## APPLIED DATA SCIENCE CAPSTONE PROJECT

# FLATTENING THE CURVE AS ECONOMIC ACTIVITIES RESUME: LOCATION DATA ANALYSIS OF LAGOS STATE IN REFERENCE TO THE COVID 19 PANDEMIC



Figure 1: Health Workers with a COVID-19 Patient. Source: Nairametrics.com

# 1.0 DESCRIPTION OF THE PROBLEM: INTRODUCTION

# 1.1 BACKGROUND

Coronavirus Disease 2019 (COVID-19) is an infectious disease caused by the novel coronavirus.<sup>1</sup> This disease which is currently a pandemic broke out in Wuhan, China in December 2019. As at 27 April 2020, there were 2,878,196 confirmed cases and 198,668 deaths recorded from the disease globally.<sup>2</sup> Due to its highly infectious nature, countries have had to take up strict social distancing measures such as imposing a lockdown. One of the adverse effects of these measures is the impact on the economy.

In Nigeria, the index case of COVID-19 was confirmed on 27 February 2020, when an Italian tested positive in Lagos two days after arrival from Milan, Italy. <sup>3</sup> Based on the Nigerian Center for Disease Control (NCDC)'s COVID-19 data, as at 27 April 2020, the country confirmed 1,337 cases, 255 cases discharged, and 40 deaths recorded. <sup>4</sup> Total active cases as at 27 April 2020, were 1,042. <sup>5</sup>

Lagos State, being the business centre of the country has been the most hit with 764 confirmed cases, 143 discharged and 19 deaths recorded. The active cases as at 27 April 2020 were 602, representing 57.77% of the country's total active cases. <sup>5</sup>

On the 29<sup>th</sup> of March 2020, the president of the Federal Republic of Nigeria announced a two-week lockdown of Lagos state, Ogun state and the Federal Capital Territory (FCT), Abuja. <sup>3</sup> The lockdown was extended for another two weeks on 13 April 2020, and for another one week on 27 April 2020; however, the president announced that the lockdown will be relaxed after the one-week extension (04 May 2020) but with an indefinite curfew from 8 pm to 6 am. <sup>3</sup>



Figure 2: The City of Lagos. Source: nomadaafricamag.com

## 1.2 THE PROBLEM

Lagos is a state in South-Western Nigeria. Although it is the smallest in area per km² among all 36 states in Nigeria, it currently has the highest population density. <sup>6</sup> According to Ekundayo (2013), if Lagos state was a country, it would be the fifth-largest economy in Africa. <sup>7</sup> Lagos, Nigeria is not the Federal Capital Territory, but it is popularly tagged as the Business/Economic Capital of Nigeria as it is home to various enterprises, investors and headquarters of multinationals/conglomerates. The search for employment and source of living is one of the reasons for its high population density, as there is a constant inflow of citizens from other states. It is obvious that a continuous or an extended lockdown of the state will largely impact the economy negatively. On the other hand, Lagos state is at high risk of a surge in the number of COVID-19 cases once the lockdown is eased due to its population and business activities. A surge in the number of cases could result in repeated lockdowns and an unpalatable cycle that can negatively affect public health and the economy.

The phrase "the new normal" has become popular in the social media space; companies and business owners have been advised to embrace it to adapt and survive the COVID-19 crisis.

The question is: How can Lagos State flatten the curve while economic activities take place?

This project will involve the use of location data of the twenty (20) LGAs in Lagos state and their COVID-19 distribution to identify venues that require strict focus in terms of public health measures to curb the spread. The idea is to curb the spread of COVID-19 in the most common venues in the LGAs as economic activities resume.

**AIM:** This project aims to develop strategies based on location data analysis, to reduce the spread of the COVID-19 virus in Lagos state as economic activities resume.

**OBJECTIVES:** The objectives of this project are as follows:

- 1. Identification of Lagos state Local Government Areas (LGAs) most affected by the COVID-19 virus.
- 2. Exploration of location data of LGAs with the highest COVID-19 cases using Foursquare API.
- 3. Segmenting and clustering of LGAs to identify defining features concerning COVID -19 spread by leveraging on Foursquare location data.
- 4. Development of strategies to reduce the spread of the virus based on the result of the analysis.

### 1.3 JUSTIFICATION

The COVID-19 pandemic has been established as a threat to both the economy of Lagos state and the health of its populace. A lot of people are scared of death from the virus or its negative impact on their income. As a Biostatistician and Public Health Data Scientist, I am worried about the possibility of a surge in the number of COVID-19 cases in the state when the lockdown is relaxed on 04 May 2020, as stipulated by the Federal Government. However, I also understand that the state cannot afford an economic shutdown as the economy already had challenges pre-COVID 19. An economic shutdown would not only affect the Lagos state economy but also Nigeria's economy at large. The need to devise strategies to flatten the curve without gross negative impact on the economy cannot be overemphasized. I believe without a doubt that this is the solution Lagos state needs at the moment.

### **1.4 TARGET AUDIENCE**

The target audience for this report includes the following:

- The Lagos State Government: This report would be beneficial for decision-makers in the areas of health and commerce in Lagos state. These decision-makers include the Governor of the state, Lagos State Ministry of Health, Lagos State Ministry of Commerce and Industry and Lagos state Ministry of Finance. The recommendations of this report will be useful to them as they strategize together to flatten the curve and keep the economy thriving.
- 2. **NCDC and Federal Government:** The NCDC and the Federal Government must work with the Lagos state government in strategizing and implementing solutions that would help flatten the curve and protect the economy.
- 3. Chief Executives and Business Owners: Decision makers of companies, conglomerates, multinationals, banks, entrepreneurs need to understand how they can protect their staff, customers and businesses. The analysis using location data would help to identify venues especially business locations that require strict social distancing measures. This report might help some businesses decide to adapt their business to work without a physical location. Ignoring the public health measures for the sake of profit is counterproductive as increased outbreak would only result in

- repeated lockdowns. This category needs all the information they can get to balance it all, part of which they will get from this report.
- 4. **Lagos State Residents:** A COVID-19 LGA analysis based on location data will be beneficial to residents of the states as it can guide them on venues to avoid or venues where they have to take extra caution.
- 5. **Data Scientist and Public Health Scientists:** This report will be informative for upcoming data scientists, data scientists and public health professionals interested in learning how to solve problems with Data Science.

# 2.0 DATA ACQUISITION AND CLEANING

The goal of the data acquisition and cleaning is to have a data frame containing the LGAs, COVID-19 cases, Population, Latitude and Longitude values and location data.

## **2.1 DATA**

Based on the definition of the problem, the data needed for this project are:

- 1. Lagos State Local Government Area (LGA) Data: Lagos state is divided into twenty Local Government Areas (LGAs) which is synonymous to boroughs. Location analysis of Lagos state can only be carried out by sourcing data on the LGAs. This data contains the list of LGAs and their population.
- 2. **Distribution of COVID-19 cases across LGAs**: This contains the number of COVID-19 cases per LGA as at 18 April 2020. This is the available data on COVID-19 distribution across LGAs as at the time of this analysis. This data is needed to compare the number of COVID-19 cases across the LGAs. Location data of LGAs with the highest number of COVID-19 cases would also be explored.
- 3. **Geographical coordinates of the LGAs**: This data is needed to extract location data of all the LGAs needed for analysis.
- 4. Location data of various venues in the 20 LGAs in Lagos state: This data is needed to identify the top venues in the LGAs of Lagos state. This will help to pinpoint venues to strictly monitor as Lagos state economic activities resume.

# **2.2 DATA SOURCES**

The data required was sourced from the following:

- "Local government areas" from Wikipedia: <a href="https://en.wikipedia.org/wiki/Lagos">https://en.wikipedia.org/wiki/Lagos</a> State#Local government areas. This data provides the list of all the 20 LGAs and their population.
- Distribution of COVID-19 cases across LGAs as at 18 April 2020 from Wiki File: COVID-19 cases in the Lagos Metropolitan Area.jpg. <a href="https://commons.wikimedia.org/wiki/File:COVID-19 cases in the Lagos Metropolitan Area.jpg">https://commons.wikimedia.org/wiki/File:COVID-19 cases in the Lagos Metropolitan Area.jpg</a>. The number of cases per LGAs were manually extracted, as the file is in JPG format; this is also the only data available with traceable source on the number of cases per LGA as at the time of the analysis.

- Geopy: A Python 2 and 3 client for several popular geocoding web services. This will be used to get the geographical coordinates of the LGAs.
- Foursquare API: to get location data about the various venues in Lagos state by LGAs.

## 2.3 DATA CLEANING

# A. Scraping of Lagos state LGA data from Wikipedia

A table containing Lagos state LGAs and their population data was scraped from the Wikipedia page: <a href="https://en.wikipedia.org/wiki/Lagos">https://en.wikipedia.org/wiki/Lagos</a> State#Local government areas. This was achieved using the Wikipedia API and data was read into a Pandas data frame.

```
import urllib.request
!conda install lxml --yes
%pip install Wikipedia
import wikipedia as wp
```

Figure 3: Code snippet showing libraries installed to scrape the data

```
url = "https://en.wikipedia.org/wiki/Lagos_State#Local_government_areas"
page = urllib.request.urlopen(url)
html = wp.page("Lagos_State#Local_government_areas").html().encode("UTF-8")
df = pd.read_html(html, header =0)[2]
df
```

Figure 4: Code used to scrape data from Wikipedia

The data frame returned, consist of 3 columns (Rank, LGA and Population) and 21 rows containing values for these columns. The last row in the data frame represents the total population of Lagos state.

[2]:		Rank	LGA	Population
	0	1.0	Alimosho	5700714
	1	2.0	Ajeromi-Ifelodun	1746634
	2	3.0	Kosofe	665421
	3	4.0	Mushin	633543
	4	5.0	Oshodi-Isolo	10621789

Figure 5: Snippet of Data frame scraped from Wikipedia

The 'Rank' column was dropped from the data as it is not necessary for the analysis. The last row which contains the total population was also removed as it is also not necessary for the

analysis and might negatively affect it if left in the data frame. The types of the data values were checked, and the 'Population' column was found to contain string type data. Population data was changed to type integer using the astype() function.

The distribution of COVID-19 cases across the LGAs was added to the data frame by manual extraction from an image file from Wikipedia by creating a column containing a list of the number of cases.



Figure 6: Map image from Wikipedia showing distribution of COVID-19 cases across LGAs

```
[8]: COVID19_Cases = [15, 0, 11, 7, 7, 0, 4, 6, 1, 2, 8, 5, 107, 41, 85, 2, 2, 2, 0, 1] df['COVID19_Cases']= COVID19_Cases df
```

Figure 7: Creation of the COVID19\_Cases column

After creating and adding the column for COVID-19 cases, the data frame now had 3 columns namely: 'LGA', 'Population', and 'COVID19\_Cases'; and 20 rows containing their values.

[8]:		LGA	Population	COVID19_Cases
	0	Alimosho	5700714	15
	1	Ajeromi-Ifelodun	1746634	0
	2	Kosofe	665421	11
	3	Mushin	633543	7
	4	Oshodi-Isolo	10621789	7

Figure 8: Snippet of data frame showing LGA, population and COVID19 distribution

# **Extraction of Geographical Coordinates:**

The Nominatim function from geopy was used to extract the geographical coordinates of the LGAs. To achieve correct extraction, an 'Address' column was first created specifying that the LGAs are in Lagos. Thereafter, the code to extract the latitude and longitude values of each LGA and merge them with the current data frame was run. The data frame returned 9 columns (LGA, Population, COVID19\_Cases, Address, location, point, latitude, longitude and altitude).

	LGA	Population	COVID19_Cases	Address
0	Alimosho	5700714	15	Alimosho, Lagos
1	Ajeromi-Ifelodun	1746634	0	Ajeromi-Ifelodun, Lagos
2	Kosofe	665421	11	Kosofe, Lagos
3	Mushin	633543	7	Mushin, Lagos
4	Oshodi-Isolo	10621789	7	Oshodi-Isolo, Lagos

Figure 9: Snippet of data frame with 'Address' column

```
from geopy.extra.rate_limiter import RateLimiter
geolocator = Nominatim(user_agent="lag_explorer")

# 1 - convenient function to delay between geocoding calls
geocode = RateLimiter(geolocator.geocode, min_delay_seconds=1)
# 2- - create location column
df['location'] = df['Address'].apply(geocode)
# 3 - create longitude, laatitude and altitude from location column (returns tuple)
df['point'] = df['location'].apply(lambda loc: tuple(loc.point) if loc else None)
# 4 - split point column into latitude, longitude and altitude columns
df[['latitude', 'longitude', 'altitude']] = pd.DataFrame(df['point'].tolist(), index=df.index)
df
```

Figure 10: Code used to extract geographical coordinates

]:	LGA	Population	COVID19_Cases	Address	location	point	latitude	longitude	altitude
	) Alimosho	5700714	15	Alimosho, Lagos	(Alimosho, Lagos, Nigeria, (6.58434295, 3.2576	(6.58434295, 3.257631097232365, 0.0)	6.584343	3.257631	0.0
	Ajeromi-Ifelodur	1746634	0	Ajeromi-Ifelodun, Lagos	(Ajeromi/Ifelodun, Lagos, Nigeria, (6.45512169	(6.455121699999999, 3.3359462518298724, 0.0)	6.455122	3.335946	0.0
	. Kosofe	665421	11	Kosofe,Lagos	(Kosofe, Lagos, 100234, Nigeria, (6.5819740000	(6.58197400000001, 3.4148359416178415, 0.0)	6.581974	3.414836	0.0
	Mushir	633543	7	Mushin, Lagos	(Mushin, Lagos, Nigeria, (6.5326324499999995,	(6.5326324499999995, 3.3520219717987674, 0.0)	6.532632	3.352022	0.0
	Oshodi-Isolo	10621789	7	Oshodi-Isolo, Lagos	(Oshodi/Isolo, Lagos, 100263, Nigeria, (6.5400	(6.540010000000005, 3.312414613183944, 0.0)	6.540010	3.312415	0.0

Figure 11: Snippet of data frame returned by geopy containing geographical coordinates

The columns for address, location and point were dropped as they are repeated variants of the LGA, and latitude-longitude values. The altitude column was also dropped as it is not necessary for the analysis.



Figure 12:New data frame containing geographical coordinates

# Foursquare data:

Using the geographical coordinates of the LGAs, location data was extracted using the Foursquare API. The goal was to use the Foursquare API to extract information about the top ten most common venues in the LGAs. The top 3 LGAs based on the number of COVID-19 cases was first explored and data on their top venues and categories were extracted. The next step was to extract all the venues in the LGAs along with their categories and geographical coordinates.



Figure 13:Top venues in Lagos Mainland with the highest number of COVID-19 cases as at 18/04/2020



Figure 14:Top venues in Eti-Osa LGA with the 2nd highest number of COVID-19 cases as at 18/04/2020

## Location Analysis of Lagos State in Reference to the COVID-19 Pandemic

:	name	categories	lat	Ing
0	University of Suya	BBQ Joint	6.606512	3.349370
1	Best Western Plus Lagos Ikeja Hotel	Hotel	6.597282	3.354239
2	Bakers' World	Bakery	6.600560	3.355194
3	Goat Hunters	African Restaurant	6.601134	3.351368
4	K.F.C.	Fast Food Restaurant	6.600272	3.355484

Figure 15: Top venues in Ikeja LGA with the 3rd highest number of COVID-19 cases as at 18/04/2020

i]:	<pre>print(Lagos_venues.shape) Lagos_venues.head()</pre>							
	(105, 7)							
i]:		LGA	LGA Latitude	LGA Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
	0	Kosofe	6.581974	3.414836	Cecilia's SteakHouse	6.583084	3.418033	Steakhouse
	1	Mushin	6.532632	3.352022	Mushin market	6.527949	3.353511	Flea Market
	2	Mushin	6.532632	3.352022	olorunsogo bus stop	6.536227	3.352369	Bus Station
	3	Mushin	6.532632	3.352022	KITS Technlogies	6.534660	3.346489	IT Services
	4	Mushin	6.532632	3.352022	CARWASH	6.539191	3.356523	Rental Car Location

Now that we have data on the venues in the LGAs, let's check the number of venues returned by each of them.

Figure 16: Snippet of venues in Lagos state returned by Foursquare

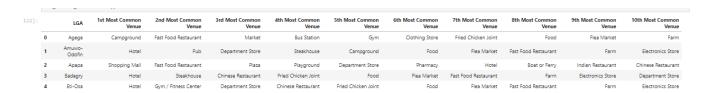


Figure 17: Snippet of Dataframe with Top 10 most common venues in each LGA

The Foursquare data on the most common venues in the LGA would be used to identify the LGAs to strictly monitor as economic activities in Lagos state reopens. This data would also be clustered to identify defining features that can be used to develop strategies based on the location data. The information on the most common venues would be the basis for identifying strategies that can be used to curb the spread of COVID19, these strategies would be highlighted in the 'Recommendations' section of this report. Codes used to generate the Foursquare data can be found in the Methodology section.

# 3.0 METHODOLOGY

The focus of this project is to identify the most common venues in the LGAs in Lagos state. This would enable us to determine the venues that require focused anti-COVID 19 measures. I have collected the data on Lagos state Local Government Areas (LGAs), their COVID-19 case distribution and geographical coordinates. I would now use the Foursquare API to get the location and type (category) of top venues in the LGAs. Afterwards, the venues will be segmented and clustered to identify defining features that would aid the development of strategies to curb the COVID19 spread. However, let's start by exploring and visualizing the data

# **3.1 EXPLORATORY DATA ANALYSIS**

# Map of Lagos State using Folium library

Folium library and coordinate data were used to draw an interactive leaflet map of Lagos State with the LGAs superimposed on it.

```
geolocator = Nominatim(user_agent="lag_explorer")
location = geolocator.geocode(address_)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinates of Lagos State are {}, {}.'.format(latitude, longitude))
The geograpical coordinates of Lagos State are 6.4550575, 3.3941795.
]: map_Lagos = folium.Map(location=[latitude, longitude], zoom_start=10)
# the codes below are to add markers to map
for lat, lng, LGA, COVID19_Cases in zip(df['latitude'], df['longitude'], df['LGA'], df['COVID19_Cases']):
label = '()',{}'.format(LGA, COVID19_Cases)
label = folium.Popup(label, parse_html=True)
folium.CircleMarker(
    [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_olor='#3186cc',
        fill_olor='#3186cc',
        fill_olor='#3186cc',
        fill_olor='#3186cc',
        fill_olor=Lagos)
map_Lagos
```

Code used to draw Leaflet Map of Lagos state



Leaflet Map of Lagos state

# **Data Distribution**

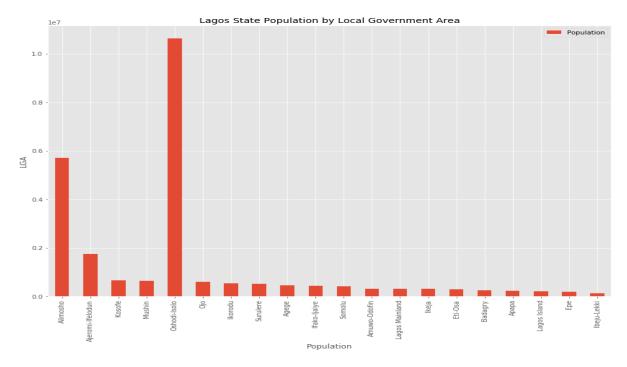
The distribution of the data was visualized using Bar charts. Starting with the population distribution, a bar plot was used to compare the population data of the LGAs. First, a copy of the data was created for plotting to avoid making changes to the original data, then LGA was set as the index. The data for LGA and population was then extracted from the data frame.



Figure 20: Snippet of LGA ad Population Data

```
[20]: df_plot_pop.plot(kind='bar', figsize=(14, 10))
plt.xlabel('Population') # add to x-label to the plot
plt.ylabel('LGA') # add y-label to the plot
plt.title('Lagos State Population by Local Government Area') # add title to the plot
plt.show()
```

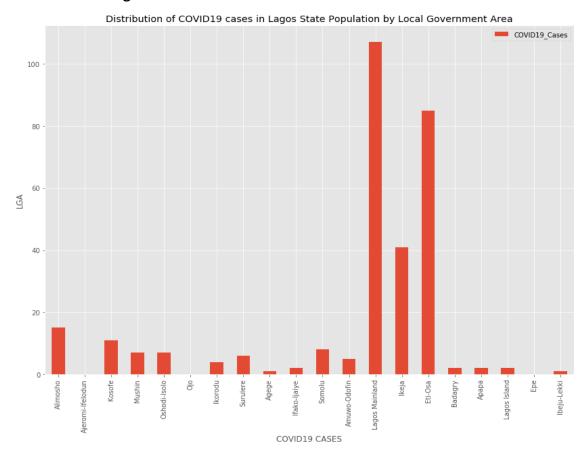
Code for Bar plot of LGA and Population



Bar Chart showing Population distribution

The bar chart shows that Oshodi-Isolo LGA is the most populated followed by Alimosho and then Ajeromi-Ifelodun LGA.

# Bar chart showing COVID19 case distribution across LGAs:



Bar Chart showing COVID19 Cases distribution

The bar chart shows that Lagos Mainland has the highest number of COVID-19 cases followed by Eti-Osa local Government and Ikeja LGA as at 18 April 2020.

It can be said that the population of the LGAs is not related to the number of COVID-19 cases just by comparing the charts; this was confirmed by plotting a regression plot.

Seaborn Regression plot of the relationship between the population and the number of COVID-19 cases

```
plt.figure(figsize=(14, 10))
sns.set(font_scale=1.5)

ax = sns.regplot(x='COVID19_Cases', y='Population', data=df_plot, marker='+', scatter_kws={'s': 200})
ax.set(xlabel='COVID-19 CASES', ylabel='POPULATION')
ax.set_title('Relationship between Population and Number of COVID-19 cases in Lagos')
```

Figure 24: Seaborn regression plot code

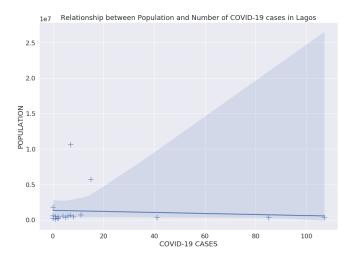


Figure 24: Seaborn Regression plot of the relationship between population and COVID19 cases

The regression plot shows that there is no linear relationship between population and COVID-19 cases. This means that the sharp difference between the number of COVID-19 cases for the three LGAs with the highest number of cases and the other LGAs must be because of other factors.

# 3.2 LOCATION ANALYSIS USING FOURSQUARE API

# Exploration of location data of LGAs with the highest COVID-19 Cases.

The LGAs with the first, second and third highest number of COVID 19 cases were explored using the Foursquare API. These LGAs were first confirmed by sorting the dataframe.



Sorted list of LGAs by COVID19 Cases in descending order

The sorted list of LGAs by the number of COVID 19 cases shows that Lagos Mainland, Eti-Osa and Ikeja LGAs have the first, second and third highest number of COVID 19 cases respectively. The location data of these LGAs were then generated using the Foursquare API.

Let's get the top 20 venues that are in Lagos Mainland within a radius of 1000 meters.

Code snippet for retrieving location data for Lagos Mainland

Data frame snippets showing the top venues in these LGAs are in the 'Data' section of this report.

## **Exploration of Location Data of all LGAs**

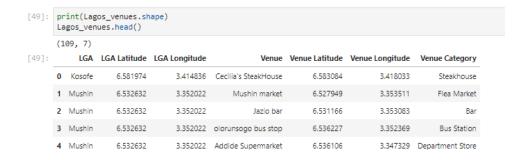
To do this, the location data of the LGAs were extracted using the Foursquare API and merged with the LGA data frame. Codes used to extract the data are shown below:

```
def getNearbyVenues(names, latitudes, longitudes):
  venues_list=[]
  for name, lat, lng in zip(names, latitudes, longitudes):
    print(name)
    # API request URL
https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&l
imit={}'.format(
      CLIENT_ID,
      CLIENT_SECRET,
      VERSION,
      lat,
      Ing,
      radius,
      LIMIT)
    # GET request
    results = requests.get(url).json()["response"]['groups'][0]['items']
# return only relevant information for each nearby venue
    venues_list.append([(
      name,
      lat,
      Ing,
      v['venue']['name'],
      v['venue']['location']['lat'],
      v['venue']['location']['lng'],
      v['venue']['categories'][0]['name']) for v in results])
  nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
  nearby_venues.columns = ['LGA',
         'LGA Latitude',
         'LGA Longitude',
         'Venue',
         'Venue Latitude',
         'Venue Longitude',
         'Venue Category']
  return(nearby_venues)
```

Code snippet on the extraction of location data for Lagos

Foursquare returned 105 venues. The first five rows of the data frame containing the venues, categories and their geographical coordinates are shown below:

#### Location Analysis of Lagos State in Reference to the COVID-19 Pandemic



Snippet of dataframe containing location data of the LGAs

To make the data easier to analyse, I used one hot coding to manipulate the data. I then grouped the rows by LGA by taking the mean of the frequency of occurrence of each category. See codes below:

```
l]: # One Hot Coding
Lagos_onehot = pd.get_dummies(Lagos_venues[['Venue Category']], prefix="", prefix_sep="")

# add LGA column back to dataframe
Lagos_onehot['LGA'] = Lagos_venues['LGA']

# move LGA column to the first column
fixed_columns = [Lagos_onehot.columns[-1]] + list(Lagos_onehot.columns[:-1])
Lagos_onehot = Lagos_onehot[fixed_columns]

#Group rows by LGA by taking the mean of the frequency of occurrence of each category
Lagos_grouped = Lagos_onehot.groupby('LGA').mean().reset_index()
Lagos_grouped
```

Pandas One hot coding for the venue categories

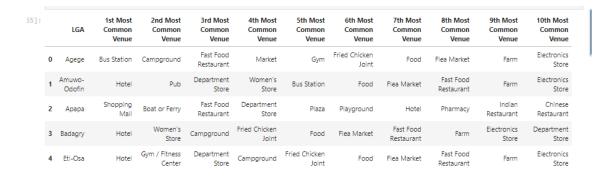


Dataframe showing one hot coding of venue categories

The next step was to generate a data frame of the top 10 venues sorted in descending order.

```
def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
row_categories_sorted = row_categories.sort_values(ascending=False)
    return row_categories_sorted.index.values[0:num_top_venues]
num_top_venues = 10
indicators = ['st', 'nd', 'rd']
# create columns according to number of top venues
for ind in np.arange(num top venues):
    try:
        {\tt columns.append('\{\}\{\}~Most~Common~Venue'.format(ind+1,~indicators[ind]))}
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))
# create a new dataframe
LGA_venues_sorted = pd.DataFrame(columns=columns)
LGA_venues_sorted['LGA'] = Lagos_grouped['LGA']
for ind in np.arange(Lagos_grouped.shape[0]):
    LGA_venues_sorted.iloc[ind, 1:] = return_most_common_venues(Lagos_grouped.iloc[ind, :], num_top_venues)
LGA_venues_sorted.head()
```

Code to generate the top ten venues in each LGA in descending order



Dataframe of the LGAs and their top ten most common venues

I can see from the data frame the top most common venues in each LGA; with this data, we can tell that the most common venues in the first 5 LGAs include hotels, restaurants, campgrounds, shopping malls, flea markets, markets, Bus stations, stores. This can help identify which venues to focus strict social distancing measures on as the Lagos lockdown eases.

# 3.3 K-MEANS CLUSTERING

The next step was to cluster the LGAs into 4 clusters.

```
kclusters = 4
Lagos_grouped_clustering = Lagos_grouped.drop('LGA', 1)
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Lagos_grouped_clustering)
kmeans.labels_[0:10]
```

K-Means clustering

I created a new data frame that includes the cluster labels as well as the top 10 venues for each LGA. This data frame was merged with the Lagos\_grouped data and the main LGA data frame to add the latitude and longitude values.



Snippet of Complete Dataframe with location data of the top venues in Lagos

# Visualisation of Clusters on Lagos state map:

Using the Foursquare location data frame with the cluster labels and Folium library, I visualised the clusters on the Lagos state map using the codes below:

```
Lagos_merged = Lagos_merged[Lagos_merged['Cluster Labels'].notna()]
Lagos_merged['Cluster Labels'] = Lagos_merged['Cluster Labels'].astype("int64")
Lagos_merged['Cluster Labels'].astype(np.int64)
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 \text{ for } i \text{ in } range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]
markers_colors = []
for lat, lon, poi, cluster in zip(Lagos_merged['latitude'], Lagos_merged['longitude'], Lagos_merged['LGA'], Lagos_merged['Cluster Labe label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
         [lat, lon],
         radius=5,
         popup=label,
color=rainbow[cluster-1],
         fill=True,
         fill_color=rainbow[cluster-1],
         fill opacity=0.7).add to(map clusters)
map_clusters
```

Code to visualise K means clustering



Leaflet map of Lagos showing cluster labels

The next step was to generate the results which include the data frames containing the top ten venues of each cluster. These results would be used to develop strategies/recommendations that can help curb the spread of COVID 19 at these venues.

```
: Lagos_merged.loc[Lagos_merged['Cluster Labels'] == 0, Lagos_merged.columns[[1] + list(range(4, Lagos_merged.shape[1]))]]
```

Code template used to generate cluster data frames

#### **CLUSTERS:**

## Snapshot of the top ten most common venues in cluster1:



Top ten venues in LGAS in Cluster 1 (please zoom in to view venues)

## Snapshot of the top ten most common venues in Cluster 2:



Top ten venues in LGAs in cluster 2(Please zoom in to view venues)

# Snapshot of the top ten most common venues in Cluster 3:



Top ten most common venues in LGA in Cluster 3

## Snapshot of the top ten most common venues in Cluster 4:



:Top ten venues of LGA in Cluster 4(Please zoom in to view venues)

### **CLUSTER INFORMATION**

The following lines of codes were used to generate details of each cluster including, the LGAs in each cluster and their shapes(numbers).

```
cluster1 = Lagos_merged[Lagos_merged['Cluster Labels'] == 0]
cluster2 = Lagos_merged[Lagos_merged['Cluster Labels'] == 1]
cluster3 = Lagos_merged[Lagos_merged['Cluster Labels'] == 2]
cluster4 = Lagos_merged[Lagos_merged['Cluster Labels'] == 3]

for i in range(4):
    x = Lagos_merged[Lagos_merged['Cluster Labels'] == i]
    print('cluster'+str(i+1) + ' shape is {}'.format(x.shape))
```

Code to generate information about clusters.

## Code output showing the size of clusters and their LGAs:

```
cluster1 shape is (11, 16)
  cluster2 shape is (4, 16)
  cluster3 shape is (1, 16)
  cluster4 shape is (1, 16)

d]: cluster1_LGAs = list(cluster1['LGA'])
  print(cluster1_LGAs)
  ['Mushin', 'Oshodi-Isolo', 'Ikorodu', 'Surulere', 'Agege', 'Ifako-Ijaiye', 'Somolu', 'Lagos Mainland', 'Ikeja', 'Apapa', 'Lagos Island']

5]: cluster2_LGAs = list(cluster2['LGA'])
  print(cluster2_LGAs)
  ['Ojo', 'Amuwo-Odofin', 'Eti-Osa', 'Badagry']

6]: cluster3_LGA = list(cluster3['LGA'])
  print(cluster3_LGA)
  ['Kosofe']

7]: cluster4_LGA = list(cluster4['LGA'])
  print(cluster4_LGA)
  ['Ibeju-Lekki']
```

Shape of Clusters and their LGAs

# 4.0 RESULTS

The results of this project would be explained based on each cluster.

#### Cluster1

Cluster 1 contains 11 LGAs. LGAs in these clusters include Mushin, Oshodi-Isolo, Ikorodu, Surulere, Agege, Ifako-Ijaiye, Somolu, Lagos Mainland, Ikeja, Apapa, and Lagos Island.

Most common venues in Cluster 1 mainly include Flea markets, Shopping malls, Markets, Stores, Food businesses (restaurants, joints, bakeries), Park, Boat/ferry, Bars, Campgrounds, Hotels, Bus stations, Movie theatre, Gym, Pharmacies, Soccer field, Baseball stadium, Rentals, and Farms. Most of these locations are known to generate high foot traffic.

It is also important to note that this cluster contains Lagos mainland and Ikeja which are the LGAs with the highest and 3rd highest number of COVID-19 cases respectively as at 18 April 2020.

## Cluster 2

Cluster 2 contains 4 LGAs: Ojo, Amuwo-Odofin, Eti-Osa, and Badagry. The most common venues in this cluster are mainly: Hotels, Gyms, Stores, Flea markets, Campgrounds, Farms and Restaurants. Hotels are the number one most common venues in 3 of the 4 LGAs in this cluster.

This cluster also contains the LGA with the 2nd highest number of COVID-19 cases (Eti-Osa). It is important to note that Eti-Osa is known to be the home to Lagos elite residents and this can be linked to the travel history of the residents. Further analysis can be carried out based on the demographic data of each LGA. However, it is still important to watch other LGAs in this cluster.

## Cluster 3

Cluster 3 contains only one LGA which is Kosofe LGA. The most common venues in Kosofe with 11 cases are mainly: Food businesses (Steakhouse, chicken joint, restaurants), Campground, Gyms, Flea market, Farm and Stores.

## Cluster 4

Cluster 4 contains only one LGA which is Ibeju-Lekki LGA. Ibeju Lekki had only 1 COVID-19 case as at 18 April 2020. The most common venues in Ibeju Lekki in include: Farm, Stores, Food businesses(steakhouse, chicken joint, restaurants), Flea Market and Gyms, and Stores. This is similar to that of cluster 3.

The only difference between the top ten venues in cluster 3 (Kosofe LGA) and cluster 4(Ibeju-Lekki LGA) is Campground which is one of the top venues in Kosofe but absent in the latter. This could be the reason why Kosofe has 11 cases and Ibeju-Lekki has only 1, as the 2 clusters are similar in terms of common venues. Ibeju Lekki is known to be mainly residential with fewer business activities in comparison to other LGAs like Lagos mainland and Ikeja.

# 5.0 DISCUSSION AND RECOMMENDATIONS

# 5. 1 DISCUSSION

Cluster 1 which appears to be a highly social cluster. It is also home to the LGAs with the highest and third-highest number of COVID19 cases as at 18 April 2020. All common venues in this cluster should be put under strict COVID-19 watch, else other LGAs in the cluster might experience a surge in the number of cases. In as much as the state needs to ease down on lockdown rules and reopen for the sake of its economy, some of these venues especially those that are non-essential, might need to remain closed till the spread of the disease reduces.

Campgrounds were differentiating features between Kosofe LGA with 11 COVID 19 cases and Ibeju-Lekki with 1 case. This venue was also featured in all clusters except cluster 4 (containing only Ibeju-Lekki). Besides, it is the fourth most common venue of Eti-Osa LGA with the second-highest COVID 19 cases and a defining factor for cluster 2, as it is a common venue in 3 out of 4 LGAs in the cluster. Certain campgrounds in Lagos are associated with large crowds usually for religious retreats. This analysis shows that it is a red spot at this time and should remain closed till COVID 19 is eliminated in Lagos.

Hotels were major defining factors for cluster 2 as it is the number one most common venue in 3 out of 4 LGAs in the cluster. This might explain why Eti -Osa was assigned to this cluster, as the perceived socio-demographic peculiarities of its residents differs from other LGAs in this cluster. Considering that Eti-Osa is the LGA with the second-highest number of COVID 19 cases, all other LGAs in this cluster must be carefully monitored.

Flea markets were featured as top venues in all clusters, thus measures must be taken to prevent the spread of the virus at these venues. Food businesses were also featured in all the clusters. The government and food business owners must adopt measures to protect staff and customers. Gyms were featured in 3 of the 4 clusters; they are non-essential and should

remain shut at this time. Farms were featured in all clusters and just like food businesses, they must adopt measures to protect their staff and consumers. Below are recommendations based on the most common venue categories in the clusters.

#### **5.2 RECOMMENDATIONS**

## Food Businesses (Restaurants, joints, steakhouses):

Outdoor joints such as BBQ joints, steakhouses, chicken joints can not afford to operate as before for now. These places are usually characterised with people standing in wait to get food, and usually found in busy areas of Lagos. Owners need to adopt new sale strategies based on the current situation. These strategies may include: movement of cooking/roasting indoors, online /digital sale of products, and making food available only as take-outs. Strict social distancing measures must be put in place during pickups.

Restaurants should only do take-outs while they observe strict safety measures among staff and during food preparation. Restaurant owners are advised to take advantage of the digital space to deliver foods to customers in their homes. Strict social distancing measures should be put in place during pickups.

Food businesses, in general, must take all safety precautions including frequent disinfection of environment, use of nose masks, sanitizer and gloves.

#### Flea Markets:

Flea markets are characterized by crowds and a lot of contact with second-hand products. This is an unsafe place to be at this time. Unfortunately, it is seen as one of the most common venues across all clusters. It is recommended that lockdown measures still apply to flea markets until the situation gets better. Relief should be provided for sellers to encourage adherence and enable them to survive this period.

#### Markets:

Markets are major economic points for Lagos state. However, their peculiarities might make them unsafe at this time. To manage the COVID-19 spread in markets while protecting the economy, they have to be segmented first based on the essentiality of products and foot traffic.

Markets with shopping plaza structures can still be managed by ensuring that strict social distancing measures are put in place. This can be done with the use of painted lines to ensure a safe distance between customers. Only a minimal number of compulsory staff should be allowed in the shops; the shop space should be considered in determining this number. Testings that require body contact e.g cloth testing should be forbidden for now.

This is the time to review the structures of crowded markets such as the popular Balogun market. The current structure and footfall of the market make it a dangerous place to be currently. Such markets should not be open until they are properly decontaminated and structured to accommodate social distancing measures.

Food markets can not be closed because they are essential, unfortunately, they are usually crowded. They also need to be decontaminated and structured in a way that will allow effective social distancing measures. Until this structuring is achieved, the opening of stalls should be rotated to decongest the markets. Curfews should also be employed and sellers should be well educated on how to protect themselves and their customers.

I recommend that a special anti-COVID-19 taskforce be assigned to food markets to ensure adherence to safety measures. Sellers should be provided with all the materials they need for free. These materials include face masks, hand sanitizers, disinfectants and every other item needed to protect themselves and their customers.

## Farms:

We need the farms at this time to produce the food we need. Farmers need to be well educated on how to safely get their products to buyers/consumers without putting any party at risk.

#### **Bus Stations:**

As economic activities resume, people would need transportation to get to work. Unfortunately, bus stations in Lagos state are characterized by large crowds, rush and 'pushing'. Bus stations are red spots at this time. This is another venue category that requires a special anti-COVID-19 task force. Companies/businesses must devise strategies to safely transport their staff to and fro work to avoid bus stations.

Strict policies must be put in place at every bus station. Painted box lines must be used to ensure a safe distance between people as they stand in wait for buses. Social distancing measures must be applied in every bus and these buses must be disinfected after every trip. Buses with no ventilation should not be allowed to move. Everyone must use face masks and gloves. Thermometers should be used to check passengers before entry. These measures must be enforced and strictly monitored.

## Stores(Departmental stores, Electronic stores, Boutiques, etc) and Salons:

Only the required number of staff should be allowed in these venues. Painted lines should be used to ensure customers adhere to social distancing measures. For salons, equipment should be frequently sterilised and sits for customers must be 6 feet apart. Thermometers should be used to check customers before entry. Face masks, gloves and sanitizers should be used by both customers and staff.

## Malls:

Malls are also known to generate a high level of foot traffic. Thermometers should be used to check customers before entry and they must be wearing face masks and gloves. Sanitisers should be used and the malls most he disinfected regularly. There must be proper ventilation and the use of air-conditioning must be discouraged. Same measures apply to staff. Only required staff should come to work. Painted box lines should be used to ensure customers adhere to social distancing measures.

#### **Hotels:**

Strict guidelines must be provided to hotels to prevent COVID-19 spread between customers and staff. Regular disinfection should be put in place and suspicions reported to the NCDC.

# Non-Essential Businesses/Facilities: Gyms, lounges, Campgrounds, etc.

These businesses/facilities are known for high foot traffic but are non-essential at this time, it is best they remain closed till the curve is flattened.

# **6.0 CONCLUSION**

This project aimed to develop strategies based on location data analysis, to reduce the spread of COVID-19 virus in Lagos state as economic activities resume. The distribution of COVID-19 cases across the LGAs and their location data were explored and analysed. The most common venues in these LGAs were found to be flea markets, restaurants and joints, malls, gyms, lounges, steakhouses, campgrounds and related facilities. Across all LGA clusters, flea markets and Food businesses were shown as consistent most common venues showing the need to focus on them.

Organisations and businesses must be innovative and embrace the 'new normal'. Companies must adopt working from home strategies. This is the time to take advantage of the digital space. The government and investors should also look at the possibility of digitalizing markets.

This is the time for investors to look into the transportation sector. There is a very limited amount of transport facilities available to citizens in comparison to the population.

Residents should also embrace the "new normal" and avoid these venues or apply extra caution if they have to visit them. Everyone should get used to ordering products online.

It is important to note that asides the venues generated by this analysis, other locations can serve as red spots such as hospitals, overcrowded communities, slums, etc. All these areas must also be considered in the fight against the virus.

The fight against COVID-19 is not just for the NCDC and Lagos state government, everyone must work together to win this fight to save public health and the economy.

In conclusion, Lagos State can win the war against COVID-19 without crumbling the economy if there are strict enforcement and adherence to safety measures.

# **Further Analysis:**

This analysis focuses on the top venues in Lagos state. Further analysis can be carried out to explore communities based on their socio-demographic characteristics or population density.

Please note: Data on COVID 19 case distribution is based on information as at 18 April 2020, which is the last available data at the time of this analysis.

## **REFERENCES**

- 1. World Health Organisation (WHO). 2020. Q & A on coronaviruses (COVID-19). Available at: <a href="https://www.who.int/news-room/q-a-detail/q-a-coronaviruses">https://www.who.int/news-room/q-a-detail/q-a-coronaviruses</a>.
- 2. World Health Organisation (WHO). 2020. Coronavirus disease 2019 (COVID-19) Situation Report-98. Available at: <a href="https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200427-sitrep-98-concl-19.pdf?sfvrsn=90323472">https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200427-sitrep-98-concl-19.pdf?sfvrsn=90323472</a> 4.
- 3. Wikipedia. 2020. 2020 coronavirus pandemic in Nigeria. Available at: <a href="https://en.wikipedia.org/wiki/2020">https://en.wikipedia.org/wiki/2020</a> coronavirus pandemic in Nigeria.
- 4. Nigeria Centre for Disease and Control (NCDC). 2020. CORONAVIRUS (COVID-19) HIGHLIGHTS. Available at: <a href="https://covid19.ncdc.gov.ng/">https://covid19.ncdc.gov.ng/</a>.
- Nigeria Centre for Disease and Control (NCDC). 2020. COVID-19 SITUATION REPORT.
   Available
   https://ncdc.gov.ng/themes/common/files/sitreps/d127b988882b6c7fb93335aa6df
   3bacb.pdf
- 6. Wikipedia. 2020. Lagos State. Available at: <a href="https://en.wikipedia.org/wiki/Lagos">https://en.wikipedia.org/wiki/Lagos</a> State
- 7. Ekundayo, J.M. 2013. *Out of Africa:* Fashola: Reinventing Servant Leadership to Engender Nigeria's Transformation. AuthorHouse. P.135. ISBN 9781481790406