The Battle of Neighborhoods

Week 1 – Part 2:

To recommend the user about the location for opening a restaurant, we need data about the entire locality.

We need

* The Neighborhoods available in the city considered, say Pittsburgh here
* Geographical coordinates for each neighborhood and each venue
* Number of the venues available in each neighborhood
* Number of Restaurants available in each neighborhood
* Type of restaurants available in each neighborhood

Data Collection

Collecting Neighborhood data:

I eventually ended up in collecting list of Neighborhoods in Pittsburgh through Wikipedia. I manually created a csv file with the list of neighborhoods in it.

Later I collected latitudes and longitudes of the Neighborhoods in Pittsburgh programmatically.

Once the required data is available, I programmatically collected different venues and categorized them. Later I used K Means to create clusters and from there, I derived observations for the best place to open a restaurant.