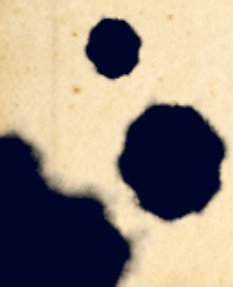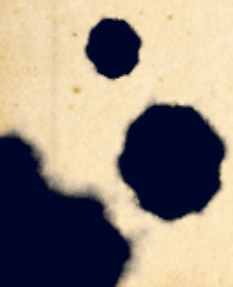# Capstone Project

The Battle of Neighborhoods

Introduction

- Business Problem

  - New York City, the largest city in the United States with a population of around 8.5 million people, accounting for over 40% of the population of New York State. New York City has been a key tourist attraction for visitors from all over the world. Making it one of the most visited cites in the world. And its no coincidence that it also has some of the best and diverse restaurants available to cater for people from all over the world.

  - Being one of the most visited cities in the world. This bring big opportunities for restaurateur to setup business in New York. However finding the best location to setup a Restaurant is a challenge with so many competition.

  - Our goal would be to find the best location to setup a Thai Restaurant by looking at how the restaurants are clustered within the city and to find the best optimal location.
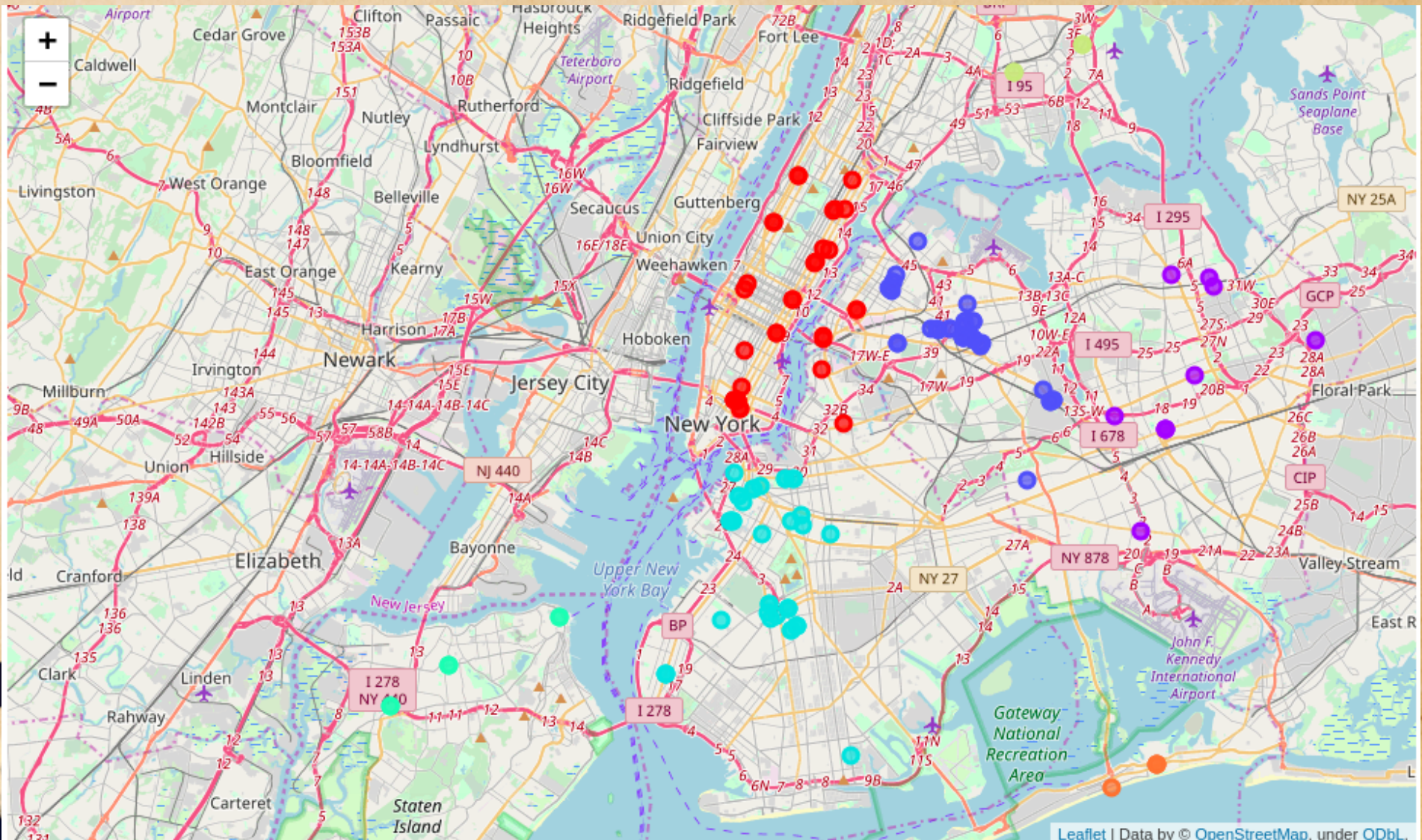
Data acquisition and cleaning

- We will be using the below data sets to give us all the Borough and Neighborhoods in New York City.

  - https://cocl.us/new_york_dataset

- We will pull from FOURSQUARE to gather all the restaurants information like location and cluster them for further analysis.

- We will filter in only Thai Restaurants in the data sets.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
55	Riverdale	40.890834	-73.912585	Nam Thai	4b94a4a2f964a520787f34e3	40.886211	-73.909991	Thai Restaurant
74	Kingsbridge	40.881687	-73.902818	Nam Thai	4b94a4a2f964a520787f34e3	40.886211	-73.909991	Thai Restaurant
85	Marble Hill	40.876551	-73.910660	Siam Square	4a739e29f964a520f5dc1fe3	40.878796	-73.916701	Thai Restaurant
337	Country Club	40.844246	-73.824099	Honey's Thai Pavilion	4b4c90d3f964a52046b626e3	40.847845	-73.831320	Thai Restaurant
351	Parkchester	40.837938	-73.856003	Thai No. 1	5585c948498e0bd8a8a706f1	40.838089	-73.863635	Thai Restaurant
388	Van Nest	40.843608	-73.866299	Thai No. 1	5585c948498e0bd8a8a706f1	40.838089	-73.863635	Thai Restaurant
452	Spuyten Duyvil	40.881395	-73.917190	Siam Square	4a739e29f964a520f5dc1fe3	40.878796	-73.916701	Thai Restaurant

K means Clustering visualization

- We select 7 clusters and run K means to see how the Thai Restaurants are spread out across the city.

Results and Conclusion

- Most of the Thai restaurants are concentrated at the heart of New York City in cluster 0, 2 and 3 and moderate numbers in cluster 1. Its no surprising clusters 4, 5 and 6 have the lowest number as they are further away from Manhattan.

- At first thought, for less competition it would mean choosing cluster 4,5 or 6 as the choice of setting up a Thai Restaurant. But given the fact that location plays an important role for visibility and accessibility as well as the population base and human traffic to support the business. Choosing the suburb area might not be the best choice.

- There is also the consideration of rent being higher at popular locations. The best compromise would be to choose a location in between. That would mean cluster 1 as its not too far from Manhattan and the restaurant concentration in that area are much spread out.