REAL ESTATE ANALYSIS IN LAGOS STATE

(Determining the best areas for Investing)

-By J. EDOH

Table of contents

Business Problem

In this project, We will be working on ways to optimize investing in the real estate sector in Nigeria (Lagos State), which has become increasingly popular over the last couple of years and has become a common investment thread.

The Real Estate sector is currently booming in Nigeria and people are making millions in it, irrespective of the economic downturn, it can be seen as a real money spinner.

We will be using the Power of Data Science to help the investor determine the areas where acquiring the properties will be beneficial or profitable while taking note of the kind of investment the investor requires, be it rentals, property development or land flipping.

Advantages of each area will then be clearly expressed so that best possible final location can be chosen by the investor.

Data

For this problem to be tackled and as every other problem is tackled in data science, we require data. One of the challenges of this project is that the report is done on Nigeria. A country where there is very little housing data for the preexisting areas, only for the new structures that have been created as a result of the resurgence of real estate in Nigeria.

So based on definition of our problem,

Factors that will help in making a good decision are:

> The location of the property(in our case the area)
>
> How close Is it close to restaurants, schools, shops, public transportation, and other neighborhood amenities?
>
> Is it in a sought-after tourist destination? The more popular and accessible an investment location is, the more valuable the property will be.
>
> The push and pull of supply and demand on that area. If there is a high demand but fixed supply, the price of properties will rise as more people attempt to buy.

The following data sources will be needed to extract/generate the required information:

The Information concerning the housing data will be scraped from the websites **https://nigeriapropertycentre.com/market-trends**

1. This contains information about the prices and demand trends on each area in question.

There are four categories for the pricing

pricing for house sales

pricing for flat sales

pricing for house rentals

pricing for flat rentals

***NOTE**: Due to lack of data only a few areas in Lagos state was used for the project*

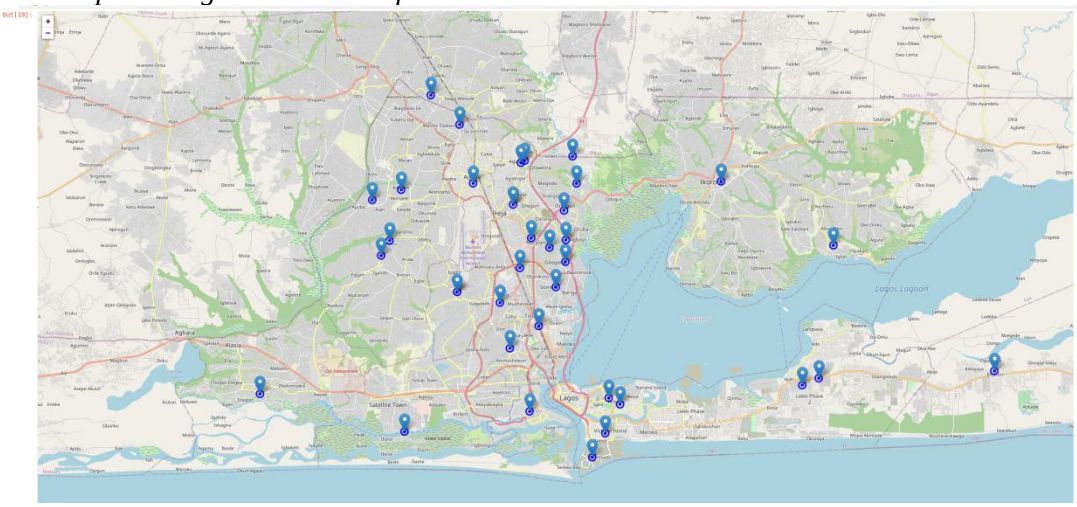
2. The geographical data of the areas will be gotten by using Google's map API.

3. The Number of restaurants, schools, shops, public transportation, and other neighborhood amenities and their types and location in every area will be obtained using **FOURSQUARE API.**

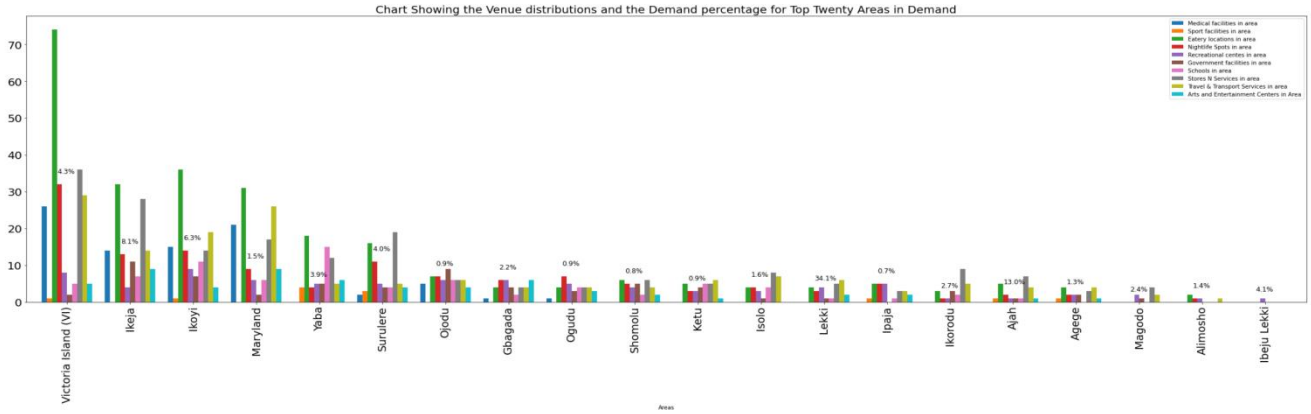
Methodology & Analysis

In this project we will direct our efforts on areas that have high number of values. The reason for this is because of Real Estate Appreciation.
Appreciation is an important terms in real estate. In simple words, appreciation is the increase of investment properties' value with time. Meaning, the longer you hold an investment property, the more value it'll gain and, thus, the more profits you'll make if you decide to sell it. As a result, appreciation determines the profitability of real estate investments over the long-term.
In order to find investment properties with a high appreciation rate, we need to look for areas where land is becoming limited, infrastructure development projects are being planned, are major business hubs, tourist destinations, and are expecting major economic growth. So It is safe to say that areas with more venues are profitable.
In first step we have collected the required data: location type (category) demand rates of every venue within 1km from each Area (according to Foursquare categorization).

Below is a map showing the locations of the areas



Below is a chart showing the demand rates for each area and the venues that can be found in each area.



The second step in our analysis will be exploration of each area, What we need to determine from our data which areas will be profitable to our investors.
So after perusing over the internet, According to a Forbes publication, the One percent rule one of the ways of determining a profitable land or housing investments.
But due to the fact that our data is a compilation of many prices(i.e the average of the prices in that area), after calculating the rent per sales price rate all areas failed the One percent rule.
So I decided to change the quota after studying the data from One percent to 0.25%, hence eliminating some few areas.

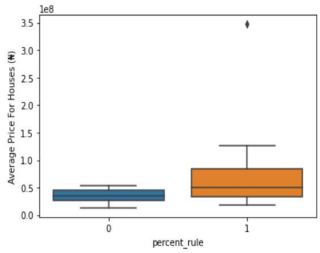
```
In [55]: calcs_df
```

Out[55]:

	Areas	Average Price For Houses (₦)	Average Price For House Rentals (₦)	Total Property Count For Houses	House Rentals Monthly (₦)	rent/price	percent_rule
0	Ikoyi	347170000	10820000	2167	901667.0	0.26	1
1	Victoria Island (VI)	125550000	5920000	831	493333.0	0.39	1
2	Magodo	103310000	3420000	931	285000.0	0.28	1
3	Apapa	102290000	4500000	24	375000.0	0.37	1
4	Eko Atlantic City	99650000	5000000	8	416667.0	0.42	1
5	Lagos Island	99550000	5930000	20	494167.0	0.50	1
6	Ikeja	91850000	4380000	1638	365000.0	0.40	1
7	Ilupeju	80350000	3760000	64	313333.0	0.39	1
8	Ogudu	74410000	2750000	143	229167.0	0.31	1
9	Lekki	71190000	3570000	27382	297500.0	0.42	1
10	Maryland	71130000	3290000	233	274167.0	0.39	1

```
In [56]: sns.boxplot(x="percent_rule", y="Average Price For Houses (₦)", data=calcs_df)
```

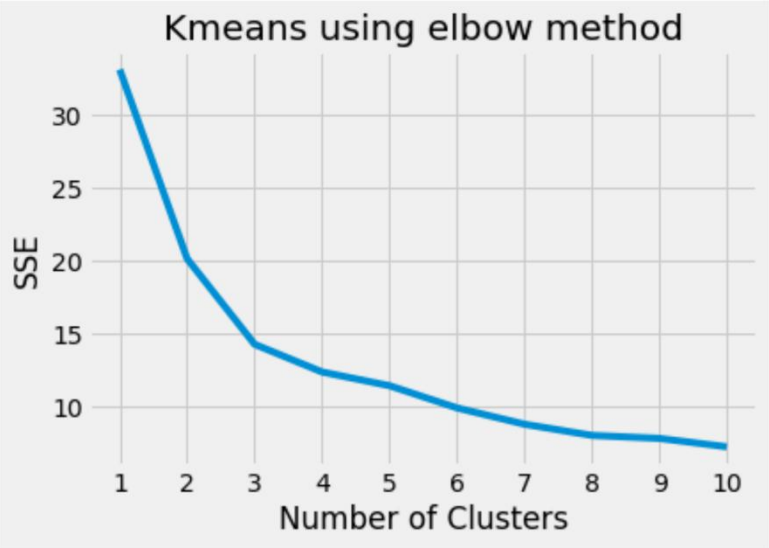
```
Out[56]: <matplotlib.axes._subplots.AxesSubplot at 0x7fbcdd932590>
```



So areas that passed this new rule 0.25% rule are deemed to be somewhat profitable to out investors.

In third and final step we focused on most promising areas and created clusters (**using k-means clustering**) of those locations to identify the optimal area location for stakeholders.

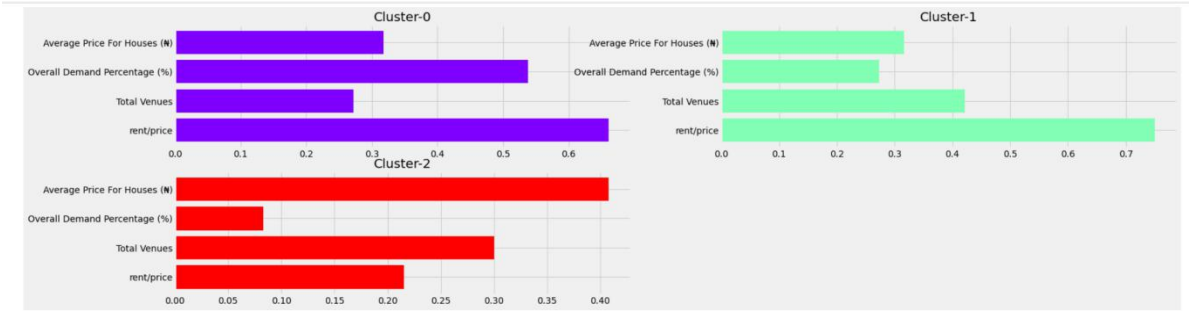
After calculating for the optimal k using the elbow method,



the k gotten was 3, and so the data was clustered into three segments. Cluster0, Cluster1, Cluster2

Result Analysis

The Clustering was based on Latitude, Longitude, obtained Venues from Foursquare API, rent to price rate, demand percentage of each areas. After doing all the necessary steps, I created a visualization that showed the mean of values in the each cluster and how they performed against the requirements.

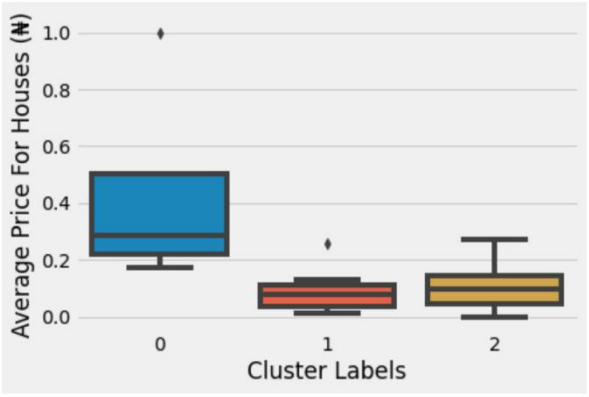


Well it is a bit hard to make a decision, but we can clearly say that the areas in Cluster 0 is out performing others.

So after looking deeper with the aid of box-plots we check how the clusters performed against each criteria.

The first one is the clusters versus the Average Price

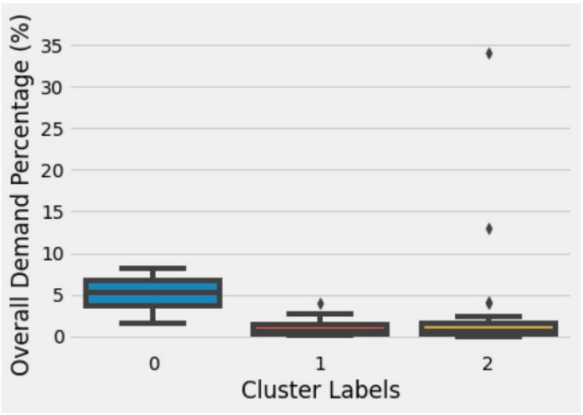
```
Out[86]: <matplotlib.axes._subplots.AxesSubplot at 0x7fbcd6f72150>
```



Areas in cluster one have better selling prices with cluster 2 following in behind and cluster 1 in last position.

The next one is the cluster versus the demand percentage.

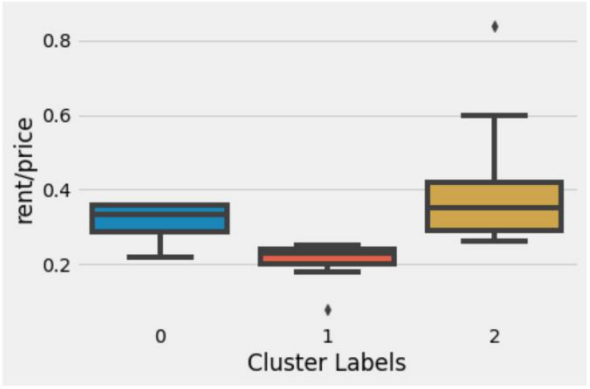
```
Out[88]: <matplotlib.axes._subplots.AxesSubplot at 0x7fbcd6f9a050>
```



Like our last report, areas in cluster 0 performs very well with a higher demand percentage range when compared to the other clusters, and cluster 2 following in behind and cluster 1 containing areas with lower demand percentage.

The next one is the cluster versus the rent to price ratio.

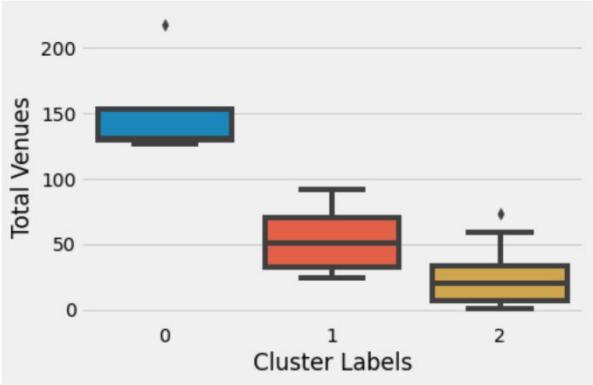
```
Out[90]: <matplotlib.axes._subplots.AxesSubplot at 0x7fbcd6cfacd0>
```



From the box-plot we can see that areas in cluster 2 outperform other clusters, this criteria is important as it depicts profitability of the areas. The areas in cluster 0 follow suit after the areas in cluster 2 and areas in cluster 1 retain their position at the bottom

The last feature to compare the clusters against is the amount venues returned by the Foursquare API

```
Out[89]: <matplotlib.axes._subplots.AxesSubplot at 0x7fbcd6de5c90>
```



In the box-plot areas in cluster 0 have more areas, with areas in cluster 1 having areas more than areas in cluster 2.

According to the box-plots, The Areas in Cluster 0 are the top priority areas to invest in, and the backup options are areas in cluster 2 while areas in cluster 1 will be avoided

In order to get the top the areas in cluster 2, the mean of the features in the cluster was used as the benchmark and all the areas that were above mean (the top eleven) was assigned to be the backup option for our investors. for the areas are also good areas for investment.

Discussions

There are other ways of improving this clustering, but unfortunately our data does not have the required features to do. One of them is the Cap Rate.
Cap Rate
This is the profit one can make from net income generated by the property, or the rate of return you'd make on a house.
The cap rate is the net income divided by the asset cost.
This helps you calculate property's potential for return on investment.
The cap rate is found by dividing the property's net operating expenses by its purchase price.
But since we can not predict monthly operating expenses (utilities, taxes, maintenance), on our data, we can’t use it.

There are also ways to determine if an investment was successful, it is a post investment metric called ROI (Return of Investment).

Return on investment measures how much money, or profit, is made on an investment as a percentage of the cost of that investment. It shows how effectively and efficiently investment dollars are being used to generate profits. Knowing ROI will allow the investors to assess whether putting money into a particular investment is a wise choice or not.

The Formula for ROI

To calculate the profit or gain on any investment, first take the total return on the investment and subtract the original cost of the investment.

Conclusion

After the detailed Analysis, I have found that the best areas to invest in now are in Cluster 0 as these areas are easily connected to the centre of the city

- CLusters for recommendations

```
 * Cluster Zero  --Top Recommendations--
     - Ikeja
     - Ikoyi
     - Victoria Island (VI)
     - Maryland
```

```
  * Cluster Two (Top Five) -- Second Tier --
     - Lekki
     - Ajah
     - Ibeju Lekki
     - Surulere
     - Magodo

* Cluster Two (Bottom Six) -- Third Tier ---
     - Gbagada
     - Ogudu
     - Ikotun
     - Ojo
     - Illupeju
     - Eko Atlantic City
```

```
* Cluster One -- Not Advisable --
     - Yaba
     - Ikorodu
     - Ojodu
     - Ipaja
     - Lagos Island
     - Kosofe
     - Mushin
     - Isheri
```