

Data Section

The data that will be used to solve is the location data of different colleges, hostels, and offices where the residence areas of different working peoples will be present.

For example the New York city has coordinates as 40.7128° N and 74.0060° W.If we take the geographical neighborhoods of that city by folium and explore them based on their similarity/nature. So it will be very easy to cluster them using KMean clustering technique in SkLearn. So our main concentrated data is on that cluster which has majority of schools , offices, business places like cinema halls, community centers, sports halls,etc.,

So to summarize we require the foursquare data of the New York city to explore its neighbor hoods and their coordinates .we also require the https://geo.nyu.edu/catalog/nyu_2451_34572 dataset inorder to access the neighborhoods of the city.

By following the above mentioned process we get the cluster with desired data, and choosing the location from that cluster would more easier and practical.