Ömer BOZAN

IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

THE BATTLE OF BOROUGHS
(ITALIAN RESTAURANT IN BERLIN, GERMANY)

IF SOMEONE IS LOOKING TO OPEN AN ITALIAN RESTAURANT IN THE CITY OF BERLIN, WHERE WOULD YOU RECOMMEND THAT THEY OPEN IT?

BUSINESS PROBLEM

The Battle of Boroughs

TARGET AUDIENCE AND DATA REQUIREMENTS:

➢ Target Audience:This project is particularly usefel for restaurant chain owner or investors looking to open an Italin restaurant in the capital of Germany.

➢ To solve this problem, we need to these data:
  - ○ We will be using below dataset for analysing Berlin:
  - ○ -List of boroughs in Berlin
  - ○ -Latitude and longitude coordinates of the all boroughs
  - ○ -Venue data (powered by Foursquare) particularly related to Italian restaruants
  - ○ -Population of Berlin according to nationalities
  - ○ -Demographics of Berlin
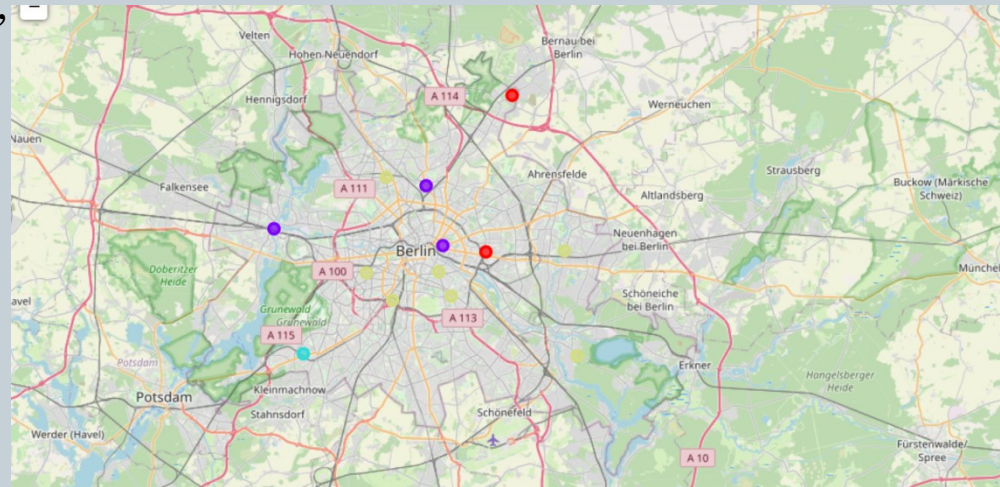
The Battle of Boroughs

Methodology:

- Berlin has a total of 12 boroughs.In order to segment the neighborhoods and explore them, it will be better to have a dataset.Unfortunately, we do not have such a dataset,that is why we will use the Wikipedia page of **Boroughs and neighborhoods of Berlin** and **Demographics of Berlin** using web scraping techniques, we will extract the required data , with the help of Python Requests and BeautifulSoup packages.

- The Python Geocoder package would be essential in finding the location data of these neighborhoods.Finally,the Foursquare API would be used to extract the data regarding existing in the city.

The Battle of Boroughs

RESULTS:

- The output from the clustering algorithms, it can be seen that the division of all the boroughs into 4 groups in terms of Italian restaurants present in that area. -Cluster 0 : Borough with a significant number of Italian restaurants. -Cluster 1 : Borough with low number of Italian restaurants. -Cluster 2 : Borough with heavy competition of Italian restaurants -Cluster 3 : Borough with very low (almost none number of Italian restaurants.

- Addition to Italian restaurants existing in specific area, we also examine the Italian immigrant where they tend to live. As can be seen in results, Italians are living more in cluster 3 and cluster 1.

- The results are visualized above: Here,

- Cluster 0=Red

- Cluster 1=Purple

- Cluster 2 =Light Blue

- Cluster 3=Yellow

The Battle of Boroughs

- **DISCUSSION AND SUGGESTIONS:**

- As observedd from clusters, opening a restaurant in Berlin could be trickyt business.
- Going to Restaurants in Category 0 and 3 can lead to less footfall,but also non-existence of competition,hence these clusters are not recommended.
- On the other hand,category 2 suffers from an intense competition and there are no Italian people around there.That is why this is not recommended.
- Finally, category 1 which corresponds to Pankow has more Italian immigrants and also has more footfall. The number of bar can be also a factor around Pankow because of the fact that Mediterranean people like for entertainment.
- In the light of these points, **Pankow** stands one step ahead for opening an Italian restuarant.