

The Battle of Neighbourhoods

IBM Applied Data Science Capstone Project

1. Background

There are fewer foreign nationals living in the United Kingdom than citizens born in other countries. About 6.2 million citizens of non-British ethnicity were living in the United Kingdom in 2019 and 9.4 million were born abroad.

The United Kingdom's migrant population is concentrated in London. About 35% of people living in the UK who were born abroad live in the capital city. Likewise, about 37% of people living in London were born outside the United Kingdom, compared to 14% in the United Kingdom as a whole.

In 2019, Canada welcomed more than 300,000 immigrants, more than one in three immigrants has chosen to settle in the Greater Toronto Area (GTA). The GTA welcomed more immigrants (118,000 newcomers) than the four Atlantic provinces, Quebec, Manitoba, Saskatchewan, and Canada's three combined territories.

Given the fact that Toronto and London are far apart from each other, resemblances between these two cities can still be found.

2. Business problem

Relocation can be considered to be a major decision for a individual. In addition to job opportunities, environmental and cultural shocks are also a major concern for migrants as they move to another area. It is also beneficial for them to consider a neighbourhood comparable to the one they used to live in. In this project, we will introduce a machine learning technique to cluster Toronto and London areas in order to suggest neighbourhoods that are the best choice for migrants based on surrounding facilities such as school, hospital, shops, etc.

As a result, people who read in this article will be someone who wants to move from one city to another. This recommendation method will also help stakeholders interested in citing a business in a new city.

3. Data

We will use the following datasets for this project:

- Toronto.csv that consists of Toronto's postcodes boroughs, neighbourhoods.
Data source: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- London.csv that consists of London's postcodes, boroughs, neighbourhoods, post town.
Data source: https://en.wikipedia.org/wiki/List_of_areas_of_London

	PostalCode	Borough	Neighbourhood
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park / Harbourfront
5	M6A	North York	Lawrence Manor / Lawrence Heights
6	M7A	Downtown Toronto	Queen's Park / Ontario Provincial Government

Fig. 1 : Toronto Dataset

	Neighbourhood	Borough	Postcode
0	Abbey Wood	Bexley, Greenwich	SE2
1	Acton	Ealing, Hammersmith and Fulham	W3
2	Acton	Ealing, Hammersmith and Fulham	W4
3	Angel	Islington	EC1
4	Angel	Islington	N1

Fig. 2 : London Dataset

- A csv file that has the geographical coordinates of each postal code for neighbourhoods in Toronto is provided.

Data source: http://cocl.us/Geospatial_data

- The Foursquare API search tool is used to obtain information on neighbourhoods venues and can be used to identify and compare Toronto and London neighbourhoods.