Open Buildings: Kampala. A proxy for wealth and economic activity

Index

Open Buildings: Kampala.
A proxy for wealth and economic activity

Buildings are the utmost expression of wealth in physical terms. A building can convey a very intimate story about its creators and inhabitants. Its materials, its design, its location, and its size are all motivated by very intricate social factors. Therefore, a building can be seen as a vector which, with high confidence, signals income and living standards.

Google recently released a "*large-scale open dataset [which] contains the outlines of buildings derived from high-resolution satellite imagery [which] current focus is … the continent of Africa*". The dataset contains the outlines of mostly all the buildings in the continent. It also provides the building's areas in square meters.

Built area can also be regarded as a proxy for jurisdictions in which there is economic activity and therefore wealth. A simple hypothesis is that visually clustering buildings in any city, will unveil the patterns that economics and social interactions have imprinted in its urban shape.

To do that, we simply applied a clustering technique called K-means to the built area field for the building of Kampala city in Uganda. This technique generates well defined, compact groups of buildings that describe Kampala's urban context under an economic perspective.

The simplest way to express it is the bigger the building, the bigger the wealth of its owners and inhabitants. This simplistic urban wealth index is an exercise that pretends to ignite a discussion about how architecture signals wealth and power.

This approach presents some limitations and challenges. Obviously, a single variable will not be able to tell the difference among buildings with similar sizes. This can be mitigated through the number of categories. Too big of a building can only be a government facility, an industry, a hotel or an apartment complex, the same goes downwards the size vector. In similar urban studies I have found that somewhere around 11 clusters can reveal stable groups in which its elements are similar enough to confidently say they belong to the same category.

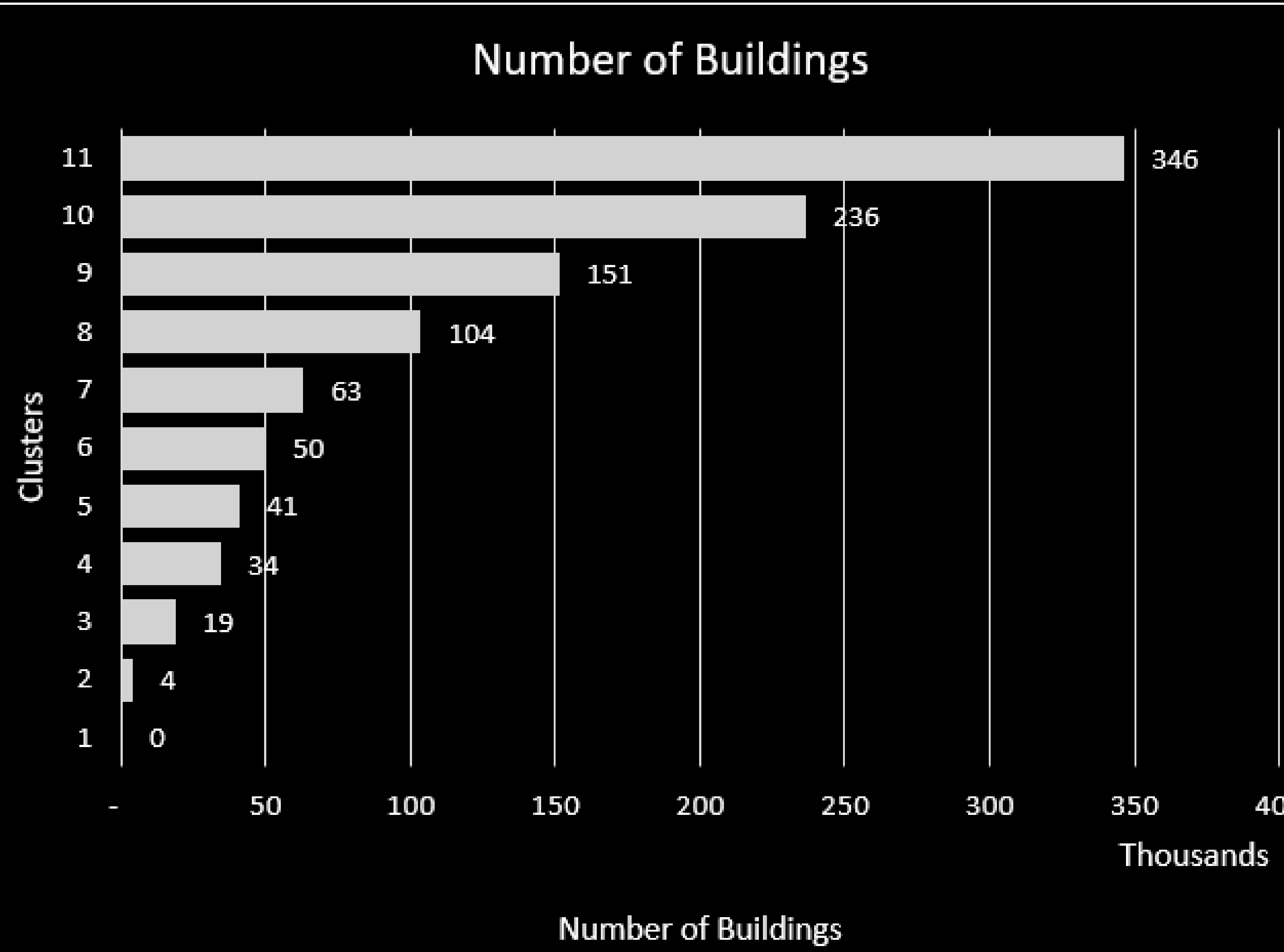
Fortunately, technology is at our service. Once again, Google comes to the rescue. The amazing power of this dataset is its granularity. This model of reality covers almost totally two single aspects of the physical urban world: buildings shape and area.

Statistics hand in hand with technological tools will provide a simple verification process. Randomly sampling clusters against Google street captions can easily reveal its nature. This article, far from being a strict academic approach is trying make a statement about the power of exhaustive data gathering process in which complex variables such as built area can minimize computation time and mitigate error at very low marginal costs.

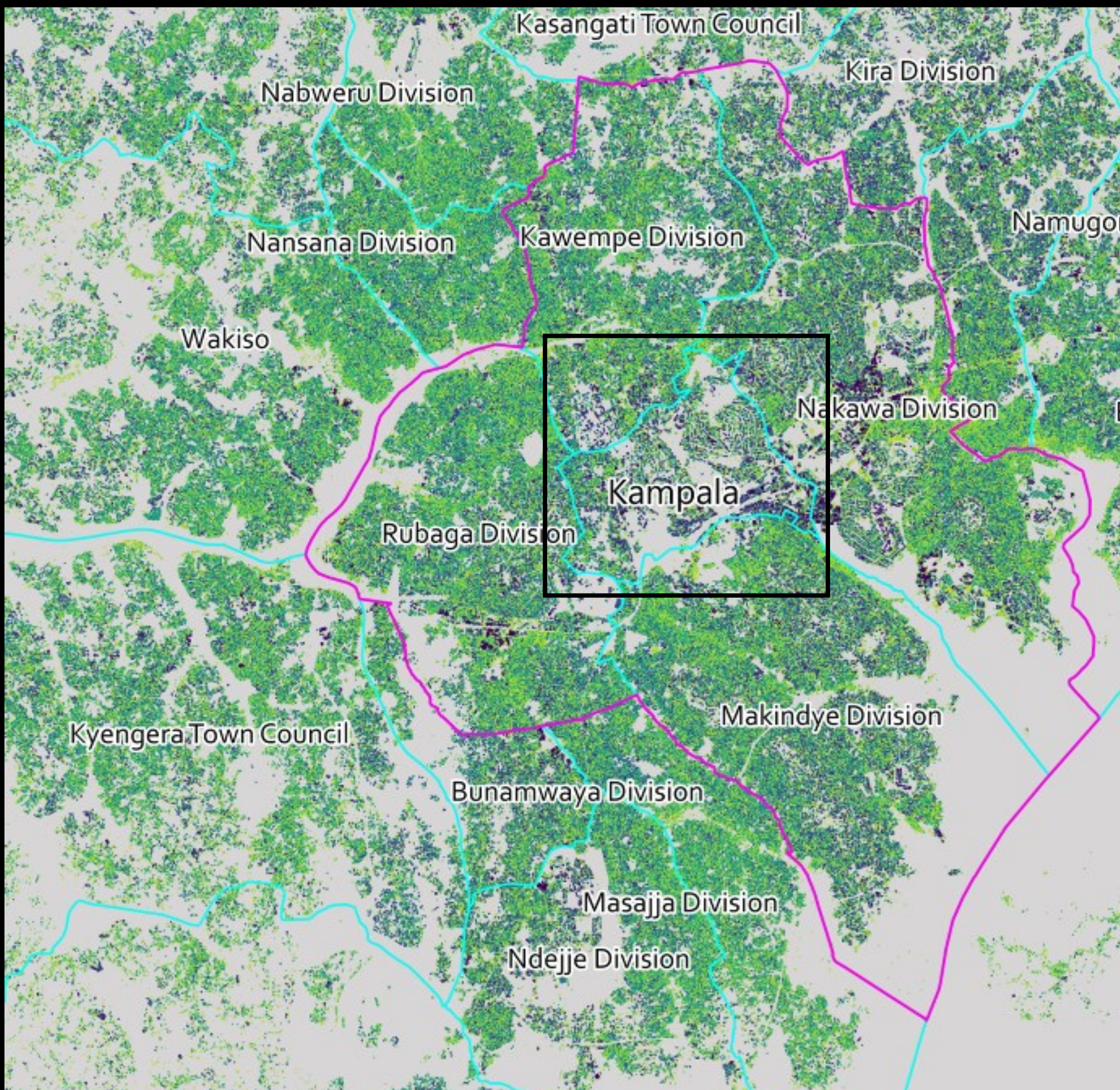
For any comments or discussion contact:
@evokare in twitter.

Open Buildings: Kampala.
Descriptive Statistics

Cluster	Number of Buildings	Average Area (m2)	Average confidence	Size
1	337	2,205.24	0.7851	
2	4,079	589.03	0.8152	Over-sized
3	18,689	286.84	0.8311	
4	34,220	191.45	0.8361	
5	40,804	142.99	0.8353	
6	49,861	111.95	0.8323	
7	63,050	87.38	0.8277	
8	103,577	65.33	0.8232	Normal size
9	151,251	45.12	0.8106	
10	236,263	25.56	0.7854	
11	346,444	9.91	0.7175	
General Statistics	1,048,575	52.50	0.7796	



Open Buildings: Kampala.



Central Kampala hosts two of the wealthiest and better-connected neighborhoods in Kampala: Kololo and Nakasero. Kololo hosts most of the embassies and private schools in the city. Nakasero on its side, hosts restaurants, government buildings and hotels.

Source: Own elaboration using data from https://sites.research.google/open-buildings/#download

Kampala Central Division

Legend

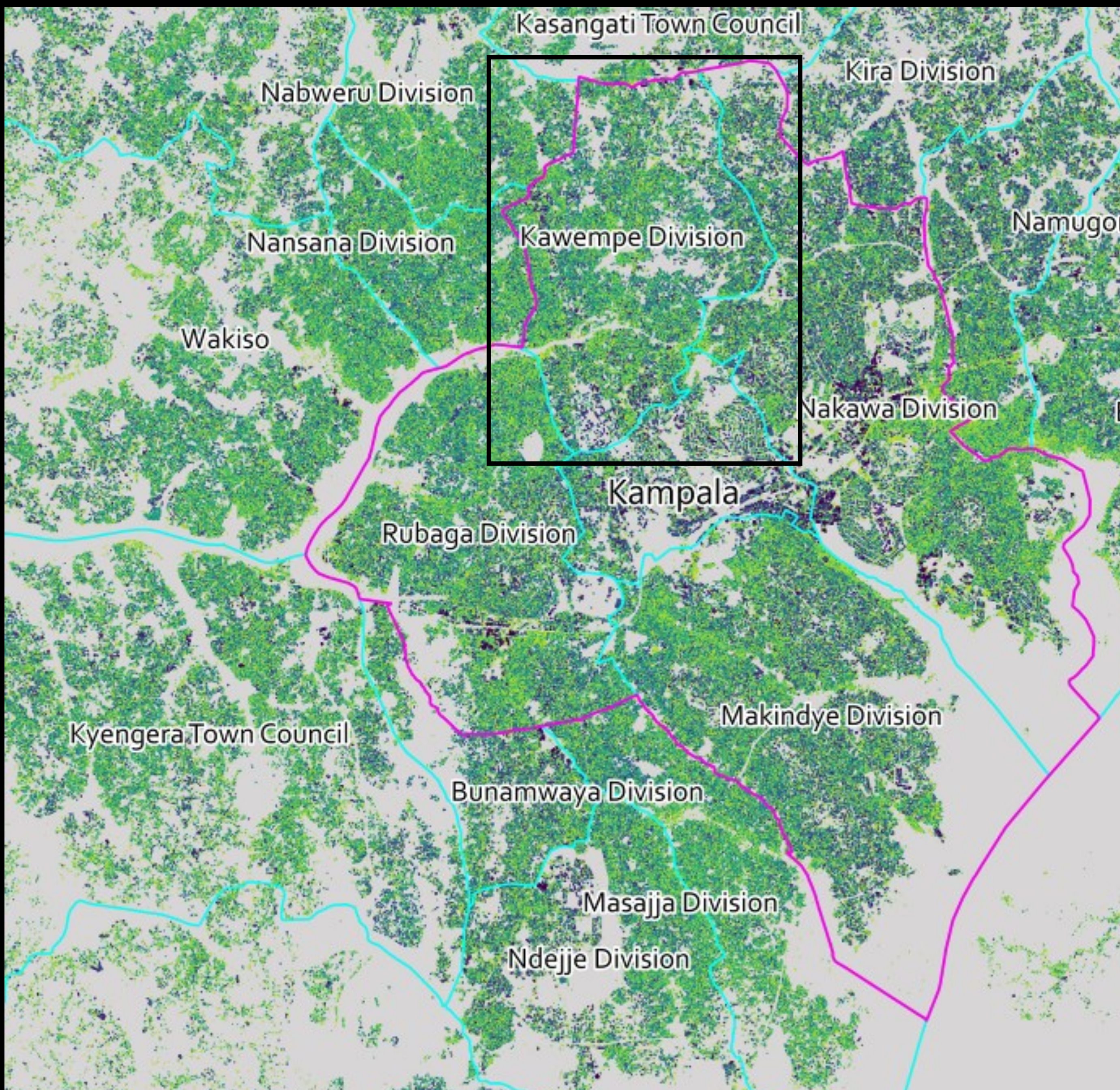
- Kampala
- Divisions

Buildings Clustered by Area (m2)

- 1 Biggest buildings
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11 Smallest buildings

The bigger the building the wealthier the area.

Open Buildings: Kampala.



Kawempe hosts a wide industrial area and the main campus of Makerere University.

Source: Own elaboration using data from https://sites.research.google/open-buildings/#download

Kawempe Division

Legend

Kampala

Divisions

Buildings Clustered by Area (m2)

1 Biggest buildings

2

3

4

5

6

7

8

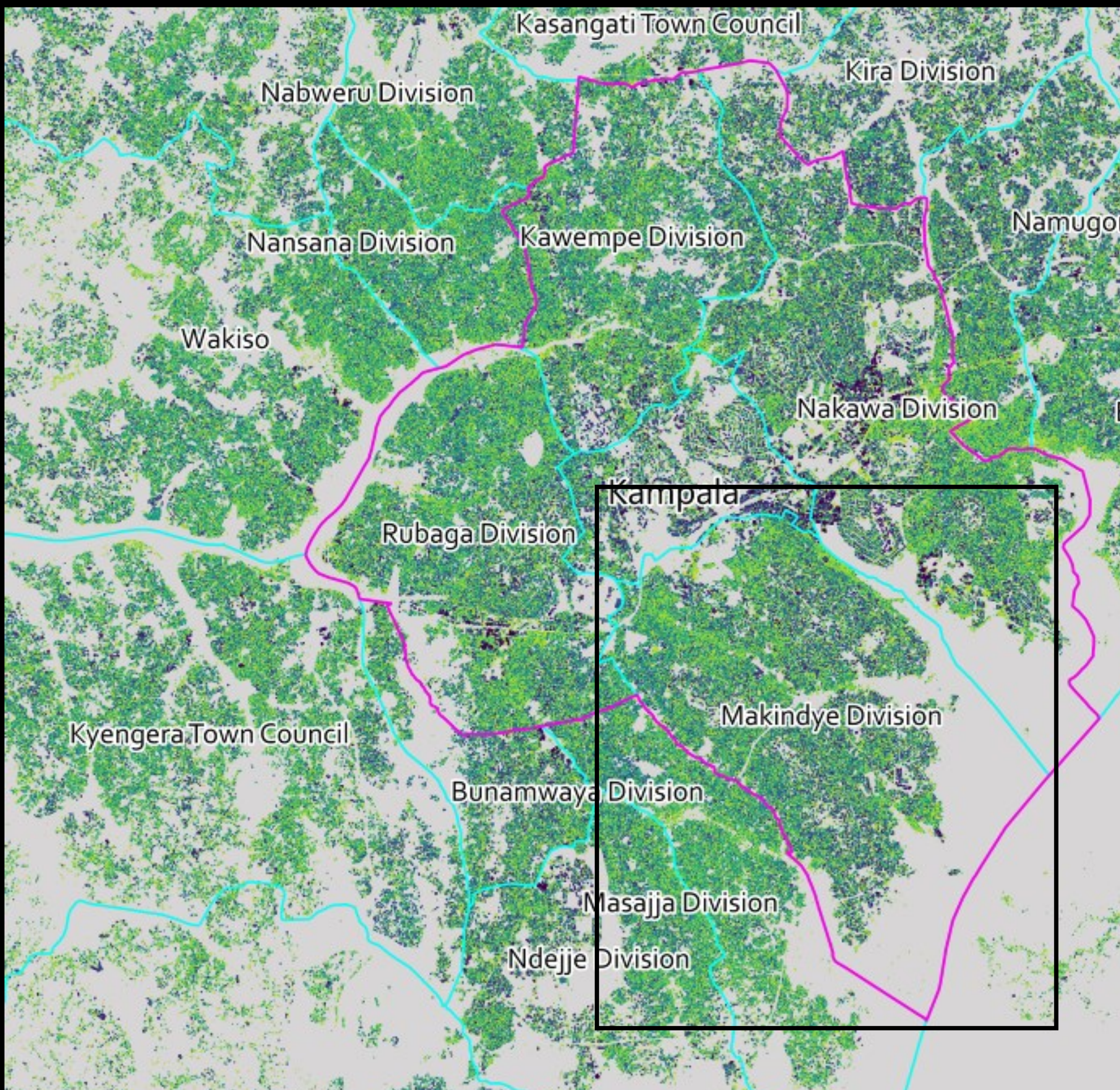
9

10

11 Smallest buildings

The bigger the building the wealthier the area.

Open Buildings: Kampala.



Makindye hosts a port, a main resort and Muyenga, a neighborhood with plenty of services and offices. Mostly a residential area.

Source: Own elaboration using data from https://sites.research.google/open-buildings/#download

Makindye Division

Legend

Kampala

Divisions

Buildings Clustered by Area (m2)

1 Biggest buildings

2

3

4

5

6

7

8

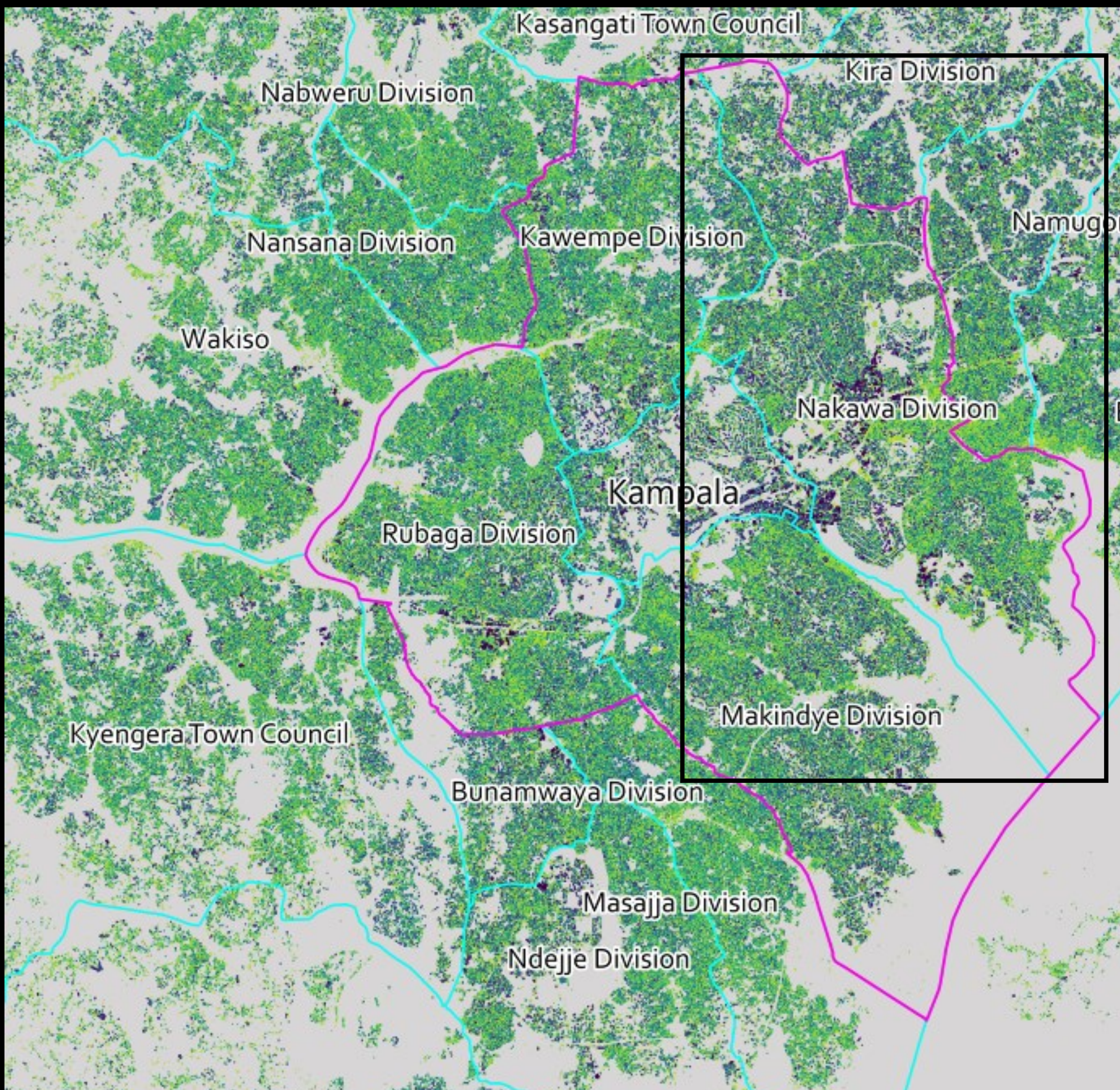
9

10

11 Smallest buildings

The bigger the building the wealthier the area.

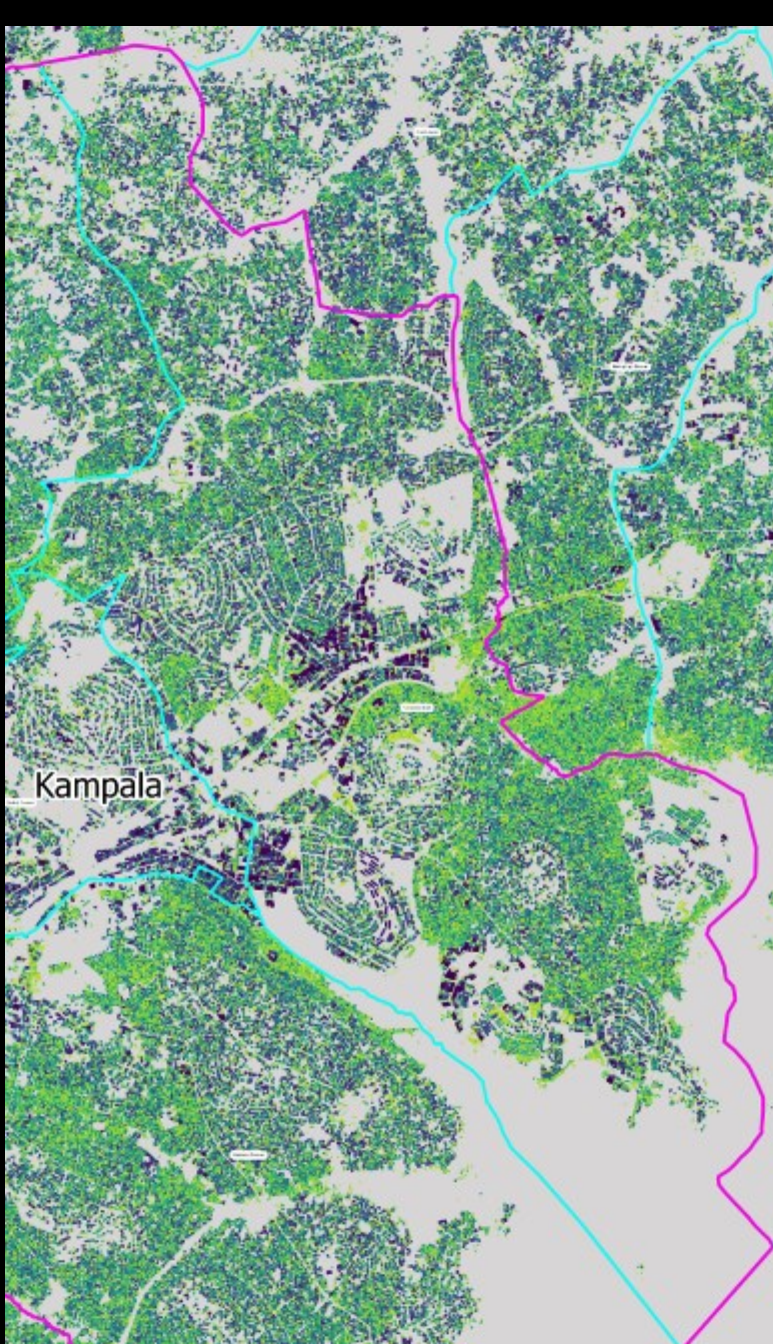
Open Buildings: Kampala.



Nakawa is home to Makerere University Business School and Kyambogo university. Bugolobi, one of its neighborhoods hosts international NGO's, churches and international representations. It also hosts one the main industrial areas in the city.

Source: Own elaboration using data from https://sites.research.google/open-buildings/#download

Nakawa Division



Legend

Kampala

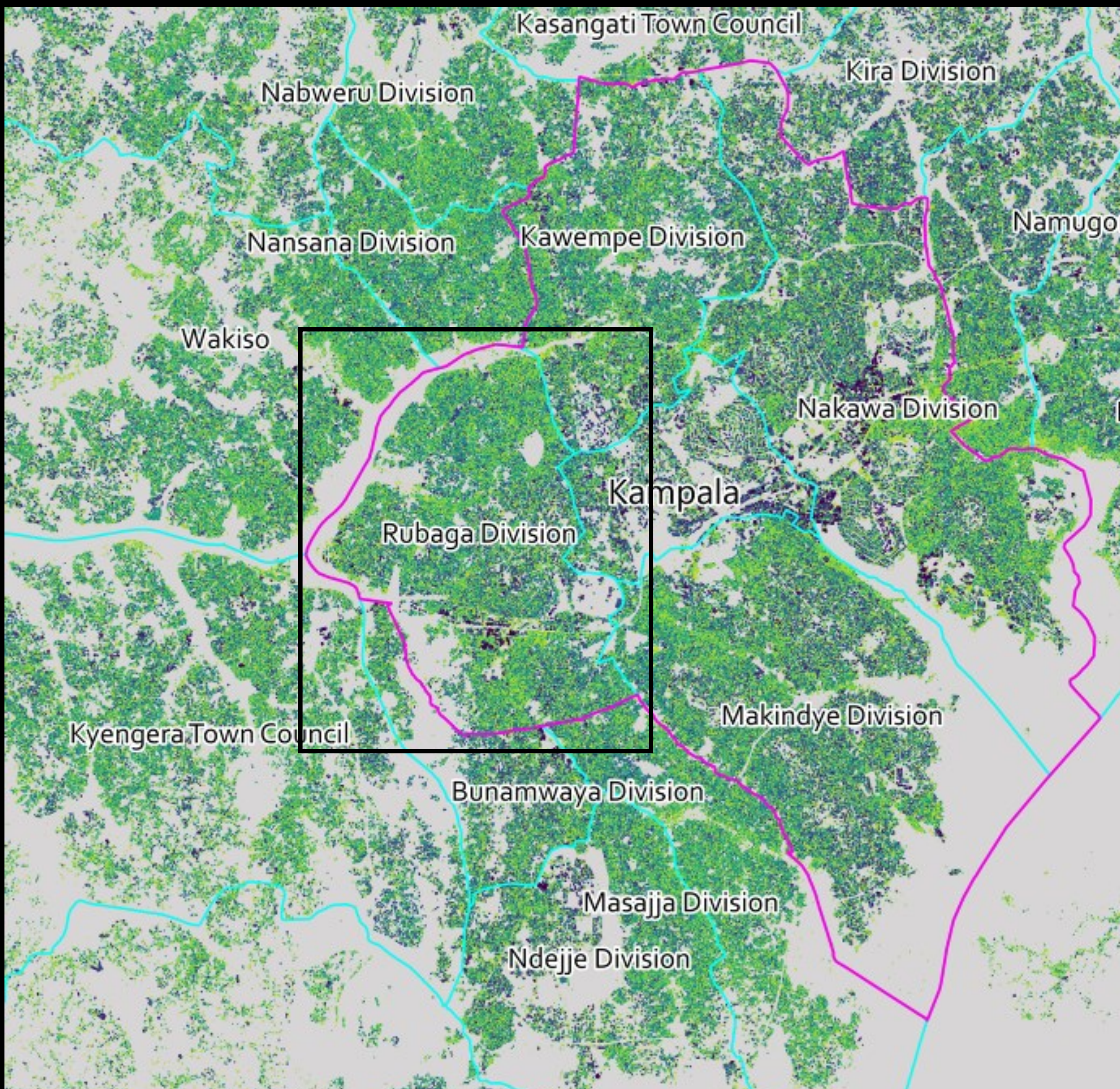
Divisions

Buildings Clustered by Area (m2)

1 Biggest buildings
2
3
4
5
6
7
8
9
10
11 Smallest buildings

The bigger the building the wealthier the area.

Open Buildings: Kampala.



Rugaba is home for an industrial railway station and multiple heavy industries. There are also a couple of stadiums and entertainment centers.

Source: Own elaboration using data from https://sites.research.google/open-buildings/#download

Rubaga Division

Legend

Kampala

Divisions

Buildings Clustered by Area (m2)

1 Biggest buildings

2

3

4

5

6

7

8

9

10

11 Smallest buildings

The bigger the building the wealthier the area.

Methodological Note

Our dataset was sourced from Open Buildings a Google's effort to map all the buildings for the whole African continent.

Building polygons

Building polygons are stored in spatially sharded CSVs with one CSV per S2 cell level 4. Each row in the CSV represents one building polygon and has the following columns:

•**latitude**: latitude of the building polygon centroid,
•**longitude**: longitude of the building polygon centroid,
•**area_in_meters**: area in square meters of the polygon,
•**confidence**: confidence score [0.5;1.0] assigned by the model,
•**geometry**: the building polygon in the WKT format (POLYGON or MULTIPOLYGON),
•**full_plus_code**: the full Plus Code at the building polygon centroid,

To produce the maps the Data Base was analyzed using R and QGIS software. The buildings for greater Kampala were segregated from the 177_buildings file using the PlusCodes.

A data clustering technique was utilized to unveil the data structure of the area_in_meters variable. This is a form of unsupervised classification method which evaluates intrinsic similarities and dissimilarities between different cases. The K-means clustering technique is a partitioning-based grouping algorithm which iteratively relocates data points (cases) among the clusters. K-means generates non-overlapping clusters based on a single dimension.

"The k-means clustering technique can also be described as a centroid model as one vector representing the mean is used to describe each cluster." (MacQueen, 1967 in Morissete et Chartier, 2013). In our case the Forgy/Lloyd algorithm was used.

For a set of cases $[x_1, x_2, \dots, x_n] \in R^d$ where R^d is the space of d dimensions, the algorithm finds a set of k clusters centers
$C = [C_1, C_2, \dots, C_k] \in R^d$ that is a solution for the following maximization problem:

$$E = \sum_{i=1}^{k} \sum_{j=1}^{n} d(c_i, x_{ij}) \quad \text{, discrete distribution}$$

$$E = \sum_{i=1}^{k} \int \rho(x) d(c_i, x_{ij}) dx \quad \text{, continuous distribution}$$

Where $\rho(x)$ I s the probability density function and d is the distance function.

Source: Morissette et Chartier, Tutorials in Quantitative Methods for Psychology. 2013, Vol. 9(1), p. 15-24.