City Hike #3 Airbnb

Gabriel Villar, villar.g@northeastern.edu

Foreword

The urban hike I will be working on is continuing in the theme of examining the differences before and after and after the Covid pandemic. While the normal analysis I have been doing is the hosts that continue to operate during the pandemic, there could be something valuable in analyzing the listings that disappeared once the pandemic began. There could be all sorts of reasons as to why a host chose to pull out of Airbnb when the pandemic struck, but we may be able to piece together information about why they might have stopped operations.

This Urban Hike will focus on the listings that have disappeared from April 2019 to April 2020. Being able to see what happened with the listings that have disappeared may shed more light as to what drives the continued success on the hosts who chose to stay in the market. I will be conducting this urban exploration online via Google Maps as I do not live in Massachusetts.

Selection Criteria

In the spirit of maintaining consistency with my previous works, I will be dividing the data up between April 2019 and April 2020 as this is one full month after took effect as well as cheap listings. The price threshold is $60 a night which makes it comparable to the average rental in Boston. While we have learned that the primary driving factor for the effects of a price of a listing in a pandemic is the number of accommodations that can be made, or how many people can stay in a unit at a time. This makes sense as the more people are staying a unit, the more likely one of them is to contract the virus if they do not take the necessary precautions. It is also important to keep in mind, that the city of Boston implemented a law at the start of 2020 that restricts who is allowed to operate an Airbnb. So, while we may be able to find buildings that fit what we expect in a discontinued listing, we also need to be aware that they might also have ceased operations due to the law that was implemented.

But what are we looking for? We know that the price of a listing is dictated primarily by the number of accommodations and the maximum nights they are staying. Cleaning fee's and other costs on the hosts, while impactful, are less of a driving force once the pandemic hit as we have examined previously. During the walk, I will be looking for buildings that may be able to host several people at once since they will need the space to safely distance themselves while someone is visiting if they are traveling with people. Another particular factor that is affecting these listings is due to the law that was implemented. If the law was intended to oust people who held onto properties but were not living in them, we might able to find properties that are in low income areas, as those would be cheaper to buy and rent out since the maintenance cost is not as high.

To begin my hike, I started by creating a map based on the subsetted information of a listing priced less than 60 and then created an anti-join to only show the listings that disappeared between 2019 and 2020. This created a list of roughly 500 units. As reference, the total number

of units that disappeared between these years is about 4400. Once I created this list, I used the geographical coordinates from the dataset to plot these units onto a map of Boston. The map is as follows:
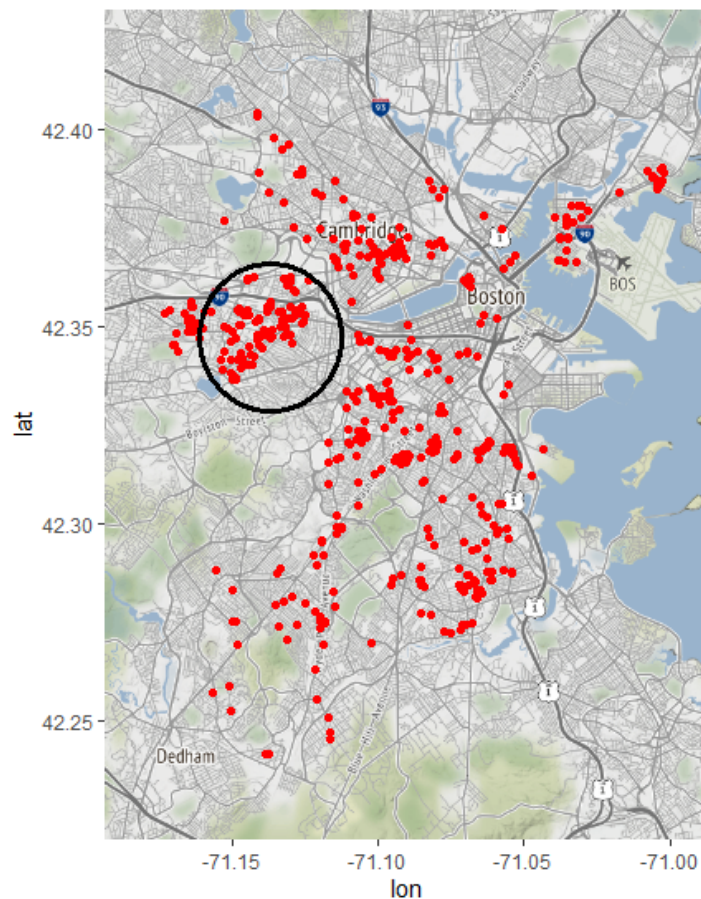
The area that is circled is where I will be exploring. Unfortunately, Airbnb does not provide addresses of listings as a privacy measure so I had no hard starting point. I will be starting from the top of the circle by the 90 freeway and working my way down the large cluster of listings.

Exploration

To begin the hike, I started with Southern Allston. What quickly caught my attention is many of the residential areas look somewhat formulaic. Many of the buildings have brick edifices and some have ivy growing on them which leads me to believe that this is an older part of Boston. However, I did stumble across a brock 4 story building with graffiti sprayed on the side of it. While I know having bars in front of a window on the ground floor is more of a safety issue, it still detracts from the overall appearance of the area which leads me to believe, along with the graffiti, that this is not a well off area (as of August 2018 Google maps).

Next, I went down to the area of Aberdeen, to the southernmost cluster of points in the circled area. This area felt to me to be the same as Allston except better off. While there were still plenty of 3 story brick buildings, they seemed to be better taken care of and there were not nearly as many bars guarding the bottom floor windows. There were also more houses strewn across the areas.

Analysis

The areas that I went exploring to me revealed more as to the area and the types of housing that were removed. We previously said that there were two possible housing types that could have

been lost, the housing that left because of Covid and the housing that left because of the new law. These areas seem to be more towards expectations I had for the listings that did not comply with the law. I decided to look into the most recent median household income of these areas and found that they were on the lower end of the metric (less than $53,000). I believe that since the median income is on the lower end, many of the properties that were held by hosts in this area, probably failed to comply with the new law rather than being concerned over the pandemic. Since these property values are low, it is likely that owners did not live here and simply rented the properties to clients while they were not there. While it is possible for hosts to rent their own properties in this area, it is unlikely that people who genuinely live here want to rent out their apartment or house to clients. So, in regards to the law that was enacted, from a cursory look, it would appear that it targeted the right areas.

Conclusion

While talking with community leaders this year, it is apparent to me that concrete actions and ideas are critical moving forward. It is very easy to get lost in the data and theories and show that an event should occur but whether or not it actually is happening requires some foot work. This city walk demonstrated that knowing information about an area, such as median income, is important to making policies that have an impact where they are needed. It reaffirms the notion that you need to know the areas you are working with in order to truly appreciate or gauge the impact on the policies you want to implement. It also helps understand what the data is already saying. As a native San Diegan, I do not know the areas of Boston well enough to visualize immediately what the data is telling me. While the cluster of data from the Allston area might be obvious to Bostonians, it is not as apparent to me so I went to explore it.

City Walk 3
Larissa Morikawa, Morikawa.l@northeastern.edu

Somewhat continuing my last city exploration, I focused on neighborhoods with outstanding foreign host listings. Last time, I explored Allston, since it was the neighborhood with the highest number of foreign hosts. After that assignment I realized that this does not necessarily mean highest concentration, so this time I made sure to account for that by calculating the proportion through 1 and 0 values defining domestic and foreign hosts, and later taking the proportion of domestic hosts within each neighborhood with the aggregate function. The following shows the results.

neighborhood	host_location_status
<chr>	<dbl>
Allston	0.9245847
Back Bay	0.9890494
Bay Village	0.9837518
Beacon Hill	0.9865712
Brighton	0.9820764
Charlestown	0.9883901
Chinatown	0.9932635
Dorchester	0.9756221
Downtown	0.9944084
East Boston	0.9941781
Fenway	0.9716544
Hyde Park	0.9807692
Jamaica Plain	0.9818383
Leather District	1.0000000
Longwood	1.0000000
Mattapan	1.0000000
Mission Hill	0.9862102
North End	0.9863567
Roslindale	1.0000000
Roxbury	0.9927374
South Boston	0.9957558
South Boston Waterfront	0.9940120
South End	0.9913110
West End	0.9872068
West Roxbury	0.9968847

Proportions of domestic hosts in Boston neighborhoods

I took the liberty of pointing out the two neighborhoods that had the highest proportion of foreign hosts. It seems like my previous worries were unnecessary since Allston still shows the highest proportion of foreign hosts by far. The second neighborhood with the highest percentage of foreign hosts was Fenway; therefore, for this city exploration I would like to explore the Fenway area.

First, I looked into the foreign-born population in Fenway. According to the 2016 Boston Planning and Development Agency report, 26% of the neighborhood were foreign born. This is more than the percentage of foreign Airbnb hosts in Fenway, but it does make sense. I wouldn't expect a lot of residents to be Airbnb hosts, especially in a high demand location for permanent/long term housing.

http://www.bostonplans.org/getattachment/3e8bfacf-27c1-4b55-adee-29c5d79f4a38

Speaking of high demand location, I thought it would be interesting to compare monthly Airbnb prices to average monthly rental costs in Fenway. This is a rather personal exploration for me since this location is one of the best locations to be living as a Northeastern student due to its close proximity to campus, green space, and shopping district. A lot of times students cannot afford to pay rent in Fenway even though they want to live there due to the high cost, so maybe Airbnb is the solution?
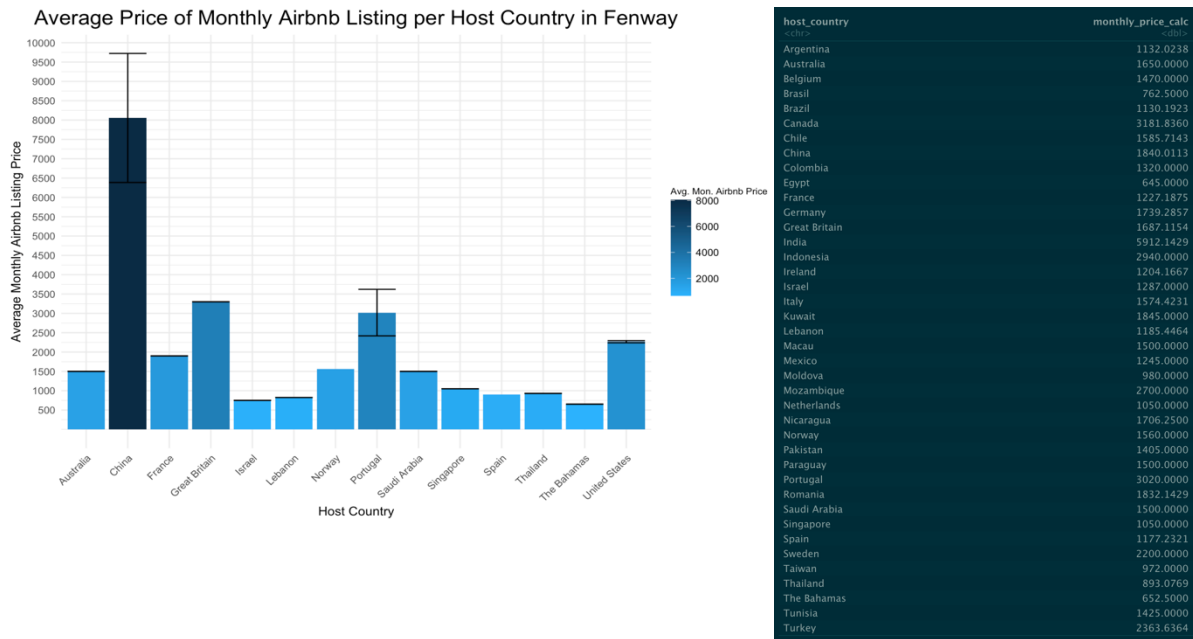
```
        One Sample t-test

data:  fenway$monthly_price_calc
t = 2.1342, df = 894, p-value = 0.0331
alternative hypothesis: true mean is not equal to 3185
95 percent confidence interval:
 3245.760 4635.826
sample estimates:
mean of x
 3940.793
```
T-test

T value was 2.1342 producing a p value of 0.0331; the p value indicates that there is a significant difference between the average monthly Fenway Airbnb listing prices and the average monthly rent in Boston. More specifically, Fenway Airbnb average monthly rental prices tend to be around 2 dollars more expensive than the average monthly rent in Boston. This finding is interesting if we consider the fact that recently, Airbnb has become a common type of long-term leases. In the last few years, hosts have been using Airbnb as a platform through which they list out their listings long-term at a competitive price. Furthermore, we have seen more and more hosts acquire rental licenses to be able to rent out their listings long-term (28 days+). So, although there are listings that are able to (or at least seemingly) compete pricewise with the general home rental market. This t test suggests that it might not actually make much of a different to rent long-term through Airbnb; therefore, Airbnb may not be a good solution for affordable housing for students in the Fenway area.

Moving on, I wanted to see if I could find a pattern in price based on host country (I will be referring to the host location's country as host country). I created a graph to visualize this analysis.

Average Price of Monthly Airbnb Listing per Host Country in Fenway

host_country	monthly_price_calc
<chr>	<dbl>
Argentina	1132.0238
Australia	1650.0000
Belgium	1470.0000
Brasil	762.5000
Brazil	1130.1923
Canada	3181.8360
Chile	1585.7143
China	1840.0113
Colombia	1320.0000
Egypt	645.0000
France	1227.1875
Germany	1739.2857
Great Britain	1687.1154
India	5912.1429
Indonesia	2940.0000
Ireland	1204.1667
Israel	1287.0000
Italy	1574.4231
Kuwait	1845.0000
Lebanon	1185.4464
Macau	1500.0000
Mexico	1245.0000
Moldova	980.0000
Mozambique	2700.0000
Netherlands	1050.0000
Nicaragua	1706.2500
Norway	1560.0000
Pakistan	1405.0000
Paraguay	1500.0000
Portugal	3020.0000
Romania	1832.1429
Saudi Arabia	1500.0000
Singapore	1050.0000
Spain	1177.2321
Sweden	2200.0000
Taiwan	972.0000
Thailand	893.0769
The Bahamas	652.5000
Tunisia	1425.0000
Turkey	2363.6364

According to this graph, it seems like on average, hosts in China have higher average monthly price listings at around $8053 per month, followed by Great Britain at around $3300, and then Portugal with a similar average monthly price at around $3020. These prices were adjusted for the number of people each listing can house; therefore, size should not be one of the factors affecting this large gap. There are other things however, that could explain this discrepancy. First, could be the amount of cases that have hosts in China compared to the other countries. Another explanation could also be that there is an outlier in the data. Previously we had seen multiple listings that listed extremely high prices on purpose without intention of actually renting it out. Putting a higher price on the listing is a way to keep the listing online but without having anyone renting it out, one could call these fake listings because they are not actually available.

Running a quick, simple regression to see if host country affects monthly pricing, and the model showed that host country has a slight predictive effect on the monthly average listing price. It seems like listings that has host country listed as China, one can expect around a $6500 dollar increase in monthly price. This may well be the result of outliers in the data, but I also included listings with host countries in United states that also holds a lot of outliers and has many more listings.

```
Call:
lm(formula = monthly_price_calc ~ host_country, data = .)

Residuals:
    Min     1Q  Median     3Q    Max
  -7244   -857    -317    433  44233

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               1.500e+03  7.928e+02   1.892   0.0586 .
host_countryChina         6.554e+03  8.547e+02   7.668 2.23e-14 ***
host_countryFrance        4.000e+02  1.121e+03   0.357   0.7213
host_countryGreat Britain 1.800e+03  1.586e+03   1.135   0.2564
host_countryIsrael       -7.500e+02  1.121e+03  -0.669   0.5036
host_countryLebanon      -6.750e+02  1.003e+03  -0.673   0.5009
host_countryNorway        6.000e+01  2.098e+03   0.029   0.9772
host_countryPortugal      1.520e+03  1.121e+03   1.356   0.1753
host_countrySaudi Arabia -1.518e-10  1.373e+03   0.000   1.0000
host_countrySingapore    -4.500e+02  1.586e+03  -0.284   0.7766
host_countrySpain        -6.000e+02  2.098e+03  -0.286   0.7749
host_countryThailand     -5.700e+02  9.710e+02  -0.587   0.5572
host_countryThe Bahamas  -8.475e+02  9.710e+02  -0.873   0.3828
host_countryUnited States 7.671e+02  7.935e+02   0.967   0.3337
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1942 on 3655 degrees of freedom
Multiple R-squared:  0.08843,   Adjusted R-squared:  0.08519
F-statistic: 27.28 on 13 and 3655 DF,  p-value: < 2.2e-16
```

After further looking into the listing, I noticed that this extremely expensive listing seemed real. It had most of the components filled out with a very thorough description. The price per night was $750 which explains the insane calculated price. It is also worth noting that this listing does not have a license and consequently is not available as long-term housing. That being said, I was curious to see what this listing looked like. To my surprise, this location was just around the corner from my own apartment and it is the coordinates to apartments in The Parkside apartment complex above the corner store. These apartments are newly renovated, luxury apartments, so in a sense the higher price range would make sense, but not as high as it states in the dataset. I looked into the average monthly price for an apartment in this complex and it is around $1500 – $4300 a month, much less than the Airbnb price. I wonder what prompted the host to list it at such a high price. My best guess is still that they probably had listed their apartment at a certain time, but then when they decided to take it off of the market, they just changed the price to a really high one in hopes that no one would rent it and if someone did, the host would get a large profit.
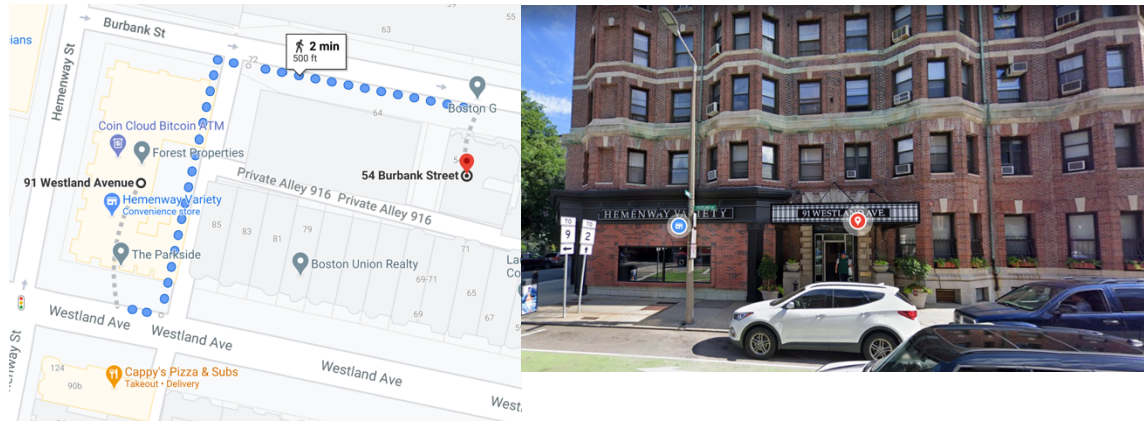
Additionally, I looked at listings that have the host in countries with the lowest average monthly price. The listings with the lowest average monthly prices in this analysis are those whose hosts are in The Bahamas. This is an apartment on Park Drive right next to the Fenway T entrance and the Fenway bus stop. It is also very close to 3 college campuses, Emerson College, Northeastern University and Boston University. Furthermore, it is only a few feet away from Boylston Street that has a lot of stores, restaurants, cafes and bars. This apartment accommodates 4 people and it is listed at $87 a night, which would amount to $21.75 per person. Considering its location, and price, I think that this is an amazing listing. I also looked at this listing's maximum nights and its 1124 nights. This is a little over 3 months, which is basically 2+ years making me think that the host probably doesn't have a true maximum. Thinking that this would be great student housing, I also looked to see if they had a license; they did not. It is hard to say whether or not this listing would be available for long term lease, even if under the table, nonetheless, it is seems like a very good listing.

While I was exploring my dataset in the beginning, I found something interesting that I thought I'd include at the end of my city exploration as a bonus. Because the Airbnb dataset that I have only shows listings in the Greater Boston area, I assumed that most of the variables would be US based. In other words, I thought that the extremes would all point towards truly domestic cases, including host location; I was proven wrong when I took the averages of days as host by host location (host country). I found that in Boston, on average, international hosts (hosts have do not have United States as their host location) have been hosts for longer. The top three average days as host is Great Britain (3036 days / 8.3 years), Chile (2899 days / 7.9 years), and Macau (2712 days / 7.4 years).

host_country	hostdays	host_country	hostdays
<chr>	<dbl>	<chr>	<dbl>
Argentina	1895.0000	Pakistan	1558.4000
Australia	2132.8571	Paraguay	1434.0000
Belgium	2535.8000	Portugal	2450.0000
Brasil	2115.0000	Romania	2285.2857
Brazil	2578.3846	Saudi Arabia	1538.0000
Canada	2142.1972	Singapore	1237.0000
Chile	2899.7143	Spain	972.1786
China	1824.0136	Sweden	743.0000
Colombia	2116.0000	Taiwan	1669.2857
Egypt	756.0000	Thailand	2018.0000
France	2252.5938	The Bahamas	1167.0000
Germany	2568.4286	Tunisia	255.0000
Great Britain	3036.9615	Turkey	2729.1818
India	1338.1429	United Kingdom	2007.3077
Indonesia	289.0000	United States	1986.1409
Ireland	1446.6667	Uruguay	2350.0000
Israel	2554.7000	Venezuela	851.0000
Italy	2430.4615	Vietnam	1971.5000
Kuwait	1699.5000		
Lebanon	1359.5893		
Macau	2712.0000		
Mexico	1677.6667		
Moldova	1943.0000		
Mozambique	1658.0000		
Netherlands	2071.0000		
Nicaragua	1756.0000		
Norway	1926.0000		

Even when narrowing this down to my specific neighborhood, Fenway, for this assignment, the host country that had the highest average host days was France with 2666 days which amounts to around 7.4 years, followed by Portugal, Thailand, and then the US. I think it would be interesting to know what a host considers when they answer the host_location prompt on Airbnb. I noticed that there were some cases in the data set that said, "Boston, from Jamaica". I am assuming that this person meant that they are based in Boston but are originally from Jamaica. So what if other hosts also had this though and just wrote where they were originally from when they were actually based on Boston, which would create misleading data. Host_location is originally supposed to state where the host of the listing is based out of, not where they are actually from, but it can be a rather confusing variable; thus making this a rather interesting phenomenon to look into. Who knows? Maybe we do just have an abundance of international hosts that have been listing their open properties/vacation properties/investment properties up for leasing while they reside in a different country!

	host_country	hostdays
1	Australia	1301.000
2	China	1678.541
3	France	2666.000
4	Great Britain	594.000
5	Israel	1591.000
6	Lebanon	332.000
7	Norway	1926.000
8	Portugal	2450.000
9	Saudi Arabia	1538.000
10	Singapore	1237.000
11	Spain	632.000
12	Thailand	2068.000
13	The Bahamas	1167.000
14	United States	1983.956