Week 4 Applied Data Science Capstone: The Battle of Neighborhoods

Business Proposal: Data Description

Data Description

This project aims to explore the neighborhoods of Lagos and carry out an analysis of the restaurants. The project requires Lagos datasets and will employ data-driven methods. Hence, this project will use the datasets from Wikipedia: https://en.wikipedia.org/wiki/Lagos_State, to explore the neighborhoods information. The dataset contains the Local Government Areas (LGAs) and the cities/towns in those Areas. Other important information will be delineated to help in the analysis. The following methods will be used to retrieve, process, and analyze all the information needed for this project.

a) Neighborhood data

The data information about the neighborhoods of Lagos will be obtained by scraping the Wikipedia page and transforming the data into a Dataframe before data cleaning.

b) **Geospatial information**

The Geocoder package will be used to generate the geospatial information which includes the latitudes and longitudes of the areas.

c) The Foursquare API

The Foursquare API will be employed to gather location information of the neighborhoods. The venue information of the LGAs will be retrieved using the Foursquare API and the restaurant's categories in these venues will be analyzed to provide preferable information for the visitors.

d) K-means algorithm

The neighborhoods and venues clustering will be done using the K-means algorithm.

e) Folium map visualization

Folium maps will be used to visualize the neighborhoods and venues cluster distributions.