**Analytical Guide to Listing on Airbnb in New York City**

Group 10: Ruiqi Xie, Wai Kin Chan, Cindy Huang

BAIS: 6050:0800

December 4th, 2020

# **Introduction**

Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present a more unique and personalized way of experiencing the world. Airbnb has become one of the world’s largest marketplace for travelers to find suitable, reliable places to stay and things to do. Since Airbnb’s opening, they offered over 7 million accommodations and 50,000 unique activities, all powered through local entrepreneurs and hosts.

We are interested in becoming a host for Airbnb in New York City. Our goal is to analyze which neighborhood group in New York City is a good place to start our business at. Furthermore, we want to have some insight on how the demand and price of Airbnb listings change over time to decide which months and price point would be the most suitable.

# **Dataset Description:**

To analyze which neighborhood group in New York City would be the best place to post our listing on Airbnb, we have a dataset that describes the listing activity in New York City for 2019 and 2020. These two datasets have all the information about hosts, listings, and geographical availability for us to draw specific conclusions on our analysis. Below are the screenshots of the first 16 rows from the dataset.

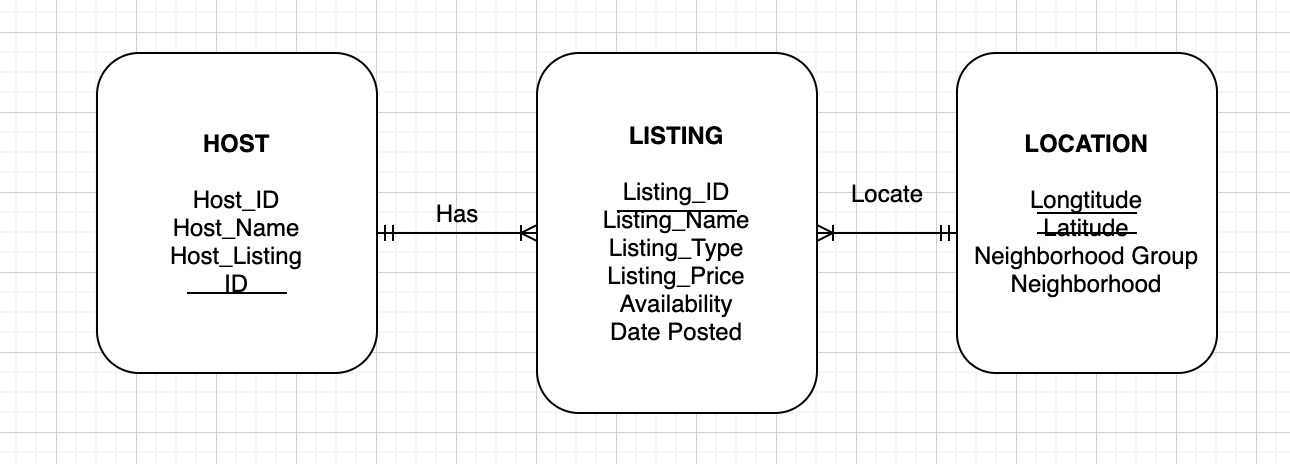


Airbnb\_NYC\_2019

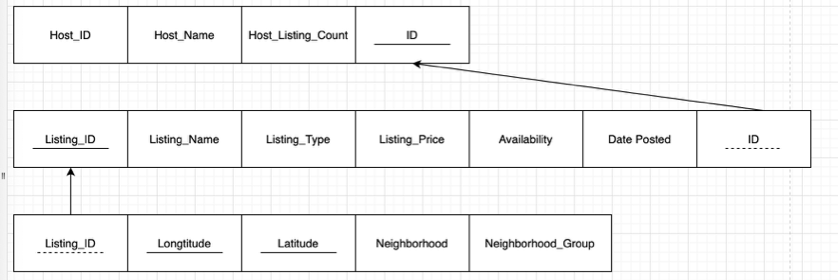


Airbnb\_NYC\_2020

**E-R diagram:**

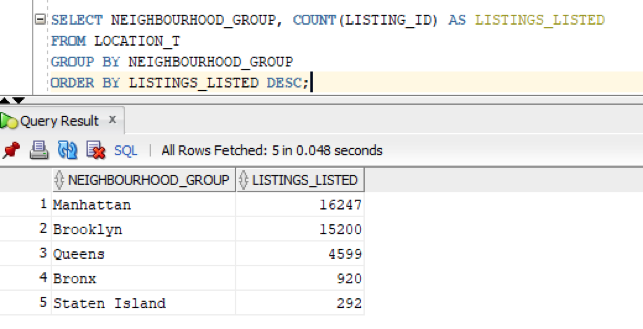


# **Relational Schema:**

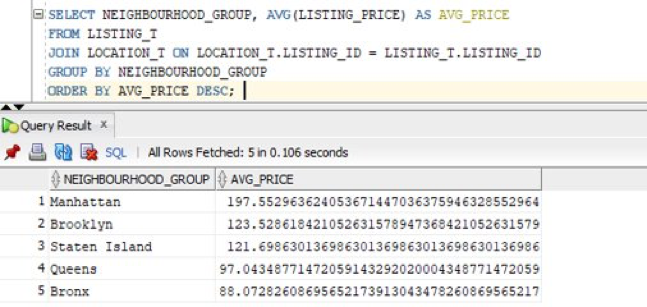


# **Data analysis with SQL queries:**

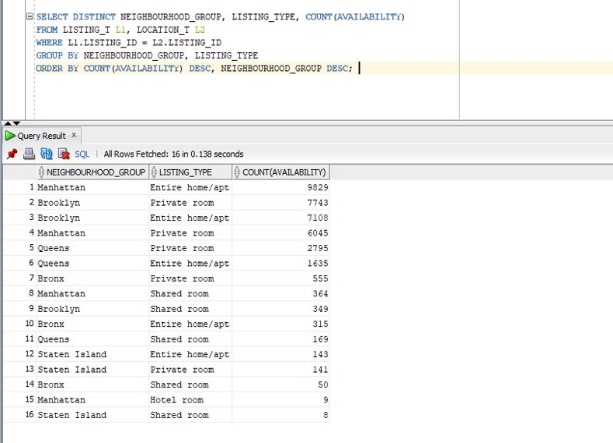
Our first two analysis question will be around number of listings and average price in each neighborhood group. The first question is which neighborhood groups are the hotspots in New York City. To solve this question, we selected neighborhood group, and counted the listings ID as listings listed in each neighborhood group from the location table. Then, we can see the SQL results in following screenshot 1. There are five neighborhood groups in New York City: Manhattan, Brooklyn, Queens, Bronx, and Staten Island. We found that the hottest neighborhood groups are Manhattan with 16,247 listings listed in Airbnb website, and Brooklyn with 15,200 listings listed. Thus, this result could give us a basic understanding of how listings separates in different neighborhood groups.



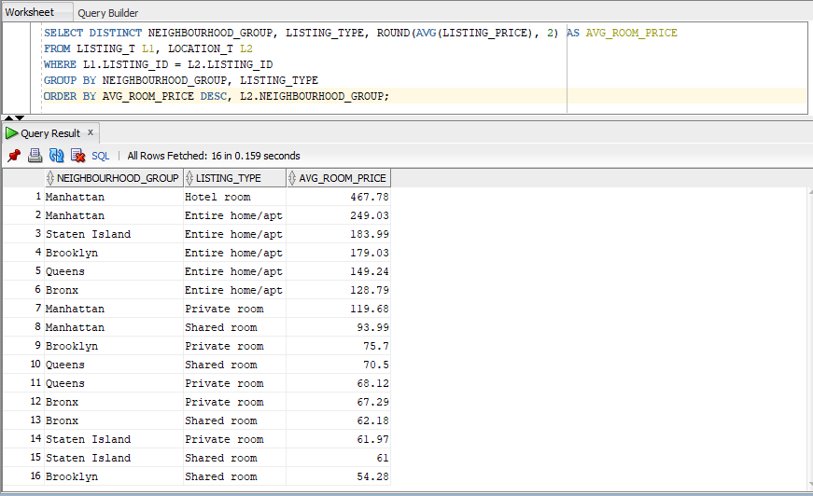
Then, we want to explore how the listings price are related to different neighborhood groups (Manhattan, Brooklyn, Queens, Bronx, and Staten Island) in New York City. Therefore, our second analyzing question is which neighborhood group has the highest average price of listings. To solve this question, we selected neighborhood group, and counted the average of listings price as average price in each neighborhood group from the listing and location table. Then, we have a descending order based on the average price in all neighborhood groups. We can see the SQL results in following screenshot, and we found that Manhattan has the highest average listing price, which is $197.55, and second is Brooklyn with $123.53 average price of listings. Consequently, we can interpret that there is some relationship between different neighborhood groups and its average listings price since Manhattan is the hotspot and it has the highest average listings price.



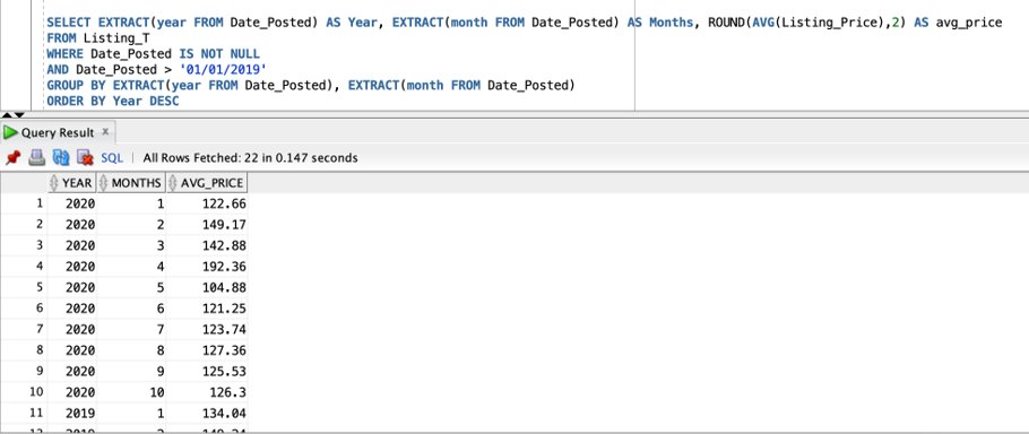
Our second analysis focused more on the demand for different room types and their corresponding average price. The first question was which room types are most popular in each neighborhood group (Bronx, Brooklyn, Manhattan, Queens, and Staten Island). To find this out, we selected the neighborhood groups, listing type (which is room type), and counted the days available from the Listing table and the Location table. We found that in Bronx, shared rooms were the least readily available, therefore the most rented out; Brooklyn, shared rooms were also the least readily available; Manhattan, hotel rooms were the least readily available; Queens, shared rooms were the least readily available, and in Staten Island, it’s a little hard to see, but shared rooms were the least readily available to users (only by like 2 points).



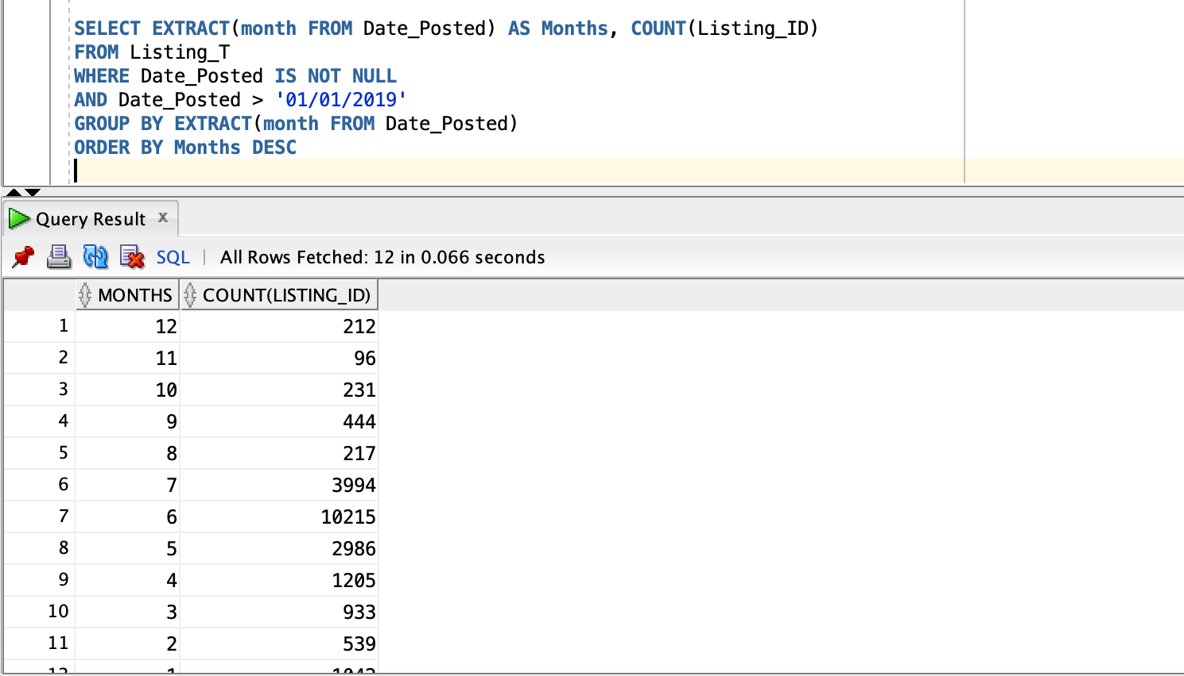
The second question was what the average price of each room type in each neighborhood is. To figure this out, we selected neighborhood groups, listing types or room type, and the average of the listing price and saved it as the variable (AVG\_ROOM\_PRICE). You can see the results here in the query. We found it interesting that Manhattan had the highest average price in all the room types out of all the neighborhoods. We hypothesis that this may be because of the elevated level of tourism in Manhattan vs the more residential areas of the Bronx, Brooklyn, Queens, and Staten Island. Manhattan is often referred to as “the City” and is often described as the cultural, financial, media, and entertainment capital of the world.



Our third analysis was focusing more on time series analysis. There are two questions answered in this section. The first question is how does the average monthly price changes from 2019 to 2020. To answer that question, first we needed to extract the relevant data from the SQL developer. We used the EXTRACT function to separate the Date\_Posted into year and month. Furthermore, we calculated the average listing price as well. All the data are taken from the Listing Table. Lastly, we grouped the data by year then month. After figuring out the SQL queries, we inserted the query into Tableau as a custom query and by using the visualization tools provided, we created a line graph. In the line graph, we can see that the average price in 2019 is a stable, hovering around $120-$140, while showing a spike at the end of the year most likely because of Christmas. In 2020, we can see a drop in average price around April, that was when COVID-19 first hit the United States.



The second question is which month has the highest number of listings posted. The overall steps are like the first question. In the SQL developer, we extracted the month from Date\_Posted and counted the number of Listing\_ID. We picked only the dates that are not null and grouped them by months. In Tableau, we used the line graph as well. From the graph, we can see that there is a significant increase around May and peaking in June at 10215 listings, which is during Summertime. It makes sense because New York is one of the hotspots for summer vacation.



# **Conclusion:**

In conclusion, we found that the hottest or most popular neighborhood group was Manhattan. Manhattan had a total of 16,247 total listings in 2019. Brooklyn was a close substitute and had 15,200 listings in 2019. Manhattan also had the highest average price of listings, around $197.55. As hopeful Airbnb hosts, an effective business decision would be based on what location has elevated levels of traffic, resulting in elevated levels of demand. These elevated levels of demand would allow us to also list at a higher price. From our analysis, we conclude that we would choose to buy a home in Manhattan because of the elevated levels in tourism, resulting in higher levels of demand for Airbnb home rentals. We do have higher competition in Manhattan, but our main competitors would be hotel rooms. While hotel rooms serve as a good substitute, the experience gained from staying at an Airbnb listing and a hotel during travel are fundamentally different. People looking to gain the full travel experience of a “New Yorker”, or “City Dweller” would more likely stay at an Airbnb home, and not a hotel. We would also choose to post the listing with a higher price in winter because there are fewer listings in that period. That creates even less competitors during the off-tourism season. This allows us to post at a higher price, gaining more profit. After an in-depth analysis at our dataset, our group would choose to buy a home in Manhattan and list the home at a higher price in winter than in summer.