

Battle of Neighborhoods

Seattle, Washington

**13 CNC Districts
Overlaid on
7 City Council Districts**



This map shows the
City Neighborhood Council's
13 Districts overlaid on top of the
newly-voted 7 City Council Districts

City Neighborhood Council
Districts are named using bold text

Seattle City Council Districts
are shown in different colors

Project Description:

Usually people tend to explore the places before moving to a new country, state, city or place for their work or to start a new life and that involves so many factors including neighborhood analysis. That calls for a search algorithm that usually returns the requested features such as population rate, median house price, school ratings, crime rates, weather conditions, recreational facilities etc. It'd be nice to have a hassle-free application that could spit out an extensive analysis of all these features for a neighborhood or a comparative analysis between neighborhoods with just sending out the names of the neighborhoods

This Project takes this requirement as its main idea to help the stake holders to achieve the desired results, so as to help them spend their time and money in some other productive means rather keep them trapped in an infinite loop of extensive search engines.

This Project would help the stakeholders take a better decision on choosing the best neighborhood out of many neighborhoods to build/buy their houses in Seattle city based on the distribution of various facilities in and around that neighborhood. As an example, this project would compare 2 randomly picked neighborhoods and analyses the top 10 most common venues in each of those two neighborhoods based on the number of visits by people in each of those places. Also, this project uses K-mean clustering unsupervised machine learning algorithm to cluster the venues based on the place category such as restaurants, park, coffee shop, gym etc. This would give a better understanding of the similarities and dissimilarities between the two chosen neighborhoods to retrieve more insights and to conclude with ease which neighborhood wins over other.

Data Sets and APIs:

Foursquare API:

This project would use Four-square API as its prime data gathering source as it has a database of more than 105 million places, especially their places API which provides the ability to perform location search, location sharing and details about a business. Photos, tips and reviews jolted by Foursquare users can also be used in many productive ways to add value to the results.

Work Flow:

HTTP requests would be made to this Foursquare API server using zip codes of the Seattle city neighborhoods to pull the location information (Latitude and Longitude).

Foursquare API search feature would be enabled to collect the nearby places of the neighborhoods. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.

Folium- Python visualization library would be used to visualize the neighborhoods cluster distribution of Seattle city over an interactive leaflet map.

Extensive comparative analysis of two randomly picked neighborhoods would be carried out to derive the desirable insights from the outcomes using python's scientific libraries Pandas, NumPy and Scikit-learn.

Unsupervised machine learning algorithm K-mean clustering would be applied to form the clusters of different categories of places residing in and around the neighborhoods.

These clusters from each of those two chosen neighborhoods would be analyzed individually collectively and comparatively to derive the conclusions.

Python packages and Dependencies:

- Pandas - Library for Data Analysis
- NumPy – Library to handle data in a vectorized manner
- JSON – Library to handle JSON files
- Geopy – To retrieve Location Data
- Requests – Library to handle http requests
- Matplotlib – Python Plotting Module
- Sklearn – Python machine learning Library
- Folium – Map rendering Library

-----THANK YOU-----