Capstone Project - The Battle of Neighborhoods

Jawharah Almulhim

10 June -2020

Introduction:

New York city also called NYC , it is one of the largest city on USA with huge number of population and diversity of people. The NYC is a main distance for visitors from all over the world . The idea of this project is to explore Manhattan which is often referred to by residents of the New York City area as the City, it is one of the most densely populated of the five boroughs of New York City.

Exploring neighborhoods and venues of this borough is handled on this project. Also , suggest a Higley rated parks on Manhattan is carried out on the project since mostly, visitors are willing to visit different places to enjoy themselves and parks are one of these places..
Finally , ending up with clustering different parks of Manhattan into different clusters with similar features.

Data

To accomplish this project , different data sources are used:

- **New York City data** that contains list Boroughs, Neighborhoods along with their latitude and longitude. Data source : https://cocl.us/new_york_dataset Explanation: the above data set is available for free and it contains main data of NYC like latitude, longitude , boroughs and neighborhoods.
- **Foursquare API service:** using API calls to get neighborhoods and venues of the selected borough as well as detailed information about venues such as tips, likes, rating and more. Such information is necessary for clustering.
- **Pandas data frames** is used to store the results of the API calls and do the operations
- **Geopy** is client which is used to locate the coordinates of addresses using third-party geocoders
- **K-mean clustering :**machine learning tool to cluster the parks on different cluster based on similarities