IBM Data Science Professional Certificate

Applied Data Science Capstone

Week # 4

Capstone Project

The Battle of Neighborhoods (Week 1)

By Romero, Richard

February 15, 2021

Description of the problem:

Before finishing his degree, a Venezuelan university student wants to study English to become a more qualified professional in the competitive job market.

Of the English-speaking countries, he always wanted to know Canada. Canada is the North American country that stretches from the United States in the south to the Arctic Circle in the north. It is a multicultural country where diversity is promoted and where the two official languages are English and French, depending on the province.

After conducting an investigation he determined that there are two English schools in two different cities, “Kaplan International Languages” in Toronto, Ontario, and “ISS Language and Career College of BC” in Vancouver, British Columbia. Both schools are of high international prestige and also have similar tuition costs, however, none of the campuses offers internal accommodation, only outsourced, therefore, the place to stay must be chosen by relatives, so in order to make a decision this It must be based on the city, that is, what areas are safe and what they can offer to their surroundings.

The student's parents take the safety and comfort of their child very seriously, for this reason they hire our services to determine which is the best option.

Description of the data and how it will be used to solve the problem:

In order to analyze the structure of cities, we need to analyze the internal distribution of the neighborhoods of each city. To know the neighborhoods we will take the information from Wikipedia, in the postal codes section:

- Vancouver: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V
- Toronto: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Once the cities have been segmented, we proceed to determine the geographical coordinates of each neighborhood and store it in a dataframe.

Before proceeding to query Foursquare, it is determined which neighborhood has the lowest crime rate based on data obtained from city police departments.

- Vancouver Police Department: https://geodash.vpd.ca/opendata/

- Toronto Police Department: https://opendata.arcgis.com/datasets/94229ca8628f43f591a74ec3a2a6d4b5_0.geojson

The neighborhoods are classified according to the data of the police departments and once carried out, the Foursquare queries are carried out to classify the neighborhood.

Once done, we proceed to buy the cities and determine which is the best option.