

First Glance at Bookings:

Though booking success rates are increasing, total number of bookings is flat



Figure 1

- Definitions:
 - Booking success (blue)*: an interaction is a success if data column `ts_booking_at` is populated
 - Booking failure (red)*: a first interaction occurs, but no booking occurs
 - Booking success rate* = $\frac{\text{number of booking success}}{\text{number of total first interactions}}$
- Overall booking success rate in these six months= 43%
 - This success rate has drastically increased per month:
Jan: 28% → Jun: 49%
 - Mainly due to red line decreasing. We should aim to increase blue line as well
 - January is a holiday season thus many more people looking, but unwilling to commit to booking
 - Relatively new product in Rio de Janeiro thus number of reviews and experienced hosts are sparse

To increase overall bookings in Rio de Janeiro, we should focus on two main objectives (in order of priority):

- Objective 1: Increase our total interactions (red + blue lines)
- Objective 2: Increase the booking success rate of our interactions (blue line / (red+blue lines))

Where should we grow our total interactions (objective 1)?

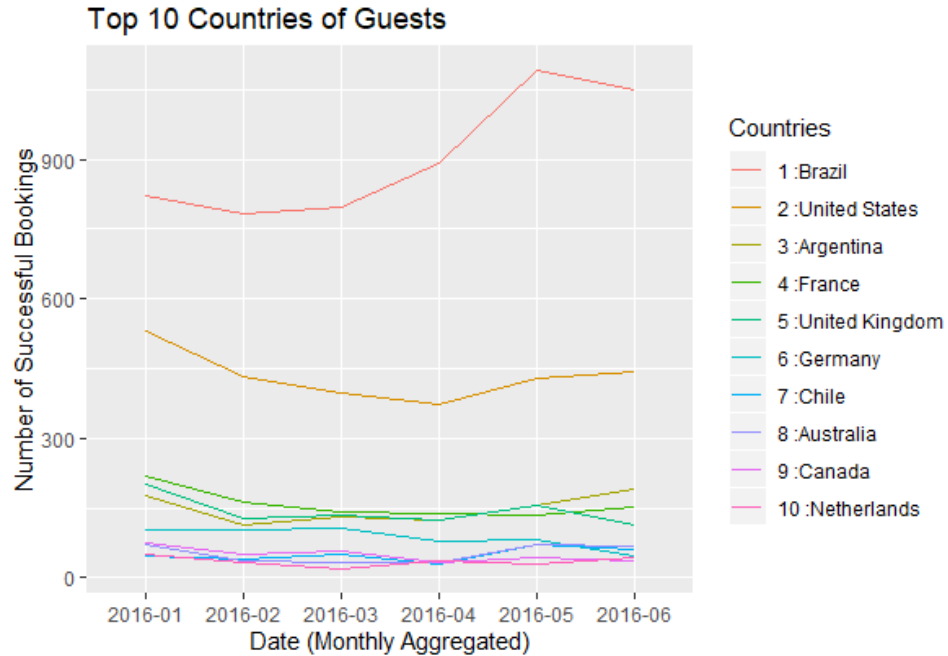


Figure 2

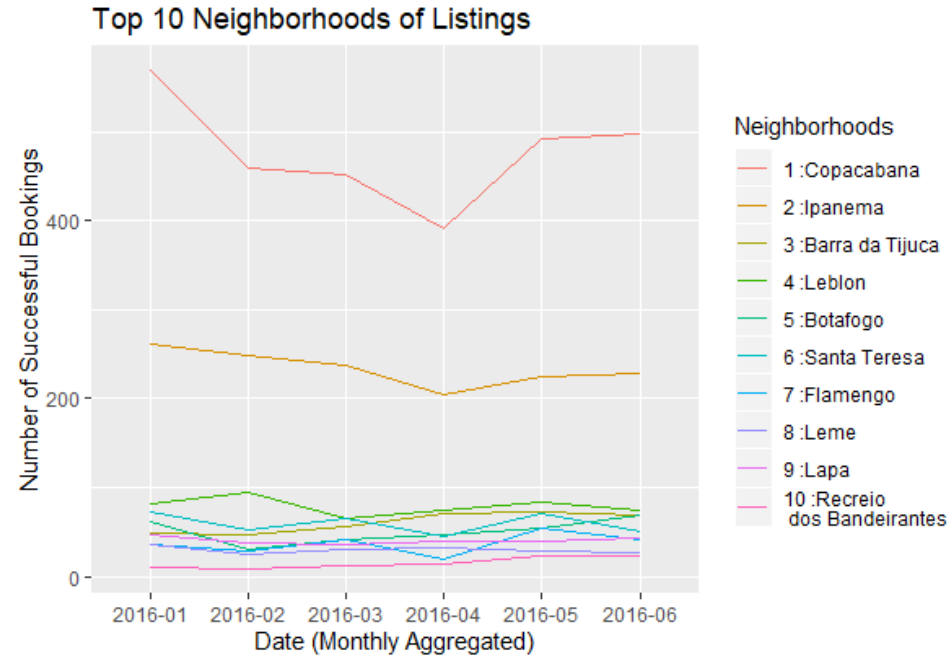
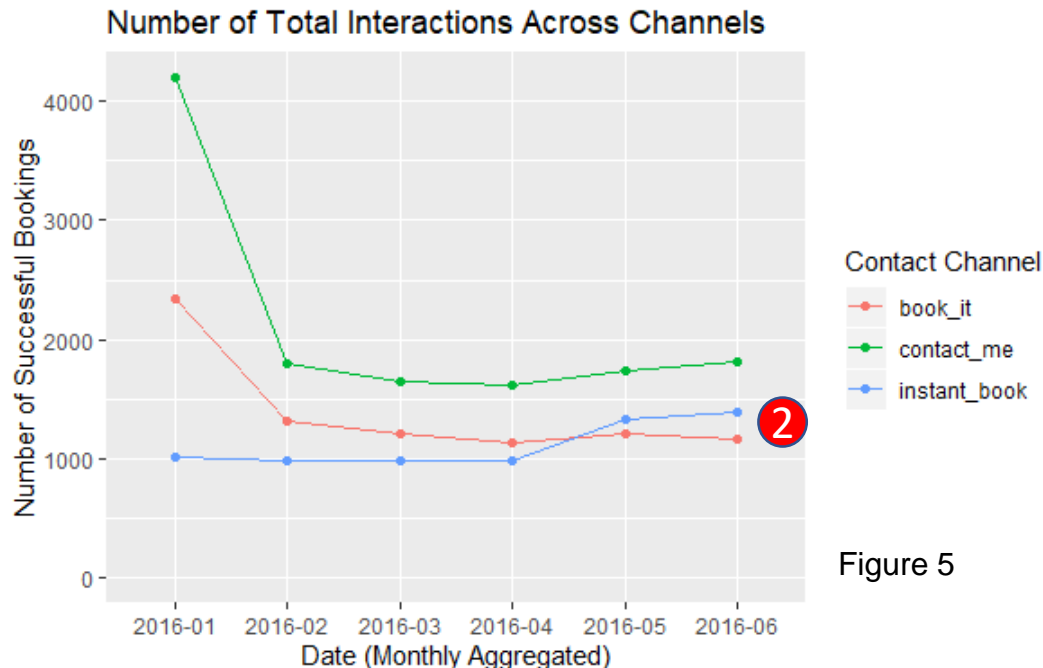
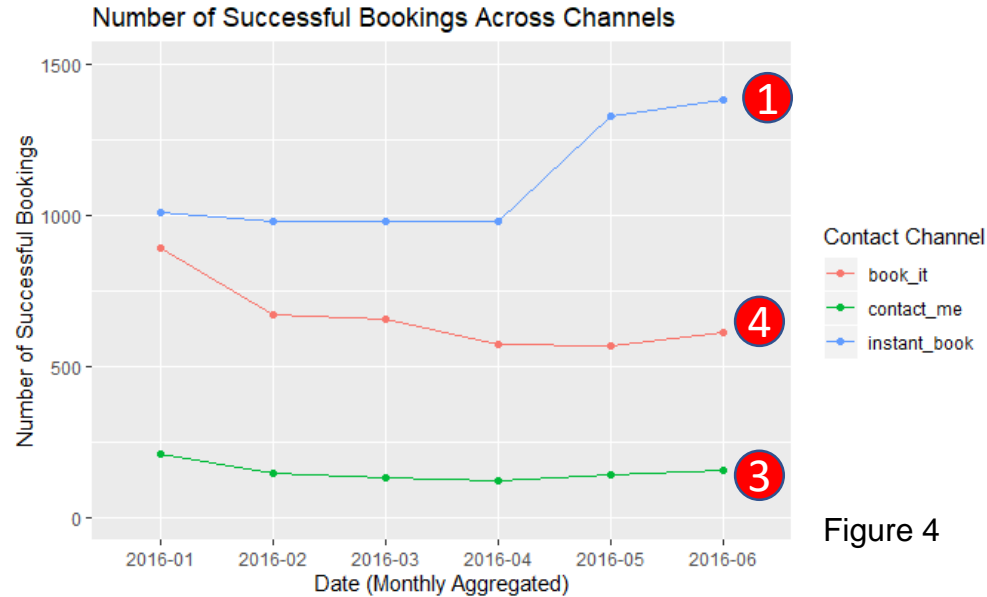


Figure 3

- The top two countries of guests and the top two listing neighborhoods dominate our market
- **[Initiative]**: Increase targeted advertising and promotional deals towards:
 - Guests from **Brazil** and **United States**
 - Brazil has experienced a large trend in growth since March, but has been leveling off since May
 - Argentina is a far third, but has a steady upwards trend
 - Listings in **Copacabana** and **Ipanema**
 - Copacabana had a severe dip in listings in April, but is on an overall upwards trend
 - Ipanema is a solid second but stagnating and slowly losing bookings over time

A closer look at how guests book



Areas doing well:

- [1] The instant_book channel has the highest success rate, guaranteeing bookings at 100%. It's also the channel of highest trending growth
- [2] It seems people are preferring to use instant_book over book_it

Areas to improve:

- [3] The contact_me channel only has 7% success rate. Though it's the channel with highest total interactions, it's by far the channel with lowest successful bookings
- [4] The book_it channel has 48% success rate, a surprisingly low rate considering the guest has offered up front money
 - Hosts may be rejecting due to sparseness/ low quality of guest reviews
 - Hosts may receive multiple offers on one listing, forcing them to decline offers

Actions to take:

- [Initiative] Create incentives for hosts to list through instant_book rather than book_it
 - Mark instant_book hosts as "premium hosts" who get priority on search results
 - Offer promotional economic incentives towards qualified book_it hosts to switch to instant_book
- Increase the success of bookings through book_it and contact_me (explored further on next slides)

How can hosts help increase booking success rates (objective 2)?

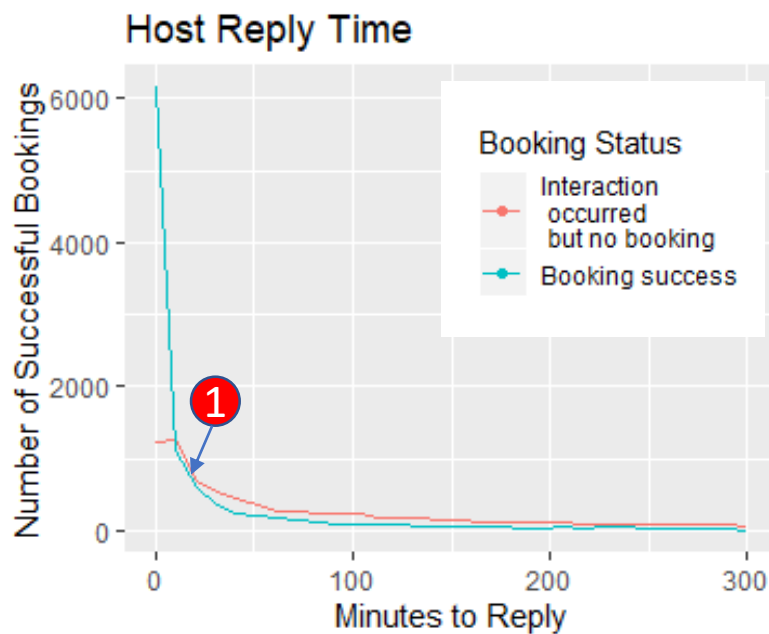


Figure 6

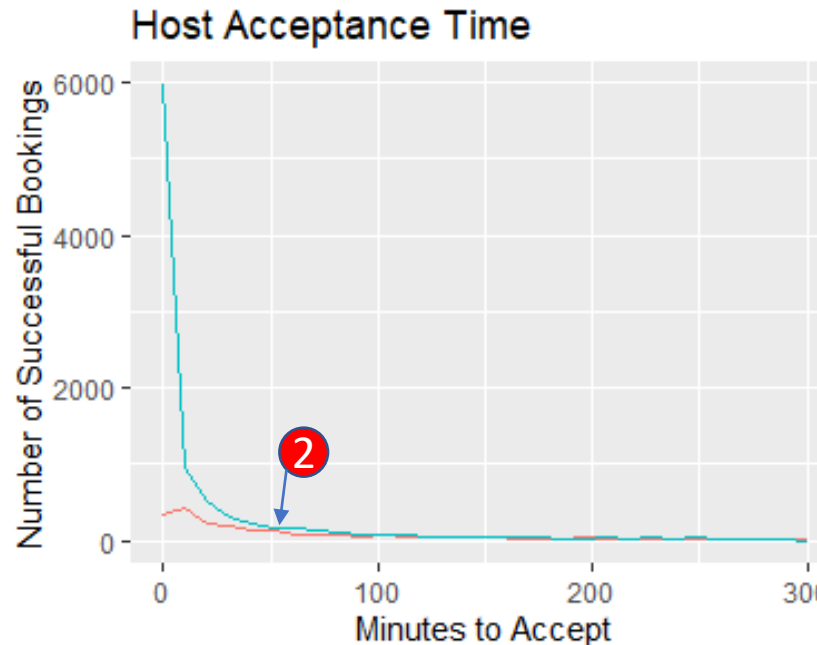


Figure 7

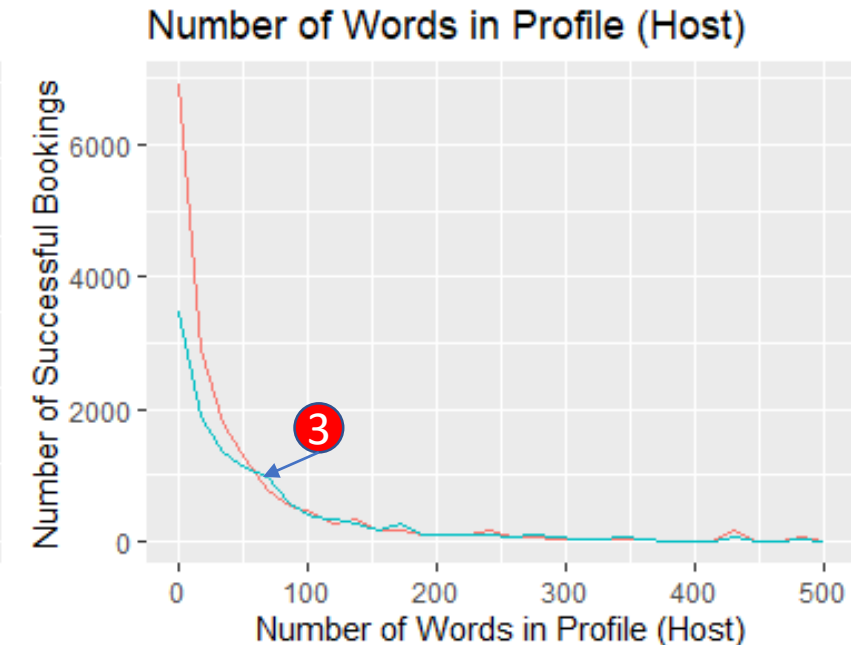


Figure 8

- [1] Hosts who reply within 30 minutes are nearly guaranteed to secure a booking. However, after 5 minutes, whether its 30 minutes or 2 hours, there isn't as much effect on the outcome (convergence of blue and red lines)
- [2] Hosts who accept guests (either `book_it` requests, or pre-approving `contact_me` requests) within 50 minutes are nearly guaranteed to secure a booking. However, after 50 minutes, whether its 2 or 3 hours, there isn't as much effect on the outcome (convergence of blue and red lines)
- [3] Hosts with at least 50 words in their profile description are much more likely to secure bookings
- **[Initiative]**: Experiment with adding following features:
 - Hosts must enter a minimum of 50 words in their profile to register a listing, and 100 words contributes towards "premium host" status. Design a template for hosts to fill out, so that it's easier for hosts to know what to write
 - Hosts can opt in to receive email or push notifications that encourage a faster response/acceptance time

How can guests help increase booking success rates (objective 2)?

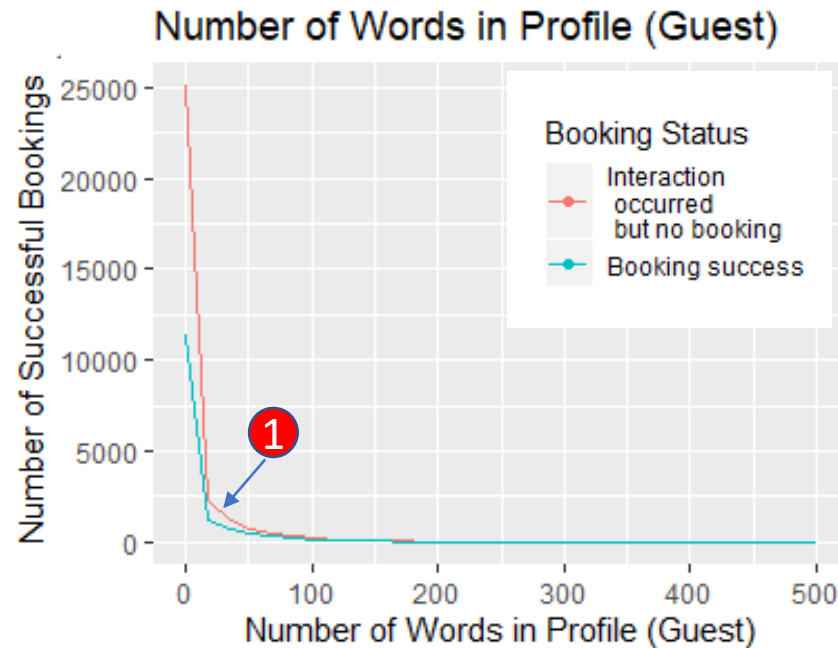


Figure 9

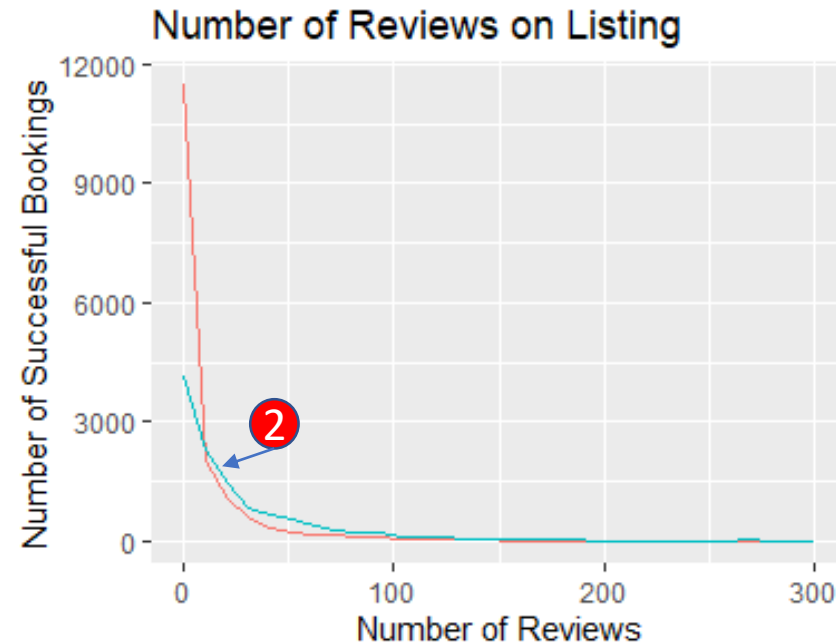


Figure 10



- [1] Guests with at least 25 words in their profile description are much more likely to secure bookings
- [2] Listings with at least 30 reviews are much more likely to secure bookings
- **[Initiative]**: Experiment with adding following features:
 - Guests must enter a minimum of 25 words in their profile to book a listing, and 50 words contributes towards “premium guest” status
 - Incentivize guests to leave reviews, making sure a follow up email is sent after a visit

What to expect

- To conclude, we should look to execute two objectives:
 - Objective 1: Increase our total interactions (red + blue lines)
 - Objective 2: Increase the success rate of our interactions (blue line / (red+blue lines))
- Things to keep track of in July 2016:
 - Number of successful bookings in guests from Brazil and US, and listings in Copacabana and Ipanema
 - Number of successful bookings across channels, number of total interactions across channels
 - Hosts' reply times, acceptance times and number of words in profiles
 - Guests' number of words in profiles, and listings' number of reviews
 - As a target guideline for how we can measure how we can measure our overall improvement, I created predictive models* (see appendix for details) to project our performance in the next month (July 2016)
 - We should expect ~**3100** total interactions in July (objective 1)
 - We should expect ~**2000** successful bookings in July, giving us a projected **64%** booking success rate (objective 2)
 - As a rough baseline, if we can improve upon 3100 total interactions and 64% booking success rate in July, we can attribute this difference to the impact our initiatives

	Total interactions (Objective 1)	Total Number of bookings	Booking success rate (Objective 2)
Baseline target to beat in July 2016	3100	2000	64%

Appendix

- R script for data analysis (if this file is not openable, code is attached below as pdf): 
- Assumptions: `rio_de_janeiro_data_analysis.txt`
 - Since `ts_interaction_first` spans Jan 1, 2016 to Jun 30, 2016, I assume July 2016 is the next month into the future as opposed to us looking back into history from present time. Other columns such as `ts_reply_at_first`, `ts_accepted_at_first`, `ts_booking_at` have very sparse data ($n < 100$) past Jun 30, 2016, thus I assume the data query pulling this dataset only conditioned on `ts_interaction_first`
 - I assume the dataset provided is the entire dataset we have on Rio de Janeiro, not a sample
- Data analysis notes:
 - One important metric missing here is prices of rooms, and revenue. This dimension of information could allow us to estimate how much budgeting we should optimally allocate for each of our initiatives
 - The dataset only spans 6 months, so it's hard to tell whether it's stationary. Most likely not as there are most likely trends over weekdays, months, holidays, quarters, and years
- Under the Listings dataset, some rows have negative `total_reviews`. There is a need to check data integrity to make sure the analysis isn't compromised – there may be a problem with either the data's storage or query
- There are no instances of rejections under `instant_book` (ie all `ts_booking_at` are populated), which doesn't seem representative of the population
- Modeling notes (slide 7): 
`model_summaries.txt`
(if this file is not openable, output is attached below code as pdf)
 - I experimented with two models, linear regression and ARIMA modeling on our dataset of 6 months, predicting onto the 7th month
 - Neither model had spectacular fit (`model_objective1` has an R^2 of 0.32 and `model_objective2` has an R^2 of .04), but I felt the served its purpose of providing a rough baseline for July. I ended up choosing linear regression over ARIMA since it had a slightly lower RMSE and visually there is a roughly linear trend on both our objectives
 - The p-values of both our models on our betas are quite high (.23 and .67), so it's hard to draw a conclusion off the beta values calibrated. However, they do agree with intuition as model 1 has a negative beta, and total interactions have been trending down since January. Model 2 has a positive beta, aligning with the fact that booking success rates have risen since January