Battle of the Neighbourhoods

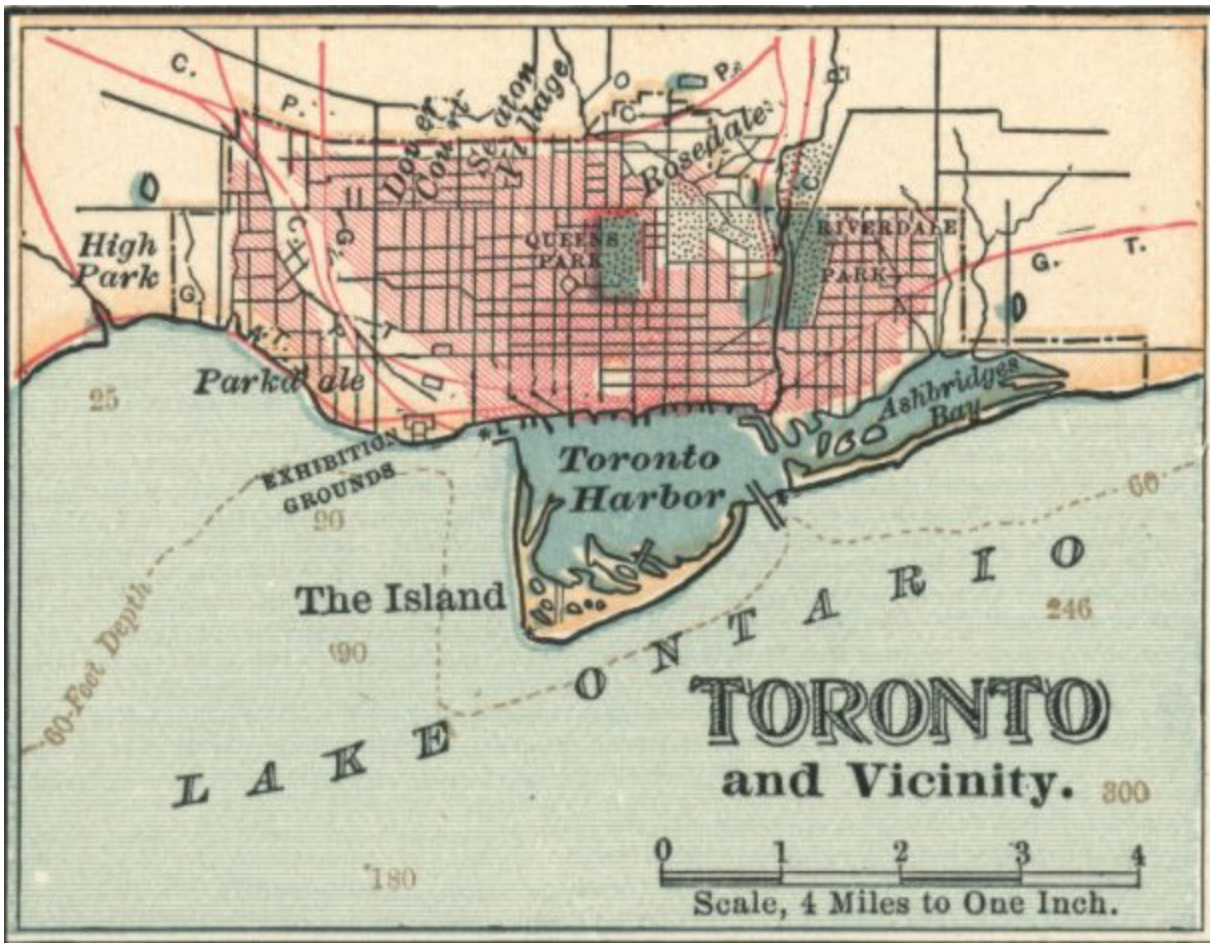

Table of Contents

Introduction - Identify the problem

We have a Canadian based Italian restaurant chain looking to expand into Toronto. The chain is primarily based in the Western Province of Alberta and British Columbia. The chain has been a success. The owner would like to build on that success by expanding into Toronto, which is known for its diversity.

Our objective is to leverage data on Toronto to help determine the best location for the restaurant. The owner is trying to find a location with little competition and as many customers as possible.

Target Audience

Our target audience is our client, the owner of the restaurant chain. Key stakeholders are the owner and his senior management team for the restaurant.

Data Overview

We will use 3 sources of data for this problem:

1. A list of postal codes, boroughs and neighbourhoods in Toronto. This list will be sourced from Wikipedia. These data points provide the information about the neighbourhoods that will be required to analyze the locations.

   Link - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

2. Geographical coordinates for the neighbourhoods and respective postal codes are the second source of data. This is required to create a map of the neighbourhoods so we can visualise the neighbourhoods. This data is provided in a CSV formatted file.

3. Data from FourSquare is the final component we require. The venue data is necessary to understand what venues and categories of venues are available in each neighbourhood in Toronto. This will help us determine the best location for the restaurant. This data will be combined with data from numbers one and two to help visualise and cluster the data points. API calls will be used to access the data from FourSquare.