## The Battle of Neighborhoods (Week 2)

## Finding a Better Place in Washington DC to open Ethiopian cultural restaurant.

## Introduction

### Background

As per the study of Ethiopian Embassy in united state more than 200,000 Ethiopian immigrants are living in Washington DC. Compared to other US states majority of the Ethiopian immigrant are living in Washington DC.

This project aims to estimate the best localization to open Ethiopian Cultural restaurant in city of Washington DC.

Prior to starting any restaurant, it’s very crucial to know the convenient location where to launch the restaurant. In order to do so, this report will try to gather data about other countries restaurant localization in Washington city neighborhood.

### Problem

As the goal is to find the best place to open Ethiopian restaurant, we need to make sure that the place should be a locality for Ethiopian. We also need to check that customer could be interested in this specific business. In order to do so, an exploration in Washington neighborhood will be done in addition to data gathering. What king of restaurant works well? This survey will allow validating the data analysis done here.

**Data**

Based on definition of our problem, factors that will influence our decision are:

* Finding the best place where Ethiopian restaurant are mostly found.
* finding the most common venues
* choosing the right neighborhood within the borough

We will be using the geographical coordinates of Washington DC to plot neighborhoods in a borough that is popular by Ethiopian immigrant, and finally cluster our neighborhoods and present our findings.

Geodata of Washington can be found <https://query.data.world/s/i3n2uliq7hfcumldfarhh2rkcxlbe5>

#### Foursquare API Data:

We will need data about different venues in different neighborhoods of that specific cluster. In order to gain that information we will use “Foursquare” locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meter.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood

2. Neighborhood Latitude

3. Neighborhood Longitude

4. Venue

5. Name of the venue e.g. the name of a store or restaurant

6. Venue Latitude

7. Venue Longitude

8. Venue Category

**Methodology**

To compare the similarities of two clusters, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like Washington. To be able to do that, we need to cluster data which is a form of unsupervised machine learning K-means clustering algorithm.

Libraries Which are Used to Develope the Project:

Pandas: For creating and manipulating dataframes.

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.

Scikit Learn: For importing k-means clustering.

JSON: Library to handle JSON files.

Geocoder: To retrieve Location Data.

Matplotlib: Python Plotting Module.

Fig 1- The 10 most common venues of all neighborhoods

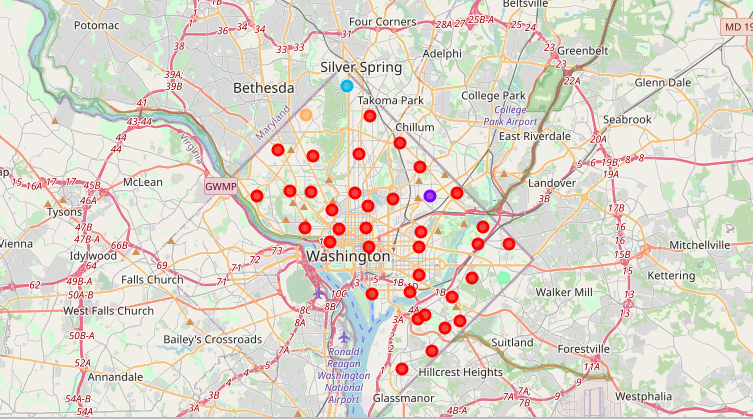


Fig 2- The first five venue of Adams Morgan



**Result and discussion**

Our analysis showed that there are a large number of establishments in Washington, A large number of hotel and restaurant. In total 817 venues are returned from each neighborhood and clustered into five categories.



**Conclusion**

The goal of this project was to identify the type of restaurant and its approximate location in a promising area of the city of Washington. Using open sources of data, we created datasets that helped identify patterns in the data. Data clustering made it possible to identify similar areas by the contingent of buyers. As a result, we identified 39 neighborhoods Adams Morgan and Down Town are selected for launching Ethiopian cultural restaurant. The selection is based on the criteria that the place where most Ethiopian immigrants are residing and the neighborhoods should have a venue for other countries hotels and restaurants.