

## Introduction

A commercial real estate firm specializing in restaurant properties is interested in finding similar neighborhoods to up and coming "hipster" neighborhoods in Brooklyn, NY. The firm has had tremendous success in Brooklyn. They are considering Washington DC as a potential city to expand to. They are looking for concrete recommendations for neighborhoods to invest in that are similar to the neighborhoods of Brooklyn that they have made their money in.

## Data

We use Foursquare data for the neighborhood data. We build a dataframe of Washington DC neighborhoods and include the Brooklyn, NY neighborhoods that are were trying to find similar neighborhoods to. A list of Washington DC neighborhoods was found here:

[https://en.wikipedia.org/wiki/Neighborhoods\\_in\\_Washington,\\_D.C](https://en.wikipedia.org/wiki/Neighborhoods_in_Washington,_D.C). The culinary neighborhoods we will be looking to find similar Washington DC neighborhoods to are here:

<http://www.bkmag.com/2015/12/22/the-best-food-neighborhoods-in-brooklyn-ranked/> We found the latitude and longitude values for these neighborhoods, cleaned the data and created a dataframe to use in our comparisons. The data frame included the following fields:

Neighborhood, Latitude and Longitude.

## Methodology

We used the kmeans algorithm to compare neighborhoods in Washington DC with the neighborhoods that the client selected as their target neighborhoods. To begin, we first select a number of classes/groups to use and randomly initialize their respective center points. The machine learning algorithm works by classifying each data point by computing distance between the point and the group center. The point is then classified to be in the group whose center is nearest to the point. K-Means is fast and perfect for our problem.

We took the Brooklyn and Washington DC neighborhood data and created a comma separated value file. We read this in to a Pandas dataframe. We then authenticated to the Foursquare api. We looped through the neighborhoods and added the clustering column to our dataframe. We then explored each cluster in order to determine the results.

## Results

The first cluster contained a group of Washington DC neighborhoods and the Brooklyn neighborhoods. Remaining clusters did not contain any Brooklyn neighborhoods. The clustering resulted in a list of 74 Washington DC neighborhoods that are similar to the Brooklyn neighborhoods the client desires.

Adams Morgan  
Columbia Heights  
Howard University  
Le Droit Park  
Mount Pleasant  
Park View  
Pleasant Plains  
Shaw

U St NW  
Downtown  
Dupont Circle  
Federal Triangle Station  
Foggy Bottom  
Georgetown  
Logan Circle  
Mount Vernon Square  
Penn Quarter  
Shaw  
Southwest Federal Center  
U St NW  
West End  
American University Park  
Cathedral Heights  
Chevy Chase  
Cleveland Park  
Colony Hill  
Forest Hills  
Foxhall  
Glover Park  
Kent  
Massachusetts Avenue Heights  
McLean Gardens  
North Cleveland Park  
Observatory Cir NW  
Tenleytown  
Wakefield  
Wesley Heights  
Woodland Normanstone  
Woodley Park  
Brightwood Park  
Chevy Chase  
Crestwood  
Fort Totten  
Hawthorne  
Petworth  
Riggs Park  
Takoma  
Arboretum  
Bloomingdale  
Brentwood  
Brookland

Carver Langston  
Eckington  
Edgewood  
Fort Lincoln  
Fort Totten  
Ivy City  
Langdon  
North Michigan Park  
Riggs Park  
Stronghold  
Trinidad  
Truxton Circle  
Woodridge  
Barney Circle  
Capitol Hill  
Judiciary Square  
Kingman Park  
Mount Vernon Triangle  
Navy Yard  
Near Northeast  
NoMa  
Shaw  
Southwest Waterfront

### **Discussion**

There are a number of neighborhoods that are similar to the Brooklyn neighborhoods the client desires. At this point it would be beneficial to target specific aspects of each neighborhood that the client desires and use the Foursquare data or other data to drill down into each neighborhood. A detailed list of features could be compiled and a decision could be made based on cost, transportation and any other factors deemed important by the client. The client should be engaged and based on preferences the list can be paired down.

### **Conclusion**

Washington DC is definitely a good candidate for expansion. The client has 74 options for neighborhoods that are similar Brooklyn neighborhoods.