MAUPP-2019, NAIROBI

POPULATION MAPS FOR URBAN PLANNING AND HEALTH MANAGEMENT

31 JANUARY 2019
APHRC CAMPUS, KITSURU, NAIROBI

PREPARED BY: MBOGA N., FORGET Y., GRIPPA T., LINARD C., SHIMONI M

African Population and Health Research Center

Contents

Table of figures

1. Introduction

These exercises are presented in the framework of the (**M**odelling and forecasting **A**frican **U**rban **P**opulation **P**atterns for vulnerability and health assessments) MAUPP project[1], a four-year research project (2014-2018) funded by the STEREO-III program of the Belgian Science Policy (BELSPO). The main objective of MAUPP is to develop models and forecasts to improve our knowledge of sub-Saharan African urban population patterns for a wider usage of urban population distribution datasets in vulnerability and health assessments. The MAUPP project aims to contribute to the WorldPop[2] open access archive of spatial demographic datasets.

Detailed population data is essential to improve evidence-based decision making by relevant authorities and organisations. One of the challenges of population data in SSA is the degree of outdatedness and the degree of reliability. Furthermore, the use of administrative units to present population data tends to blur the real underlying spatial patterns by creating a false impression of homogeneity within the entities. Indeed, this makes the aggregation and subsequent analyses dependent on the choice of the administrative units. In many cases, the administrative units tend to change in different census years. Consequently, the use of gridded population products, where each pixel represents the (estimated) number of inhabitants can provide useful estimation of population counts. Population products such as WorldPop are produced through Dasymetric mapping (Stevens, Gaughan, Linard, & Tatem, 2015; Wu, Qiu, & Wang, 2005).

For this workshop, two exercises are designed to demonstrate the use of Worldpop data and other open geodata using opensource QGIS GIS Software[3]. Population data is obtained from the Worldpop website, while administrative data is obtained from the database of global administrative areas (GADM)[4]. Shuttle Radar Topography Mission (SRTM) elevation data is freely available from the EarthExplorer website[5].

A quick overview of the process of searching for and retrieving data is presented for information purposes. We encourage you to explore the websites and the available datasets for your various applications.

1.1 Accessing the population data

As already mentioned, population data is obtained by searching on the WorldPop website which has an easy-to-use interface. The data has been produced using a Dasymetric mapping approach utilizing the Random Forest estimation technique

[1] https://maupp.ulb.ac.be/.
[2] http://www.worldpop.org.uk/
[3] https://qgis.org/en/site/about/index.html
[4] https://gadm.org/
[5] https://earthexplorer.usgs.gov/

that is described in detail in (Stevens et al., 2015). The data can be accessed in the following simple steps:

a) Search for the website http://www.worldpop.org.uk/data/get_data/

b) In the DATA menu, click GET DATA

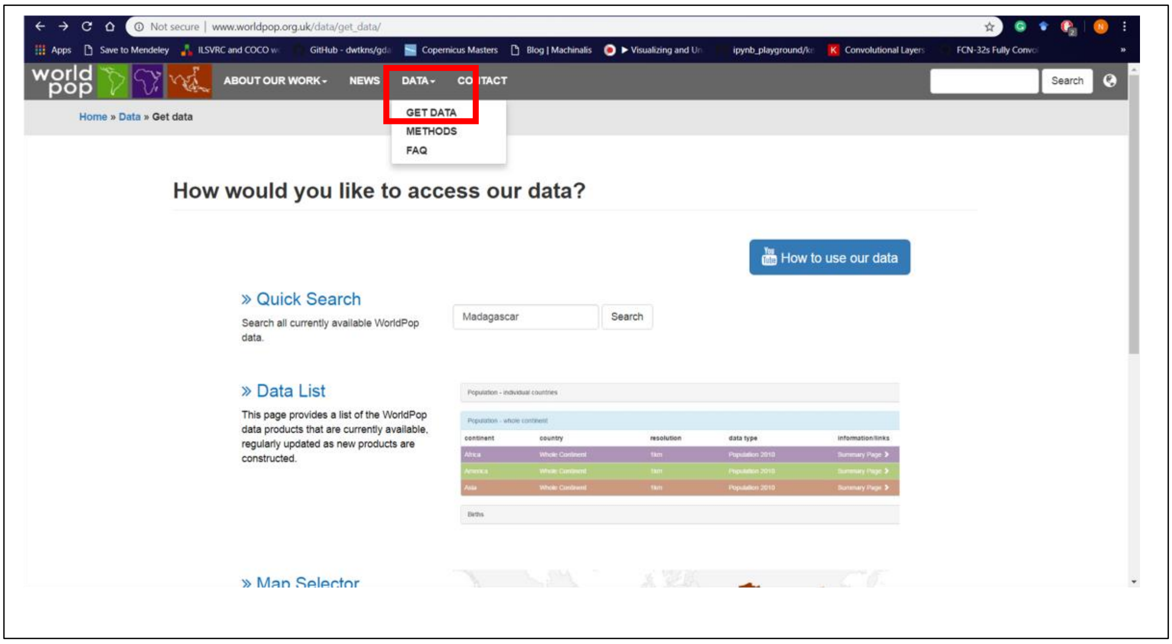


Figure 1: Searching for data for Madagascar

Search>>, <<Data List>>, <<Map Selector>> and <<Data portal (Beta)>>. Here, we use the <<Quick search>> dialogue box whereby you fill in the name of the country of interest. Then apply the search criteria. A Page with the search results pops up.

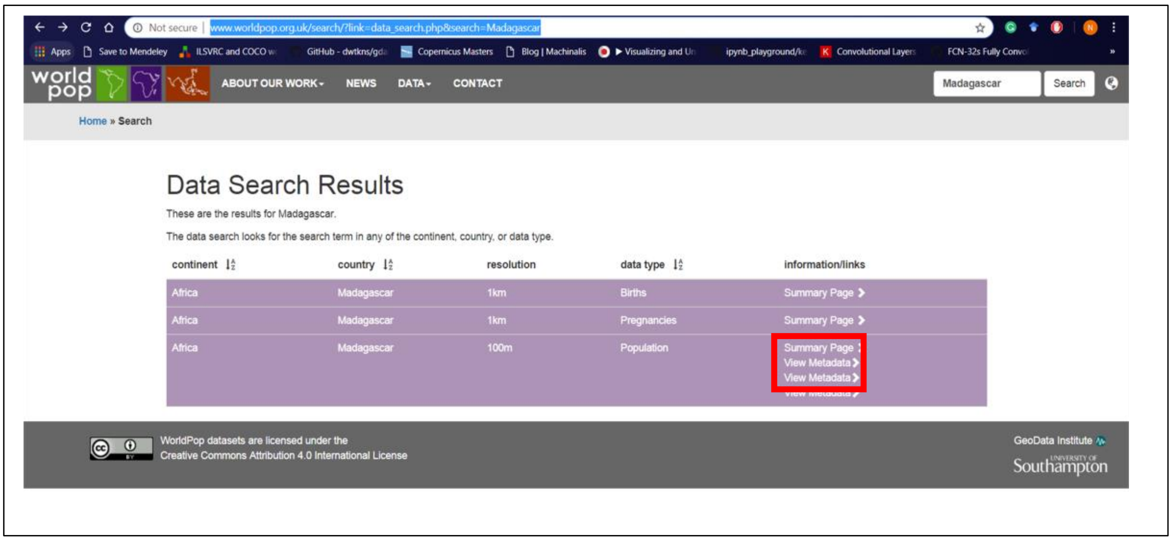


Figure 2: Population Data Search results

d) You need to select the data that meets the spatial resolution requirements of your mapping needs. In this exercise, we use the data provided at a spatial resolution of 100m. Then click on the Summary page to view the metadata. The map of Madagascar is displayed, showing the estimated number of people per square grid. As indicated in the metadata, this gridded population layer was produced using the Random Forest estimation technique described in detail in (Stevens et al., 2015). Depending on the country, other methods could have been used.

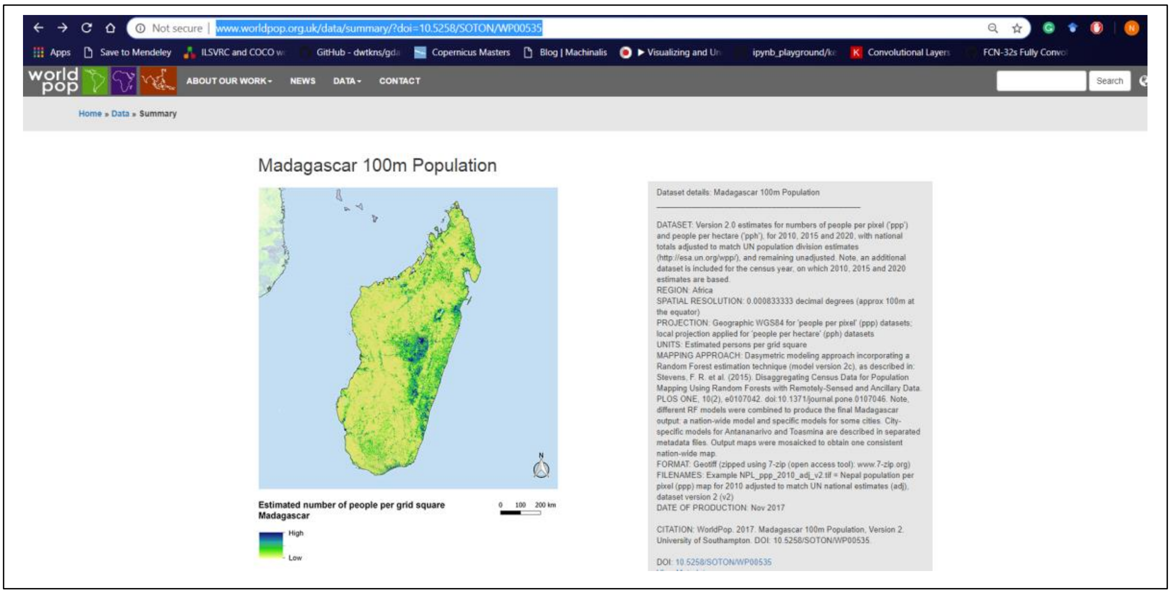


Figure 3: Map of estimated population per grid for Madagascar

e) Browse to the bottom of the page to view the downloading options. You could voluntarily fill in your details in the provided form as they help in adapting to the needs of the diverse data users. However, it is not a mandatory requirement for downloading the data. One option is to download the entire dataset. However, it is possible to download only specific files by clicking on “Browse Individual Files” which is a useful feature in case of poor internet access.

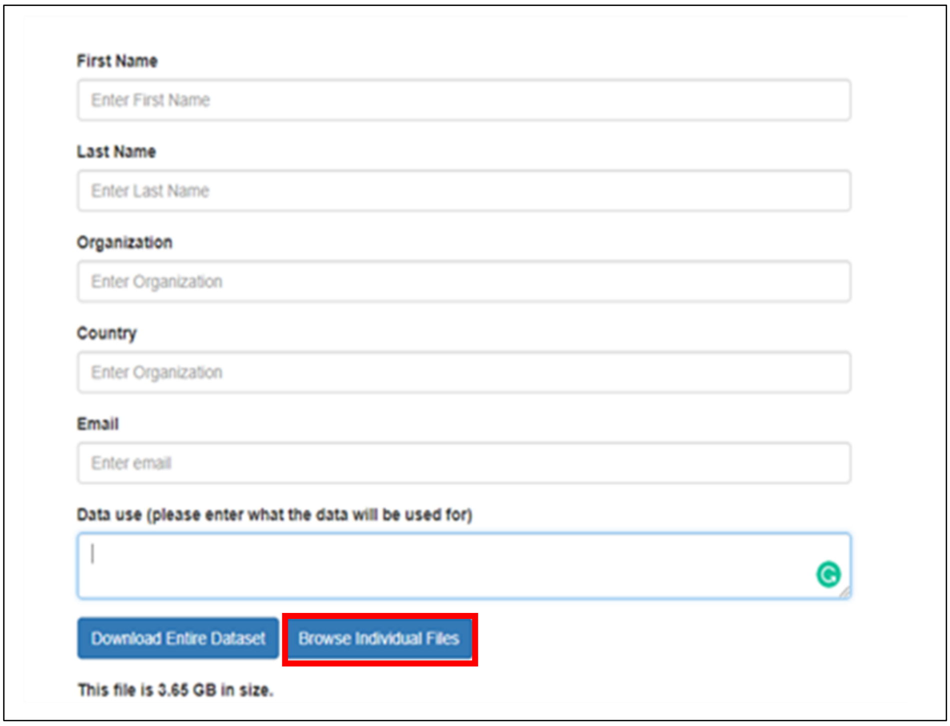
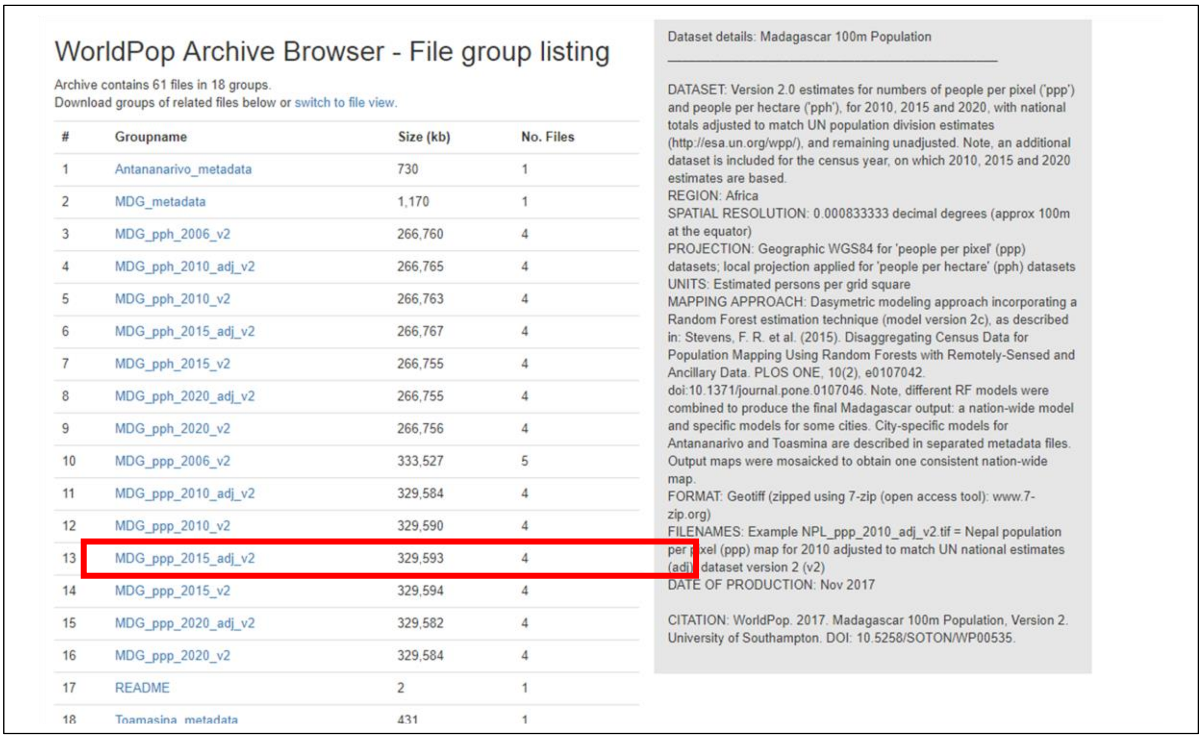


Figure 4: Contact form and data download options

Download the MDG_ppp_2015_adj_v2.tif which is the Madagascar population per pixel (ppp) map for 2015 adjusted to match UN national estimates (adj), dataset version 2 (v2)



WorldPop Archive Browser - File group listing

Archive contains 61 files in 18 groups.
Download groups of related files below or switch to file view.

#	Groupname	Size (kb)	No. Files
1	Antananarivo_metadata	730	1
2	MDG_metadata	1,170	1
3	MDG_pph_2006_v2	266,760	4
4	MDG_pph_2010_adj_v2	266,765	4
5	MDG_pph_2010_v2	266,763	4
6	MDG_pph_2015_adj_v2	266,767	4
7	MDG_pph_2015_v2	266,755	4
8	MDG_pph_2020_adj_v2	266,755	4
9	MDG_pph_2020_v2	266,756	4
10	MDG_ppp_2006_v2	333,527	5
11	MDG_ppp_2010_adj_v2	329,584	4
12	MDG_ppp_2010_v2	329,590	4
13	MDG_ppp_2015_adj_v2	329,593	4
14	MDG_ppp_2015_v2	329,594	4
15	MDG_ppp_2020_adj_v2	329,582	4
16	MDG_ppp_2020_v2	329,584	4
17	README	2	1
18	Toamasina_metadata	431	1

Dataset details: Madagascar 100m Population

DATASET: Version 2.0 estimates for numbers of people per pixel ('ppp') and people per hectare ('pph'), for 2010, 2015 and 2020, with national totals adjusted to match UN population division estimates (http://esa.un.org/wpp/), and remaining unadjusted. Note, an additional dataset is included for the census year, on which 2010, 2015 and 2020 estimates are based.
REGION: Africa
SPATIAL RESOLUTION: 0.000833333 decimal degrees (approx 100m at the equator)
PROJECTION: Geographic WGS84 for 'people per pixel' (ppp) datasets; local projection applied for 'people per hectare' (pph) datasets
UNITS: Estimated persons per grid square
MAPPING APPROACH: Dasymetric modeling approach incorporating a Random Forest estimation technique (model version 2c), as described in: Stevens, F. R. et al. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. PLOS ONE, 10(2), e0107042. doi:10.1371/journal.pone.0107046. Note, different RF models were combined to produce the final Madagascar output: a nation-wide model and specific models for some cities. City-specific models for Antananarivo and Toasmina are described in separated metadata files. Output maps were mosaicked to obtain one consistent nation-wide map.
FORMAT: Geotiff (zipped using 7-zip (open access tool): www.7-zip.org)
FILENAMES: Example NPL_ppp_2010_adj_v2.tif = Nepal population per pixel (ppp) map for 2010 adjusted to match UN national estimates (adj), dataset version 2 (v2)
DATE OF PRODUCTION: Nov 2017

CITATION: WorldPop. 2017. Madagascar 100m Population, Version 2. University of Southampton. DOI: 10.5258/SOTON/WP00535.

Figure 5: List of data files and metadata

1.2 Accessing the administrative unit's data

The analysis needs to be localised to the administrative units. GADM provides maps and spatial data for all countries and their subdivisions. The following steps can be used to download the administrative data:

a) Search for the website by pasting the URL https://gadm.org/download_country_v3.html in your browser.

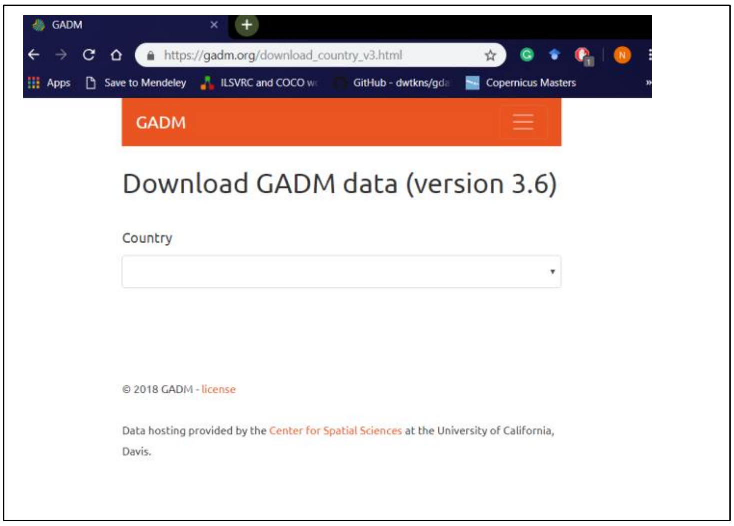


Figure 6: GADM data search interface

b) Search for the country of interest, in this case Madagascar.

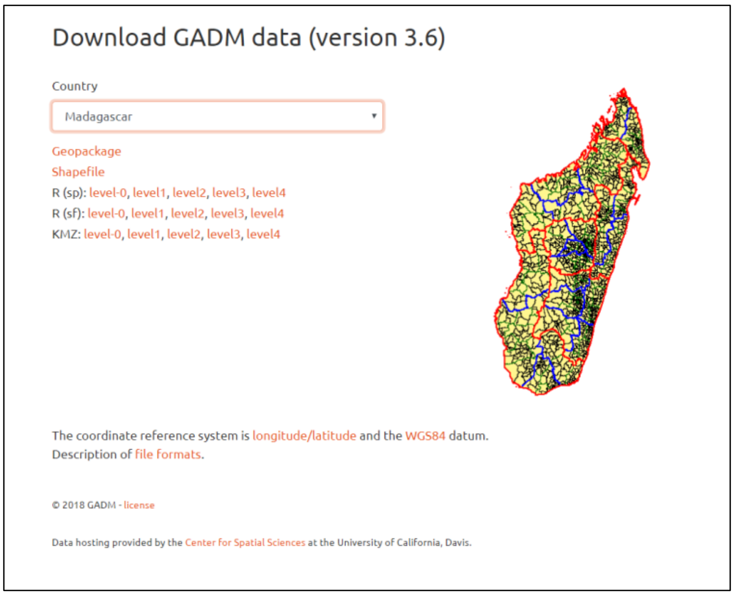
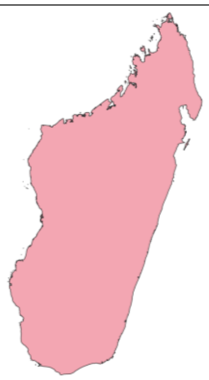


Figure 7: Search results for Administrative units of Madagascar

You can go ahead and download either the <<Geopackage>> or <<Shapefile>> based on your preference. Now, <<level-0>>, <<level-1>>, <<level-2>>, <<level-3>> and <<level-4>> indicate how fine the administrative units are, with <<level-0>> being at country level and <<level-4>> being the lowest administrative unit. In our exercise, we work with the administrative units at <<level-3>>.

		
Level-0	Level-1	Level-2

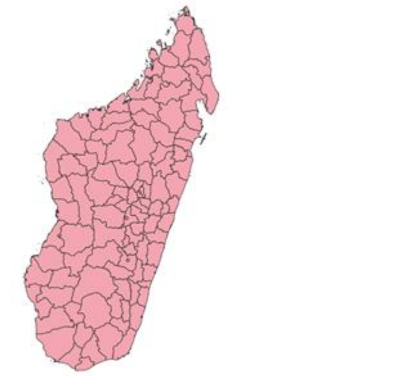
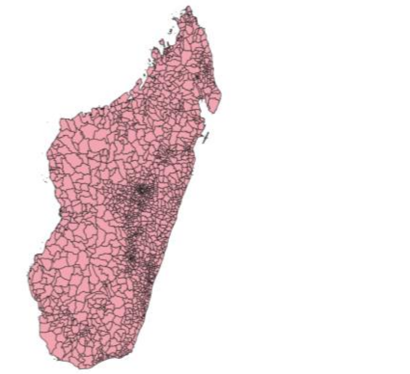
		
Level-3	Level-4	

Figure 8: the different levels of administrative units provided by GADM

1.3 Accessing the elevation data

Elevation data (Digital elevation model- DEM) can be downloaded from the earth explorer website. <https://earthexplorer.usgs.gov/>. We make use of the SRTM 1-arc resolution (30m) DEM. Note that registration (free) is required before one can download DEM or any other data from the USGS website.

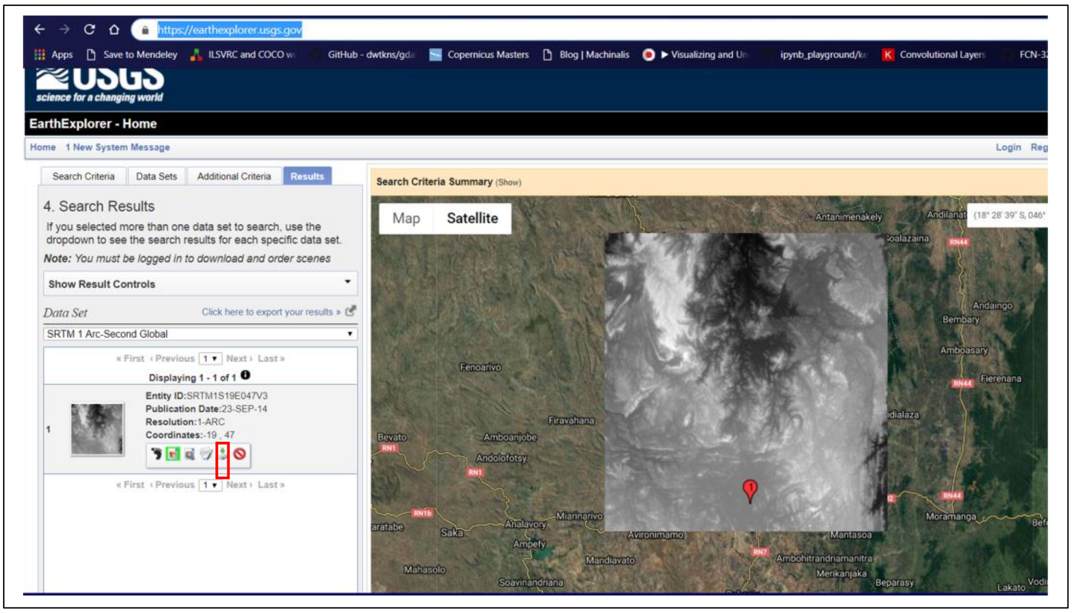


Figure 9: Accessing SRTM data from USGS website

1.4 Generation of the travel time isochrone

The code used to develop the travel time isochrome can be found at this github repository.

https://github.com/tgrippa/openrouteservice_api_script

2. Tasks

The exercises have been designed for two cities in sub-Saharan Africa (SSA) namely Antananarivo, Madagascar and Dakar, Senegal. These tasks are described respectively as follows:

1. To determine the **ratio of the population having accessibility of 15 minutes to hospitals** in Antananarivo, Madagascar.
2. To determine the **ratio of the population that is at risk from flooding** in Dakar, Senegal

2.1 To determine the ratio of the population having accessibility of 15 minutes to hospitals in Antananarivo, Madagascar.

2.1.1 Background

This is an instructor led exercise. The objective is to determine the ratio of the population in Antananarivo that has an accessibility of less than 15 minutes to hospitals. This type of information can be useful in planning the location of health centres with respect to the population of each considered administrative unit. Moreover, it could serve to improve the accessibility of the people to hospitals.

In this task, we will first determine the boundaries that intersect with the 15 minutes travel time layer (isochrone). Afterwards, the population i.e. the number of people per pixel occurring within each intersection will be extracted. Also, the total number of people per pixel for each original administrative unit will be extracted and stored in a geodatabase. Finally, a ratio of the population that has an accessibility time of 15 minutes from the hospitals will be calculated. A visual display of the results will then be displayed.

2.1.2 Data

The data used for this exercise are:

a) Point location of health centres
b) Administrative units
c) Population data
d) Travel time isochrone

2.1.3 Software

> Q-GIS 3.4.1 -Madeira

2.1.4 Methodology

a) Launch QGIS 3
b) In the QGIS browser, navigate to the folder containing data for the exercise***…\MAUPP_WORKSHOP_NAIROBI_2019\Workshop\Antananarivo***
c) Add the data:
   - The level 3 administrative data
   - The hospitals layer
   - The travel time isochrone
   - The population layer.

For better visualisation, right click on the population layer, click on layer properties and navigate to Symbology. Select the <<Paletted/Unique values Render type>> and for the colour ramp, select the <<'viridis color ramp'>>

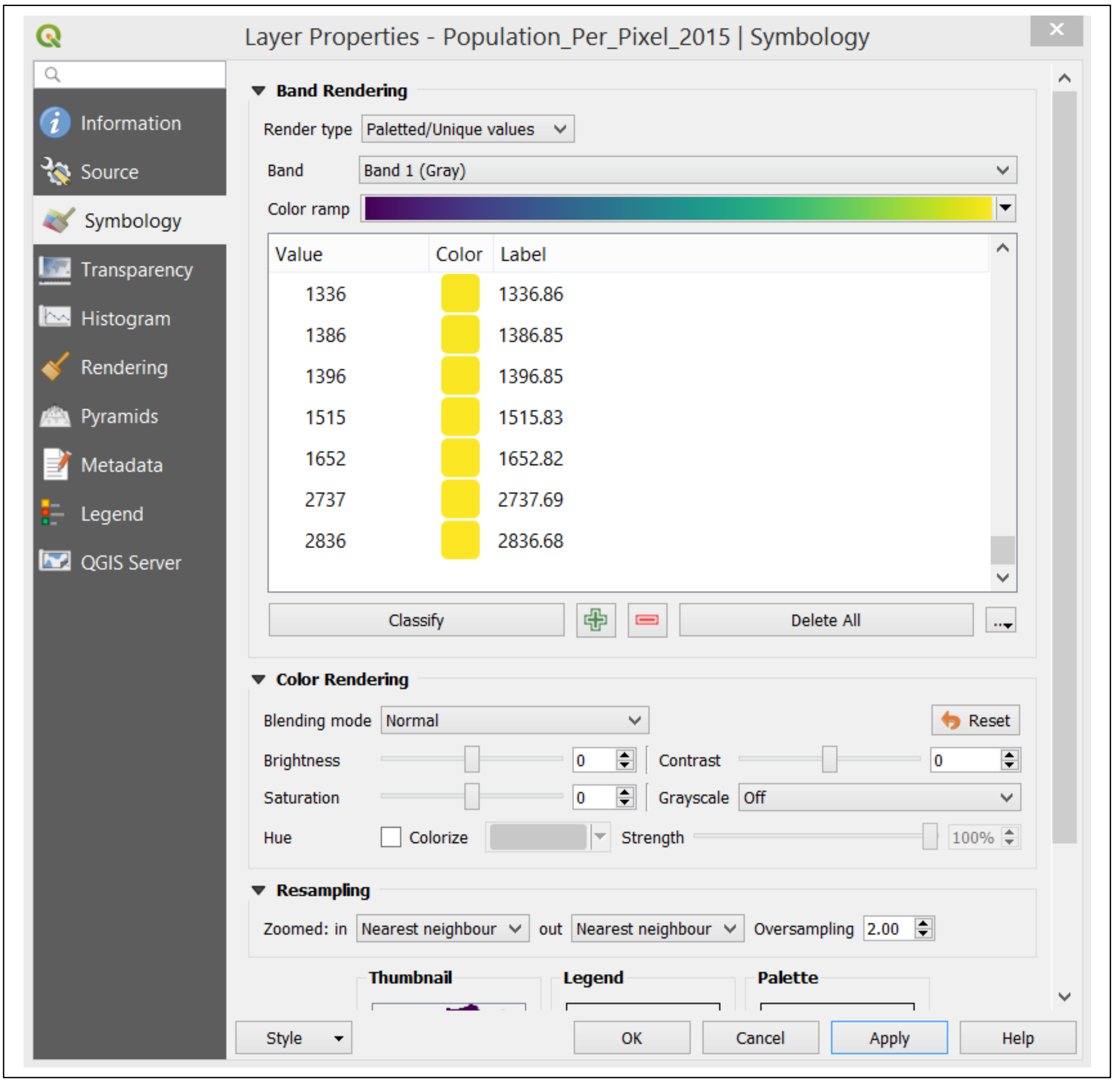


Figure 10: Changing symbology for the population data

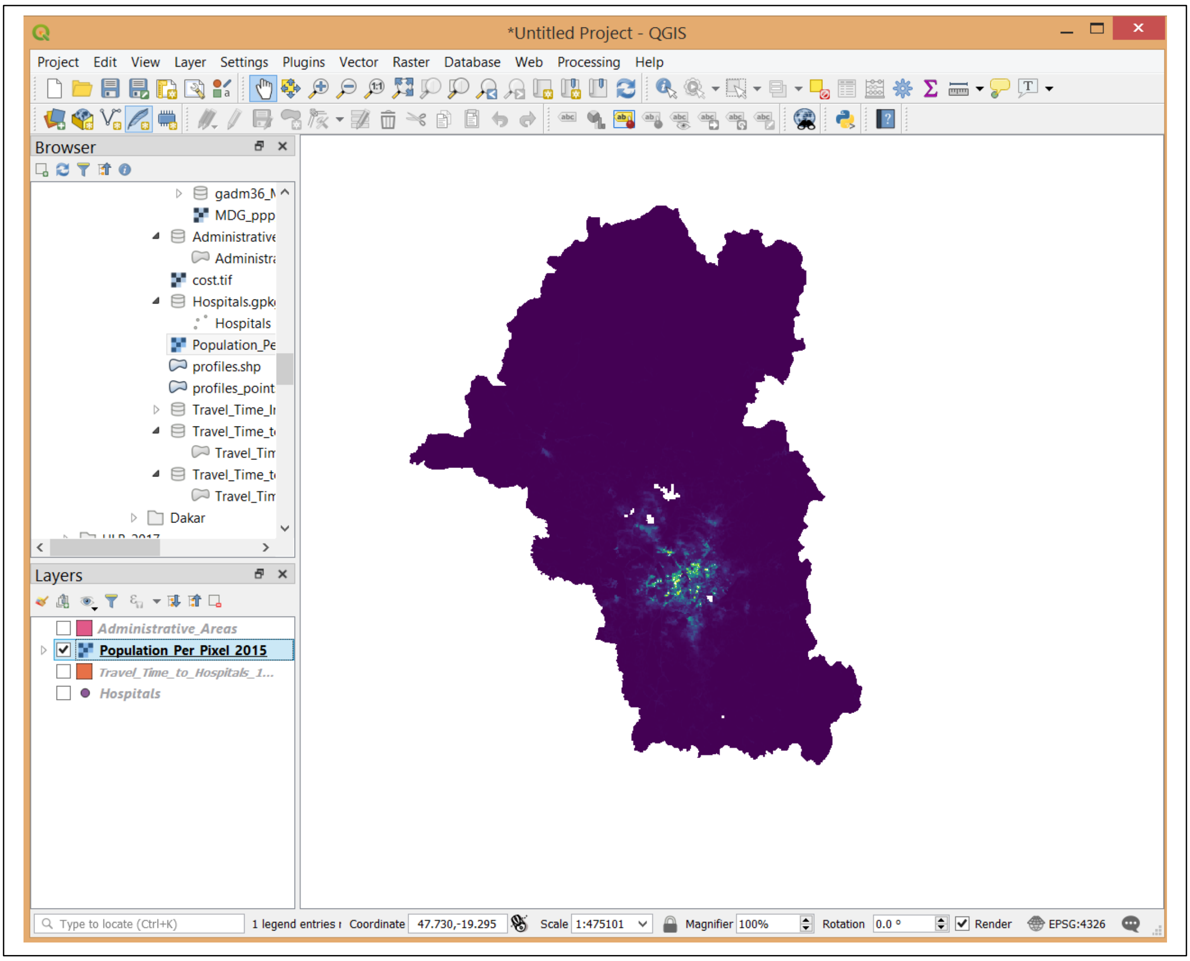


Figure 11: Population per pixel for Antananarivo, Madagascar

d) We use the <<intersection>> function to find the portion of the administrative layers that is overlaid by the travel time isochrone. From the QGIS Toolbar, click <<Vector-Geoprocessing-Intersection>>. The input layer is the Administrative units layer, while the overlay layer is the Travel time cost raster.

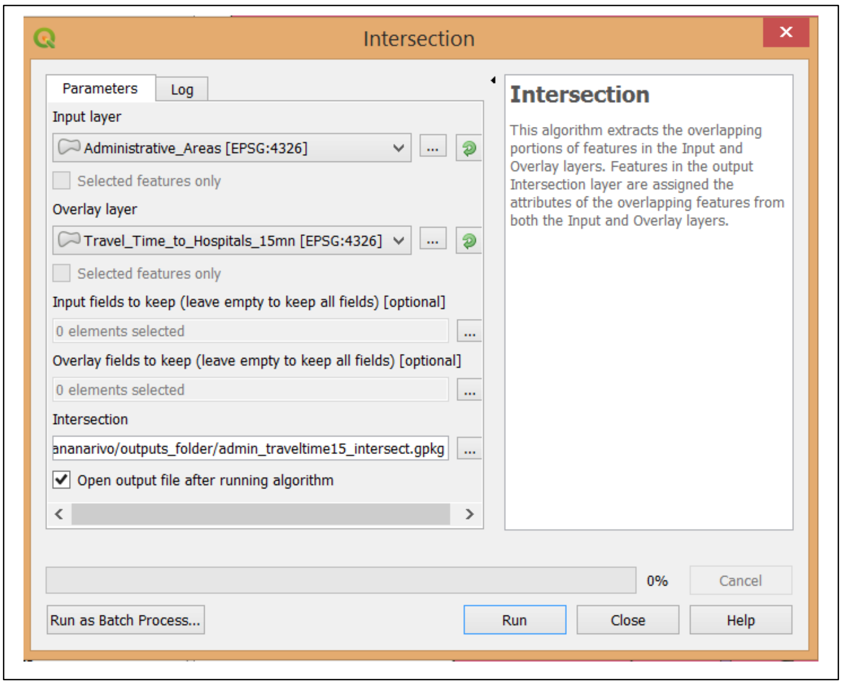


Figure 12: The Intersection tool dialogue box in QGIS

The output of the intersection operation is shown by the highlighted orange region.

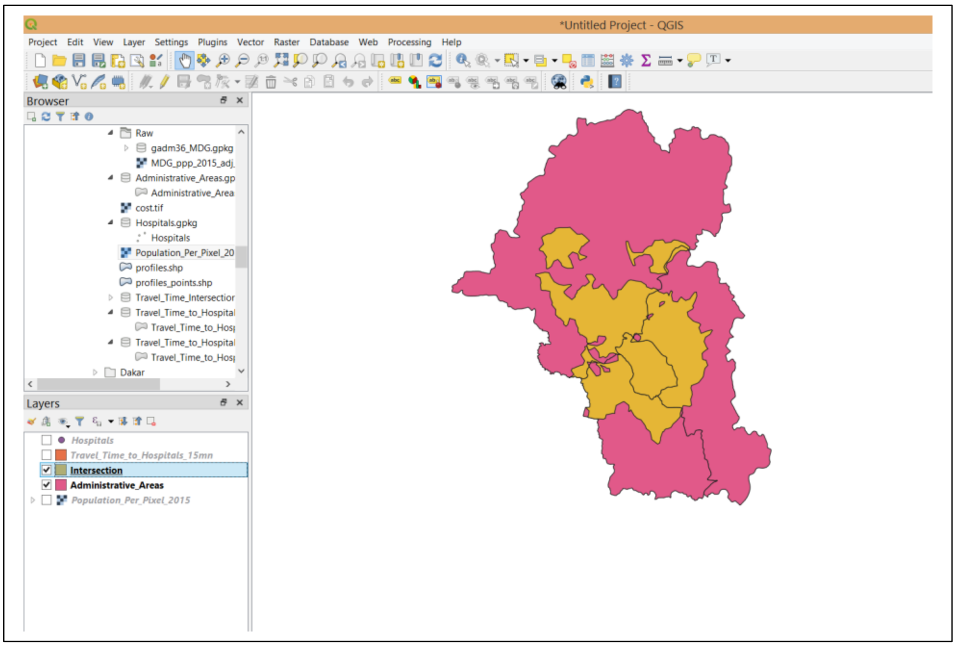


Figure 13: Intersection results between the travel time layer and the administrative units for Antananarivo, Madagascar

e) Next, we need to extract the population data from the population raster. This should be done twice because we have two vector layers (the original administrative unit and the zone that is less than 15 minutes away from hospitals). To do this, we will use the <<Zonal Statistics>> tool. This tool allows for the extraction of statistics of raster values from a raster layer into a vector layer. Examples of statistics that can be extracted include mean, sum or mode of the raster values. From the QGIS toolbox, click on the <<Processing toolbox – Raster analysis- Zonal statistics>>. The Zonal statistics dialogue window opens.

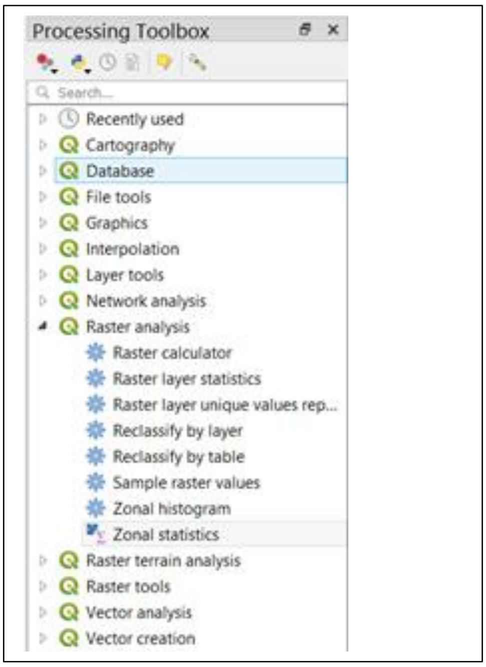


Figure 14: Accessing the Zonal Statistics tool

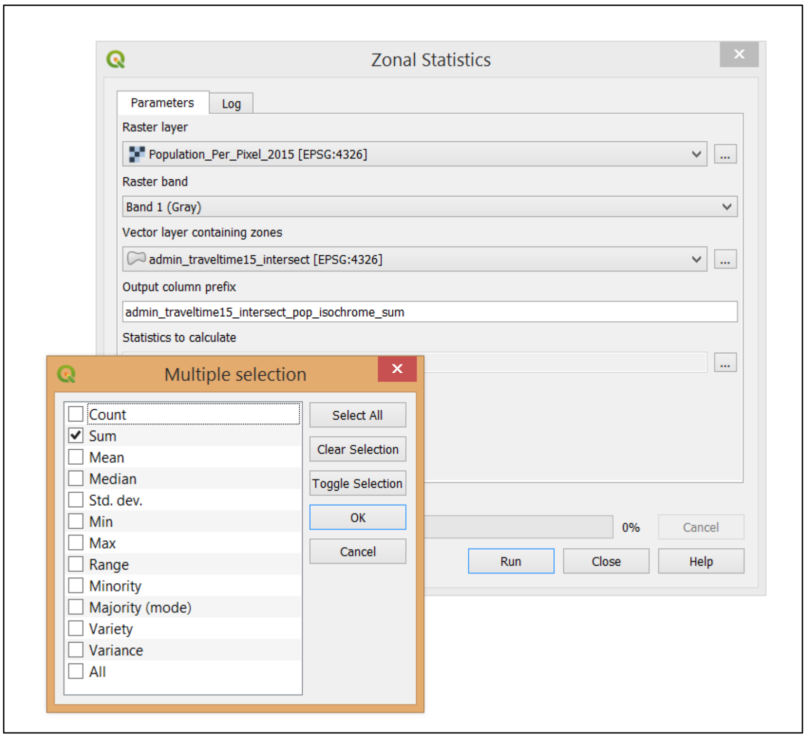


Figure 15: Computing the number of people who have accessibility of 15 minutes in each administrative unit in Antananarivo, Madagascar.

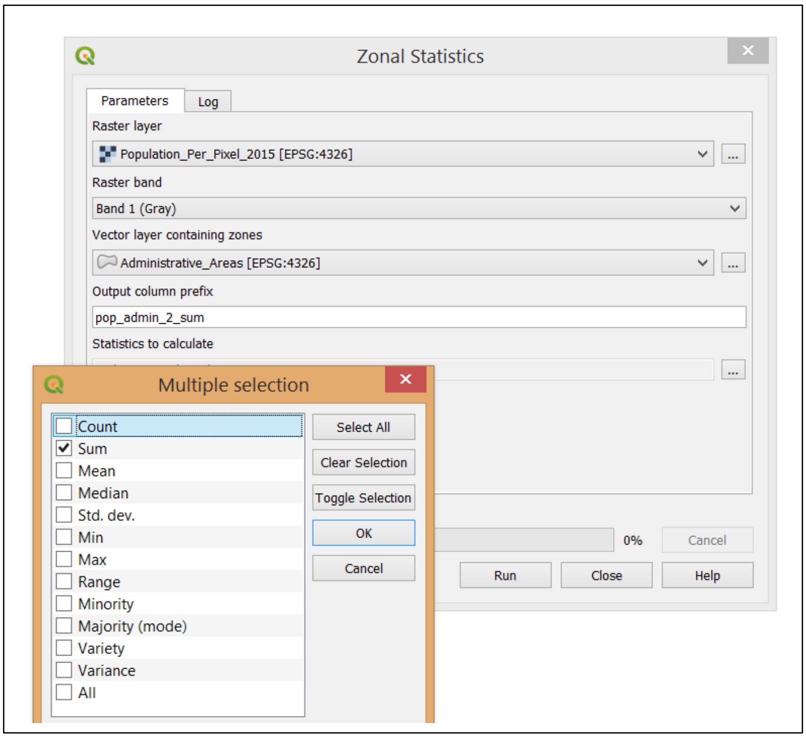


Figure 16: computing the number of people in each administrative unit in Antananarivo, Madagascar

f) We need to join the attribute tables to allow for computation of the population ratio. Right click on the administrative layer, Select <<Join>>. Select the <<Join

layer>> followed by the <<Join field>> and the <<Target field>>. In this example, the join field and the target field have the same id, but this is not necessarily always the same.

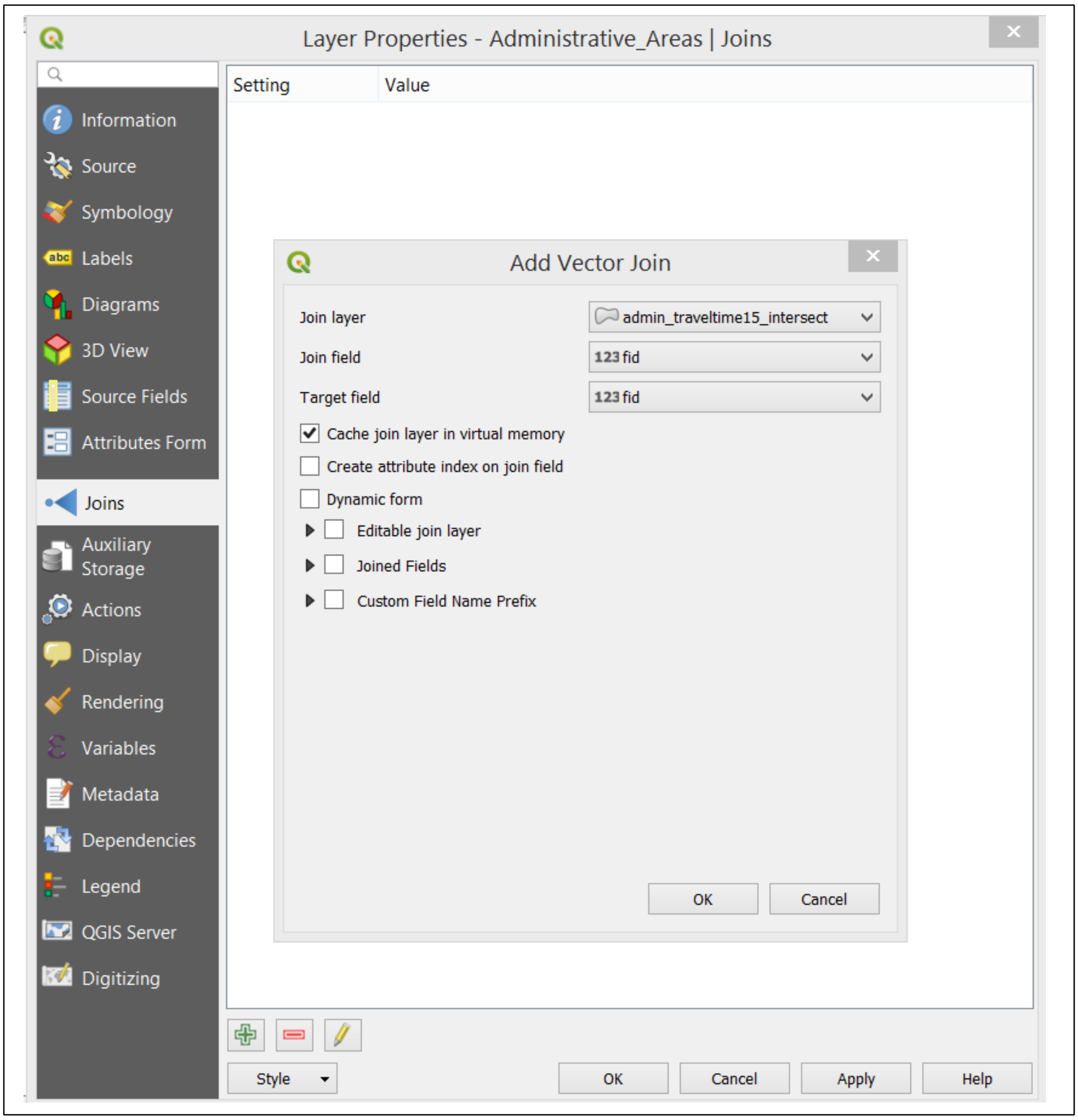


Figure 17: Performing a join between the tables containing the total population per administrative unit and the table having the population per each isochrone in each administrative unit in Antananarivo, Madagascar.

g) In this step, we make use of the field calculator. From the QGIS toolbar, search for <<Processing Toolbox-Vector Table-Field Calculator>>

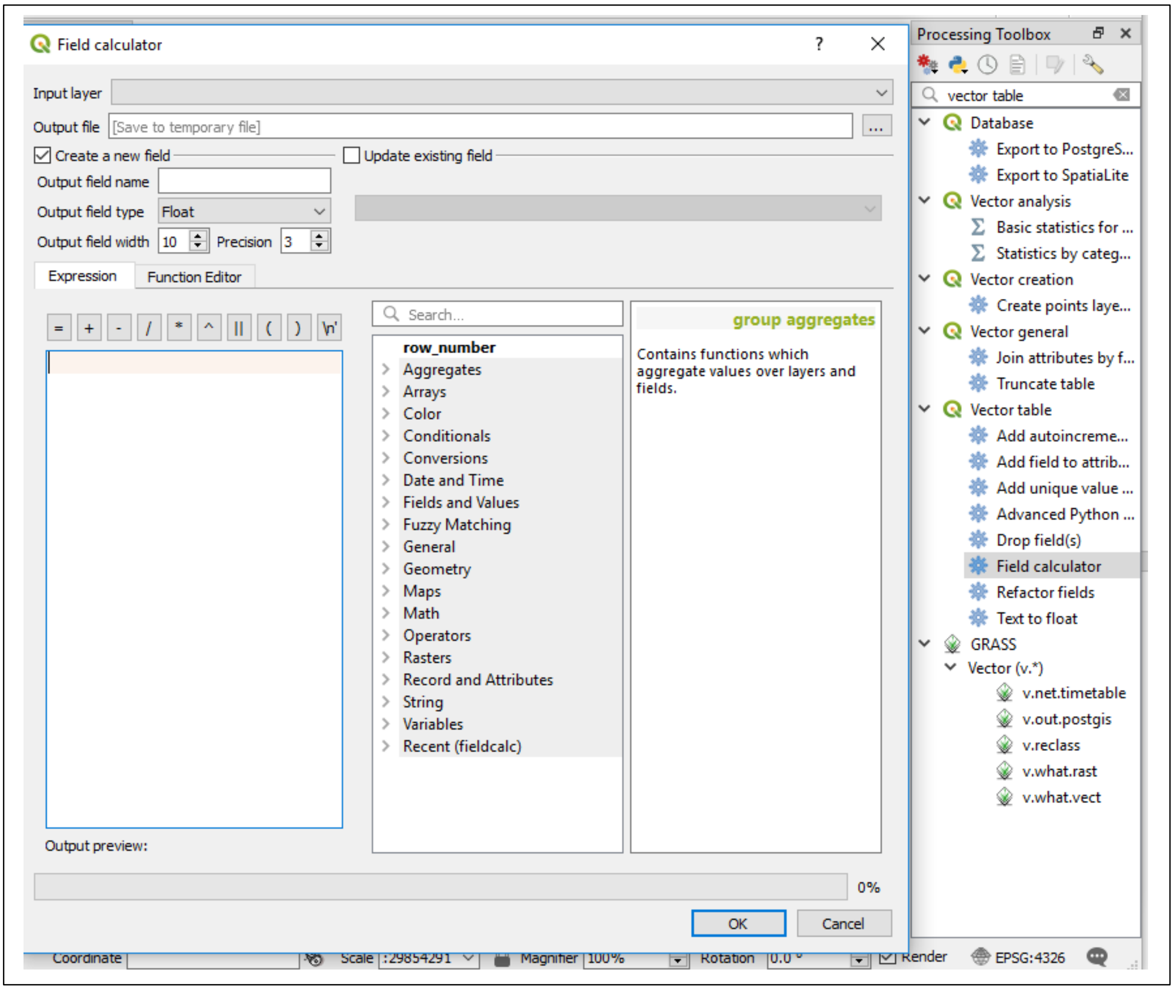


Figure 18: Using the field calculator to calculate the ratio of the population having accessibility of 15 minutes to hospitals in Antananarivo, Madagascar.

Remember to create a new field and since the ratio will be a decimal number, set the <<output field type>> to Decimal number (real). In the Show values pane, Search for <<Fields and Values>>. Navigate to the attribute columns that contain your values and create an expression in the <<expression>> pane.

The Expression used to compute the population ratio:

= "admin_traveltime15_intersect_population_isochrome2_sum" / "pop_admin_2_sum"

A new column with of the population rations is should be created in your layer of administrative units.

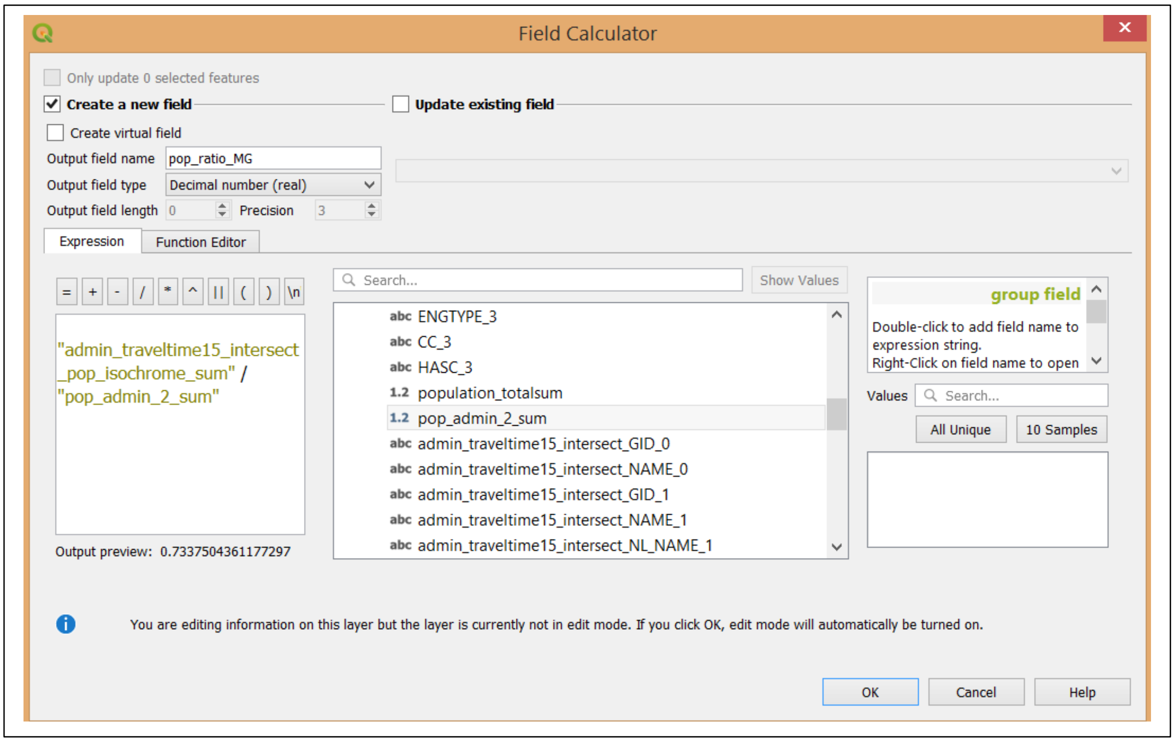


Figure 19: Field calculator dialogue box with the column values and calculation expression for the ratio of the population having accessibility of less than 15 minutes to hospitals in Antananarivo, Madagascar

The highlighted column shows the computed ratios.

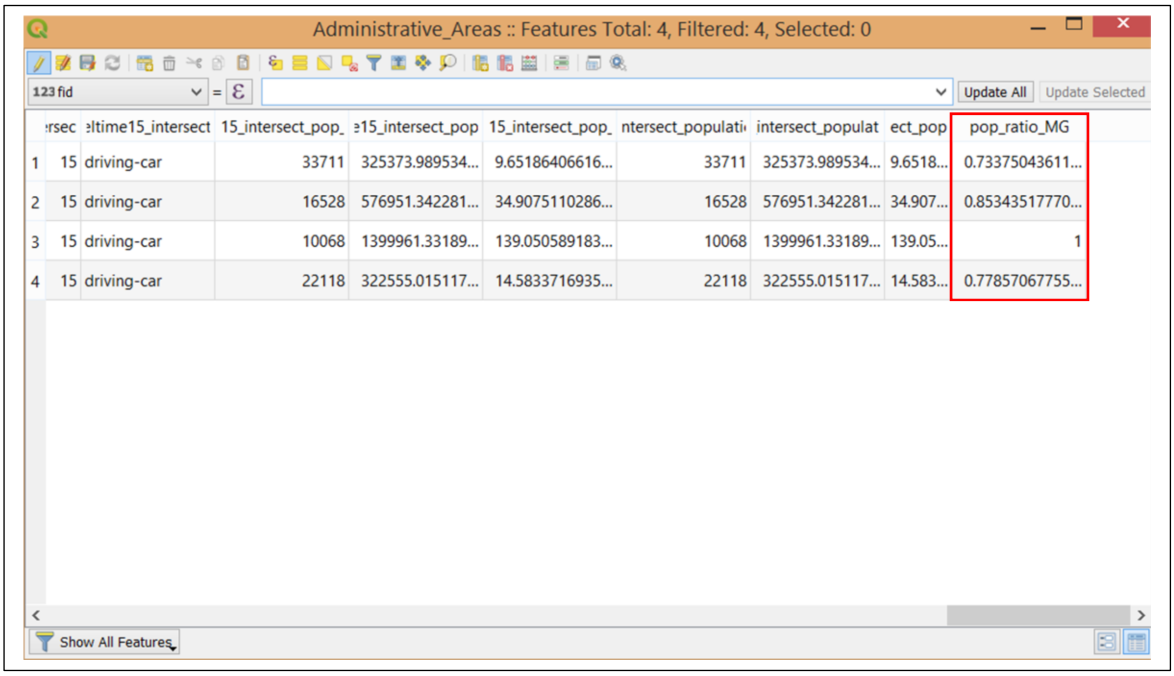


Figure 20: Attribute table of the administrative units showing the new column containing the calculated ratio in Antananarivo, Madagascar.

h) Lastly, we can use the symbology tools to efficiently display our results.

2.1.5 Results and Display

a) To access the layer properties, right click on the "Administrative_Areas" layer from the layer tree and click on<< Properties>>. An array of options is available that you can manipulate.

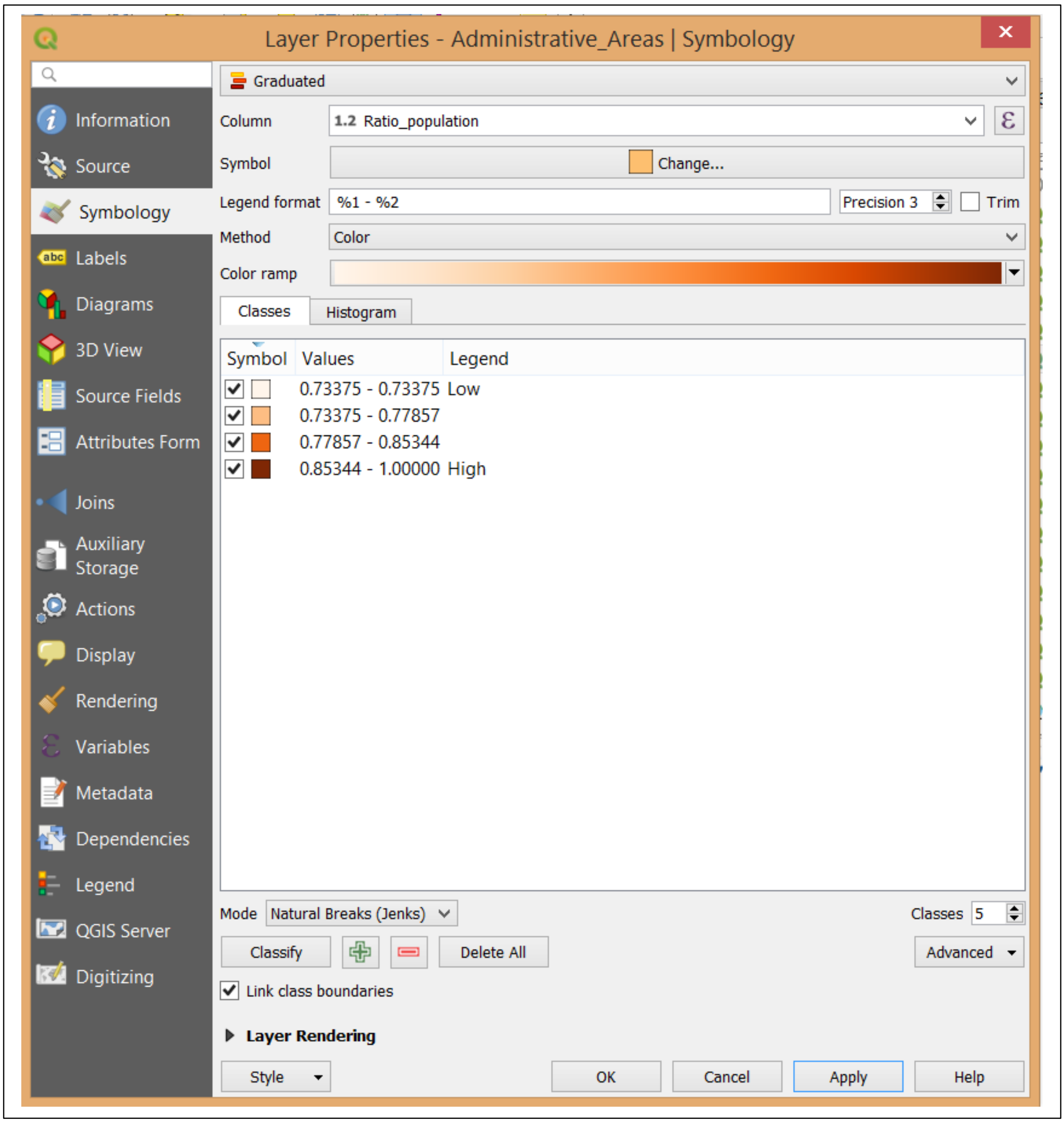


Figure 21: Changing the symbology of the Administrative layer using the population ratio column to be displayed in the legend of the Antananarivo map.

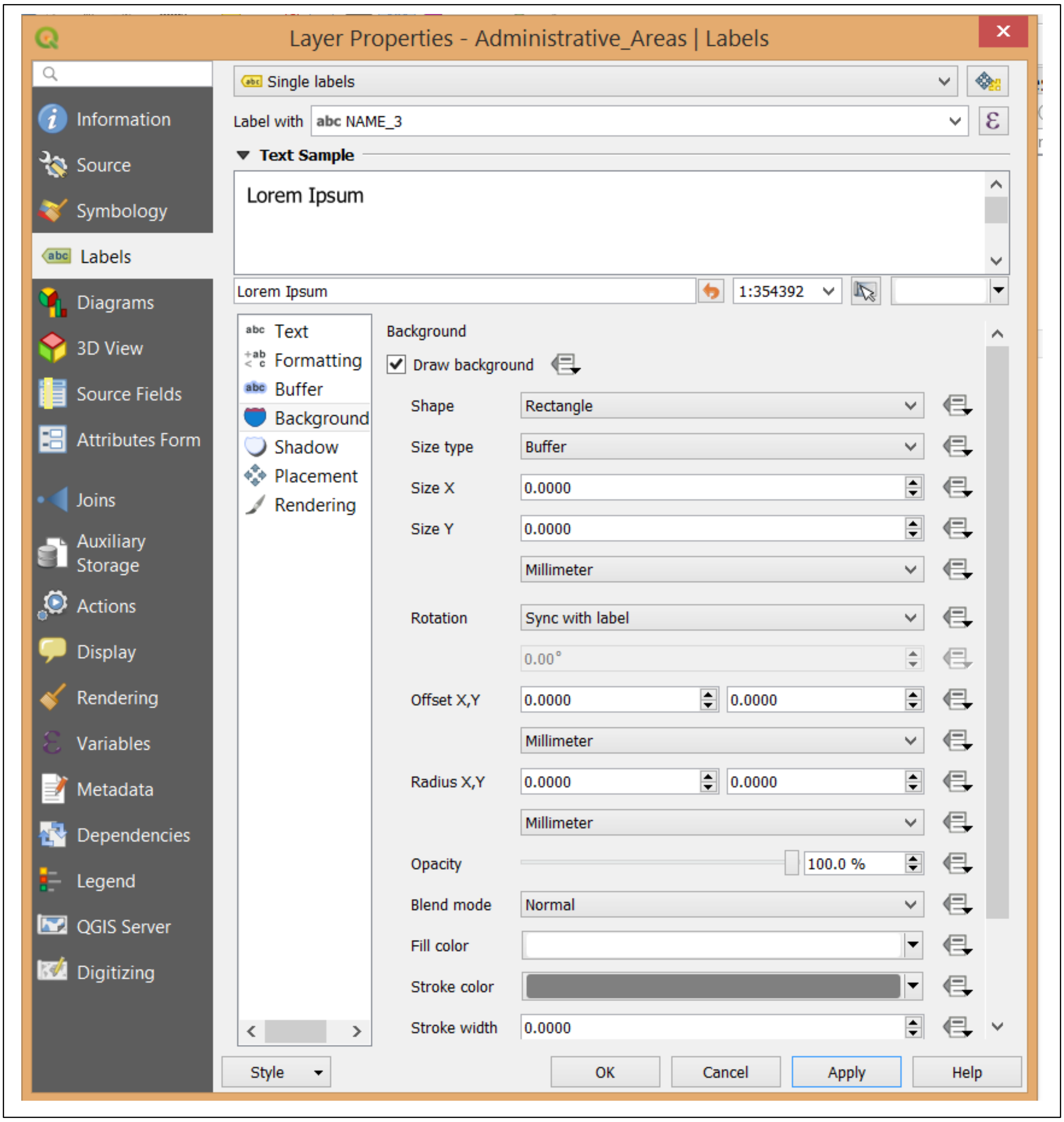


Figure 22: Formatting the texts used for the labels in Antananarivo

Map design and semiology if geographical data is too large to fit the scope of this workshop. However, interested participants can follow on their own one of the many tutorials available online. A more detailed tutorial on QGIS' Map Composer to make a good map can be followed[6]. Access the Layout Manager by clicking <<Project-Layout Manager>>. Then give a name to the <<print layout>>. The map composer launches. In the map composer, Add a new map to the layout – then drag a square shape on the display to load the data layers. <<Move item or Content >>– enables you to drag your map, zoom in/out. Double click outside the highlighted area to escape the command

[6] https://www.qgistutorials.com/en/docs/making_a_map.html

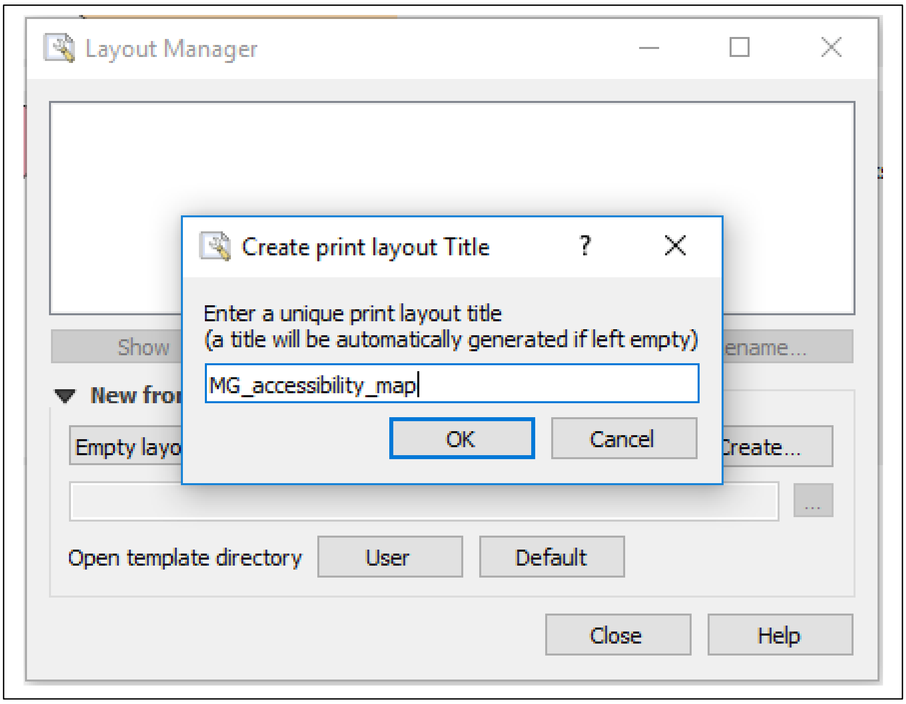


Figure 23: Accessing the Layout Manager

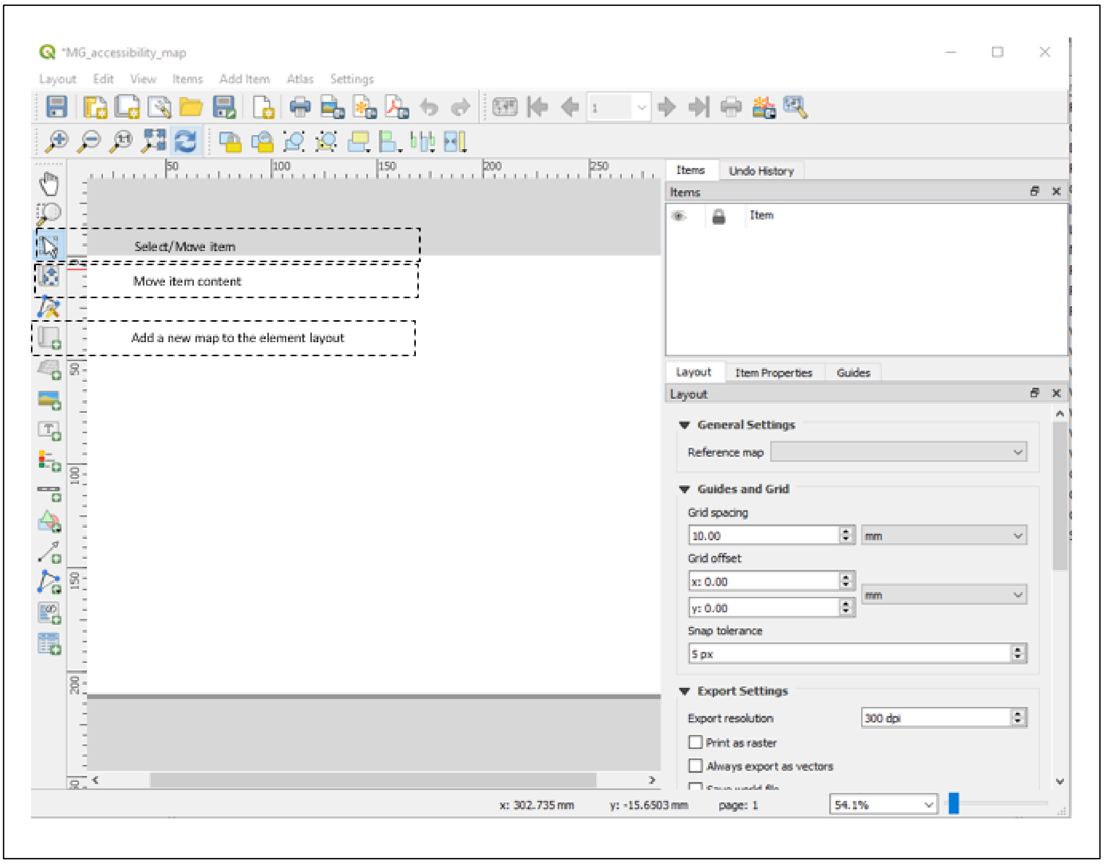


Figure 24: Important commands for manipulating the map in the QGIS Map Composer

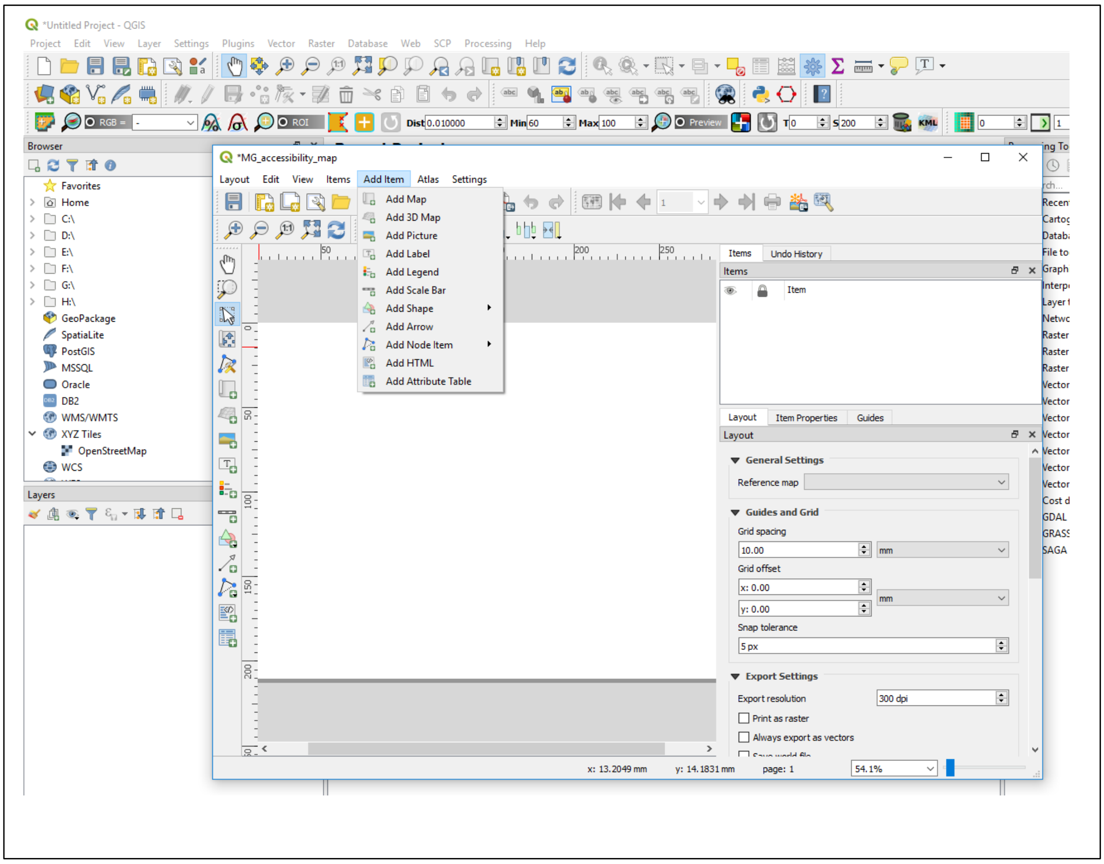


Figure 25: Showing the "Add Items" command for adding the legend, scale and other map elements

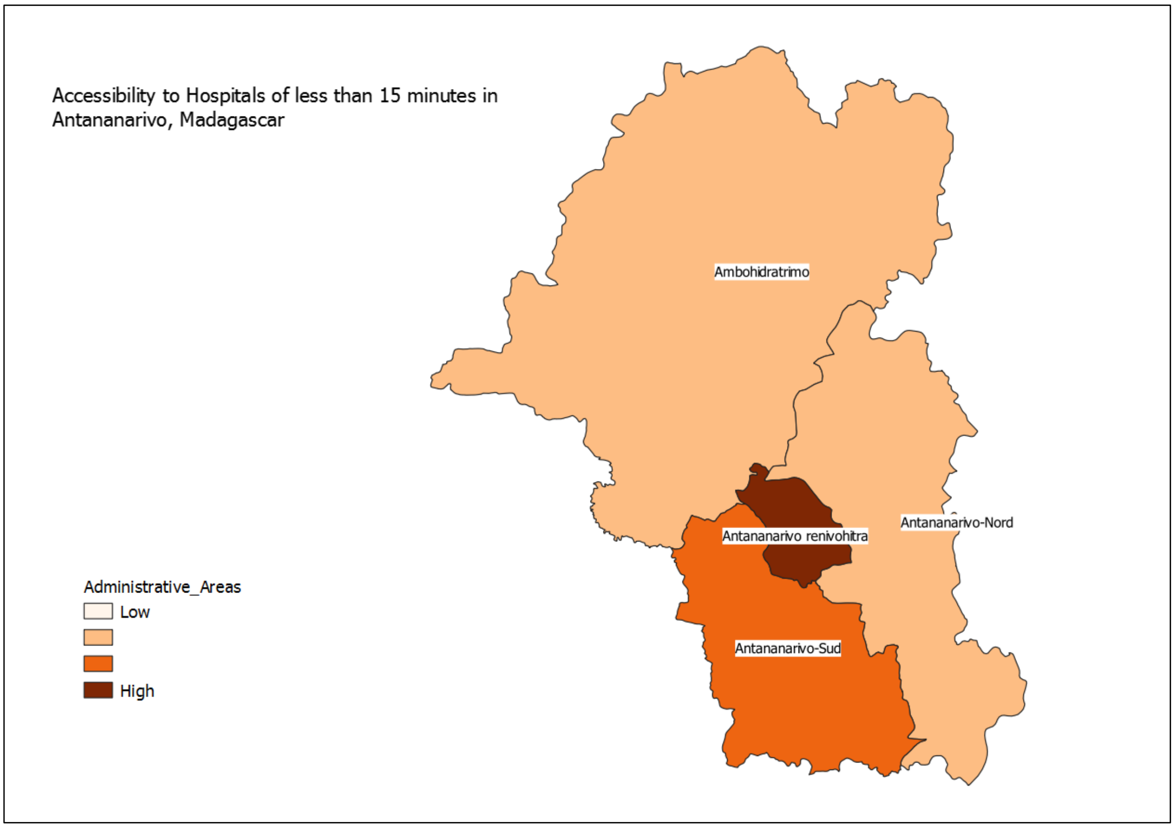


Figure 26: A map showing the ratio of the population that has accessibility of hospitals of less than 15 minutes to Hospitals in Antananarivo, Madagascar

2.1.6 Conclusion

A map of the ratio of the population that has accessibility of less than 15 minutes to hospitals in Antananarivo, Madagascar has been prepared. As expected, there is a high accessibility because it is a capital city. However, the situation might be quite different in the rural areas which are characterised by fewer hospitals and poor road networks. Indeed, this kind of analysis can influence decision and policy makers in construction and improving accessibility of hospitals to the people, thereby improving their livelihoods.

2.2 To determine the ratio of the population that is at risk from flooding in Dakar, Senegal

2.2.1 Background

The second exercise explores the concept of population at risk of flooding. Flooding is a complex phenomenon with many variables to be considered as is the case in (Ndiaye et al., 2016). But in this exercise, we restrict our focus on the contribution of elevation to flood risk. Essentially, we make use of a digital elevation model obtained from SRTM that has a spatial resolution of 30m. A mask is created on the digital elevation model whereby areas that have an elevation of less than five meters are considered to have high susceptibility of flooding. Next, the number of people per pixel that are within the flood risk areas are determined. Similarly, the total population per pixel for each administrative unit are extracted. Finally, the proportion of people at risk of flooding is then computed for each administrative unit. We envisage that this exercise will prove slightly more challenging but interesting. The participants are invited to try out the exercise and pose questions when in doubt.

2.2.2 Data

The data made use in this exercise is:

a) Elevation data
b) Population data
c) Administrative units

2.2.3 Software

> Q-GIS 3.4.1-Madeira

2.2.4 Methodology

The following steps may be followed for the execution of this task:

a) Launch QGIS 3 from your windows explorer.
b) In the QGIS browser, navigate to the folder containing data for the exercise***…\MAUPP_WORKSHOP_NAIROBI_2019\Workshop\Dakar***
c) Add the data:
- The level 3 administrative data
- The DEM
- The population layer.

d) Aligning the rasters. Now, since the population raster and the DEM have mismatching spatial resolution, we perform a resampling of the DEM to that of the population raster. The <<Align Rasters>> tool is used. To access the tool from the QGIS toolbar, click on <<Rasters-Align Rasters>>. Remember to name the output raster files for the aligned rasters. You can load the new rasters which have now a corresponding spatial resolution.

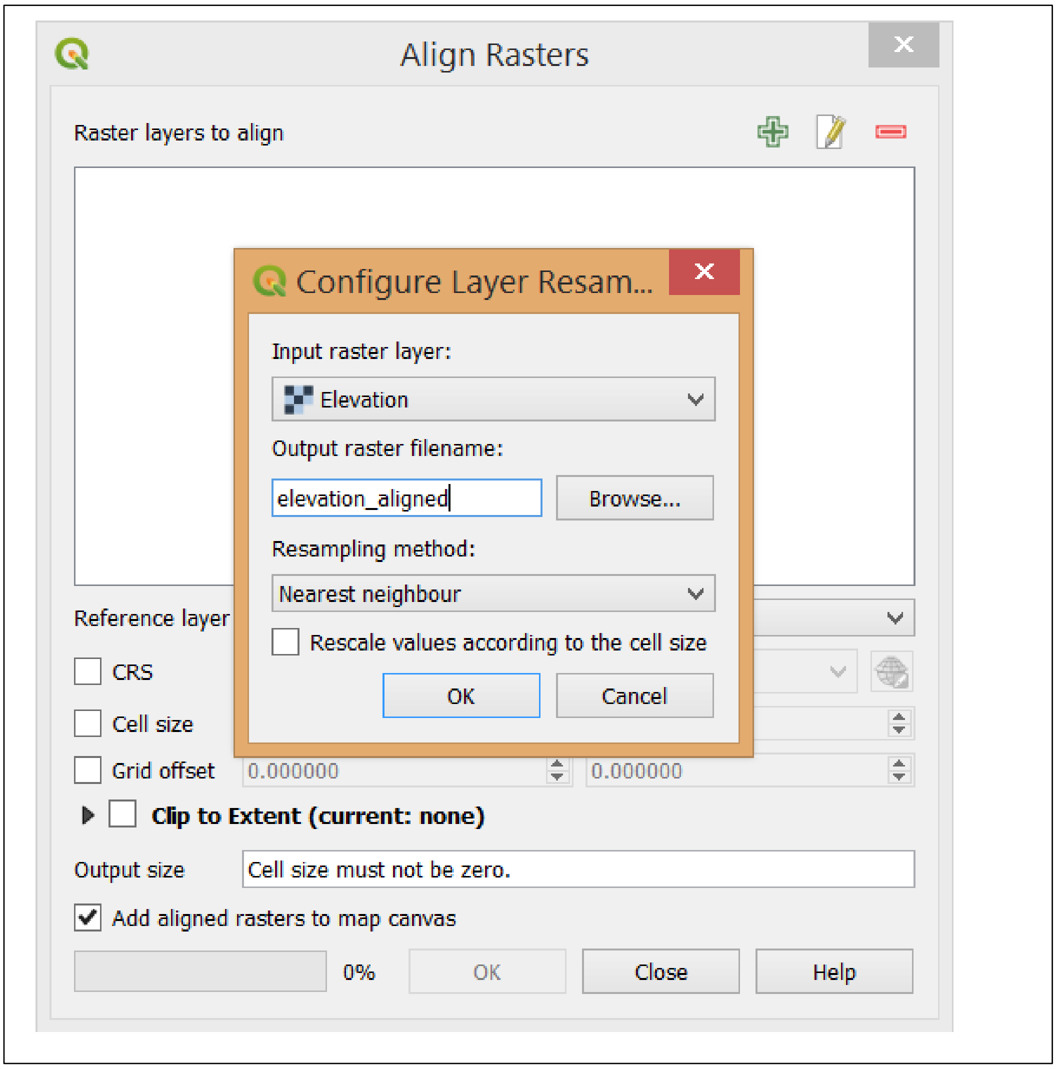


Figure 27: The "Align Rasters" dialogue box

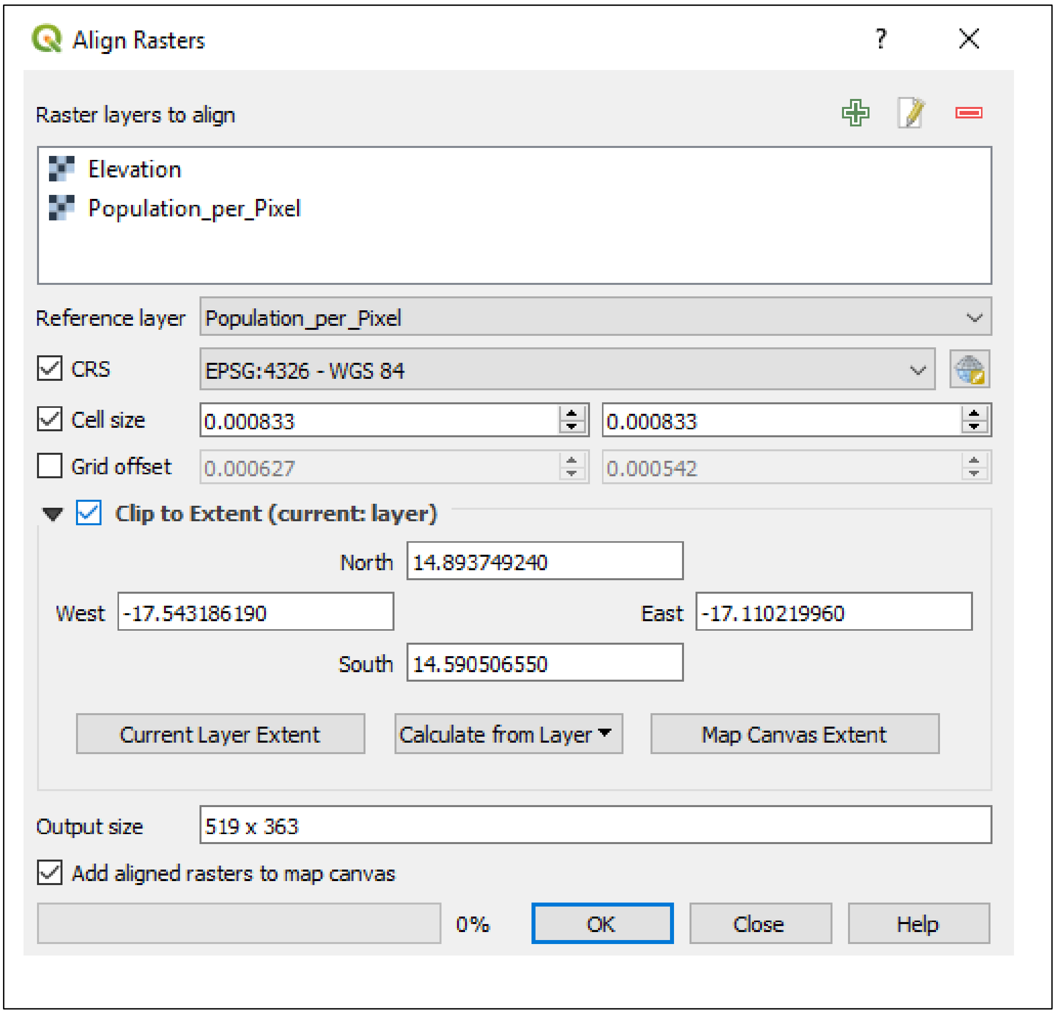


Figure 28: The "Align Rasters" dialogue box showing the used parameters

e) Next, it is useful to assign <<nodata>> values to the background pixels. This step is useful because it ignores the background pixels from the raster calculations. An easy step is to specify the nodata value in the <<Layer Properties-Transparency>>.

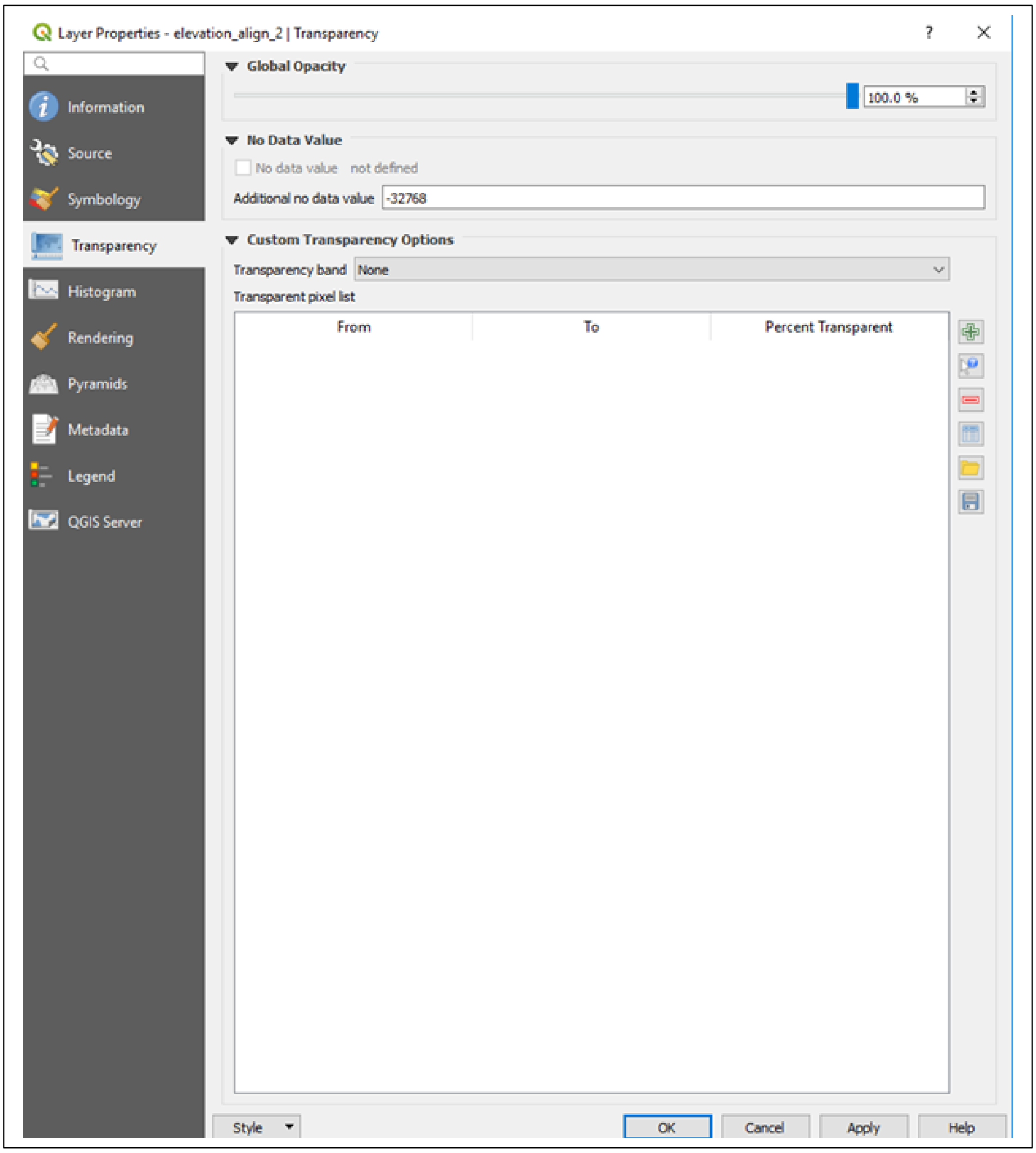


Figure 29:Specifying the null value using the Transparency option of the layer properties

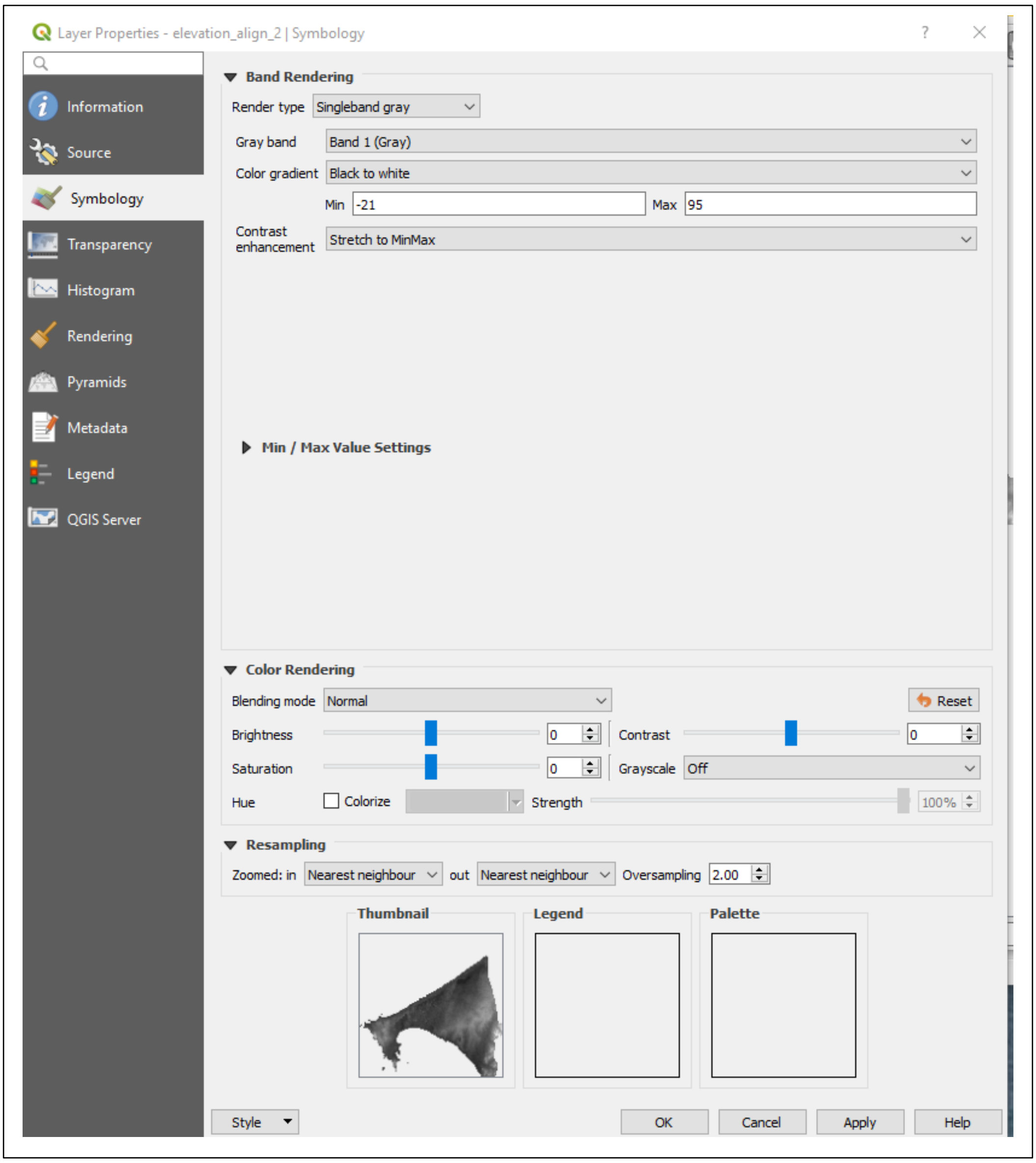


Figure 30: viewing the new range of pixels values after eliminating the null values

Alternatively, you can determine the background pixels by clicking on the region bordering the map. Extreme values, in our case, -999.534 (for population) and -32768 for the DEM are observed. Assign these values to nodata. We make use of the <<r.null>> . Access the tool from the QGIS toolbar by clicking on <<processing-proceesing toolbox-r.null>>

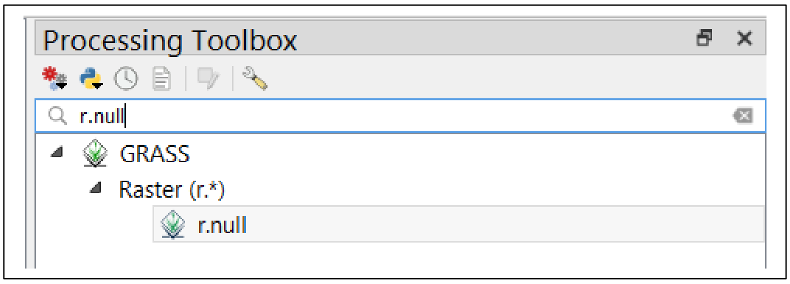


Figure 31: Accessing the "r.null" tool from the <<Processing Toolbox>> used to create nodata values

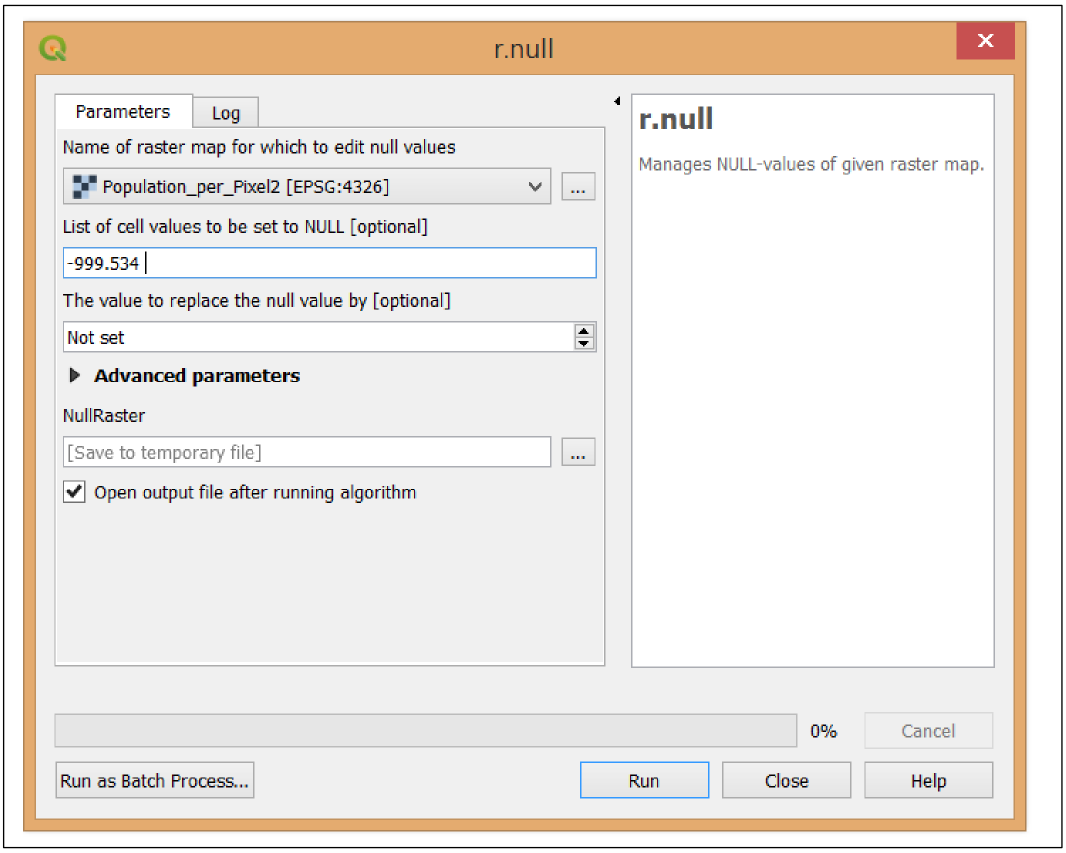


Figure 32: "r.null" dialogue box for the population raster of Dakar, Senegal

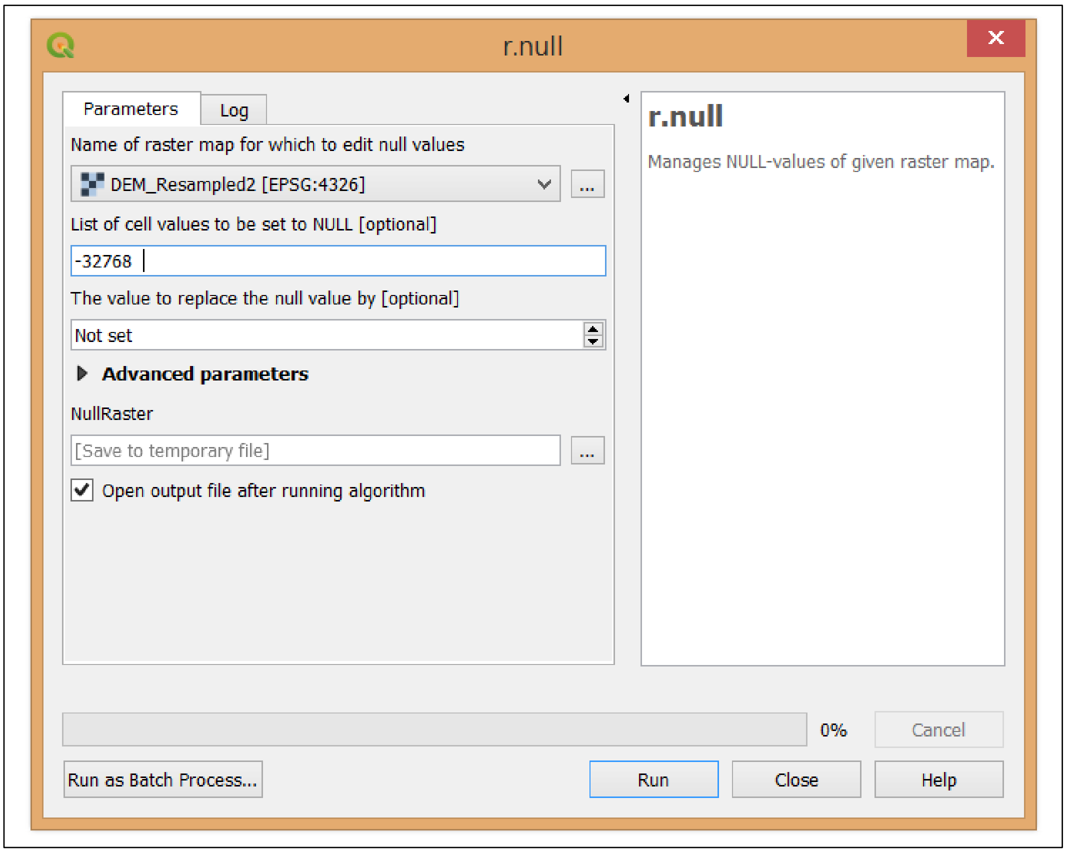


Figure 33: "r.null" dialogue box for the Dakar digital elevation model

f) This step involves the creation of a flood risk mask from the elevation layer. A threshold is applied such that the areas that are less than five meters are at a flood risk.

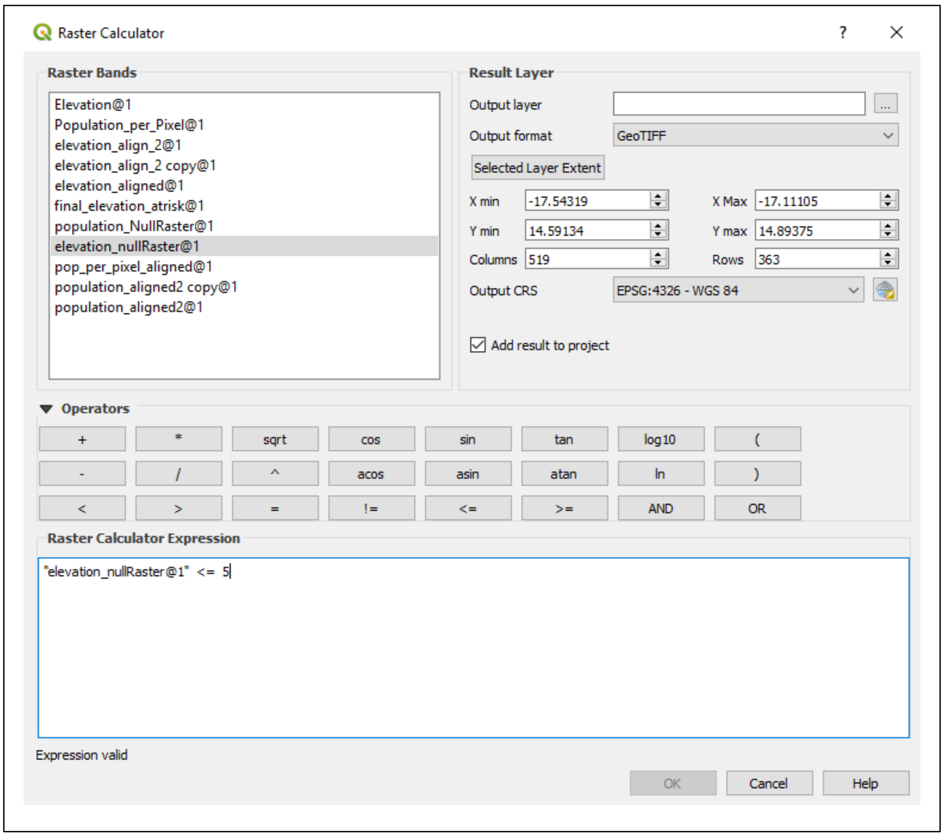


Figure 34: Creating a mask where pixels with an elevation less than 5m are categorised as flood risk pixels from the Dakar elevation raster

Next, we perform a raster multiplication to obtain a population layer that is at risk of flooding

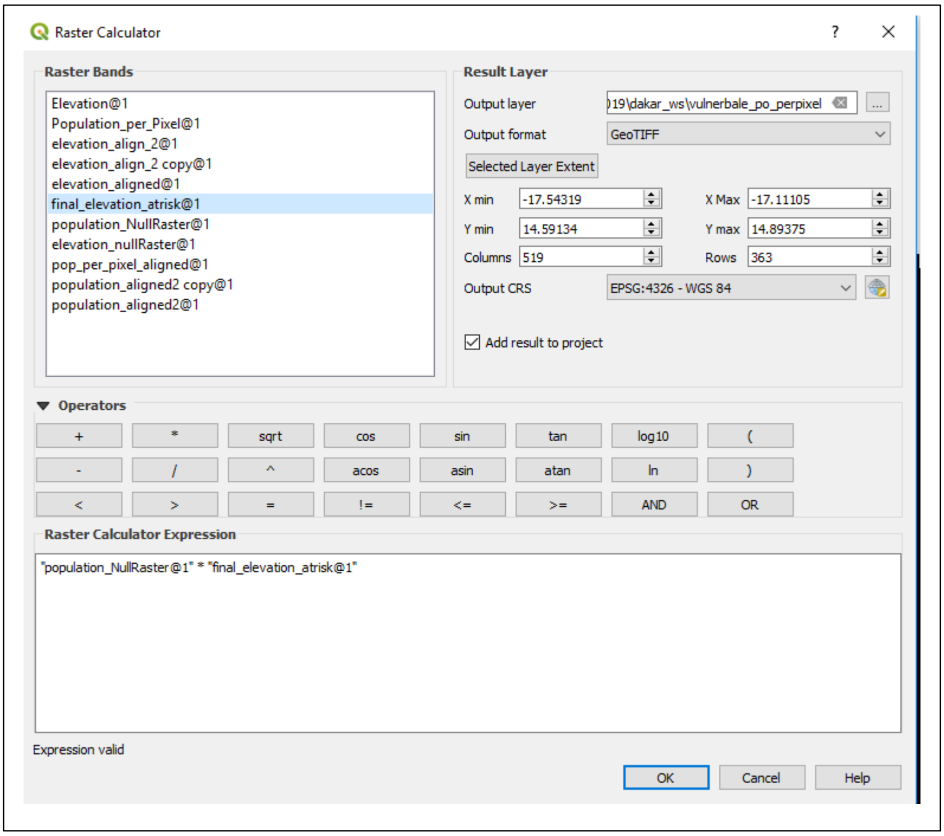


Figure 35: Multiplication of the population raster and the masked elevation layer to determine the pixels in the population raster at risk of flooding in Dakar, Senegal

g) The next step involves the extraction of population at risk and the total population for each administrative unit using the Zonal statistics tool. For this exercise, we are interested with the <<sum>> statistic.

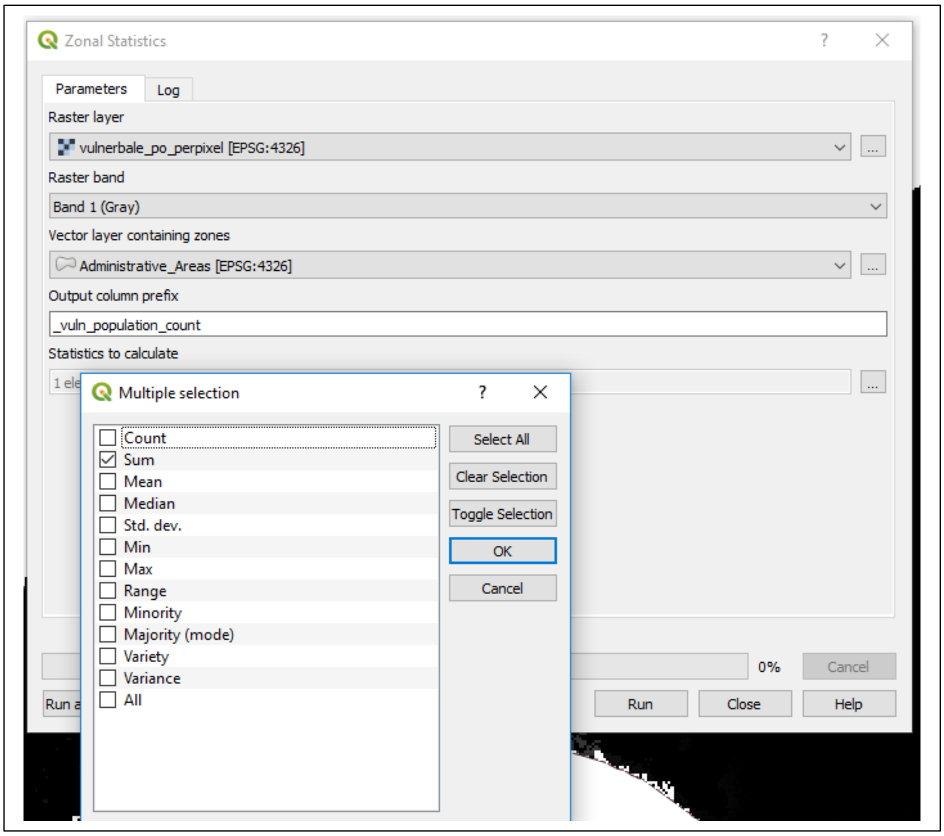


Figure 36: Extraction of the population at risk of flooding in each administrative unit in Dakar, Senegal

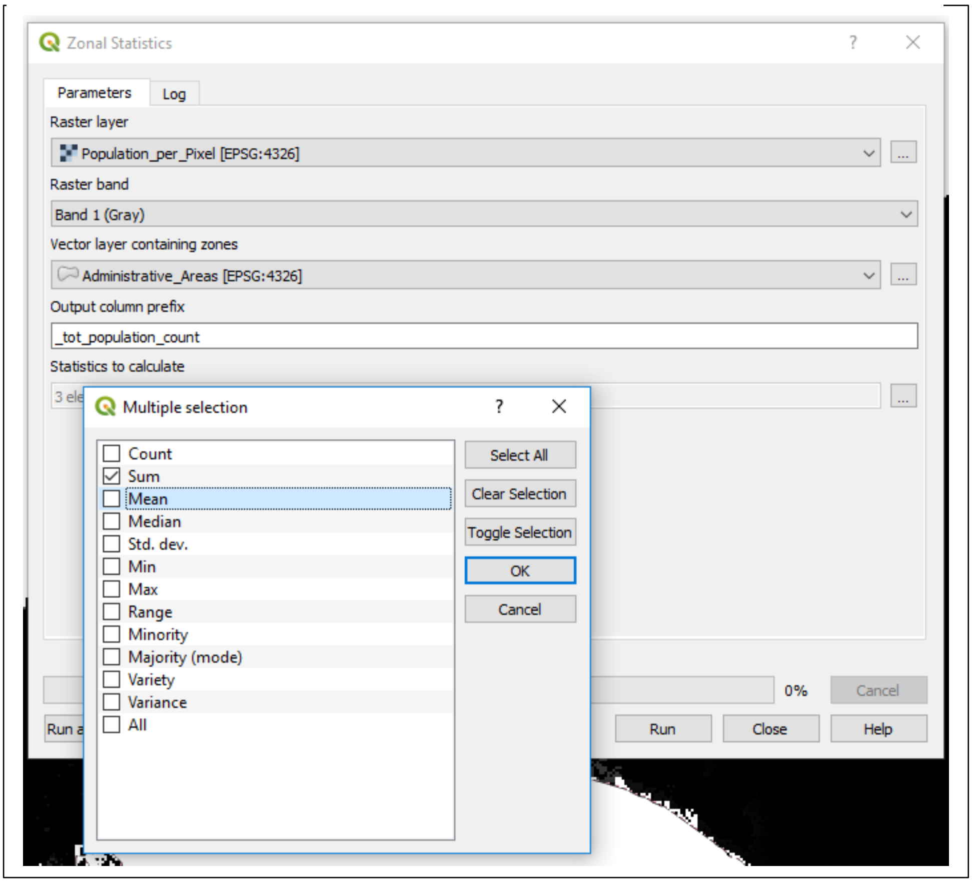


Figure 37: Extracting the total population for each administrative unit in Dakar, Senegal

	_tot_population_countsum	_vuln_population_countsum
1	234788.90577030182	19183.450340270996
2	539231.6595821381	103521.19355010986
3	483849.79023742676	13583.037769317627
4	452601.9633407593	44005.64831542969
5	258851.70512604713	37391.09171259403
6	308745.1742992401	8797.02336883545
7	373867.49364089966	7258.238965988159
8	380448.1623764038	13737.424074172974
9	163138.50786691904	15083.27917933464
10	303360.20456027985	33174.79956674576

Figure 38: An open attribute table showing the extracted population counts. The total population is indicated by "_tot_population_countsum" and the population at risk of flooding is given by "_vuln_population_countsum"

h) Lastly, we launch the field calculator to compute the ratio of the population that is at risk of floods in Dakar.

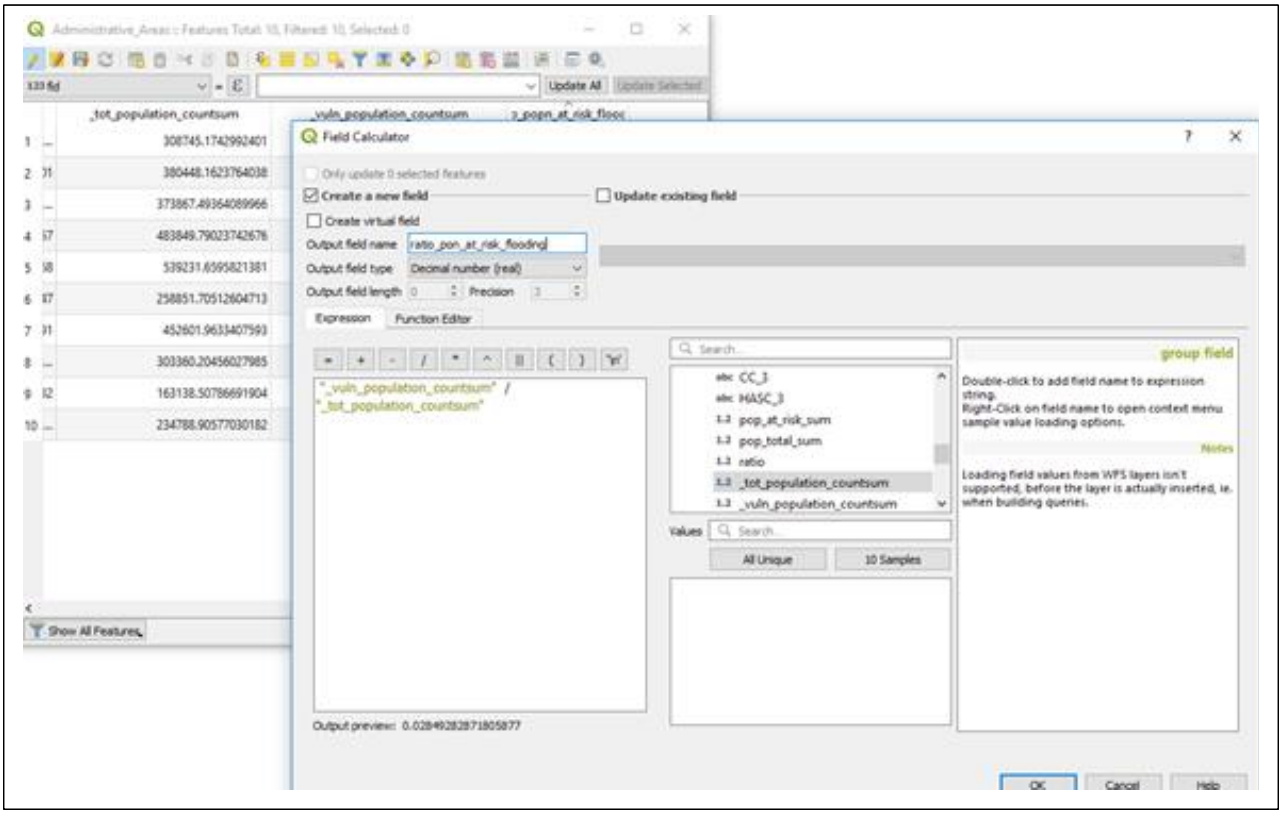


Figure 39: Using the "field calculator" to compute the ratio of the population that is at a risk of flooding in each administrative unit in Dakar, Senegal

2.2.5 Results and Display

A map is created to visualize the results using appropriate symbology.

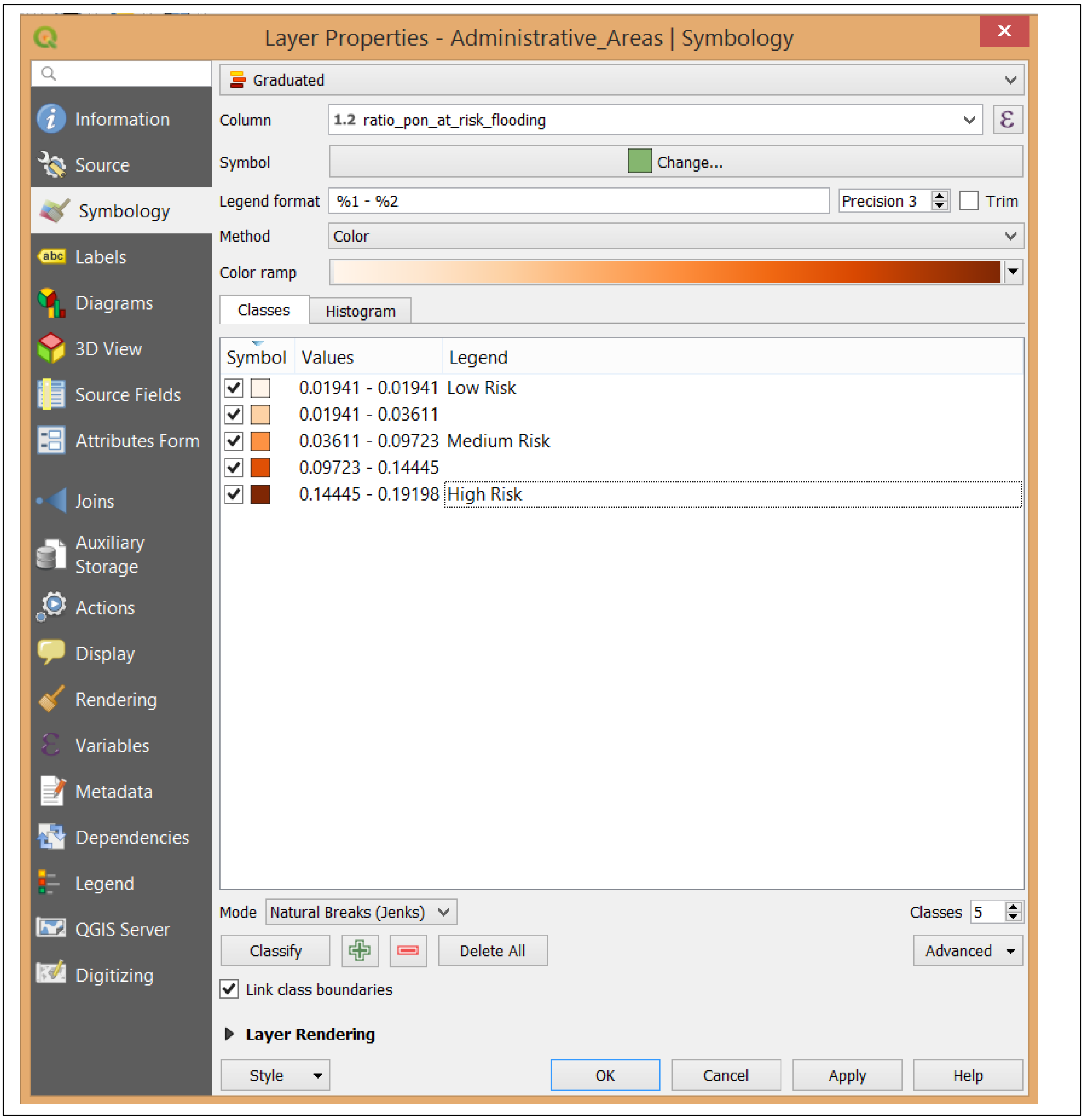


Figure 40: The Layer properties manager - changing the symbology to be based on the ratio of population at risk of flooding in Dakar, Senegal

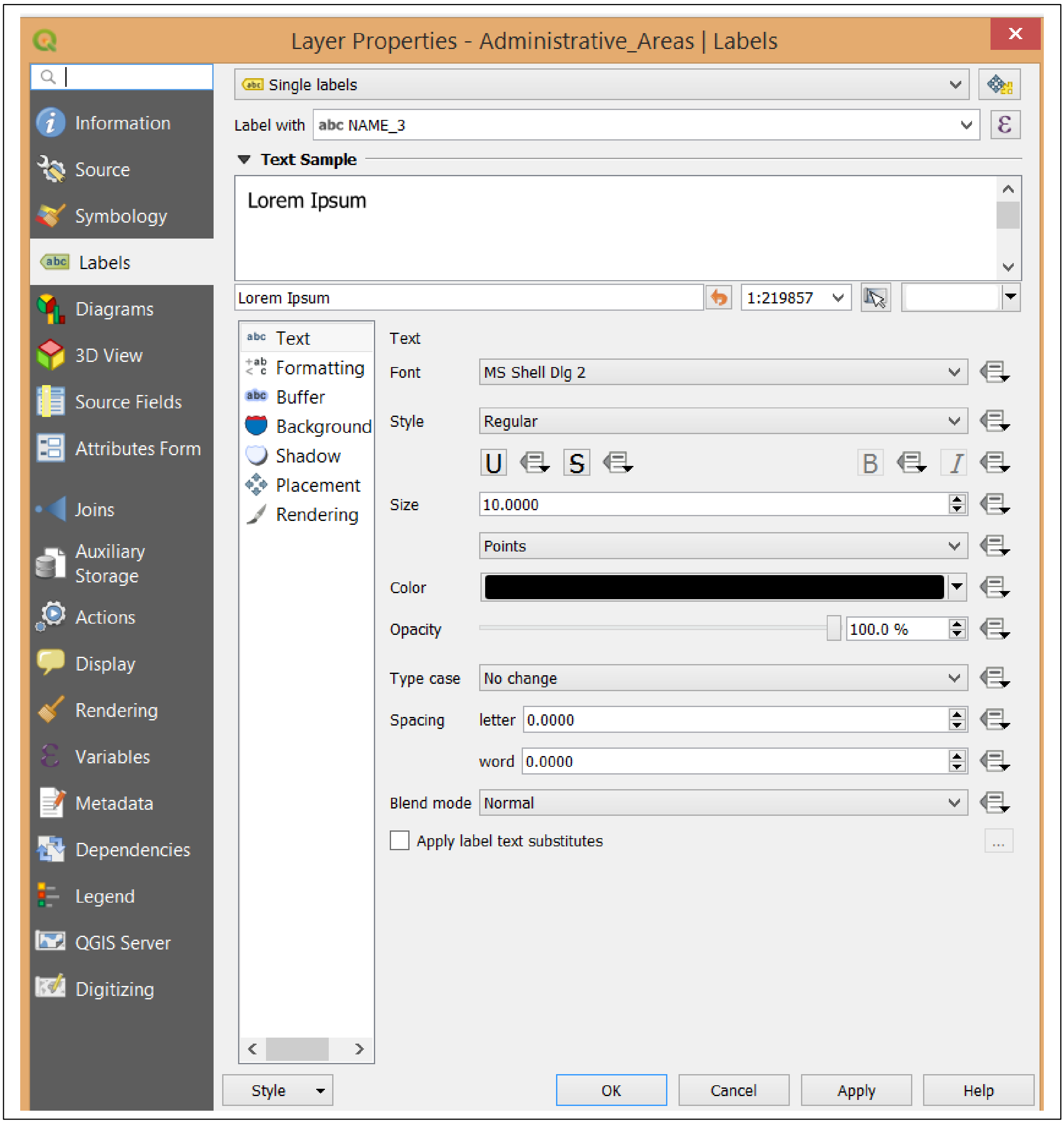


Figure 41: Layer Properties - Modifying the text properties of the labels for Dakar

A more detailed tutorial on Map Composer to make a good map can be followed[7]. Access the Layout Manager by clicking <<Project-Layout Manager>>. Then give a name to the <<print layout>>

[7] https://www.qgistutorials.com/en/docs/making_a_map.html

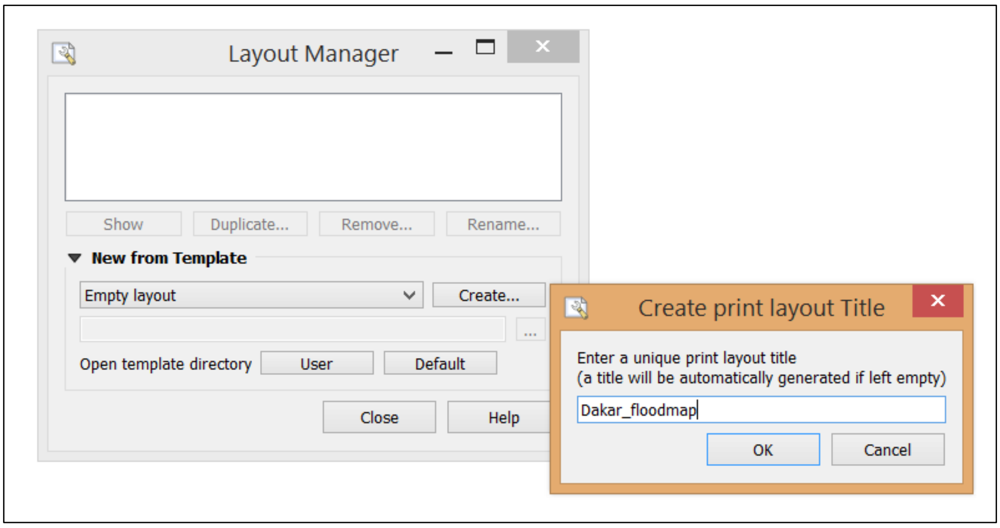


Figure 42: Layout Manager for creating the Dakar Map

Since the computed ratio of population at risk of flooding is a ration number, a graduated symbol is used with Natural breaks applied to the data. This is of course subjective but there are books that discuss ways of displaying population data.

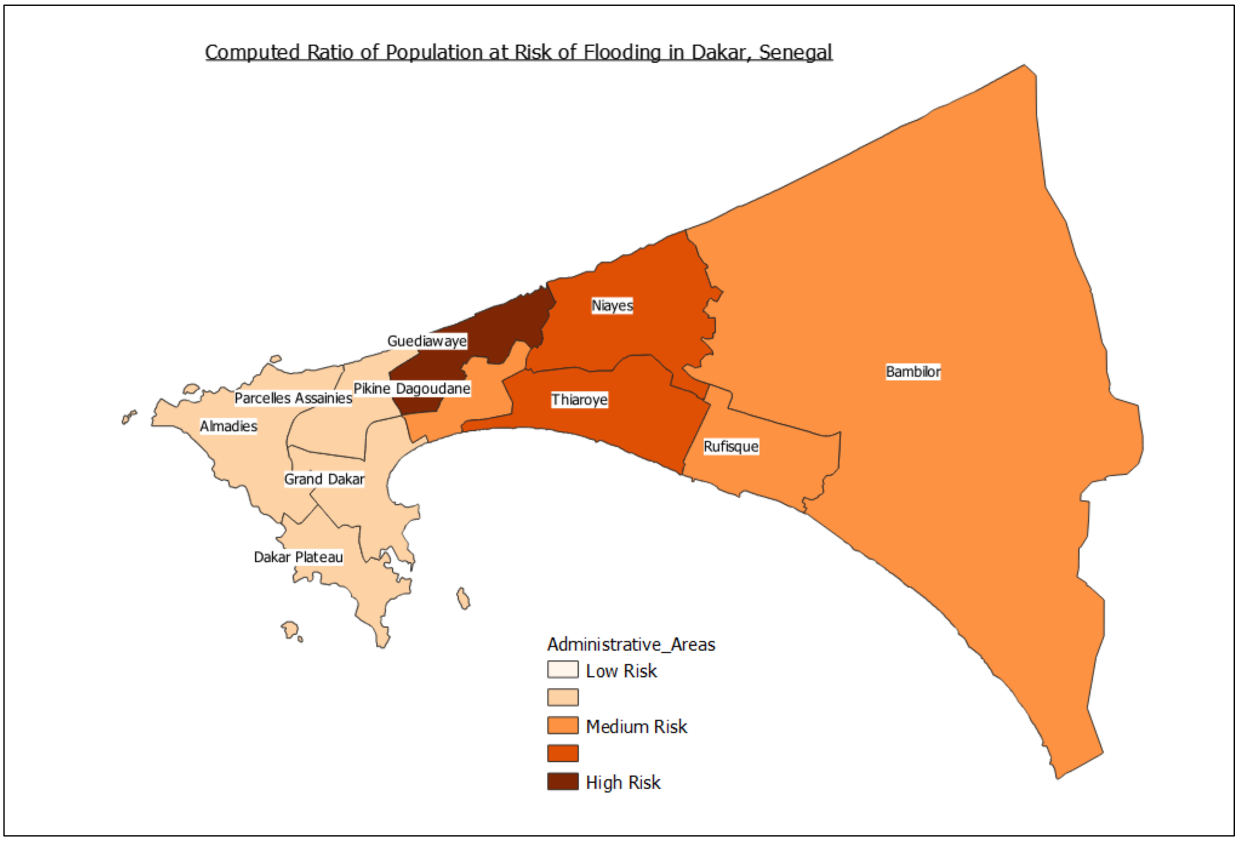


Figure 43: A map of the ratio of population that is at risk of flooding in Dakar, Senegal.

2.2.6 Conclusions

The administrative units with a high risk of flooding is Gueidiawaye according to the analysis carried out. Indeed, the four administrative units of Gueidiawaye, Pikine Dagoudane, Niayes and Thiaroye have experienced notable flooding events in Dakar. Measures are being taken to address flooding in these neighbourhoods. Production of such a map is a significant step towards deployment of mitigation strategies. The WorldBank website[8] describes how Guediawaye and Pikine were flooded in 2012 and how a strategy for flood mitigation and management was set up.

3. Conclusions

This tutorial has demonstrated the utility of open data and open-source geospatial tools. Two different themes of flooding and access to health care have been explored. Step-by-step instructions on how to access WorldPop data and other open data have been described. The methodology described in here is easy to follow, the work by no means exhaustive and indeed can be improved further. Indeed, geoinformation processing is potent in the providing information that can guide decision making and shape policy for the improvement of the welfare of communities.

References

Ndiaye, M. L., Traore, V. B., Toure, M. A., Sambou, A., Diaw, A. T., & Beye, A. C. (2016). Detection and Ranking of Vulnerable Areas to Urban Flooding Using GIS and ASMC (Spatial Analysis multicriteria): A Case Study in Dakar , Senegal. *International Journal of Advanced Engineering, Management and Science*, *2*(8).

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data, 1–22. https://doi.org/10.1371/journal.pone.0107042

Wu, S., Qiu, X., & Wang, L. (2005). Population Estimation Methods in GIS and Remote Sensing: A Review. *GIScience & Remote Sensing*, *42*(1), 80–96. https://doi.org/10.2747/1548-1603.42.1.80

[8] http://www.worldbank.org/en/news/feature/2016/02/03/sustainably-managing-flood-risks-in-dakars-outer-suburbs