Capstone Project - The Battle of Neighborhoods Discuss:

1- introduction:

description of the problem :

If some one want to open coffee shop in a large and rich of coffee culture city like costa cafe, Viet Nam, it will be competitive to start up coffee business. In this case my contractor is a humble Vietnamese man who has contacted me to give advises and draw up essential lines of business prediction and back-up plans (but this part we will just discuss about predicting hot spot).

Problem Which Tried to Solve:

The major purpose of this project, is to suggest a better neighborhood in a new city for the person who are shiffting there. Social presence in society in terms of like minded people. Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby.

One interesting idea would be to compare the neighborhoods of the two cities and determine how similar or dissimilar they are in a city of my choice

In a city of my choice, if someone is looking to open a restaurant, where would you recommend that they open it? Similarly, if a contractor is trying to start their own business, where would you recommend that they setup their office?

Foursquare API:

This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 50 and the radius parameter would be set to 1000..

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.

XML: To separate data from presentation and XML stores data in plain text format.

Geocoder: To retrieve Location Data.

Beautiful Soup and Requests: To scrap and library to handle http requests.

Matplotlib: Python Plotting Module.

First of all we need to collect Data of all coffee shops in Da Nang including their name, id, location (address, latitude, longitude) then pick up the "hot" neighbor where locates most of the venue. In order to asset Data we use FourSquare and apply folium for visualizing a particular neighbor in which that we will observe customer "traffic" and predict an appropriate location of new coffee shop in town. In this case you will find its temporary name on the folium map, "O Day Roi!"(Meaning "Here It Is!" in Vietnamese)

2. Conclusion

We will need a location where we can catch out customers from "hot" location we have picked up from the map and stay in a certain distance so as to lessen the competivity of business.