Investigation of Mega Hosts in NYC: How to handle 50+ listings on Airbnb?

Xinyu Xiong,[1] Feng,[2] n Yang,[3] and Y[illegible]an Deng[4]

[1] *UC Davis Statistics Department,*

[2] *UC Davis Statistics Department,*

[3] *UC Davis Statistics Department,*

[4] *UC Davis Statistics Department,*

Abstract:

Airbnb, a platform providing an alternative way of accommodation, has been explored for its business model in much research. A fast-growing group of hosts who are professional multi-operators - also called mega hosts - has again changed the pattern of Airbnb rental business, owning an extraordinarily large number of listings and beginning to supersede personalized and unique "home" by hotel-like rooms on the Airbnb market. Many researchers have recognized this problem, while most of which were tracking some typical mega hosts who earn large profits to figure out how they run their business, little research has focused on the overall performance of mega hosts. The focus of this paper would be an exploration of New York City mega hosts' behaviors, including their choices of listings' locations, naming strategies, pricing, and other basic settings of the listings. As a result, it has been concluded that mega hosts have specific preferences in the above aspects, which somehow leads to the difference in the number of reviews between mega hosts and normal hosts.

I. Introduction

Airbnb has been regarded as an alternative way for travelers to accommodate during their trips than living in hotel rooms or resorts. While the idea behind this new pattern of accommodation is to give travelers a chance to live as a local, some hosts list a large number of hotel-like properties on the Airbnb platform and manage them through specific operations in order to make profits[1]. Those hosts are usually called "mega hosts." If they manage properties properly, they will receive an enormous amount of profits each year. In this case, Airbnb has been used as a commercial hotel booking site to get standard and nondistinctive rooms instead of as an intermediary to connect with local hosts.

We can get to know about the three main types of hosts on Airbnb: bed-and-breakfast hosts, "feel like home" Hosts, and mega hosts[2]. The biggest difference between mega hosts and other hosts is that mega hosts should have multiple listings and certain intelligence in investments. In other words, mega hosts should have the ability to make their listings popular among customers' choices, and they might be the most "professional" hosts on Airbnb, a lot of examples can help verify this. According to a report from the New York state attorney general's office, by 2014 there were more than 100 hosts with at least ten properties in New York state. The annual revenue of a particular host who owns 272 unique Airbnb listings is about $7 million, without paying taxes and costs that a hotel operator is supposed to pay[3]. Though there is a lot of news about the increasing

[1] R. Lorena. "New "mega hosts" help level up Airbnb's brand and traveler experience." (2017, June 30). Retrieved from: https://blog.turnkeyvr.com/new-mega-hosts-help-level-airbnbs-brand-traveler-experience/.
[2] BnB Hosts Team. "The 3 Main Airbnb Host Categories: Which One Are You?"(2020, April 21). Retrieved from: https://www.bnbhosts.com.au/3-airbnb-host-categories/.
[3] M. Chris. "NY Attorney General: 72% Of Airbnb Rentals Violate State, City Laws. Consumerist." (2016, September 27).Retrieved from: https://consumerist.com/2014/10/16/ny-attorney-general-72-of-airbnb-rentals-violate-state-city-laws/index.html.

trends of mega hosts[4] entering Airbnb, few data analysis has been established to discover how mega hosts are operating their properties differently by selecting, advertising, pricing, and listing their properties to make maximum profits. Hence, in the following content, we will discover and analyze behavior and performance of mega hosts.

II. Related Works

Ameya Deshmukh, with his thorough investigation on the effect of super hosts on the Airbnb market, has provided a formula of occupancy rate which is an explicit measurement of the profitability of each host.[5] The formula states that the occupancy rate can be increased by the average length of accommodation and the number of reviews per month, but have a negative relationship with the percentage review rate. The former two can be connected with variables *minimum_nights* and *number_of_reviews* provided in the dataset to have a rough idea of profitability, though we don't have the exact variable for calculation. The percentage review rate, denoted as the percentage of guests who write reviews among the total number of guests who have visited the listings, may have some variation between mega hosts guests and normal hosts guests since we found that they are facing different groups of guests.

Yoon Koh, Jaewook Kim, and Yue Vaughan[6] have explored the effect of listing names on their financial performance. As a linguistic experiment, they have found that there are specific words which can provide satisfaction to customers. Mega hosts may use the 50-word limited title as a useful tool for advertisement, generating formulaic patterns of titles which can attract more

[4] "From Air Mattresses to Unregulated Business: An Analysis Of The Other Side Of Airbnb." (A Comprehensive Survey of 14 Metropolitan Statistical Areas , 2016).Retrieved from: https://www.ahla.com/sites/default/files/Airbnb_Analysis_September_2016_0.pdf

[5] D. Ameya. "The Effect Of Superhost Status On Airbnb In Berlin Using Occupancy Rate And Revenue Per Available Listing."(2020). Retrieved from: https://doi.org/10.13140/RG.2.2.16576.56325

[6] K. Yoon , K. Jaewook , and T. V. Yue " How you name your Airbnb's title matters: comparison of seven countries." (2021).Journal of Travel & Tourism Marketing, 38(1), 93–106. Retrieved from: https://doi.org/10.1080/10548408.2021.1875105

customers. Interestingly, some words which have specific positive meanings may turn out to have a negative or uncertain impact on financial performance, such as "family", "social", and "friendly". Some words like "drive" and "kitchen", however, can have significant positive impacts.

Mann-Whitney test has been applied by Ameya Deshmukh to test for occupation rate differences between mega hosts and normal hosts by neighborhoods. He uses the Kurtosis and Skewness of data to check the assumption of non-normality and found that the revenues per available listings of mega hosts are significantly higher than that of normal hosts.[5] Hence, we can use the Mann-Whitney test to have a preview of differences between mega hosts and normal hosts and try to dig deeper for more information by plotting and analyzing.

Standing at the view of mega hosts, another worth considering problems would be location, since properties in different neighborhoods may have different initial values. According to data from furmancenter.org, even though the rental price in NYC may not have a large variance across different neighborhood groups, we can still notice that the median rental price at Stuyvesant Town, Manhattan ($2,490 per month) is more than twice as many as the median price at Mott Haven, Bronx ($940 per month) in 2018.[7] It indicates that mega hosts in different areas may have different benchmarks of pricing their houses to meet the largest group of target consumers. Also, it reminds us to exclude the effect of different neighborhood groups on listings prices when comparing mega hosts and normal hosts.

Considering the consumer's budget is also important for mega hosts in pricing, since properties in different neighborhood groups may have different target tenants. The paper "Why people choose

[5] D. Ameya. "The Effect Of Superhost Status On Airbnb In Berlin Using Occupancy Rate And Revenue Per Available Listing."(2020). Retrieved from: https://doi.org/10.13140/RG.2.2.16576.56325

[7] "Greenpoint/Williamsburg Neighborhood Profile." (NYU Furman Center) Retrieved from: https://furmancenter.org/neighborhoods/view/greenpoint-williamsburg.

Airbnb over Hotel?" by Yujia Chen and Markus Schuckert[8] shows that the cheaper price and trip budget are two main reasons that people choose Airbnb instead of hotel. Hence there might be some information that we can discover from the pricing of mega hosts. For example, their prices might be more standardized and will have less extremely large values.

These works shed light on the potential preferences of mega hosts in many parts of their businesses which will be considered in our method section. Also, to have a better understanding of the location differences that may determine mega hosts' behavior, we have added data related to crime rates and local rental prices in consideration. More details of data will be introduced in the next section.

III. Data

The original data that we have for this project is "Airbnb_NYC_2019", which includes information about every host in NYC. The description of variables is shown below:

Variable names	Description
id	**Listing ID**
name	**Listing names**
host_id	**Host ID**
host_name	**Host names**
neighbourhood_group	**Neighborhood group**
neighbourhood	**Neighbourhood**
latitude	**Latitude of the house**
longitude	**Longitude of the house**
room_type	**Private/Entire/Shared Room**
price	**Prices of listings**

[8] C. Yujia , and S. Markus. "Why people choose Airhnb over Hotel?" (2016). Retrived from: https://www.researchgate.net/publication/305225839_Why_people_choose_Airbnb_over_Hotel

minimum_nights	**Minimum nights required**
number_of_reviews	**Number of reviews**
last_review	**Date of the last review**
reviews_per_month	**Reviews per month**
calculated_host_listings_count	**Number of listings of the host**
availability_365	**Number of days when listing is available for booking**

The raw data contains some columns with more than 10,000 invalid entries, so we firstly drop two columns from the dataset, "last_review" and "reviews_per_month," which contain a great number of NA values. Usually, reviews per month give us a reference of the popularity level (sometimes a dislike level, but it would be very unlikely for a terrible room to receive a lot of bad reviews since the Airbnb algorithm always shows the best possible listings. Hence those with bad reviews may not have many opportunities of being chosen) of a particular listing, but here we have another variable "number_of_reviews" to show the popularity in an accumulative period of time. After dropping two columns, we still have some NA values for "name" and "host_name," so we drop the remaining NA values. We have 48,858 listings after preprocessing the data.

Investigating the existence of mega hosts in this dataset, we found that there are 514 out of 37,425 hosts with listing numbers that are equal or greater than 5, hosting a total amount of 5,640 out of 48,858 listings. If we define those who own more than 5 listings as mega hosts, then the average number of listings each mega host owns would be more than 10; whereas the rest, ordinary hosts, own 1.17 listings for each on average. So far we can tell that mega hosts do exist in this dataset. Indeed, some people may have more than 5 houses without an investment purpose, and hence the threshold for identifying a mega host is obscure. We can still roughly distinguish the group

of mega hosts and normal hosts, however, since the two groups have very different features which will be discussed in the analysis section.

IV. Methods

While the ultimate goal for each mega host is maximizing profits, every single factor may affect one's decision. Therefore, we plan to make a broad investigation of our data to find these factors.

Firstly, we need to know how mega hosts get their properties. There are two possible ways: by managing other's properties as agents, or by renting a group of properties in a certain area. Given the latitude and longitude of each listing, we can create a map to show the locations of the listings held by some typical mega hosts. If the listings are rented by mega hosts, then we are expected to notice many of the listings are aggregated together, since there are cases in which timeshare owners rent out more than 1,000 listings in hotels and resort chains for Airbnb business[9]. The next thing that we are going to discover is how mega hosts select locations of properties. There is a sharp contrast across different neighborhood groups - Manhattan has a number of mega listings which is much larger than other neighborhood groups. We can find more specific details in location differences to investigate the particular neighborhood groups that mega hosts prefer. Also, the median prices of house rents and crime rates vary from different neighborhood groups, which may affect mega hosts' decisions on selecting properties.

The next step is to discover how mega hosts name their listings, as there is evidence that the usage of some words may have a positive relationship with the financial performance of listings.[6] There are a lot of formulas for Airbnb titles which tell hosts to get attractive names. These formulas

[6] K. Yoon , K. Jaewook , and T. V. Yue " How you name your Airbnb's title matters: comparison of seven countries." (2021).Journal of Travel & Tourism Marketing, 38(1), 93–106. Retrieved from: https://doi.org/10.1080/10548408.2021.1875105

[9] K. Ashlee. "Some Airbnb 'Mega Hosts' Are Renting Out More Than 1,000 Properties At Once. Consumerist." (2017, June 9) Retrieved from: https://consumerist.com/2017/06/09/some-airbnb-mega-hosts-are-renting-out-more-than-1000-properties-at-once/.

are not only composed of fancy words like "perfect" and "comfortable", but also contain necessary information which can inform guests about the convenience and features of houses in a concise way. For example, showing a close distance to a landmark or famous site is a way to attract interested travelers, such as "Eiffel Tour — 10 Min Walk" or "Oceanside View".[10] To discover the features of mega hosts naming formula, we are going to find some information from the distribution of word uses.

Besides, we are interested in the level of availability and minimum night for listings from mega hosts. Usually, we will consider the number of listings as the one's "market share" on Airbnb, and the number of reviews as the popularity level of listings. However, it is very possible that listings held by normal hosts are only available for 30 days in a year and have the minimum night of only one night; whereas mega hosts, who own a large number of empty properties in a whole year, may have hundreds of their listings with 365-day availability and set a minimum night to be 10 days. In this case, we can't help doubting that mega hosts may have a market share much larger than the proportion of their listings; and the number of reviews of mega hosts listings may not exceed that of normal hosts because of their difference in minimum night requirement. Hence, the analysis of this section is to give us some references for a better understanding of the popularity level.

Price is a factor that will directly affect the demands. We will analyze whether the price set by mega hosts will be lower than the average price of normal listings so that they can be ahead of the demand curve because of their cheaper prices; or whether they have higher prices so that they can guarantee their income from each listing. We may also try to figure out the reason behind their pricing strategy and its outcomes. For example, the price may be largely dependent on the room type, since the price of an entire house is usually higher than a single room, in common sense.

[10] "Writing Catchy Airbnb Titles: Proven Formulas That Attract 5x More Bookings. Medium."(iGMS, 2019, June 25). Retrieved from: https://medium.com/@airgms/writing-catchy-airbnb-titles-bc7ba0f3a0ee.

The final step is to make a contrast of the number of reviews between the two types of hosts. With all of the information above, we can have a general idea of how mega hosts run their business and make profits.

V. Analysis

1. *How mega hosts get their properties*

First of all, we want to know how mega hosts get their properties. As we stated upon, there are two possible ways - managing other's properties or renting a group of properties. To show more accurate results, we studied the behaviors of the top five largest mega hosts and tried to find the patterns of their properties over the whole population. The five graphs are listed in Figure 1 from the largest mega host to the fifth largest host. From the graphs, we can see many of their listings are aggregated together. It provides us an idea that these listings might be rented by mega hosts since their locations would be more spread out if mega hosts are managing others' houses. It is very unlikely to find owners of more than 10 properties who are looking for agents in such a small range of areas. Besides, as mentioned in the method section, we have noticed examples of timeshare owners.[9] So we can predict that mega hosts get their properties by renting a group of properties.

The next thing we would like to explore is where the mega hosts rent properties, depending on the features of neighborhood groups. Notice that here we are using Tableau to draw a map for better visualizations. In Figures 2 and 3, we compared the location and price distribution of both population and mega hosts. The graph is a scatter plot with a background of the NYC map, with each scatter point representing a listing. The locations of scatter points are automatically adjusted by latitudes and longitudes, and the price of each listing is represented by the size of each point. From Figure 2, we can tell that the overall price of listings in Manhattan and some areas of Brooklyn are

much higher than others. For mega hosts, this trend is more extreme. In Figure 3, we can tell that the majority of mega hosts are located in Manhattan and Brooklyn, and their prices are significantly higher than that of the remaining three. It indicates that mega hosts are willing to choose locations to rent out their listings with higher prices, so that they may have higher revenues. Of course, considering the economic relationship between price and demands, mega hosts' aggregation at these two neighborhood groups can't be captured only by the price since higher prices usually cause less demand. Other factors like population densities, GDPs, and more features of neighborhood groups also influence mega hosts' decisions on location.

For now, we had a rough picture of mega hosts' preparation works for their businesses. The next step is to have a deeper look at their operating strategies. Mega hosts with more listings may have a larger-scale business, and hence are more representative and profitable. In the following analysis, we contrasted between the top mega hosts with more than 50 listings (the top 11) and normal hosts who are the rest. It shows that the top eleven mega hosts have numbers of listings ranging from 327 to 52 (Table I). Different from normal hosts who have properties distributed all around New York City in the Bronx, Brooklyn, Manhattan, Queens, and Staten Island, the top eleven mega hosts have only listings distributed in three neighbourhood groups in NYC -- Brooklyn, Manhattan, and Queens.

2. Mega hosts's naming strategy

To study the difference of listing names between normal hosts and mega hosts, we find the top 25 words that appear most frequently in listing names for normal listings and mega listings (Figure 4 and 5). Among mega listings, we could find that symbols like "|" and "+" and abbreviations like "1br" and "w/" are frequently used, while for normal listings, we could hardly find symbols and abbreviations in the top 25 words. It is known that Airbnb gives hosts up to 50 characters for listing

names, so symbols and abbreviations are used to save up characters. Clearly, mega hosts are more familiar with how to save up characters. Another difference is that descriptive words like "cozy," "spacious," and "sunny" appear frequently in normal listings, and mega hosts do not use those descriptive words so often. This could prove that normal listings are more personalized and could provide a more unique experience.

In addition, we also see some specific words in mega listings like "sounder" and "blueground," through checking the first 10 listings of the top 2 mega hosts, we find that those are the mega hosts' names of listings. With 327 listings of Sounder and 232 listings of Blueground, such host names are more like brands. Those mega hosts marked listing names with their names to promise the quality of the listings, and their listings are more like commercial products of standardized management.

3. How do mega hosts set "minimum nights"?

To check the non-normality of the minimum night of normal and mega listings, we calculated that the kurtosis for normal and mega listings' minimum nights are 860.32 and 27.84, and the skewness are 21.92 and 1.98, which proves non-normality here (normal distribution's kurt should be 3 and skewness should be 0). With the assumption of non-normality and independent observations of data, we conducted a simple Mann-Whitney test to see if there are significant differences between mega hosts and normal hosts in the minimum night. The result shows that there is a big difference between the two types of hosts (statistics=8307742.0, p-value= 0.0). From the boxplots of the minimum night set by normal hosts and mega hosts (Figure 6), we could observe that the minimum night set by mega hosts has smaller variance and fewer outliers than the minimum night set by normal hosts, indicating that those top mega hosts have similar preference in setting minimum night. As we observed a great number of large value outliers in the boxplot, we compare the median of the minimum night set by mega hosts and normal hosts. The median of mega hosts is 30 days,

while the median of normal hosts is 2 days. Thus, mega hosts generally target the group of guests who could stay for a longer time.

Since mega listings are only located in three out of five neighborhood groups, it is also necessary to exclude the effect of neighborhood group differences on minimum night settings. From the boxplots of the minimum night of normal listings in different neighborhood groups (Figure 7), we could get little information as there are too many outliers. In this way, we calculated the five-number summary (Table II), and it shows that there is no big difference in the minimum night set by normal hosts among five neighborhood groups. Most minimum nights required by normal hosts are approximately within the range of one to six days, but there are large value outliers above hundred days in all neighborhood groups. From the boxplots of mega listings (Figure 8), we could observe that the minimum night's medians are similar among three neighborhood groups, and there are some outliers in Manhattan. Then we checked the five-number summary for mega hosts (Table III), and we surprisingly found that the minimum nights of mega listings in Brooklyn and Queens are all 30 days as there is no variation. The minimum nights of mega listings in Manhattan also centered about 30 days although there are outliers, and the variation might be due to more mega listings in Manhattan. In this way, the minimum nights required by mega hosts are also similar among three neighborhood groups, which are around 30 days. We could prove that minimum nights will not be influenced by the different neighborhood groups, and mega hosts generally require larger minimum nights than normal hosts.

4. How do mega hosts set "availability"

Similarly, we check the non-normality of the availability first, and the kurtosis for normal and mega listings' availability are -0.86 and 2.09, while the skewness are 0.83 and -1.59, which shows non-normality here. We conducted a Mann-Whitney Test to have a preview of availability

difference, which shows a significant result (statistic=10419674.0, p-value=0.0). The median of total available days of mega listings is 318.5 and the median of normal listings is 38. From the boxplots of availability of normal listings and mega listings (Figure 9), we could clearly see that mega listings generally have more available days than normal listings. For mega hosts, more than half of properties have availability between 300 and 350 days, and more than three quarters of houses and apartments have availability larger than 200 days, but there are still some outliers which are below 100 days. The availability of mega listings meets our expectation as mega hosts are supposed to rent a lot of properties for commercial purposes on Airbnb, and the properties are empty and available for use at most times. In contrast, for a half of normal listings, the properties might not be empty for most of the year as the median is only 38. However, the distribution of normal listings' availability is highly skewed according to Figure 9. It indicates that there also exist some normal hosts who own empty properties with a very high availability (the 75th percentile is above 200).

With great difference in availability of listings, it also shows that the market share of mega hosts might be many times larger than their portion of listing quantities on the entire Airbnb market. This would be very considerable in the next few years, according to data in Penn state, the amount of mega hosts has increased "from 1,171 in September 2014 to 2,193 in September 2015, an 87.3% increase".[11] In this way, the market share of mega hosts grows fast on the Airbnb.

Also, the distributions of available days in different neighborhood groups are analyzed to see the association between availability and neighborhood groups. From the boxplots of normal listings' availability in different neighborhood groups (Figure 10), we could see that Staten Island and the Bronx have relatively high availability, while Brooklyn and Manhattan have relatively low availability. The data of 2012 NYC Boroughs GDP in Wikipedia may help explain this trend. Among the five

[11] E. Christopher. "Airbnb Runs 'Illegal Hotels,' Hotel Industry Study Claims" (2016). Retrieved from: https://fortune.com/2016/01/20/airbnb-illegal-hotels-study/

neighborhood groups, Staten Island and the Bronx are the two with lower GDP. Different from the other three neighborhood groups which have "a thriving hub of entrepreneurship and high technology startup firms," Staten Island and the Bronx are mainly famous for "the world's largest metropolitan zoo" and "the last undisturbed forests in the city."[12] Therefore, people who work in NYC are more likely to live in Manhattan, Brooklyn, and Queens, which are close to their working places. Hence, they leave their properties in Staten Island and Bronx empty, which helps explain why normal listings have higher availability in the two neighborhood groups. In addition, this also helps explain why mega hosts do not run their business in Staten Island and the Bronx. From the boxplots of mega listings' availability (Figure 11), it shows that mega listings in Brooklyn and Manhattan have relatively higher availability than mega listings in Queens. However, the difference is not that big, and mega listings have high availability in all three neighborhood groups. In this way, we can conclude that mega hosts have their listings with exceptionally more available days than normal hosts on the Airbnb market. This could be an advantage of mega hosts, as we mentioned, their large market share will allow their listings to be chosen by guests with higher probabilities, especially when customers' demand is high, and the supply of normal listings is low.

5. Which room type mega hosts prefer?

Room type is also one of the important factors in mega hosts' operation. From the histogram of room type of normal listings (Figure 12), we could see that the entire homes and private rooms dominate the normal listings, while there are a few normal listings that provide shared rooms. The mega listings (Figure 13), however, show that the number of entire homes is about six times greater than the number of private rooms. In addition, there is no shared room provided in mega listings.

[12] "Boroughs of New York city" (Wikipedia). Retrieved from: https://en.wikipedia.org/wiki/Boroughs_of_New_York_City

The biggest difference between the two types of hosts is that mega hosts set more entire homes than the other two room types, while normal hosts have similar proportions of private rooms and entire homes. For normal hosts, a great part of listings are provided in private rooms as normal hosts could live together with customers and provide a unique homestay experience, and this is primarily the purpose of Airbnb. In some cases, the listing provided by the normal host is the only property they live in, so the normal host is not able to provide the entire home. Mega hosts, on the other hand, manage their listings in a commercial mode, and the entire home type could allow them to maximize profits. However, private rooms account for a certain proportion of aggregate demands. According to the data from Airbnb Resource center, there are more than 161,000 "superhosts" with private rooms listings on Airbnb, which is 29% of all "superhosts" listings.[13] "Superhosts", a group of distinguished hosts qualified officially by Airbnb, are automatically evaluated every 3 months in their performance standards and other qualifications.[14] Since superhosts listings are usually in high demand and are updated regularly (market demands are dynamic), the data provide us a rough idea of the popularity of private rooms.

In conclusion, mega hosts' provision of more entire houses probably is based on their idea of maximizing profits. While for normal hosts, the needs of different hosts are diverse and the properties they owned are different, which allows them to provide a high proportion of both private rooms and entire homes.

[13] A. Laura. "Laura's Letter: What I learned from staying in private rooms" (2019) Retrieved from: https://www.airbnb.com/resources/hosting-homes/a/lauras-letter-what-i-learned-from-staying-in-private-rooms-60

[14] "Superhost Terms and Conditions." (Airbnb Website). Retrieved from: https://www.airbnb.com/superhost/terms

6. Listing prices

We are also curious about whether mega hosts have specific pricing strategies, the kurtosis for normal and mega listings' prices are 579.94 and 8.56, while the skewness are 19.11 and 1.76, which again proves non-normality here. The result of Mann-Whitney test shows that there is a significant difference in prices between the two types of hosts (statistics=15245214.0, p-value= 0.0). From the boxplots of prices of normal listings and mega listings (Figure 14), we could see that a great number of outliers with high prices exist in normal listings. The median price of normal listings is 101 while the median price of mega listings is 215. The overall higher price of mega listings is reasonable. For most mega listings, mega hosts have to pay the rental costs as they do not own those properties, while most normal listings are properties owned by normal hosts, so they do not have the rental costs. To ensure profits, mega hosts have to set a higher price. Although mega hosts set higher prices overall, they don't have any listings with extremely large prices. From Figure 18, normal hosts have certain portions of listings with prices above 2,000 dollars per night, while all mega listings are below 1,200 dollars. It is also reasonable since it is difficult to rent out listings with higher prices, and mega hosts will have a higher risk of investment if they set very high prices, as the rental costs might be greater than the profits.

We also want to compare listing prices of different room types between normal and mega listings. From the five-number summaries of listing prices grouped by room type for normal listings and mega listings (Table IV & V), we could see that the price of entire homes is much higher than private rooms and shared rooms. For the type of entire homes, mega listings overall have higher prices than normal listings, which is reasonable because of rental costs, but for the type of private rooms, the prices of normal listings are generally higher. This is probably because normal hosts could provide a more unique homestay experience while mega hosts could not.

To compare prices of different neighborhood groups (Figure 15 & 16, Table VI & VII), mega listings overall have higher prices than normal listings in all neighborhood groups except for Queens. In addition, Manhattan has a higher listing price than other neighborhood groups for both normal listings and mega listings. Manhattan is also known as the most expensive place in NYC. According to data from furmancenter.org, the median rental price for Manhattan is the highest which is $1,770.[7] This might lead to the higher listing price of Manhattan. Besides, the effect of other factors such as tourism and commercialism also contributed to higher listing prices in Manhattan.

In this way, mega hosts overall set higher listing prices mainly because of the rental costs, but mega listings provided in private rooms and listings in Queens have lower prices than normal listings of the same type, we still need more information to figure out the reason.

7. Number of reviews

To check the popularity of listings, we want to check the number of reviews, the median of the number of reviews of mega listings is 0 (at least half of mega hosts listings don't have any review) and that for the normal listings is 6.0. From the boxplots of reviews (Figure 17), we could see that the number of reviews varies a lot among normal listings and has a great number of large value outliers, and mega listings have relatively a small range. This could tell us that mega listings overall have very few reviews, and half of those listings do not have any reviews.

To compare the number of reviews of different neighborhood groups among normal listings, the boxplots and five-number summary (Figure 18 and Table VIII) show that the numbers of reviews are overall similar for different neighborhood groups, which are approximately within the

[7] "Greenpoint/Williamsburg Neighborhood Profile." (NYU Furman Center) Retrieved from: https://furmancenter.org/neighborhoods/view/greenpoint-williamsburg.

range of one to forty. In the boxplots and data summary of mega listings (Figure 19 and Table IX), Brooklyn overall has more reviews, and there are several large value outliers in Manhattan.

The number of reviews as mentioned is considered as important information of mega hosts' profitability. However, the smaller number of their listings' reviews doesn't necessarily mean that mega listings have less popularity than normal listings. The fewer reviews among mega listings might be associated with the minimum night requirement of mega listings. From the previous study, we found that mega listings have higher minimum night requirements, so their numbers of reviews are also smaller as the listing could be booked by fewer customers. Also, the room type distribution of mega hosts is also worth considering, since the majority of their listings are entire homes that have relatively higher prices. What's more, the standardized rooms of mega listings lack features, which may initially have less attraction to guests and hence receive fewer reviews, as people are more likely to write reviews for impressive and unique accommodations. These three factors hence explain why mega hosts may encounter the amount of reviews less than normal, but once customers check in, mega hosts can guarantee that they will earn stable, long-term, and considerable revenues compared to normal hosts. This might be one of the mega hosts' strategies: under the limitations of singular room type, they are trying to adjust the minimum length of accommodation to filtrate long-term guests and let their listings be occupied as long as possible.

VI. TABLES AND FIGURES

A. Tables

TABLE I. Top eleven hosts who have most amount of listings in NYC and their # listings

Host ID	Number of listings
219517861	327

107434423	232
30283594	121
137358866	103
12243051	96
16098958	96
61391963	91
22541573	87
200380610	65
7503643	52
1475015	52

352

353

354 TABLE II. Data summary of minimum nights required by normal hosts

neighbourhood_group	count	mean	std	min	25%	50%	75%	max
Bronx	1089	4.5647	15.646	1.0	1.0	2.0	3.0	365.0
Brooklyn	20028	5.9848	17.615	1.0	2.0	3.0	5.0	999.0
Manhattan	20461	7.5250	23.141	1.0	1.0	3.0	5.0	1250.0
Queens	5585	4.8319	14.842	1.0	1.0	2.0	3.0	500.0
Staten Island	373	4.8311	19.728	1.0	1.0	2.0	3.0	365.0

355

356

357 TABLE III. Data summary of minimum nights required by mega hosts

neighbourhood_group	count	mean	std	min	25%	50%	75%	max

Brooklyn	61.0	30.000	0.00000	30.0	30.0	30.0	30.0	30.0
Manhattan	1182.0	26.078	13.1626	2.0	29.0	30.0	30.0	180.0
Queens	79.0	30.000	0.00000	30.0	30.0	30.0	30.0	30.0

TABLE IV. Data summary of prices of different room types of normal hosts

room_type	count	mean	std	min	25%	50%	75%	max
Entire home/apt	24181.0	210.479	290.119	0.0	119.0	159.0	225.0	10000.0
Private room	22196.0	89.9268	160.567	0.0	50.0	70.0	96.0	10000.0
Shared room	1159.0	70.0759	101.754	0.0	33.0	45.0	75.0	1800.0

TABLE V. Data summary of prices of different room types of mega hosts

room_type	count	mean	std	min	25%	50%	75%	max
Entire home/apt	1212.0	238.307	106.237	70.0	171.5	227.0	275.25	1170.0
Private room	110.0	63.0728	77.2814	24.0	37.0	41.5	52.00	616.0

TABLE VI. Data summary of prices of different neighborhood groups of normal hosts

neighbourhood_group	count	mean	std	min	25%	50%	75%	max
Bronx	1089.0	87.4692	106.799	0.0	45.0	65.0	99.0	2500.0
Brooklyn	20028.0	124.350	187.200	0.0	60.0	90.0	150.0	10000.0
Manhattan	20461.0	194.417	298.457	0.0	90.0	145.0	209.0	10000.0
Queens	5585.0	100.323	168.171	10.0	50.0	75.0	112.0	10000.0
Staten Island	373.0	114.812	277.620	13.0	50.0	75.0	110.0	5000.0

TABLE VII. Data summary of prices of different neighborhood groups of mega hosts

neighbourhood_group	count	mean	std	min	25%	50%	75%	max

Brooklyn	**61.0**	**144.262**	**47.95203**	**35.0**	**129.0**	**149.0**	**159.0**	**312.0**
Manhattan	**1182.0**	**239.844**	**108.9678**	**31.0**	**175.0**	**229.0**	**278.0**	**1170.0**
Queens	**79.0**	**43.9241**	**10.135963**	**24.0**	**37.5**	**41.0**	**50.0**	**70.0**

TABLE VIII. Data summary of reviews of different neighborhood groups of normal hosts

neighbourhood_group	count	mean	std	min	25%	50%	75%	max
Bronx	**1089.0**	**26.0184**	**42.2449**	**0.0**	**1.0**	**9.0**	**32.0**	**321.0**
Brooklyn	**20028.0**	**24.2623**	**44.3868**	**0.0**	**1.0**	**6.0**	**25.0**	**488.0**
Manhattan	**20461.0**	**22.0864**	**43.5269**	**0.0**	**1.0**	**5.0**	**21.0**	**607.0**
Queens	**5585.0**	**28.0838**	**52.2298**	**0.0**	**1.0**	**8.0**	**32.0**	**629.0**
Staten Island	**373.0**	**30.9410**	**44.8308**	**0.0**	**1.0**	**12.0**	**42.0**	**333.0**

TABLE IX. Data summary of reviews of different neighborhood groups of mega hosts

neighbourhood_group	count	mean	std	min	25%	50%	75%	max
Brooklyn	**61.0**	**4.08197**	**3.094592**	**0.0**	**2.0**	**4.0**	**6.0**	**15.0**
Manhattan	**1182.0**	**1.87564**	**3.109179**	**0.0**	**0.0**	**0.0**	**2.0**	**20.0**
Queens	**79.0**	**0.68354**	**0.884979**	**0.0**	**0.0**	**0.0**	**1.0**	**3.0**

B. Figures

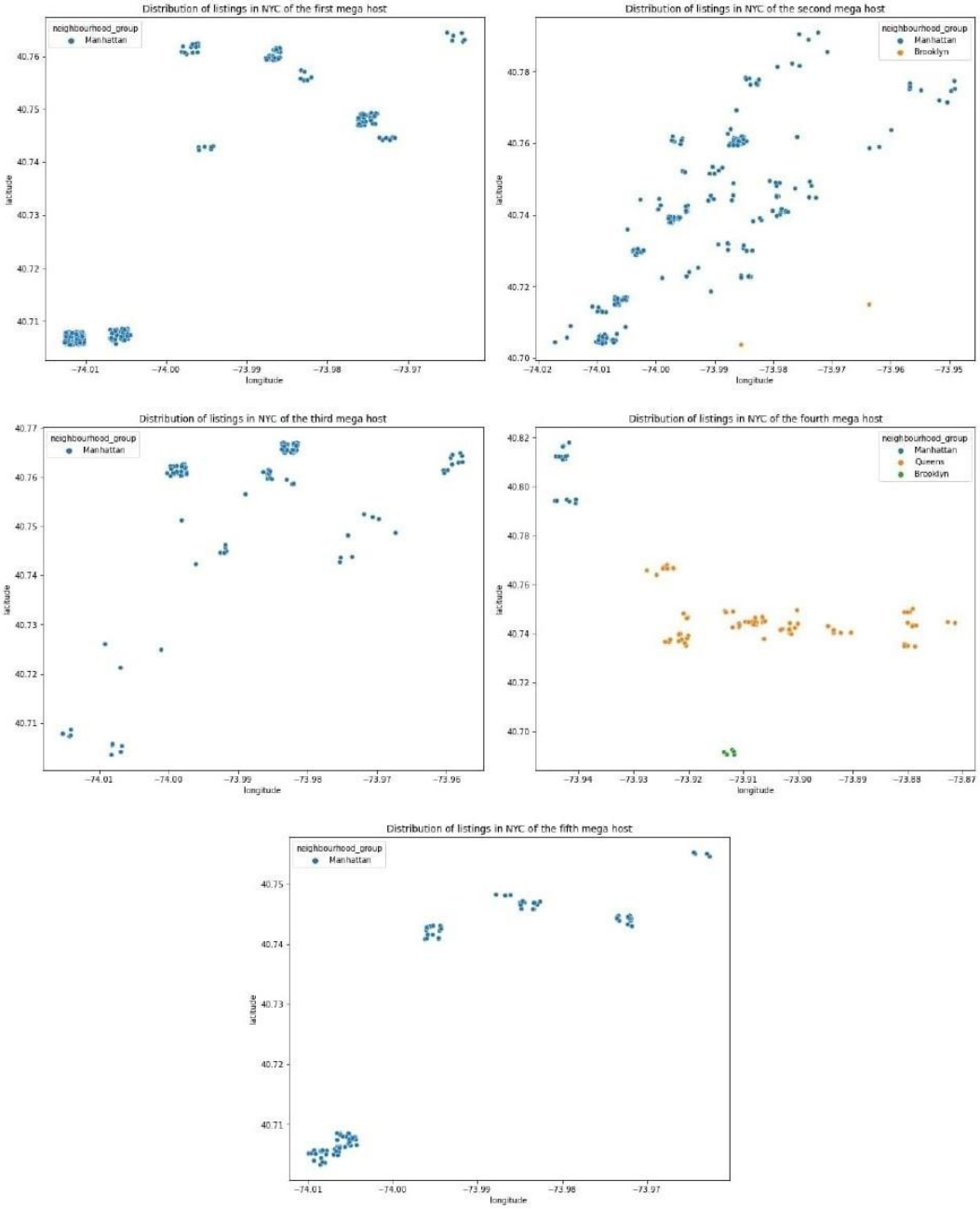
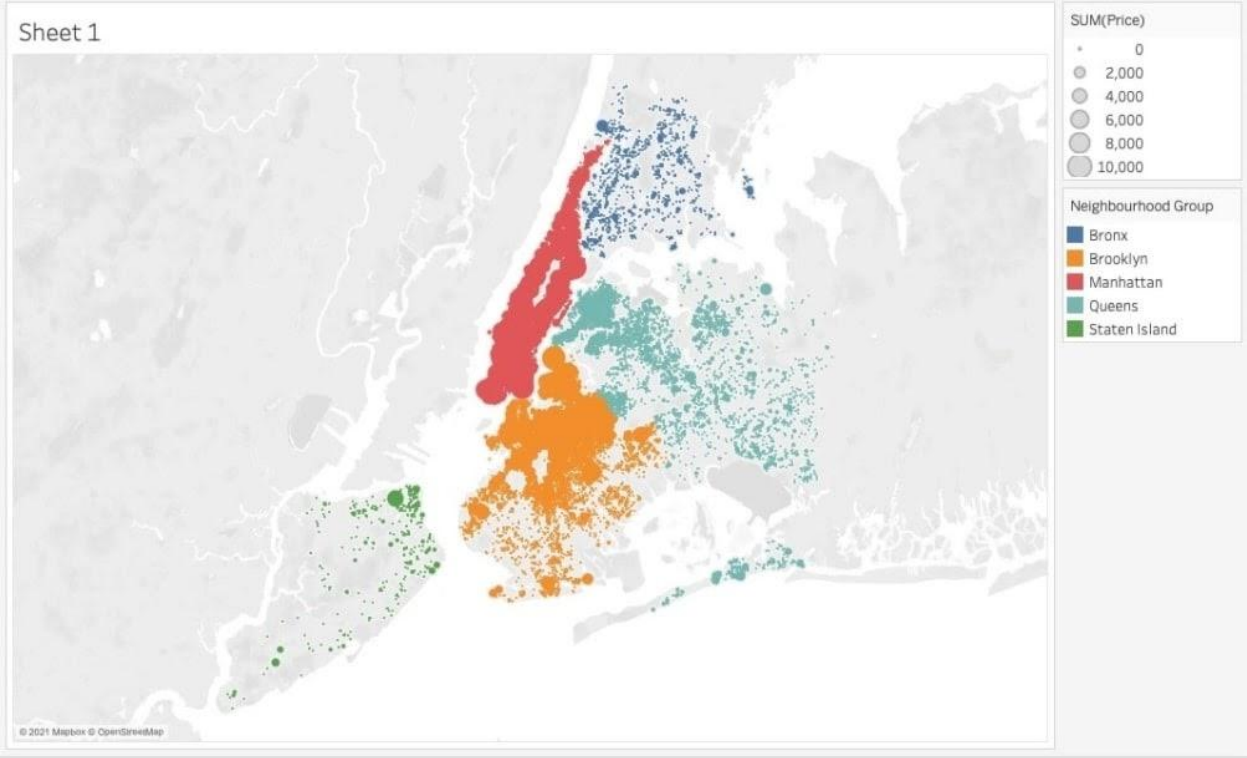
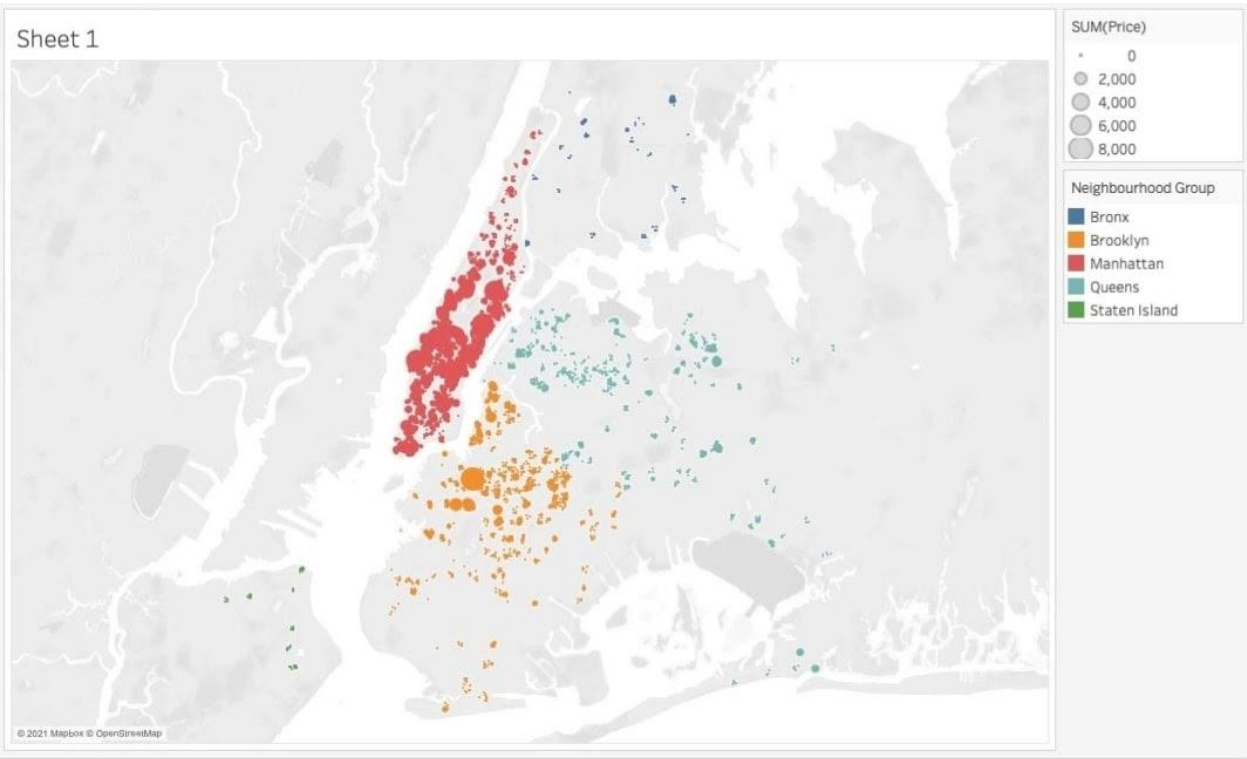


FIG1.jpg: Distribution of listings in NYC of 1st to 5th mega host



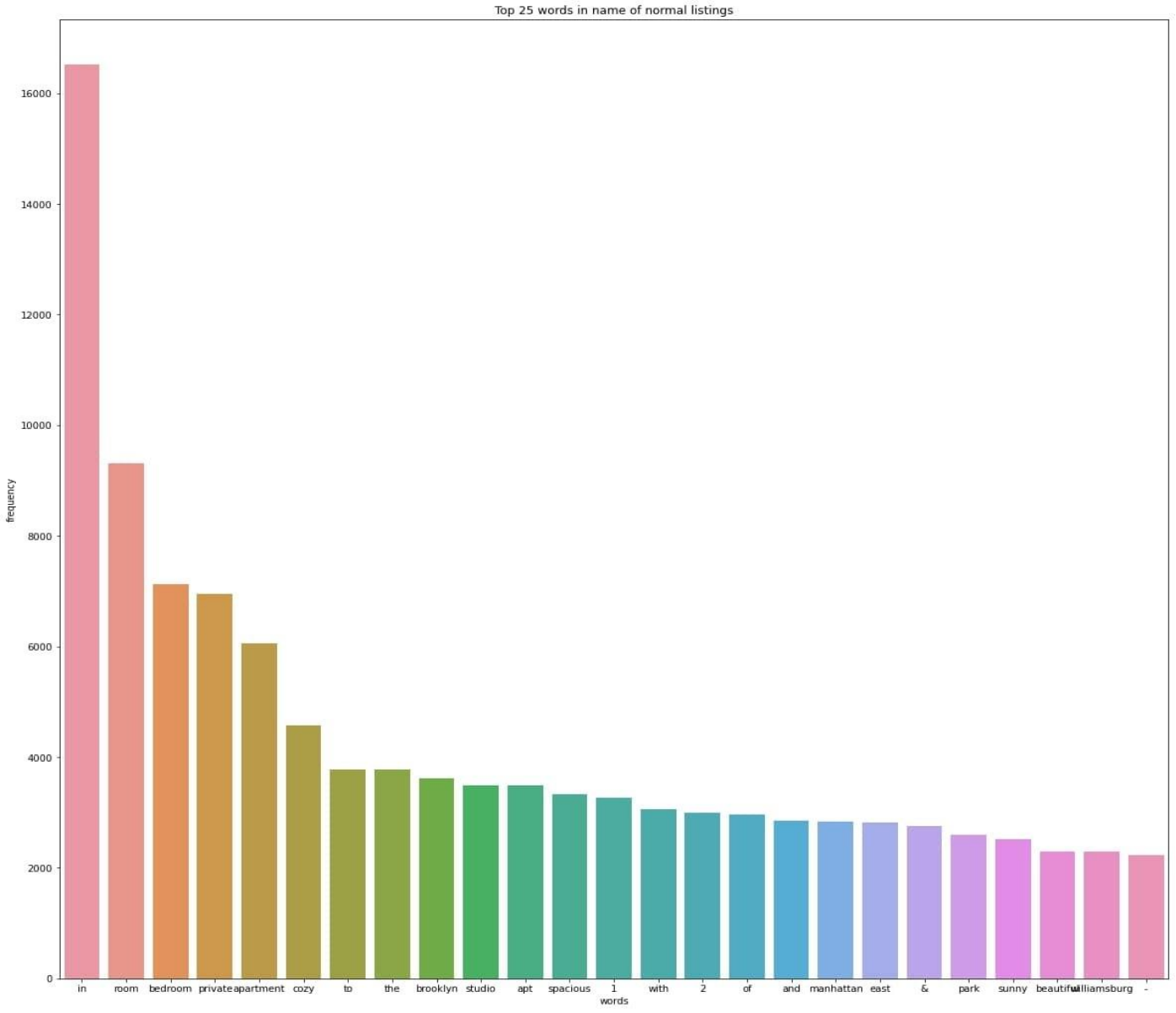
382

383 FIG2.jpg: Distribution of neighborhood group and price(population data)



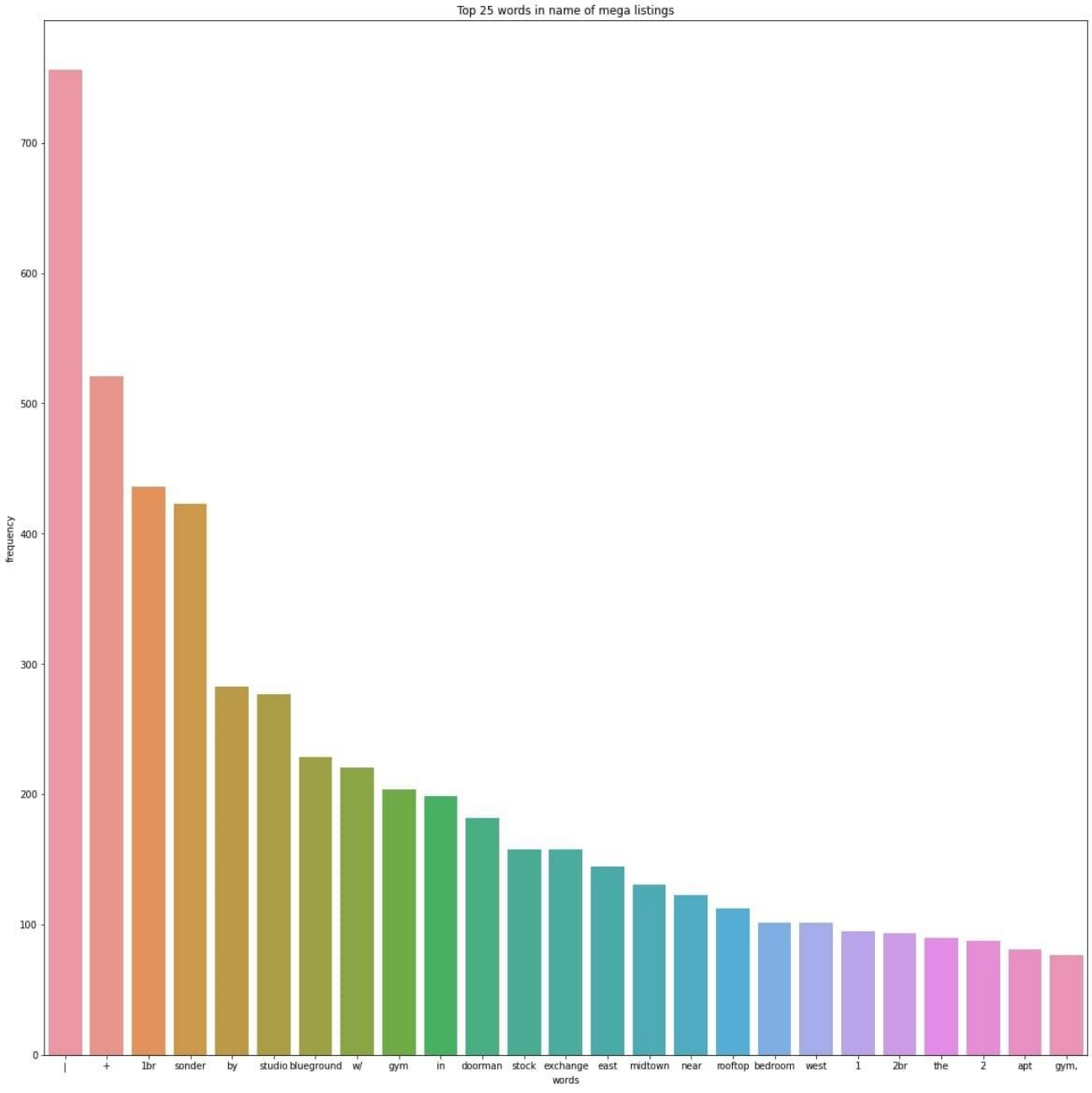
384

385 FIG3.jpg: Distribution of neighborhood group and price(Mega Host data)



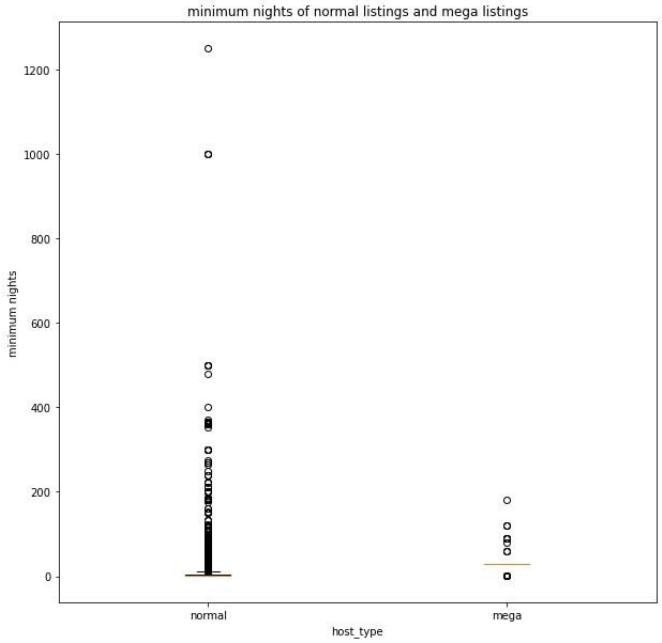
386

387 FIG4.jpg: Top 25 words in listing names of normal hosts



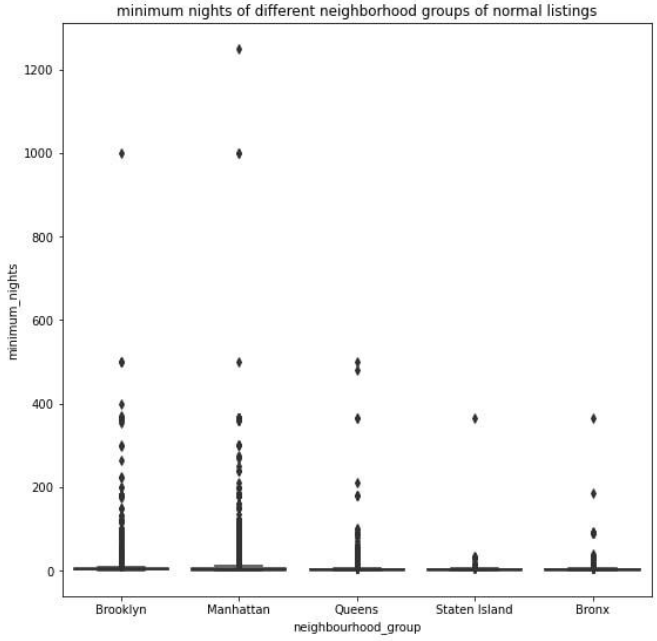
388

389 FIG5.jpg: Top 25 words in listing names of mega hosts



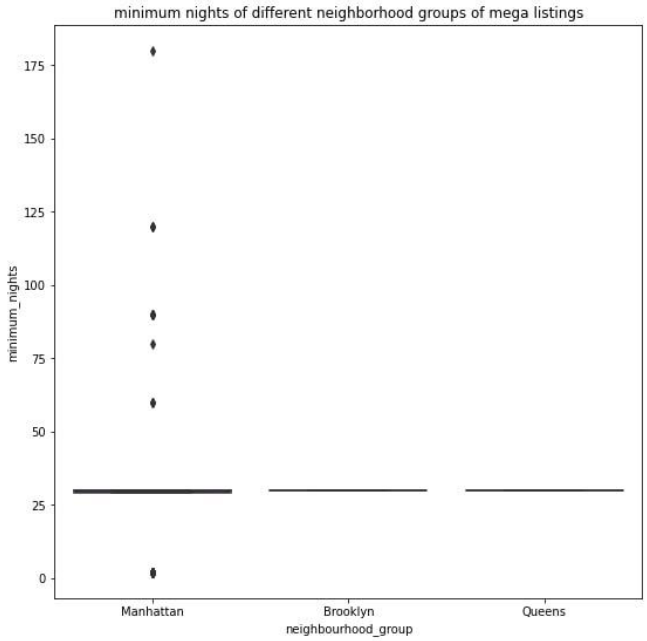
390

391 FIG6.jpg: Minimum nights required between normal hosts and mega hosts



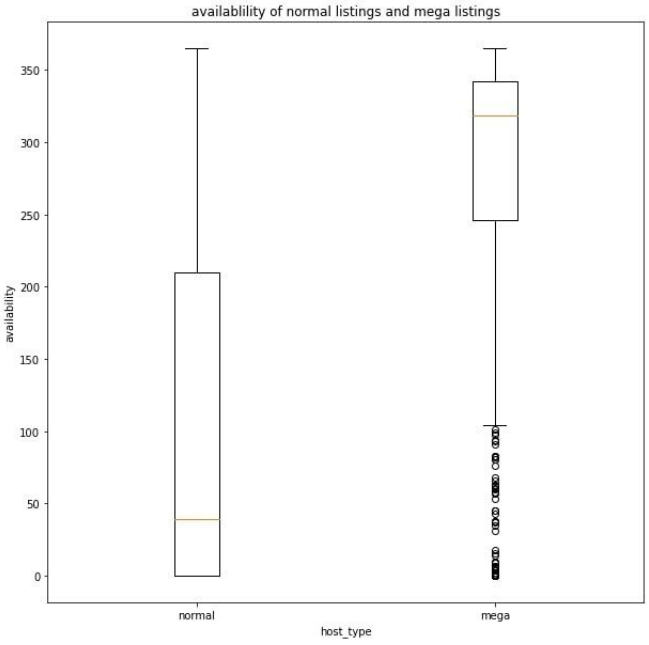
392

393 FIG7.jpg: Minimum nights required by normal hosts among five neighbourhood groups



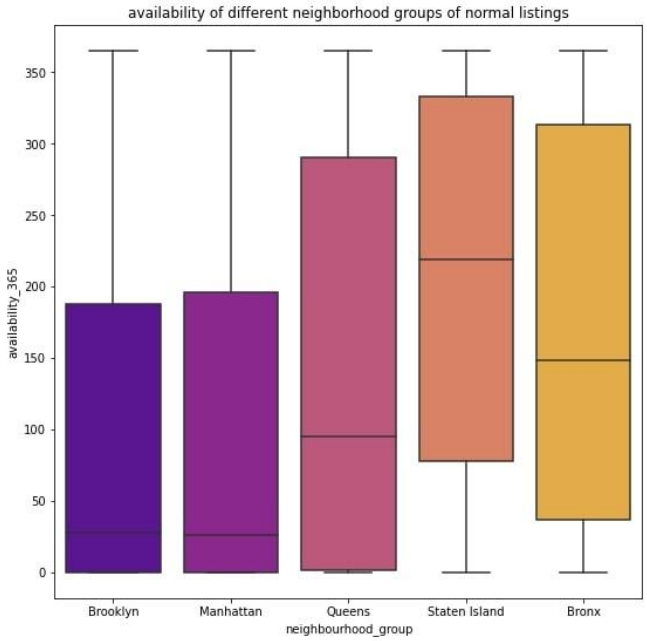
394

395 FIG8.jpg: Minimum nights required by mega hosts among five neighbourhood groups



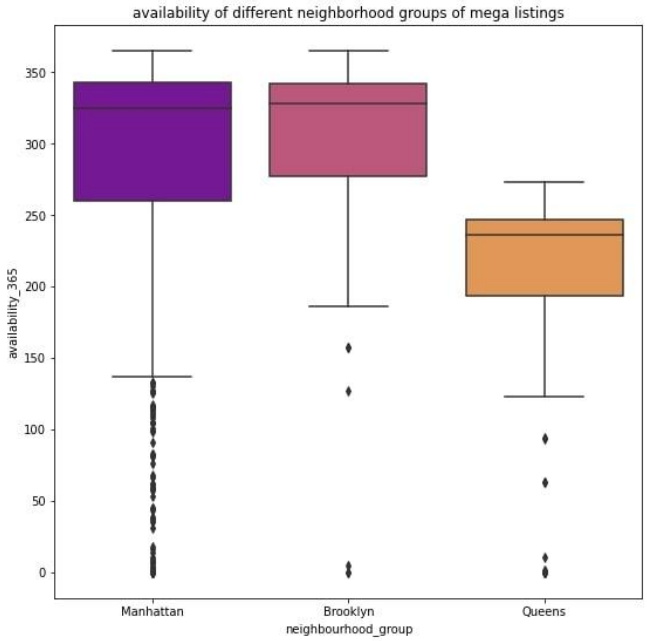
396

397 FIG9.jpg: Available days between normal hosts and mega hosts



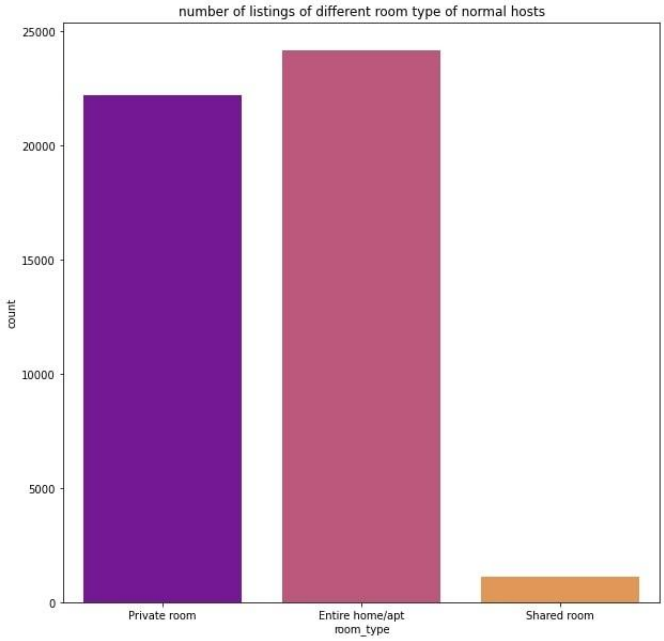
398

399 FIG10.jpg: Availability of listings among five neighbourhood groups of normal hosts



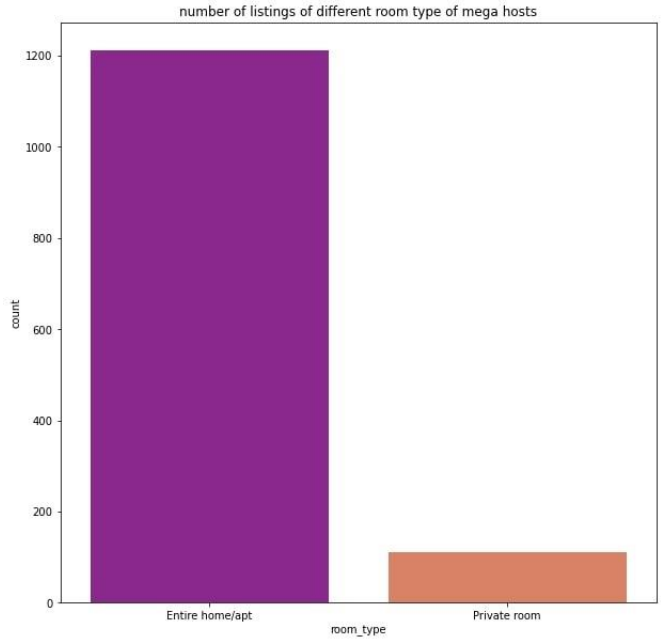
400

401 FIG11.jpg: Availability of listings among five neighbourhood groups of mega hosts



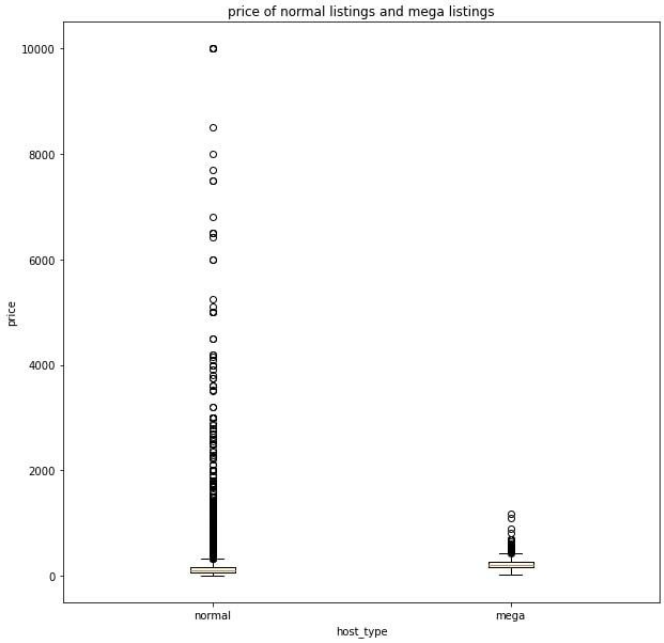
402

403 FIG12.jpg: Number of listings of different room types among normal listings



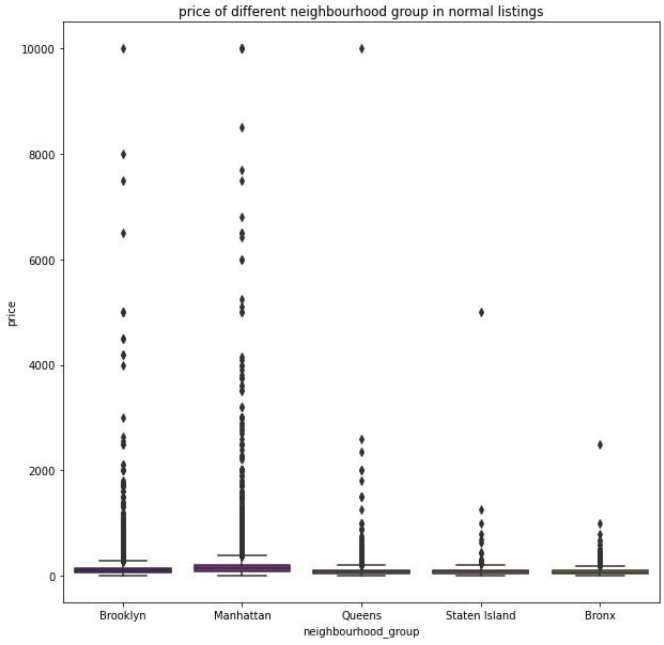
404

405 FIG13.jpg: Number of listings of different room types among mega listings



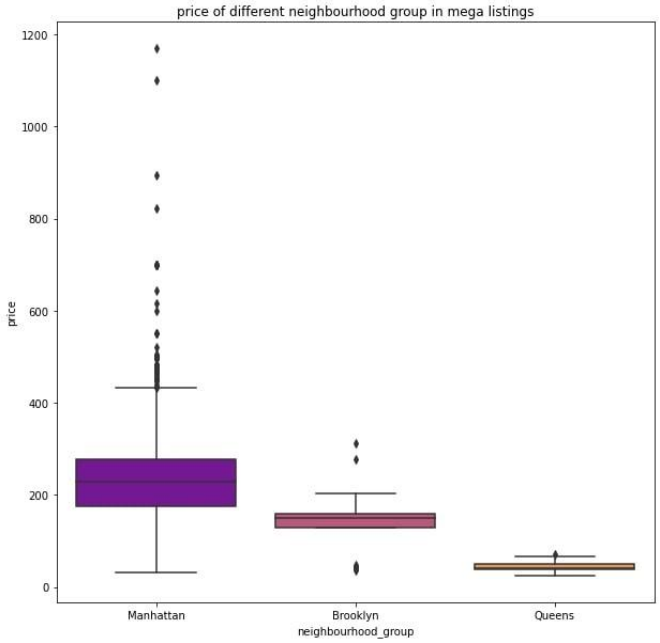
406

407 FIG14.jpg: Price of normal listings and mega listings



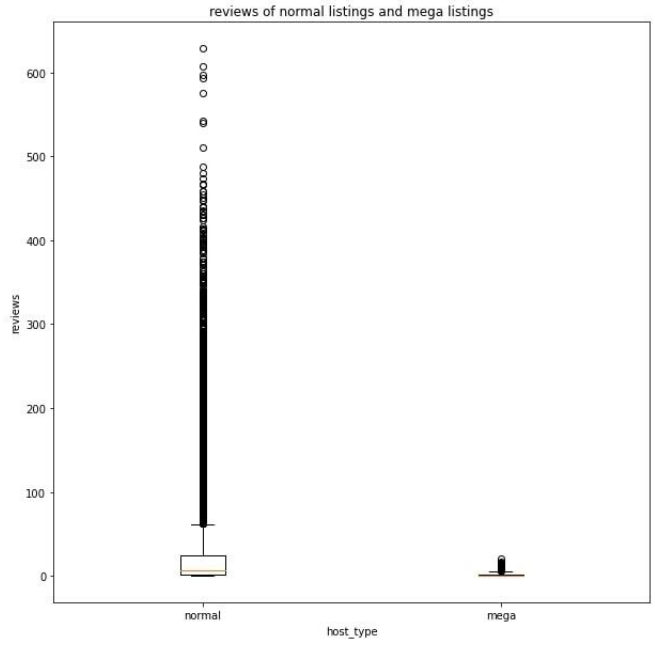
408

409 FIG15.jpg: Price of different neighborhood groups in normal listings



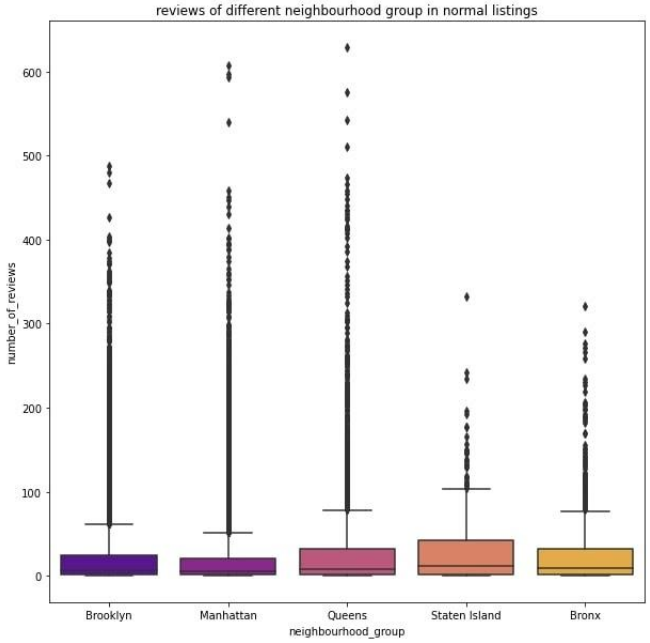
410

411 FIG16.jpg: Price of different neighborhood groups in mega listings



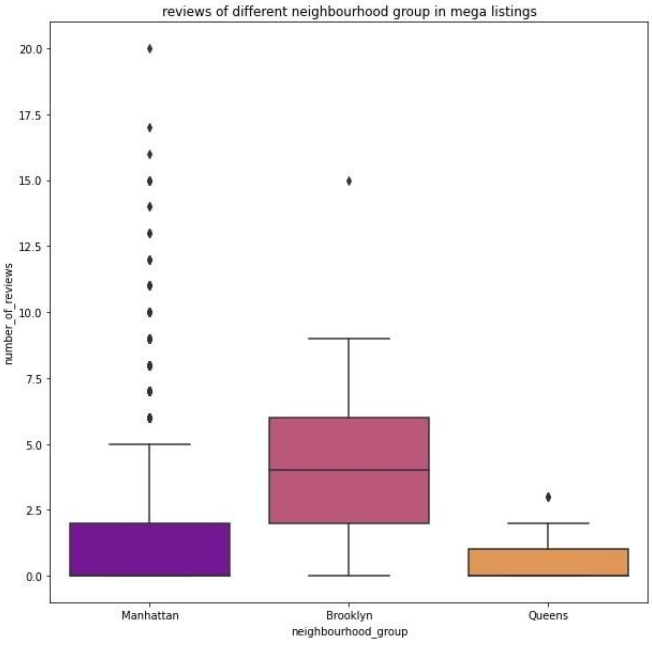
412

413 FIG17.jpg: Number of reviews of normal listings and mega listings



414

415 FIG18.jpg: Number of reviews of different neighborhood groups among normal listings



416

417 FIG19.jpg: Number of reviews of different neighborhood groups among mega listings

418

419

VII. CONCLUSION

In conclusion, we came up with an overview of mega hosts' business. Mega hosts usually rent several listings in certain areas where the average price of listings is high. The aggregation of listings in one area provides them conveniences to manage their properties efficiently and spend less costs on investigating new mega-hosts areas. Mega hosts are professional at advertising their listings, with formulaic listings names and certain brand names for promotion. In comparison to normal hosts, mega hosts have longer time requirements on minimum night. It can be considered as a crucial strategy for them to ensure long-term profits, since they are only being chosen by long-term tenants. The room types of mega hosts might be a limitation since they mainly provide entire houses, while private rooms have a certain number of demands in the market. The median availability of mega hosts is above 300 days, which is an apparent advantage in market share, compared to normal hosts with median availability lower than 50. It means that mega hosts will have higher probabilities to be occupied at some period of time when the total listing supplies are low. Affected by the room type and locations, mega hosts set higher prices than normal hosts. Considering the price together with minimum nights and availability, it shows that the strategy of mega hosts is targeting guests who can stay for a longer time and are willing to pay for a higher price, which explains why these listings are mostly located in Manhattan - visitors to Manhattan may have relatively higher consumption levels. At least half of the mega host listings have zero reviews, which indicates that their listings are not so popular as normal listings. The room types and minimum nights set by mega hosts may cause a smaller number of people to have opportunities to visit. Also, the lack of features reduces their attractions to potential customers.

In our research, we didn't explicitly evaluate the profitability of mega hosts. More data are needed for further analysis, such as the rate of occupancy and costs that mega hosts need to tolerate. Our hypothesis of mega hosts profitability is that mega hosts businesses are profitable. Even though

they have really low numbers of reviews, it cannot be equivalent to a low occupancy rate. What if they just have a low percentage review rate?

For further discussion, the trend of mega hosts may cause many problems. Currently, the market share of mega hosts is dwarfed by that of normal hosts. However, with the growth rate of nearly 90% in the number of mega hosts in 2015, we can't help imagine whether mega hosts will finally dominate the market in one day, and completely change the platform into a hotel reservation system. It is probably time for Airbnb to implement some new policies to stop this trend, such as placing an upper limit of listing numbers that an individual host can hold. Of course, the outbreak of COVID-19 may have caused a large and negative impact on these mega hosts, which can slow this trend from deteriorating.

VIII. REFERENCES (NUMERICAL)

[1] R. Lorena. "New "mega hosts" help level up airbnb's brand and traveler experience." (2017, June 30).Retrieved from: https://blog.turnkeyvr.com/new-mega-hosts-help-level-airbnbs-brand-traveler-experience/.

[2] BnB Hosts Team. "The 3 Main Airbnb Host Categories: Which One Are You?"(2020, April 21). Retrieved from: https://www.bnbhosts.com.au/3-airbnb-host-categories/.

[3] M. Chris. "NY Attorney General: 72% Of Airbnb Rentals Violate State, City Laws. Consumerist." (2016, September 27).Retrieved from: https://consumerist.com/2014/10/16/ny-attorney-general-72-of-airbnb-rentals-violate-state-city-laws/index.html.

[4] "From Air Mattresses to Unregulated Business: An Analysis Of The Other Side Of Airbnb." (A Comprehensive Survey of 14 Metropolitan Statistical Areas , 2016).Retrieved from: https://www.ahla.com/sites/default/files/Airbnb_Analysis_September_2016_0.pdf

[5] D. Ameya. "The Effect Of Superhost Status On Airbnb In Berlin Using Occupancy Rate And Revenue Per Available Listing."(2020). Retrieved from: https://doi.org/10.13140/RG.2.2.16576.56325

[6] K. Yoon , K. Jaewook , and T. V. Yue " How you name your Airbnb's title matters: comparison of seven countries." (2021).Journal of Travel & Tourism Marketing, 38(1), 93–106. Retrieved from: https://doi.org/10.1080/10548408.2021.1875105

[7] "Greenpoint/Williamsburg Neighborhood Profile." (NYU Furman Center) Retrieved from: https://furmancenter.org/neighborhoods/view/greenpoint-williamsburg.

[8] C. Yujia , and S. Markus. "Why people choose Airhnb over Hotel?" (2016). Retrived from: https://www.researchgate.net/publication/305225839_Why_people_choose_Airbnb_over_Hotel

[9] K. Ashlee. "Some Airbnb 'Mega Hosts' Are Renting Out More Than 1,000 Properties At Once. Consumerist." (2017, June 9) Retrieved from: https://consumerist.com/2017/06/09/some-airbnb-mega-hosts-are-renting-out-more-than-1000-properties-at-once/.

[10] "Writing Catchy Airbnb Titles: Proven Formulas That Attract 5x More Bookings. Medium."(iGMS, 2019, June 25). Retrieved from: https://medium.com/@airgms/writing-catchy-airbnb-titles-bc7ba0f3a0ee.

[11] E. Christopher. "Airbnb Runs 'Illegal Hotels,' Hotel Industry Study Claims" (2016). Retrieved from: https://fortune.com/2016/01/20/airbnb-illegal-hotels-study/

[12] "Boroughs of New York city" (Wikipedia). Retrieved from: https://en.wikipedia.org/wiki/Boroughs_of_New_York_City

[13] A. Laura. "Laura's Letter: What I learned from staying in private rooms" (2019) Retrieved from: https://www.airbnb.com/resources/hosting-homes/a/lauras-letter-what-i-learned-from-staying-in-private-rooms-60

[14] "Superhost Terms and Conditions." (Airbnb Website). Retrieved from: https://www.airbnb.com/superhost/terms