Capstone Project- Battle of the Neighbourhoods in the city of Mumbai, India

Data Cleaning:

The table present in the website was procured using the Pandas ‘read_html()’ method:

DataFrame gathered from website

The DataFrame was then cleaned to only include names of the post offices which are the

The names of neighbours were then extracted from the above DataFrame and a new dataframe was created consisting of the neighbourhood names and their latitude and longitude.

ThT prTvious map has bTTn appropriatTly zoomTd to highlight thT relTvant nTighbourhoods in Mumbai.

Thus, thT dataframT is shortTnTd to only thosT neighbourhoods which arT within 25Kms from

The map of Mumbai with relevant neighbourhoods is as follows:

The dataframe created after getting the response from the Foursquare API as nearest venues within 1000m from a neighbourhood is as follows:

Analysis

A bar graph was plotted using the number of unique venue categories along the Y axis with the corresponding neighbourhood

Results and Discussion

Bas3d on th3 bar graphshowing th3 numb3r of uniqu3 v3nue cat3gori3s for each

Then the Foursquare data was again used in order to find venues pertaining to food for all the neighbourhoods to determine the saturation of a neighbourhood with respect to number of food joints. Optimal locations were identified as those neighbourhoods which are part of the best clusters (having neighbourhoods with most diverse amenities) yet have the least