

# Capstone Project Final

## The Battle of Neighborhoods

### Introduction

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 cities in the world. And it's no coincidence that it also has some of the best and diverse restaurants available to cater for people from all over the world.

### Problem

Being one of the most visited cities in the world. This brings big opportunities for restaurateurs to set up business in New York. However, finding the best location to set up a restaurant is a challenge with so many competitors.

Our goal would be to find the best location to set up a Thai restaurant by looking at how the restaurants are clustered within the city and to find the best optimal location.

### Data

We will be using the below datasets to give us all the Borough and Neighborhoods in New York City [https://cocl.us/new\\_york\\_dataset](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.

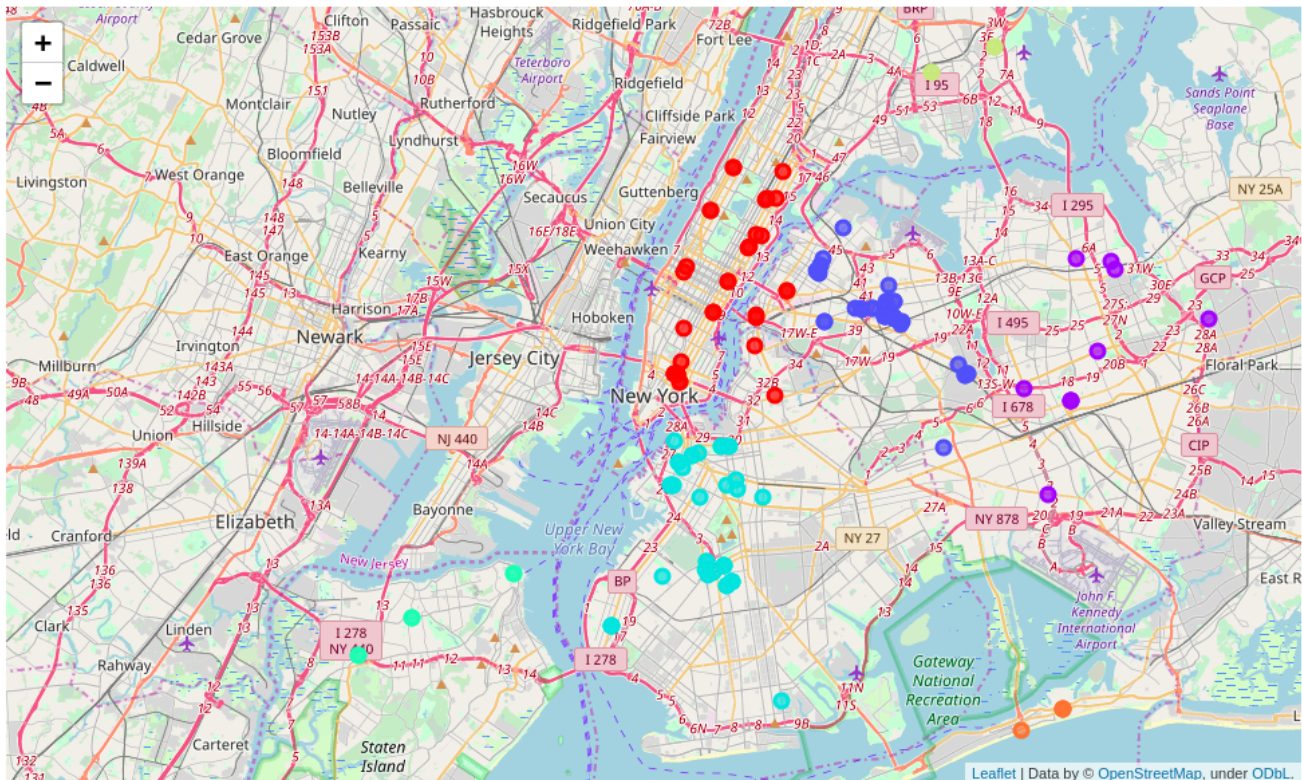
### Methodology or Approach

- Load the datasets from [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)
- Using FourSquare API to find venues of each neighborhood
- Filter in Thai Restaurants
- Do a Kmeans clustering of the Thai Restaurants location
- Do a visual analysis on the clusters

## Results and Discussions

From the filtered datasets, we used K means clustering to have a visual on how all the Thai Restaurants are spread across the city.

We can see 3 main clusters (red, cyan and blue) are situated near the city center.



The breakdown of the number of Thai Restaurants in all the clusters.

```
: for c1 in range(kclusters):  
    print("There are " + str(len(ny_merged2[ny_merged2['Cluster Labels'] == c1])) + ' Thai restaurants in cluster '
```

```
There are 35 Thai restaurants in cluster 0  
There are 10 Thai restaurants in cluster 1  
There are 37 Thai restaurants in cluster 2  
There are 38 Thai restaurants in cluster 3  
There are 6 Thai restaurants in cluster 4  
There are 9 Thai restaurants in cluster 5  
There are 4 Thai restaurants in cluster 6
```

## **Observation 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.