Battle of Neighborhoods

INTRODUCTION

- New York City review
- Optimal location for new restaurant business
  - Relatively low competition
  - Relatively high demand

NEW YORK CITY

- One of the most populous city in the USA
- Cultural diversity
- Financial Capital of the USA
- Highly competitive market
- Strategic entry for new ventures to:
  - Reduce risk
  - Maintain a good return on investment

DATA DESCRIPTION

- There are a total of 5 boroughs and 306 neighborhoods
- Dataset available here: https://geo.nyu.edu/catalog/nyu_2451_34572

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
2	Wakefield	40.894705	-73.847201	Walgreens	40.896528	-73.844700	Pharmacy
3	Wakefield	40.894705	-73.847201	Rite Aid	40.896649	-73.844846	Pharmacy
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop

DATA DESCRIPTION

- Data from Foursquare provide coordinates and detailed categorization of the different restaurants in New York City

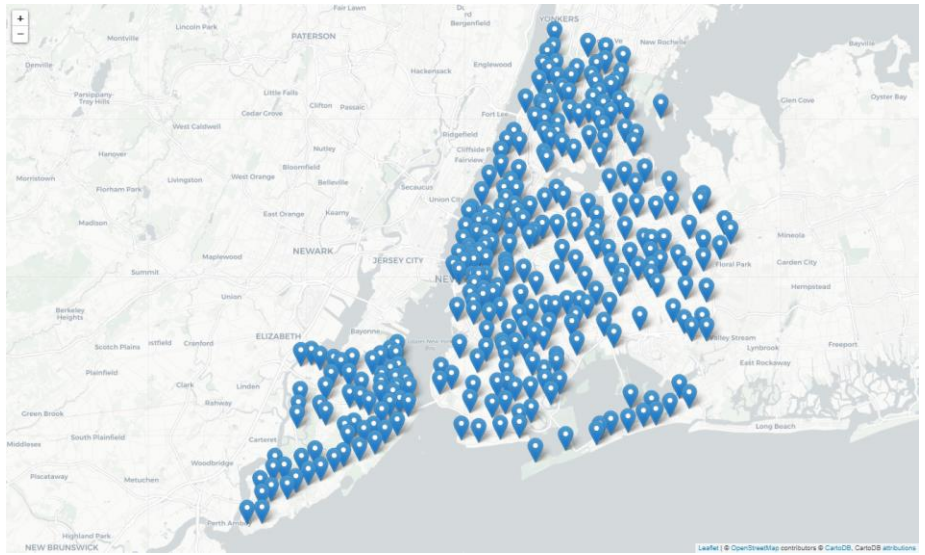
	Neighborhood	BBQ Joint	Bagel Shop	Bakery	Bed & Breakfast	Beer Bar	Beer Garden	Beer Store	Bistro	Breakfast Spot	...	Mac & Cheese Joint	Pet Café	Pie Shop	Pizza Place	Poke Place	Southern / Soul Food Restaurant	Steakhouse	Taco Place	Vegetarian / Vegan Restaurant	Wings Joint
0	Wakefield	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	Wakefield	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	Wakefield	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
9	Co-op City	0	0	0	0	0	0	0	0	0	...	0	0	0	1	0	0	0	0	0	0
13	Co-op City	0	1	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

ANALYTICAL APPROACH

- New York City has a total of 306 neighborhoods
- Subset only catering-related results
- Clustering on only the catering distribution

METHODOLOGY

- New York City geographical coordinates data
- Transformation into GeoPandas DataFrame
- Coordinates used for Foursquare queries
- Folium used to create map of New York City



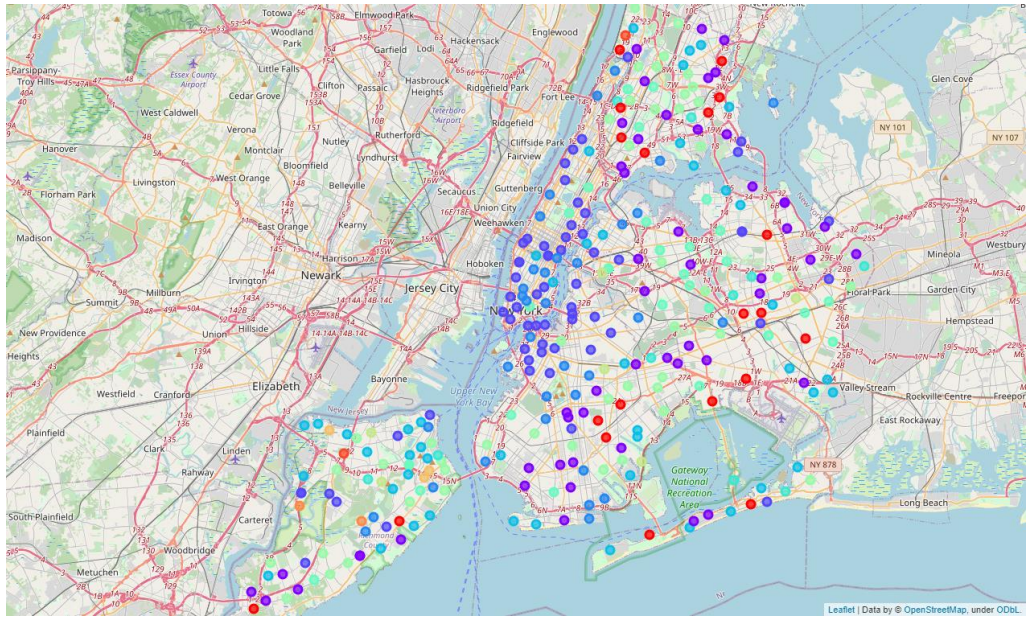
METHODOLOGY

- Obtain data from Wikipedia using Pandas
- Pandas uses html5lib to scrape any table structures online
- The scraped data will be saved in a Pandas DataFrame format

	Jurisdiction	Jurisdiction.1	Population	Gross Domestic Product	Gross Domestic Product.1	Land area	Land area.1	Density	Density.1
1	The Bronx	Bronx	1418207	42.695	30100	42.10	109.04	33867	13006
2	Brooklyn	Kings	2559903	91.559	35800	70.82	183.42	36147	13957
3	Manhattan	New York	1628706	600.244	368500	22.83	59.13	71341	27544
4	Queens	Queens	2253858	93.310	41400	108.53	281.09	20767	8018
5	Staten Island	Richmond	476143	14.514	30500	58.37	151.18	8157	3150

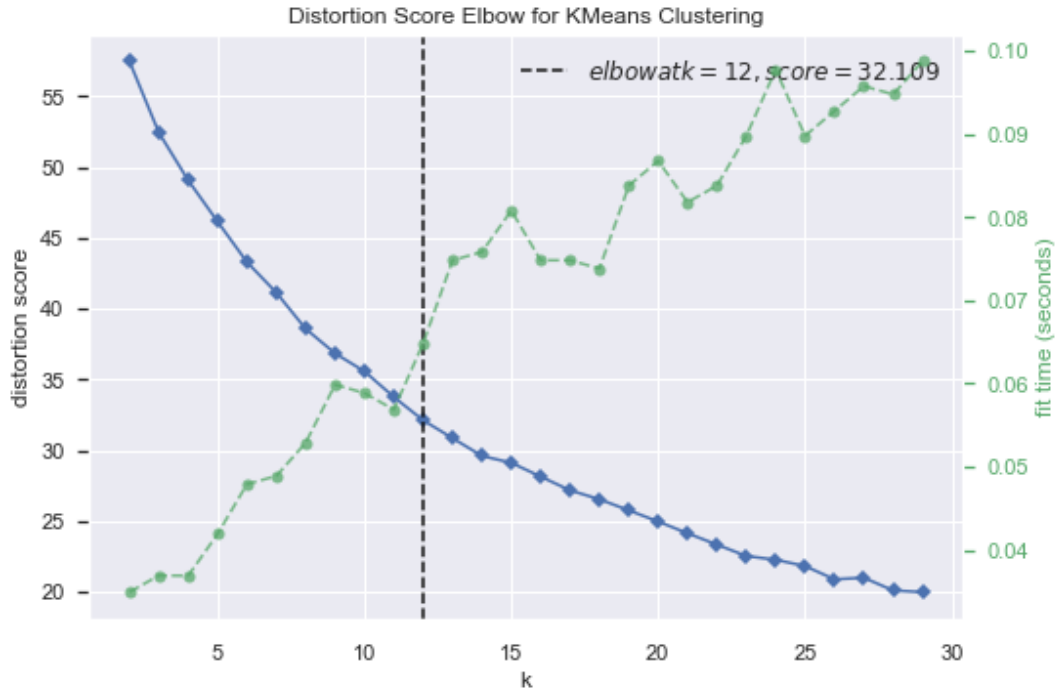
METHODOLOGY

- Query Foursquare using the locations
- Subset only the catering related data
  - E.g. BBQ Joint, Bagel Shop, Buffet, Bar, Pizza Place, etc.
- Determine optimal clustering using elbow methods
- Identify restaurant composition for different cluster of neighborhoods



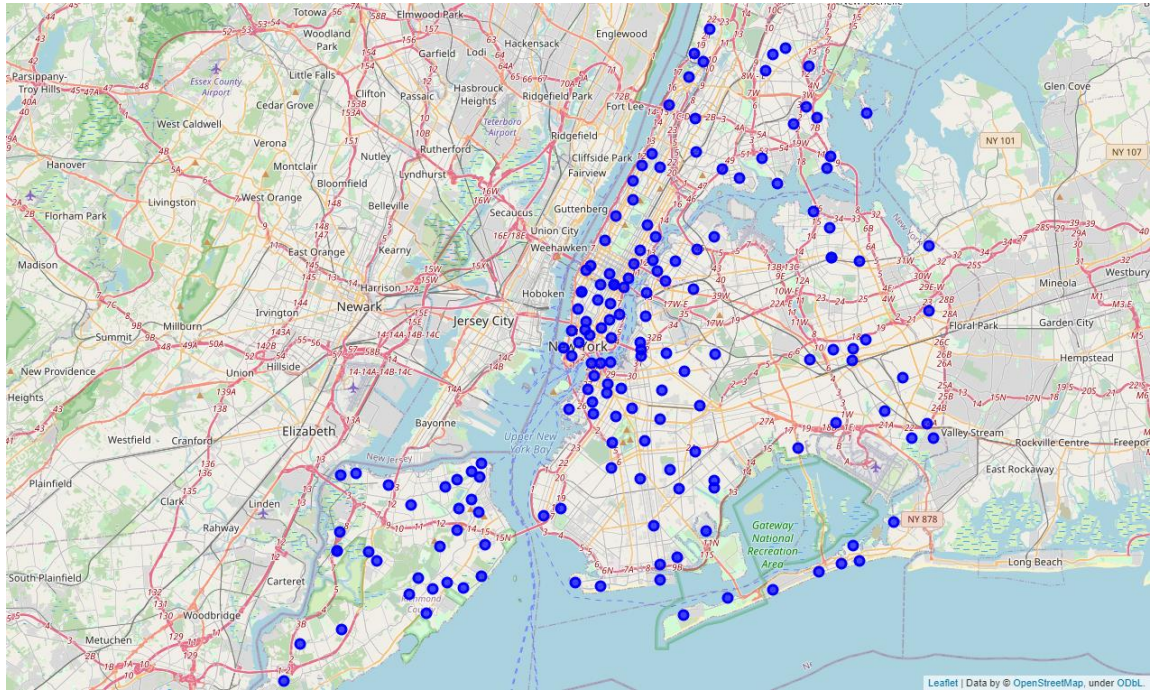
RESULT

- Elbow method indicates a total of 12 clusters for the 306 neighborhoods



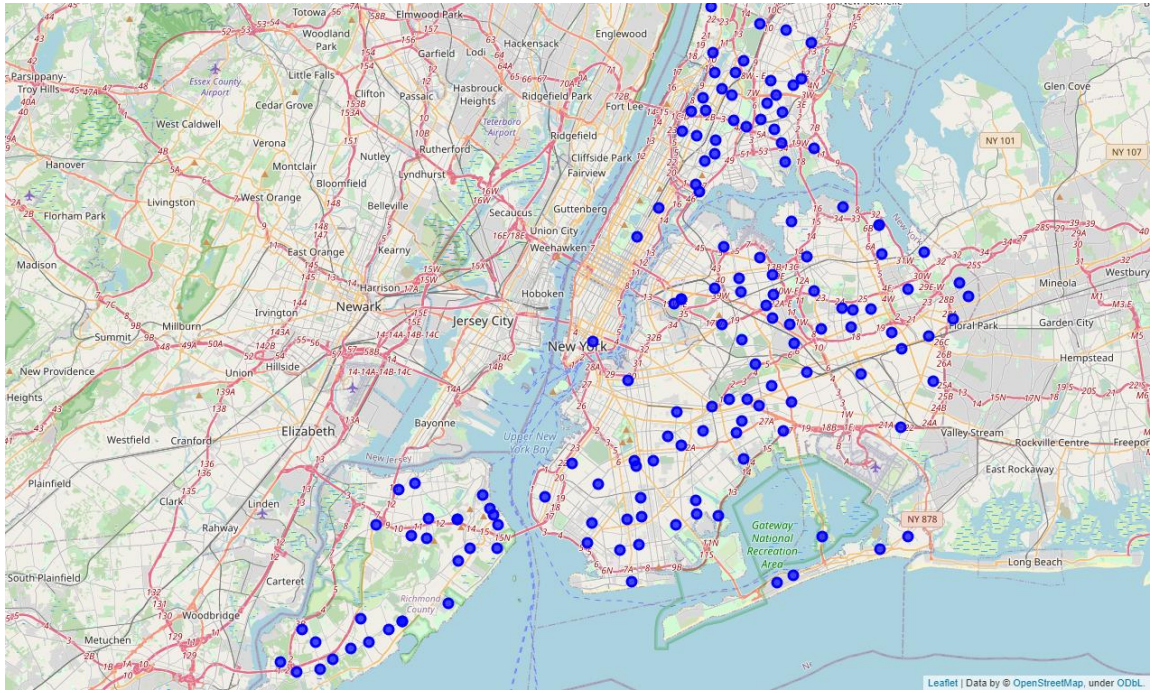
RESULT – Cluster 0, 2, 3, 4

- For neighborhoods within these clusters, the algorithm is suggesting to have an expected proportion for café to be over 10%
- For any neighborhoods that currently has a lower proportion of café, opening a new café can be a decent entry strategy



RESULT – Cluster 1, 5, 6, 7, 8

- For neighborhoods within these clusters, the algorithm is suggesting to have an expected proportion for Pizza place to be over 30%
- For any neighborhoods that currently has a lower proportion of café, opening a new Pizza place can be a decent entry strategy



DISCUSSION

- Scope of refinement: Revise Foursquare categories
  - Some of the categories may be too specific and can be grouped into a more general category to avoid overfitting
- Introduce work demographics and residential demographics to the algorithm
- Introduce NLP elements on restauran t menu to extract affordability and c uisine related data

CONCLUSION

- Café seems to be a more viable strategy for the city center while Pizza place is more towards the outskirt of it
- The analysis has been done on limited data and yet is already able to provide a rough insight on which kind of catering services to focus on at different part of the city