The Battle Of Neighborhoods - Mumbai

Introduction/Business Problem

We have data about the areas in Mumbai City, India. Our goal is to cluster venues in Mumbai according to their categories for eg. cafe. book shop, etc. We need to use FourSquare API to retrieve data about these venues from the locations data we have and use K means Clustering to form clusters of these venues and potray them on a map of the city.

Data

The wikipedia page of the city contains location data with pincodes for the areas in Mumbai City.

Mumbai neighbourhood coordinates [edit]

Area	Location	Latitude	Longitude
Amboli	Andheri,Western Suburbs	19.1293	72.8434
Chakala, Andheri	Western Suburbs	19.111388	72.860833
D.N. Nagar	Andheri,Western Suburbs	19.124085	72.831373
Four Bungalows•	Andheri,Western Suburbs	19.124714	72.82721
Lokhandwala	Andheri,Western Suburbs	19.130815	72.82927
Marol	Andheri,Western Suburbs	19.119219	72.882743
Sahar	Andheri,Western Suburbs	19.098889	72.867222
Seven Bungalows	Andheri,Western Suburbs	19.129052	72.817018
Versova	Andheri,Western Suburbs	19.12	72.82
Mira Road	Mira-Bhayandar,Western Suburbs	19.284167	72.871111
Bhayandar	Mira-Bhayandar,Western Suburbs	19.29	72.85
Uttan	Mira-Bhayandar,Western Suburbs	19.28	72.785
Bandstand Promenade	Bandra,Western Suburbs	19.042718	72.819132
Kherwadi	Bandra,Western Suburbs	19.0553	72.8314
Pali Hill	Bandra,Western Suburbs	19.068	72.826
I.C. Colony	Borivali (West),Western Suburbs	19.247039	72.84983

We will use BeautifulSoup library to scrape this data from the webpage, as we can see we already have the coordinates data available as well.

Further we will use FourSquare API to get venues close to the locations we have retrieved from the

	Location	Location Latitude	Location Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Andheri,Western Suburbs	19.1293	72.8434	Cafe Arfa	19.128930	72.847140	Indian Restaurant
1	Andheri,Western Suburbs	19.1293	72.8434	5 Spice , Bandra	19.130421	72.847206	Chinese Restaurant
2	Andheri,Western Suburbs	19.1293	72.8434	Subway	19.127860	72.844461	Sandwich Place
3	Andheri,Western Suburbs	19.1293	72.8434	Cafe Coffee Day	19.127748	72.844663	Coffee Shop
4	Andheri,Western Suburbs	19.1293	72.8434	Apple Service Centre	19.128036	72.842718	IT Services

This is the final data that will be used for k means clustering