



TAKE-HOME CHALLENGE: Data Science Analytics

Metrics¹



1. Major Metric²: booking rate³
 - Definition: portion of the inquiries leading to a booking
 - Reason: I pick booking rate metric cause it is related to conversion rate and, if I move it, I can realistically improve company growth and revenue.
2. Quantity Metric: 7-days booking quantity
 - Definition: quantity of bookings within a week
 - Reason: This represents a metric for quantity of bookings through the website which determines the revenue. The previous metric is in terms of rate and this one gives a sense of absolute number of bookings.
3. Engagement Metric: 28-days active guests
 - Definition: number of unique guests who are inquiring within 4 weeks
 - Reason: This type of metric is to indicating the user engagement through the website. Although not directly determining the booking rate, this metric has correlation with booking rate and revenue so we like having more active users on our website.
 - Side note:
 - ✓ We have two types of users; power users and casual users. The first group are people like marketing speakers who use Airbnb frequently. Casual users are the ones with few number of trips per year.
 - ✓ These two communities have different needs. If a change decreases the booking quantity of casual users by 20% but increases the booking quantity of power users by 10%, this may reflect negatively in the engagement metric but increase the booking quantity metric.

1. Please refer to the appendix to see how the metrics are exactly computed.

2. This is the metric we will use through the analysis.

3. The total booking rate in the dataset is 41.51% in 27790 inquiries (after removing outliers).

Metrics

4. Detailed Metric: “contact me” reply rate

- Definition: portion of “contact me” inquiries by guests that receives a response from hosts
- Reason: this metric indicates the quality of guest-host contact relations

5. Detailed Metric: “contact me” acceptance rate

- Definition: portion of “contact me” inquiries by guests lead to a booking, i.e., accepted to the booked ratio of the “contact me” channel
- Reason: this indicates how effective is the “contact me” option in terms of ultimately resulting in a booking.
 - We will see our weakness and the potential opportunities through this field later

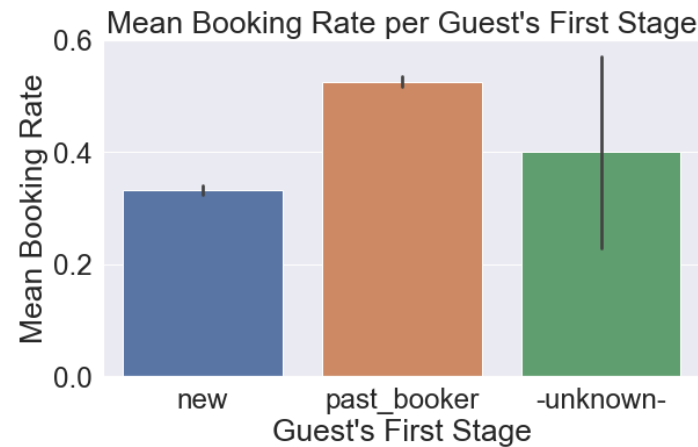
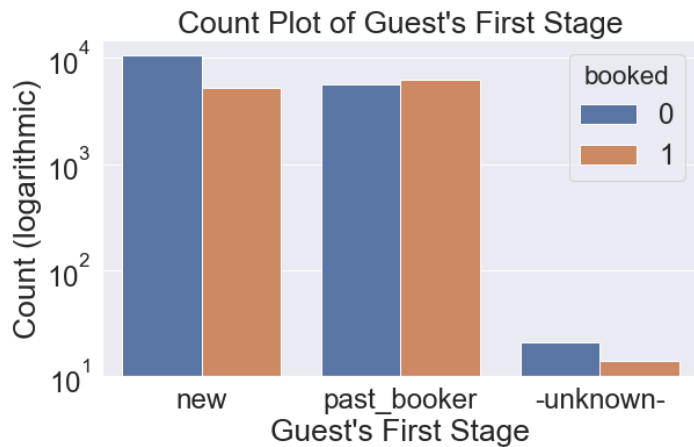
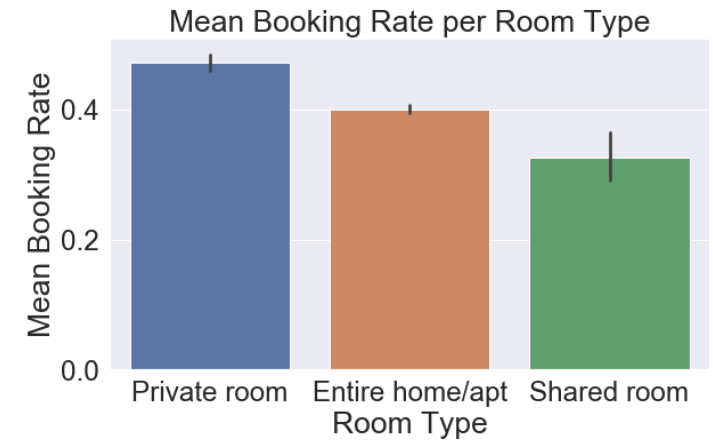
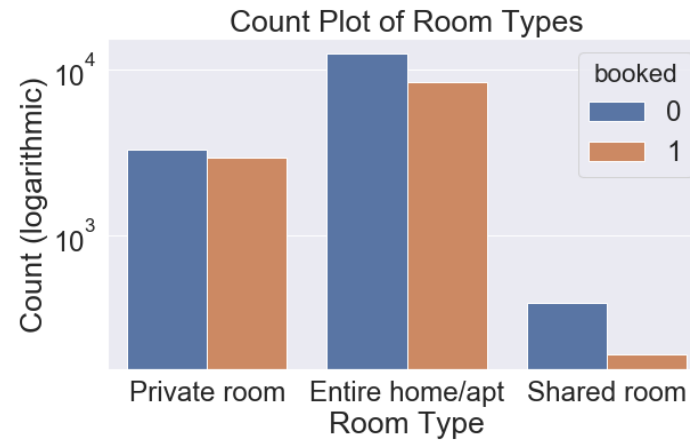
6. Detailed Metric: “book it” acceptance rate

- Definition: portion of “book it” inquiries by guests accepted by hosts
- Reason: this indicates how effective is the “book it” option in terms of contributing to the booking



Analysis 1: basic

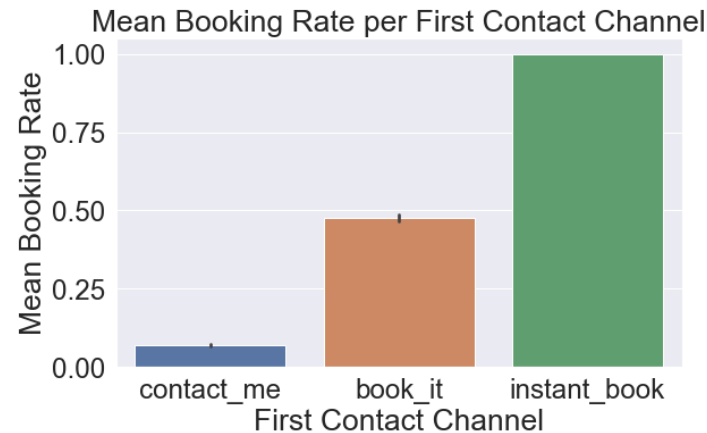
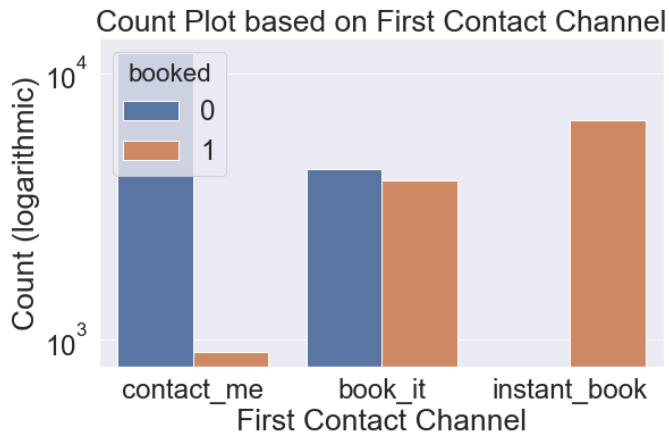
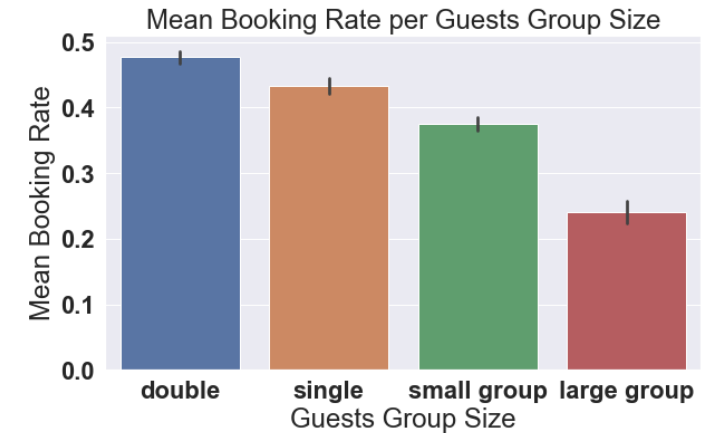
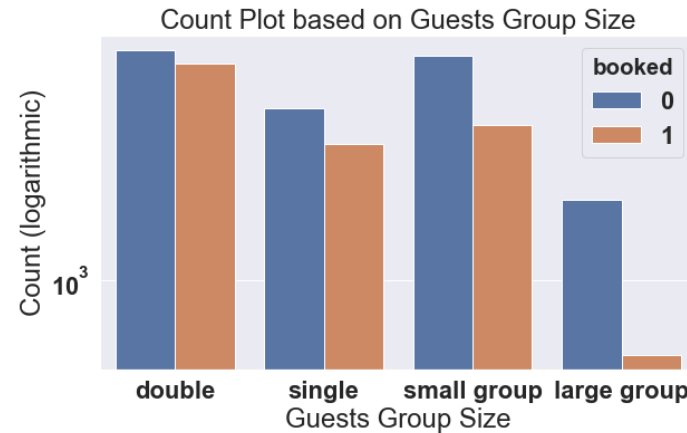
- Private rooms have the best and shared rooms have the worst booking rates.
- Two main reasons can be mentioned for lower booking rate of shared rooms:
- People interested in shared rooms are typically more price sensitive and also they care about who are their roommates. Hence, they search more to rent a shared room which results in a lower booking rate.



- Lower booking rate of new users is normal
- There may be an opportunity here that will be discussed later.

Analysis 1: basic

- Double travelers are the ones with the highest booking rate and large groups have the lowest booking rate by far.
- This makes sense as hosts don't tend to rent their property to large groups of people. On the other hand, as the number of persons in a group grows, it gets harder to reach a consensus.
- Listings suitable for large groups are typically more expensive. There may be an opportunity here that will be discussed later.



- We see that there are lots of users making use of "contact me" option, but their booking rate is very low.
- Booking rate for guests using "instant book" option is 100% as anticipated.
- "Book it" option has fairly low booking rate (<50%)

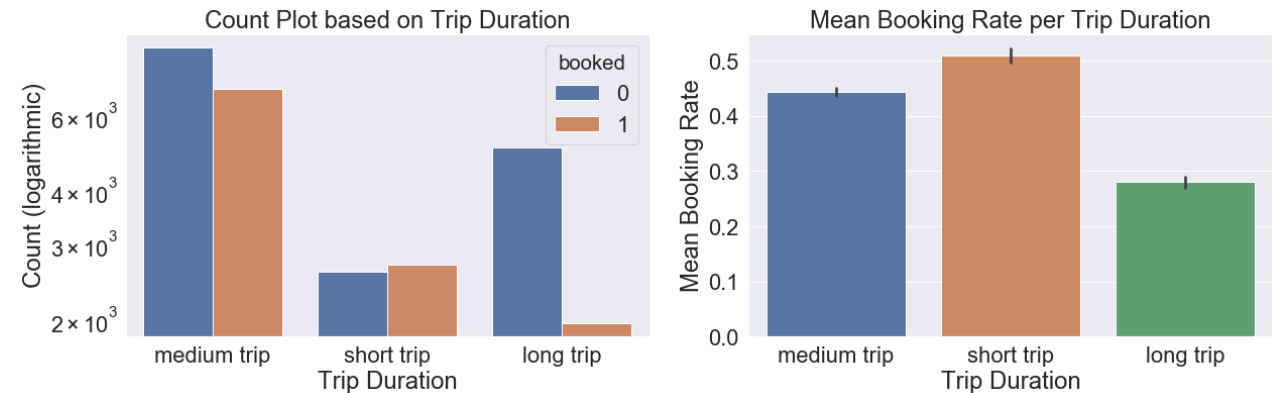


The same analysis can be done on the other features if there is more time.

will be used in analysis

Analysis 3: feature engineering

- Creating new features like
 - Trip duration (check-in to check-out time):
 - Short trips (<3 days), medium trips (3 to 7 days), long trips (>7 days)
 - Trip types:
 - Leisure Trips: check-in time on Friday and check-out time on the following Saturday or Sunday, or check-in time on Saturday and check-out on the following day
 - Work/business trips: all the other trips by assumption¹
 - Check-in time is typically 3 p.m.; hence, trips with check-in time on Friday and check-out on the same weekend will be considered as leisure trips.

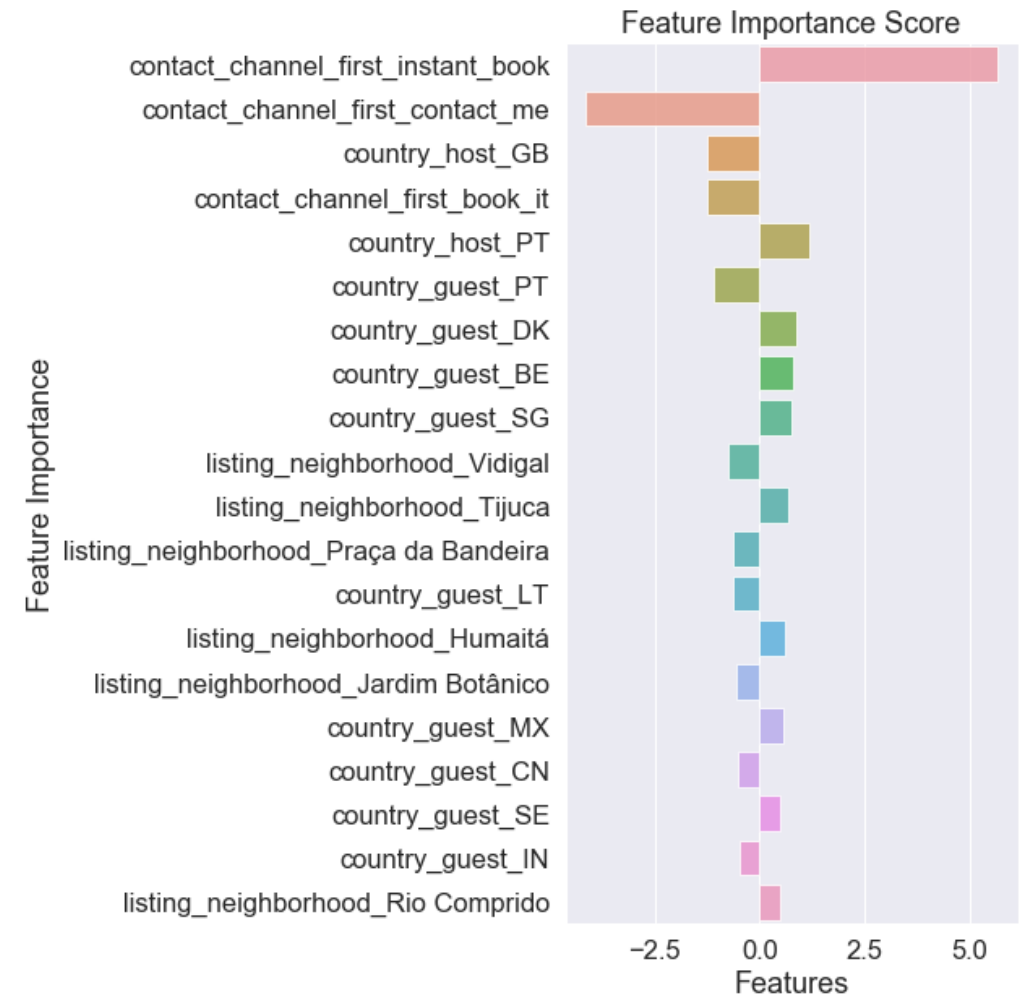


For example, plotting the booking rate versus trip duration groups will tell us that the best conversion rate is for short trips and the worst is for long trips. This is expected as people travelling for a long time usually select the residence place more carefully.

1. If I had more time, I could work to strengthen this assumption and make it more accurate.

Analysis 4: ML classification

- ❖ The plot shows the 20 most important variables on the booking rate (either with positive or negative impact) extracted from Logistic Regression classification method (technical details in the attached Jupyter notebook)
- Instant "book feature" is the most important positive variable (trivial).
- "contact me" and "book it" options have the highest negative impacts on the booking rate.
- Hence, "Trust" is a challenging issue that should be addressed. I give recommendations on this later.
- The other important observation is the high negative impact of India and Ipanema as the guest country and listing neighborhood respectively.



Product ideas & business insights



✓ Opportunities in contact channels and booking options

▪ (priority 1 with highest impact)

- Low booking rates for “contact me” and “book it” options indicates a “Trust issue”.
- Considering how much guests are interested in “contact host” option, there is a huge opportunity if we target these users and improve the booking rate. We should consider whether our website has good guest-host contact features and satisfies users' needs.
- How about adding a guest-host in-app messenger? Should be assessed by an A/B test.
- There isn't a lot of guests using "instant book" feature. It's a huge opportunity for us to target hosts and incentivize them to opt-in this feature and the guests to use this. For example, put a discount for both sides separately when using instant book.
- An interesting idea will be improving the binary scheme in instant book selection as follows:
 - As described, hosts can opt-in this feature for guests or not, there is no flexibility.
 - I recommend optimizing the instant book opt-in option based on the guest quantity and rate.
 - Example: A host may choose to opt-in the instant book option if the guest is single or couple, and the guest's rate is above or equal to 3 from 5.

Product ideas & business insights



✓ Opportunities in new users and large group guests (priority 2)

- I will reach the UI team to check whether there is something confusing on the website at the first look that makes the booking process difficult for new users.
- In this case, including a tutorial will help.
- There may be an opportunity in investigating the conversion funnel as well.
- Low booking rate for large guests groups may be a sign of search algorithm weakness to match the users inquiring for large number of guests with the best possible listings.
- As listings for large group of guests have more prices, there is an opportunity. If the low booking rate of large group listings increases slightly, it may contribute to the revenue considerably.
- Software development team may assign a tag to listings with the rent history by a large group. Then, prioritize them in the search results for the guests in looking for large group rentals.

✓ Opportunities in countries, neighborhoods, and private rooms (priority 3)

- I would recommend the cyber security team to inspect the activity log of users from India and other countries with negative impact to see if there is any spam, fake or malware behavior.
- As Ipanema is a popular beach neighborhood in Rio, its negative impact on the booking rate does not make sense. We have to investigate the listings in this neighborhood to fix any issue like unreasonable pricings. Also, the guest countries with a negative impact should be considered for possible fixes like weaknesses in website translation, culture issues, etc.
- We can invest on attracting more hosts to advertise private rooms like by making the process easier, offering discounts. Also prioritize private rooms recommendations in the search results (when guest is not filtering out private rooms which is typically the case).

Appendix

❖ How to compute metrics 1, 4, 5, and 6:

- booking rate: number of items in “ts_booking_at” divided by the number of items in “ts_interaction_first”
- “contact me” reply rate: number of items in “ts_reply_at_first” divided by the number of items in “ts_interaction_first” where “contact_channel_first” is “contact_me”
- “contact me” acceptance rate: number of items in “ts_booking_at” divided by the number of items in “ts_interaction_first” where “contact_channel_first” is “contact_me”
- “book it” acceptance rate: number of items in “ts_booking_at” divided by the number of items in “ts_accepted_at_first” where “contact_channel_first” is “book_it”