COURSERA CAPSTONE Project

Coursera IBM Data Science Certification

Jose Mª Huertas Garrido

April 5th, 2020

Report Content

1- Description of the Problem and Background

    1.1- **Introduction**
- **Description of the problem and a discussion of the background**

    1.2- **Business Problem**
- **Description execute data processing**

2- Data Section

    2.1- **Description of the Data**
- **The following data is required to answer the issues of the problem**

    2.2- The data will be used to solve the problem
- **The data will be used as follows**

1. Description of the Problem and Background

1.1- Introduction

- **The intention of the study is the evaluation of the situation of the educational centers, the degree of crime or racial classification, the sexual influences on the educational level of the schools of the Toronto neighborhoods.**

- **We will identify the most conflictive neighborhoods, with the highest criminality, these data on the well-being of Toronto - Security.The data was provided by the Toronto Police Services, the Fire Services, the Paramedical Services (formerly EMS) and the Community Housing Corporation.most conflictive neighborhoods**

- **Using a study carried out in the different schools in Toronto, which identify the academic levels of each Toronto Neighborhoods Education Center. This dataset contains information related to the education of the 140 neighborhoods that make up the city of Toronto. Education includes 9th grade graduation rates, college applications, attendance at library, you etc.academic levels of each educational center**

    - Toronto's secular public school system of English language education is administered by the Toronto District School Board (TDSB). In total, this system manages 427 schools (an estimated 300,000 students study there annually, making the TDSB the largest school district in Canada and the fourth largest in North America).

1. Description of the Problem and Background

- Toronto has three universities. The central campus of the University of Toronto (U of T) is located in the city center and is the largest university in Canada, employing nine thousand people, and responsible for the education of more than 53,000 students. The U of T also has two smaller campuses, one in Scarborough and one in Mississauga. The three boards serve about 71,500 students a year. Two other universities are located in Toronto: York University, located in North York, and Ryerson University, located in the city center. The first serves 45,000 students a year, and the second serves 20,000.

- **We can classify schools according to neighborhood profiles. Obtained in the Population Census completed 5 years and collects data on age and sex, families and households, language, immigration and internal migration, ethnocultural diversity, aboriginal peoples, housing, education, income and work. Neighborhood Profiles of the city of Toronto use these census data to provide a portrait of demographic characteristics and economic aspects of people and households in every Toronto city neighborhood. The profiles feature selected highlights of the data, but these attachment files provide the complete data set gathered for each neighborhood.**
Population census carried out 5 years

1.2- **Business Problem**

- **The challenge is to demonstrate whether or not the academic level of each educational center may be influenced by aspects of the neighborhood such as age, sex, language, immigration or internal migration,and mainly, the most conflictive neighborhoods, with the highest crime rates.**

2. Data Section

2.1- Description of the Data

- **The following data is required to answer the issues of the problem**
    - List of Boroughs and neighborhoods of Toronto with their geodata (latitud and longitud).
      I used Forsquare API to get the neighbourhoods of Toronto
    - List of the most conflictive neighborhoods, with the highest criminality, these data on the well-being of Toronto - Security
      https://open.toronto.ca/dataset/wellbeing-toronto-safety/
    - List of data to identify the academic levels of each.

      Toronto Neighborhoods Education Center
      https://open.toronto.ca/dataset/wellbeing-toronto-education/
    - List Neighborhood Profiles of the city of Toronto use these census data to provide a portrait of demographic characteristics
      and economic aspects of people.
      https://open.toronto.ca/dataset/neighbourhood-profiles/

2. Data Section

2.2- The data will be used to solve the problem:

- **The datawill be used as follows:**
  - Use Foursquare and geopy data to map the best and worst neighborhoods of the Toronto Neighborhood Education Center.
  - Use homer and geopy data to map the location of Toronto city neighborhood profiles use this census data
  - Use Foursquare and geopy data to map the location of the most conflictive neighborhoods, with the highest criminality.
  - The addresses of the locations of the Education Center will be converted to geodata (lat, long) using Geopy-distance and Nominatim.
  - Data will be searched from open data sources, if available, from real estate sites, libraries or other government agencies such as the City of Toronto Open Data.
  - The processing of these DATA will allow answering the key questions to make a decision:
  - What is the Education Center with the best level in areas with high levels of crime?
  - What is the Education Center with a major racial variety?
  - Which are the best Educational Centers in less favored areas? What are the worst educational centers in areas with low crime and worse educational levels?
  - Any other interesting finding of statistical data from Educational Centers.