

# Capstone Project

## The Battle of Neighborhoods

Toronto is a fast pace growing city of Canada, a place to live, to work and to enjoy. In the early 1980 Toronto had surpassed Montreal as Canada's most populous city and chief economic hub. And the diverse population of Toronto reflects its current and historical role as an important destination for immigrants to Canada. More than 50 percent of residents belong to a visible minority population group, and over 200 distinct ethnic origins are represented among its inhabitants While the majority of Torontonians speak English as their primary language, over 160 languages are spoken in the city

### Problem Description:

Don Senior immigrated from Latin-American in early 1990 and after a successful carrier, he would like to give back to the community that help him succeed. He would like to invest in a new business venture specialized in products and services for Latin-Americans. Also knowing how difficult it was at the beginning, Don Senior would like to focus on the new generation.

### Target Audience:

The objective is to locate and recommend which neighborhood will be the best choice to start off their first Latin-American center.

### Success Criteria:

The success criteria of this project will be a good recommendation of the neighborhood choice in Toronto based on several key factors: higher number of households with Latin-American background, lack of specialize stores or lack of cultural venues/restaurant from Latin-America.

### The Data:

1. The City of Toronto provides a fair amount of public information in its open-data portal [<https://www.toronto.ca/city-government/data-research-maps/>]. The Data Portal catalogue includes over 295 data sets such as Festivals and Events, Licensed Child Care Centers, Priority Investment Neighbourhoods, Wellbeing Neighbourhood index and transportation data.
2. Toronto neighbourhood-profiles [<https://open.toronto.ca/dataset/neighbourhood-profiles/>] a summary of the 2016 Canada Census from Statistics Canada will be used to analyze socio-economical parameters such as population, immigration, households....
3. Toronto neighbourhoods real estate [<https://trreb.ca>] data can be access at The Toronto Regional Real Estate Board and will provide information about the affordability of the neighbourhoods

4. Boundaries of City of Toronto Neighbourhoods can be access at [<https://open.toronto.ca/dataset/neighbourhoods/>] and provide geolocation and shape of the neighbourhoods
5. FourSquare will provide the location and type of venues per neighbourhood

After analysis Feature engineering variables will be derived from the different dataset:

- Population Score/Ratio
- Immigration Score/Ratio
- Latin American Immigration Score/Ratio
- Household characteristics
- Average income
- Venues

We will also use clustering techniques such as the K-Clustering to segment and cluster these neighborhoods so that we can group them together to understand their similarities and what best we can do for these types of neighborhoods.

With all these features, techniques and data, we will then be able to come up with a best recommendation to the Don Senior to answer the which is/are the best neighborhood for them to first start off to offer their services.