COUSERA CAPSTONE PROJECT

IBM DATA SCIENCE

PROFESSIONAL SPECIALIZATION

LOCATING AN AFRICAN RESTAURANT IN TORONTO

BY MOSES ADINFONO

INTRODUCTION

BACKGROUND

Have you ever relocated to a new country and wanted a taste of the meal back home? A number of people, after arriving in a new country, look for where they can get meals close to what they experience back home (the country they came from). This, among many other reasons, gave birth to the establishment of restaurants like: Chinese Restaurants, Italian restaurants and the likes.

Canada is one of the leading countries in policies that favor immigration. When immigrants arrive in Canada, the prime city of choice is usually Toronto.

I have a few friends in North York, Toronto. They tell me time and again that they wish they had a restaurant that would serve them some African dishes.

BUSINESS PROBLEM

This project aims to find the best neighborhood in North York Toronto to setup a New African Restaurant for immigrants from Africa and others who want to have a taste of delicious African delicacies.

DATA

For this project, we combined three datasets to achieve our goal.

The first source of data is Wikipedia. We will scrap the Wikipedia website for neighborhoods in Toronto using Beautiful Soup. The attributes we collected from this data sources were:

- Postal Code
- Borough
- Neighborhood

Here's is the link to the data source:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

The next dataset we employed was data of geographical coordinates gotten from previous exercises in the capstone program. It was a *Geospatial Coordinate CSV* file that we downloaded from IBM Cloud. Attributes contained and utilized from this dataset includes:

- Postal Code
- Longitude
- Latitude

Finally, we made use of the Foursquare Location data to solve to execute our idea. The dataset was obtained from querying the Foursquare API. Among the many attributes contained in this dataset, the following were utilized:

- Neighborhood
- Neighborhood Latitude
- Neighborhood Longitude
- Venue
- Venue Latitude
- Venue Longitude
- Venue Category

The combination of these data were the ingredients we use in cooking our location for an African Restaurant in North York, Toronto.

METHODOLOGY

The process began with collection and gathering of relevant data. Web-Scraping the Wikipedia website with Beautiful Soup was and reading of the CSV file into a DataFrame.

Subsequently, the gathered data was explored, cleaned and manipulated to optimize its performance in the analysis and reduce errors.

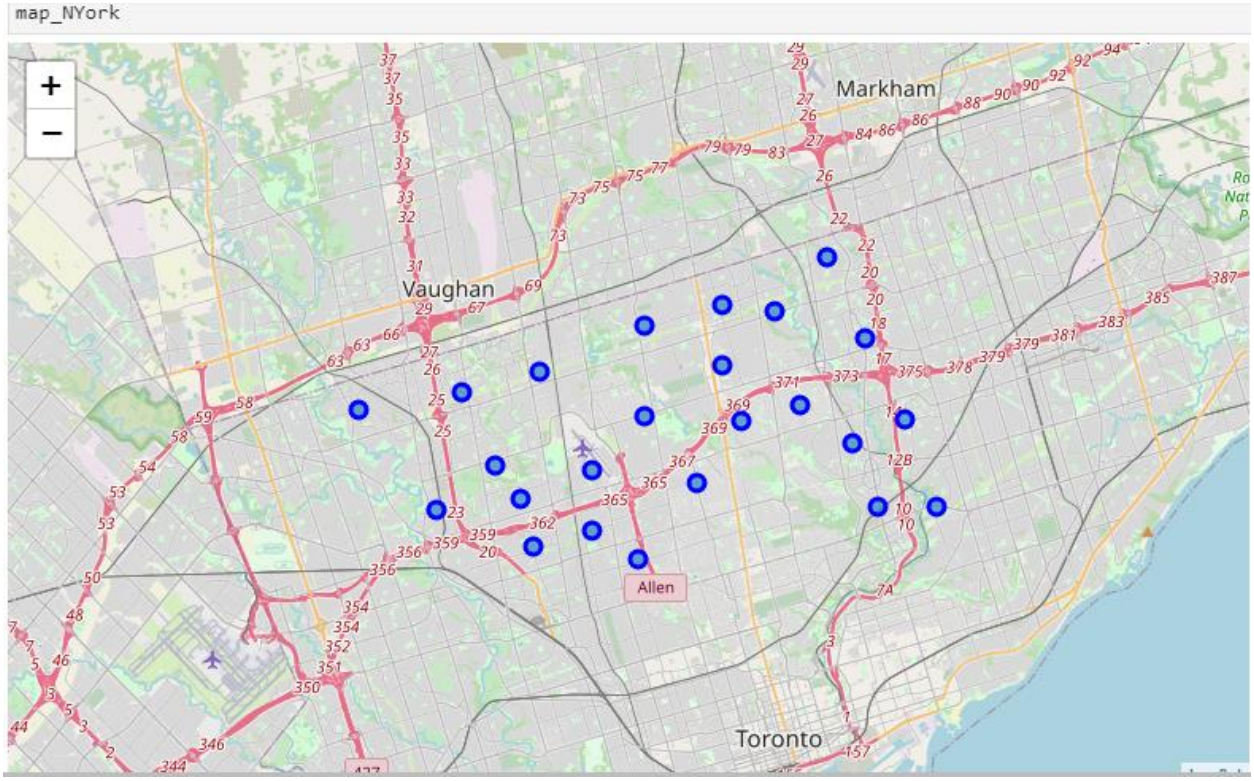
With our data looking good, we plugged it into the mix with our query of the Foursquare API to explore the venues of North York, Toronto; with emphasis on areas with restaurants or African restaurants.

K-means clustering was used to group similar locations and the Folium package was used to visualize the clusters.

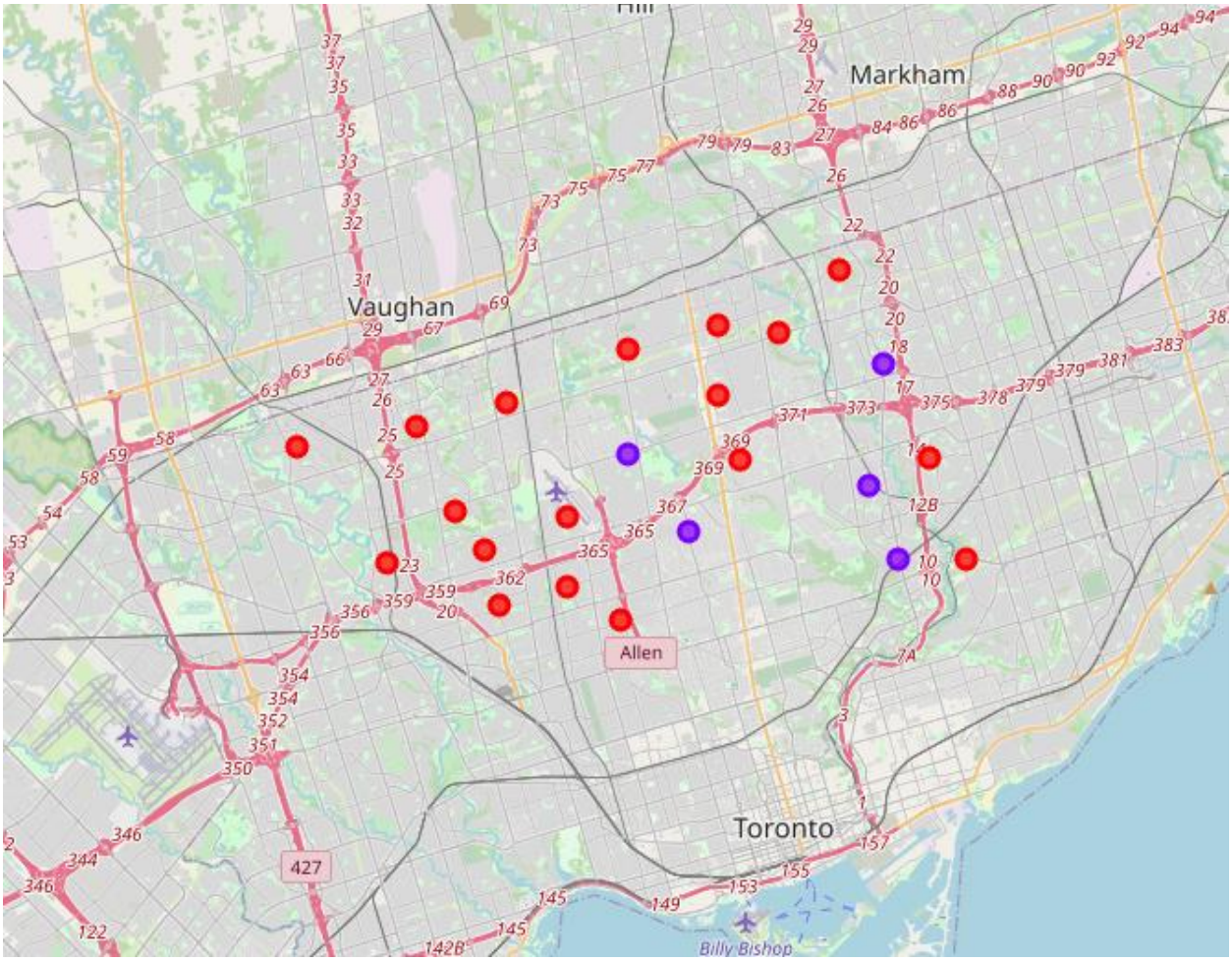
The optimal number of clusters, k, was gotten using the “Elbow Method”.

RESULTS

The image below is the visualization the locations in North York, Toronto before clustering.



The next image shows the result of the clustering.



DISCUSSION

We notice very easily that Cluster 1 happens to be more saturated, with 18 similar points, than Clusters 2 with just 5 similar points.

We had also initially observed that there were no African Restaurants in this Neighborhood of Toronto.

Neighborhoods in Cluster 2, like Don Mills, Bedford Park, Oriole, etc., would be my recommendation for anyone looking to set up an African Restaurant in North York, Toronto. This is because the restaurant would very easily stand out and be noticed as there isn't much completion in these locations.

CONCLUSION

In this project, we have gone through the process of identifying a business problem, specifying the data required, extracting the data, preparing the data, cleaning the data and performing machine learning by clustering the date based on similarities.

We have provided insight and recommendations to business developers as it concerns locating an African restaurant in North York, Toronto.