**Analyzing New York City Neighborhoods for the purpose of starting an African**

**Restaurant**

**IBM Applied Data Science Capstone Project**

**By: Emmanuel Ekanem**

**Date: November 28, 2020**

**INTRODUCTION/BUSINESS PROBLEM**

New York City (NYC) is the capital of the state of New York and one of the most densely populated cities in the world. New York City is the most populous city in the United States, with an estimated 8,336,817 people living in the city, according to U.S. Census estimates dating July 1, 2019 (up from 8,175,133 in 2010; 8.0 million in 2000; and 7.3 million in 1990). This amounts to about 40% of New York State's population and a similar percentage of the metropolitan regional population. New York's two key demographic features are its population density and cultural diversity. According to a 2001 study by Claritas, four of the city's five boroughs ranked among the twenty most diverse counties in USA. Queens ranked 1st, Brooklyn 3rd, Manhattan 7th, and The Bronx 17th. In addition, Hudson County and Essex County, New Jersey, both of which are part of the New York Metropolitan Area, ranked 6th and 15th, respectively. NYC lies on the west coast of USA and attracts heavy tourism from all over the globe every year. Personally, I have friends and family members in New York City whom I visit from time to time, and a major part of my stay is relishing the diverse cuisines. As a major socioeconomic centre in the world, NYC is extremely diverse with people from various culture and cultural affiliations residing here. The multi-cultural nature of New York City is one reason why it has numerous cuisines from all over the world. Immigrant Africans, Caribbeans, and African Americans make up 25.1% of New York City's population. According to the U.S. Census Bureau, there were 2,086,566 black people residing in New York City. Africans love food a lot, and it is often said that the way to a person's heart is through the stomach. Compared to other parts of USA, more Africans reside in NYC. So, having an African Restaurant will give not only Africans but also African Americans and other lovers of African foods; a feeling of home away from home. However, from my stay in NYC and other research, we do not have much African Restaurants there. The purpose of this project is to

resolve this problem through the study of NYC neighbourhoods to determine possible locations for starting an African Restaurant. This project will be useful for entrepreneurs and angel investors who are looking to invest and open an African Restaurant in NYC. Our project will consist of empirical data analysis to provide rewarding recommendations for these stakeholders.

## DATA COLLECTION

New York City has a total of 5 boroughs and 306 neighborhoods. In order to segement the neigh borhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and th e neighborhoods that exist in each borough as well as the the latitude and logitude coordinates of each neighborhood. Our data sources will include the following, https://en.wikipedia.org/wiki/New_York_City , https://en.wikipedia.org/wiki/Demographics_of_New_York_City , https://cocl.us/new_york_dataset , Python geopy library and Foursquare API data. The Wikipedia data will be scraped using BeautifulSoup package to get population demographics and cultural diversity of NYC. The New York dataset and python geopy library will be used to get the geographical coordinates of NYC boroughs and neighbourhoods. This data will also be useful for plotting the map of NYC with its neighbourhoods superimposed on top by means of the python folium library. The Foursquare API platform will be used to extract and explore venue data. This data which contains venue recommendations for the boroughs and neighbourhoods in NYC will be used to analyse African Restaurants in the different neighbourhoods and build the unsupervised learning model to cluster neighbourhoods. There will be a limit of 100 venue recommendations per neighbourhood and a radius of 1000 meters around the geographical coordinates of each neighbourhood.