Capstone Project - The Battle of Neighborhoods

1. **Description of the Problem**

   A few business owners I know have a successful Vietnamese restaurant in the United States. Their restaurant is more of a fast-casual type of restaurant which is similar to Chipotle's model. The owners believe they have a unique advantage because most Vietnamese restaurants are traditional family run restaurants and have not broken out into chains. This is one of the primary reasons for this business's success.

   So far, this fast-casual Vietnamese chain has had a great success in parts of the United States. However, the business owners would like to expand into Canada specifically Toronto since one of the owners has ties there. Since Toronto is such a large busy city, the restaurant scene can be very competitive. However, the owners believe with the right intel they can setup shop in Toronto and at least have some form of success.

   The business owners want to focus on the core central neighborhoods of Toronto, because they thrive in busy areas where business professionals can grab a quick bite to eat on their lunch breaks. However, they don't want to detract from other Vietnamese restaurants in the area since they believe that the entire Vietnamese community should work together so that all can have an equal shot at success.

2. **Description of the data and how it will be used to solve the problem**

   The data for this project will first come from public sources from Wikipedia and Foursquare.

   Wikipedia will be used to gather location and neighborhood data for the central Toronto area.
   This data can be found here in the Postal Codes of Canada Wiki page:
   https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

   Foursquare will be used to gather the restaurant data for the area determined by the Wikipedia data.

   This data will be used to map the neighborhoods in central Toronto and cluster them into areas of analysis. Within in those clusters restaurants will be rated by most popular and their type. Not only do we want to see where Vietnamese restaurants are, it will be good to get a full picture of all the types in that area and how popular they are.

3. **Methodology**

   The first step in the methodology for the analysis was to gather and clean the data. This was done by first scraping the Toronto Postal Code data from the Wikipage. Once this was done the data was put into a data frame. In addition, data that was

null or not assigned was dropped from the table. The data was connected to the geospatial data using Folium and the neighborhoods mapped visually.

The next step was to collect the Foursquare API data. This data was connected to the geospatial data of the central Toronto neighborhoods. Data in this included restaurant data from the area.

A series of checks were done to make sure the restaurants had neighborhoods attached and to make sure the restaurant categories were represented.

Moving forward the specific restaurant names were not used since the goal was to get the actual categories of restaurants in a neighborhood or area.

The neighborhoods were then grouped and had rankings of categories of venues so that we could get an overall view of a neighborhood's top type of restaurants or venues.

The last piece was to create clusters of neighborhoods using K-Means Clustering to determine which neighborhood groups and restaurants can be combined. This is done to help keep the report information digestible.

Finally, a mapping of the clusters was created, and the reports were generated for the clusters.

4. **Results**

The First, Third, Fourth, and Sixth Clusters all had a Vietnamese restaurant within their top 5 most common venues.

The Fifth Cluster seems to wrap around the busy areas and is not ideal.

The Second Cluster is in the heart of the core of Toronto, which includes the financial district. Financial districts are usually some of the busiest areas in cities. It also appears to have very few Vietnamese restaurants in the area.

The Seventh Cluster also appears to not have a lot of Vietnamese Restaurants in the area. It also seems to not be too far from the University area so there is probably a lot of housing, however it is not directly in the University area so it might be difficult to attract students.

5. **Recommendations**

I would recommend a great location for a fast-casual Vietnamese restaurant would be in the second cluster area identified. More specifically the Commerce Court, Victoria Hotel neighborhood since this is the financial district and a famous hotel is there. Traditionally restaurants that are fast and reasonably priced do well with customers who are on the go and in a rush, but still want a good lunch or dinner. I think this would be the best place to look for a location to rent and start a pilot fast-casual Vietnamese restaurant in the Toronto area.

6. **Conclusion**

This analysis sought to find a solution for what area in Toronto would be a good place to launch a fast-casual Vietnamese restaurant, that won't impact other local Vietnamese restaurants and will get a lot of traffic.

The data used was from Foursquare and Wikipedia public data sources to get geospatial and restaurant information.

The analysis was done to bring this data together in a meaningful way.

This resulted in an understanding of the Toronto restaurant scene and what types of venues are popular in certain neighborhood locations.

Finally the recommendation was to look for sight in the Commerce Court, Victoria Hotel neighborhood area since that area matches all the main characteristics the owners were looking for to start a new pilot store.