CAPSTONE PROJECT -

INTRODUCTION

- Explore opening an authentic Thai restaurant in Toronto area
- It is a great opportunity for this entrepreneur who is based in Canada
- Finding the location to open such a restaurant is one of the most important decisions for this entrepreneur
- Designing this project to help him find the most suitable location

DATA AND DATA EXTRACTION

- List of Neighborhoods in Toronto
- Latitude and Longitude data geocoder package
- Foursquare API to get venue data

METHODOLOGY

- Extracting the list of neighborhoods from Wikipedia page
- Web scraping by utilizing panda's html table scraping method
- csv file to match the coordinates of Toronto neighborhood
- Visualized the map of Toronto using Folium package to verify coordinates
- Foursquare API to pull the list of top 100 venues within 500 meters radius
- analyze each neighborhood by grouping the rows by neighborhood and taking the mean on the frequency of occurrence of each venue category

METHODOLOGY

- Perform clustering method by using k-means clustering. K-means clustering algorithm identifies k number of centroids
- 3 clusters based on their frequency of occurrence for "Thai food" for location to open a restaurant

RESULT

- The results from k-means clustering show that we can categorize Toronto neighborhoods into 3 clusters
  - Cluster 0: Neighborhoods with little or no Thai restaurants
  - Cluster 1: Neighborhoods with no Thai restaurants
  - Cluster 2: Neighborhoods with high number of Thai restaurants

RECOMMENDATIONS

- Most of Thai restaurants are in Cluster 2
- Lowest (close to zero) in Cluster 1
- Cluster 1 might be a good location as there are not a lot of Asian restaurants in these areas
- Recommendation to entrepreneur is to open an authentic Thai restaurant in these locations with little to no competition