# The battle of neighborhoods Project (IBM Data Science Capstone)¶

## By: BRADJI Mahdi

## 1) Project Description ( Business Problem )

Business : Opening a restaurant .
Location : Algiers , Capital of Algeria.
Problem :
somoene wants to start a restaurant business in Algiers but he is confused about the **Best Place** to get the max amount of clients , and the **Type of restaurant** ( turkich ..ext) to target the right clients. So the purpose of this project is to help **business-men** with the right insights based on algerian venues data. it aims to make analysis to exsisting restaurants to expand.

### 1-1) 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.

### 1-2) Workflow:

Using credentials of Foursquare API features of nearby places of the neighborhoods would be mined. places per neighborhood parameter would reasonably be set to 5000 and the radius parameter would be set to 1000.

### 1-3) Libraries :

Pandas: For creating and manipulating dataframes.
Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.
JSON: Library to handle JSON files.
Geocoder: To retrieve Location Data.

Beautiful Soup and Requests: To scrap and library to handle http requests.
Matplotlib: Python Plotting Module

# 2) Data section :

## 2-1) Data Link :

https://en.wikipedia.org/wiki/Algiers#Geography

## 2-2) The process :

1. extracting algiers Neighborhoods through scrapping a wikipedia page.
2. exploring nearby venues to each Neighborhood with a 10k radius using foursquare API and then extracting venues categories and we can get the right insights from them.

## 2-3) Foursquare API Data

We will need data about different venues in different neighborhoods of algiers. 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.

## 2-4) we will need :

1.venues location.
2.venues categories.
3.each category count.

## 3) Methodology :

- ✓ Setting up the right developing environment ( Installing needed packages , libraries) .
- ✓ Scrapping data from Wikipedia .
- ✓ Loading Algiers neighborhoods Data and creating a dataframe that contains the best shape of data to make necessary analysis .
- ✓ Retrieving data locations (longitudes, altitudes ) through geocoder.
- ✓ Connecting to foursquare API through personal credentials.
- ✓ Creating a function to retrieve nearby venues to each neighborhood with it’s category and specific location.
- ✓ Mapping locations to get the data more clear and efficient.
- ✓ Categorizing venues per neighborhood to get the diversity of each one .
- ✓ Refining venues and letting just restaurants data , to get more accurate analysis and target the problem directly.
- ✓ Categorizing restaurants per category to specify the right restaurant type.
- ✓ Visualizing results to make analysis clear and readable.

- ✓ Grouping neighborhoods and getting the frequency of each restaurant category in the specific neighborhood.

## 4) Results:

We've got some insights at the middle of analysis but they had no effect on our data-driven decision , so the main compenent of our choice will be the final result of analysis which is restaurants frequencies in each neighborhood , we can get a clear answer based on it. We've found that French and turkich restaurants are the most frequent and others with small frequencies, and in each neighborhood the same number of restaurant is found.

## 5) Discussion:

The results we've got are very helpful but also kind of limited, because we still miss the number of everyday clients, and satisfaction rate .. ext.

So our conclusion is based on inferences with a good accuracy.

## 6) Conclusion:

So as we've explored our dataset in several ways we came to a conclusion that in each neighborhood there is almost the same restaurants count,so it doesn't matter where the business owner will launch it's restaurant. As For Type of restaurant , according to the data we see that French restaurant are the most frequent, if the business owner has a Competitional mindset we recommend that he start a french restaurant, otherwise there is indian and turkich food domains that can be very beneficial.