Capstone Project:

Data Analysis on Industrial Cluster Effect in Manhattan

1 Introduction

Many businessmen have trouble finding a location for their new business. Therefore, I'm going to conduct a project to prove whether there is **Industrial Cluster Effect** in Manhattan. If Industrial Cluster Effect does exist, merchants should choose locations with similar stores instead of setting up stores in isolation.

2 Data Description

I use the **Foursquare location data** as well as a **json** file from this course to solve this problem. The json file has some neighborhood information. Foursquare location data has a great amount of data including venues information in each neighborhood.

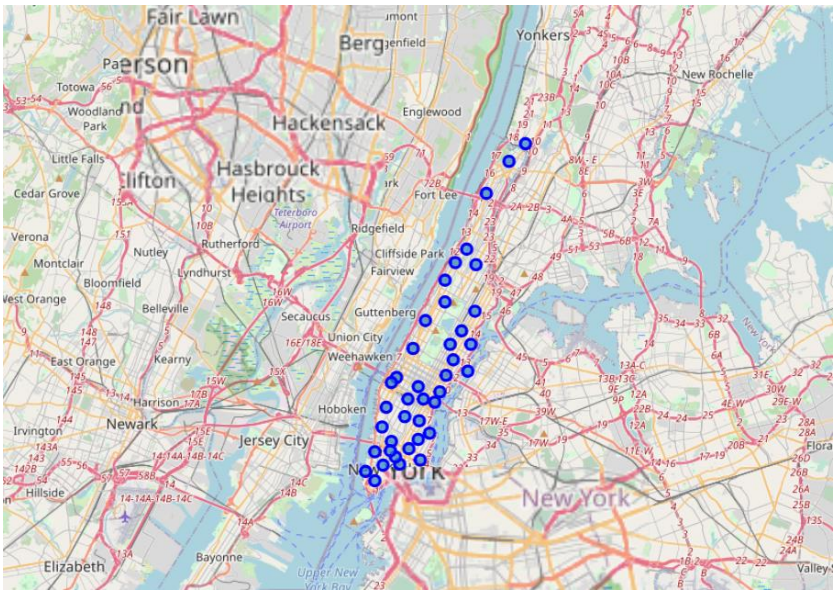
3 Methodology

3.1 Dealing with json file

This is the dataframe I transformed into, showing the Latitude and Longitude of Neighborhoods in Manhattan.

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

As you can see, the markers on the map are different neighborhoods in Manhattan.



3.2 Dealing with Foursquare location data

In this dataset, there are more details of the neighborhoods, for example, some venue information of each neighborhood.

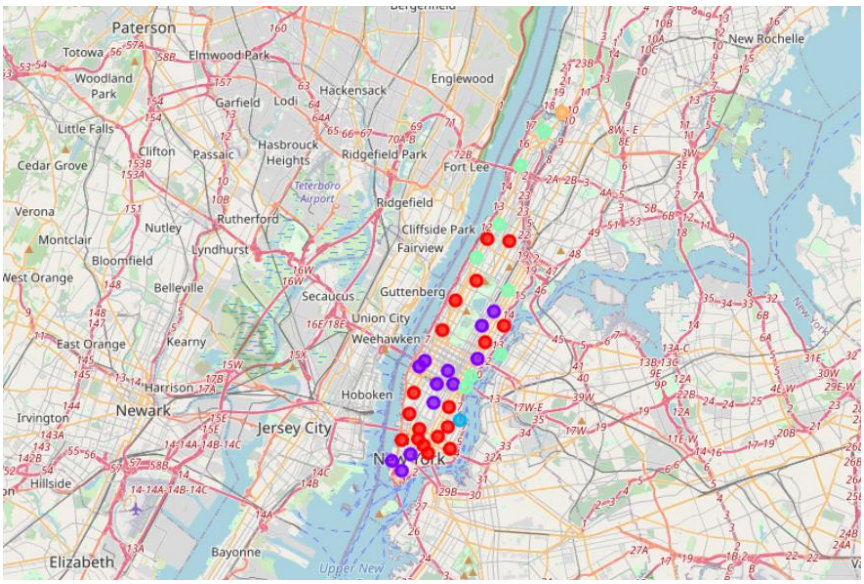
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop
4	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop

This dataframe shows the mean score of different characteristic in each neighborhood. With this dataframe, we can use K-means to cluster the neighborhoods.

	Neighborhood	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Antique Shop	Arcade	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Rest
0	Battery Park City	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
1	Carnegie Hill	0.000000	0.000000	0.00	0.000000	0.011905	0.000000	0.000000	0.000000	0.011905	0.000000	0.011905	0.000000	0.0
2	Central Harlem	0.000000	0.000000	0.00	0.065217	0.043478	0.000000	0.000000	0.000000	0.000000	0.021739	0.000000	0.000000	0.0
3	Chelsea	0.000000	0.000000	0.00	0.000000	0.030000	0.000000	0.000000	0.000000	0.000000	0.130000	0.000000	0.000000	0.0
4	Chinatown	0.000000	0.000000	0.00	0.000000	0.030000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
5	Civic Center	0.000000	0.000000	0.00	0.000000	0.044444	0.011111	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0
6	Clinton	0.000000	0.000000	0.00	0.000000	0.030000	0.000000	0.000000	0.000000	0.000000	0.010000	0.000000	0.000000	0.0
7	East Harlem	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
8	East Village	0.000000	0.000000	0.00	0.000000	0.010000	0.000000	0.000000	0.010000	0.010000	0.010000	0.000000	0.010000	0.0
9	Financial District	0.000000	0.000000	0.00	0.000000	0.050000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
10	Flatiron	0.000000	0.000000	0.00	0.000000	0.020202	0.000000	0.000000	0.000000	0.000000	0.010101	0.000000	0.020202	0.0
11	Gramercy	0.000000	0.000000	0.00	0.000000	0.025641	0.000000	0.012821	0.000000	0.000000	0.012821	0.000000	0.000000	0.0
12	Greenwich Village	0.000000	0.000000	0.00	0.000000	0.010000	0.000000	0.000000	0.000000	0.000000	0.010000	0.000000	0.000000	0.0
13	Hamilton Heights	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
14	Hudson Yards	0.000000	0.000000	0.00	0.000000	0.056604	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
15	Inwood	0.000000	0.000000	0.00	0.000000	0.033333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
16	Lenox Hill	0.000000	0.000000	0.01	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.010000	0.000000	0.000000	0.0

4 Results

Neighborhoods of different color belongs to different clusters.



This dataframe shows the details of neighborhoods in Cluster 1. Neighborhoods in this single cluster have a lot of similar cafes, bars and restaurants, proving that they show high similarity with each other.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Chinatown	Chinese Restaurant	Bakery	Bubble Tea Shop	Cocktail Bar	Salon / Barbershop	Vietnamese Restaurant	Coffee Shop	American Restaurant	Spa	Optical Shop
5	Manhattanville	Coffee Shop	Seafood Restaurant	Chinese Restaurant	Mexican Restaurant	Park	Italian Restaurant	Café	Bike Trail	Sushi Restaurant	Boutique
6	Central Harlem	Chinese Restaurant	African Restaurant	American Restaurant	Bar	Cosmetics Shop	French Restaurant	Seafood Restaurant	Café	Market	Bookstore
9	Yorkville	Coffee Shop	Italian Restaurant	Gym	Sushi Restaurant	Bar	Deli / Bodega	Wine Shop	Diner	Mexican Restaurant	Japanese Restaurant
10	Lenox Hill	Coffee Shop	Pizza Place	Italian Restaurant	Sushi Restaurant	Cocktail Bar	Café	Gym	Gym / Fitness Center	Burger Joint	Salon / Barbershop
12	Upper West Side	Wine Bar	Italian Restaurant	Dessert Shop	Coffee Shop	Bar	Pizza Place	Vegetarian / Vegan Restaurant	Bagel Shop	Mediterranean Restaurant	Sushi Restaurant
13	Lincoln Square	Plaza	Italian Restaurant	Café	Theater	Gym / Fitness Center	Concert Hall	Performing Arts Venue	Indie Movie Theater	Wine Shop	American Restaurant
17	Chelsea	Art Gallery	Coffee Shop	Café	Ice Cream Shop	American Restaurant	Cycle Studio	Cupcake Shop	Nightclub	Market	Hotel
18	Greenwich Village	Italian Restaurant	Gym	Coffee Shop	Café	Pizza Place	Sandwich Place	Bakery	Ice Cream Shop	Indian Restaurant	Clothing Store
19	East Village	Cocktail Bar	Pizza Place	Coffee Shop	Mexican Restaurant	Bar	Ramen Restaurant	Juice Bar	Wine Bar	Japanese Restaurant	Italian Restaurant
20	Lower East Side	Chinese Restaurant	Art Gallery	Ramen Restaurant	Café	Cocktail Bar	Grocery Store	Latin American Restaurant	Mediterranean Restaurant	Pharmacy	French Restaurant
21	Tribeca	Park	Italian Restaurant	Wine Bar	Spa	Café	Steakhouse	Playground	Bakery	Scenic Lookout	Men's Store

5 Discussion

As shown above, in Manhattan, merchants usually open similar stores, bars, cafes and restaurants near each other, showing that maybe the gathering of stores and restaurants will attact more customers in a certain neighborhood.
So I suggest that you find a place where many similar firms locate when you want to start your business in Manhattan.

6 Conclusion

There is a significant Industrial Cluster Effect in Manhattan.