Coursera Capstone Project
IBM Applied Data Science Capstone

Opening a new coffee place in Seattle, Washington

Business Problem

- The objective is to analyze and select the best neighborhood to open a new coffee shop in Seattle.

Data

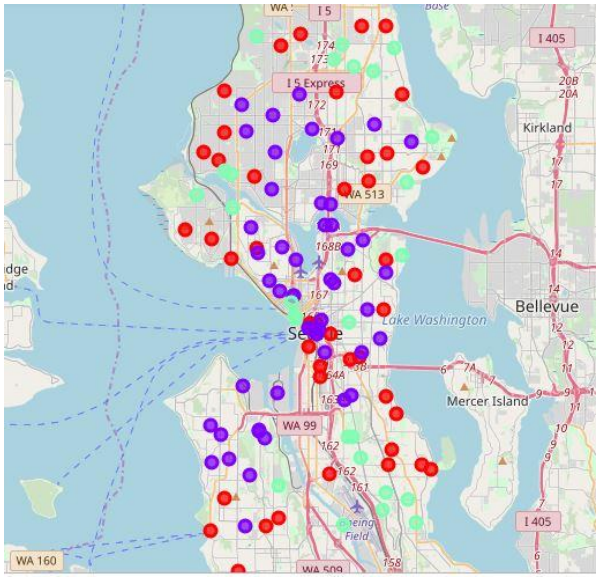
- To solve the problem, we'll need following data-
- List of neighborhoods in Seattle
- Latitude and Longitude of those neighborhoods. This will help us make a map and visualize the data
- Coffee places data in different neighborhoods. This will help us cluster the data

Methodology

- Get the neighborhood data from Wikipedia page- https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Seattle
- Using Foursquare API, we'll get the data for venues, it in a data frame and run k-means clustering.
- Create a map for better visualization and in the end, analyze the data.

Results

- We cluster the neighborhoods into three clusters based on the frequency of occurrence of coffee shops. The results can be visualized as shown in the map below- Cluster 0 is in red color, cluster 1 in purple color, cluster 2 in mint green color.



Discussion

- As observed in the results, there are more coffee places concentration in the neighborhoods mentioned in Cluster 1 and cluster 0, and the least in Cluster 2. Hence, the neighborhoods in Cluster 2 has a good potential of opening a new profitable coffee place, given less competition.

Conclusion

- In this project, we identified the business problem, find out the data that we'll required, extract the data, perform machine learning methodologies, form clusters and then analyze the results. The findings will help the stakeholders to identify the best neighborhood to open a new coffee place.