Battle of Neighborhoods of -Port Elizabeth & **Pretoria in South Africa**

Capstone Project

2. Data Requirements

Required data can be gathered from:

- **Port Elizabeth** and **Pretoria City** information, including districts and neighborhoods, can be obtained from Wikipedia:

o (Source - https://en.wikipedia.org/wiki/Port_Elizabeth) &

o (Source - https://en.wikipedia.org/wiki/Pretoria)

- Photos and Picture of **Port Elizabeth** and **Pretoria** City used for Presentation from

(Source - https://afrotourism.com/travelogue/)

- The data used for this project will be acquired from http://www.sapostalcodes.info. The datasets consists of the postal codes and suburb names of each city.

- In order to obtain venues and their categories we will use Foursquare API search feature [FOURSQUARE] (https://foursquare.com/) will be used to collect neighborhood venue information as well as the longitude and latitude details of each suburb. Details about local venues and locality will provide insight into the qualities of a neighborhood.

- In addition to Foursquare, various python packages will be used to create maps and machine learning models to gather further insights and provide efficient recommendations and results into our neighborhood battle project.

These packages includes:

1. Pandas - Library for Data Analysis

2. NumPy – Library to handle data in a vectorized manner

3. JSON – Library to handle JSON files

4. Geopy – To retrieve Location Data

5. Geocorder - For geolocation of neighborhoods

6. Requests – Library to handle http requests

7. Matplotlib – Python Plotting Module

8. Sklearn – Python machine learning Library

9. Folium – Map rendering Library