Report Coursera capstone

**By : Mounir Ourekouch**

**Software Engineer Student, ENSIAS ,Morocco**

1. Introduction/Business Problem :

A contractor want to start a new business in fast food in London. Unfortunately, he has no idea about the right area for this project.

Therefore, he decided to rely on the science of data analysis in order to find the appropriate area for this new project,

Especially the population density in various neighborhoods of London, as well as the distribution of different venues and facilities in the city of London.



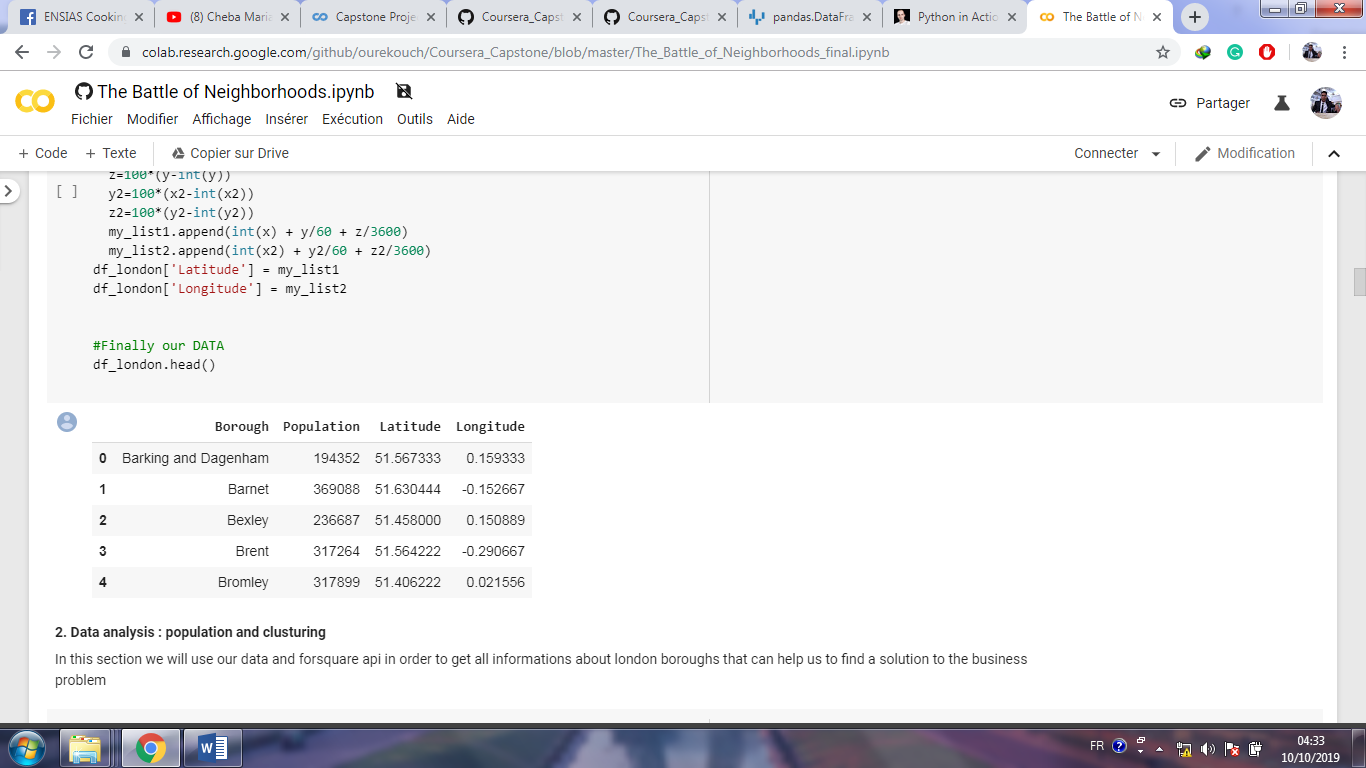
1. Data Section :

In order to solve our business problem, we will use two datasets:

* **List\_of\_London\_boroughs :**

This dataset is constructed by scraping the Wikipedia page : "https://en.wikipedia.org/wiki/List\_of\_London\_boroughs", the final dataset will help us to get boroughs Coordinates and also the populationof each borough .

Figure 1 : London Boroughs



* **Foursquare location data :**

After getting information’s about boroughs coordinates, we will use this coordinates to discover venues in each borough and cluster all boroughs according the most frequent venues in each borough that will help us to make the final decision.

1. Methodology Section :

In the project we will mainly use two methodologies in order to have enough information has to make a decision:

* Data visualization :

We will visualize first our data in order to get information about population and the top boroughs according to population:

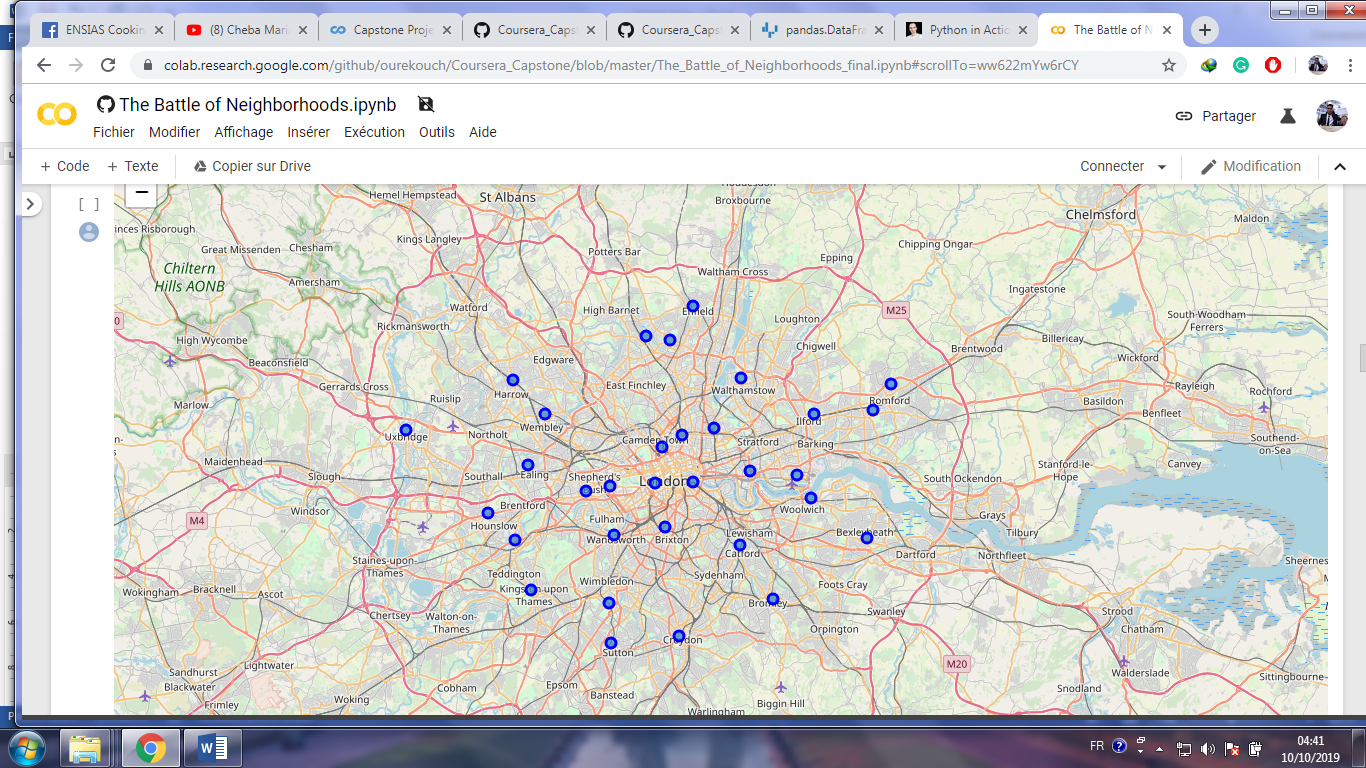


* Data clustering :

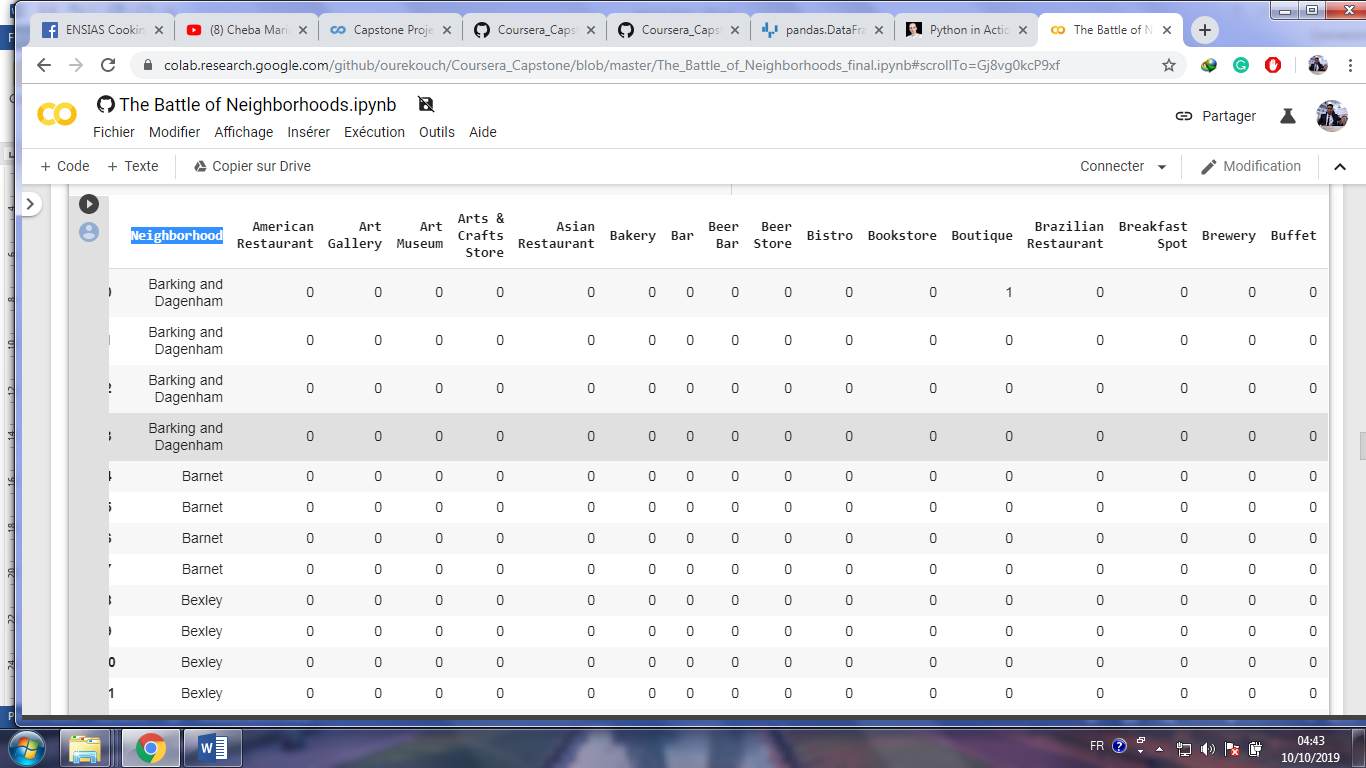
Because the first result is not enough the help us to make a decision and have a solution for our business problem , for that we need to use also foursquare api , folium in order to get more information about venues and make clusters for London boroughs

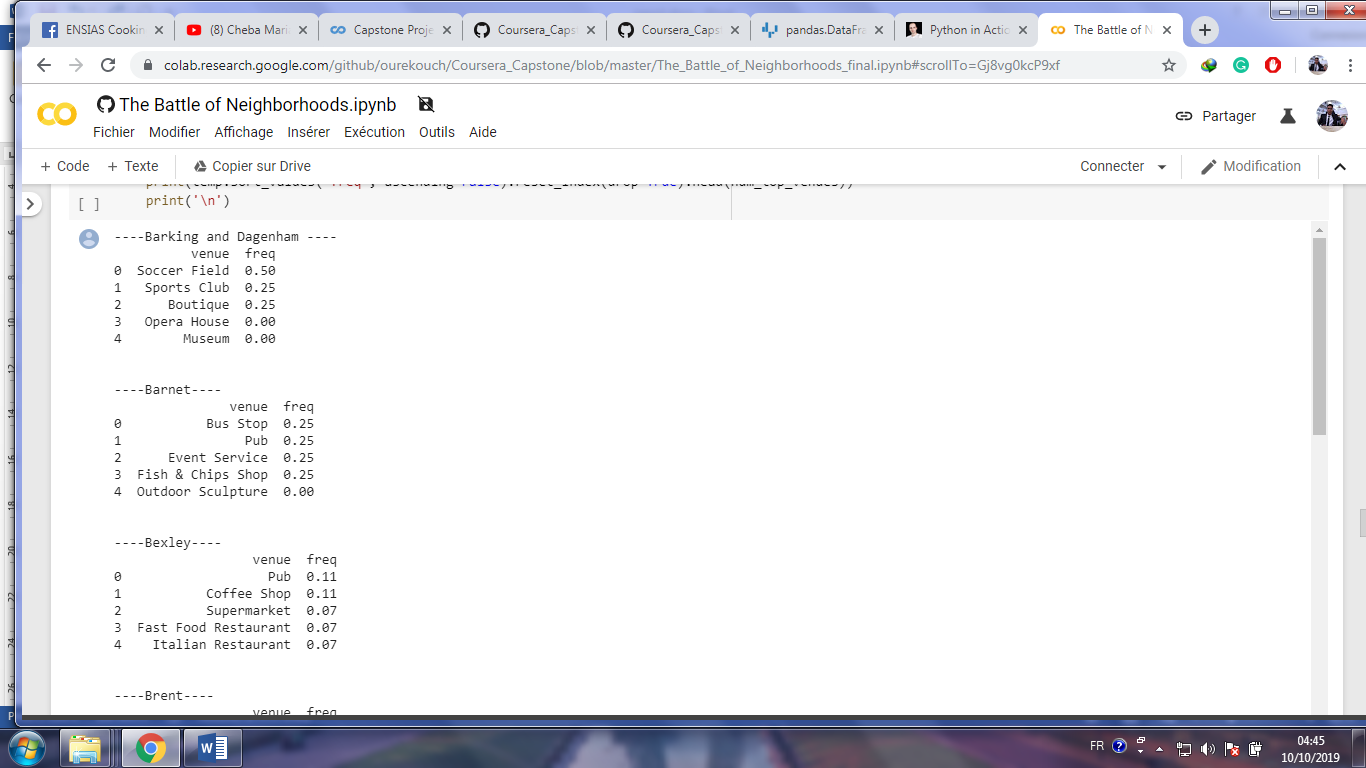
* Checking collected data about London boroughs in map :

Figure 2 : London

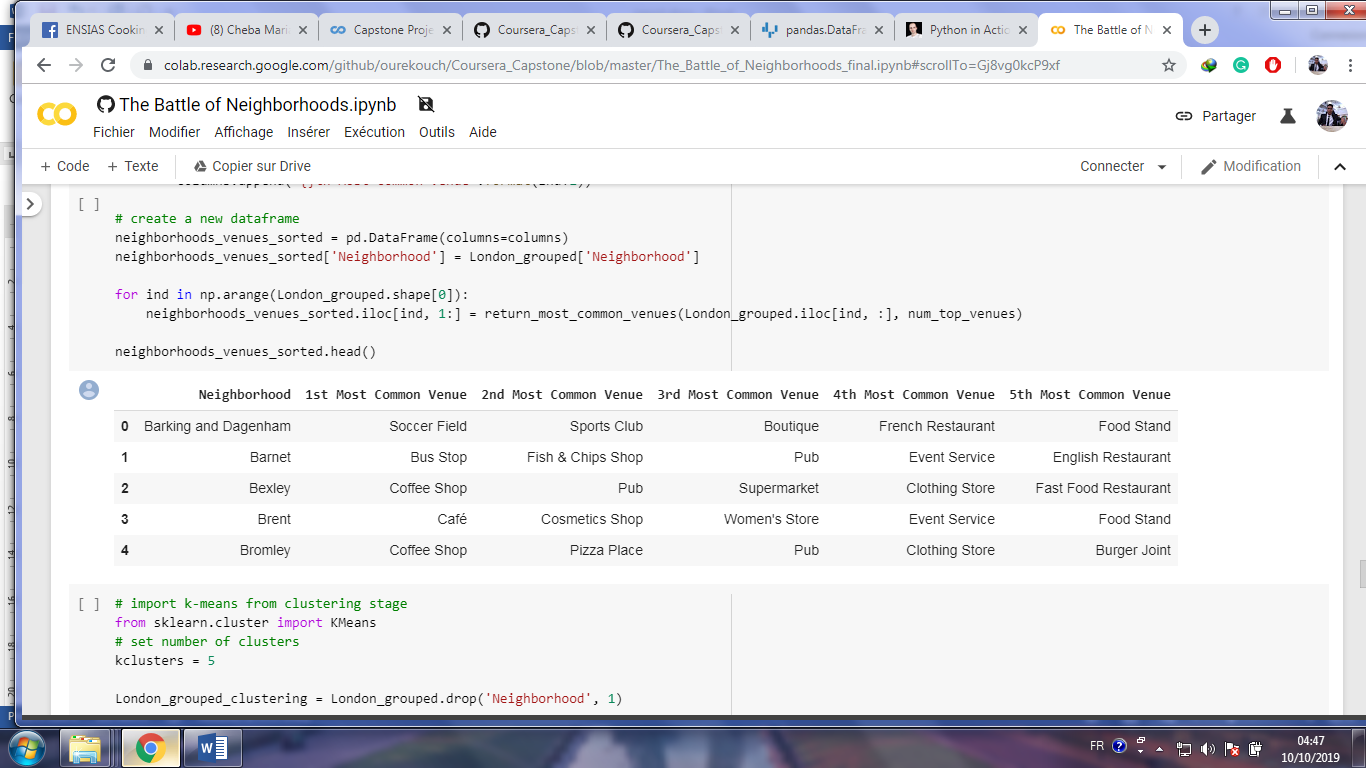


* Calculate the sum of venues for each borough or neighborhood and get the frequencies of venues in neighborhoods :

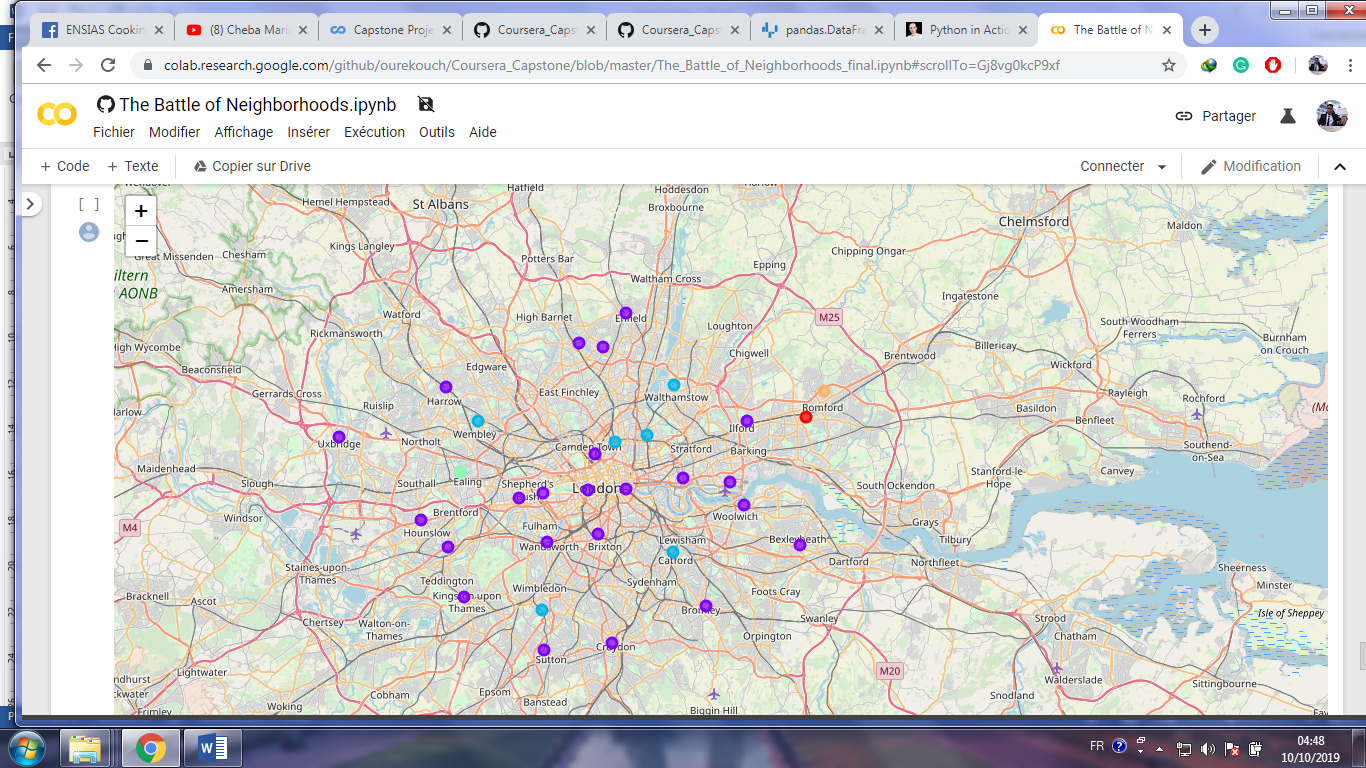




* Get the top venues for each neighborhood :



* **London maps with clusters :**



1. Results :

After the first visualization, we get that the first 5 boroughs according to population are:

* Croydon (372 K )
* Barnet (369 K )
* Ealing (342 K )
* Enfield (320 K )
* Newham (318 K )

After analyzing and clustering boroughs we get that :

----Croydon----

Venue freq:

* Coffee Shop 0.17
* Clothing Store 0.17
* Pub 0.10
* Bookstore 0.07
* Women's Store 0.03

----Barnet----

Venue freq:

* Bus Stop 0.25
* Pub 0.25
* Event Service 0.25
* Fish & Chips Shop 0.25
* Outdoor Sculpture 0.00

----Ealing----

Venue freq:

* Park 0.75
* Pharmacy 0.25
* American Restaurant 0.00
* Outdoor Sculpture 0.00
* Museum 0.00

----Enfield----

Venue freq

* Pub 0.5
* Indian Restaurant 0.1
* Grocery Store 0.1
* Coffee Shop 0.1
* Sandwich Place 0.1

----Newham----

Venue freq

* Light Rail Station 0.25
* Supermarket 0.25
* Gym / Fitness Center 0.12
* Bus Station 0.12
* Pub 0.12

So in order to make the final decision between those 5 boroughs who has the max of population we will add another score (Competitive rate) :

**Competitive rate :**

The sum of frequencies of venues in boroughs multicoated by +1 (if the venue will be benefit for our business for example: +1 for schools , stadium ,...) , by -1(if the venue make a danger for our business for example another restaurant )

|  |  |
| --- | --- |
| Borough | rate |
| Croydon | -0.17+0.17+0.10+.07+0.03=**0.2** |
| Barnet | 0.25+0.25+0.25+0.25=**1** |
| Ealing | 0.75+0.25=**1** |
| Enfield | 0.5-0.1+0.1-0.1-0.1=**0.3** |
| Newham | 0.25+0.25+0.12+0.12+0.12=**0.86** |

1. Discuss Section :

In the end, the best borough for starting a fast food restaurant in London with the big frequencies of population, and less competitive venues are:

1. **BARNET OR EALING**
2. **NEWHAM**
3. Conclusion :

Thanks to foursquare api and our collected data we had succeeded to get good results that helped us to make a decision and have a solution for our business problem.