Battle of Neighborhoods

Pablo Fernandez Cid

26/04/2020

1. Introduction

1.1. Background

Toronto is a Canadian city with a population of more than 2.6 million and divided in 140 neighborhoods. Currently, for this study case, a person is living in Etobicoke borough, more specifically in the neighborhood of Thistletown, with 10,360 residents and multiple services, such as three public school boards and three surrounded parks.

The majority of the residents live in houses, are married and have at least one children, according to the data from Neighborhood Profiles. The person is not married and doesn't have any child, and currently is living alone. He was born in Europe and he enjoys the food from different countries, but specially the Italian food.

1.2. Problem

Due to a job opportunity in Toronto East, our subject is thinking in moving to a new neighborhood, closer to this new possible job.

As this job opportunity is a improvement in his career, he expects to have also a improvement in the neighborhood, having in mind services, such as restaurants or entertainment, and also be surrounded by people in the same situation as him (single people, currently working and from Europe). Besides, he expects the same security at least as in his previous neighborhood.

1.3. Interested People

This analysis could be interesting also for people in the same situation as our subject or people interested in the city of Toronto.

2. Data

2.1. Data Used

For this study, the following data is going to be used:

- Foursquare API: Extraction of information regarding the venues per each Neighborhood

- Wikipedia: for subtracting information regarding the relation between postal codes and Neighborhoods (link here). This data will help to link the neighborhoods with their services.

- Neighborhood Profiles: information in order to understand the population, their income and their origin for each neighborhood. Data is extracted from multiple sources and merged by Open Data Toronto. See link here for more information

- Police Report information: Information regarding police report from 2014 to 2019 (link here). This data will be used to understand the security of each neighborhood, having in mind the number per each type of crime (Assault, Auto Theft, Break and Enter, Robbery and Theft Over) per year.

2.2. Data Preparation

Per each database, it is necessary a different preparation:

- Foursquare:
    1. Get data regarding the location per each Neighborhood. As the data we have it is the location of each Postal Code, Foursquare data will be linked to each Postal Code
    2. Calculate the frequency per each kind of venue per each Postal Code

- Neighborhood profile:
    1. Extract only the necessary rows:
        - Population Data
        - Age of the population
        - Type of dwelling: apartment, single-detached house among other
        - Family Characteristics, couples with/without children,
        - Income of individuals in 2015
        - Immigrants by continent
    2. Format the numbers and convert them to integer in order to operate with them

    Example of data for Neighborhood profiles

Topic	Characteristic	City of Toronto	Thistletown-Beaumond Heights
Immigrants by selected place of birth	Asia	674490	2655
	Americas	212010	1205
	Europe	298270	1140
	Africa	77445	460
	Oceania and other places of birth	3780	10
	Total	1265995	5470

- Police Report information:
    1. Group by Neighborhood
    2. Calculate per each Neighborhood the quantity of crimes per year from 2016 to 2019 and per categories

    Example per data for the sum of all the years

Neighbourhood	Assault	Auto Theft	Break and Enter	Robbery	Theft Over
Agincourt North	449	178	321	181	28
Agincourt South-Malvern West	707	220	479	164	80
Alderwood	218	97	148	41	41
Annex	1478	132	885	245	177
Banbury-Don Mills	483	131	439	90	62