

Data Used for Capstone Project: Two Sets of Data were used to complete this project.

Mapsofindia.com is a website that lists location across every city in India. I used this website to extract different neighbourhoods in Bangalore. I scrapped the page:

<https://www.mapsofindia.com/pincode/india/karnataka/bangalore/>

This page gave the lists of various pin codes in Bangalore, which is kind of proxy of areas with reasonably equivalent population. Along with the pincode, the url gave major areas belonging to that pincode.

For this project. I later used Nominatim to receive latitude and longitude of all the areas. Since I was focussing only on South East Bangalore, I selected only those areas whose latitude was lesser than average latitude of all listed areas, and longitudes higher than average.

API calls from Four Square to extract the number of venues that are available in each location are made. Instead of name of the data since we are more interested in categories of data, categories of venues was specifically extracted. Number of instance of particular categories in given area gives idea about how frequent is this kind of category in given area. Top 20 categories were later identified and rest categories were discarded as these datas were too sporadic to give any character.