IBM Applied Data Science Capstone Project

# Asian Breakfast Shop: Neighborhood Battles in Edison, NJ

Nan Feng

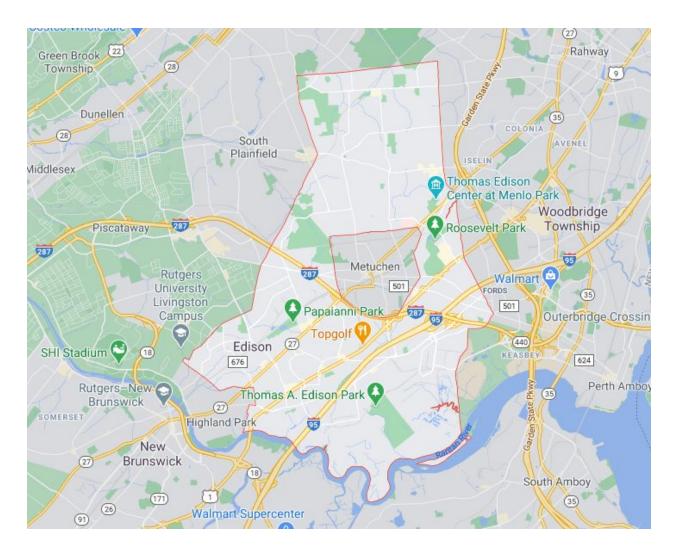


## Introduction

A friend wants to start her dim sum restaurant in one of the 31 neighborhoods in Edison. She hired me as her data scientist to help her decide which neighborhood/location she should open the business.

Based on my research, Edison is the fifth-most populous municipality in New Jersey, with a population of a little more than 100K, 29.27% of Asians are residing in Edison township. It has many Asian supermarkets that Asian residents go to every week. There are at least 25 (popular) Chinese restaurants in town, only one is a dim sum place.

Edison is a huge town with a total area of 30.69 square miles (79.49 km2) and 31 sections/neighborhoods. We want to select the most popular place for the restaurant business.

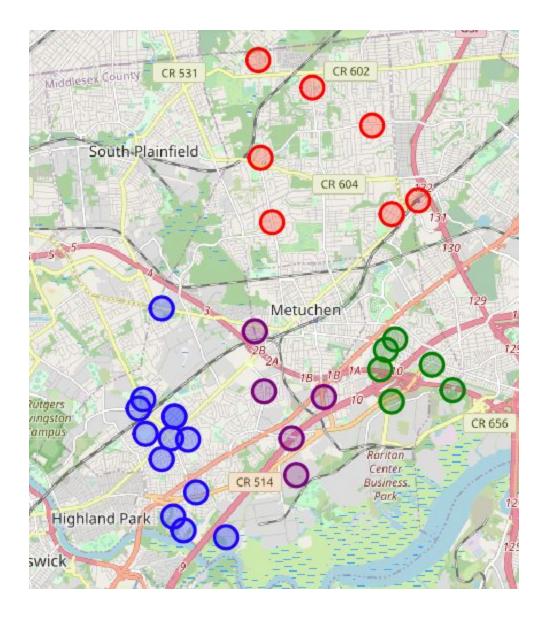


# Data used in this report

Edison neighborhood data is web scraped from wiki link with neighborhood names, latitudes, and longitudes.

	Neighborhood	Notes	Latitude	Longitude			
0	Bonhamtown	The neighborhood began as colonial village tha	e tha 40.523330 -74.3				
1	Camp Kilmer site	Abutting the Livingston Campus (Rutgers Univer	outting the Livingston Campus (Rutgers Univer 40.520830				
2	Clara Barton	Named for Clara Barton. Located in the eastern	40.534170	-74.339440			
3	Greensand	Site of the Kin-Buc Landfill, where Edmonds Cr	40.491670	-74.387220			
4	Haven Homes	NaN	40.509440	-74.406390			
5	Lahiere	Lahiere is a neighbohood in the Stelton sectio	40.518890	-74.402500			
6	Lincoln Park	For other places named Lincoln Park, New Jerse	40.518890	-74.402500			
7	Lindenau	Lindenauis a neighborhood of the Stelton secti	40.515000	-74.411110			
8	Martins Landing	Located on the northern banks of the Raritan R	40.493060	-74.399720			
9	Menlo Park	Menlo Park is where, in 1876, Thomas Edison se	40.565000	-74.337500			
10	Menlo Park Mall	Super regional shopping mall on U.S. Route 1 a	40.538330	-74.378610			
11	Metropark	Largely located in neighboring Iselin, the off	40.568080	-74.329795			
12	Middlesex County College	Sited 159 acres (64 ha) of the former Raritan	40.505830	-74.366110			
13	New Dover	The neighborhood began as colonial village tha	40.585000	-74.343330			
14	New Durham	NaN	40.543610	-74.406110			
15	Nixon	NaN	40.514170	-74.367500			
16	North Edison	NaN	40.600000	-74.377500			
17	Oak Tree	An area of the battlefield of the Oak Tree Eng	of the battlefield of the Oak Tree Eng 40.577780				
18	Oak Tree Road	lak Tree An area of the battlefield of the Oak Tree Eng		-74.411674			
19	Phoenix	Located in the eastern part of the township, P	40.529720	-74.341110			
20	Piscataway	This is the oldest neighborhood in Edison. Alo	40.501670	-74.395830			
21	Potters	In the northern part of the township Potters w	40.593610	-74.361390			
22	Pumptown	Pumptown is a residential neighborhood with a	vn is a residential neighborhood with a 40.563060				
23	Raritan Center	Sited on part if the former Raritan Arsenal,[4	ed on part if the former Raritan Arsenal,[4 40.522500				
24	Raritan Manor	NaN	40.536390	-74.336670			
25	Sand Hills	Sand Hills is a neighborhood straddling Edison	40.525000 -74.319440				
26	Silver Lake	A small lake of same name borders the neighbor	mall lake of same name borders the neighbor 40.496209 -74.				
27	Stelton	One the oldest parts of Edison.	40.514170	-74.403610			
28	Valentine	Valentine is adjacent to the Fords section of	40.530830	-74.325560			
29	Vineyard Village	Adjacent to former Edison Assembly. Edward Ant	40.524728	-74.375613			

Using KMEANS clustering with n\_clusters = 4, just based on each neighborhood's latitude/longitude from the above data frame, we can divide the 31 neighborhoods of Edison into 4 different regions/clusters. This is just a way to show the 31 neighborhoods in Edison.



Geo divided clusters are not enough for our business analysis purpose.

We will make calls to Four Square APIs to explore the most popular business (venues) in each neighborhood

- 1. Radius = 1000 m
- 2. Limit = 100 venues

## Methodology

#### **KMeans Clustering on Venues and Cluster Analysis**

We used KMeans Clustering with n\_clusters = 6, n\_init = 12, algorithm = "auto". The clustering is based the popular venues in the neighborhood.

The 31 neighborhoods are divided into 6 different regions.



#### Each cluster has different top popular venues

- Cluster 1: American Restaurant; Italian Restaurant; Bar, Deli
- Cluster 2: Pizza:
- Cluster 3: Indian Restaurant
- Cluster 4: Park

We will focus on cluster 0---

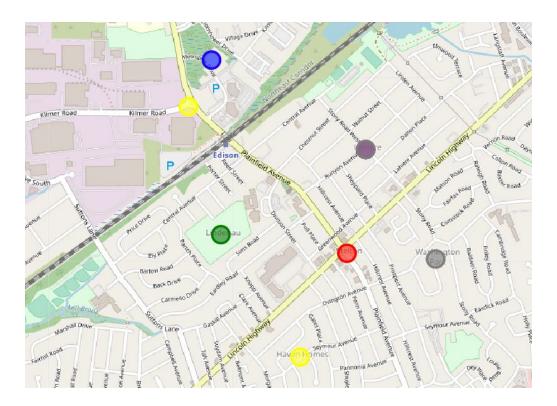
10th Most Popula Venu	9th Most Popular Venue	8th Most Popular Venue	7th Most Popular Venue	6th Most Popular Venue	5th Most Popular Venue	4th Most Popular Venue	3rd Most Popular Venue	2nd Most Popular Venue	1st Most Popular Venue	Venue Label	Neighborhood
Sandwich Place	Fast Food Restaurant	Farmers Market	Pharmacy	Bubble Tea Shop	Asian Restaurant	Dessert Shop	Korean Restaurant	Bakery	Chinese Restaurant	0	Washington Park
Ban	Asian Restaurant	Korean Restaurant	Construction & Landscaping	Bubble Tea Shop	Indian Restaurant	Bakery	Dessert Shop	Spa	Chinese Restaurant	0	Lindenau
Bubble Tea Sho	Lawyer	Malay Restaurant	Fast Food Restaurant	Bank	Bakery	Dessert Shop	Asian Restaurant	Korean Restaurant	Chinese Restaurant	0	Lincoln Park
Bubble Tea Sho	Lawyer	Malay Restaurant	Fast Food Restaurant	Bank	Bakery	Dessert Shop	Asian Restaurant	Korean Restaurant	Chinese Restaurant	0	Lahiere
Supermarke	Bank	Fast Food Restaurant	Spa	Sandwich Place	Indian Restaurant	Korean Restaurant	Chinese Restaurant	Dessert Shop	Bakery	0	Stelton
Food & Drink Sho	Fried Chicken Joint	Frozen Yogurt Shop	Farmers Market	Yoga Studio	Fast Food Restaurant	Rental Service	Auto Garage	Lake	Business Service	0	Greensand
Chinese Restaurar	Playground	BBQ Joint	Deli / Bodega	Dentist's Office	Indian Restaurant	Hotpot Restaurant	Bakery	Korean Restaurant	Spa	0	Camp Kilmer site
Supermarke	Hookah Bar	Pharmacy	Spa	Wings Joint	Bubble Tea Shop	Dessert Shop	Bakery	Korean Restaurant	Chinese Restaurant	0	Haven Homes
Farmers Marke	Yoga Studio	Lawyer	Spa	Convenience Store	Deli / Bodega	Indian Restaurant	Playground	Dentist's Office	Business Service	0	Oak Tree Road

- Cluster 0 include 9 different neighborhoods: Washington Park, Lindenau, Lincoln Park,
  Lahiere, Stelton, Greensand, Camp Kilmer site, Haven Homes, Oak Tree Road
- Top Venues:
  - o Chinese Restaurant, Korean Restaurant, Asian Restaurant
  - o Bakery, Bubble Tea, Dessert shop
  - Deli/ Bodega

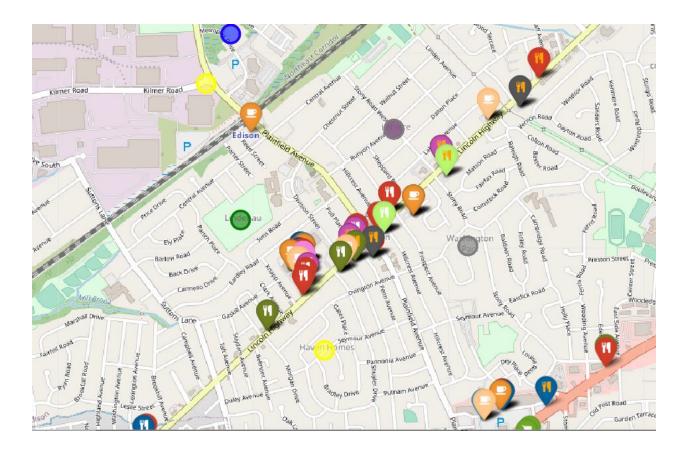
### Further Analysis on Cluster 0: "Edison Chinatown" and Results

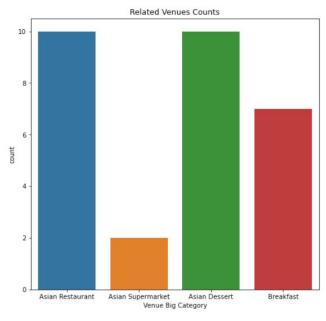
Using KMeans clustering again.

- n\_clusters = 6
- n\_init = 12



Now Stelton neighborhood (red circle marker in the map) is singled out as the center of "Edison Chinatown".





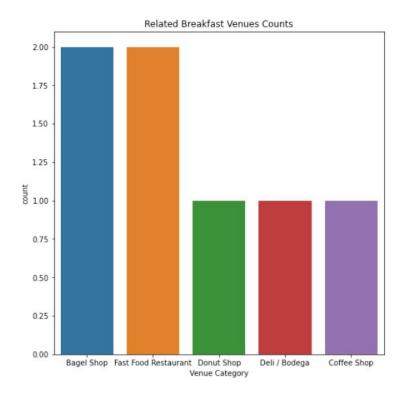
#### In Stelton Neighborhood, there are

Aisan Supermarket: 2

• Asian Restaurants: 10

Asian Dessert: 10

• Breakfast Places: 7 (Are they Asian breakfast type?)



We further examin the breakfast related venues in Stelton, we found none of the seven breakfast related venues are Asian style.

# **Conclusion and Suggestions**

We performed detailed analysis in 31 Neighborhoods in Edison.

We pinned our final neighborhood in Stelton, Edison.

The reason is Asian food related venues are thriving in this area.

They have Asian supermarkets, Asian restaurants and Asian dessert shops. But Asian stylpe breakfast shops are missing in this neighborhood,

We suggested to the client that she should open her dim sum in the H-Mart Mall in Stelton.

