I currently have a friend in Toronto who lives in the Rosedale neighborhood. They have opened a successful business (yoga studio) in this area. They are currently analyzing expanding the business and opening several new locations. They feel that the biggest advantage that they have had in successfully running their business has been the lack of competition around them. Given that Toronto is a very large city, it would impossible to know every single neighborhood, let alone the potential spaces that a business could be opened in. In order to do this my friend has approached me to narrow down the search area and perhaps even provide a best recommendation for other neighborhoods to possibly open another location in.

In order to solve this problem we will first need to figure out what neighborhoods are similar to Rosedale since the business owner feels this is the largest contributor to their success. In order to do this we will use Foursquare mapping data in order to pull venues from the neighborhoods. Once this is done for all neighborhoods in Toronto we will determine a frequency for each type of venue. Finally we will use k means clustering in order to determine which neighborhoods are similar to Rosedale and thus should be explored further for a potential business expansion. There are many features that you can extract from this foursquare data. These include things like venues, comments, photos, etc. In order to accomplish this task we will simply be using the venue and the categories of the venues.