Vegetarian-Friendly Cities in Toronto, Canada

Introduction

My name is Anthony Splendor and I have been a vegetarian for over 4 years. While society has made a lot of advancements in being more inclusive for those with vegetarian diets, there are still some places that are better than others. After I graduated college, one of my best friends moved to Toronto and she is also a vegetarian. Before her move we spoke about vegetarian options in Toronto and would have benefited from knowing the top neighborhoods with vegetarian-friendly dining.

People with vegetarian diets can often find themselves in a bad situation if they do not plan out their moves because there is great variance between cities in terms of vegetarian-friendly dining options. I experienced this first-hand when I moved to Baton Rouge, Louisiana for a summer internship. The city I lived in was not vegetarian-friendly and I often ended up cooking at home.

There are multiple stakeholders in this project who would benefit from the results produced. This would benefit those with a vegetarian diet currently living in Toronto and would possibly be planning a move to a different part of the city (perhaps for a new job opportunity or lifestyle change), those with a vegetarian diet planning on moving to Toronto, business and restaurants who want to target a vegetarian-based audience in Toronto, and the City of Toronto to provide resources to those with a vegetarian diet.

Data

In order to perform the analysis, we need two main sources of data:

1. List of neighborhoods in Toronto,
2. List of vegetarian-friendly restaurants in Toronto.

The neighborhood data can be retrieved from Wikipedia using a webscraper (BeautifulSoup) to transform into a format that is digestible by Python. This data is quite stagnant and would not be expected to change drastically over-time. However in order to accommodate changes made by the municipalities in zoning, any changes made to the Wikipedia page will update the script's dataset when run.

The vegetarian-friendly restaurants in Toronto can be extracted using Foursquare's API. Foursquare has compiled a vast database filled with businesses and how people interact with them. Although, in this analysis we will only be exploring restaurants and other dining establishments with vegetarian-friendly options using the vegetarian query. The data from the Foursquare API is received as a JSON file and must be transformed prior to use in Python.

For this analysis to be successful, the critical data fields required are latitude and longitude coordinates for not only the neighborhoods but also the dining establishments. This is key for mapping the various point on a graph and performing k-means clustering.

Methodology

With both necessary datasets identified, we can begin analyzing the data and performing exploratory data analysis. However, it is hard to perform this and identify trends when the data is in a tabular format. With this in mind, we can visualize the data by plotting the coordinates on a map using Folium.

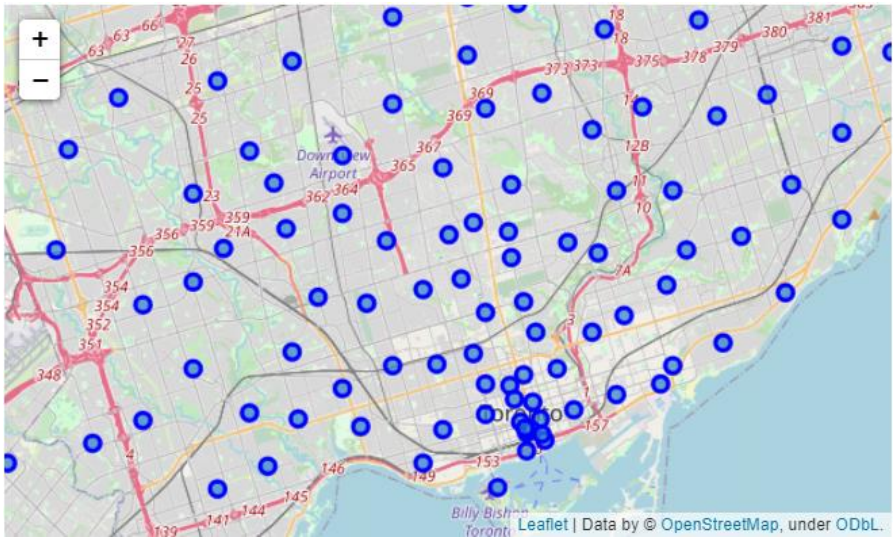

Fig. 1: Map of Neighborhoods in Toronto, Canada

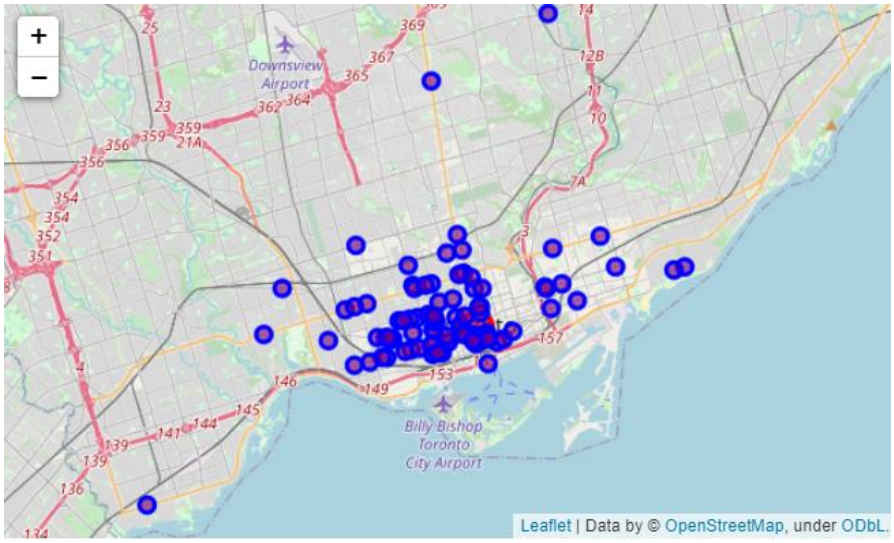

Fig. 2: Vegetarian-Friendly Restaurants in Toronto, Canada

Each blue dot as shown in Fig. 1 represents a neighborhood in Toronto. This is helpful to understand the data retrieved from the Wikipedia page. Fig. 2 shows purple dots with a blue outline representing a vegetarian-friendly restaurant or dining establishment. This data was retrieved from the Foursquare API. These initial Folium maps allow us to perform exploratory data analysis. From the maps, it is observed that most vegetarian-friendly restaurants are located closer to downtown Toronto, and that the density of restaurants is quite high there. While, on the outskirts of Toronto, there are practically no vegetarian-friendly dining options. It can be hypothesized that the k-means cluster analysis will reveal a cluster located within the downtown Toronto area. Exploratory data analysis is helpful because it allows for validation of the model deployed.

To find the optimal pair of coordinates with the highest density of vegetarian restaurants, a k-means cluster model was used. The various restaurants' coordinates were clustered to form regions. The purpose of this clustering model to create regions where the distance between the data points within the region (intra-) is minimized and the distance between regions (inter-) is maximized. Clustering models are unsupervised and do not require a training dataset, however they are dependent on the user to specify the number of clusters. The number of clusters can greatly affect the model's practicality, so the elbow method was used to validate the number of clusters. Elbow method relies on sum of squared errors (SSE) to calculate the squared difference between the observation (in this case the coordinates of the vegetarian restaurant) and the mean of that observation's cluster summed for all observations.

$$SSE = \sum_{i-1}^{n} (x_i - \bar{x})^2$$

Fig. 3: Sum of Squared Errors (SSE) Equation

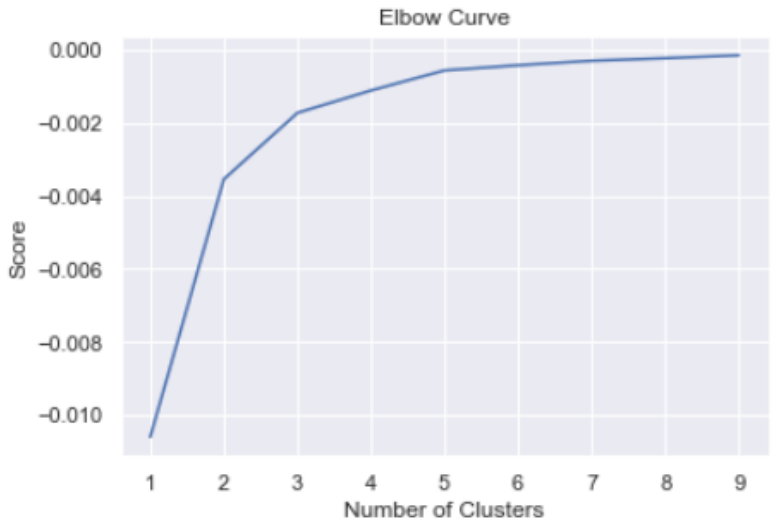

Fig. 4: Elbow Method Curve

Fig. 4 shows a graph created by using the elbow method to test different numbers of clusters in the model and calculating the SSE. The goal is to minimize the SSE, however after 3 clusters the value added is decreased and will create too many clusters, so 3 clusters were chosen as the optimal number.

The ensuing step is to find the densest cluster and calculate the distance from the cluster coordinates and all the neighborhoods in Toronto. This will result in the top 10 most vegetarian-friendly cities in Toronto, Canada (categorized by distance to the optimal cluster).

Results

Now that the data is organized into dataframes and the optimal number of clusters has been calculated, the k-means model can be deployed. This was done using the scikit-learn library. Using the matplotlib library, the results of the k-means model can be visualized as shown in Fig. 5 below.

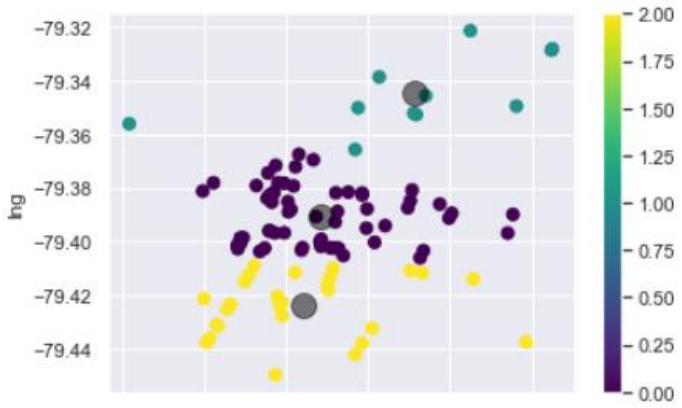

Fig. 5: Results of K-Means Clustering

From Fig. 5, the purple and yellow clusters are the densest with vegetarian-friendly restaurants. The clusters can be classified as 0, 1, and 2 where there are 28, 11, and 61 points within the clusters, respectively. Since these clusters refer to a specific coordinate, they can be plotted on top of Fig. 1, the list of neighborhoods in Toronto, Canada, to produce Fig. 6.

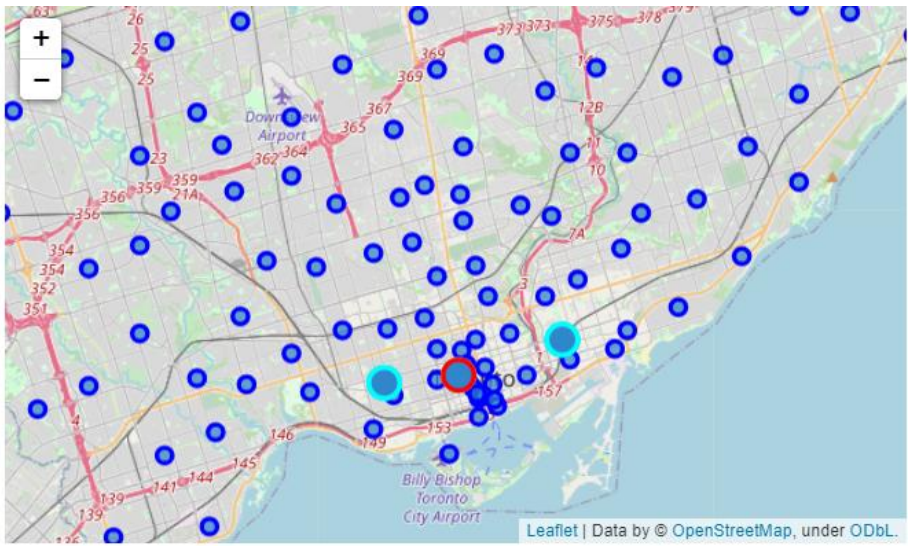

Fig. 6: Map of Neighborhoods in Toronto, Canada with Clusters

To reiterate, the 3 clusters are representations of regions where vegetarian-friendly restaurants are located. The red cluster had the greatest number of restaurants at 61. Using the coordinates of this cluster, we can append the neighborhoods dataframe with a new column that calculates the distance from the neighborhood to the cluster. Doing this, it is found that the following neighborhoods are the top 10 most vegetarian-friendly in Toronto, Canada:

1. Central Bay Street
2. Richmond, Adelaide, King
3. Ontario Provincial Government
4. Kensington Market, Chinatown, Grange Park
5. First Canadian Place, Underground city
6. Toronto Dominion Centre, Design Exchange
7. Garden District, Ryerson
8. Commerce Court, Victoria Hotel
9. University of Toronto, Harbord
10. Church and Wellesley

Discussion

The results were in alignment with the observations made earlier during the exploratory data analysis portion of the project. All the neighborhoods that made their way on the top 10 list were in the Downtown Toronto region. Using k-means clustering, it should be straightforward to figure out cities for other types of attractions and dining given the wide-range of data available in the Foursquare dataset.

Conclusion

Based on the results, we can conclude the top 10 most vegetarian-friendly restaurants in Toronto, Canada. This was derived using a k-means clustering model utilizing the

Foursquare dataset and a list of neighborhoods from Wikipedia. Those who adopt a vegetarian lifestyle should consider looking at this list of neighborhoods if they like to eat at restaurants that accommodate their diet to ensure they are not left stranded in terms of options. Business can also use this information to perform additional analyses on optimal business placement for vegetarian-centric establishments (perhaps yoga, crystal shops, etc.).