BATTLE OF NEIGHBORHOODS

Salvador Ollero

Introduction

In previous projects we have seen and analyzed the Neighborhoods in cities such as Toronto and New York. This time, and as I live in Chicago, the main goal is to determine which neighborhoods have which type of food in Chicago. In addition, we will group the by clusters to see which of them are similar.

Imagine you want to run a new brand restaurant and you are trying to figure it out which is the best place to do it. In order to do it in the best way you will try to collect some information about the neighborhoods such as number of restaurants, type of food and their location. Thanks to this project you will have this information and in addition it will be clustered, then, once you find a suitable neighborhood for your restaurant you will be able to find others with similar conditions inside the cluster.

But this project is not only good for restaurants, it is also for costumers. In fact they will look for one kind of food to eat, and they will be able to know which neighborhood would be the best to find it.

Data

The data used for this project is in the next link: https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago, where a table like this is provided:

	Neighborhood	Community area
0	Albany Park	Albany Park
1	Altgeld Gardens	Riverdale
2	Andersonville	Edgewater
3	Archer Heights	Archer Heights
4	Armour Square	Armour Square

However, to make it easier we will change the name of the Community area column by Borough, as we did in previous projects. In addition, the table will be sorted by boroughs in order to know if there are many neighborhood in just one borough:

	Neighborhood	Borough
0	Albany Park	Albany Park
1	Mayfair	Albany Park
2	North Mayfair	Albany Park
3	Ravenswood Manor	Albany Park
4	Archer Heights	Archer Heights

As we can see, each bough may have many neighborhoods so we will group them together:

	Neighborhood	Borough
0	Albany Park, Mayfair, North Mayfair, Ravenswoo...	Albany Park
1	Archer Heights	Archer Heights
2	Armour Square, Chinatown, Wentworth Gardens	Armour Square
3	Wrightwood, Scottsdale, Crestline, Beverly Vie...	Ashburn
4	Auburn Gresham, Gresham	Auburn Gresham

Now we have a dataframe with 78 rows and two columns, but before usind foursquare we need the Latitudes and Longitudes. We use geolocator, and we get the following data:

	Neighborhood	Borough	Latitude	Longitude
0	Albany Park, Mayfair, North Mayfair, Ravenswoo...	Albany Park	41.9719	-87.7162
1	Archer Heights	Archer Heights	41.8114	-87.7262
2	Armour Square, Chinatown, Wentworth Gardens	Armour Square	41.84	-87.6331
3	Wrightwood, Scottsdale, Crestline, Beverly Vie...	Ashburn	41.7475	-87.7112
4	Auburn Gresham, Gresham	Auburn Gresham	41.7505	-87.6643

Now we can use Foursquare to get venues. In this project we will only focus in restaurants, the we use query = ‘restaurant’ and we process the data to get the following table:

	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Albany Park, Mayfair, North Mayfair, Ravenswoo...	Albany Park	41.9719	-87.7162	4	Mexican Restaurant	Sandwich Place	Korean Restaurant	Latin American Restaurant	Pizza Place	Chinese Restaurant	Fast Food Restaurant	Diner
1	Archer Heights	Archer Heights	41.8114	-87.7262	4	Mexican Restaurant	Wings Joint	Seafood Restaurant	Hot Dog Joint	Italian Restaurant	Pizza Place	Chinese Restaurant	Sandwich Place
2	Armour Square, Chinatown, Wentworth Gardens	Armour Square	41.84	-87.6331	0	Chinese Restaurant	Asian Restaurant	Bakery	Italian Restaurant	American Restaurant	Hot Dog Joint	Indian Restaurant	Breakfast Spot
3	Wrightwood, Scottsdale, Crestline, Beverly Vie...	Ashburn	41.7475	-87.7112	3	Food	BBQ Joint	Pizza Place	Italian Restaurant	Wings Joint	Eastern European Restaurant	Czech Restaurant	Deli / Bodega
4	Auburn Gresham, Gresham	Auburn Gresham	41.7505	-87.6643	2	Fast Food Restaurant	American Restaurant	Greek Restaurant	BBQ Joint	Dim Sum Restaurant	Wings Joint	Falafel Restaurant	Deli / Bodega

This table is the one that we will use to cluster and to map the neighborhoods. It provides position of the borough, its neighborhoods, its cluster, and the most common restaurants in them. This table have enough information to help people to take a decision however, clustering and mapping and its analysis would provide them with better tools.

All the pictures here belongs to the first five rows of the data frame.