Coursera Capstone

IBM Applied Data Science Capstone

Opening a new coffee shop in Sumaré, Brazil.

- Business problem

  Choosing the location is the most important task before the money investment, this will determine if the investiment will give back a good return.

- Objective: Opening a new coffee shop in Sumaré, through analysing and selecting the best place.

Data

- Data required
  List of neighborhood in Sumaré, Brazil.
  Geo location like latitude and longitude of neighborhood.
  Venue data, related to Coffee Shop.

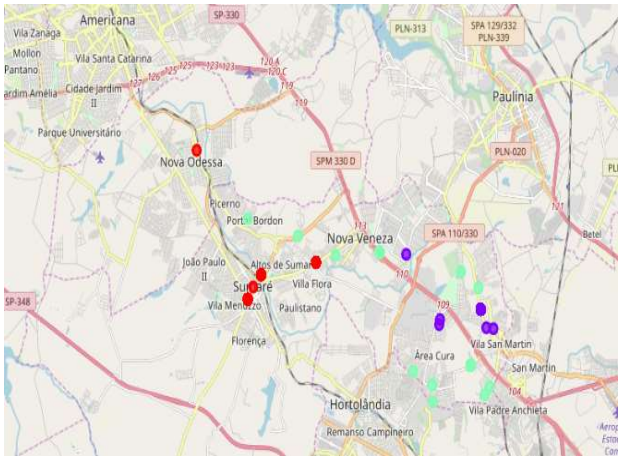
- Source of data
  Wikipedia neighborhood
  https://pt.wikipedia.org/wiki/Lista_de_bairros_de_Sumaré
  Foursquare API for venue data.
  Geocoder package for coornates.

Methodology

- Web scraping wikipedia
- Geocoder coordinates
- Venue data from Foursquare
- Get the mean of the frequency of occurrence of each venue
- Filter coffee shop by venue
- Execute k-mean clustering from the data
- Visualize the cluster in the map using folium

Results

- Categorize in 3 clusters:
  - Cluster 0: neighborhood with moderate number of coffee shop.
  - Cluster 1: neighborhood with low number to no existence of coffee shop.
  - Cluster 2: neighborhood with high concentration of coffee shop.



Discussion

- As noted results, most of the coffee shops are high concentrated is in green color of Sumaré city, but also, purple color have very low number to no coffee shops in the neighborhoods of Sumaré city and represents a good business opportunity to open a coffee shop.
  And coffee shops in green color there are too much competition and it’s not a good place to open.

Recommendations

- In this project, we only consider frequency of occurrence of coffee shop, there are other factors such as traffic of people and cars, etc.

Conclusion

- In this capstone we identified the business problem, specifying the data required, extracting and preparing the data, performing machine learning by clustering the data into 3 clusters based on their similarities, and provided recommendations to the investors about the best location to open a new coffee shop.