**IBM Data Science Capstone**

**Jordanian Neighborhood Clustering**

**And Café’ Opening Opportunity**

**Prepared By: Alaa Hamdan**

## I. Project Description:

### 1- Introduction:

Jordan is one of the Middle-East countries that is having a small Geographical area and around 10 Million population, it is famous of its excellent weather and of the historical places (such as Petra) Jordan Consist of 12 Governorates and around 89 Major Neighborhoods distributed across the country. those major Neighborhoods are having different characteristics and what is popular in one Neighbor is different than the others having a private small business is a little bit challenging due to the economic situation. the percentage of small businesses that ends by closing out is around 40%. that is why intensive studies and analysis should be done in prior one of the popular venues in Jordan is Cafe's. but it is only popular in limited Neighborhoods. opening Cafe is very beneficial if done in the right place

### 2- Business Question:

In This Project, the aim is to find the best 5 Neighborhoods to open a Cafe that is having high probability of being successful. the approach is to utilize Foursquare data to explore Jordanian Neighborhoods in terms of:

* What are the venues exist in each Neighborhood
* what are the similarities among Neighborhoods
* is there any kind of Neighborhoods that is having high probability of being successful if we open a Cafe
* is there a Neighborhood with low number of competitors but with high probability of being a successful location?

## II. Description of the data:

the main data used for this project will be from two sources:

* The geolocation data of Jordanian Major Neighborhoods (collected from the different websites)
* The venues in each neighborhood. (FourSquare API)
* Other supporting data:
  + Coordinates (Geocoder Python)
  + GeoJson ([http://data.beta.nyc](http://data.beta.nyc/))

### 1. Data collection process:

* the location of each Neighborhood will be loaded through an excel file.
* For Jordan Coordinates, call Geocoder Python to get its coordinate.
* For each neighborhood's coordinate, call FourSquare API to get the surrounding venues.
* Count the occurrences of each venue type and attach that information to each neighborhood.
* The output of the data collecting process will be a 2 dimensions dataframe:
* Each row represents a neighborhood.
* Each column will be the count of one type of venue in that neighborhood

### 2. Using data to solve the question:

* First, Correlation between Neighborhoods data and Venues.
* Second, explore each Neighborhood with its most existing Venues
* Third, Cluster Neighborhood and find Similar Neighborhoods
* Forth, Find the cluster with highest number of Cafe's
* Inside this cluster, as all Neighborhoods are having similar characteristics, find the Neighborhood with lowest number of Competitors to be a candidate for opening a cafe
* extract the top 5 candidate Neighborhoods

## III. Methodology:

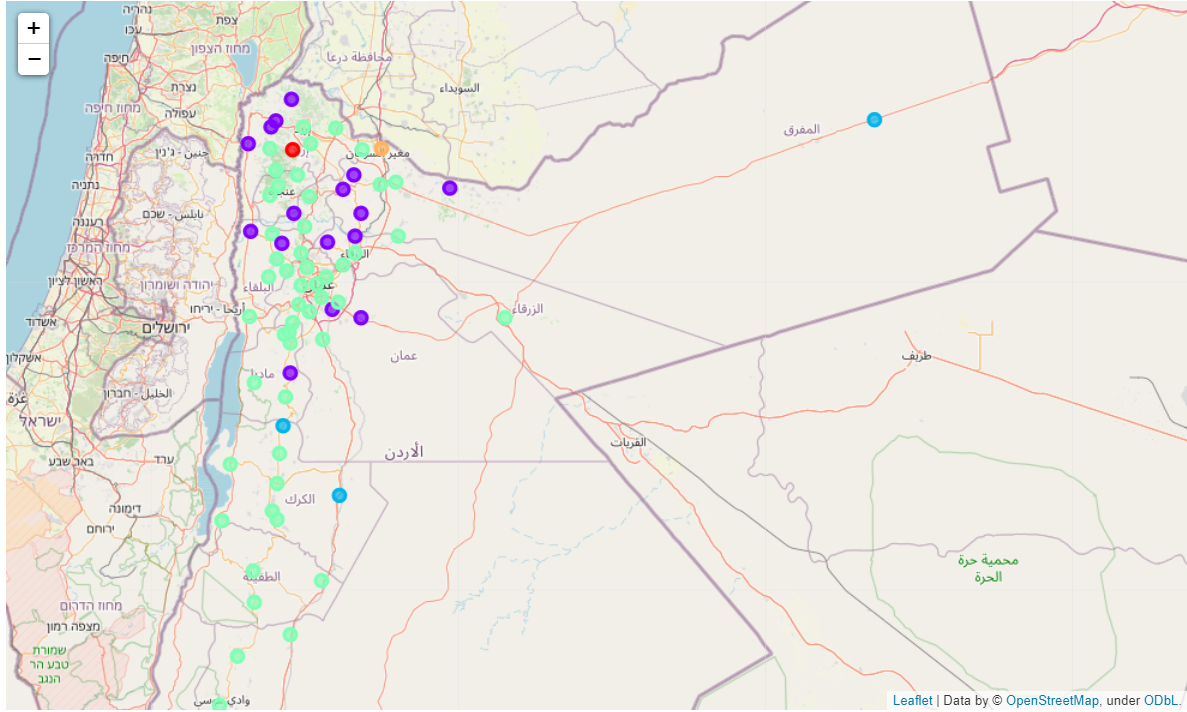
* Venues in each major Neighborhood will be collected from Foursquare API limited to 100 Venues and in a range of 2Km from the Neighborhood coordinates
* After Combining all the data, Clustering Methodology will be used to find the Similar Neighborhoods
* KMeans clustering will be choose to cluster Neighborhoods into 5 Clusters
* Visualization over Folium to be used for visualizing the results
* a Neighborhood popular with Cafe's is the target cluster
* the Neighborhood belong to this cluster with lowest Number of Cafe's is a target Neighborhood

## IV. Results:

* Number of Major Neighborhoods analysed was 89, Venues returned for 74 Neighborhoods (15 Neighborhoods were Droped
* 74 Neighborhoods were clustered in 5 clusters
* Cluster 3 was considered as the cluster with highest number of Cafe's
* inside Cluster 3: we have 86 Cafe's distributed over 27 Neighborhoods
* the following Neighborhoods are having High Competetion in terms of Cafe's so it is not benificial to open Cafe there:
  + Center\_Irbid
  + Marka
  + Center\_Amman
  + Jamaa
  + Wadi\_seer

#### Best Neighborhoods to open a cafe:

* the Followig Neighborhoods in Cluster 3 are having the lowest Number of Cafe's, they are the best candidates to open a cafe:
  + Neighborhood
    - Yarqa 1
    - Arda 1
    - Arjan 1
    - Center\_Aqaba 1
    - Petra 1



## IV. Conclusion:

**In order to have High Probability of success, we suggest the following Neighborhoods to open a Café**

* + - **Yarqa**
    - **Arda**
    - **Arjan**
    - **Center\_Aqaba**
    - **Petra**