**Capstone Project: Final Report**
**Helping home buyers choose which neighborhood will best meet their needs.**

## Intro/Background

This project is intended for real estate agents who frequently work with out of town clients who may be unfamiliar with the area. Please imagine that you are a real estate agent in Toronto and often have a common problem when helping out of town clients. Many of your clients, who are unfamiliar with the Toronto area, will narrow down their choices to 2 or 3 houses that they liked. In making a final decision, your clients often tell you what types of venues and/or businesses they would like in their ideal neighborhood. The clients then often ask you to help them decide which of the 2 or 3 houses they are looking at are in a neighborhood with their ideal "wants." You have been learning data science, in the hopes of getting an upper hand in your industry and decide that you can use your newly acquired data science skills to help your clients.

You have a new client that has recently landed a lucrative data science job at a university in Toronto. You have shown the client and their partner several houses and they have narrowed their choice down to 3 lovely homes. Each home is in a different neighborhood in Toronto and, as usual, the clients have asked for your help in deciding which neighborhood would fit their needs the best. Your clients have picked 3 homes in the neighborhoods of 1) Berczy Park, 2) Queen's Park, and 3) Rosedale. The clients have told you that they MUST have 3 things in the neighborhood they choose that are extremely important to them. They tell you that these 3 things are 1) a park, 2) coffee shops, and 3) a gym. Your project is to now utilize your new skills in python, data analytics, and Foursquare to determine which neighborhood is best for your clients.

## Data

To solve this problem, we will use 2 data sources. First, we will use the following Wikipedia page, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. We will scrape this page to get the postcode data from the table and then transform it into a dataframe. The resulting dataframe will contain 3 columns for 'Postcode', 'Borough', and 'Neighborhood.' We will clean and analyze this data and use Folium to visualize the home locations and neighborhoods of interest. Next, we will use the Foursquare location data source to determine what venues (specifically those that our clients are interested in) are in the neighborhoods. We will then be able to utilize this data to make a final determination of which neighborhood is the best choice for our clients.