Capstone Project - The Battle of Neighborhoods

# Introduction

This essay looks into data and to identify if there are any shifts of the most popular venues of the New York City over the past 3 years. The data is based on records from the Foursquare.com. A better understanding of the venue popularity developments will help entrepreneurs to have a better understanding of the potential investment opportunities. The essay can mainly provide insights about the type of venues and the locations.

# Data Selection

The essay selects three time points as samples to study the most popular venues: 2016/3/1, 2017/3/1, 2018/3/1, 2019/3/1 and 2020/3/1. It is interesting to see if there is any shifts of the most popular venues over the years. This may provide a hint about people’s taste change, and selection change, or their spendable income change. For each year, 20575 entries were

# Methodology

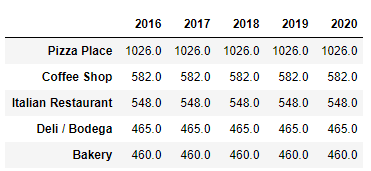
The data is pulled from the Fourquarters.com. For each venue in each Neighborhoods, a venue category is assigned. The data is then grouped by venues to calculate the top 10 most popular venues of the selected year.

By adding the popularity scores of the venues of each year, it is then possible to make some comparisons.

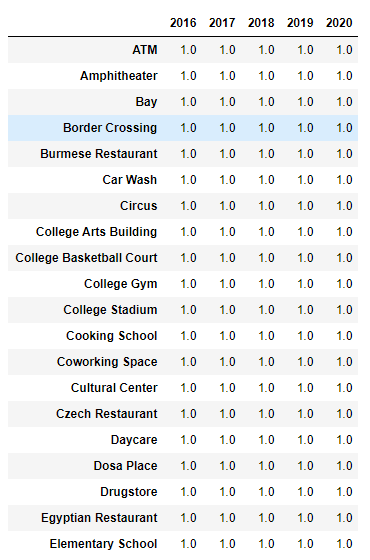
# Result

## 4.1 Calculate the Most and Least Popular Venues

It is found that the top 5 most popular venues over the years are Pizza Places, Coffee Shops, Italian Restaurants, Deli / Bodega and Bakeries. The popularities remain static over the years.

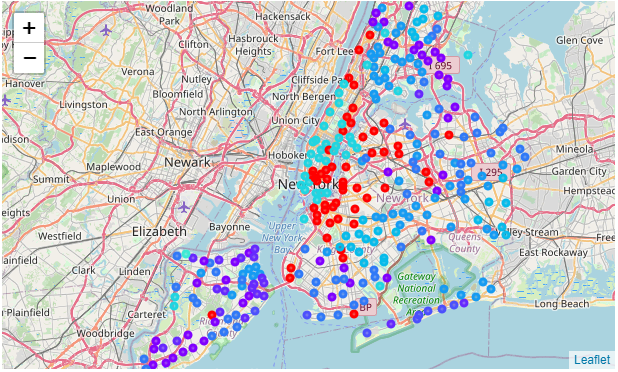
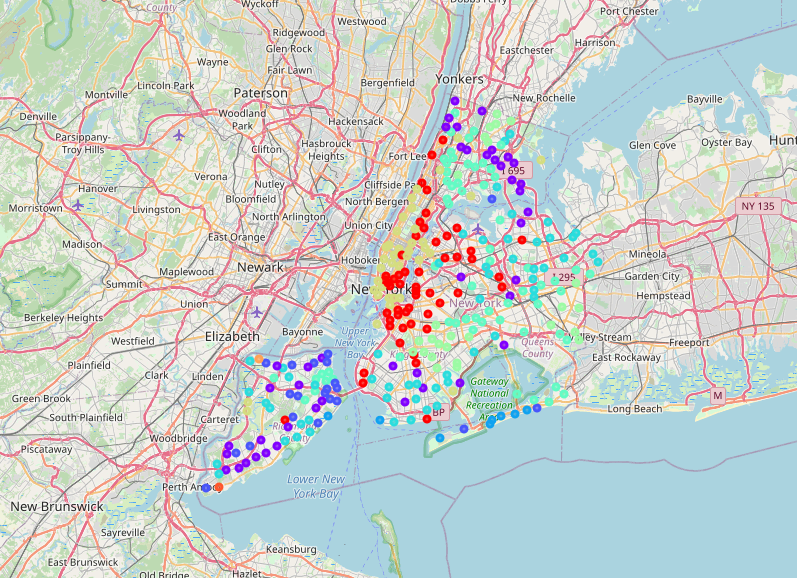


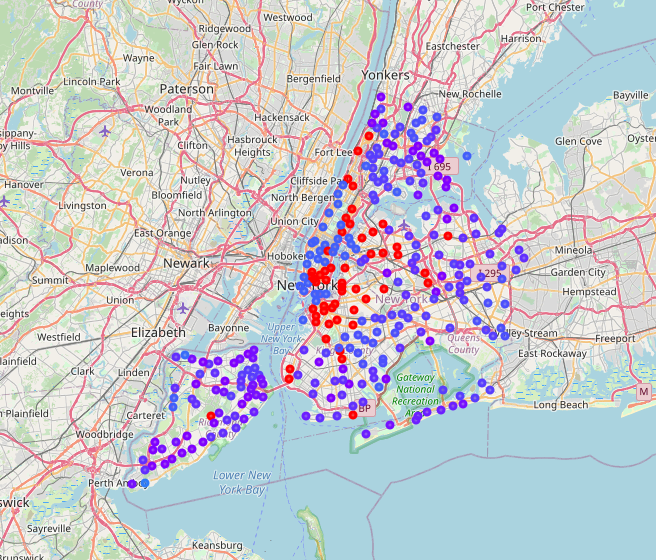
The least popular venues are ATMs, Amphitheater, Bays, Border Crossings, Burmese Restaurants, Car Washes, Circus, College Arts Buildings, College Basketball Courts, College Gyms, Cooking School, Czech Restaurants, Daycares, Dosa Places, Drugstores, Egyptian Restaurants and Elementary Schools.



## Clustering the Neighborhoods

Take the map of NY2016 as an example, when K = 10, K = 20 and K = 50. The map look like below.





# Discussion

It is interesting to find that, the most popular venues over the past 5 years remain the same. One possible explanation is that, the NY city has such a matured and stable culture that the default choices are at the steady state.

For the least popular venues, it can be interpreted as, some of the venues are very “specific”, such as college facilities; some are for minorities of the population, such as Egyptian Restaurants. Others are very functional based, such as ATMs and Drug Stores. Decisions about which ATM or Drug Stores to go for is normally distance based and with less preferences to the special “promotions”, therefore are not searched as many times as the popular venues.

About the choice of the number of Clusters and the outcome.

It seems when K = 10, it gives a more segmented clustering comparing with K = 50.

# Conclusion

NYC is a matured and well-developed city with a stable taste among its citizens. Over the past five years, the popularities of venues remain the same. If someone decide to move from one Neighborhood to another, he or she doesn’t need to worry about the development of the local venues will bring any big uncertainties to the new home. Pizza, Coffee and Italian Restaurants are the most favorite places to visit.