**The Battle of the Neighborhoods   
Report By Eviatar Shemesh**

**Introduction & Business Problem**

The City of New York is the most populous city in the United States.

It is diverse and is the financial capital of USA.

It is multicultural.

It provides lot of business opportunities and business friendly environment.

It has attracted many different players into the market. It is a global hub of business and commerce.

The city is a major centre for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theatre, fashion, and the arts in the United States.

This also means that the market is highly competitive.

As it is highly developed city so cost of doing business is also one of the highest.

I was hired by a coffee shop named Devocion, a small company that makes the best and freshest coffee in New York.

We have 4 shops and we open a new one in Brooklyn, Cause the shop we opened there at Williamsburg was a hit, despite all the competition we made huge profits.

We want to collect data and get some few neighborhoods that coffee is popular at, but we believe that it'll be like the coffee shop we opened at Williamsburg.

**Target Audience**

The objective is to locate and recommend to the Devocion which neighborhoods of Brooklyn will be best choice to start a Coffee Shop.

The Management also expects to understand the rationale of the recommendations made.

This would interest anyone who wants to start a new coffee shop in Brooklyn,   
in neighborhoods where coffee is very popular, and the competition is intense

**Data**

#### Our data will be collected from 2 sources:

1.JSON File with New York Neighborhoods and borough, where we will extract only the relevant data, of Brooklyn.

The JSON file will be collected at: <https://cocl.us/new_york_dataset> .

We will organize it and extract only the Brooklyn data. This File contains 4 columns:

1)Borough

2)Neighbourhood

3)Latitude

4)Longitude

Only Brooklyn Data



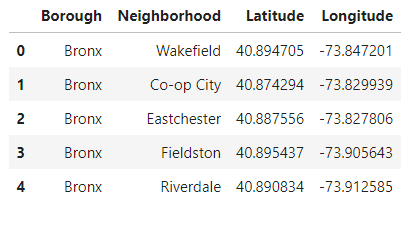
2.Foursquare API, to search for common venues around each neighbourhood, and cluster them into groups. Using it, we get the top 10 most common venues to each neighbourhood, which looks like that



**Methodology**

In this section we will talk about the data processing and methods to get the wanted result.

First, we collect the New York data using the JSON file.

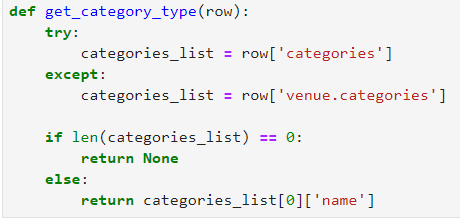


After that, we clean it by getting only the data where the Borough is Brooklyn, and drop the borough column cause it’s irrelevant, all the boroughs are Brooklyn.

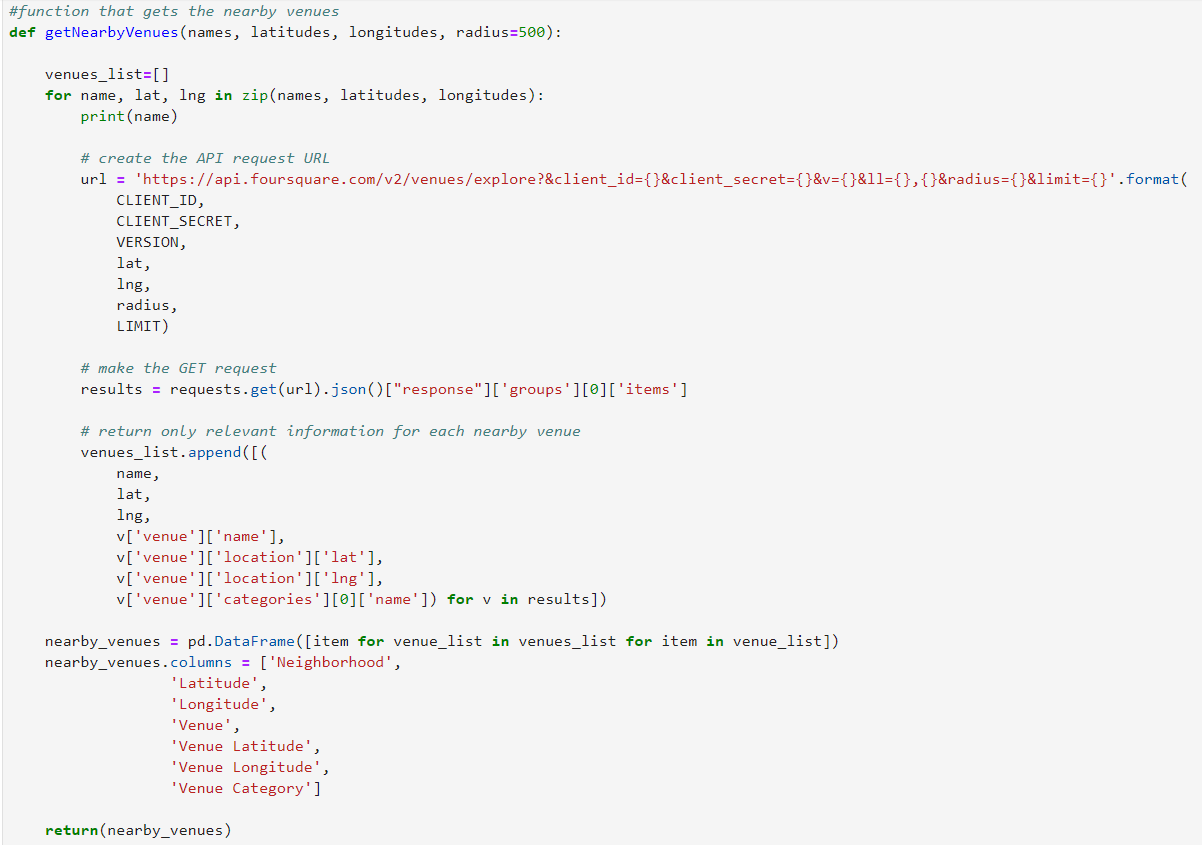


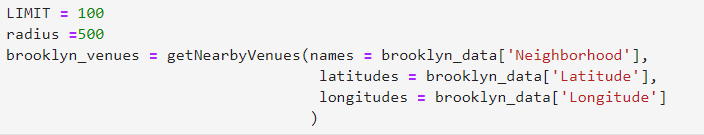
Then we move to our second resource, the Foursquare API, we right few functions.

The first one is to extract the category out of each venue



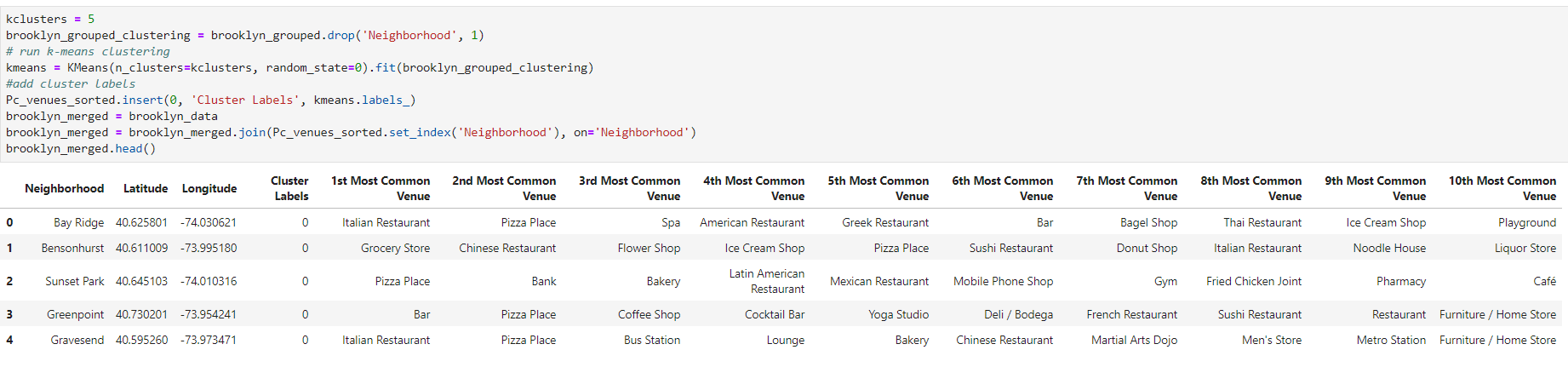
The second one is used to get nearby venues of each location



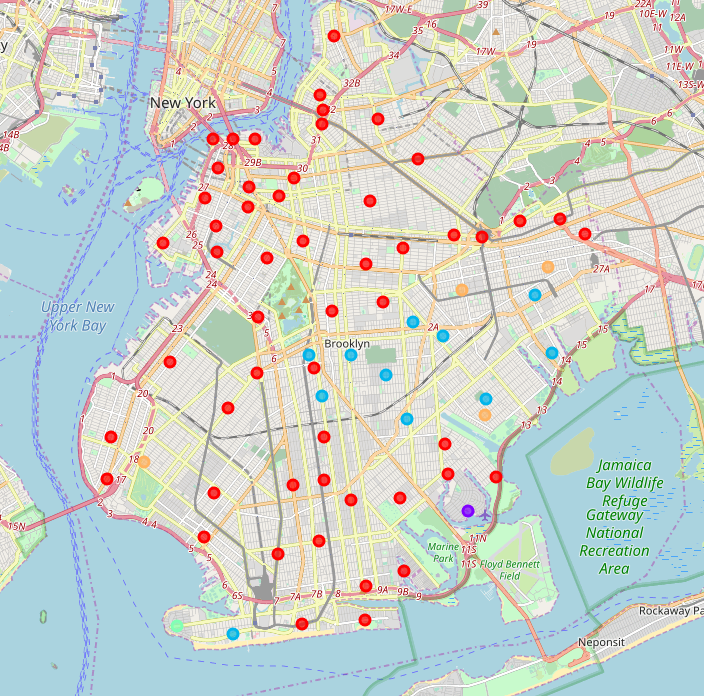
We apply the methods on our data of Brooklyn neighborhoods, max 100 per neighborhood, and maximum distance of 500 meters.  
The result of this run will be inserted into a new Data Frame

We use the one hot encoding method and inserting into a new Data Frame the top 10 most common category venues for each neighborhood

After we have This Data Frame, we use the KNN to cluster all of our Brooklyn neighborhoods into 5 groups, to find which neighborhoods are similar to our best shop neighborhood, Williamsburg

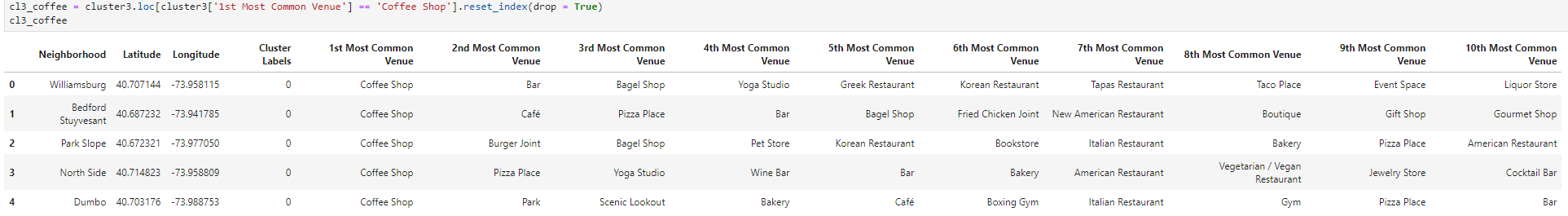


We put it on a map to show the cluster output



After that, we check which cluster group our Williamsburg neighborhood is, and inserting it into a new Data Frame.

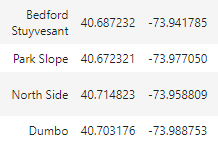


As I mentioned above, our wanted neighborhoods are ones similar and where coffee shops are very popular, so into a new Data Frame we insert only the neighborhoods that are at the same cluster as our Williamsburg neighborhood is, and the most common venue category is coffee shop

So, those 4 neighborhoods are potential neighborhoods to open new Devocion Coffee Shop.

**Result**

As I mentioned above, the result is list of 4 neighborhoods that are potential locations to open new Devocion Coffee Shop.



**Discussion**

Based on the results, I’m recommending our company to open a new coffee shop at Bedford Stuyvesant, and I’ll explain why.

Our Goals where to find similar neighbourhood to Williamsburg, where coffee is very popular.

As you can see in the results, in this neighborhood, the top 2 venues categories out there are connected to coffee, so this will be my recommendation.

**Conclusion**

To conclude, I’m very happy with the results.

They came after a lot of work, clean data and the most important thing, a lot of data.

The list of neighborhoods is very small(Only 4 neighborhoods), so no much research will be needed to select the new location for our Devocion Coffee Shop.