ACCESSING THE POPULATION SIZE, DENSITY AND VENUES INFORMATION OF LAGOS NEIGHBOURHOODS

OBAFEMI ADEPOJU

DATA DESCRIPTION

The data for this project was obtained in unstructured format. A table was produced manually from information sourced from official documents and from the internet. The data requirement was divided into four parts based on how they were obtained.

1. Postal codes of the different neighbourhoods (Local Government Areas);
2. Lagos State official population data listing the Local Government Areas (neighbourhoods) with their population densities;
3. Geo-coordinates of all neighbourhoods of Lagos;
4. Venues information of the different neighbourhoods.

Postal codes of the neighbourhoods were obtained from different internet sources but mostly from nigeriapostcodes.com The codes were important as they distinguish one neighbourhood from the other and help in generating coordinates.

A crucial data for this project is the population data which gives all statistics about population size and density (per square kilometer, sq.km) of the 20 neighbourhoods. This data was obtained from the Lagos State Bureau of Statistics 2017 official document. The 2017 version of the document was the most recent version available for download at the time of working on this project. This data will be used on the map where it will be placed superimposed for all neighbourhoods on the map.

Geo-coordinates of the 20 neighbourhoods were all obtained online from different sites. There was no single official source for this data, hence, they had to be scraped from different sites one by one. This data will be used to locate the neighbourhoods on the map of Lagos State.

Lastly, venues information about the neighbourhoods were obtained using Foursquare API. The information obtained include name of the venues and the categories. The data will be cleaned, processed and passed through unsupervised machine learning for segmentation and clustering analysis.

All datasets obtained were entered into a table to generate a .csv file which was used to create a pandas dataframe for descriptive analysis and machine learning modeling.

	Postal Code	Neighbourhoods	Population	Population density	Latitude	Longitude
0	100283	Agege	1415547	60768.47	6.619830	3.322165
1	102103	Ajeromi Ifelodun	1966700	103258.63	6.455480	3.333860
2	100275	Alimosho	2804919	14855.05	6.610556	3.295830
3	102102	Amuwo Odofin	719337	2931.16	6.450000	3.266667
4	102241	Apapa	715792	13568.42	6.449998	3.366665

Fig 1. A view of the first 5 rows of the dataset in a pandas dataframe