**SEGMENTING AND CLUSTERING COMMERCIAL ENTERPRISES IN NAIROBI, KENYA**

**INRODUCTION: BUSINESS PROBLEM**

Due to the strategic location of certain venues for commercial enterprises in Nairobi city, Kenya near the central business district, there is a rising number of restaurants and recreational facilities in these area and this has increased the population of Nairobi City County rapidly, which has led to traffic congestion and air/land pollution.

Therefore this project analyses the data about the geographical location of these venues in order to visualize the distribution of these facilities by clustering them into different categories that will inform investors on specific location in Nairobi to set up facilities that are less found in those areas.

Also this analysis will inform the government on the strategy to improve city connectivity via better roads to and from these venues, whereas have a better understanding of the living standards of citizens from various locations in the city and reroute traffic towards other constituencies at rush hours.

**DATA**

The data used in this project is a table of the current constituencies of Nairobi city county, Kenya with the corresponding wards in each constituency. The geographical location of these constituencies and the foursquare data on the venues of different facilities located in these constituencies. The data is collected from a web page through web scrapping using the python pandas library to read the table from the website. The data about the venues located in each of these constituencies is from foursquare which is requested using foursquare API calls in python libraries; *requests and json.normlize.*

The link where I found the constituency data is: <http://www.infohub.co.ke/2016/04/list-of-nairobi-constituencies-and-wards.html>.

The data will be used to cluster constituencies based on the category and number of facilities located in each constituency in order to determine the type of economic activity that is predominant in most places of Nairobi in order to inform the decision of investors on what kind of commercial enterprises are still lacking and could be lucrative if set up in certain locations of Nairobi.

The government can use this data and analysis to identify congested constituencies and those that still have spaces so that some of this facilities could be set up in those constituencies hence diverting traffic towards those locations and reducing air pollution in the city central business district.