Introduction / Business Problem

Vegetarian and vegan restaurants are becoming more and more popular, especially in high population cities like New York City, Los Angeles, and Chicago. Because there are fewer vegetarians in smaller cities, opening vegetarian/vegan restaurants in smaller cities can still be a risky proposition.

We can decrease this risk by identifying neighborhoods that are most similar to neighborhoods in big cities with high numbers of vegetarian/vegan restaurants.

Data

For this project, I plan to train a machine learning model to identify neighborhoods that are likely to support a new vegetarian / vegan restaurant.

To train this model, I will use neighborhood data from NYC, LA, and Chicago, which all have many vegetarian vegan restaurants. This neighborhood data will consist of:

- Venue data from Foursquare
- Income and demographic data from:
    - NYC: http://app.coredata.nyc
    - Chicago: https://guides.lib.uchicago.edu/ChicagoStudies/Data
    - LA: http://la.myneighborhooddata.org/data/

Then, I will use data from three smaller cities to identify new neighborhoods where a new vegetarian / vegan restaurant is likely to do well. These three smaller cities I will use are:

- Buffalo, NY
- Cleveland, OH
- Omaha, NE

The neighborhood data I will use can be found at:

- Venue data from Foursquare
- Income and demographic data from
    - Buffalo: https://data.buffalony.gov/Economic-Neighborhood-Development/ACS-2017-Data-Profile-5-Year-Estimates-Erie-County/nc2w-3vy2
    - Cleveland: https://www.communitysolutions.com/resources/community-fact-sheets/cleveland-neighborhoods-and-wards/
    - Omaha: http://www.city-data.com/nbmaps/neigh-Omaha-Nebraska.html