



# AirBed&Breakfast

Book Rooms With Locals, Rather than Hotels.

Author: Mei Yee Chong  
Date: 2022/09





# About Airbnb

Airbnb, Inc is an American company that operates an **online marketplace** for lodging, primarily homestays for vacation rentals, and tourism activities.



Founded in 2008, based in San Francisco, California, the platform is accessible via **website** and **mobile app**.



Airbnb does not own any of the listed properties; instead, it profits by **receiving commission** from each booking.



Now, it has over **1.5 million** listings in **34,000** cities and **190** countries.



# About dataset

## Context:

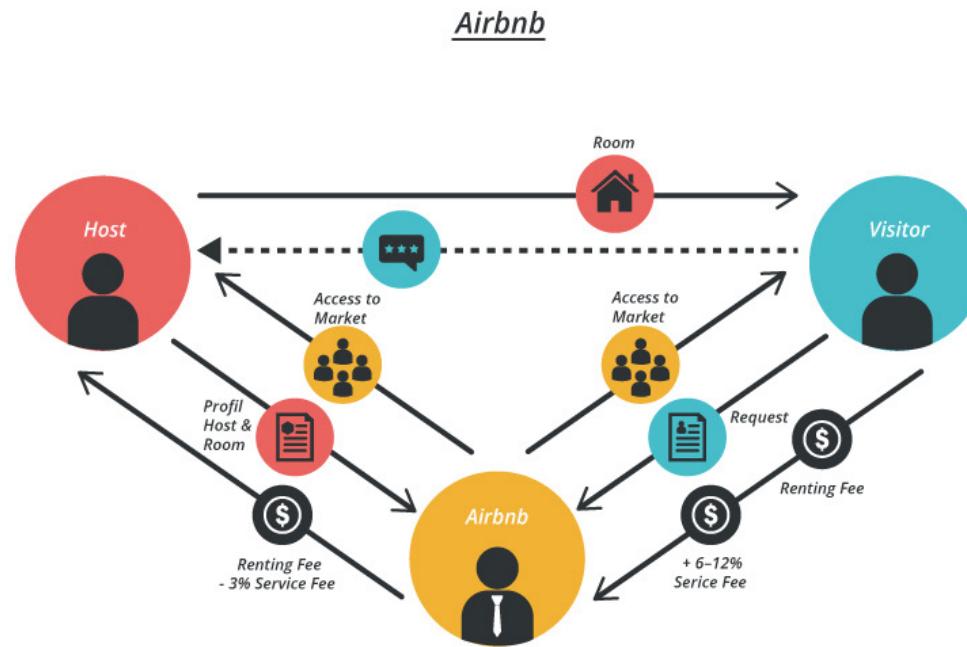
Since 2008, guests and hosts have used Airbnb to travel in a more unique, personalized way. As part of the Airbnb Inside initiative, this dataset describes the listing activity of homestays in New York City.

## Content:

- The following Airbnb activity is included in this New York dataset.
- Listings, including full descriptions and average review score Reviews, including unique id for each reviewer and detailed comments Calendar, including listing id and the price and availability for that day.

# Questions

- To figure out the feature of listings of Airbnb in New York City from this dataset.
- What kind of factors can decide the price of listings?





# Data Cleaning

Python



# Data cleaning

- Check & remove duplicates columns
  - Remove the incorrect data
  - Check & remove NA
  - Drop the redundant columns
  - Remove dollar sign
- 
- Github: <https://github.com/chongmeiyee/Data-Analysis-Project-Airbnb.git>



# Analyse

Python



# Descriptive Statistics

In [23]:

```
df.describe()
```

Out[23]:

	<b>id</b>	<b>host id</b>	<b>lat</b>	<b>long</b>	<b>Construction year</b>	<b>price</b>	<b>service fee</b>	<b>minimum nights</b>	<b>number of reviews</b>	<b>review rate number</b>	<b>calculated host listing since</b>
<b>count</b>	1.001470e+05	1.001470e+05	100147.000000	100147.000000	100147.000000	100147.000000	100147.000000	100147.000000	100147.000000	100147.000000	100147.000000
<b>mean</b>	2.928002e+07	4.925791e+10	40.728054	-73.949638	2012.487603	625.445086	125.089878	8.114921	27.335677	3.280448	7
<b>std</b>	1.622241e+07	2.853783e+10	0.055862	0.049548	5.763301	331.668555	66.337094	30.712162	49.268821	1.284300	3
<b>min</b>	1.001254e+06	1.236005e+08	40.499790	-74.249840	2003.000000	50.000000	10.000000	-1223.000000	0.000000	1.000000	1
<b>25%</b>	1.523217e+07	2.458389e+10	40.688700	-73.982590	2007.000000	340.000000	68.000000	2.000000	1.000000	2.000000	1
<b>50%</b>	2.930837e+07	4.911774e+10	40.722260	-73.954450	2012.000000	625.000000	125.000000	3.000000	7.000000	3.000000	1
<b>75%</b>	4.333596e+07	7.398688e+10	40.762760	-73.932320	2017.000000	913.000000	183.000000	5.000000	30.000000	4.000000	2
<b>max</b>	5.736024e+07	9.876313e+10	40.916970	-73.705220	2022.000000	1200.000000	240.000000	5645.000000	1024.000000	5.000000	332



# Correlation

In [24]:

```
df.corr()
```

Out[24]:

	<b>id</b>	<b>host id</b>	<b>lat</b>	<b>long</b>	<b>Construction year</b>	<b>price</b>	<b>service fee</b>	<b>minimum nights</b>	<b>number of reviews</b>	<b>review rate number</b>	<b>calculated host listings count</b>
<b>id</b>	1.000000	-0.001249	-0.009004	0.041173	0.000452	0.007246	0.007247	0.005697	-0.036371	0.035916	0.022915
<b>host id</b>	-0.001249	1.000000	0.000802	-0.008458	0.004548	0.003463	0.003445	-0.002212	-0.005042	0.004002	0.001517
<b>lat</b>	-0.009004	0.000802	1.000000	0.073841	0.005272	-0.004374	-0.004347	0.015285	-0.025225	-0.003467	0.032357
<b>long</b>	0.041173	-0.008458	0.073841	1.000000	0.001501	0.003564	0.003558	-0.039433	0.070669	0.014364	-0.105283
<b>Construction year</b>	0.000452	0.004548	0.005272	0.001501	1.000000	-0.004070	-0.004048	-0.000154	0.002099	0.005294	-0.002344
<b>price</b>	0.007246	0.003463	-0.004374	0.003564	-0.004070	1.000000	0.999991	-0.003491	0.004763	-0.005172	-0.001001
<b>service fee</b>	0.007247	0.003445	-0.004347	0.003558	-0.004048	0.999991	1.000000	-0.003473	0.004742	-0.005176	-0.000992
<b>minimum nights</b>	0.005697	-0.002212	0.015285	-0.039433	-0.000154	-0.003491	-0.003473	1.000000	-0.049967	-0.002542	0.083976
<b>number of reviews</b>	-0.036371	-0.005042	-0.025225	0.070669	0.002099	0.004763	0.004742	-0.049967	1.000000	-0.017656	-0.080612
<b>review rate number</b>	0.035916	0.004002	-0.003467	0.014364	0.005294	-0.005172	-0.005176	-0.002542	-0.017656	1.000000	0.024292
<b>calculated host listings count</b>	0.022915	0.001517	0.032357	-0.105283	-0.002344	-0.001001	-0.000992	0.083976	-0.080612	0.024292	1.000000



# Correlation Matrix





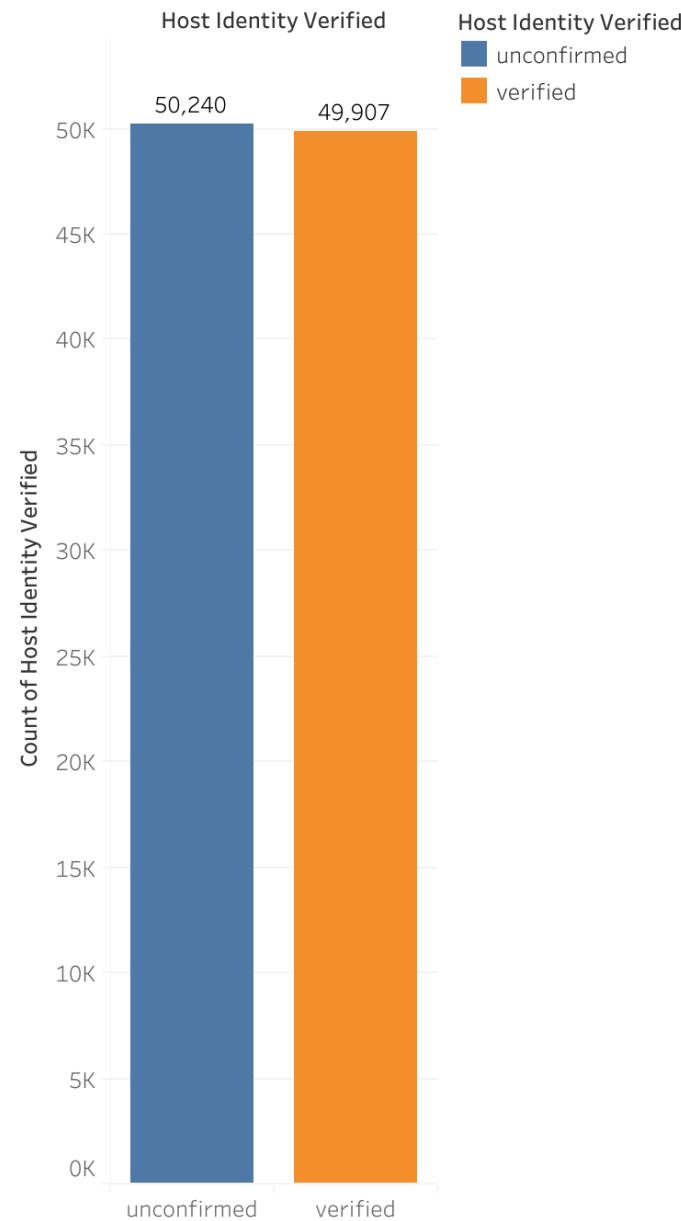
# Visualization

Tableau



## Count of Host Identity

Verified

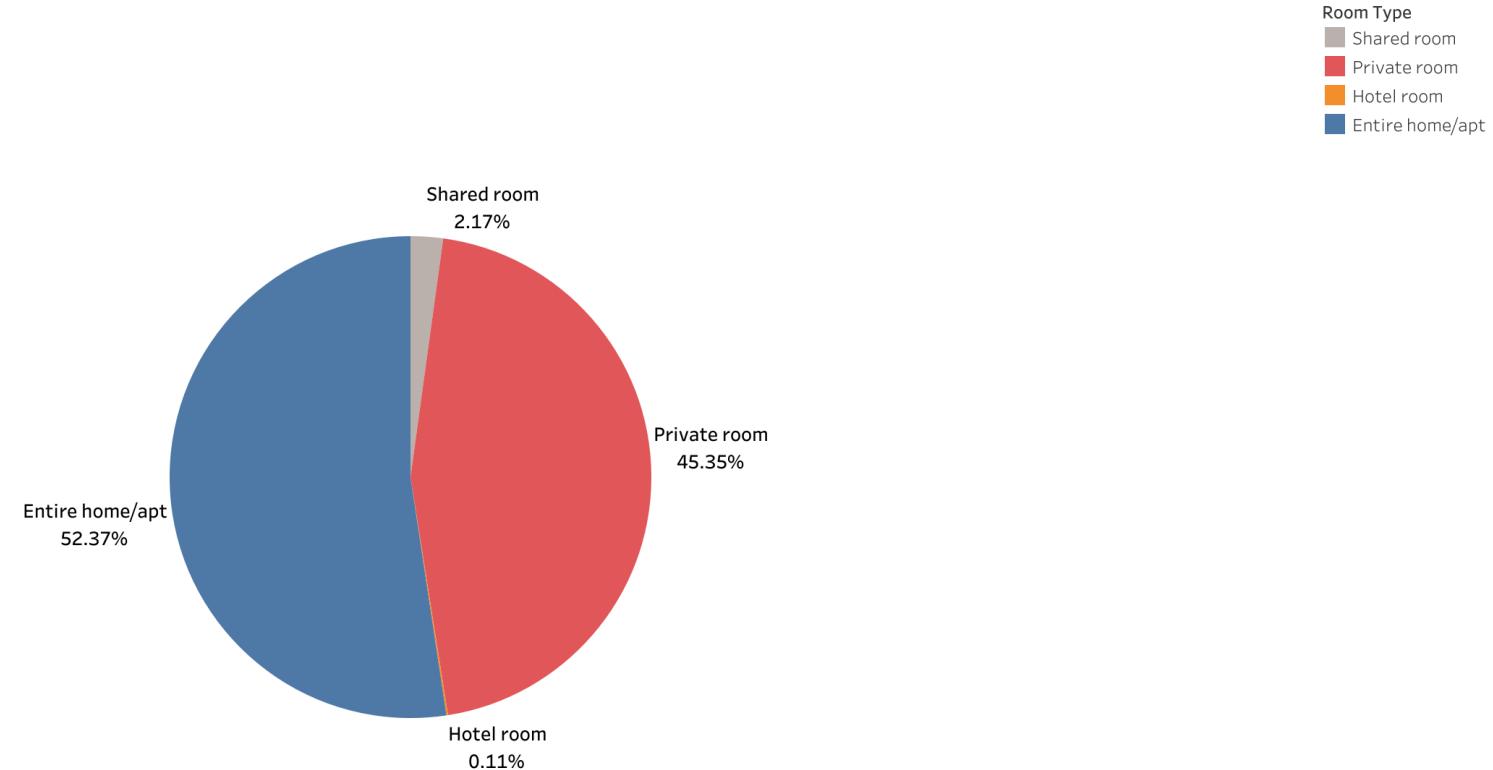


There are many host in Airbnb didn't confirm their identity.



# Room Type distribution

Room Type distribution

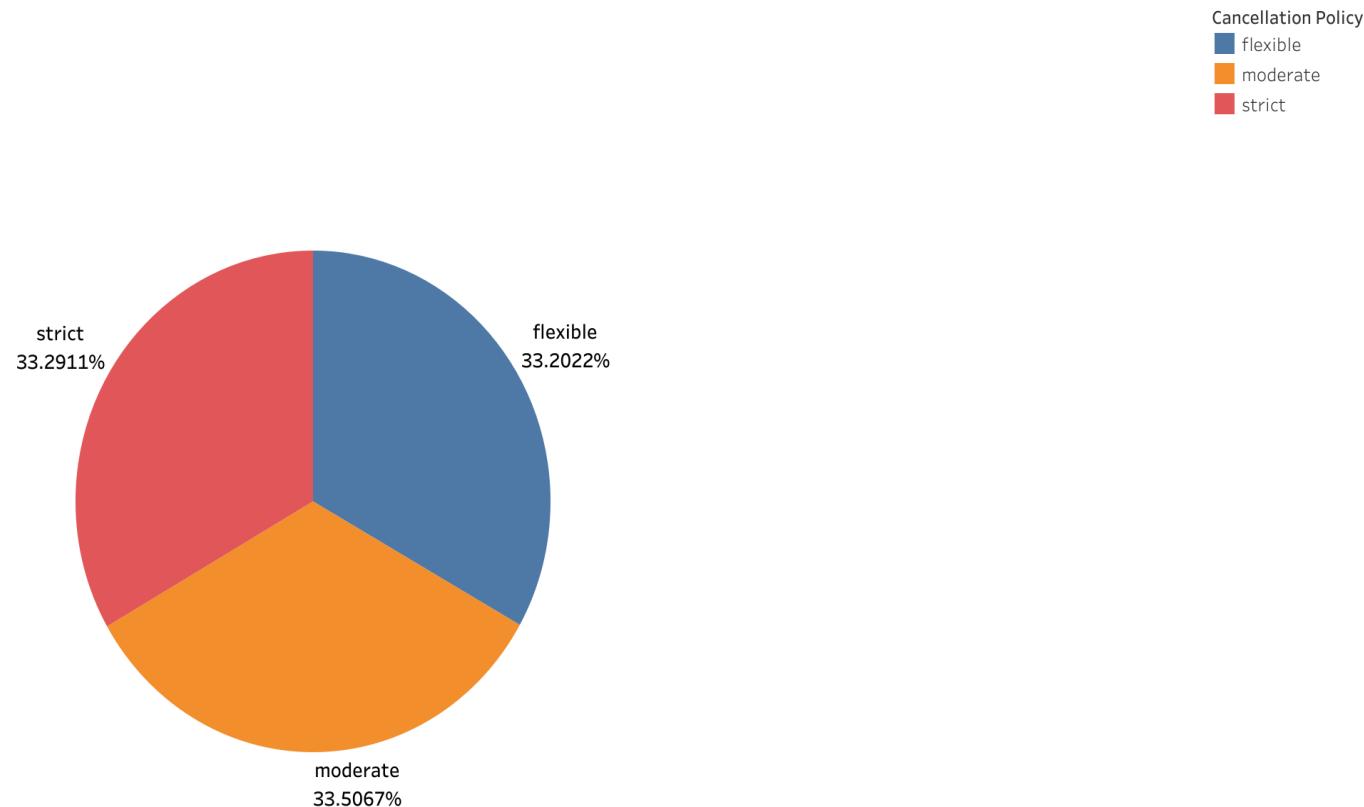


- There are **52%** of entire home or apartment and **45%** of private room offered in Airbnb.
- Host provide a few of shared room and hotel room in Airbnb platform.



# Cancellation Policy distribution

Cancellation Policy distribution

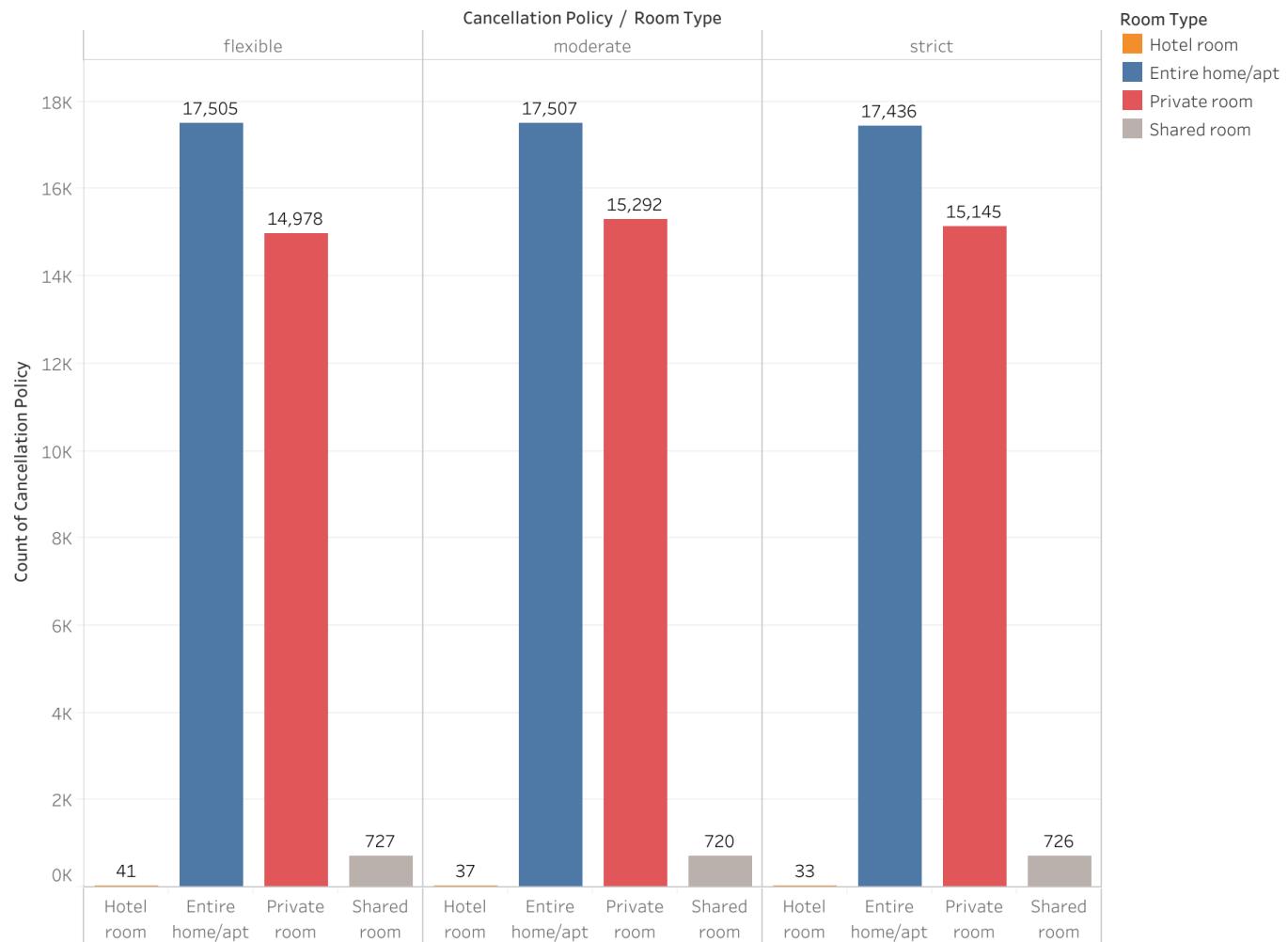


The distribution of cancellation policy are almost the same in Airbnb.



# How does the Cancellation Policy varies across room type?

Count of Cancellation Policy across different Room Type

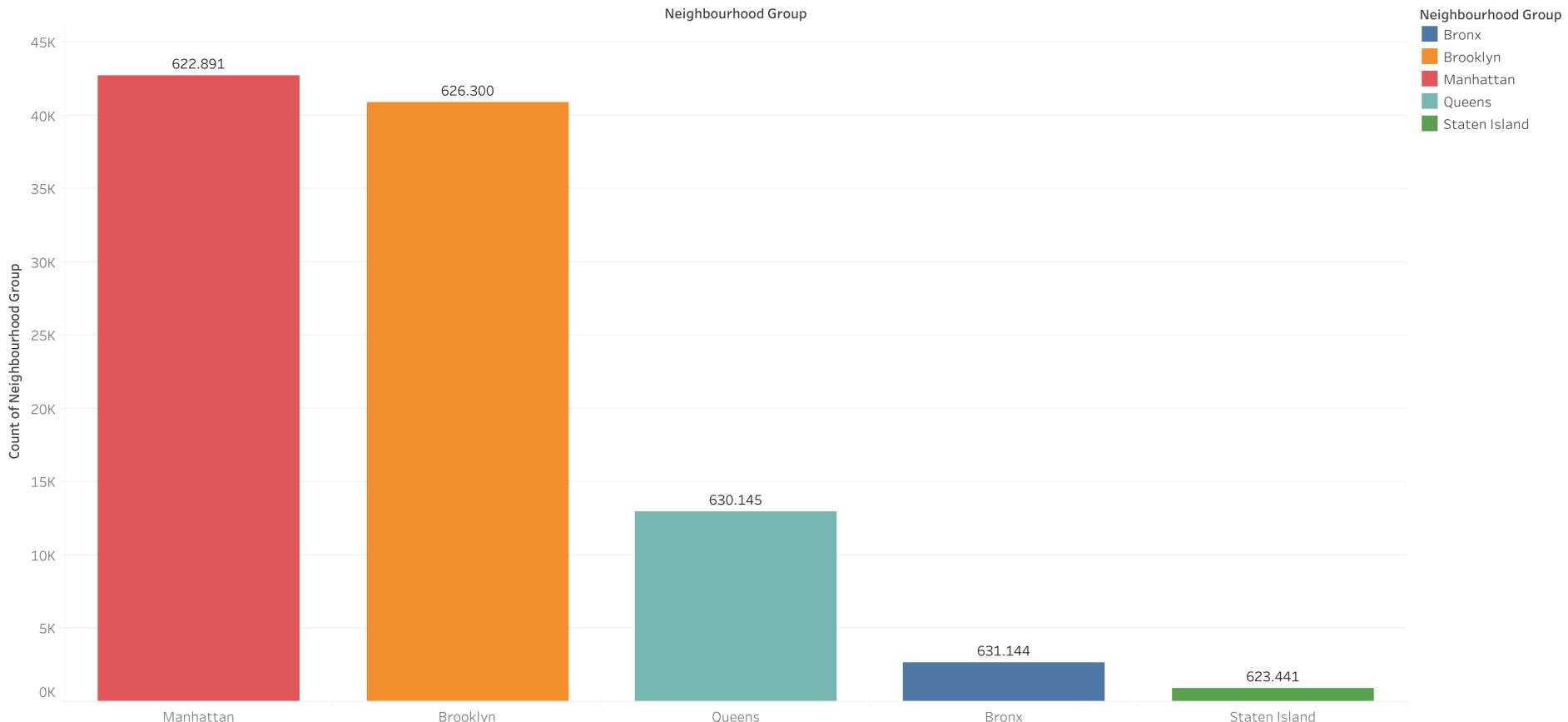


The cancellation policy are uniformly distributed across different room type.



# Neighbourhood group distribution and average price

Most popular neighbourhood group

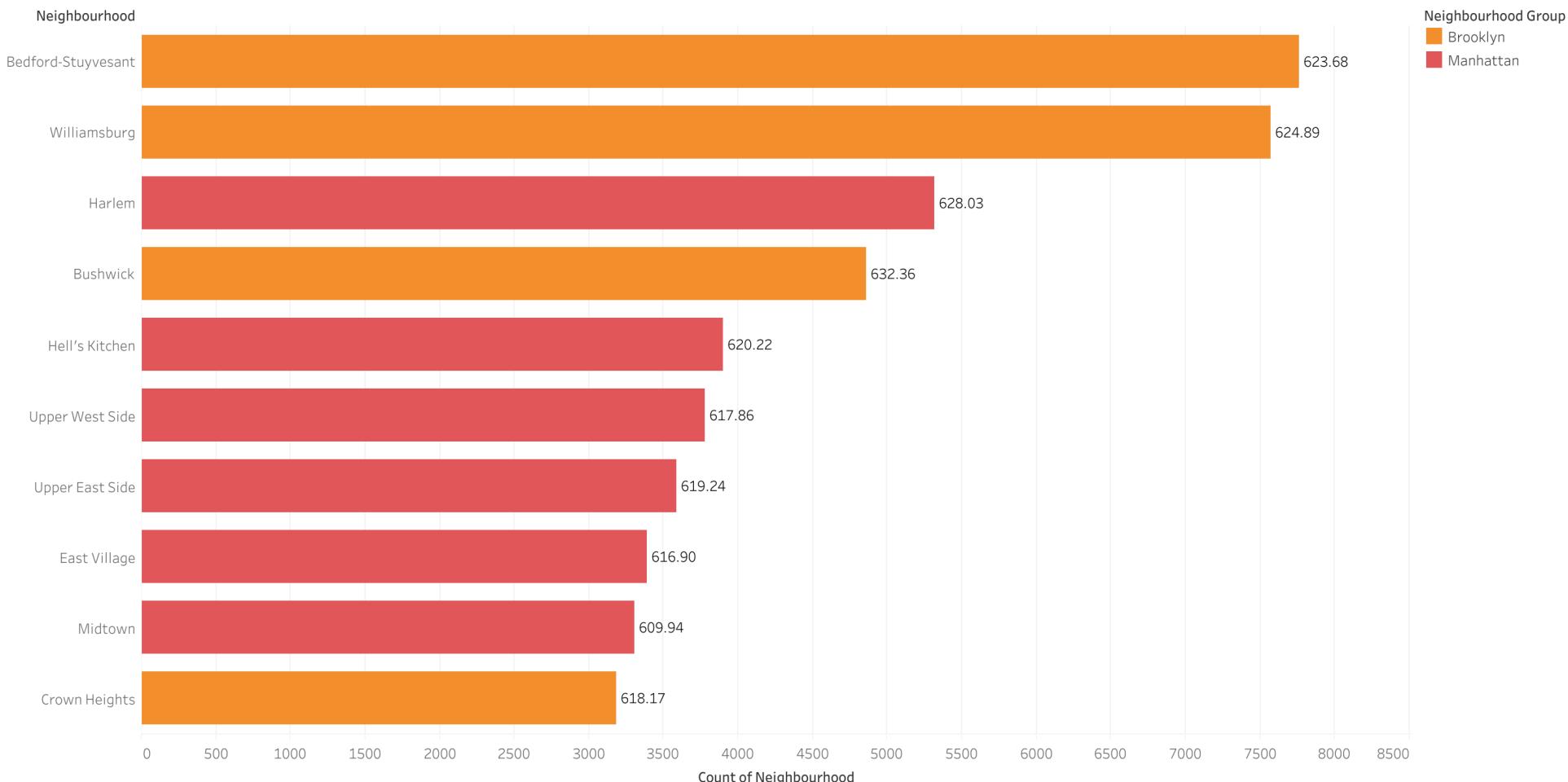


- The neighbourhoods of listings in New York are classify into five group which is Manhattan, Brooklyn, Queens, Bronx and Staten Island.
- Most of the listings of Airbnb in New York are distributed into **Manhattan and Brooklyn**.
- The **average price** of these neighbourhood groups are around **\$630**.



# Neighbourhood distribution and average price

Most popular neighbourhood

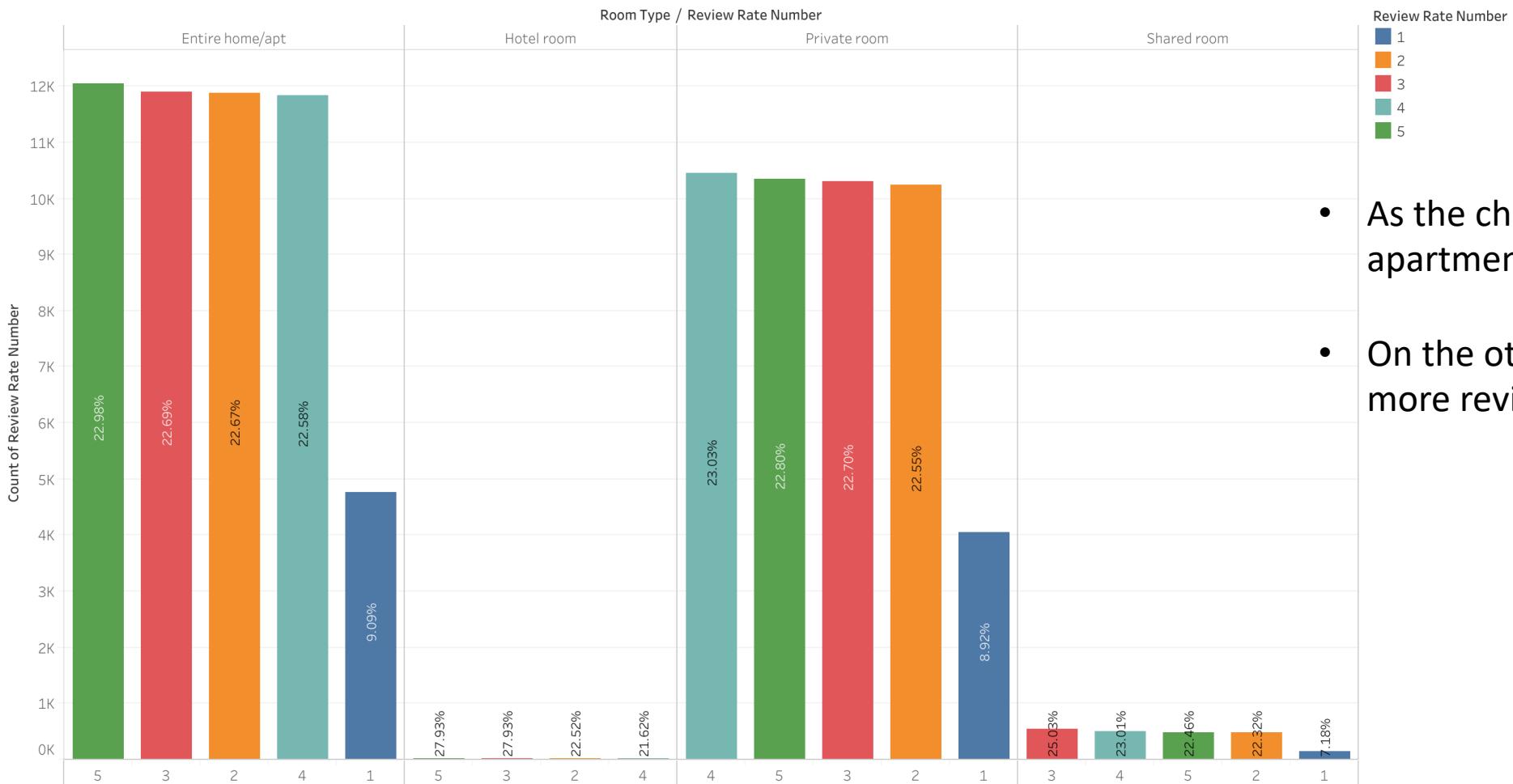


- The horizontal bar chart above shows the Top 10 popular neighbourhoods of listings in New York.
- Consistent with the observation of the previous slide, these popular neighbourhoods belong to Brooklyn and Manhattan.



# How does the review rate varies across room type?

Count of Review Rate Number across different Room Type

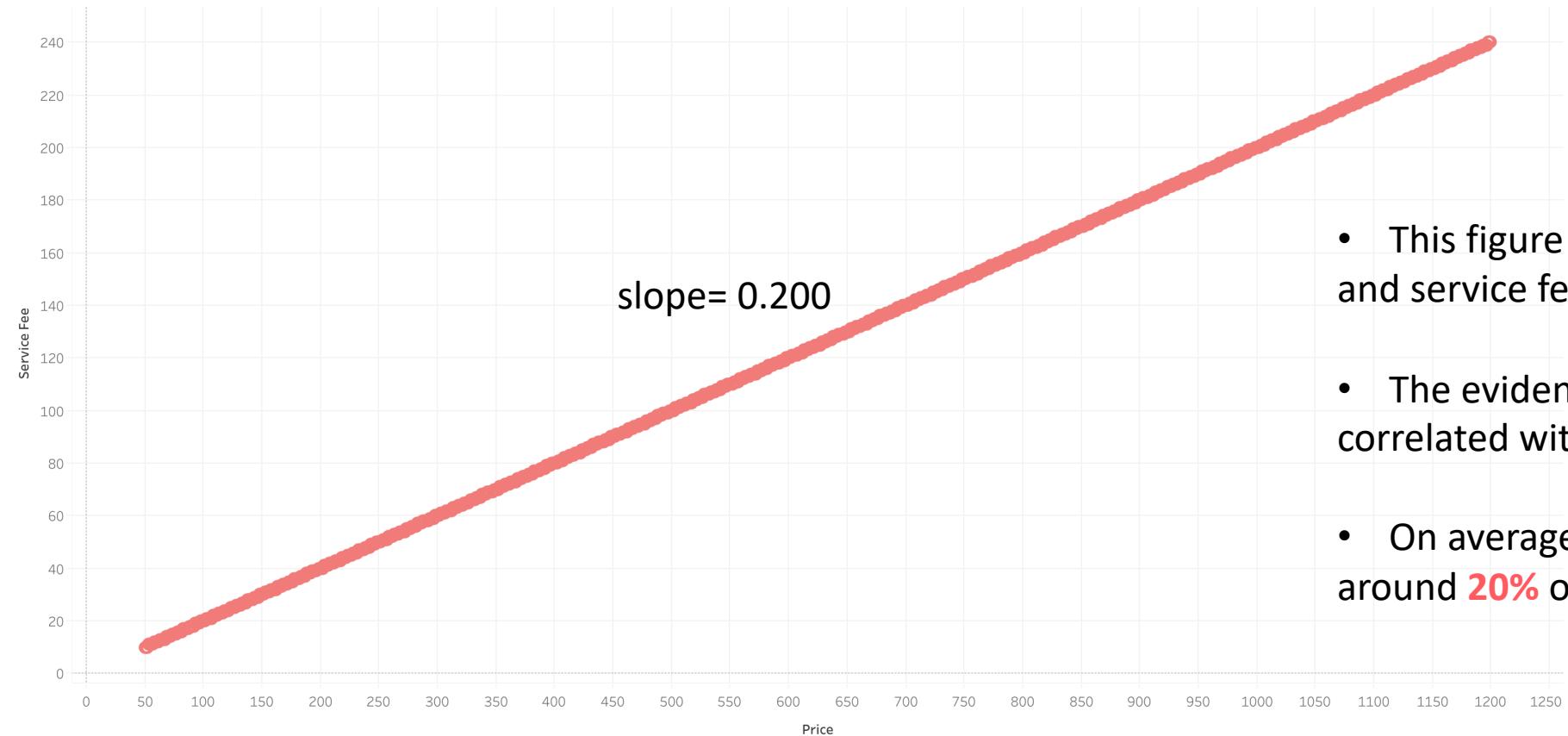


- As the chart shows, entire home or apartment receive more review rate of five.
- On the other hand, private room receive more review rate of four.



# Does price affect the service fee?

Relation btw Price and Service Fee

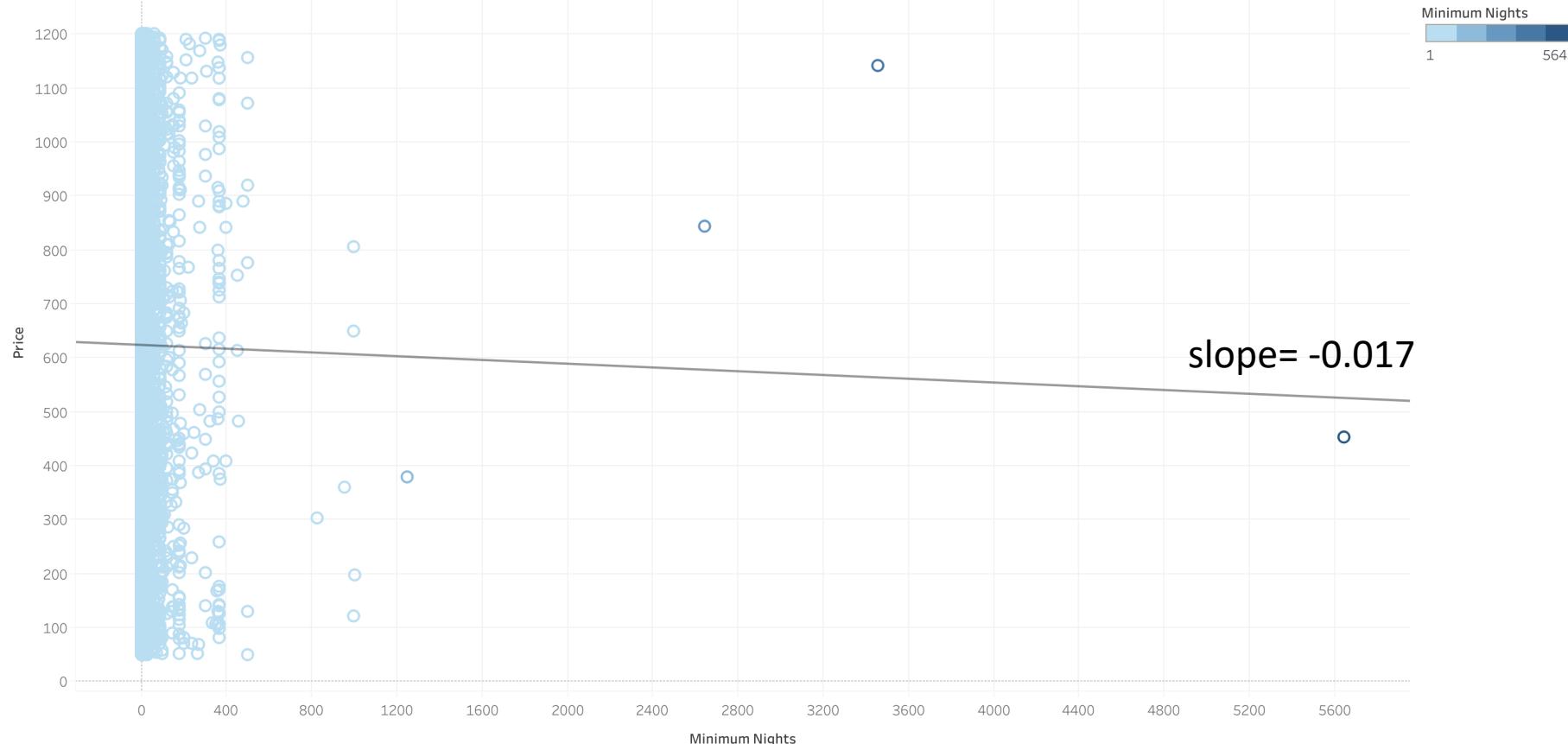


- This figure represents the scatter plot of price and service fee.
- The evidence shows that the price is highly correlated with the service fee.
- On average, airbnb will charge service fee around **20%** of the price of the listings.



# Does minimum nights affect the price?

Relation btw Minimum Nights and Price

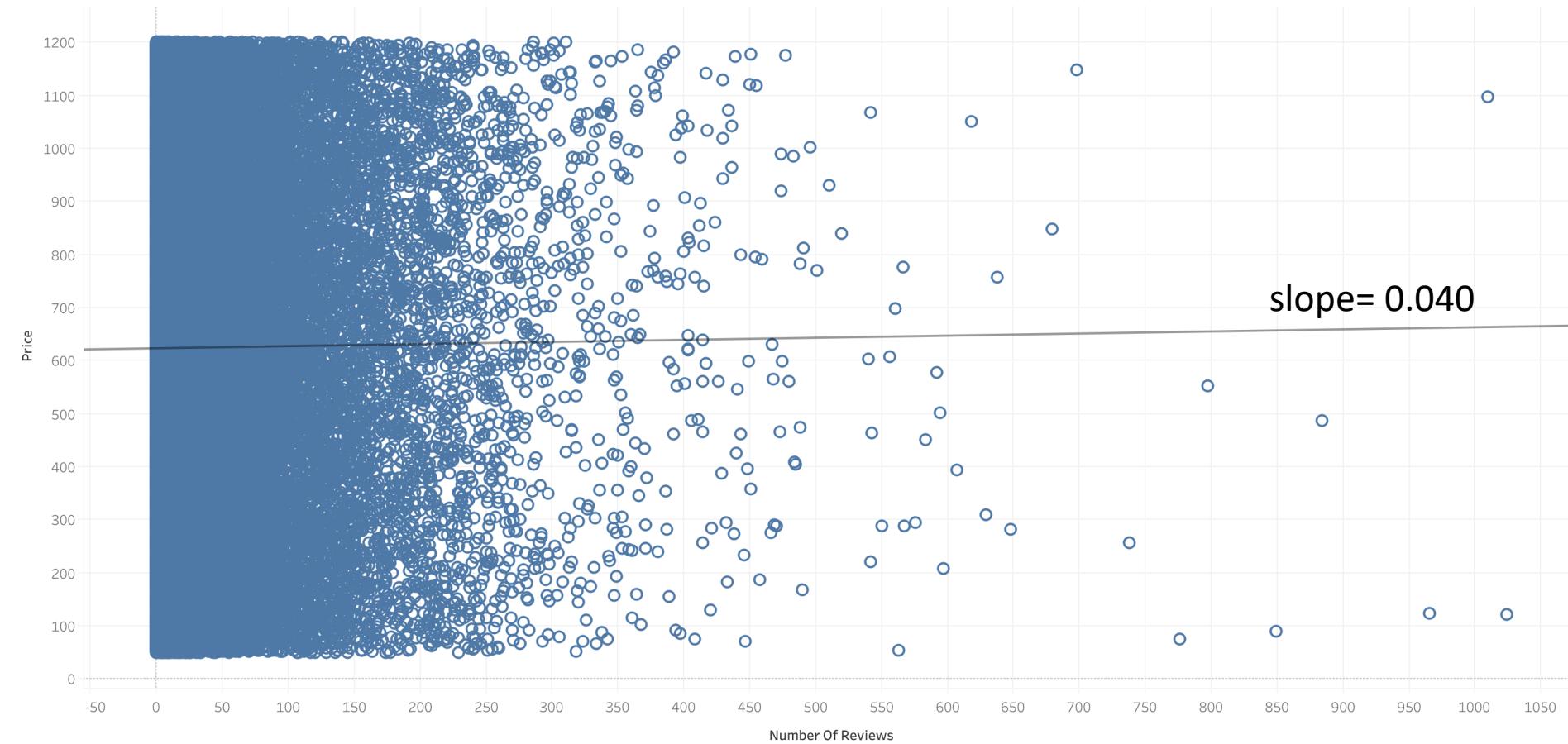


- The graph above represents the scatter plot of price and minimum number of night stay for the listings.
- The trend line is negative sloped, which means that **minimum nights is negatively correlated with the price**.



# Does number of reviews affect the price?

Relation btw Number of Reviews and Price

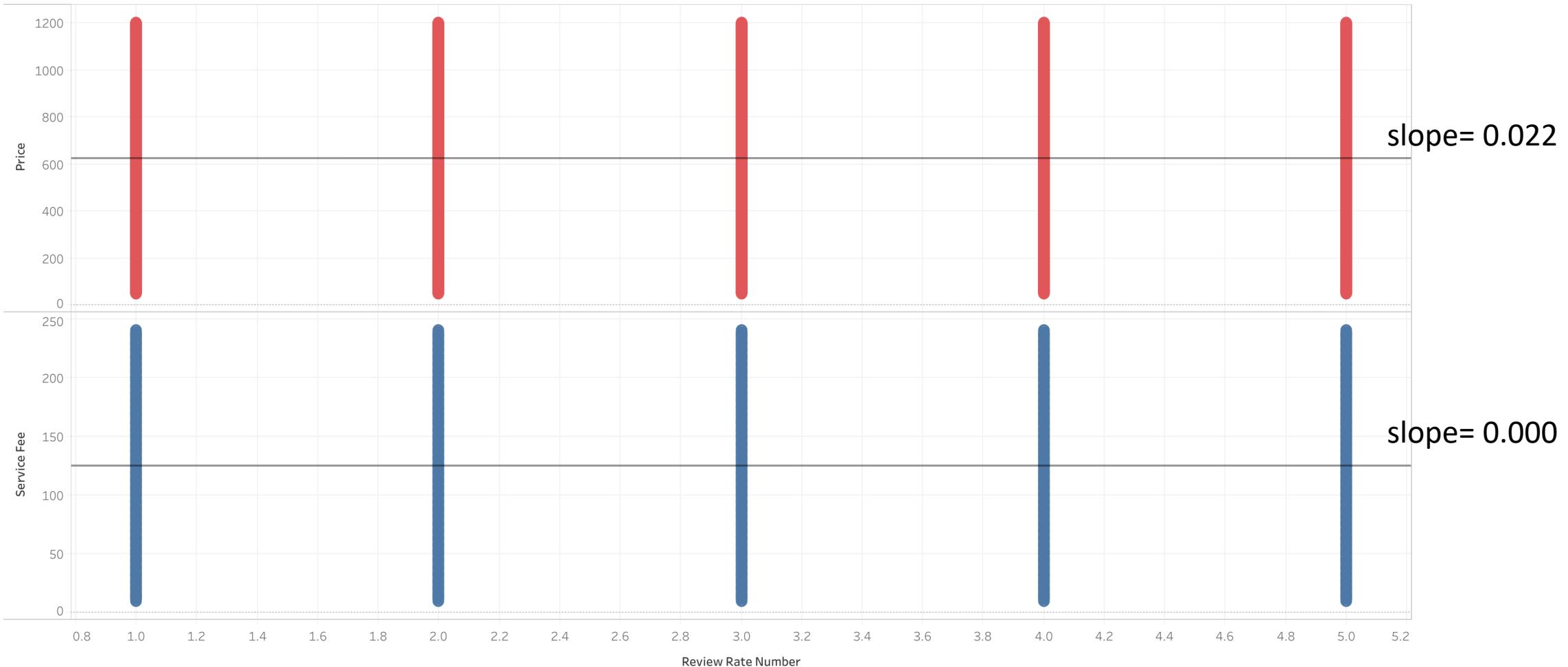


- The graph above represents the scatter plot of price and the number of reviews the listing has.
- The trend line is positive sloped, which means that **the number of reviews is positively correlated with the price**.



# Does review rate affect the price/service fee?

Relation btw review rate, price and service fee



- The trend line of price has a positive slope while has no significant trend for service fee.
- In other words, **higher review rate of listings** of Airbnb may **positively associated with its price**.



## Airbnb in New York

# Dashboard-1

#listings: 1,00,147

Avg.price: \$625/day

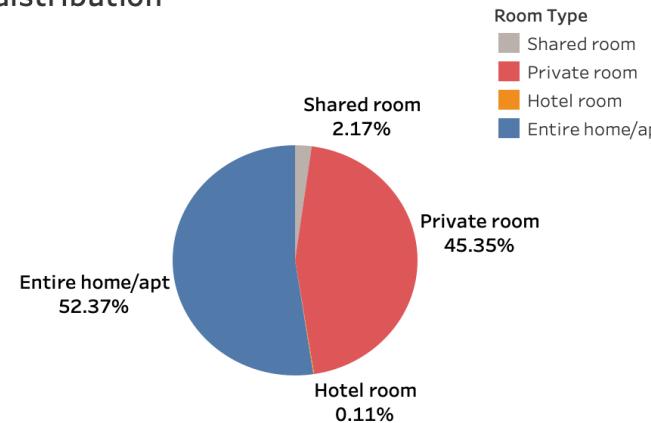
popular neighbourhood: Brooklyn&Manhattan

Room Type (全部)

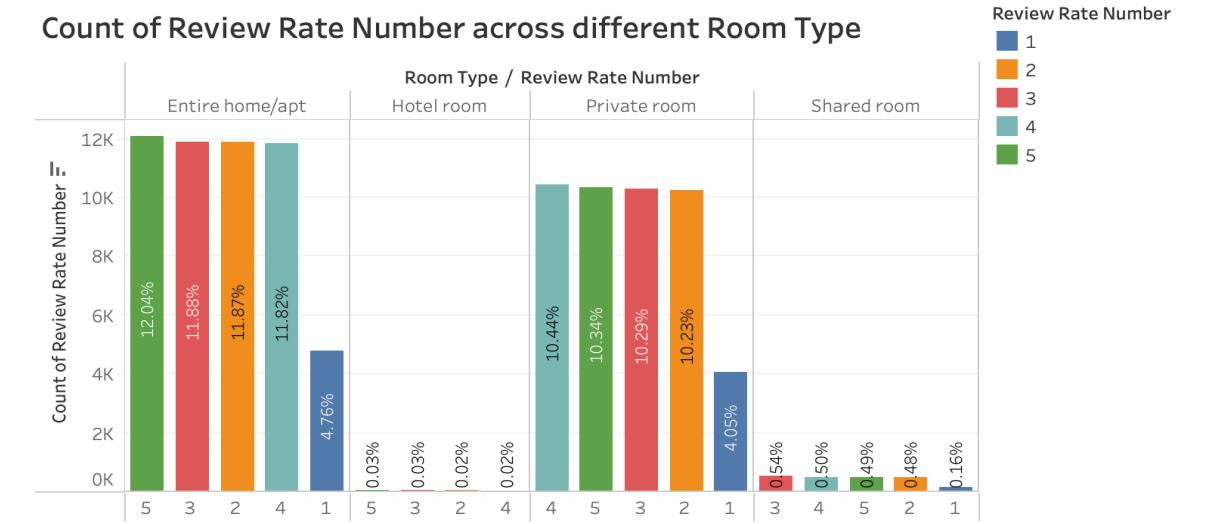
Cancellation.. (全部)

Neighbour.. (全部)

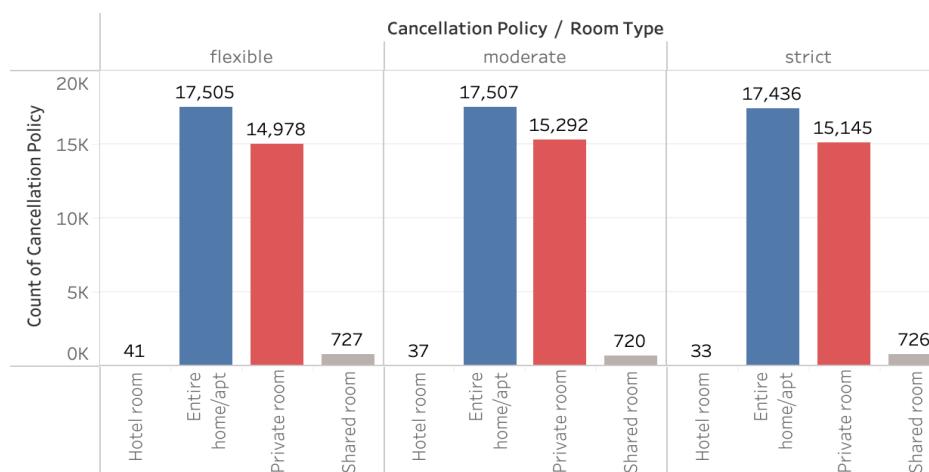
Room Type distribution



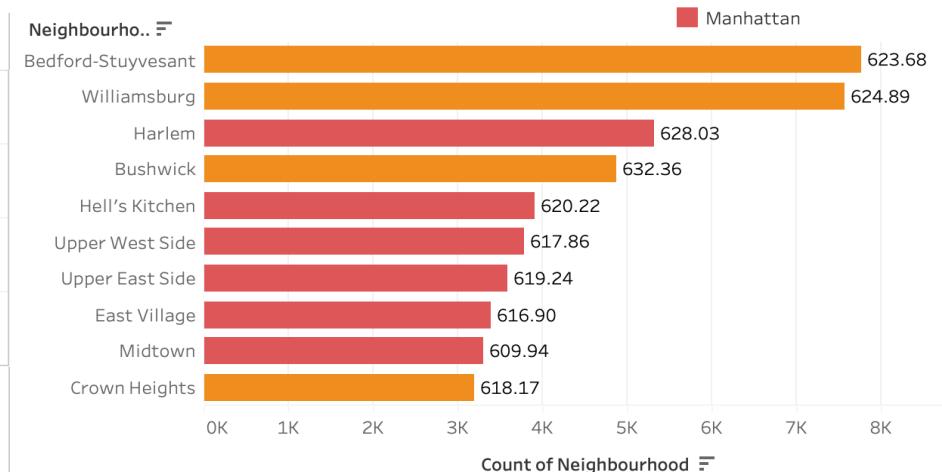
Count of Review Rate Number across different Room Type



Count of Cancellation Policy across different Room Type



Most popular neighbourhood



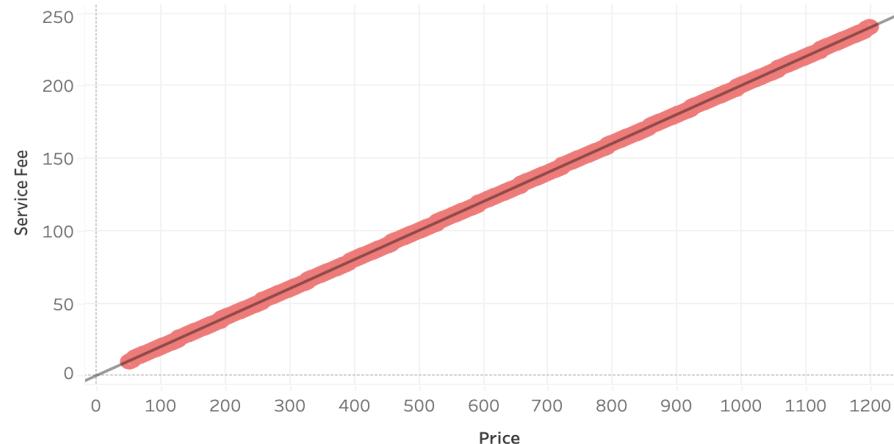


# Dashboard-2

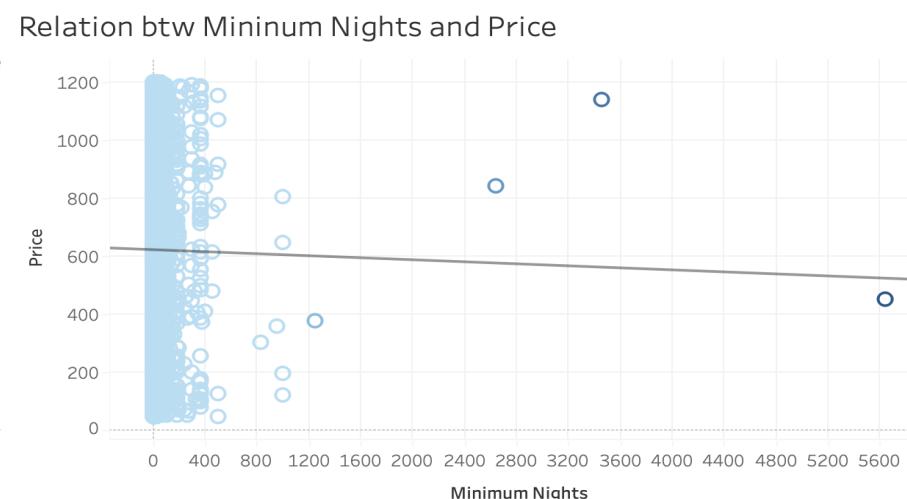


Airbnb in New York #listings: 1,00,147 Avg.price: \$625/day popular neighbourhood: Brooklyn&Manhattan

Relation btw Price and Service Fee



Relation btw Mininum Nights and Price



Minimum Nights  
1 5645

Neighbourhood Group  
 (全部)  
 Bronx  
 Brooklyn  
 Manhattan  
 Queens  
 Staten Island

Room Type  
 (全部)  
 Entire home/apt  
 Hotel room  
 Private room  
 Shared room

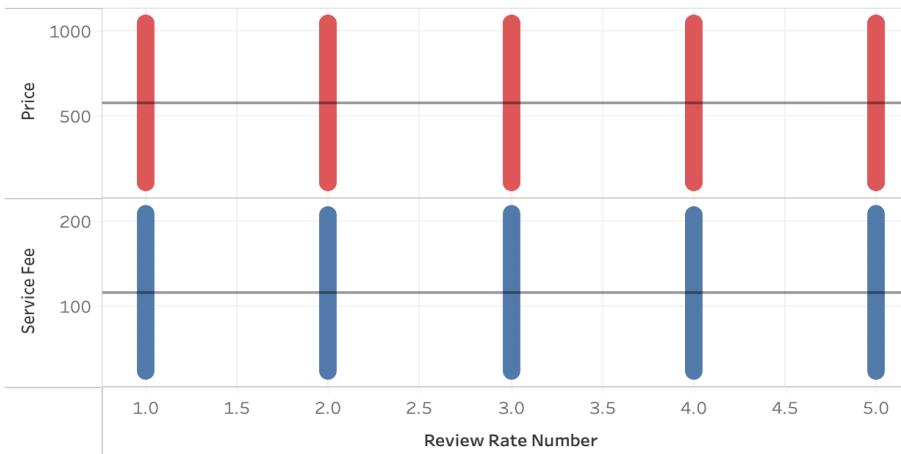
Price  
50 1200

Service Fee  
10.0 240.0

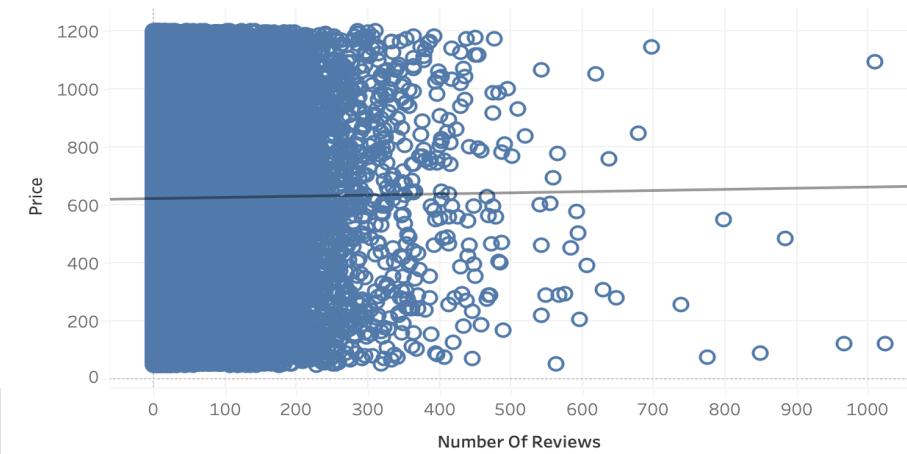
Review Rate Number  
1.000 5.000

Number Of Reviews  
0 1024

Relation btw review rate, price and service fee



Relation btw Number of Reviews and Price





# Summary & Recommendations



# Summary & Recommendations

Summary:

1. Most of the listings of Airbnb in New York are entire room, apartment and private room. Most of it are distributed in **Manhattan and Brooklyn**.
2. The **entire room/apt** receive the **highest** review rate number.
3. The number of reviews of listings is **positively associated with its price**.

Recommendations for others potential host:

1. It would be a good choice for you to start your renting business in the other neighbourhood **rather than Manhattan and Brooklyn**.
2. You are recommended to rent the room type such as **entire room/apt**.
3. You can encourage your travelers to **write a review for your listings**.