CLUSTERING AND VISUALISATION OF NAIROBI CITY DIVISIONS BASED ON MOTOR RISK AND RELATED DATA -24th JANUARY 2021.

Contents

List of Figures

List of Tables

1.0 INTRODUCTION

1.1 Background Information

Nairobi city is the largest and capital city of Kenya. the city derived its name from the famous fierce Maasai tribe which is stereotypically known for killing lions for one to graduate from a boy to a Moran or tribe warrior. Nairobi means “cool water” usually a reference from the Nairobi river flowing through the city.

Nairobi is the most populous city in Central and Eastern Africa with a population of **4,397,073 inhabitants** living within 696 km2 (269 sq mi) according to 2019 national Kenyan census. (Wikipedia, 2021). This is a population density of **6,317.6/km2.**

The city has 13 divisions, and each division has a head quarter with a police division head quarter where all accidents and other crimes are reported.

1.2 Problem Description

What are the safest Nairobi City Divisions to work and live in?

Due to the populous nature of Nairobi city and other demographic details like unemployment and low income, insecurity becomes a challenge especially for the middle class and the affluent. The study is meant to inform new employees or business owners operating motor vehicles on the motor risk of each division to help them set company base or residential homes.

Insurance companies also would like to give discounts depending on the risk factor of the region or division of operation. Customers operating from less densely populated areas with low risk will get premium discounts on their motor vehicle insurance covers.

2.0 PROJECT DATA

2.1 Data Description

In this analysis we shall analyse motor vehicle accidents and crimes (Theft and vandalism) data to help motorists in deciding which divisions are riskier than others while choosing divisions to live, work or do business in. Results findings can also be used by insurance actuaries in creating models for discounting motor vehicle insurance premiums depending on the region or division where a motorist operates or lives.

Divisions will be clustered based on a five-year average combined motor accident and crime rate data and we will further analyse the common venues from each division using four square API data.

a) **Motor vehicle accidents and crimes (theft and vandalism) related data**. This will be collected from police division headquarters. Insurance companies in Kenya require a police report to pay any motor related claims. We will take an average of five-year incidences. This data will assist in clustering of Nairobi city divisions into regions of varying motor risk status.

b) **Google Map search to get the Geo co-ordinates of each police division headquarters** as the central point for the city divisions.
Using police division headquarters geo-location, we can use four square data to find out the most common venues within each division and we can relate this data with population and average motor vehicle accident and crime incidences to infer correlations.

c) We will use **four square API** to get **most common venues** within a division using the division police headquarters as the central point.
We will review the connections between motor vehicle accident incidences with the four square rated common venues in each area.

d) Population census data
Using this data, we will infer if there is any correlation between motor vehicle accident incidences and population data of a region. Four square API data on common areas will be reviewed against the population density and common areas returned.

Nairobi city divisions will be clustered using a non-supervised clustering algorithm K-Means and most common areas from these divisions fetched from Four square API will be analyzed. Clustering will be based on high to low risk. From the results motorists (businesspeople and workers) will be able to make decisions on which divisions to live or work while insurances companies can use the findings to develop insurance premium discounting models.

2.2 Data Feature Selection.

Figure 1 below represents a Jupyter notebook screenshot of the main data frame consisting the data for the analysis. Columns represents main data features selected for the analysis. The main features we will use for the overall analysis will comprise; 5-year average incidences (Service, 2019), Police division name, Division population (statistics, 2019), division Geo co-ordinates (Maps).

Table 1: Jupyter Notebook screenshot sample of main data frame with data features

	Police Division	2019	2018	2017	2016	2015	5 yrs Total	5 yrs Average	Population	Latitude	Longitude
0	Central	207	257	236	198	246	1144	228.8	206,564	-1.279451	36.818635
1	Kilimani	231	189	210	336	255	1221	244.2	185,777	-1.291653	36.795152
2	Gigiri	173	206	223	420	176	1198	239.6	308,854	-1.236273	36.808864
3	Buru Buru	145	242	208	241	228	1064	212.8	545,200	-1.279361	36.878418
4	Embakasi	163	316	289	304	248	1320	264.0	988,808	-1.309654	36.913263

3.0 METHODOLOGY

3.1 Exploratory Data analysis

In this section we explore our data further to prepare for visualisations and analysis.

We use the describe function to get values of common statistical values like mean, average, max, min and the like. These values help in data cleaning for instance if there are columns with null row values, we can replace those nulls with the mean value of those features.

Table 2: common Statistical values of the dataset

	2019	2018	2017	2016	2015	5 yrs Total	5 yrs Average	Latitude	Longitude
count	13.000000	13.000000	13.000000	13.000000	13.000000	13.000000	13.000000	13.000000	13.000000
mean	186.153846	230.230769	213.000000	232.153846	198.615385	1060.153846	212.030769	-1.291337	36.740661
std	99.762089	87.894021	91.191191	84.941986	82.541443	358.282441	71.656488	0.255220	0.202474
min	67.000000	13.000000	13.000000	111.000000	6.000000	304.000000	60.800000	-1.901200	36.287000
25%	129.000000	206.000000	148.000000	166.000000	157.000000	909.000000	181.800000	-1.309654	36.724126
50%	173.000000	242.000000	210.000000	205.000000	209.000000	1064.000000	212.800000	-1.279451	36.808864
75%	219.000000	284.000000	269.000000	261.000000	248.000000	1221.000000	244.200000	-1.266942	36.835605
max	467.000000	342.000000	381.000000	420.000000	337.000000	1770.000000	354.000000	-0.667319	36.913263

To do visualisations on the data we need to check on the column data types to ensure that they are compatible with analysis process. For instance, you cannot do a calculation formular on a type string against an integer or float.

Table 3 below contains the column data types information.

Table 3: Basic information about data set column types

#	Column	Non-Null Count	Dtype
0	P_Division	13 non-null	object
1	2019	13 non-null	int64
2	2018	13 non-null	int64
3	2017	13 non-null	int64
4	2016	13 non-null	int64
5	2015	13 non-null	int64
6	Total	13 non-null	object
7	Average	13 non-null	int64
8	Population	13 non-null	object
9	Latitude	13 non-null	float64
10	Longitude	13 non-null	float64

Column 'Total' and 'Population' are of type "object" which is not good for visualisation or calculation analysis, so we need to convert both to type "int". After converting Total and population columns to integers our data was ready for visualisations. Figure below contains the cleaned and formatted data for visualisations.

Table 4: Cleaned data sample for visualisation and analysis

	Division_Code	P_Division	2019	2018	2017	2016	2015	Total	Average	Population	Latitude	Longitude
0	1	Central	207	257	236	198	246	1144	229	206564	-1.279451	36.818635
1	2	Kilimani	231	189	210	336	255	1221	244	185777	-1.291653	36.795152
2	3	Gigiri	173	206	223	420	176	1198	240	308854	-1.236273	36.808864
3	4	Buru Buru	145	242	208	241	228	1064	213	545200	-1.279361	36.878418
4	5	Embakasi	163	316	289	304	248	1320	264	988808	-1.309654	36.913263
5	6	Kayole	124	264	146	166	209	909	182	426482	-1.270870	36.912845
6	7	Ong'ata Rongai	82	284	146	111	125	748	150	177800	-1.406750	36.714800
7	8	Ngong	67	13	13	205	6	304	61	70800	-1.901200	36.287000
8	9	Industrial Area	178	211	305	203	157	1054	211	189536	-1.294559	36.834616
9	10	Kasarani	467	324	381	261	337	1770	354	780656	-0.667319	36.324019
10	11	Kabete	219	209	195	166	146	935	187	434208	-1.257636	36.724126
11	12	Pangani	235	342	269	261	274	1381	276	626482	-1.266942	36.835605
12	13	Lang'ata	129	136	148	146	175	734	147	197489	-1.325719	36.781253

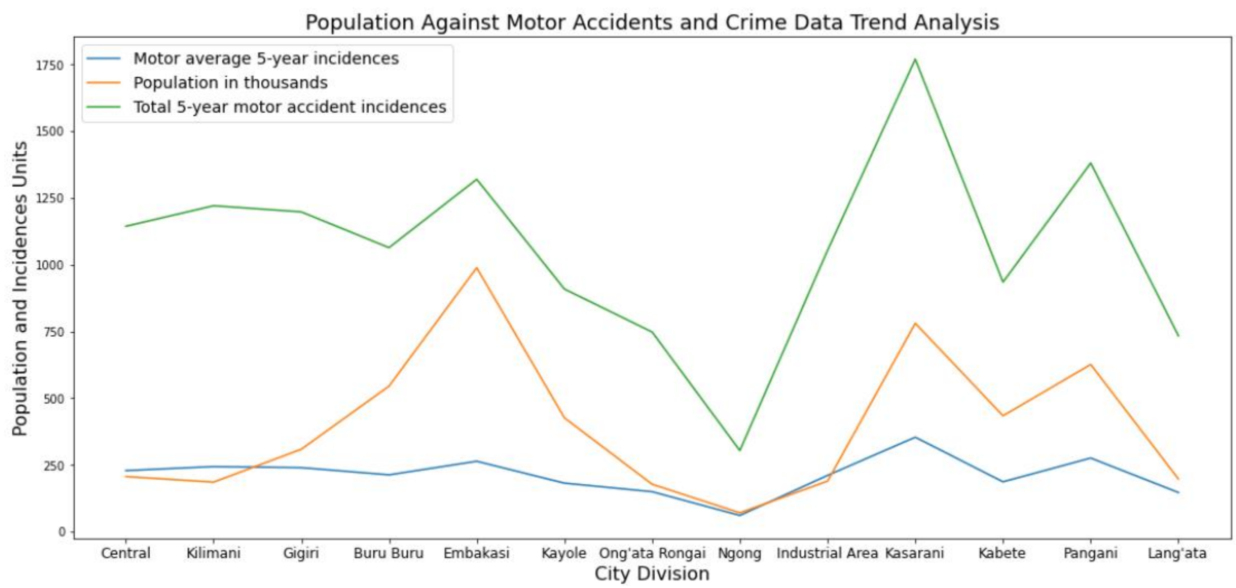
3.2 Data Visualisations and correlation Analysis

In this section we will do some visualisations to determine the main features correlations that will help us derive some inferences that we can use to advise the business.

Relationship between total 5yr incidences, 5yr average incidences, and Division Population

In this visualisation we will plot police divisions against, five-year average motor accidents and crime related data, and City division population as per the latest census. We would like to find out if trend wise the number of accidents and motor crime incidences is related to the size of the population.

Figure 1: Relationship between, Total incidences, average incidences, and City Division Population.



Observations

There is a direct relationship between motor accident and motor crime incidences with the size of the population in each division. The number of incidences increases with the size of the population.

3.3 Data Statistical analysis

To cluster the Nairobi city Divisions based on Average motor five-year average incidences and Population size, we used K-means algorithm. K=5 was optimized at 5 to enable creation of an optimal business model. This would give us 5 clusters where which will be labeled accordingly, for the business and clients operating motor vehicles within Nairobi city divisions.

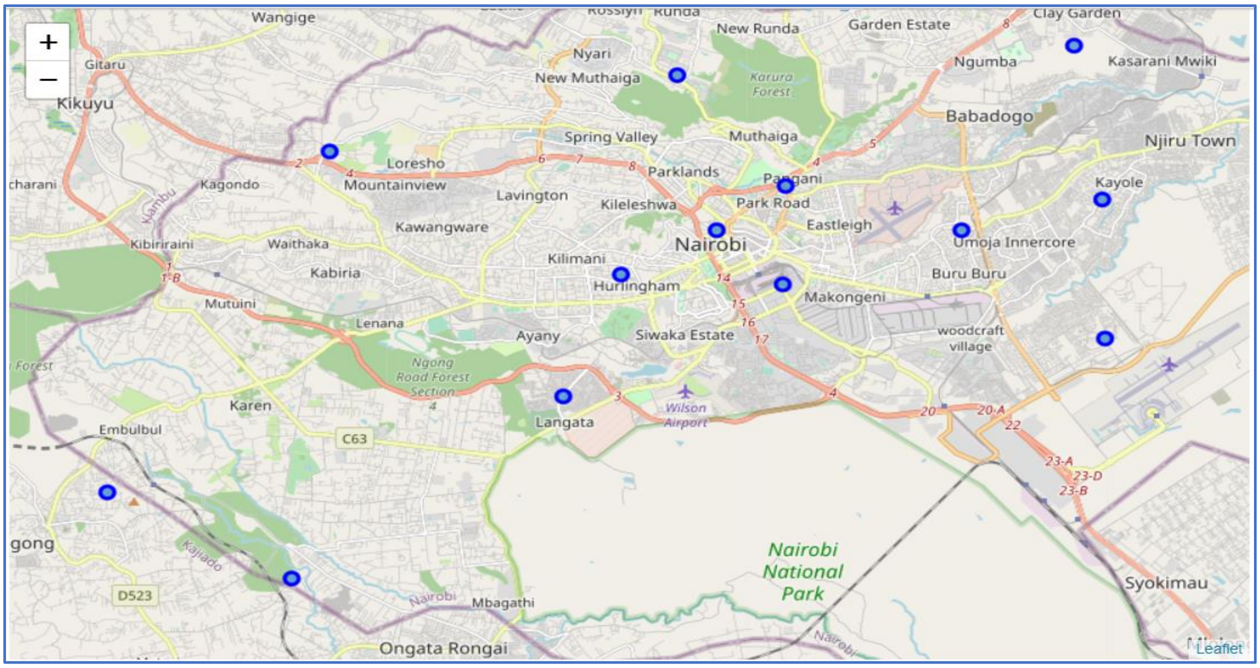
a) Statistical Analysis and Modeling

To enable K-Means clustering algorithm interpret equally features with different magnitudes and distribution, we normalized the data frame over standard distribution. **StandardScaler ()** was used.

3.4 Nairobi City Divisions and Four-square Common places and Clustering

In this section we evaluate Nairobi City and its divisions through a map. I will further query from 4 square to find the common venues within the clustered divisions.

Figure 2: A map of Nairobi city with the divisions superimposed.



To cluster Nairobi City Divisions, I initialized four square API to get the common venues within the Divisions. Following data was used.

Table 5: Nairobi City Divisions Geo co-ordinates

	City_Division	Latitude	Longitude	Incidences
0	Central	-1.279451	36.818635	228.8
1	Kilimani	-1.291653	36.795152	244.2
2	Gigiri	-1.236273	36.808864	239.6
3	Buru Buru	-1.279361	36.878418	212.8
4	Embakasi	-1.309654	36.913263	264.0
5	Kayole	-1.270870	36.912845	181.8

I designed a radius a limit of **100 venues** and a radius of **800 Meters** from the geographic co-ordinates of each division and below is a table containing the head of the venues returned. A total of **185 venues** were returned.

Table 6: Nairobi City Divisions with the top 10 common places returned by Four square API.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Buru Buru	Lounge	Gym	African Restaurant	Flea Market	Bakery	Bus Station	Soccer Stadium	Food Court	Fast Food Restaurant	Food
1	Central	Coffee Shop	African Restaurant	Café	Hotel	Bar	Italian Restaurant	Fast Food Restaurant	Performing Arts Venue	Museum	Chinese Restaurant
2	Embakasi	Fast Food Restaurant	Convenience Store	Gym / Fitness Center	Bar	Café	Women's Store	Food Court	Flea Market	Food	Food & Drink Shop
3	Gigiri	Burger Joint	Café	Ethiopian Restaurant	Ice Cream Shop	Pizza Place	Thai Restaurant	Italian Restaurant	Hotel	Fast Food Restaurant	Spa
4	Industrial Area	Fast Food Restaurant	Food Court	American Restaurant	Ice Cream Shop	Bus Station	Fried Chicken Joint	Pizza Place	Soccer Field	Food & Drink Shop	Ethiopian Restaurant

The venues presented a total of **70** unique categories. From these unique categories I created a table of the divisions containing the top ten common places as returned by Four square API (Square, 2021).

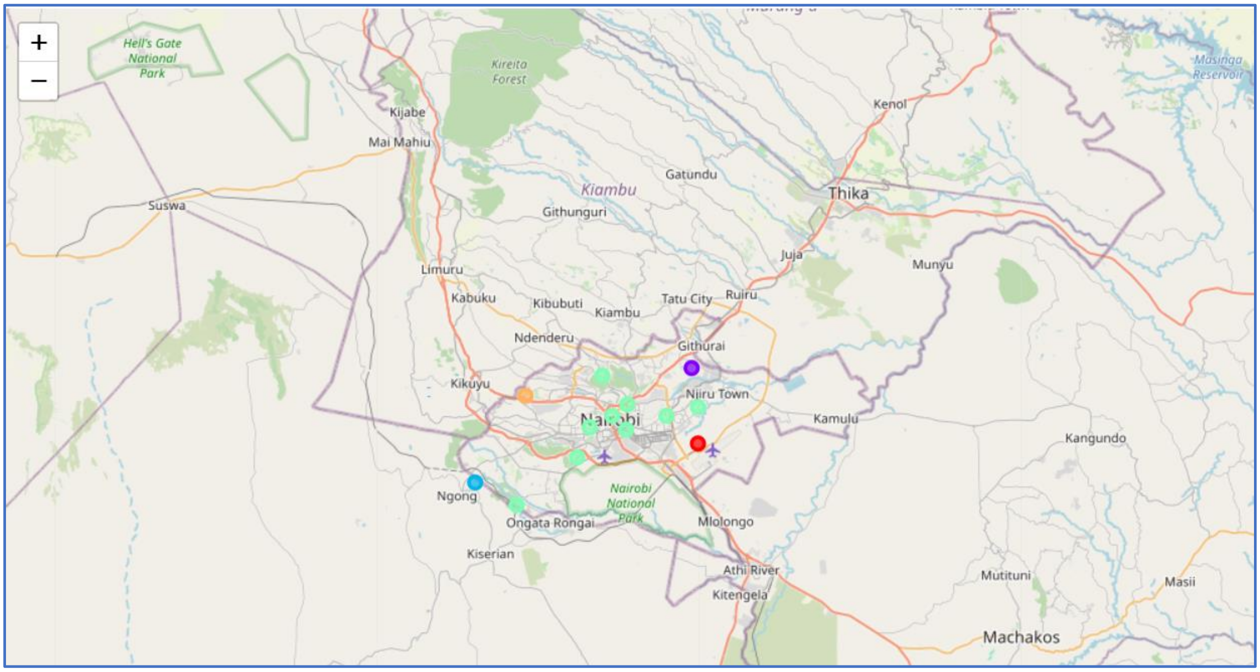
K-means Clustering algorithm was used to cluster the divisions into **5** clusters as K was optimised at 5. Below is the head of the table containing the divisions with theirclusters

Table 7: Nairobi city divisions containing their clusters in cluster Labels column

	City_Division	Latitude	Longitude	Incidences	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Central	-1.279451	36.818635	228.8	3.0	Coffee Shop	African Restaurant	Café	Hotel	Bar	Italian Restaurant	Fast Food Restaurant	Performing Arts Venue
1	Kilimani	-1.291653	36.795152	244.2	3.0	Bar	Hotel	Coffee Shop	African Restaurant	BBQ Joint	Beer Garden	Middle Eastern Restaurant	Lounge
2	Gigiri	-1.236273	36.808864	239.6	3.0	Burger Joint	Café	Ethiopian Restaurant	Ice Cream Shop	Pizza Place	Thai Restaurant	Italian Restaurant	Hotel
3	Buru Buru	-1.279361	36.878418	212.8	3.0	Lounge	Gym	African Restaurant	Flea Market	Bakery	Bus Station	Soccer Stadium	Food Court
4	Embakasi	-1.309654	36.913263	264.0	0.0	Fast Food Restaurant	Convenience Store	Gym / Fitness Center	Bar	Café	Women's Store	Food Court	Flea Market

Figure below show the map of Nairobi City containing the five clusters of the city divisions as clustered by K-means algorithm. Five clusters are represented by, Green, Red, Purple, Brown and Blue

Figure 3: Nairobi City Divisions K-Means Clusters



3.5 Examining the city division clusters.

a) Cluster 1 Labeled **'0'** and represented by red bubble on the map above.

This cluster contains Embakasi Division with an average of 264 incidences and table below show us the top 10 venues.

	Latitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4	-1.309654	Fast Food Restaurant	Convenience Store	Gym / Fitness Center	Bar	Café	Women's Store	Food Court	Flea Market	Food	Food & Drink Shop

b) **Cluster 2** Labeled **'1'** and represented by purple bubble on the map above.

This cluster contains Kasarani with the highest number of crime and accident incidences averaging at **354** incidences in a year. Below table contains the top ten common places for this cluster.

	Latitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
9	-1.227841	African Restaurant	Fried Chicken Joint	Ethiopian Restaurant	Fast Food Restaurant	Flea Market	Food	Food & Drink Shop	Food Court	Garden	Department Store

c) **Cluster 3** Labeled '2' and represented by Blue bubble on the map above.

In this cluster it contains Ngong with the lowest number of accident and crime incidences averaging at 60 incidences in 5 years. Below table contains the top 10 common places in this cluster

	Latitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
7	-1.3527	Bus Station	Women's Store	Department Store	Ethiopian Restaurant	Fast Food Restaurant	Flea Market	Food	Food & Drink Shop	Food Court	Fried Chicken Joint

d) **Cluster 4** Labeled '3' and represented by Green bubbles on the map above.

This cluster contains majority of the divisions with incidences ranging from 146 to 239 on average. Below table contains a table containing the top 10 common venues in these divisions.

	Latitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	-1.279451	Coffee Shop	African Restaurant	Café	Hotel	Bar	Italian Restaurant	Fast Food Restaurant	Performing Arts Venue	Museum	Chinese Restaurant
1	-1.291653	Bar	Hotel	Coffee Shop	African Restaurant	BBQ Joint	Beer Garden	Middle Eastern Restaurant	Lounge	Hookah Bar	Motel
2	-1.236273	Burger Joint	Café	Ethiopian Restaurant	Ice Cream Shop	Pizza Place	Thai Restaurant	Italian Restaurant	Hotel	Fast Food Restaurant	Spa
3	-1.279361	Lounge	Gym	African Restaurant	Flea Market	Bakery	Bus Station	Soccer Stadium	Food Court	Fast Food Restaurant	Food
5	-1.270870	Convenience Store	Bus Station	Gym / Fitness Center	Gym	Restaurant	Fast Food Restaurant	Pizza Place	Women's Store	Food	Ethiopian Restaurant
6	-1.376750	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	-1.294559	Fast Food Restaurant	Food Court	American Restaurant	Ice Cream Shop	Bus Station	Fried Chicken Joint	Pizza Place	Soccer Field	Food & Drink Shop	Ethiopian Restaurant
11	-1.266942	Restaurant	Shopping Mall	Food & Drink Shop	Playground	Arcade	Ethiopian Restaurant	Fast Food Restaurant	Szechuan Restaurant	Moving Target	Department Store
12	-1.325719	Gym	Bar	Convenience Store	Pizza Place	Shopping Mall	Pub	Women's Store	Bus Station	Video Store	Lake

e) **Cluster 5** Labeled '4' and represented by Orange bubbles on the map above.

This cluster represents Kabete Division which has a relatively low number of crime incidences. Below table represents the top 10 common venues in this division. Table below contains common venues in this cluster.

	Latitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
10	-1.257636	Ice Cream Shop	BBQ Joint	Gym	Garden	Hookah Bar	Hotel	Gym / Fitness Center	Hotel Bar	Gas Station	Department Store

4.0 RESULTS AND DISCUSSION

The table below contains the finals results merged of Nairobi City divisions data containing, Population, average 5-year motor vehicle incidences, K-means Cluster Number and ten most common venues in each division.

Table 8: Final analysis results containing K-means Cluster Number, Common venues from Four Square, Population size and motor vehicle incidences 5-year average.

	City_Division	Latitude	Longitude	Incidences	Population_size	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Central	-1.279451	36.818635	228.8	206564	3	Coffee Shop	African Restaurant	Café	Hotel	Bar	Italian Restaurant	Fast Food Restaurant	Performing Arts Venue	Museum	Chinese Restaurant
1	Kilimani	-1.291653	36.795152	244.2	185777	3	Bar	Hotel	Coffee Shop	African Restaurant	BBQ Joint	Beer Garden	Middle Eastern Restaurant	Lounge	Hookah Bar	Motel
2	Gigiri	-1.236273	36.808864	239.6	308854	3	Burger Joint	Café	Ethiopian Restaurant	Ice Cream Shop	Pizza Place	Thai Restaurant	Italian Restaurant	Hotel	Fast Food Restaurant	Spa
3	Buru Buru	-1.279361	36.878418	212.8	545200	3	Lounge	Gym	African Restaurant	Flea Market	Bakery	Bus Station	Soccer Stadium	Food Court	Fast Food Restaurant	Food
4	Embakasi	-1.309654	36.913263	264.0	988808	0	Fast Food Restaurant	Convenience Store	Gym / Fitness Center	Bar	Café	Women's Store	Food Court	Flea Market	Food	Food & Drink Shop
5	Kayole	-1.270870	36.912845	181.8	426482	3	Convenience Store	Bus Station	Gym / Fitness Center	Gym	Restaurant	Fast Food Restaurant	Pizza Place	Women's Store	Food	Ethiopian Restaurant
6	Ong'ata Rongai	-1.376750	36.714803	149.6	177800	3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	Ngong	-1.352700	36.669900	60.8	70800	2	Bus Station	Women's Store	Department Store	Ethiopian Restaurant	Fast Food Restaurant	Flea Market	Food	Food & Drink Shop	Food Court	Fried Chicken Joint
8	Industrial Area	-1.294559	36.834616	210.8	189536	3	Fast Food Restaurant	Food Court	American Restaurant	Ice Cream Shop	Bus Station	Fried Chicken Joint	Pizza Place	Soccer Field	Food & Drink Shop	Ethiopian Restaurant
9	Kasarani	-1.227841	36.905729	354.0	780656	1	African Restaurant	Fried Chicken Joint	Ethiopian Restaurant	Fast Food Restaurant	Flea Market	Food	Food & Drink Shop	Food Court	Garden	Department Store
10	Kabete	-1.257636	36.724126	187.0	434208	4	Ice Cream Shop	BBQ Joint	Gym	Garden	Hookah Bar	Hotel	Gym / Fitness Center	Hotel Bar	Gas Station	Department Store
11	Pangani	-1.266942	36.835605	276.2	626482	3	Restaurant	Shopping Mall	Food & Drink Shop	Playground	Arcade	Ethiopian Restaurant	Fast Food Restaurant	Szechuan Restaurant	Moving Target	Department Store
12	Lang'ata	-1.325719	36.781253	146.8	197489	3	Gym	Bar	Convenience Store	Pizza Place	Shopping Mall	Pub	Women's Store	Bus Station	Video Store	Lake

From Table 8 above the following can be deduced,

a) The number of motor accident and crime incidences are directly proportional to the size of the population in the divisions. We also observed this from trend analysis figure 1.

b) K-means algorithm was able to cluster the divisions into the following clusters that are labeled accordingly.

   i) **Cluster 1(Red bubble)**: 264 average incidences: **MEDIUM TO HIGH RISK**

   ii) **Cluster 2(Purple bubble)**: 354 average incidences: **HIGH RISK**

   iii) **Cluster 3(Blue bubble):** 60 average incidences: **LOW RISK**

   iv) **Cluster 4(Green bubble):** 187 -239 average incidences: **MEDIUM RISK**

   v) **Cluster 5(Orange bubble):** 146 average incidences: **LOW TO MEDIUM RISK**

c) Looking at the population demographics Kasarani and Embakasi have the highest population size and have been clustered by K-means algorithm as **HIGH RISK** and **MEDIUM TO HIGH RISK,** respectively. Gong with the lowest population has been clustered as **LOW RISK**

d) **From Four square API common venues from Nairobi City Divisions, we can deduce the following.**

   i) In divisions ranked as being medium to high and high risk, Fast food, African Restaurant and Fried chicken joints rank as the highest common venues.

   ii) In divisions ranked as low risk, Coffee shops, Lounge, Burger joint, ice cream shops and gyms rank as the highest common venues.

From the result above we can see a relationship between the common venues in divisions of high and low risk as being different. Future studies could be done to find out whether the type of restaurants being set up in different divisions is related to the earning or spending potential of the population in those divisions.

5. OBSERVATIONS AND RECOMMENDATIONS

Observation-population size has a direct correlation with the number of accident and crime incidences hence affecting the risk factor in the same manner.

The study through use of an algorithm, K-means clustering algorithm was able to cluster the division in varying risk clusters.

Type of common venues sprouting up in the neighborhood can indicate the level of motor risk that can be experienced.

Recommendations – motor operatives choosing a division to operate in needs to consider the population density. Also, insurance companies wishing to develop premium discounting models will need to factor in the population density variable.

Motorists and insurance companies can also use the study findings to select their division of operation and development of motor premium insurance discounting models, respectively.

Neighborhood with coffee shops, lounges, burger joints, ice cream and gyms appear to be of relatively low risk and can be used to help in making a choice of the division to operate or live in

6. CONCLUSION

K-means unsupervised clustering algorithm and data visualisation techniques are powerful and easy to use data science tools that can be used to carry out crucial research studies as this one.

Nairobi City is in a developing world and the city has seen an exponential growth of human population due to rural to urban migration. Unemployment rate is also high as jobless graduates are flocking the market in millions every year. Opportunities for jobs are also limited and this has seen the city have a skyrocketing insecurity cases motor accidents and crime related incidences being among the cases.

In this study we have managed to show that the city can be clustered into different divisions depending on the risk factor using motor vehicle related incidences which forms the highest percentage of reported insecurity cases.

Motorists wishing to reside or operate within the city divisions can use the studies recommendation to choose the safer divisions. Insurance companies also wishing to reward their customers operating within the city, can come up with insurance premium discounting models that discount in variation to the level of risk within the divisions that the insured operates.

References

Maps, G. (n.d.). Geo Positioning system.

Service, K. P. (2019). *Motor accident and motor crime incidences.* Nairobi: KPS.

Square, F. (2021, January 24). *www.foursquare.com/*. Retrieved from https://foursquare.com/: https://foursquare.com/

statistics, K. B. (2019). *National census.* Nairobi: KBS.

Wikipedia. (2021, January 11). *Nairobi*. Retrieved from en.wikipedia.org: https://en.wikipedia.org/wiki/Nairobi

Article published on LinkedIn.

https://www.linkedin.com/pulse/clustering-visualisation-nairobi-city-divisions-based-samuel-ruugia