The Battle of Neighborhoods

Sebastian Toledo

March 08, 2020

1. Introduction

1.1 Background

Today, technology represents an incredible and fundamental tool for society that goes beyond what was previously considered possible and one of the most exciting and requested areas is data science, since thanks to various techniques such as Data Analysis, Machine Learning and Deep Learning have allowed to solve innumerable problems in a very efficient and fast way. That is why my interest and motivation in contributing to society in a responsible way, starting with this first contribution that focuses on analyzing and comparing neighborhoods in two of the most exponential and important cities in the world, New York and Toronto.

1.2 Problem

Through Data Analysis and some Machine Learning algorithms we will look for patterns that help us understand the similarities and differences that the neighborhoods of both cities have. Information that is very valuable for tourists, business people or interested in living there.

2. Data acquisition

For this project we will use information provided by two sources that I will cite below.

Dataset containing relevant information from the city of New York: https://geo.nyu.edu/catalog/nyu_2451_34572

Website containing relevant information about the city of Toronto: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.

This information is essential to respond to our analysis as it contains longitude, latitude and corresponding neighborhood of each city, we also have the Foursquare API that provides us with additional information as more relevant places that make up each neighborhood and their respective information provided by visitors.

To respond to our problem, we will use the information provided and we will categorize and segment it to later graph it and be able to explain through a statistical summary the main factors that occur in each neighborhood and how they are related, we can also give a clear perspective of the similarities and differences between New York and Toronto.