**Relocation to another city**

Alexey Sytnikov

April 24, 2020

**1. Introduction**

In this project we will try to solve how to find the best place to move from your hometown. Particularly, this work will be targeted to people willing to move from Toronto (Canada) to New York (the USA).

There are many reasons why people decide to relocate, e.g. job offers, weather conditions, relationships and etc. Moving to another city, people still want to feel at home. That is why it is important to find the same conditions for living and choose a house with similar places next to it that people used to visit before they moved. So, the main beneficiaries of this problem solution are considered to be people who are willing to relocate for some reasons and to find comparable urban environment at the same time.

We will use several data science approaches and tools to find out the best places for relocation to New York for people from Toronto. To reduce the amount of computation power used we will explore one borough in each city. As our example task we can consider the following situation: there is a very successful trader who lives in Downtown of Toronto and who got a job offer from one of the best firms of Wall Street, and now he has to decide which Manhattan neighborhood to choose to move to.

**2. Data**

To implement the following analysis, we need information about both of the cities (New York and Toronto) representing as tables with the data about boroughs and neighborhoods of the cities and their geographical coordinates. Specifically, we are going to focus on one borough in each city (Manhattan and Downtown Toronto) Also, the venues that are located next to every neighborhood in selected boroughs should be included in the tables.

To create these tables, we will use some information sources:

* The dataset about New York City is extracted from the spatial data repository of New York University (<https://geo.nyu.edu/catalog/nyu_2451_34572>). This dataset contains all the necessary information: boroughs, neighborhoods and geographical coordinates of neighborhoods.
* The data about boroughs and neighborhoods of Toronto with the corresponding postal codes is scraped from the Wikipedia page (<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>).
* There is a file that contains all the postal codes of Toronto with the corresponding geographical coordinates.
* Information about centers of each city, borough and neighborhood is obtained using Nominatim, open-source geocoding tool.
* Data about all venues for chosen neighborhoods in each city is retrieved using Foursquare API.

As mentioned above, the data retrieved and scraped from different sources should be represented in the form of tables that can be used for analytical purposes. Two types of tables are formed. The first type of table represents each neighborhood being analyzed (Manhattan and Downtown Toronto) in the following shape (Fig. 1).

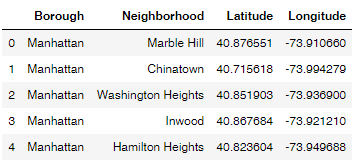


Fig. 1. Data about Manhattan neighborhoods and their coordinates

The other type of tables shows all the venues with their corresponding geographical coordinates and categories that are found for each neighborhood in a borough.

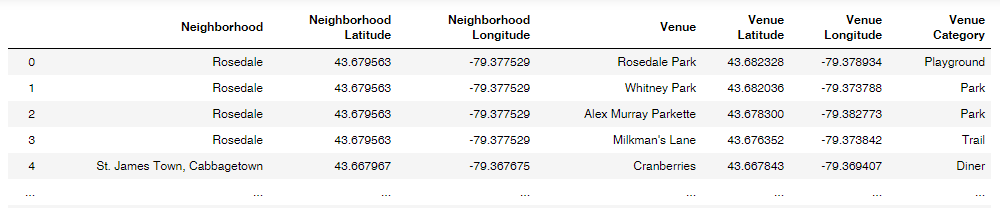


Fig. 2. Data about all venues pertaining to Downtown Toronto neighborhoods

It should be considered that there were some restrictions taken. First, all the venues were found within the radius of 500 meters from the center of a neighborhood. Second, there was a limit to the number of venues selected that was set on 50 by default.