Recommendations for people moving between two cities

Clustering neighborhoods of New York and Toronto

Having handy list with similar neighborhoods may assist those moving between two cities

- People invest a lot of time and effort to decide which neighborhood to chose when moving from city to city
- Generally people would differentiate neighborhoods by housing prices, rent rates, crime rate, public schools availability and their rate, etc.
- Generating such a list will reduce the hassle around the process of researching neighborhoods and deciding where to move.

Data acquisition and cleaning

- New York census data is obtained from http://www.city-data.com/nbmaps/neigh-New-York-New-York.html%23N16
- Toronto census data obtained from https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/
- Venues data is obtained from Foursquare.com
- Data cleaning involved formatting obtained data, removing unnecessary rows and normalizing it to reduce noise from absolute values.

Similarity of other parameters

Toronto

	Venues	Number of venues
0	Coffee Shop	518
1	Café	269
2	Park	227
3	Pizza Place	227
4	Bakery	185
5	Italian Restaurant	184
6	Sandwich Place	167
7	Grocery Store	149
8	Fast Food Restaurant	136
9	Restaurant	135

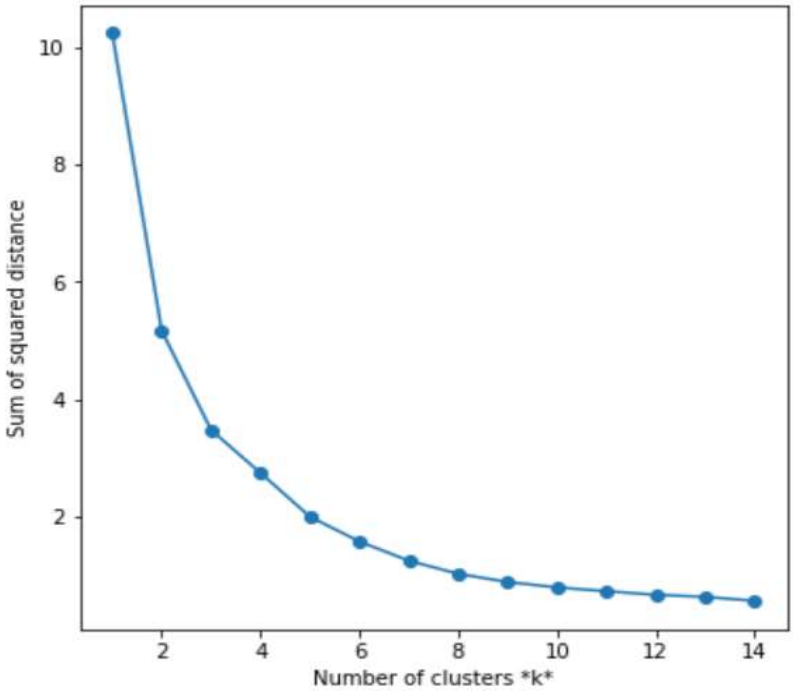
New York

	Venues	Number of venues
0	Coffee Shop	169
1	Park	163
2	Italian Restaurant	158
3	American Restaurant	128
4	Pizza Place	128
5	Theater	127
6	Bakery	112
7	Gym	108
8	Hotel	96
9	Café	90

Similarity of other parameters

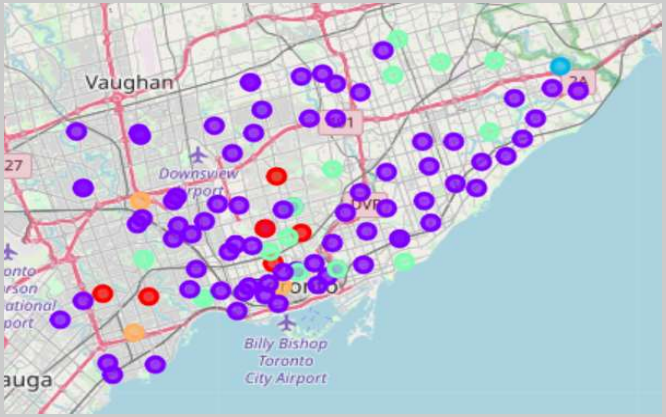
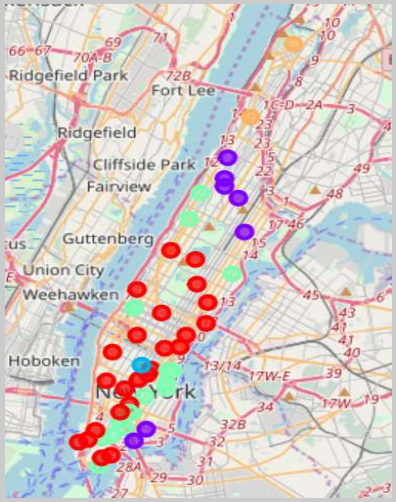
Toronto	New York
Median normalized population	
0.6% of the total population per neighborhood	0.7% of the total population per neighborhood
Median normalized household income in a neighborhood	
18.7% of the most wealthy neighborhood	52.5% of the most wealthy neighborhood
Median normalized area of a neighborhood	
8% of the largest neighborhood	3.2% of the largest neighborhood

K-means clustering



After k = 5 the slope of the graph notably changes, thus, for the purpose of clustering the neighborhoods of New York and Toronto I picked k = 5.

Mapped clustered neighborhoods

Toronto	New York
	

Conclusion and future direction

- Built a model to cluster neighborhoods in different cities
- Model is highly sensitive to the parameters picked for clustering
- Capture more information which is crucial for any person in deciding which neighborhood to move in
- Parameters that might be useful:
  - Housing prices;
  - Rent rates;
  - Crime rates;
  - Public school info.