

An aerial photograph of the New York City skyline, showing a dense cluster of skyscrapers. The image is split vertically down the middle. The left half is in a dark, muted color palette, while the right half is in full color. A thick white horizontal bar is positioned in the upper right corner, and a large white plus sign is located in the lower right corner.

Airbnb in New York

502466 Ying Chen

499318 Jiaqi Yin

502461 Che Zhang



Catalog

- Dataset Description
- Neighborhood
- Housing
- Host
- Data Insights



01 *Part One*

Dataset Description



Dataset Description



Dataset Introduction:

The data utilizes public information compiled from the Airbnb website including the availability calendar for 365 days in the future, and the reviews for each listing. Data is verified, cleansed, analyzed, and aggregated.



Data collection:

The data was collected by Inside Airbnb from 2008.8 to 2021.12



Data Source:

<http://insideairbnb.com/get-the-data/>



Data Size & Dimension:

32MB

38277 rows and 69 columns



Dataset Description



Dataset Processing:

Airbnb was founded in August 2008 and the data was collected for all Airbnb's from inception to December 2021. We chose data from New York for our study because it is a metropolitan area in the U.S., which allowed us to better focus on Airbnb in metropolitan areas in the U.S.

The original data had Price preceded by the \$ symbol. Since Tableau could not recognize it and could not change the type of Price to number, we removed all the \$ symbols from the Price column.

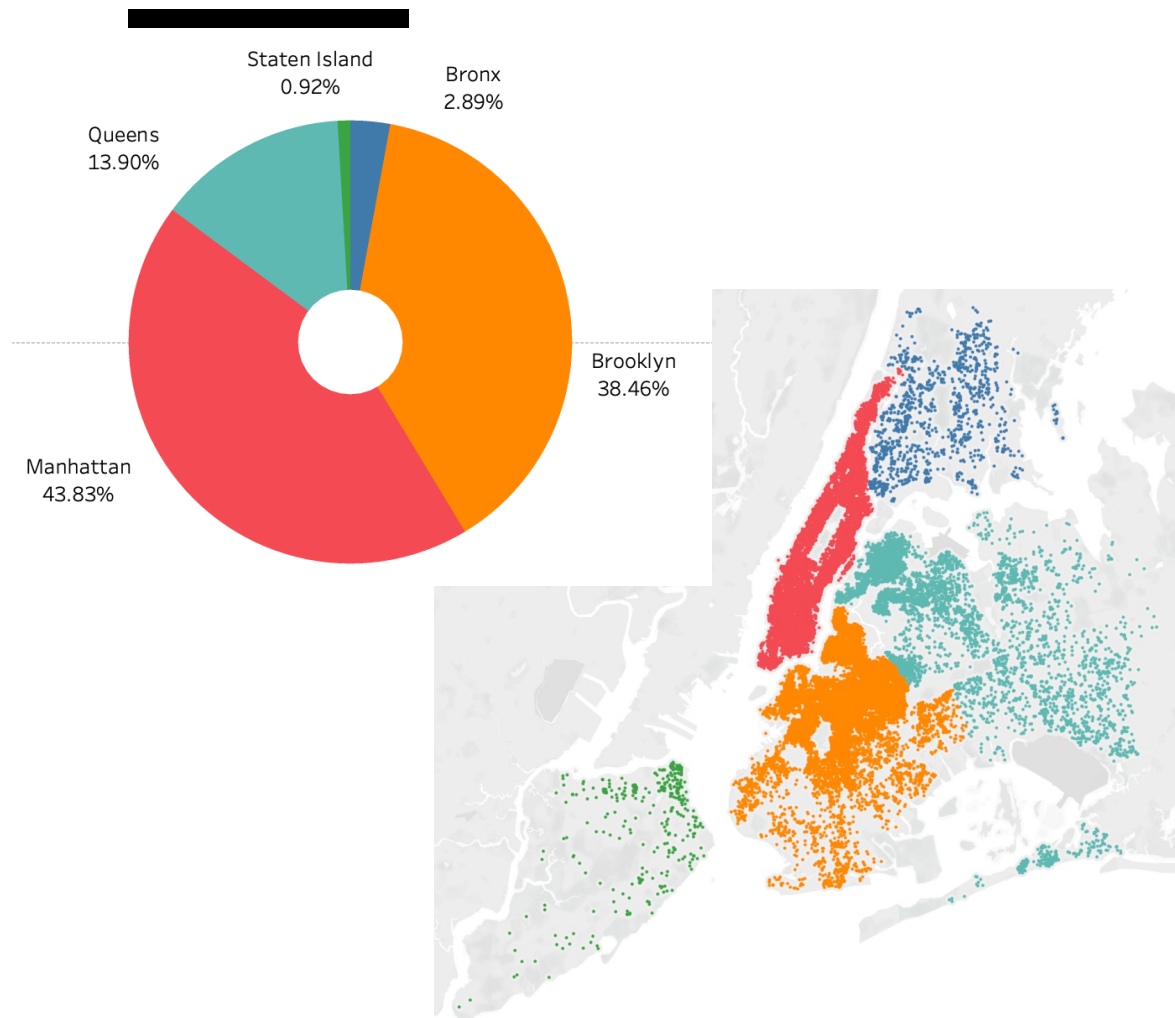
Since there are some abnormal values in Price, such as 10,000 and 9,999, which obviously shows that those properties are not for rental in general, we removed those extreme values and restricted Price to \$0-\$8,000.

We also created another calculated field called "Review Scores Group" based on the ones digit (0 - 5) to see the distribution of Review Scores in "Likert Chart for Neighborhood".

02 *Part Two*

Neighborhood



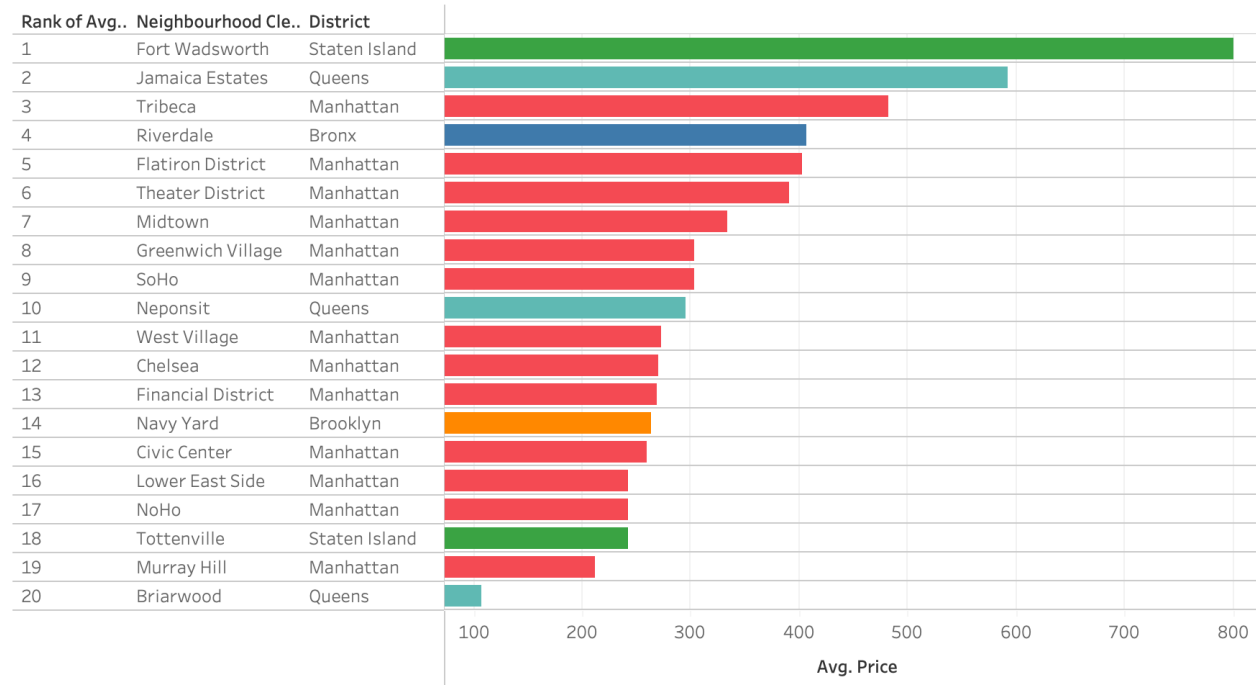


Red: Manhattan
Orange: Brooklyn
Light Blue: Queens
Dark Blue: Bronx
Green: Staten Island

Distribution of Airbnb by Districts

- We created a donut chart and a density map to show the distribution of Airbnb by Districts in New York.
- From the donut chart, we can tell that Manhattan(43.83%) and Brooklyn(38.46%) own the majority of Airbnb in New York, counting for more than 80% in total, following Queens(13.9%), Bronx(2.89%), and Staten Island(0.92%).
- From the density map, we can see that the closer distance to Manhattan, the higher density.

Top 20 Most Expensive Neighborhood

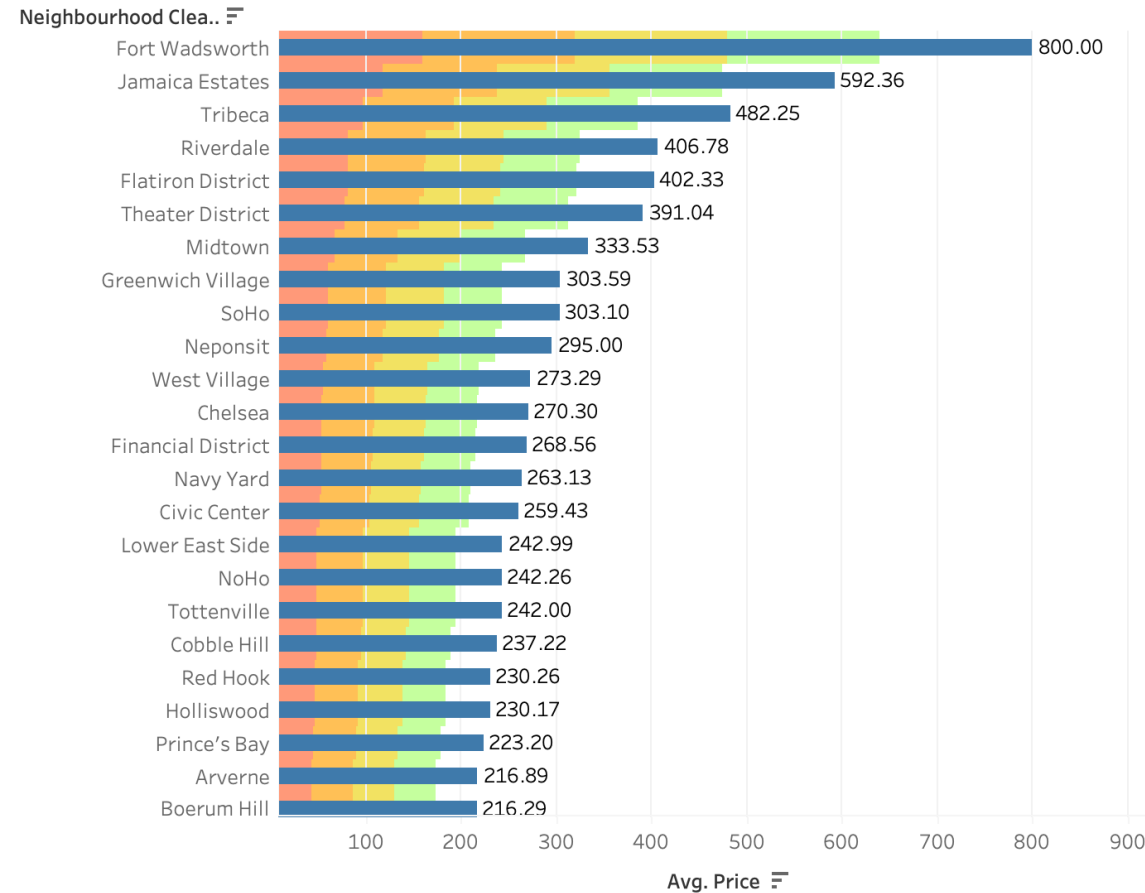


Red: Manhattan
Orange: Brooklyn
Light Blue: Queens
Dark Blue: Bronx
Green: Staten Island

Top N Most Expensive Neighborhoods

- We create a bar chart sorted by the average price by neighborhoods and colored by districts to show the top N most expensive neighborhoods in New York. “N” can be adjusted by the parameter “Top N most expensive”. We use 20 here as example.
- We add a “Rank” column upfront, so people can see the rank number directly when they use “filter” or “highlight” action in dashboard.
- Top 1 and 2 in the list is from Staten Island and Queens, but not from Manhattan, the most popular district. So we think that’s because they might be vacation house areas. The majority Neighborhood in this list is from Manhattan as expected, since Manhattan is always the most expensive area for housing.

Average Price Range by Neighborhood

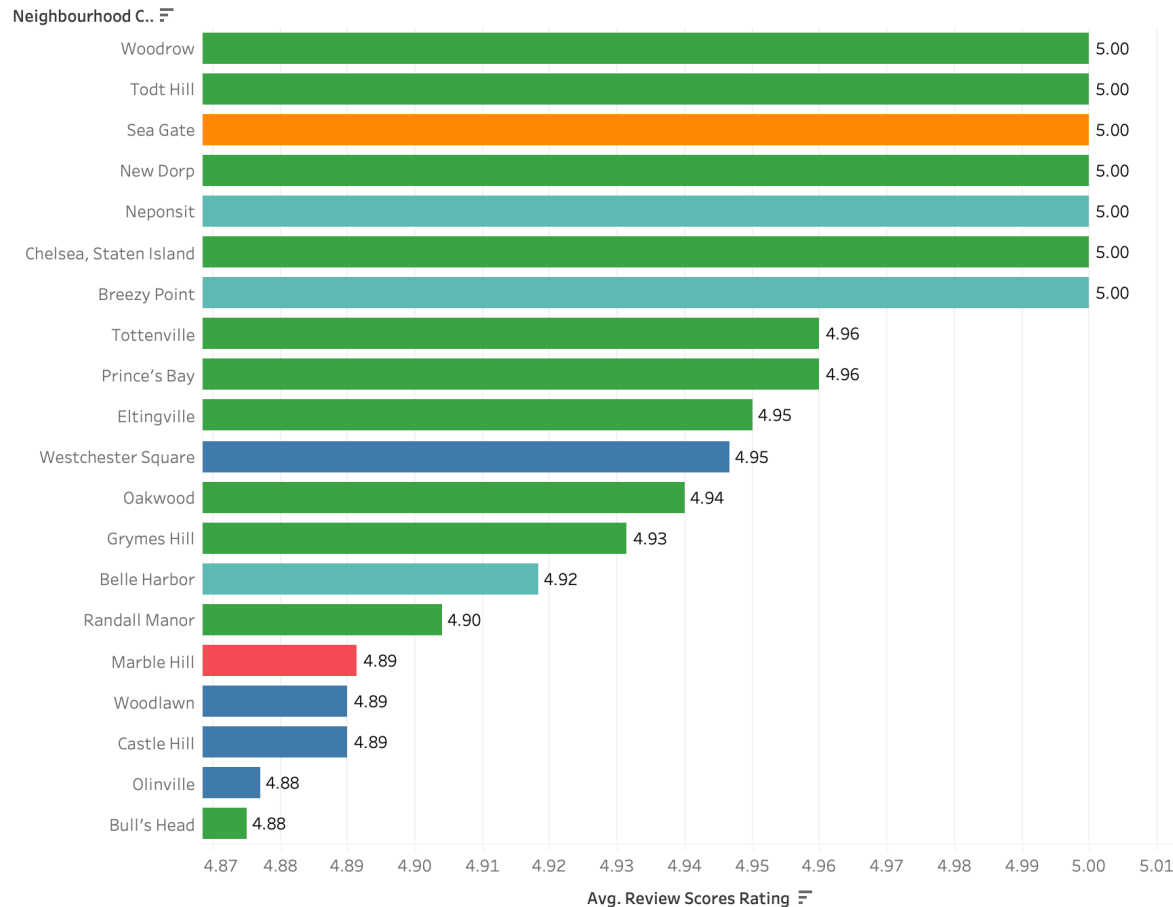


Average Price Range by Neighborhoods

- We created this bullet chart for the average price of each neighborhood, and filled the 20%, 40%, 60%, 80% of average price by different colors below.
- This chart can give us a detailed distribution of price for each neighborhood. As we have "filter" action on this chart, people can see the exact price and the distribution in this chart.



Top 20 Highest Review Score Neighborhood

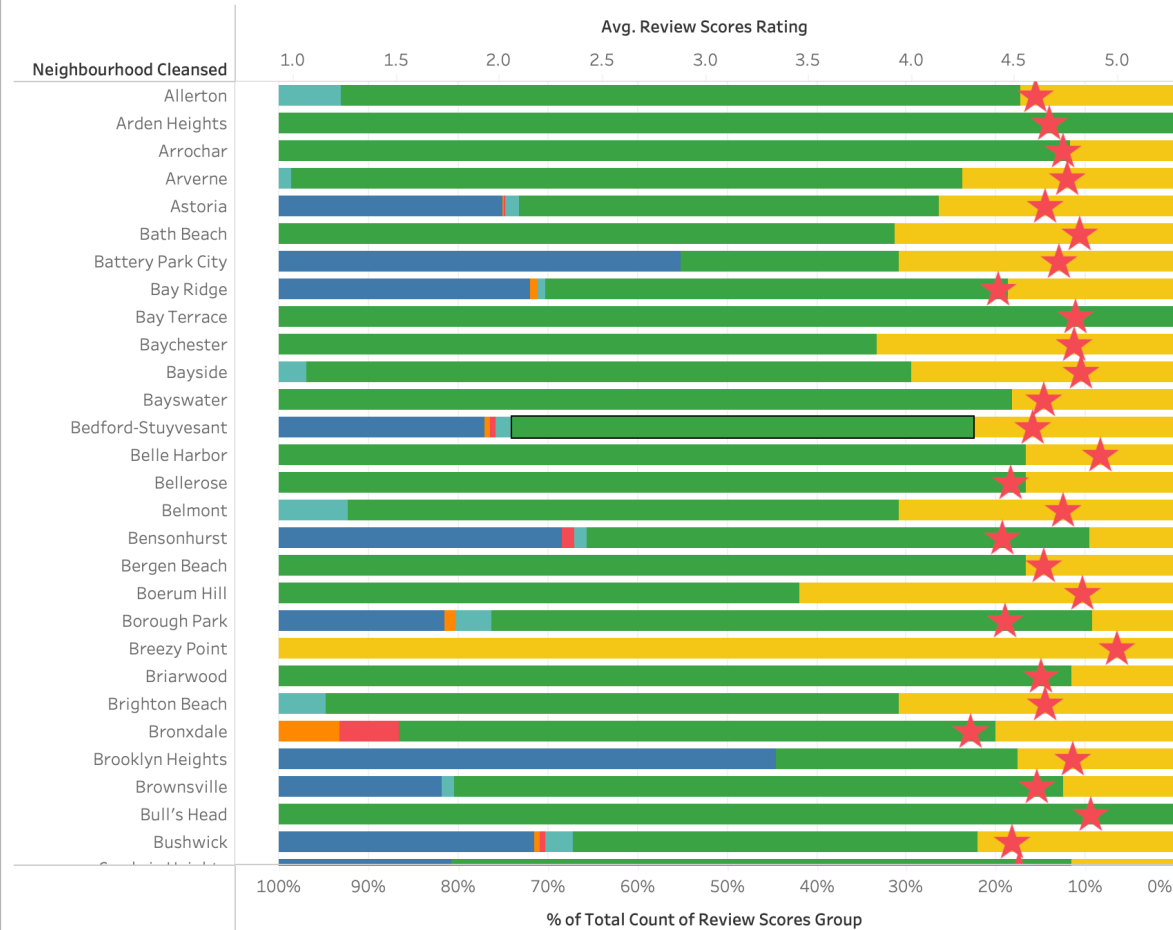


Red: Manhattan
Orange: Brooklyn
Light Blue: Queens
Dark Blue: Bronx
Green: Staten Island

Top N Highest Review Score Neighborhoods

- We created a bar chart for the average review score sorted by neighborhood and colored by district so that we can show the top N neighborhoods having the highest average review score. "N" can be adjusted by the parameter "Top N highest Score". We use 20 here as example.
- We can see that 11 out of top 20 highest review scores neighborhoods come from Staten Island, so we might think that Staten Island has some qualified Airbnb services there.

Likert Chart for Neighborhood



Likert Chart for Neighborhoods

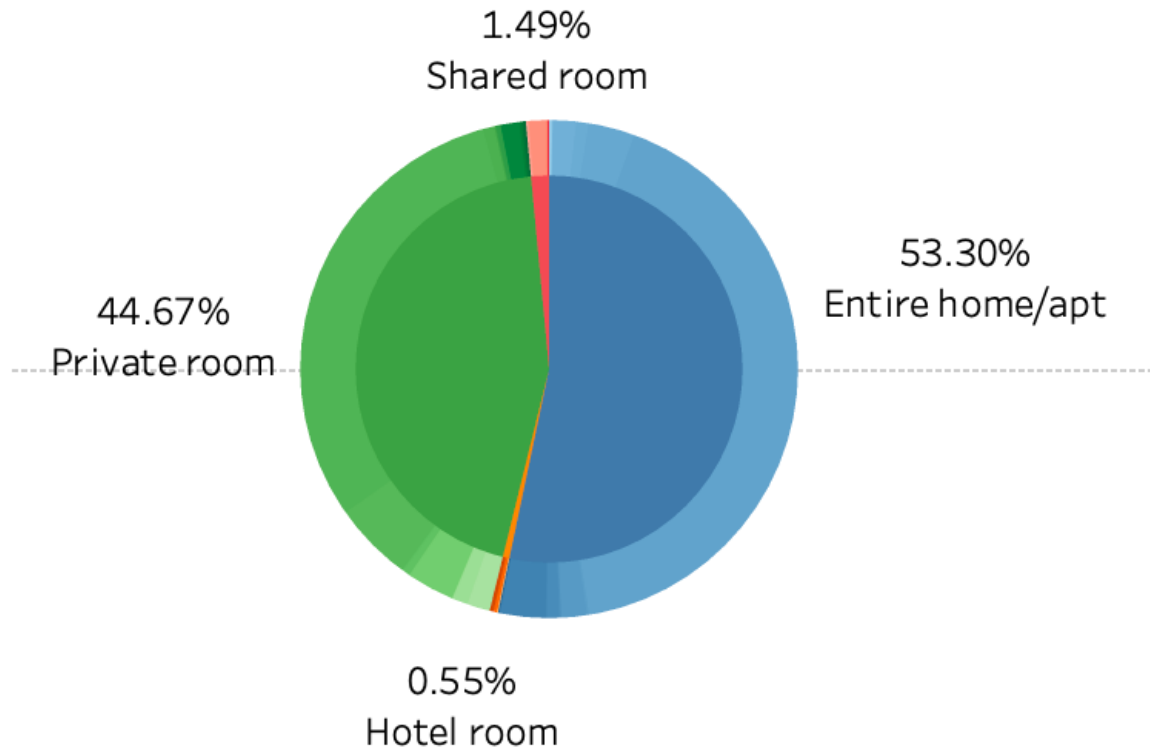
- We created a new calculated field based on the ones digit(i.e., if the score is 4.5, the new score will be 4).
- We first mark the average review scores for each neighborhood by the red stars. Then, we created a Likert chart to show that the percentage of each score of each score range.
- As we can see from the chart, 4 or 5 account for the majority of scores for many neighborhoods. Others have some 0 score. It is possible that people who gives rating online might extremely like or dislike the properties.



03 *Part Three*

Housing



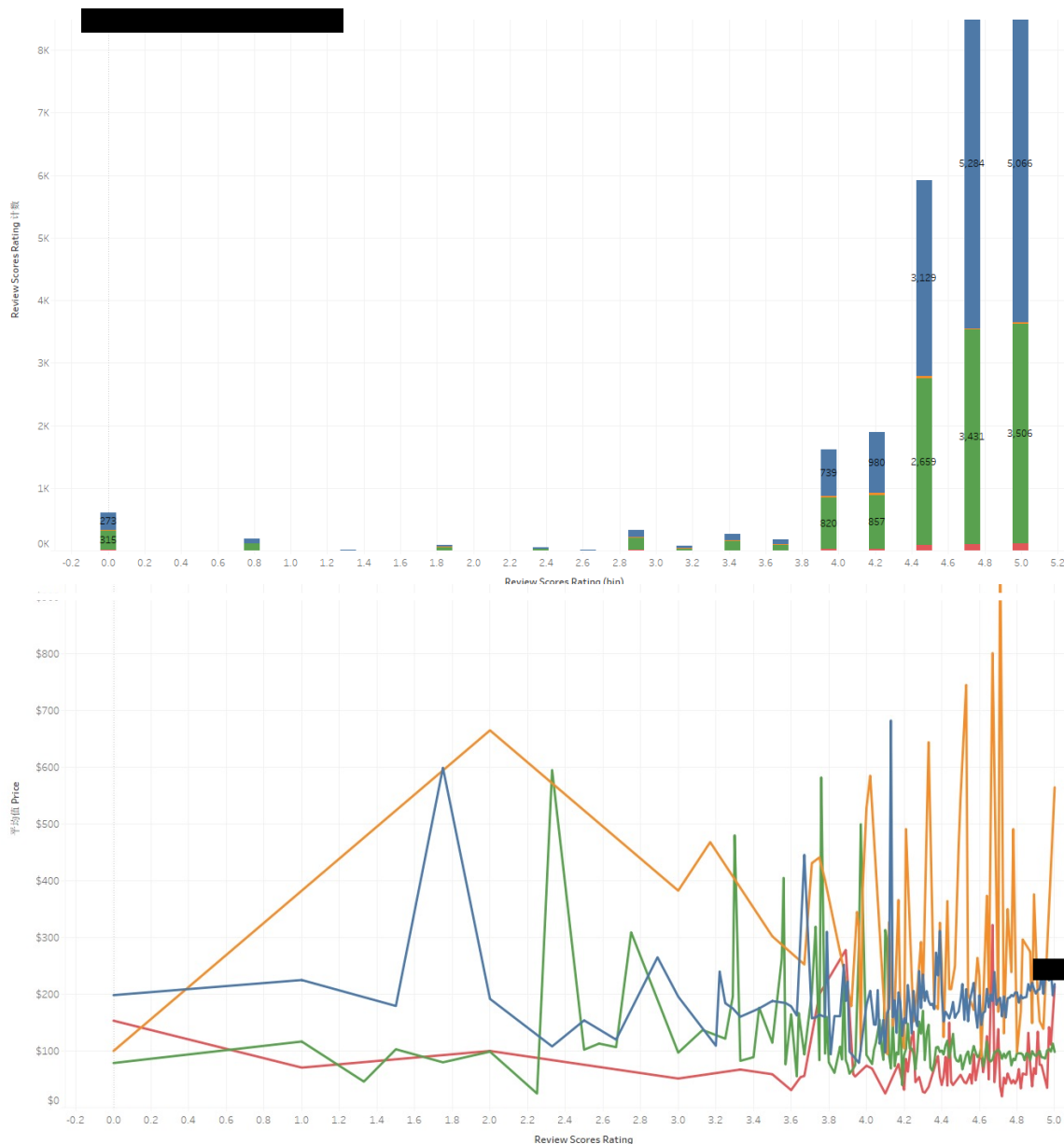


Blue: Entire home/apt
Orange: Hotel room
Green: Private room
Red: Shared room

Sunburst Chart of Type

- We first create a hierarchy with room type and property type because room type contains property type. Then, we create a sunburst chart for the number of houses by room type and property type.
- The majority of Airbnb room types are entire home/apt, accounting for 53.3%, followed by Private room, accounting for 44.67%.
- The largest number of property types are Entire home/apt rental units, accounting for 41.84% of the total, followed by Private room in rental units, accounting for 30.48%. This shows that the number of rental units in Airbnb is the largest.



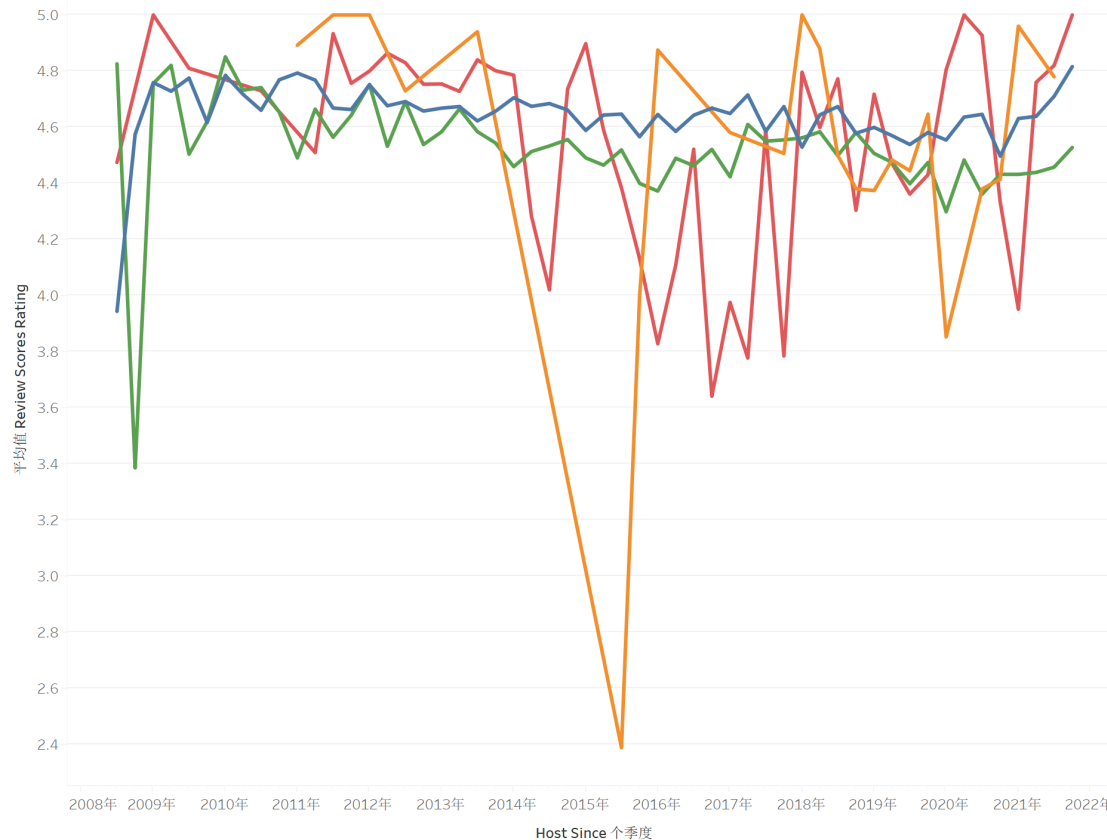


Number of rating for each room type & Score due to AVG Price

- For entire home/apt and private rooms, the average price of a high-scoring room is relatively stable, between \$200 and \$250. However, for shared rooms, the price of high-scoring rooms is the lowest among the four types of rooms, but relatively high among shared rooms. The price of hotel room fluctuates the most.
- Most of the rooms were rated 4.5 to 5 with the greatest number of scores being 4.7. This shows that most of the rooms are good and that the residents are giving the rooms good scores.

Blue: Entire home/apt
Orange: Hotel room
Green: Private room
Red: Shared room

Score based on the housing



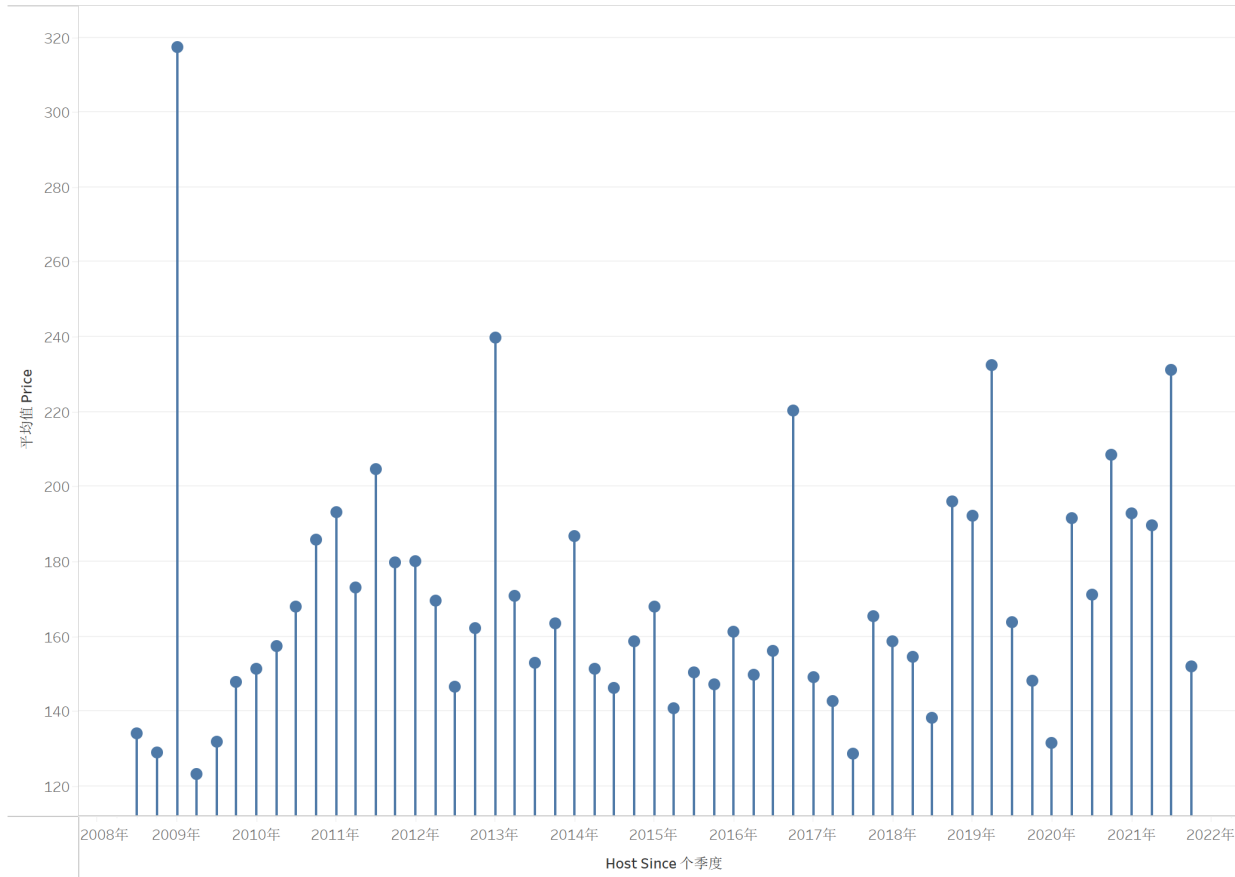
Blue: Entire home/apt
Orange: Hotel room
Green: Private room
Red: Shared room

Score based on the housing

- According to the date that the host joined Airbnb, Entire home/apt and Private room have more even scores without great ups and downs because of their high number. In general, Entire home/apt has a higher rating than Private room. On the other hand, Hotel room and Shared room have low number of houses, so when there is a low score, it leads to a big ups and downs in the average score. 2021, the overall average score has increased to some extent, except for Hotel room where the score has decreased in the fourth quarter of 2021.



AVG Price Based on the Host Time



AVG Price Based on the Host Time

- The average rate fluctuates from quarter to quarter, except for the first quarter of 2009, when the average rate was below \$250, indicating that the overall Airbnb rate in New York is still reasonable. During the year, the average price is usually the lowest in the third quarter, probably also because there are not many holidays, and it is not a peak travel period.

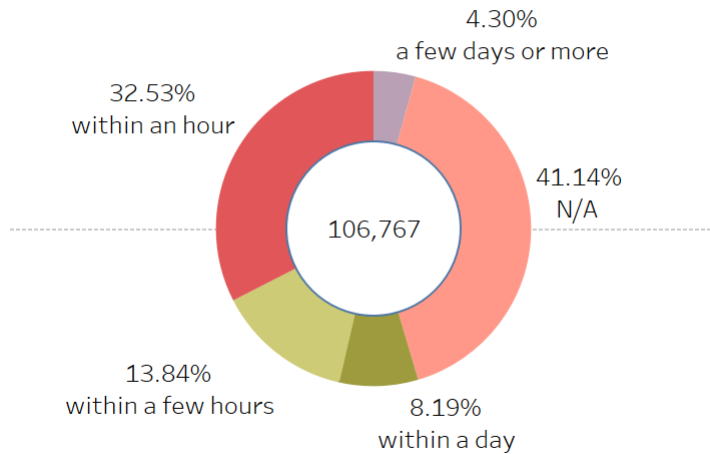


04 *Part Four*

Host



Host response time and host accommodate

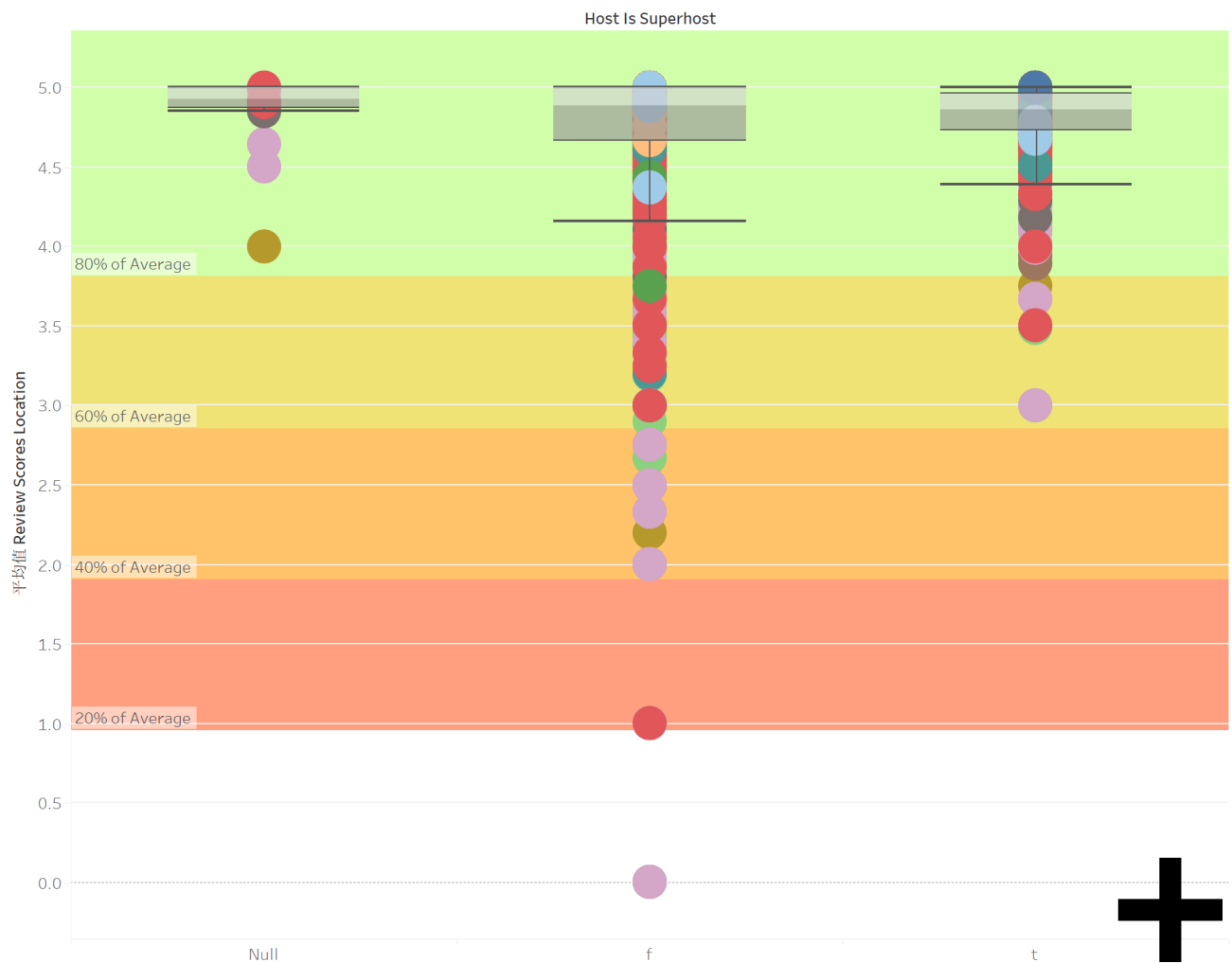


Purple: a few days or more
Pink: N/A
Green: within a day
Yellow: within few hours
Red: within an hour

- First, we use pie chart to find the correlation between accommodation and host response speed. From the result, we find that response speed within an hour have the most popular customer volume. The slowest simultaneous message return time is probably a few days or more, and this response speed results in them having the fewest number of clients. From these data we can conclude that the faster hosts reply to messages, the more promptly they solve customers' problems, the more customers will choose them.



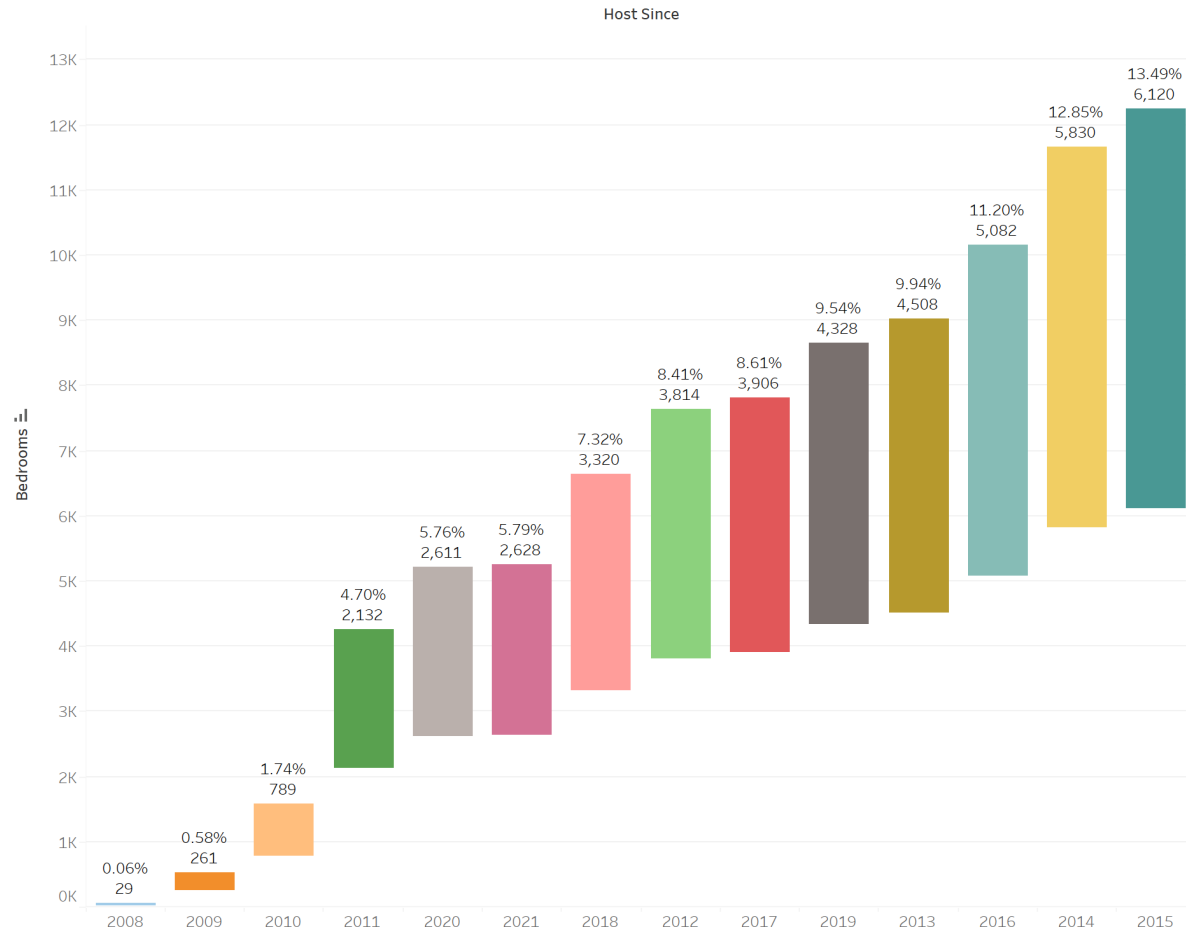
Superhost and location score



Superhost and location score

- Second, we use box-and-whisker plot to determine the relationship between whether the host is a superhost and the location score distribution. Also use different colors to distinguish the types of Airbnb properties. We find that the average location score of superhost is more concentrated than the average location score of non-superhost. The majority of superhost's location scores are in the 80%average range, while the majority of superhosts are in the 60% average range.

Year since host join Airbnb and bedroom numbers



Year since host join Airbnb and bedroom numbers

- Final, we use Gantt Bar to calculate the correlation between the year of host join Airbnb and the bedrooms number each year has. Hosts who joined Airbnb in 2015 had 14.32% of all rooms, with 6,296 rooms, followed by 14.04% in 2014. This shows that Airbnb is growing better in 2015 and 2014, with more hosts joining and providing close to 30% of the rooms.



05 *Part Five*

Data Insight





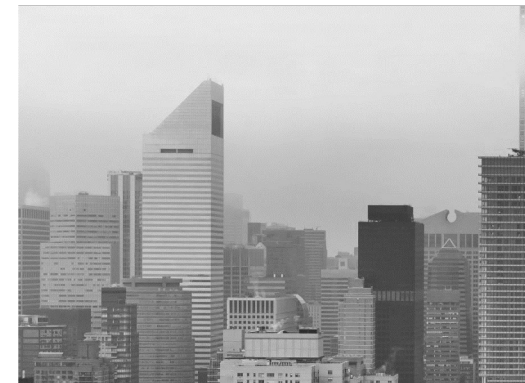
Neighborhood

- Manhattan and Brooklyn are the most popular Airbnb Locations.
- Manhattan has the most expensive neighborhoods. And the Airbnb service in the Staten Island gained lots of praise.



Housing

- Entire home/apt and private room have the highest numbers, and the third quarter is usually the lowest vacation quarter of the year.



Host

- Host improve the response speed can get more customer resources
- The geographical location has a certain influence on becoming a Superhost



Thank You

