Explore the Impact of Location on Airbnb Price in New York City

IBM DATA SCIENTIST CERTIFICATE
CAPSTONE PROJECT – THE BATTLE OF
NEIGHBORHOODS

BY TING LIU
OCTOBER 2019

INTRODUCTION

- Background: Airbnb Inc. is an online marketplace for arranging or offering lodging, primarily homestays, or tourism experiences. Airbnb listing prices are suggests by Airbnb and determined by hosts.
- Problem: How to pricing/how to choose a place in budget
- Interests: Tourist who wants to find an Airbnb for stay; Airbnb hosts; Airbnb to suggest a price.

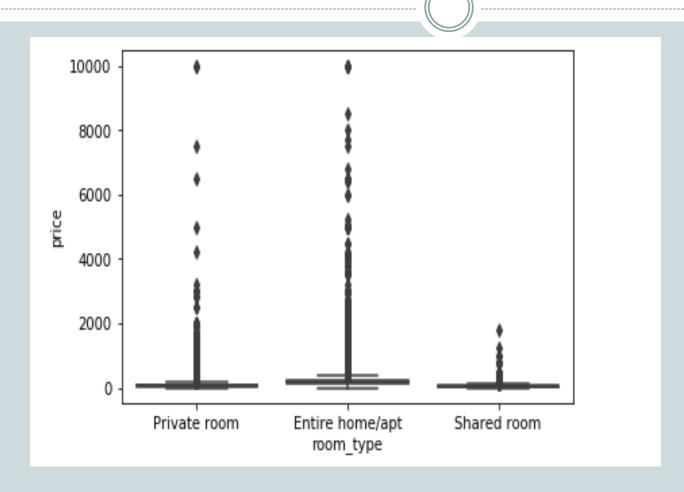
Data

- The New York City Airbnb Open Data (Kaggle.com)
 - o 48895 rows (listings)
 - o 16 variables in
 - × Host
 - x Location
 - × Price
 - × Reviews
 - Availability
- Foursquare location data

Methodology

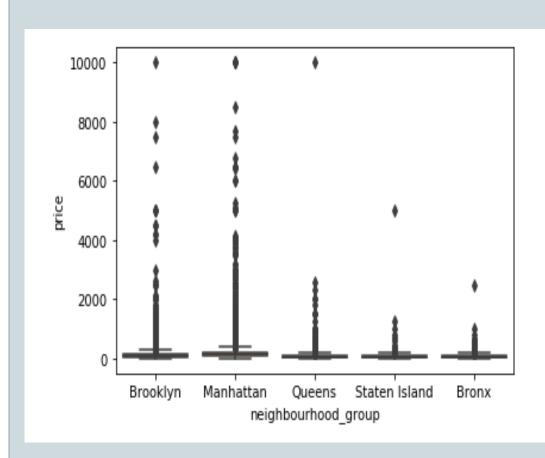
- Exploratory Analysis
- K-means clustering
- Visualization
- Foursquare location data

Exploratory Analysis I



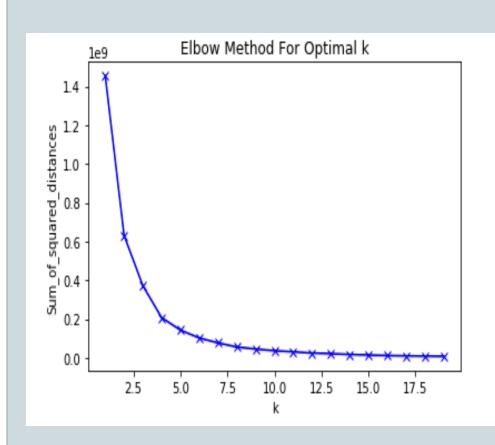
- Most expensive: Entire home/ apt
- Most diverse: Entire home/ apt

Exploratory Analysis II



- Most expensive: Manhattan
- Most diverse: Manhattan

Cluster Analysis I



Elbow Method is used to find the optimal k.

Optimal k = 4

Cluster Analysis II

Cluster	Number of Listing	Average price
0	12139	\$195.20
1	64	\$2697.97
2	983	\$663.00
3	13	\$7366.92

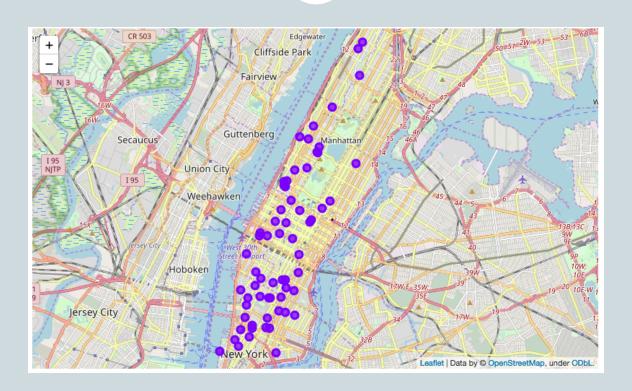
- Cluster o: Most listing, least expensive
- Cluster 3: Least listing, most expensive

Cluster o



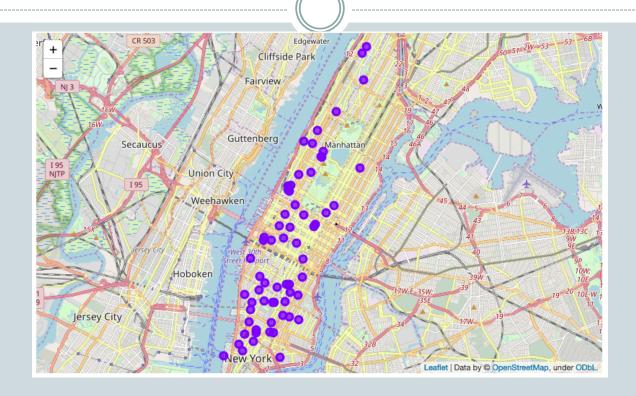
- 12139 listings (100 show on the map)
- \$195 average listing price
- Located in all neighborhoods

Cluster 1



- 64 listings
- \$1731 average listing price
- Located in almost every neighborhood

Cluster 2



- 983 listings (100 show on the map)
- \$663 average listing price
- Located in also all neighborhoods

Cluster 3



- 13 listings
- \$5100 average listing price
- Located in 10 neighborhoods. No concentrate.

Common Venues

Cluster 3 – Most expensive		Cluster 1 – 2 nd expensive		
Coffee Shop	62	Italian Restaurant	61	
Italian Restaurant	51	Theater	52	
Art Gallery	44	Coffee Shop	39	
Hotel	30	Hotel	34	
Café	30	American Restaurant	30	
Park	29	Gym / Fitness Center	27	
Bakery	29	Clothing Store	26	
Gym	28	Gym	25	
Wine Shop	24	Boutique	25	
Gym / Fitness Center	24	Café	24	
Cluster 2 – 3 ^{ra} expensive		Cluster 0 – least expensive		
Italian Restaurant	56	Hotel	41	
Coffee Shop	45	Italian Restaurant	38	
Gym / Fitness Center	35	Pizza Place	38	
Hotel	34	Coffee Shop	37	
Bakery	33	Theater	35	
American Restaurant	31	Mexican Restaurant	28	
Wine Bar	28	Bar	26	
Park	28	Bakery	23	
Café	28	American Restaurant	21	
French Restaurant	23	Seafood Restaurant	20	
			_	-

Result

- As the price decreases, the density of venues decreases as well. For example, for Cluster 3 listings, there are 51 Italian restaurants nearby. So for each listing, there are about 4 Italian restaurants nearby. But for Cluster 0 listings, same as Cluster 3, Italian restaurants are the 2nd most common venues, but the total number is 38. So for every listing, there are 3 Italian restaurants nearby.
- Art Galleries are very common near Cluster 3 listings. For each Cluster 3 listing, there are more than 3 Art Galleries nearby. Also, Wine shops are common for Cluster 3 listings only. There are about 2 Wine shops near each Cluster 3 listing.
- Gym/Fitness centers are common for Cluster 1,2 and 3. For the least expensive cluster Cluster 0, you can find different restaurants around, but not so much fitness centers.
- Clothing stores and Boutiques are only common near Cluster 1, which is the 2nd expensive cluster.

Conclusion

- Airbnb prices differed from town to town, but in a town, such as Manhattan, the prices are not determined by the neighborhood, as expensive ones and affordable ones are mixed together.
- In areas that Art Galleries, clothing stores and bouquets concentrated, the room prices are high.
- In areas that with a lot of restaurants but no of few gym, the prices are low.
- In the future: to add house/apt information