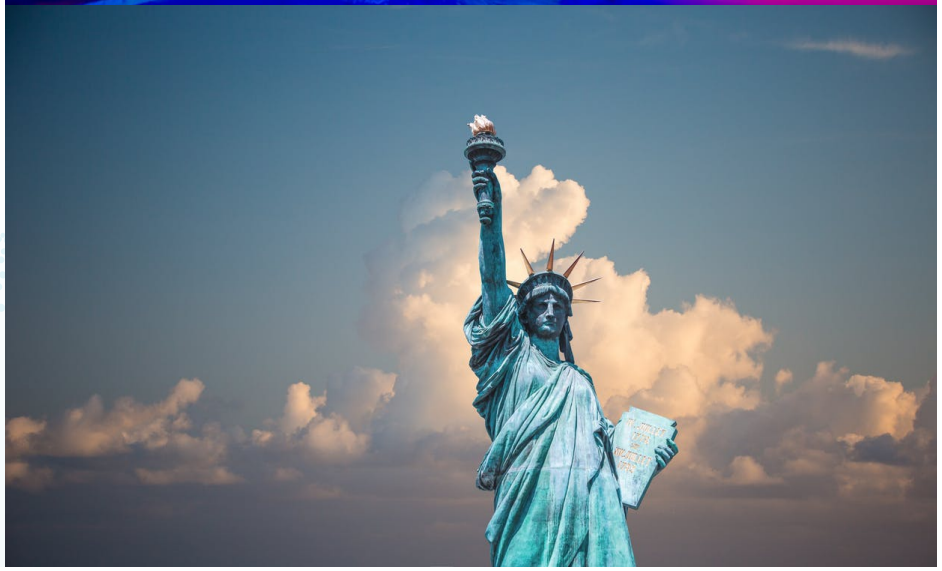
Locating A New Chinese Restaurant in New York City or Toronto

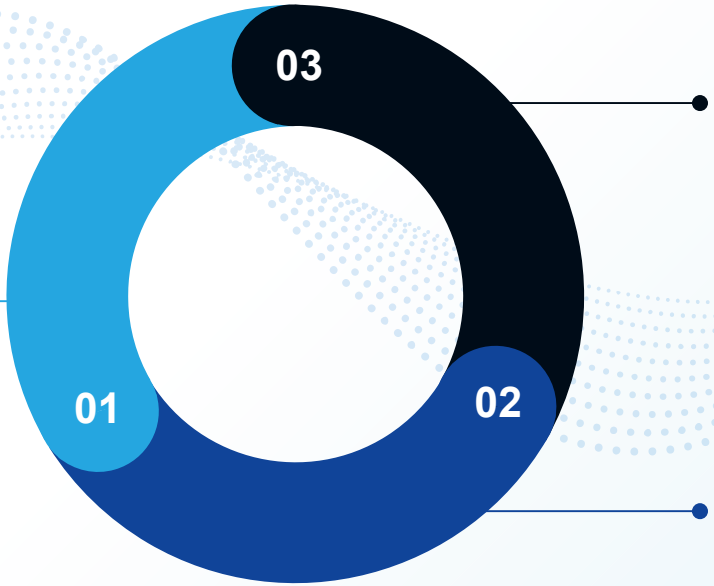
Overview

Determine the best neighborhood for a new Chinese restaurant in TO or NYC base on the following:

- Above-average income
- Undersaturated by Chinese restaurants
- Have above-average population in the city



Key Execution Steps



Data Sourcing & Cleaning

Sourcing & understanding data

Cleaning data

Exploratory data analysis

ML Cluster Analysis

Cluster mapping

Oversaturation evaluation

Neighborhood Analysis

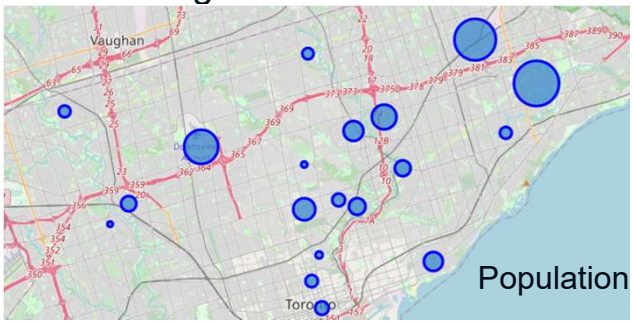
Comparative analysis

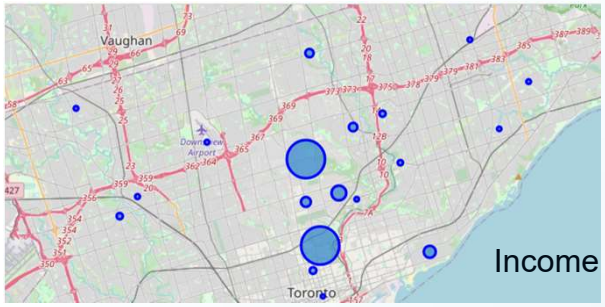
Data Sourcing

- Toronto neighborhood postal codes: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- Toronto neighborhood geospatial data: https://cocl.us/Geospatial_data
- Toronto neighborhood population and income data: https://en.wikipedia.org/wiki/Demographics_of_Toronto_neighbourhoods
- New York City neighborhood location data: https://cocl.us/new_york_dataset
- New York City neighborhood population data: https://data.cityofnewyork.us/api/views/swpk-hqdp/rows.csv?accessType=DOWNLOAD
- 5. New York City neighborhood income data: https://ny.curbed.com/2017/8/4/16099252/new-york-neighborhood-affordability

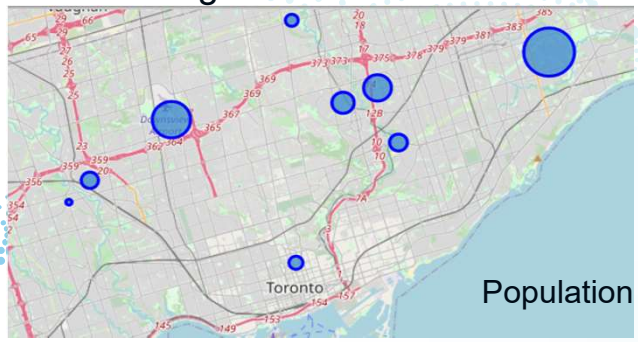
Data Cleaning – Toronto

Toronto neighborhood data before cleaning



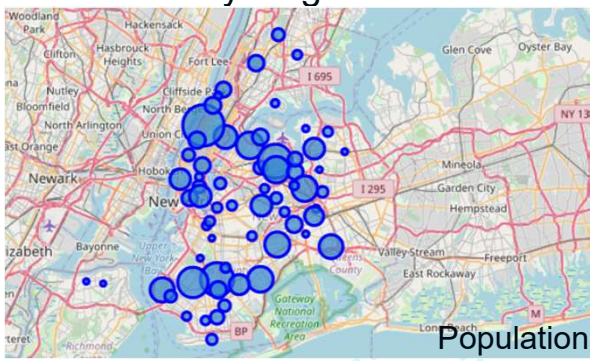


Toronto neighborhood data after cleaning with inter-quartile filter

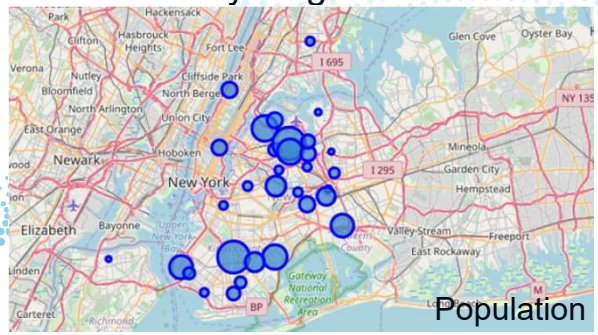


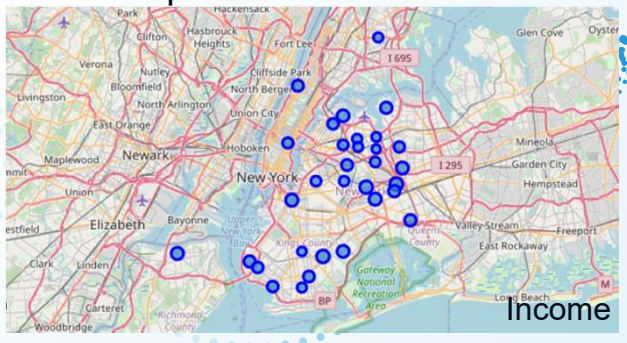
Data Cleaning – New York City

New York City neighborhood data before cleaning



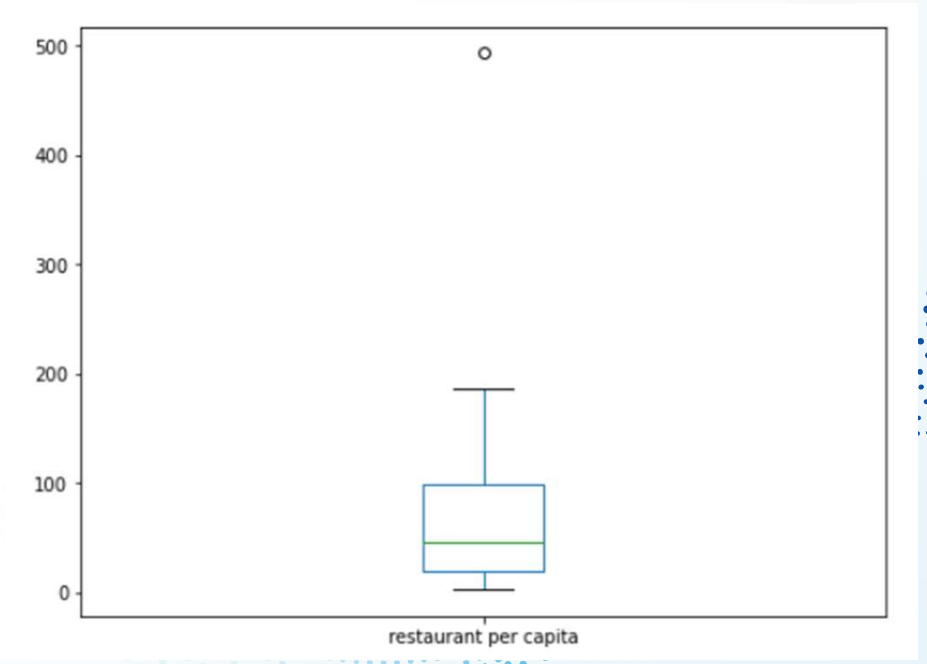
New York City neighborhood data after cleaning with inter-quartile filter





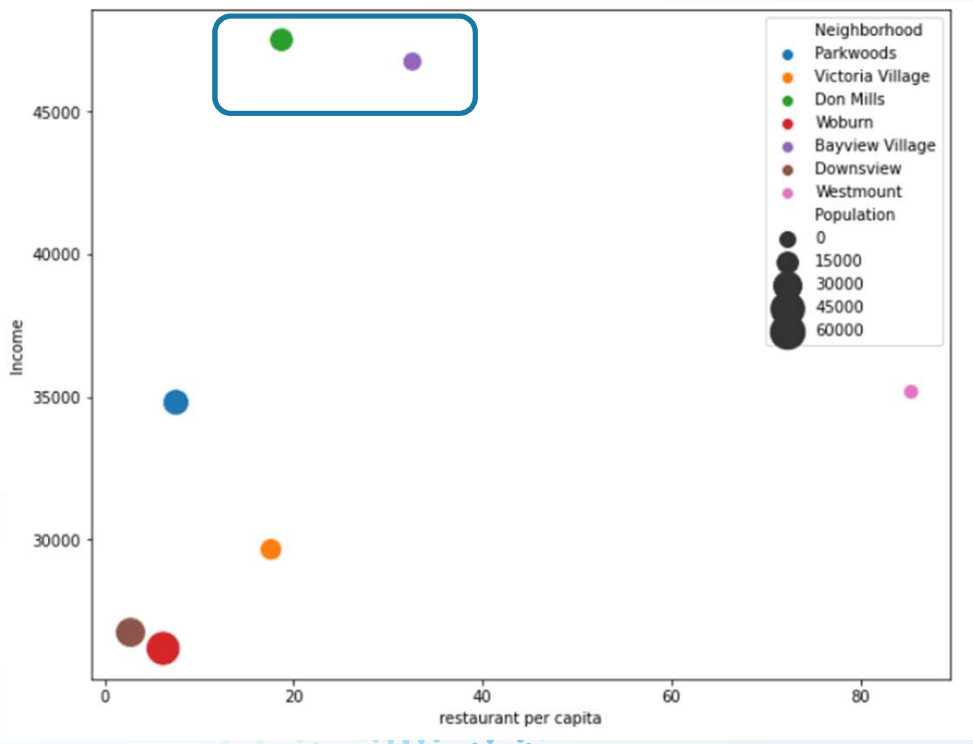
Data Exploration - Outliers

- Filtered data from both cities were investigated for outliers
- Restaurant per capita column was added to data frame of both cities
- An outlier found in the Toronto data was eliminated from the data frame



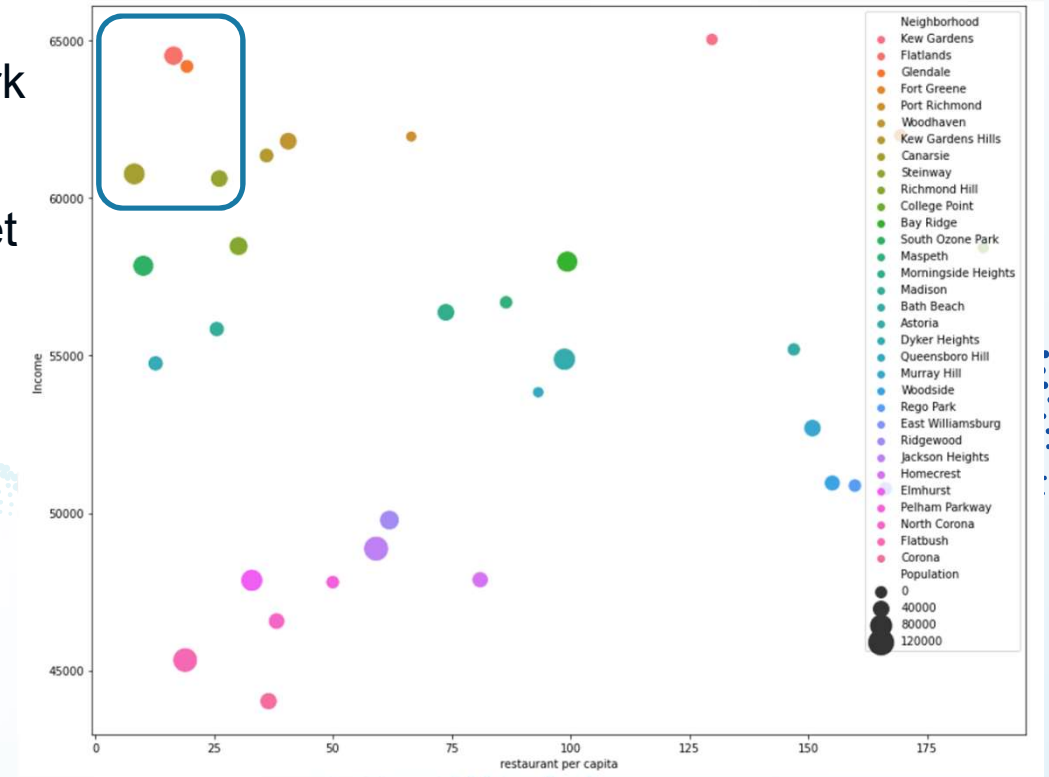
Exploratory Analysis – Toronto

- A scatter-bubble plot of Toronto neighborhoods
- Neighborhoods of Don Mills and Bayview Village met most of the criteria for an ideal location in Toronto



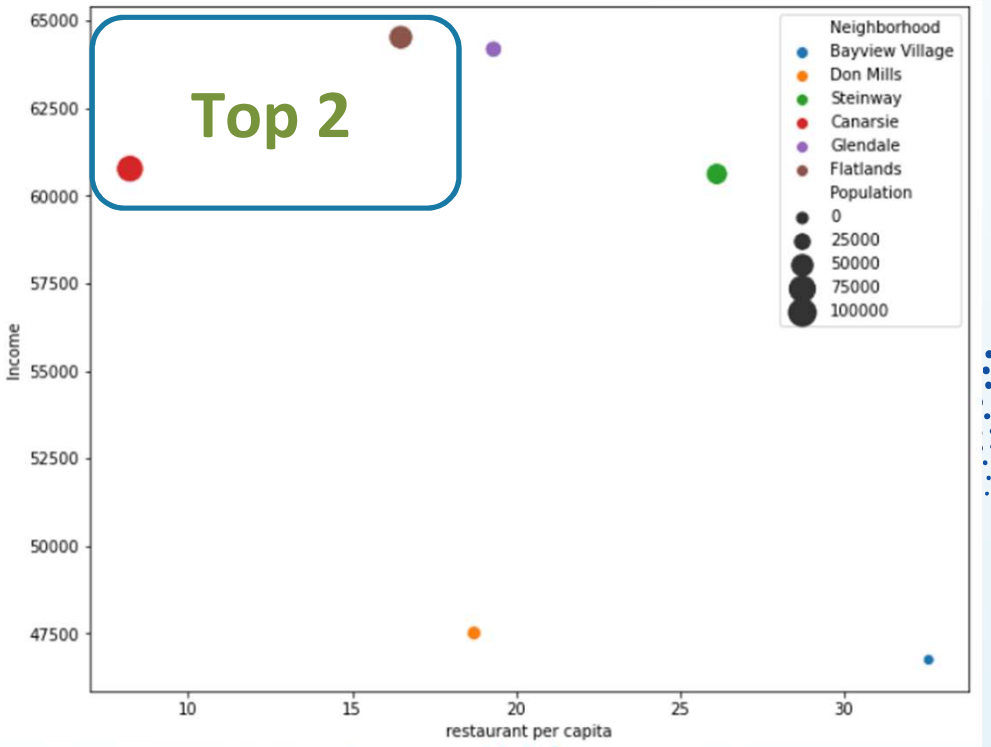
Exploratory Analysis – New York City

- A scatter-bubble plot of New York City neighborhoods
- The following neighborhoods met most of the criteria for an ideal location in New York City:
  - Canarsie
  - Flatlands
  - Glendale
  - Steinway



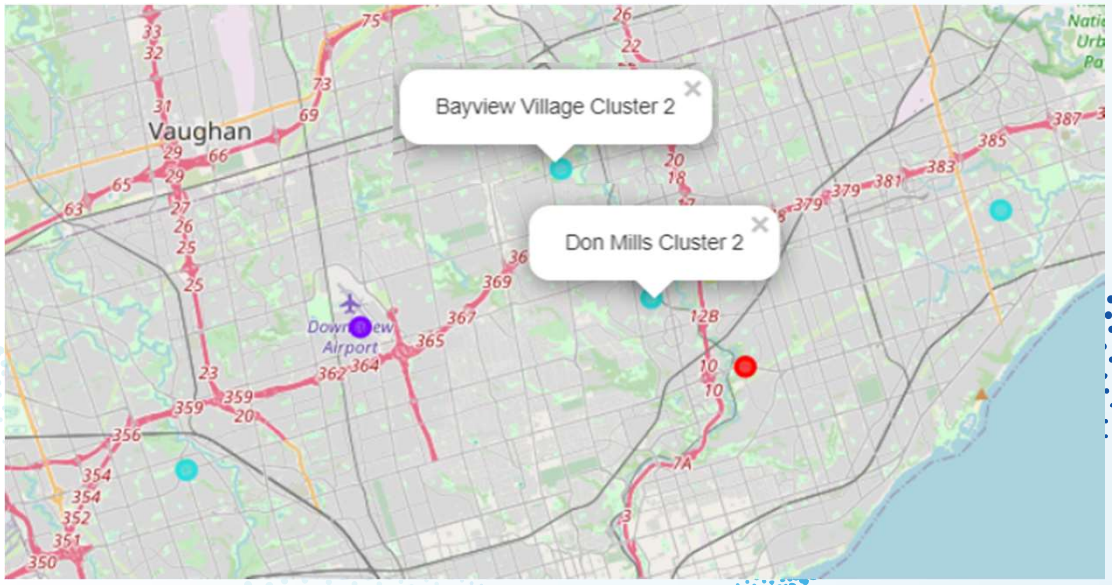
Results – Comparative Analysis

- Compared four top neighborhoods in New York City and two top neighborhoods in Toronto
- Neighborhood of Canarsie and Flatlands met the criteria of an ideal location to locate a new Chinese restaurant
- What about potential competition from other restaurants?



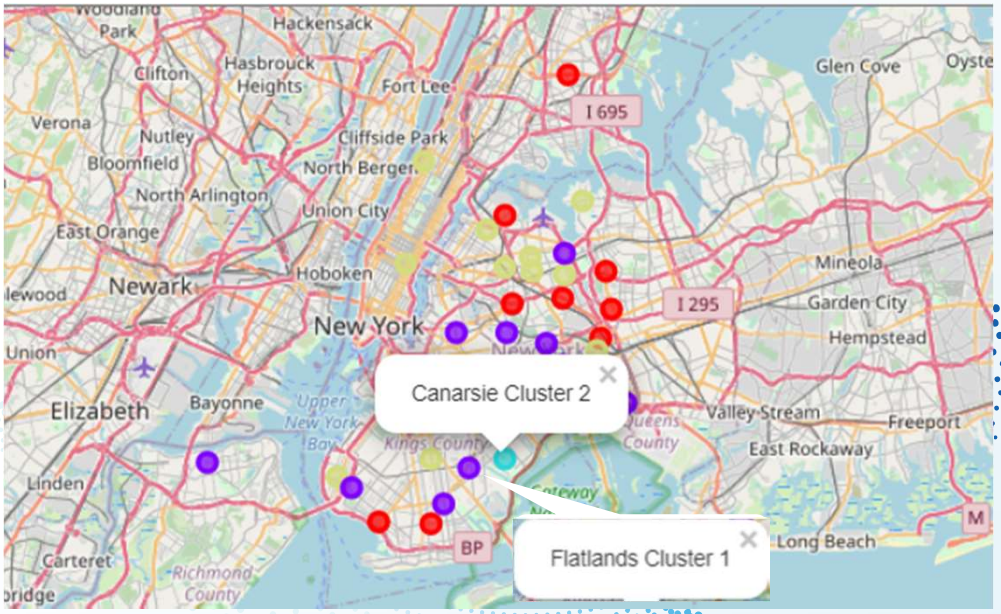
Results – Clusters Mapping (TO)

- K-means was used to cluster Toronto neighborhoods into four clusters
- Both top Toronto neighborhoods fall into the same clusters



Results – Clusters Mapping (NYC)

- K-means was used to cluster New York City neighborhoods into four clusters
- Canarsie, the most ideal neighborhood of all, falls into its own lone cluster
- Flatlands falls into a different cluster, but geographically close to Canarsie



Results – Restaurants Saturation Rank

Neighborhood	City	Restaurant per Capita	Population	Income	Chinese Restaurant Rank
Canarsie	**New York City**	**8.2**	**85,058**	**$60,766**	**5th**
Flatlands	New York City	16.5	66,726	$64,519	6th
Don Mills	Toronto	18.7	21,372	$47,515	8th
Bayview Village	Toronto	32.6	12,280	$46,752	2nd
New York City Average			49,269	$56,866	
Toronto Average			19,641	$56,311	

- Canarsie, the top neighborhood, has Chinese restaurants as the 5th most common restaurant in the neighborhood
- Canarsie not oversaturated by Chinese restaurants

Conclusion

- Built a solid model that predicted an ideal location for a restaurant in New York City or Toronto
- Model is general enough to be used for locating restaurants in other cities, given the neighborhoods location, population and income data
- Model could be improved with data on real estate prices and demographics data on other neighborhoods excluded from the modeling exercise