating, there are specific hotels make higher than 150k in less than 3 months the users are willing to pay even in the most busy part of the year.

Hotels with ratings 4 then 3 are the most visitable compared to hotels with the other ratings. In addition to the other notes I have clarified previously and in the next sections, kindly read the interpretation below each part for detailed insights. “HTML file attached”

The most frequent duration of staying in the hotel is 1 day. The booking traffic happens all week except Saturday and Sunday. Both days have the least recorded values.

What recommendations would you give to the Product Owner?

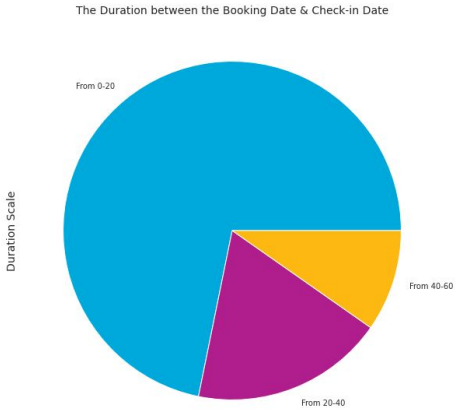
1. As shown in the figures: If Agoda user booked his hotel in less than 20 days prior. He will be saved from prices exposure of 60.38%. Unlike if he decided to book the hotel between 20 to 40 days prior to his check-in, then he's actually would be saved from prices exposure of just only 18.49%.

   The main idea here is:

   Booking your hotel with sufficient days before your check-in date, the sufficient time prior the check-in you give yourself, means more money into your pocket.

3. High resorts with impressive revenue in less than 3 months are:

   a. The resort with id = 21720 recorded a revenue close to 140k in less than 3 months.
   b. The resort with id = 508110 recorded a revenue close to 60k in less than 3 months.
   c. The resort with id = 3644 recorded a revenue close to 35k in less than 3 months.
   d. The resort with id = 1968534 recorded a revenue close to 15k in less than 3 months.



Duration in Days Between the Booking & Check-in	Save Money Percentage of
Between 0-20 Days	60.38%
Between 20-40 Days	18.49%
Between 40-60 Days	9.71%

What recommendations would you give to the Product Owner?

4. Highest Hotels with impressive revenue:

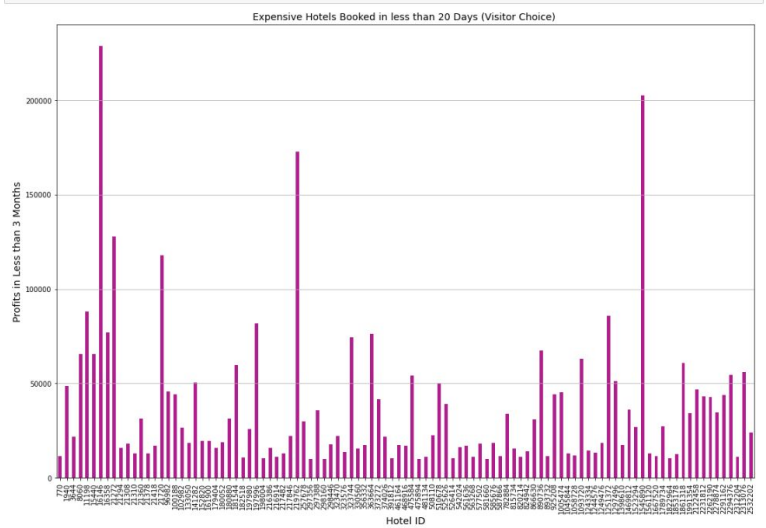
For the city 9395 we find that:
- Hotel(21272) has revenue close to 160k for less than 3 months.
- Hotel(197996) has revenue more than 100k for less than 3 months.

For the city 17193 we find that:
- Hotel(1243334) has revenue more than 20k for less than 3 months.
- Hotel(1244576) has revenue more than 17.5k for less than 3 months.

For the city 5085 we find that:
- Hotel(1251372) has revenue close to 175k for less than 3 months.
- Hotel(100188) has revenue more than 100k for less than 3 months.
- Hotel(2231812) has revenue close to 100k for less than 3 months.

For the city 16808 we find that:
- Hotel(16146) has revenue more than 300k for less than 3 months.
- Hotel(1545890) has revenue more than 205k for less than 3 months.
- Hotel(219762) has revenue more than 180k for less than 3 months.

For the city 8584 we find that:
- Hotel(323744) has revenue more than 94k for less than 3 months.
- Hotel(374026) has revenue more than 20k for less than 3 months.



Thank you