Introduction/Business Problem

Business Problem:

In Bogota City, Colombia, someone is looking to open a restaurant and wants to know where is the best place to open it.

We will use data analytics to find a neighborhood with a similar offer.

Data

From foursquares the venues nearby each of the city neighborhoods.

create the API request URL

```
    url =
'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
        CLIENT_ID,
        CLIENT_SECRET,
        VERSION,
        lat,
        lng,
        radius,
        LIMIT)
```

From local open databases the coordinates for the neighborhoods.

https://bogota-laburbano.opendatasoft.com/explore/dataset/georeferencia-puntual-por-localidad/download?format=csv

Methodology

Geographic data for restaurants in the City's neighborhoods will be obtained and clustered by venue cathegories.

Then a one hot encoding process will be run in order to use a clustering algorithm, which is used because it is needed to characterize each region and determine if it is a good place to open the restaurant.

Results

Five clusters were found that can be named as:

“Fast Food”

“Sea Food Restaurant”

“Latin American”

“Italian”

“Local Restaurant”

Depending on the type of restaurant that is to be opened, it is recommended to place it on the neighborhood of the proper cluster.

Discussion

More data is needed to fine-tune the process. There were just a hundred venues found. And further analysis can be made using neighborhood segmentation.

Conclusion

Open data and web-scraping constitute a good resource for data analytics.

Using open source tools like follium and panda it is possible to create professional GIS analysis.