

I rented out my apartment in Brooklyn on airbnb since 2012. Back then, there was no option called “smart pricing” on their site. I tried to look for similar listings, but couldn’t. Mine was one of the firsts that offering a dorm style airbnb, a hostel hybrid called Homestel.

I made up a number, and people chose to book it because it was a great price. I made money because there wasn’t much supplies back then.

Nowadays, things have changed. Airbnb has an option called smart pricing that it will set a price, and another option called instant booking. At first I thought it was a great idea: smart and instant. But it turned out to be a disaster.

From the host side, I kept getting these bookings that left huge gap in my calendar for a fraction of what I want to rent out. I lost a significant amount of income for 2 months that I tried out their smart pricing in 2015. Couldn’t trust the ‘smart’ pricing, I went back to old school handling the bookings myself.

Now, I also use airbnb a lot, not just for travelling, but for long term rental for work as well. I found a really good deal in Sunset, San Francisco for around \$2000/month for a 1-bedroom in-law unit. I thought that was normal in 2017. However, after 2 weeks booking, about to arrive to San Francisco, the host messaged me and said the price was wrong and asked for \$3300/month. I know San Francisco price is ridiculously, but it is in Sunset, and I never experienced a host asking for more money like this before. I have no option but let them cancel it. They claimed it was airbnb’s mistake.

Do you start to see how ‘smart’ this pricing system is?

Now, I planned an epic trip for my friends and I to be in South America for 3 months in 3 cities: Buenos Aires, Bogota and Lima, thought it would be great to learn Spanish while studying Data Science. Plus costs there is a fraction comparing to the Bay Area or New York.

I am not going to complaining about how the airbnb listing’s photos are much better in photos than in person. I will focus on their smart pricing system.

I booked airbnb for Buenos Aires and Bogota, booked the flight, counting down the days, then my airbnb host in Bogota messaged, again 2 weeks after my booking. She said she was travelling, and got home and saw that the price was wrong. Now the price on the airbnb was \$600/month for Bogota, which is normal there I thought. She asked me to pay guess what? \$100/day, \$3000/month in Bogota, Colombia. Her exact words: ‘I can give you the same rate for the first week, but then it is \$100/day after the first week. Or she will cancel it.’ I am not sure how that happened, how could the smart pricing keeps making these kinds of mistakes.

Recently I noticed, airbnb kept sending suggestions to ask me to lower my price to get my bookings, and funny thing is my place is all booked till Christmas.

I decided then that I want to tackle this. And apparently I am not the only one with the same thought.

There are companies out there such as BeyondPrice and Wheelhouse pricing software working on helping hosts to maximize their incomes. Then, there is eliotandme.com that is free to use and kind of fun.

Here is a quick screen recording for eliotandme, they use holidays, locations, size and weather to suggest the price: <https://cl.ly/7bac0d15bb5c>

BeyondPrice use holidays, and days of the weeks prediction, they show it in a breakdown over a calendar. <https://cl.ly/b151e728a01a>

Wheelhouse might be the most sophisticated with 10 billions data points they claimed. But the loading is so slow. <https://cl.ly/48da7103f32d>

Seeing how people use things such as days of the week, holidays, and weather for dynamic pricing, I am wondering how can I apply a day time system.

One of the challenge I encountered as an airbnb host is how to limit the gaps between reservations, and I haven't seen anyone approach it in automation way yet. Something to keep my mind in as I explored my data.

Just from preliminary data exploring I saw some interesting data.

1. The max of the host listing counts was over 500. Meaning someone with more than 500 listing counts are running rental business on airbnb.
2. The count of non-null for License was 0. Meaning Seattle doesn't require hosts to have a license. Rules are more relaxed here than in New York or San Francisco, which is great news for businesses such as Zeus Living.
3. The number of bedroom that is most listed is 1: I want to look into if that was a spare room, or was that an entire 1-bedroom apartment. These will give some insights to real estate as well as the sharing economy culture. Airbnb started as a sharing platform, and now people run businesses on it.
4. I do not have enough numeric data for a scatterplot, the numeric data I got are all related to price.
5. I want to make a time-series plot, but I don't have enough data with just listings.csv
6. One correlation that I see clearly is the number of bedrooms and the price.
7. A question I want to look more into is which is the most expensive neighborhood.

For presentation: I would want to tell a story, a personal story and my motivations and why I want to do this project.

If I am ambitious, I want to make a video to document all of this. But first, I want to focus on the EDA.