Battle of Neighborhoods

Long Island City (Queens), NY and Harbourfront East (Toronto), ON comparison

Long Island City, NY

Harbourfront East

Description of the data:

The data used for this project will be acquired from Wikipedia, publicly available geospatial data, as well as from FourSquare via their API. The sources are listed below:

Toronto Data: <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

New York Data: <https://ibm.box.com/shared/static/fbpwbovar7lf8p5sgddm06cgipa2rxpe.json>

Foursquare: [https://api.foursquare.com/v2/venues/explore?&client\_id={}&client\_secret={}&v={}&ll={},{}&radius={}&limit={}](https://api.foursquare.com/v2/venues/explore?&client_id=%7b%7d&client_secret=%7b%7d&v=%7b%7d&ll=%7b%7d,%7b%7d&radius=%7b%7d&limit=%7b%7d)

These data will be ingested, transformed, and assessed within IBM Watson Studio using a Python v3.5 notebook. The planned libraries for use are the following:

Numpy-Library for handing vectored data

Pandas- Library for data analysis

JSON- Library for managing JSON file types

Geopy- Library for location retrieval

Requests- Library for handing http queries

Matplotlib- Library for Python plotting

Sklearn- Library for Python machine learning

Folium- Library for map rendering

First the New York data will be assessed to find the specific location parameters for Long Island City. Then the Foursquare API will be leveraged to bring in information pertaining to venues within Long Island City. At the end of the assessment we’ll see the top ranked venue types within Long Island City utilizing k-mean clustering. Once complete, the same process will be applied to the Toronto data. In this case though we will be focusing our assessment on Harbourfront East.

The end result will be a list of most common venue types within both locations for us to compare and contrast.