**Data – Battle of Neighbourhoods**

For our analysis we need an overall idea of New Jersey state in general and Essex county in particular.

Use 'https://www.zipcodestogo.com/New%20Jersey/' link to get the postal codes of counties and cities.

Then use "Beautiful soup" to retrieve the postal codes, county and city information.

Perform exploratory analysis on data to check if it suits the requirements.

Data with Null values should be removed.

Data should not contain duplicate values.

Data should be formatted correctly and have the correct domain.

Once the data is standardized the qualified data is ready to be processed.

The geo location of the city-county should be obtained from Geocoder.

The venues for each city-county should be obtained from FourSquare.

The city-county for which geo location and venues are not available in Geocode and FourSquare can be deleted from the dataset as it doesnot add any value.

Then cluster the nearby venues to find out the top ten amenities available for each location and extract those features. Based on the amenities cluster the neighbourhoods which have similar characteristics. These venues can be shown to the clients to choose a location.

Advantage is that given a location, we get an overview of the facilities available and it gives us the knowledge to choose the best options.