ANALYSIS OF CULTURAL BELIEFS BASED ON POPULAR VENUES

Esha Tulsian

Background

◦ More than **195** countries in the world and **1000s** of different cultures

◦ A city's layout is planned beforehand to **reflect cultural traditions** and values

Focus on 3 Cities of the World

NEW YORK CITY, USA

TORONTO, CANADA

SEOUL, SOUTH KOREA

Business Problem

1. Compare the 3 cities and how each represents their unique culture
2. Analyze which venue type is most common

Interest

1. **Businesses**: understand the culture and what the population likes

2. **General Public**: learn about cultural comparisons

Data Sources

NYC
- [Postal Code and Geospatial Data](#)
- Foursquare API

Toronto
- [Postal Code](#)
- [Geospatial Data](#)
- Foursquare API

Seoul
- [Postal Code](#)
- Latitude and Longitude data from Geopy
- Foursquare API

DATA CLEANING

Data Cleaning Steps

NYC

1. Select columns "Borough", "Neighborhoods", "Latitude", and "Longitude"

2. Create a dataframe

Toronto

1. Scrape Wikipedia and create dataframe

2. Clean data

    1. Remove NA values
    2. Make sure each row is a different postcode

Seoul

1. Scrape Wikipedia

2. Use Geopy Library to get Latitude and Longitude Data

Features

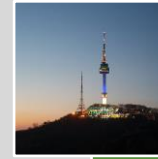

NYC
- Borough
- Neighborhood
- Latitude
- Longitude

- 5 Unique Boroughs
- 306 Unique Neighborhoods

Toronto
- Borough
- Neighborhood
- Postcode
- Latitude
- Longitude

- 10 Unique Boroughs
- 103 Unique Postcodes

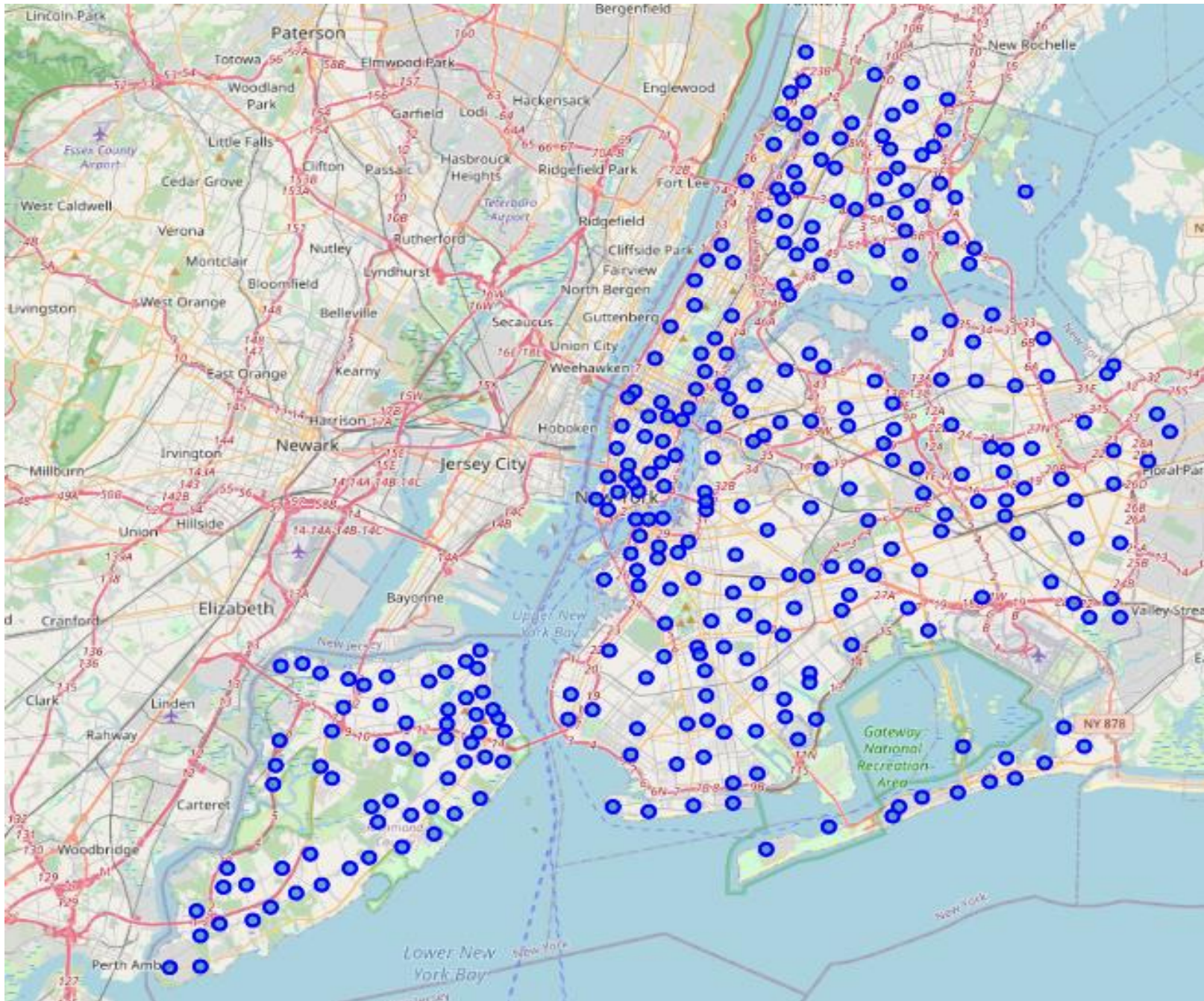
Seoul
- Postcode
- District
- Latitude
- Longitude

- 25 Unique Districts

EXPLORATORY DATA ANALYSIS

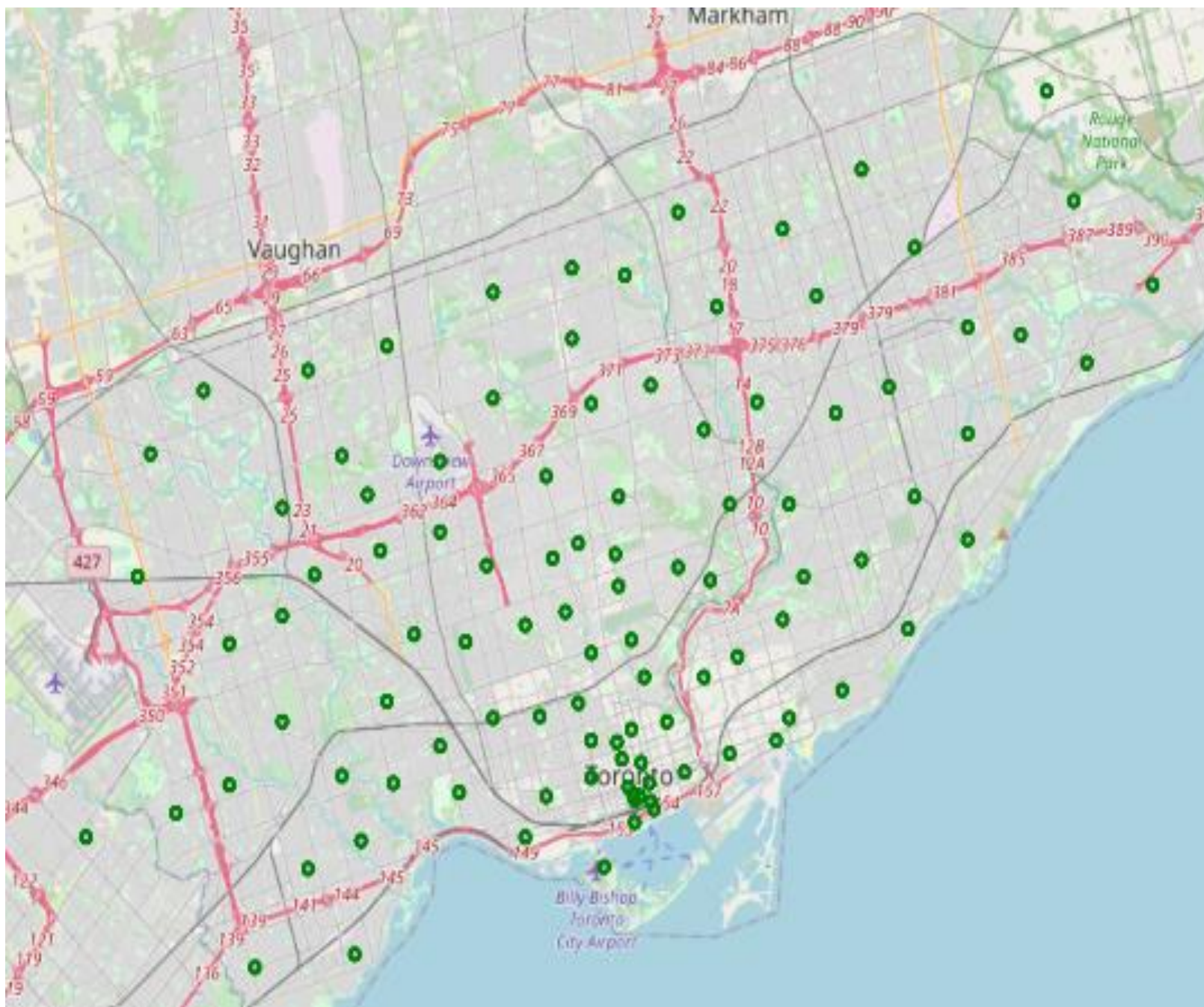# Steps

1. Plotted Neighborhoods on a Map
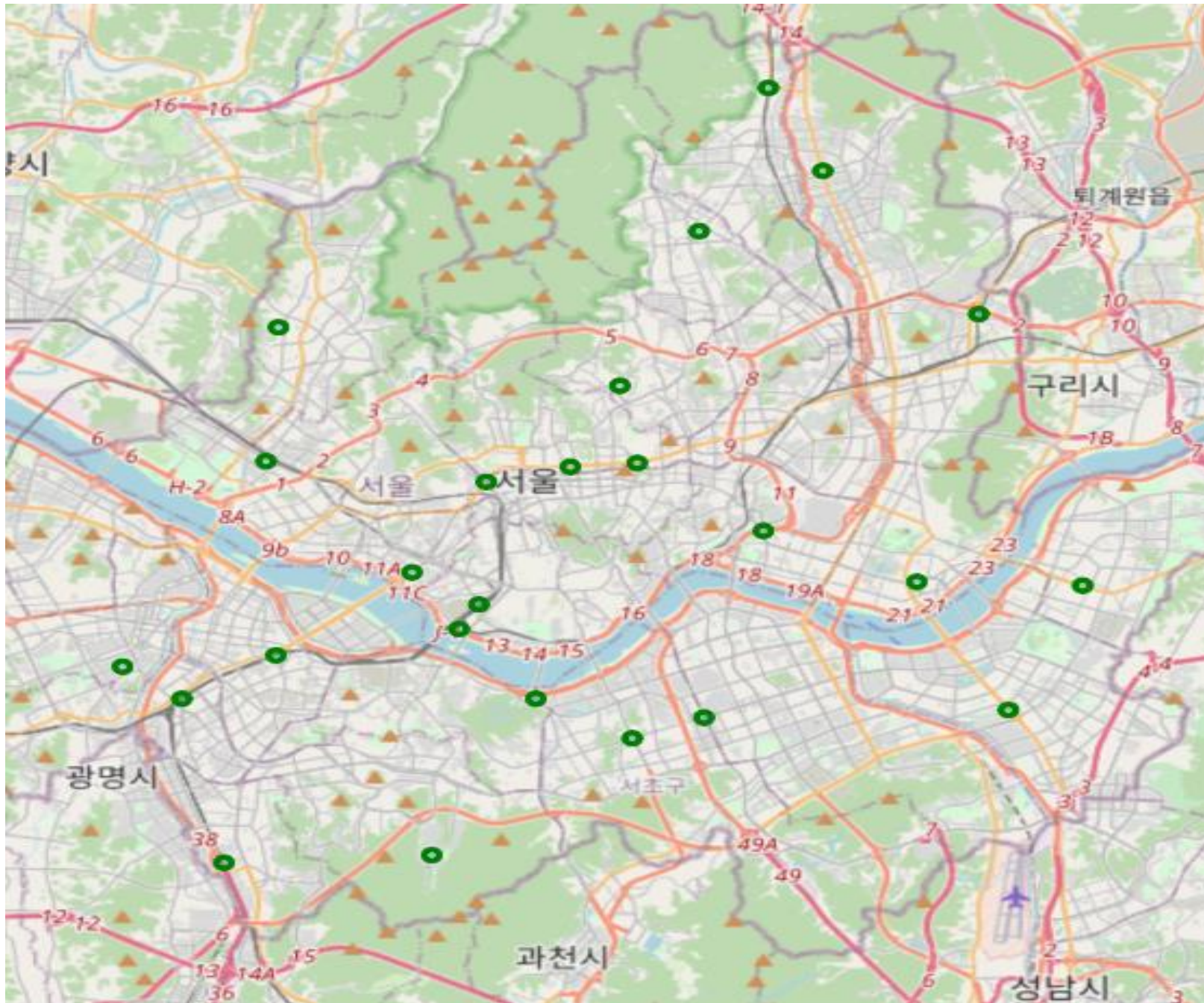2. Used Foursquare API to get venue data

New York City

- 10032 Total Venues
- 421 Unique Venue Categories

Toronto

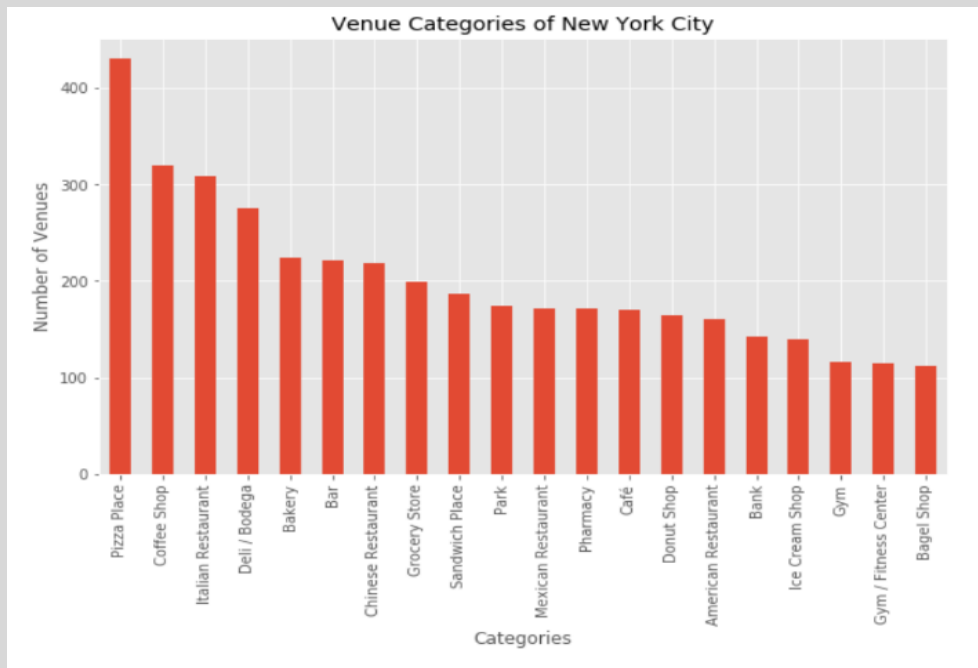- 2124 Total Venues
- 269 Unique Venue Categories

Seoul

- 791 Total Venues
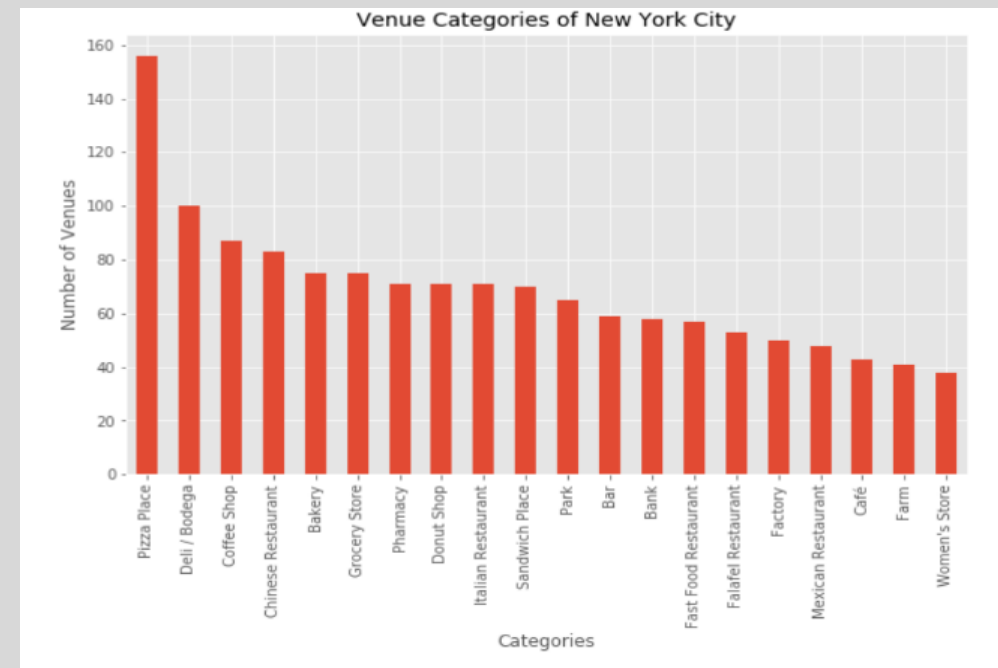
- 124 Unique Venue Categories
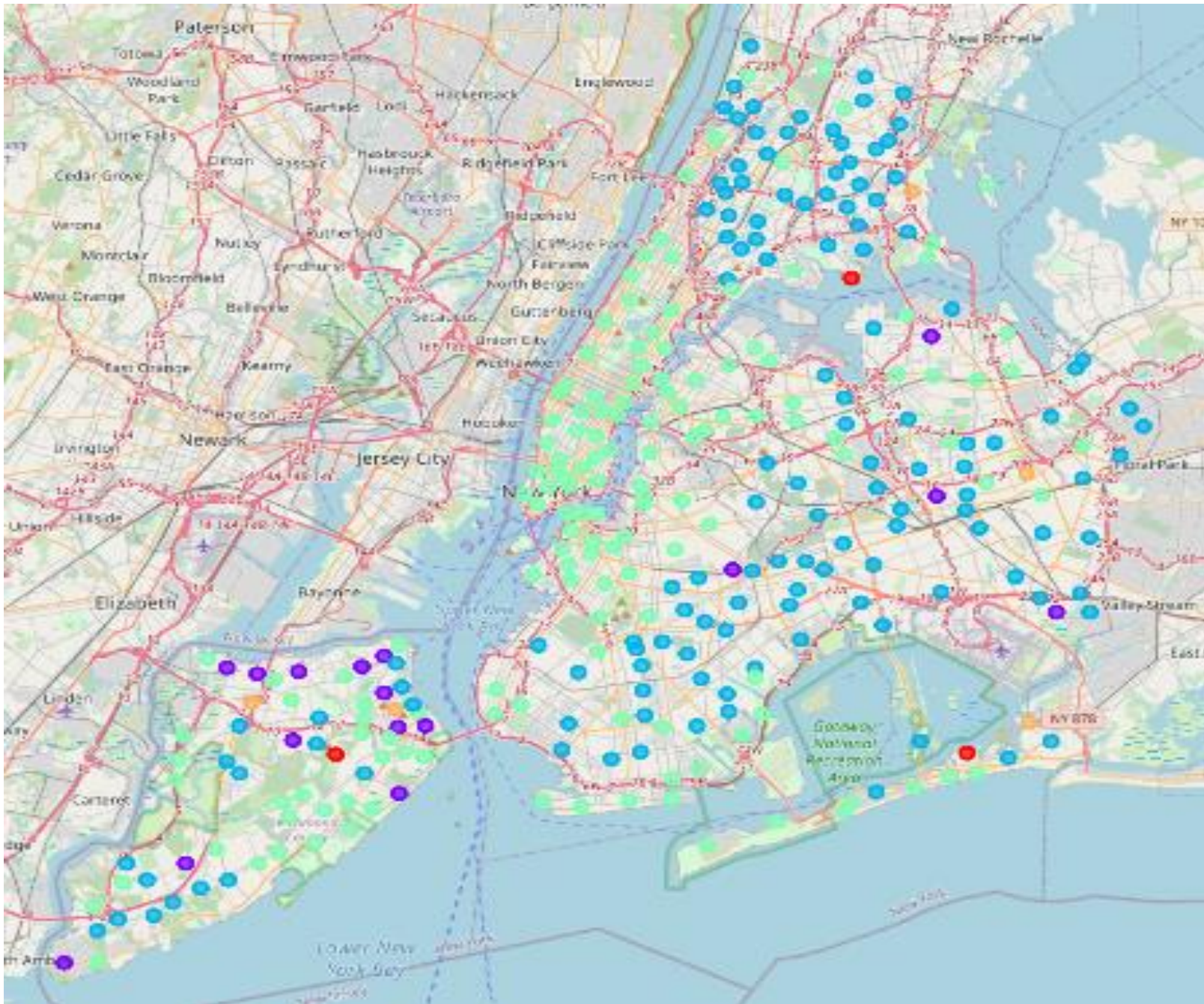
RESULTS

New York City

Top 20 Most Common Overall Venue Types

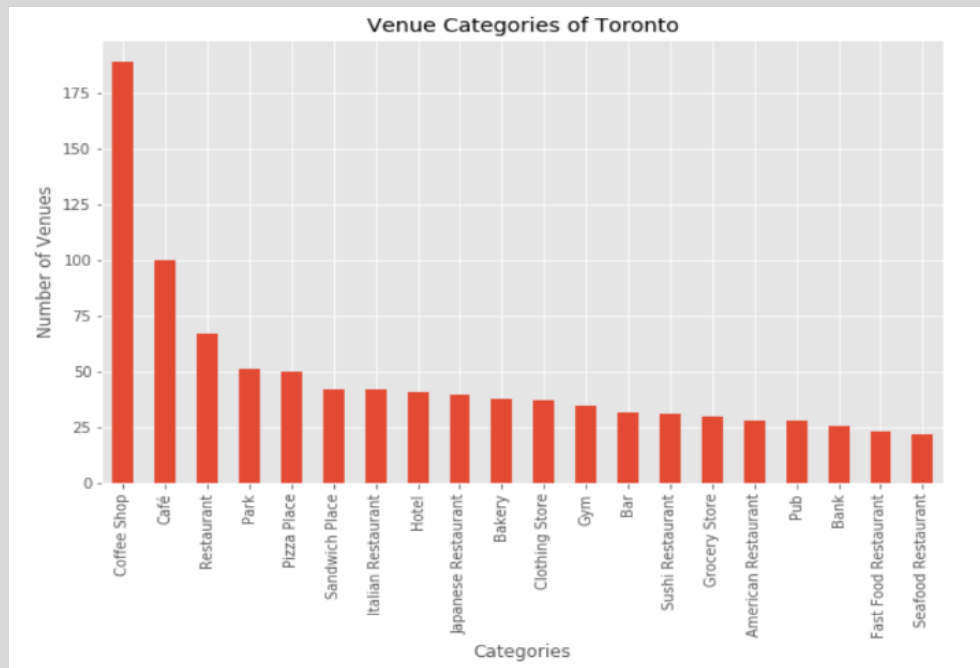**Top 20 Venue Types after looking at Top 10 Venues of Each Neighborhood**
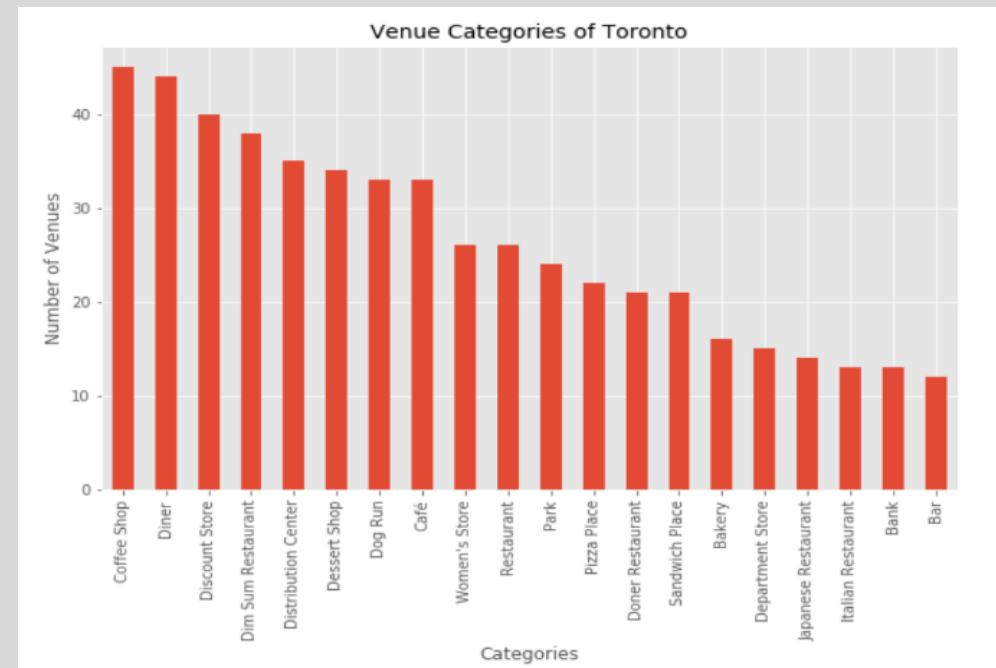
New York City

- Cluster 0
  - Red
  - 3 Neighborhoods
- Cluster 1
  - Purple
  - 16 Neighborhoods
- Cluster 2
  - Blue
  - 133 Neighborhoods
- Cluster 3
  - Green
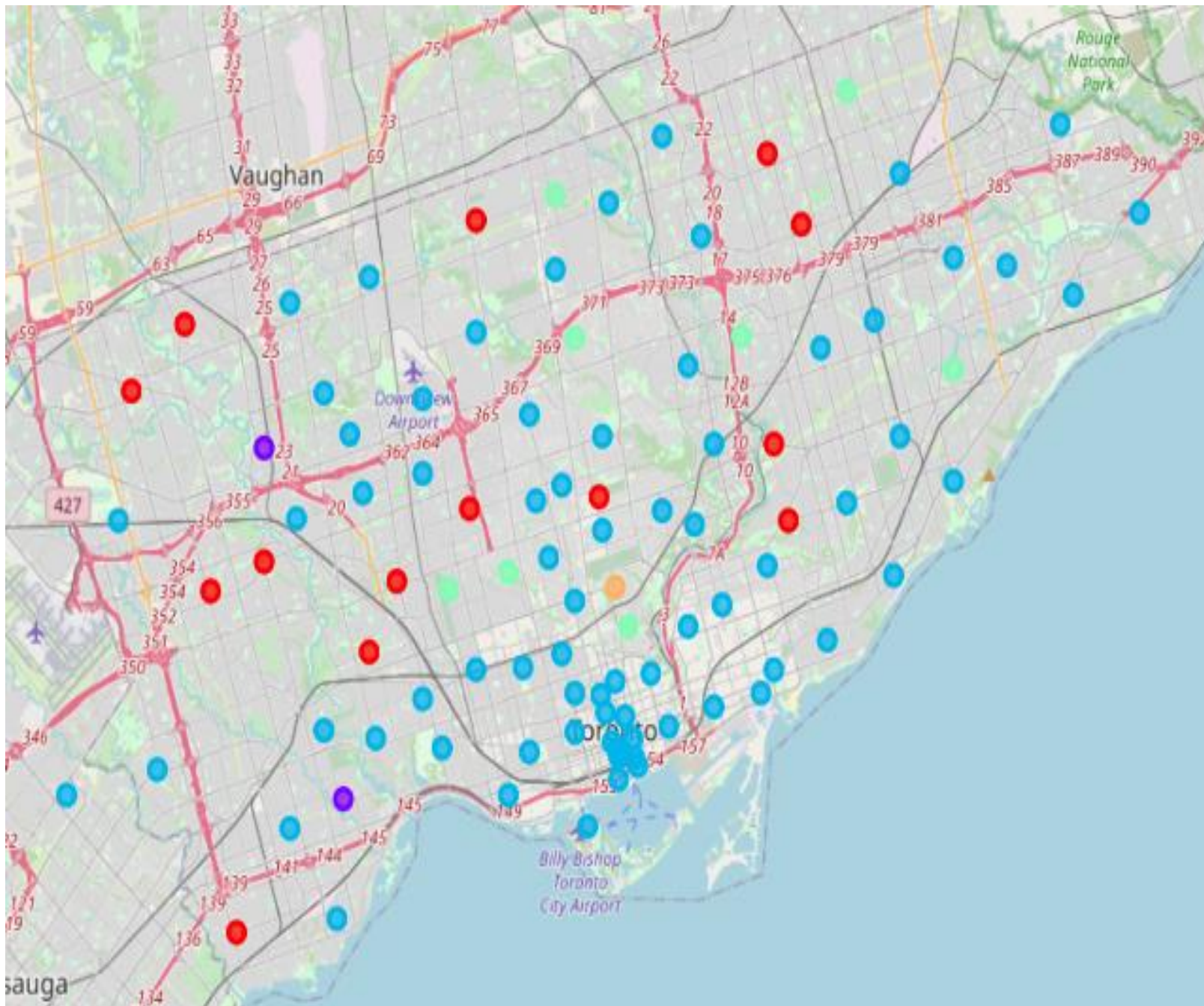 - 148 Neighborhoods
- Cluster 4
  - Orange
  - 5 Neighborhoods

Toronto

Top 20 Most Common Overall Venue Types

Top 20 Venue Types after looking at Top 10 Venues of Each Neighborhood
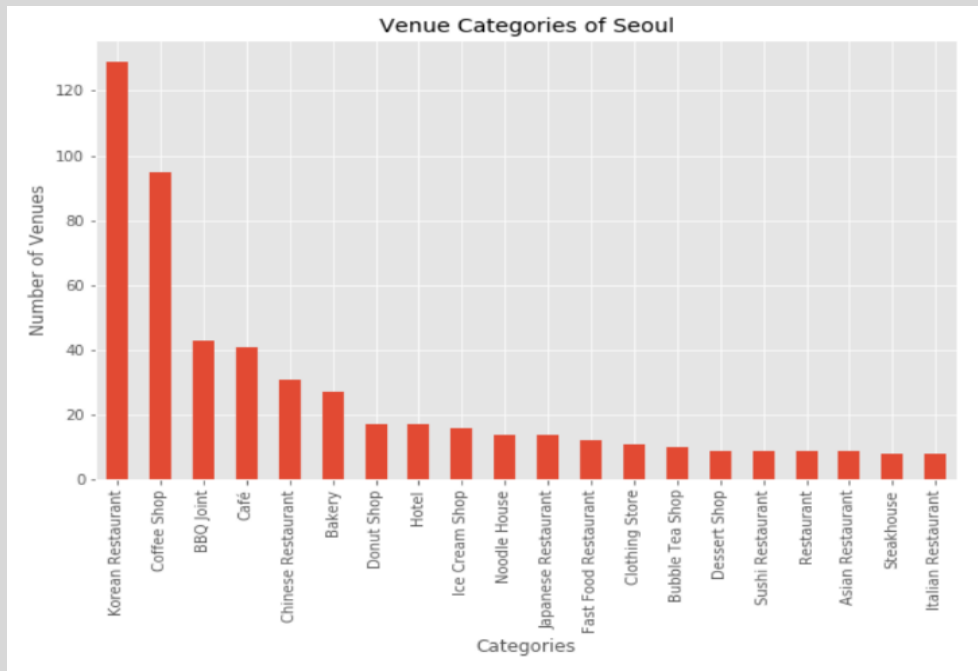
Toronto

- Cluster 0
  - Red
  - 14 Neighborhoods

- Cluster 1
  - Purple
  - 2 Neighborhoods

- Cluster 2
  - Blue
  - 74 Neighborhoods

- Cluster 3
  - Green
  - 8 Neighborhoods

- Cluster 4
  - Orange
  - 1 Neighborhood

Seoul

Top 20 Most Common Overall Venue Types

Top 20 Venue Types after looking at Top 10 Venues of Each Neighborhood
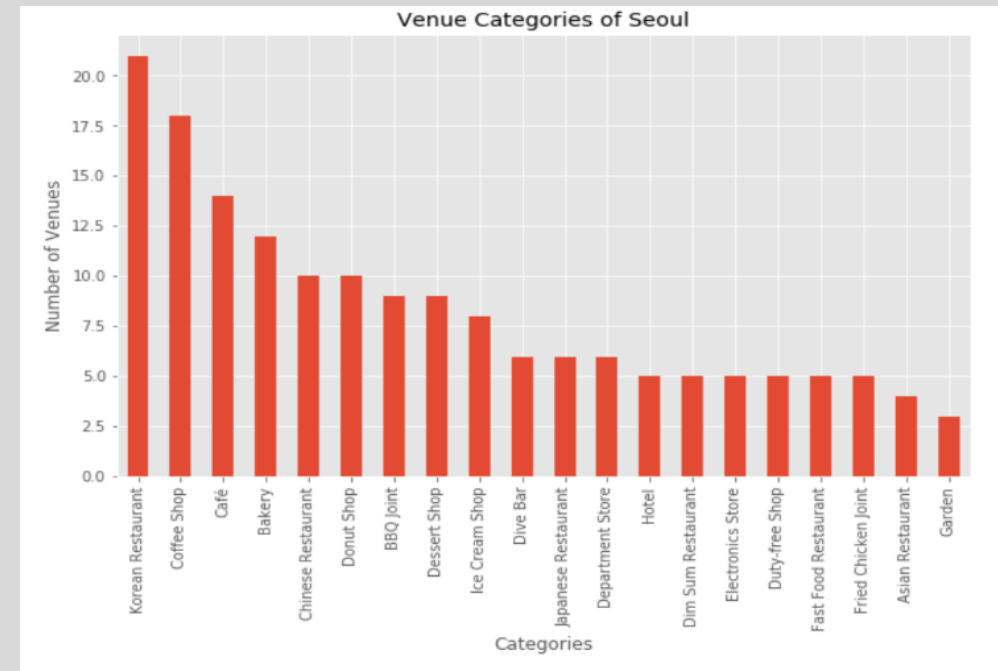
Seoul

- Cluster 0
  - Red
  - 1 Neighborhood
- Cluster 1
  - Purple
  - 10 Neighborhoods
- Cluster 2
  - Blue
  - 12 Neighborhoods
- Cluster 3
  - Green
  - 1 Neighborhood
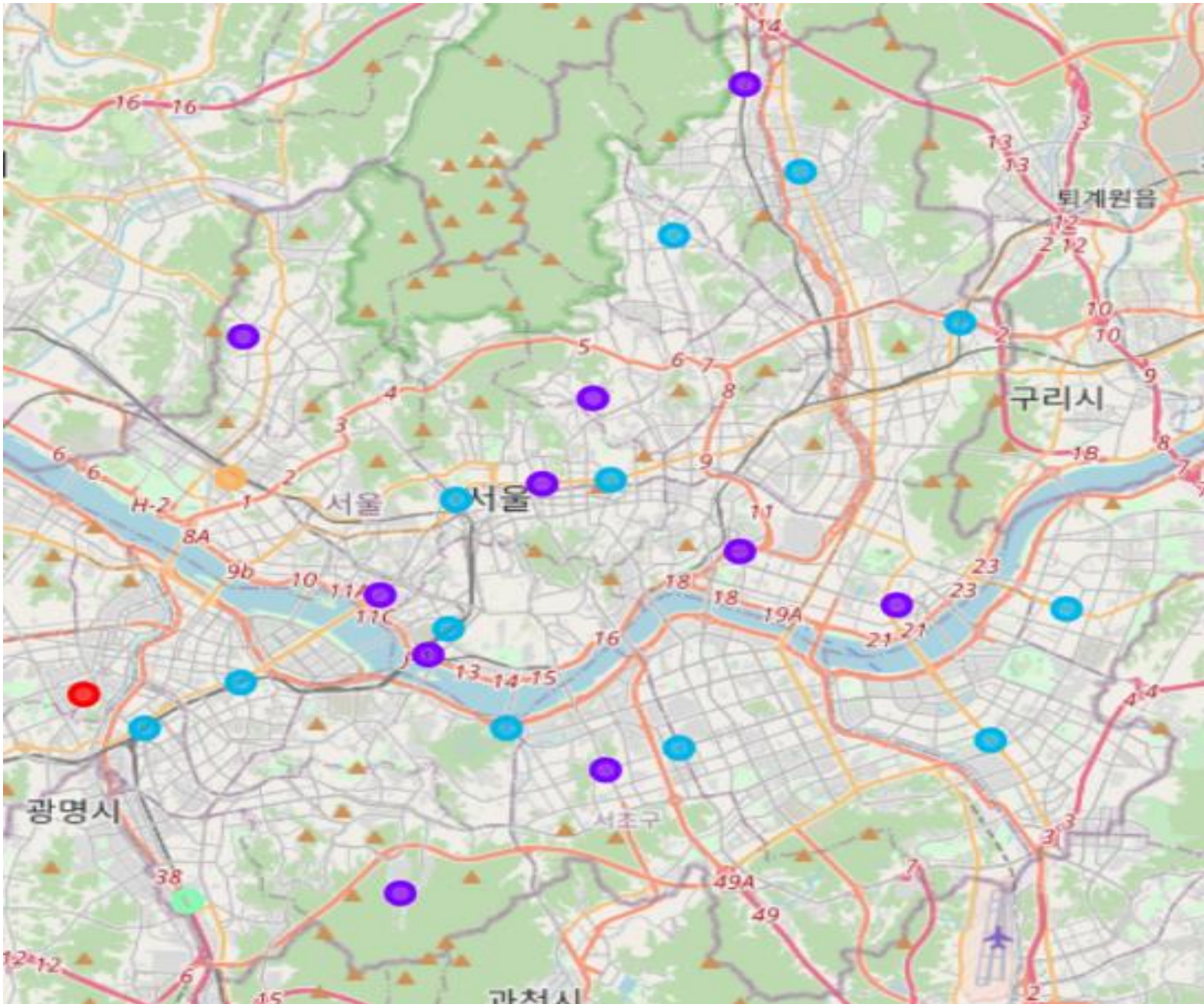- Cluster 4
  - Orange
  - 1 Neighborhood

Discussion

- **Coffee shops or cafes** are frequented by all the cultures
  - Suggests that coffee shops are important in all 3 cultures
  - Provides for a small comfy place for customers to enjoy their coffee while doing work or socializing
- **Cultural Aspect:** Though they all frequent them, it does not mean the décor of the coffee shop or items in the coffee shop are the same
  - For instance, in NYC, they have coffee shops like Starbucks. However, in Seoul, they have themed cafes based on popular cartoons or locations.
- All the cities had one or more food and drink places as their top 3
- Toronto: only city with a non-food venue - **discount store**
  - **Shopping** plays a big role in their culture
- NYC: makes sense that **pizza places, Italian restaurants, and deli/bodegas** are frequented because NYC is known for its delicious pizza, pasta, and subs.
- Seoul: **more Korean restaurants** as the general population is Korean. Korean food is known for having delicious BBQ, which is reflected by **BBQ joints** being frequented often as well.

Conclusion

◦ The whole data science process was followed

  ◦ The steps were identifying the business problem, specifying and cleaning data required, performing exploratory data analysis, performing machine learning by clustering, and analyzing the results

◦ To answer the business question,

  ◦ each city is different and the top venue categories reflect the culture

  ◦ For instance, coffee shops are popular in all cities but decorations and menu items should be related to the culture

THANK YOU FOR YOUR TIME!!

Hope you enjoyed!