

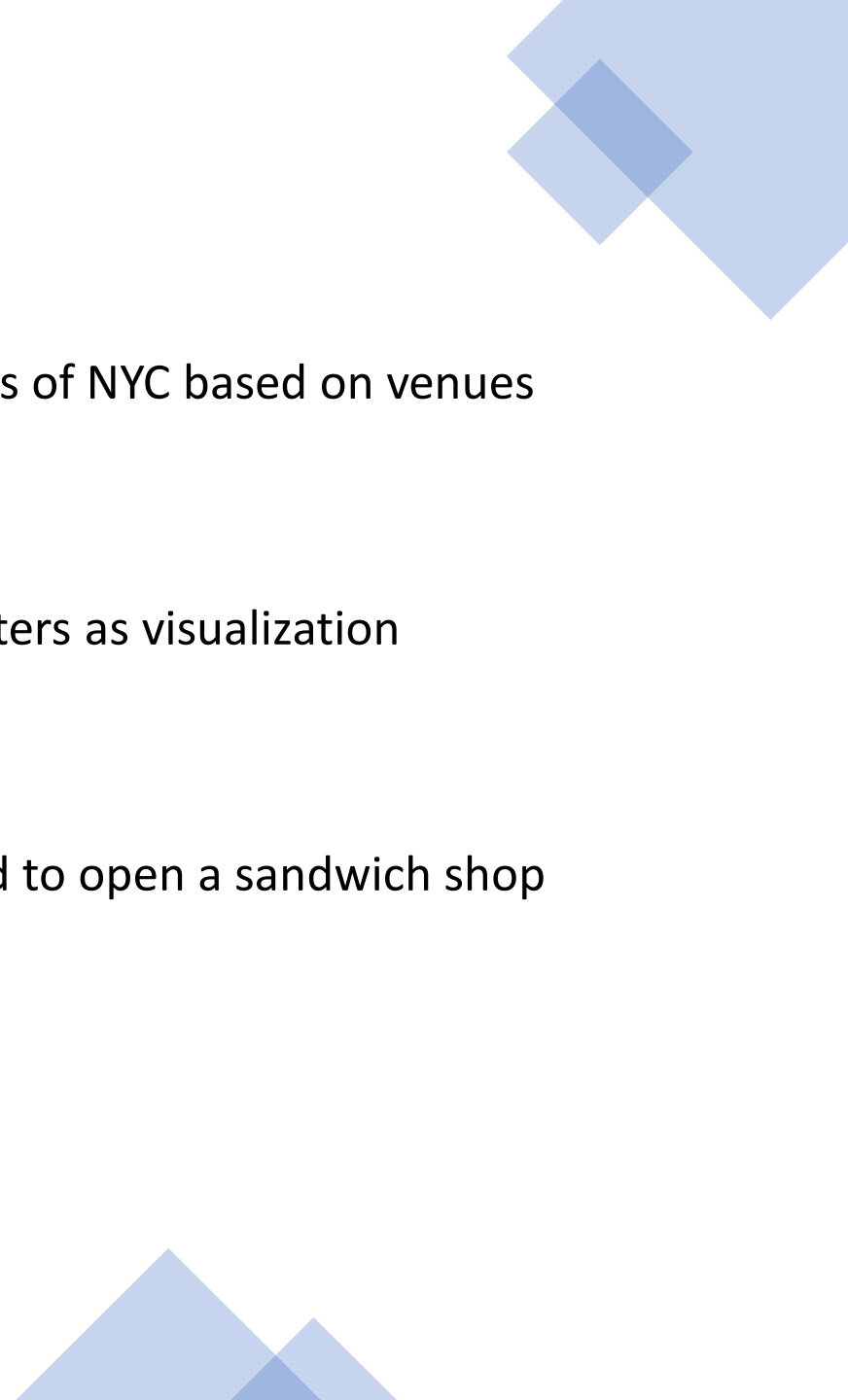


Coursera Capstone Project

BY:
Michiel Schinkel

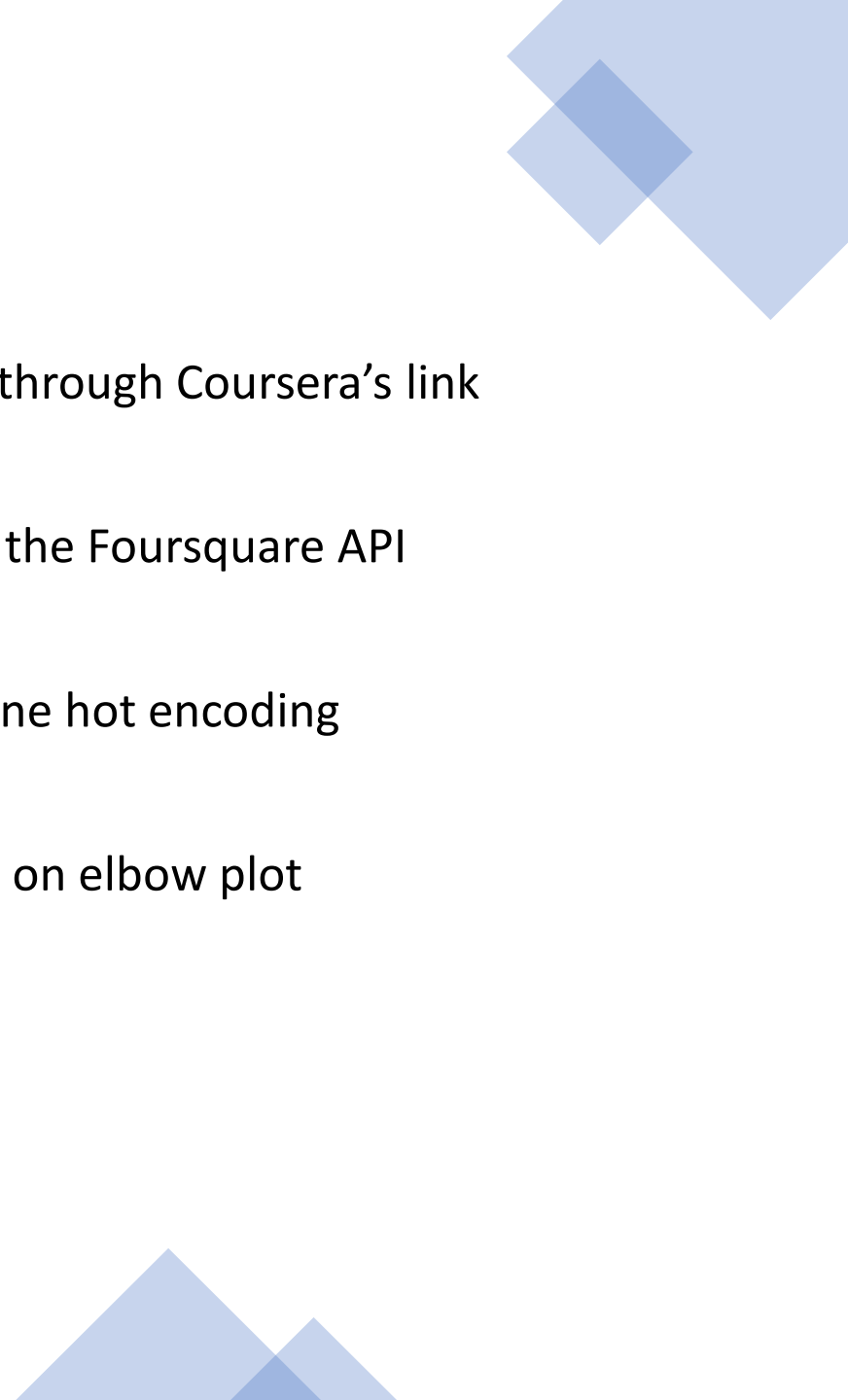


Aims

- Cluster the neighborhoods of NYC based on venues
 - Present a map of the clusters as visualization
 - Select ideal neighborhood to open a sandwich shop
- 

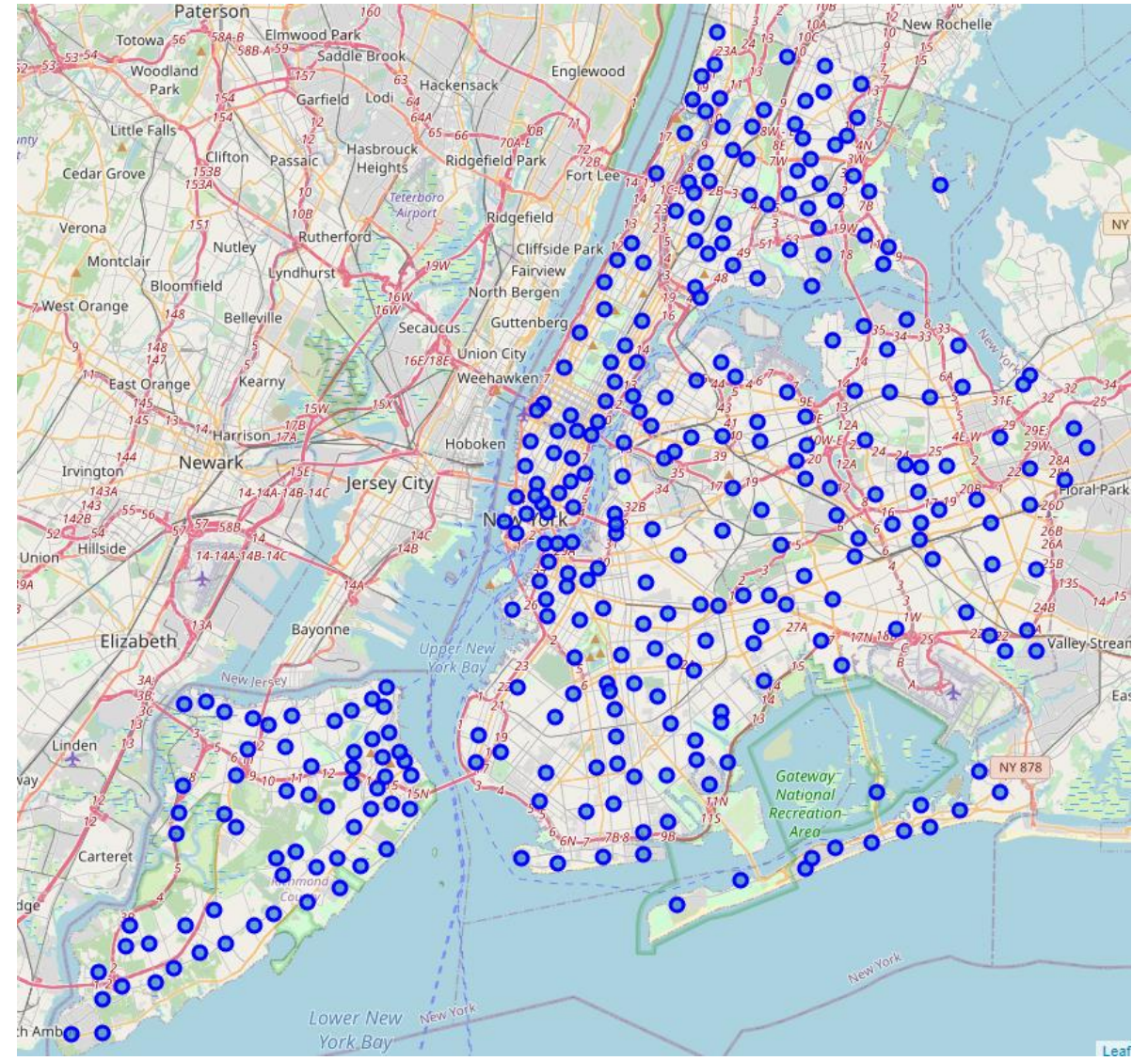


Methods

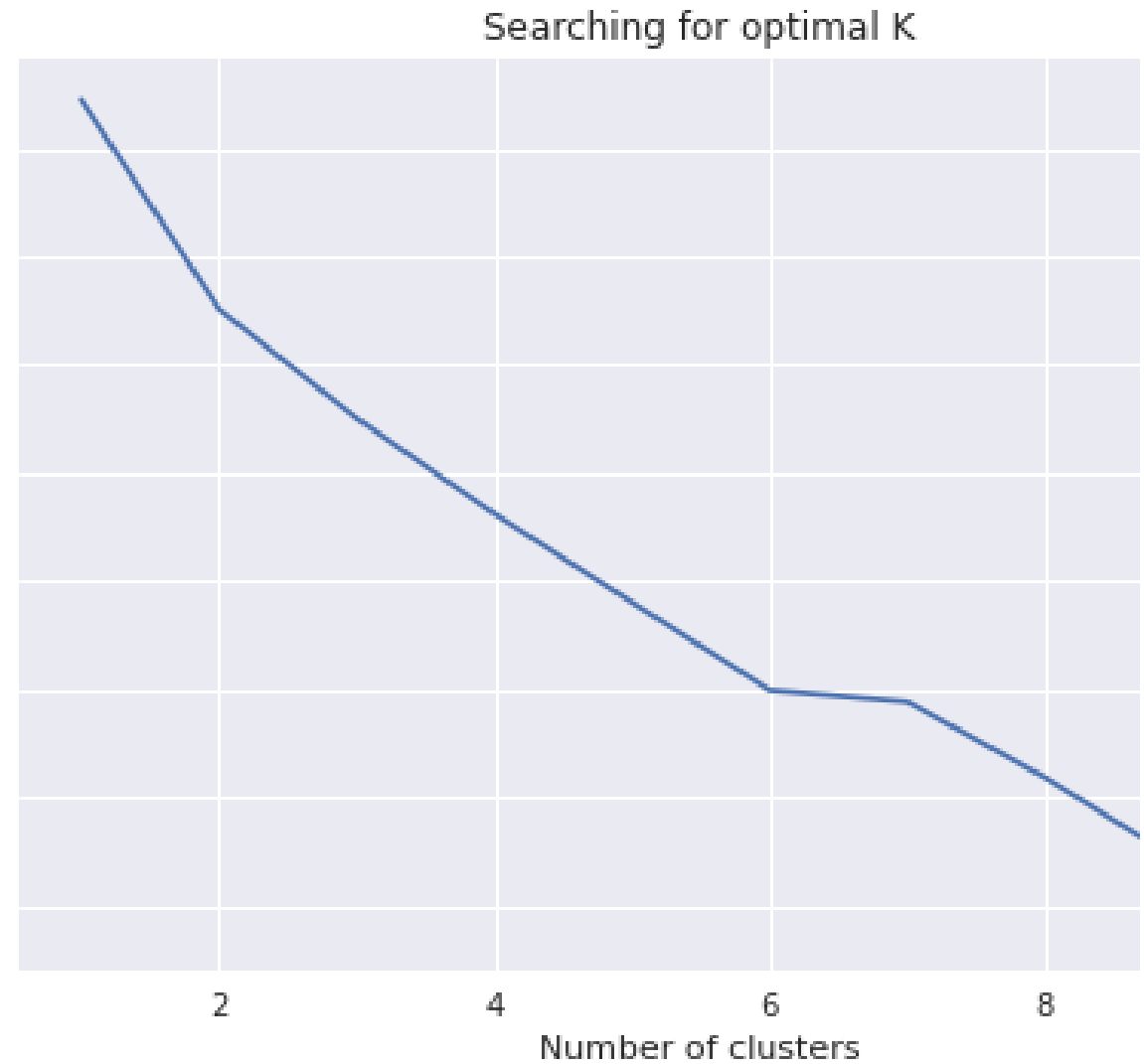
- Neighborhoods obtained through Coursera's link
 - Venues obtained through the Foursquare API
 - Venues as features with one hot encoding
 - Selecting optimal K based on elbow plot
 - Create K clusters
- 



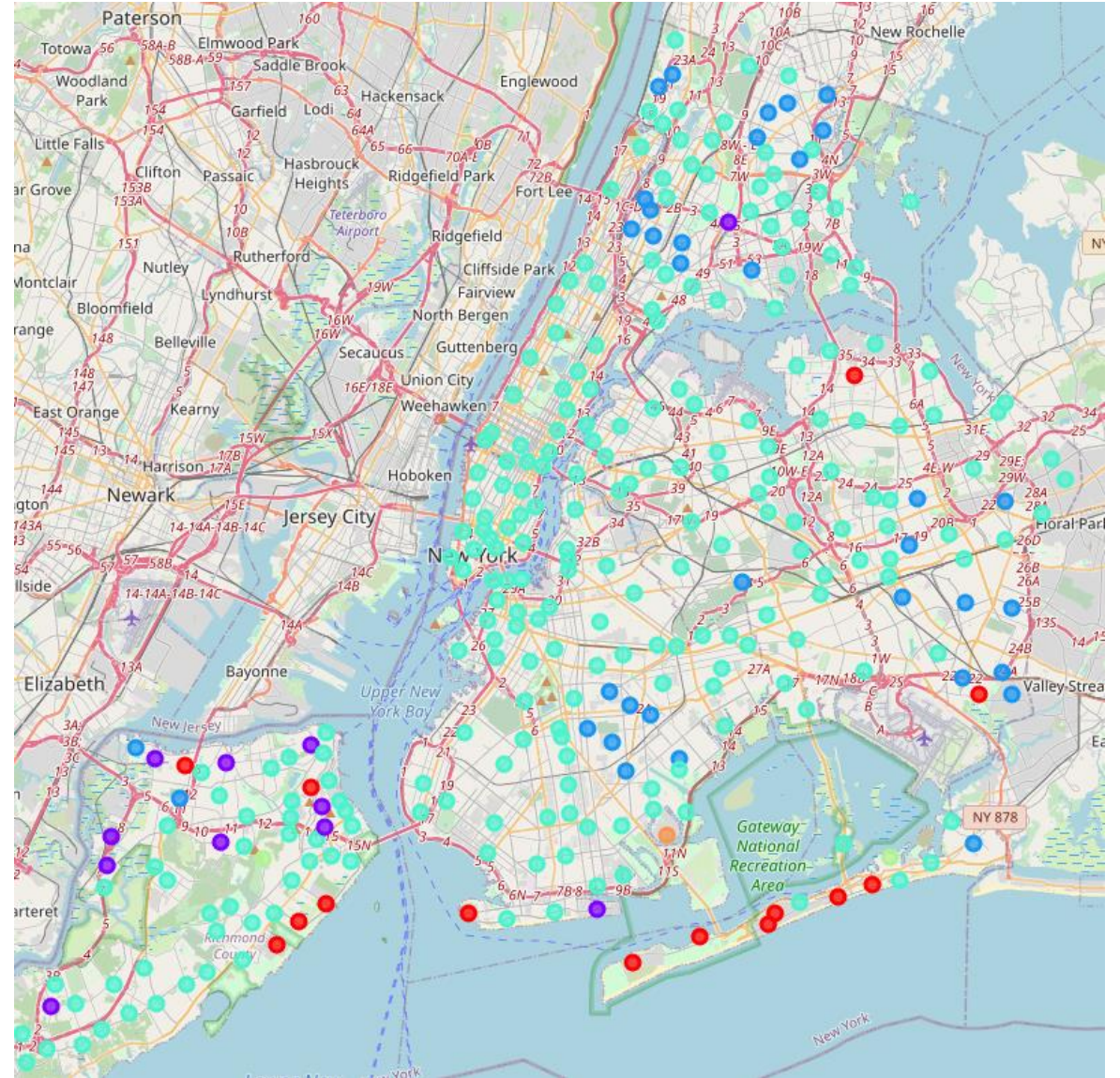
Map of New York with neighborhoods



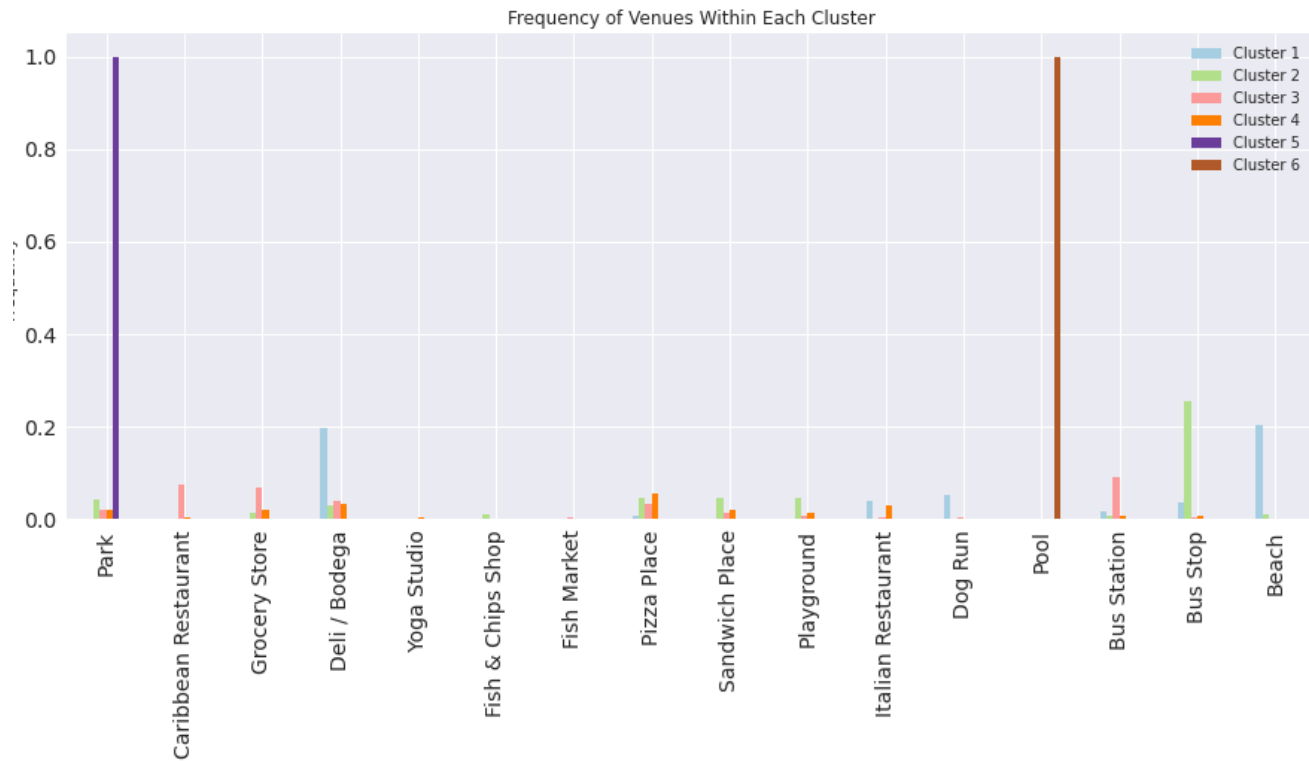
Elbow plot to select K (k=6)



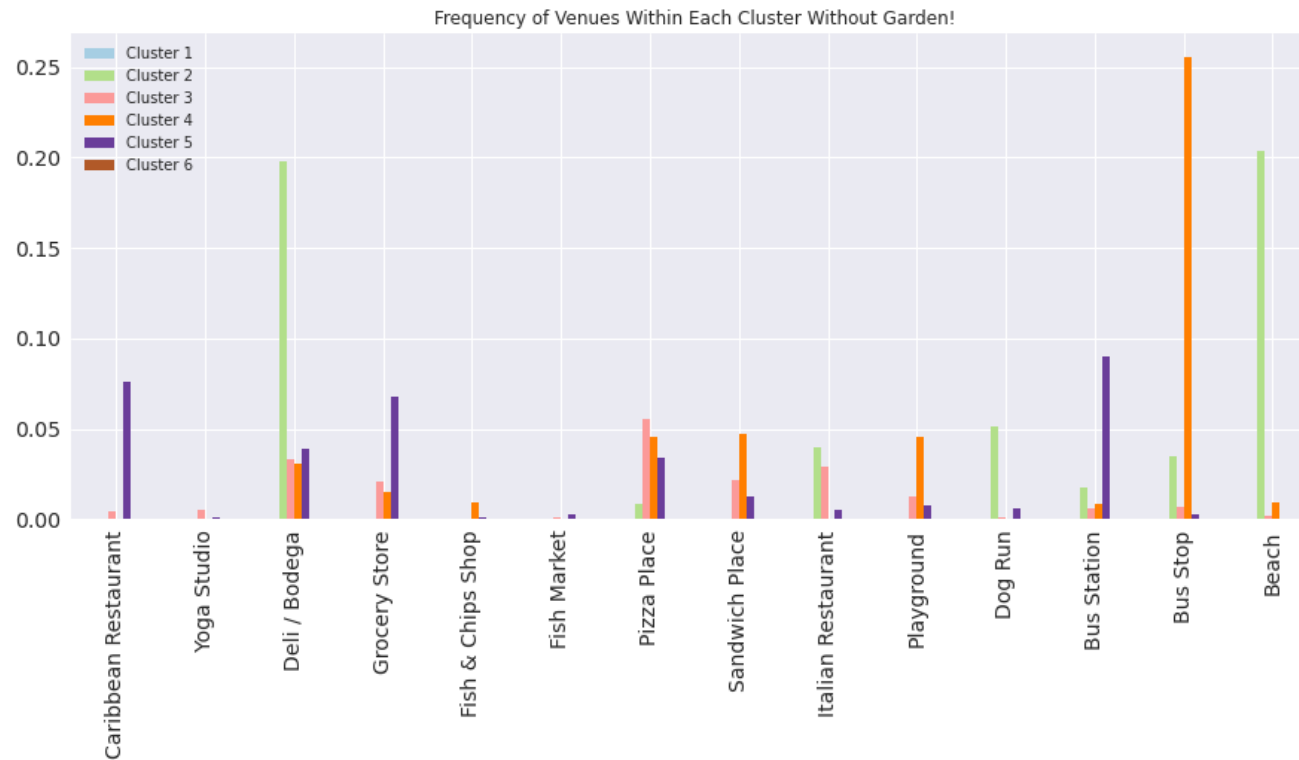
Map of NYC with clustered neighborhoods



Occurrence of venues in clusters



Occurrence of venues in clusters (No pool or parks)





Conclusion



- NYC neighborhoods cluster in 6 distinct groups based on venues
- Several clusters have sandwich places in them (cluster 3,4,5)
- Best neighborhood to open up a new sandwich shop has:
 - Most sandwich places in the similar neighborhoods (cluster 4)
 - But no sandwich places in that particular neighborhood

BEST PLACE: Allerton, the Bronx

