THE BATTLE OF THE NEIGHBORHOODS

Author: Paul Weirich

Outline

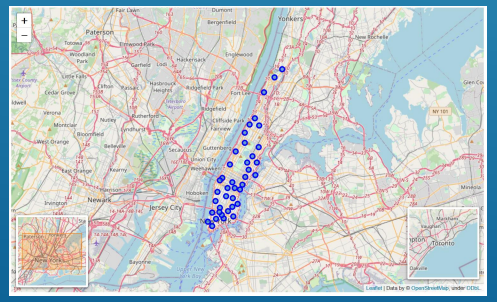
- Motivation
- Objectives
- Method
- Results
- Conclusion

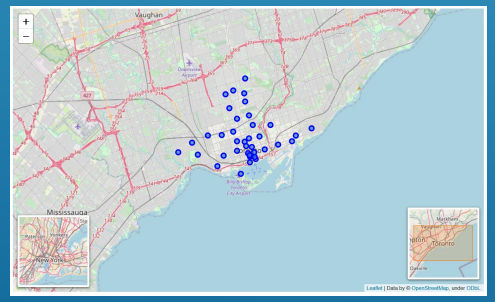
Motivation

- Person wants to move from New York to Toronto

- Assumption:
  - The person does not much about Toronto
  - Wants to find a neighborhood with similar venues
  - Does not want to spend time with collecting information, comparing etc.

- Is there a way to to make a smart and efficient decision for finding a suitable neighborhood in the new city?



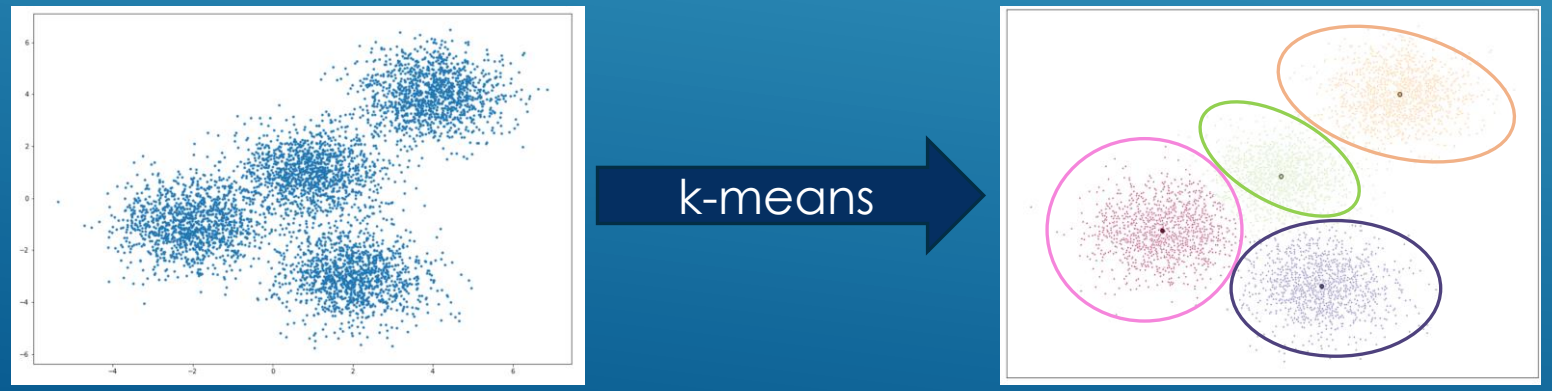


Objectives

- Development of system to find similarities of neighborhoods
- Order neighborhoods by smiliarities
  - → Clustering
- Help User to make a good decision
  - Clear graphical representation

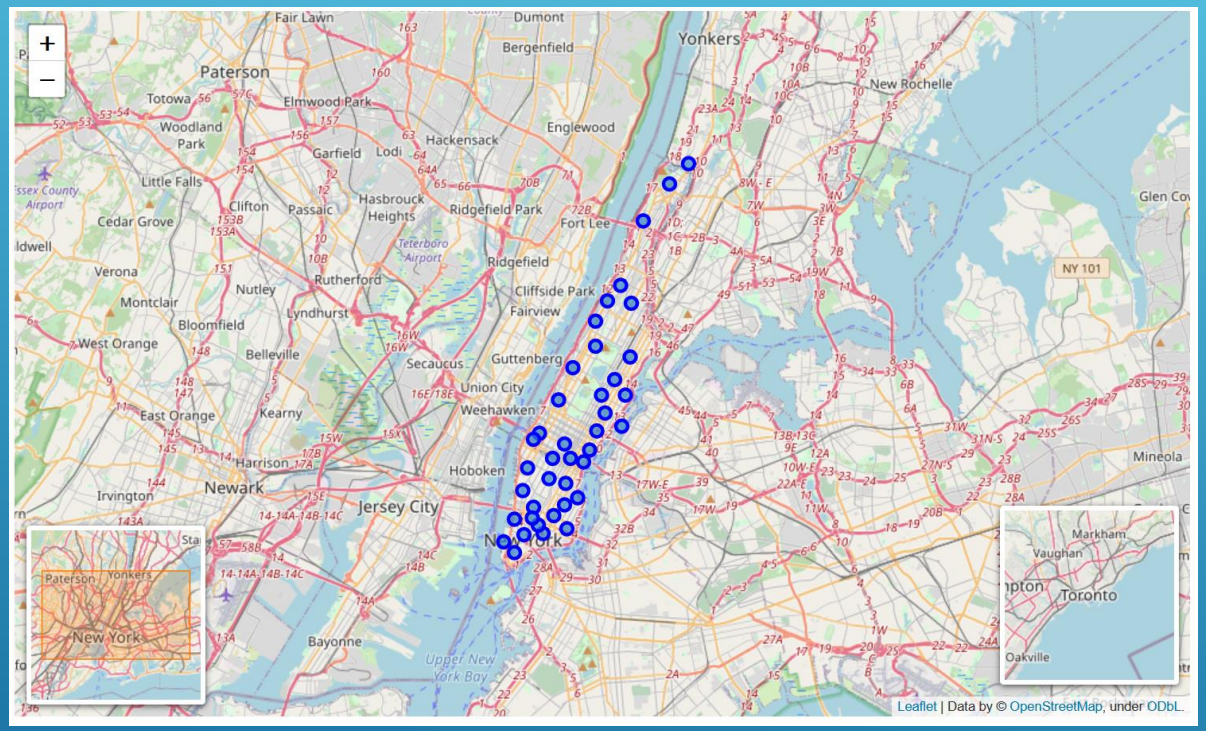
Method

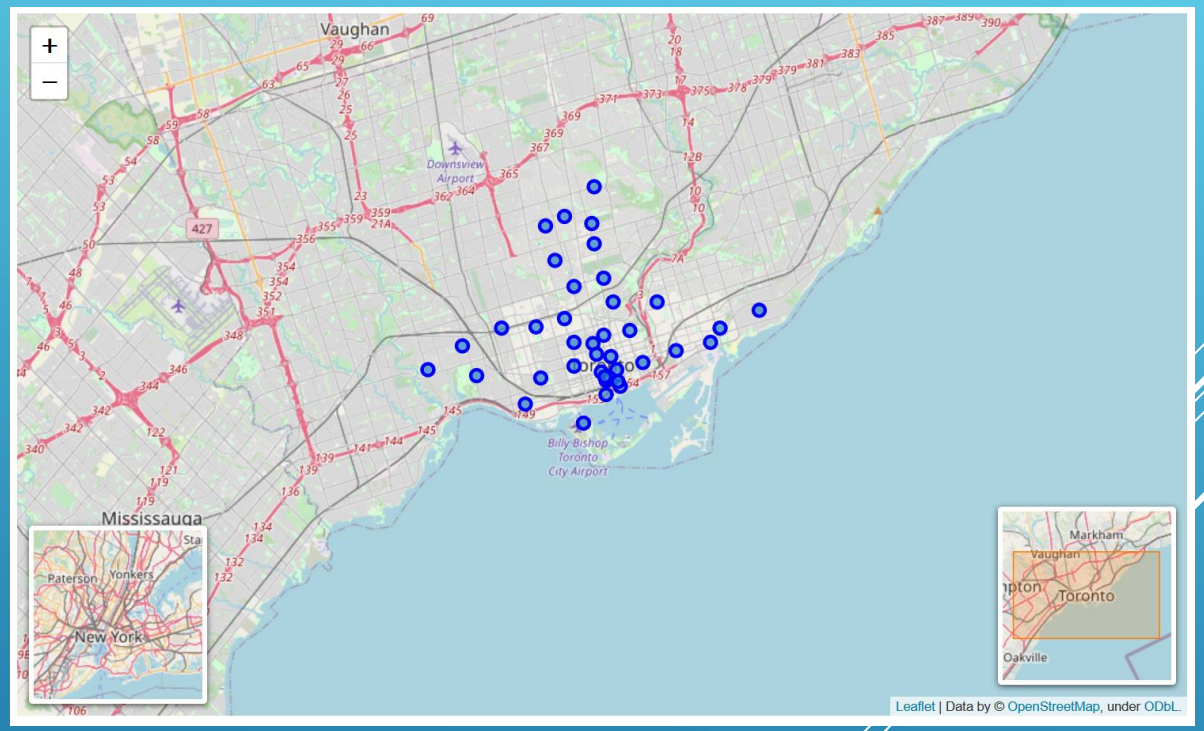
- Acquire neighborhood data via downloading
- Use Foursquare API information about venues for each neighborhood
- Encode categories of venues using One Hot
- Use k-means algorithm for finding similarities
  - Find number of k‘s for clusters with elbow method



Results

Geographical Location

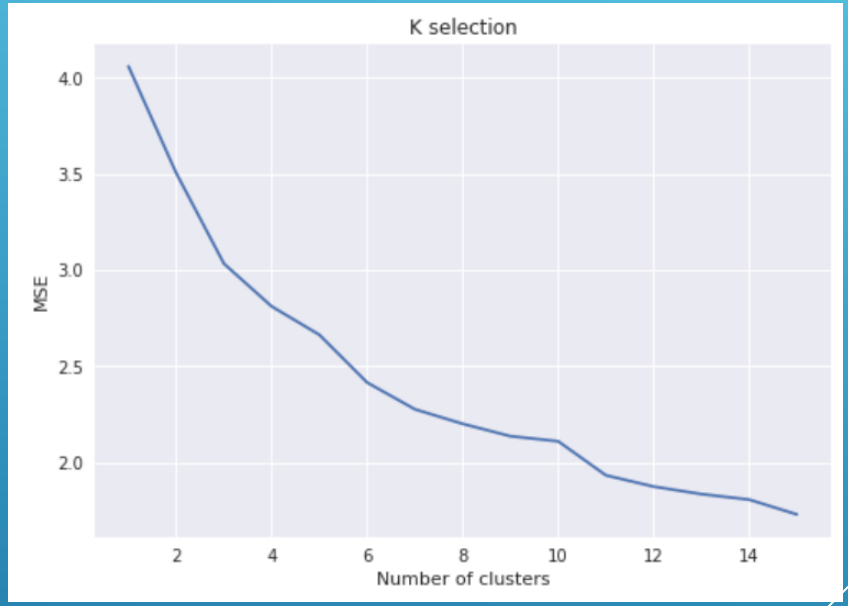




Results

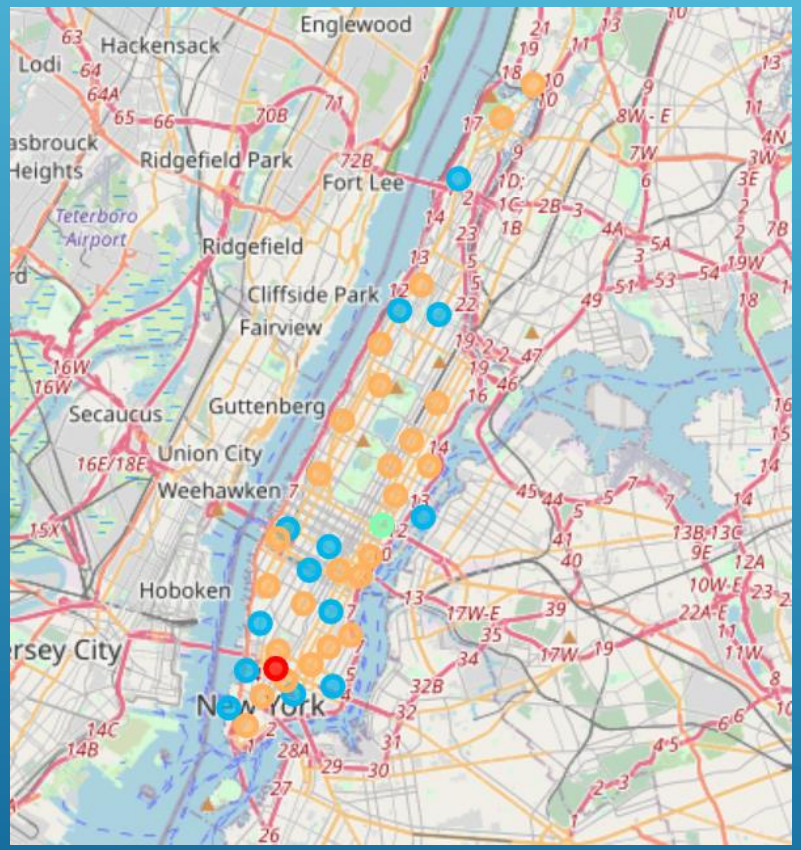
Finding appropriate number of clusters with k-means

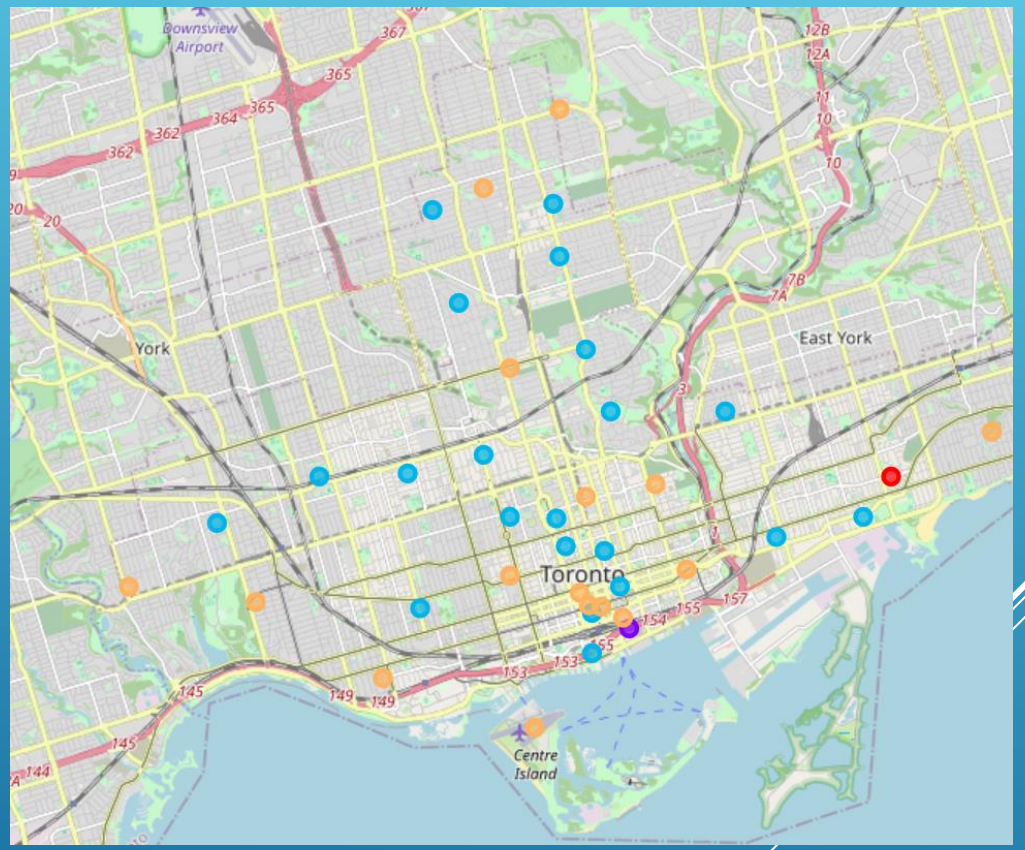
- Elbow is located around 5, therfore it is a value for the number of clusters



Results

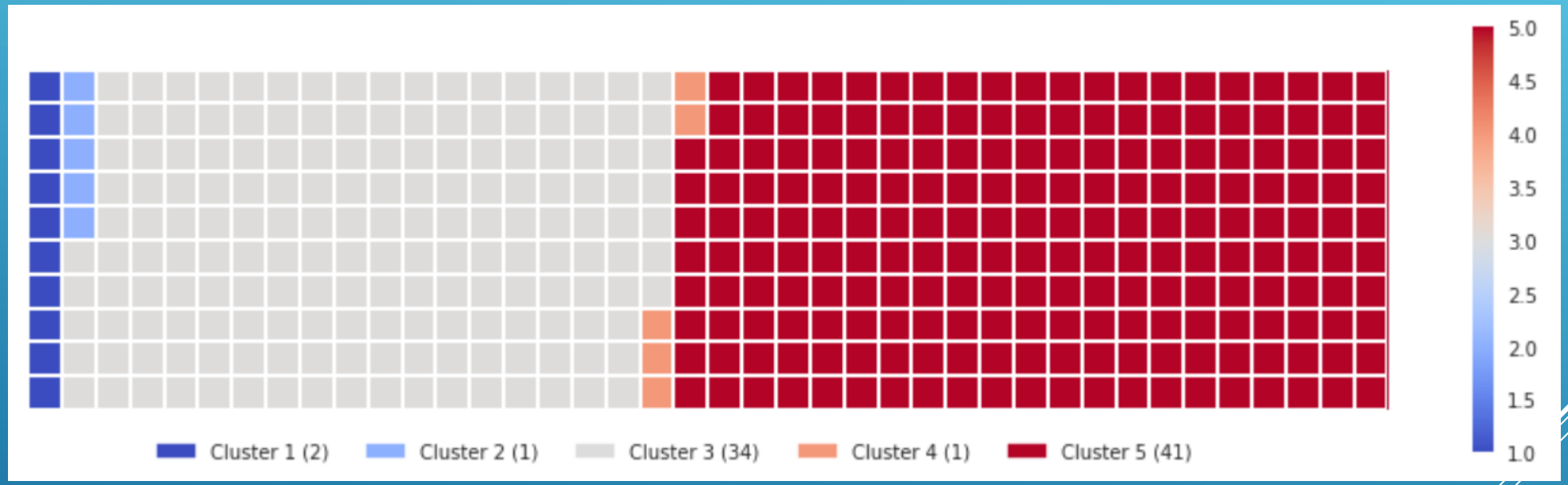
Geographical location of clustered neighborhoods





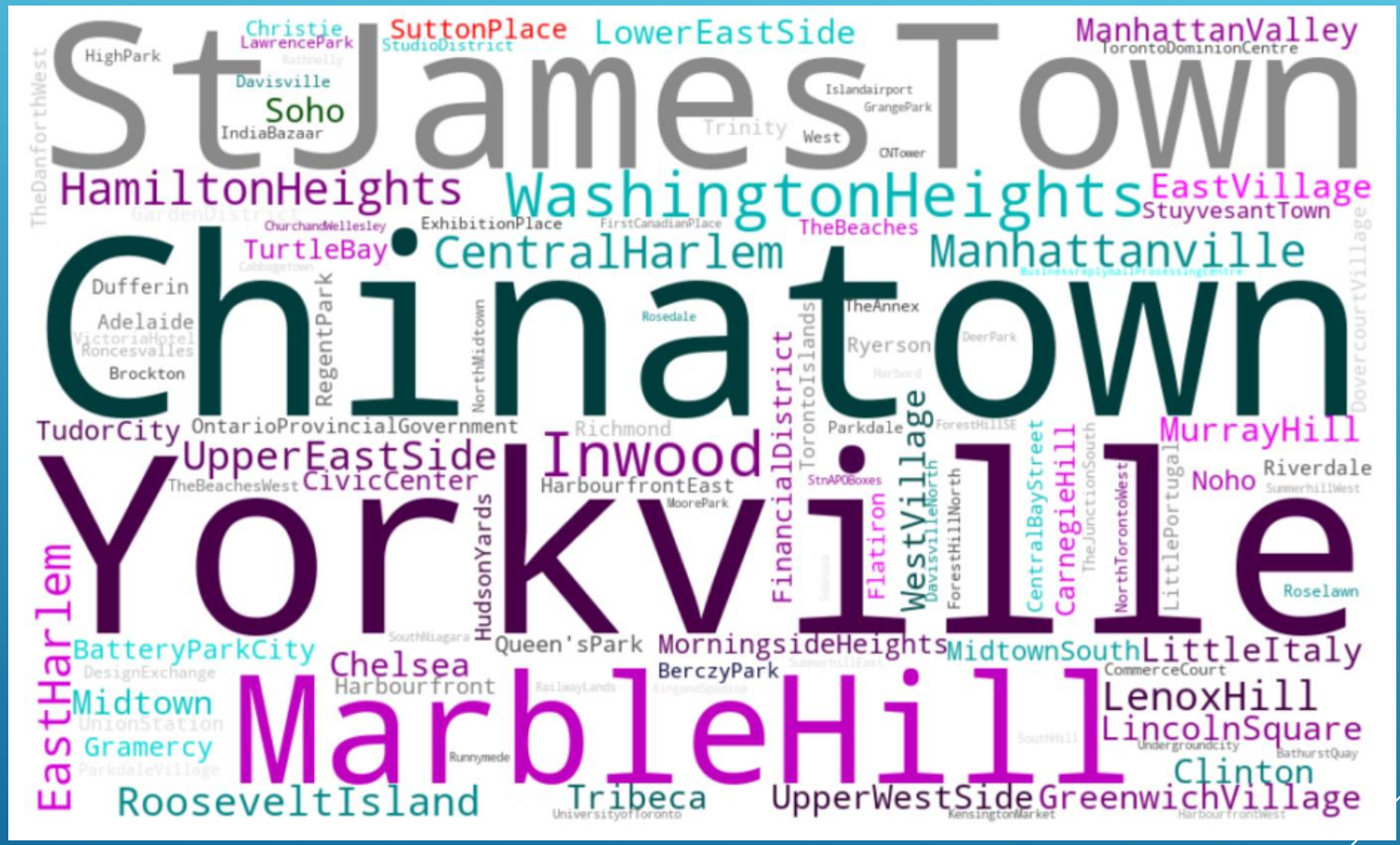
Results

Ratio of segmented Data



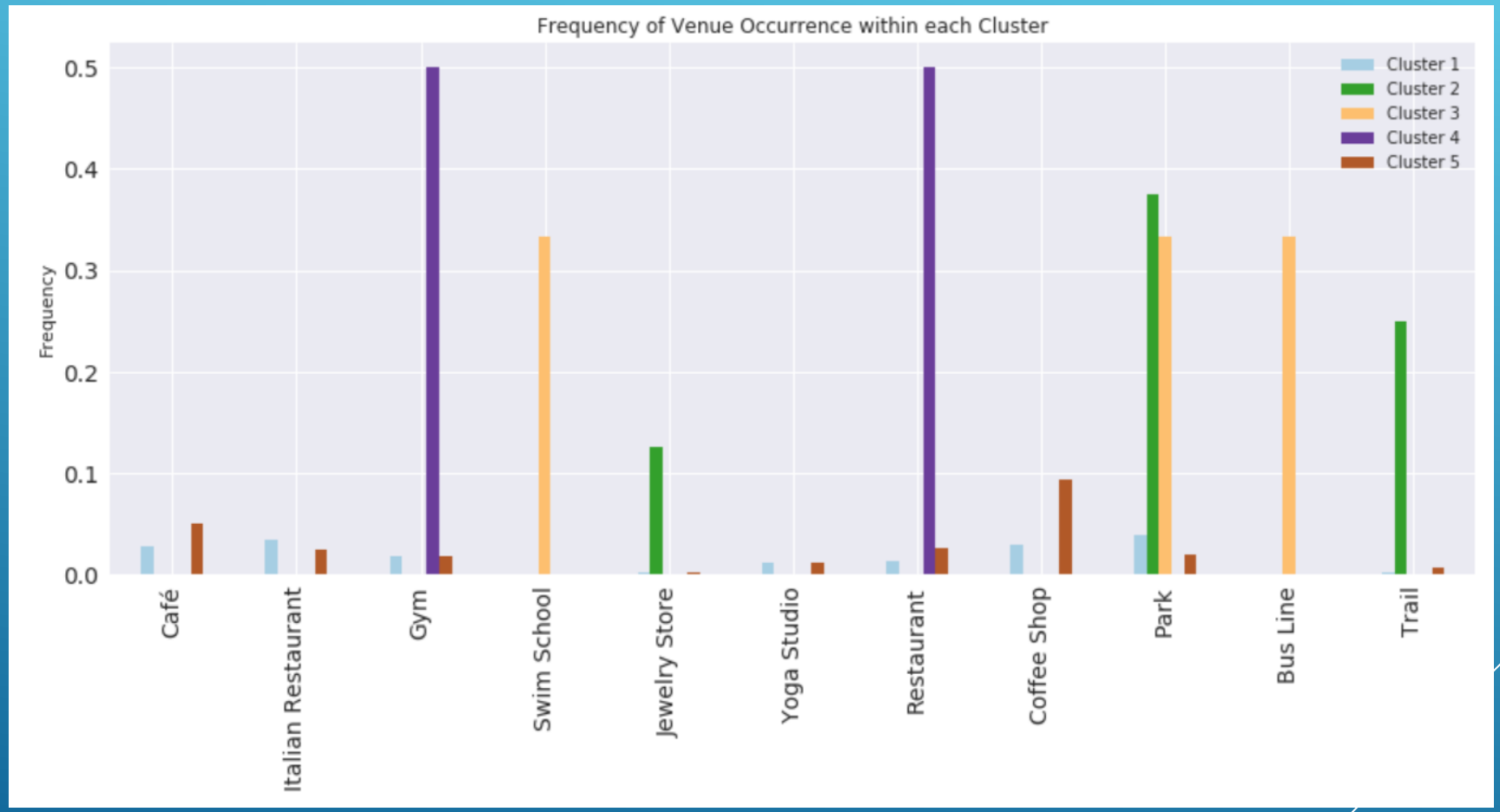
Results

Word Cloud with segmented Neighborhoods sorted by color



Results

Bar Chart with frequency of Venues in cluster



Conclusion

- **Cluster 1**: Neighborhoods that have Cafés, Italian Restaurants, Coffee Shops and Parks around
- **Cluster 2:** Neighborhoods that have Jewelry Stores, Parks and Trails around
- **Cluster 3:** Neighborhoods that have Swim Schools, Parks and Bus lines around
- **Cluster 4:** Neighborhoods that have Gyms and Restaurants around
- **Cluster 5:** Neighborhoods that have Cafés and Coffee Shops around