



GeoHealth and QuickOSM, two QGIS plugins for health applications

Vincent Herbreteau, Christophe Révillion, Etienne Trimaille

► To cite this version:

Vincent Herbreteau, Christophe Révillion, Etienne Trimaille. GeoHealth and QuickOSM, two QGIS plugins for health applications. Nicolas Baghdadi; Clément Mallet; Mehrez Zribi. Earth Systems - Environmental Sciences : QGIS in Remote Sensing Set, 1, ISTE, pp.257-286, 2018, QGIS and Generic Tools, 978-1-78630-187-1. <<http://www.iste.co.uk/book.php?id=1288>>. <hal-01787435>

HAL Id: hal-01787435

<https://hal.archives-ouvertes.fr/hal-01787435>

Submitted on 7 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapter 7

GeoHealth and QuickOSM, two QGIS plugins for health applications

7.1. Background of the use of GIS for health and the development of dedicated plugins in QGIS

Maps and spatial analyses are increasingly used in the field of health to illustrate the spatial and temporal organization of data and to provide a geographical perspective on health data. Geographic Information Systems (GIS), and geomatics tools in general, are thus used by health organizations (local or international) to inform on the distribution of diseases, by managers to monitor and control epidemics, and by doctors and researchers to understand disease distributions. Scientific publications in the field of health also show the widespread use of spatial maps and analyses. Major organizations such as the World Health Organization, ministries of health in many countries, and research organizations have set up GIS laboratories with geomathicians or health geographers to conduct these tasks. However, many organizations do not have in-house expertise and still wish to conduct such analyses. More and more people in the field of health are thus being trained in geomatics, but implementation remains difficult due to the complexity of the tools and the often sporadic use.

Chapter written by Vincent HERBRETEAU¹, Christophe REVILLION² and Etienne TRIMAILLE^{3,4}.

¹ IRD, UMR Espace-Dev (IRD, UAG, UM, UR), Saint-Pierre, La Réunion, France.

² Univ. Réunion, UMR Espace-Dev (IRD, UAG, UM, UR), Saint-Pierre, La Réunion, France.

³ Kartoza, Le Cap, Afrique du Sud.

⁴ Developer of GeoHealth and QuickOSM plugins.

The goal of the GeoHealth plugin, developed for QGIS software, is therefore to provide a simple and intuitive interface so that non-specialist users (i.e. non-geomaticians) who rarely use GIS softwares can meet the main need to map disease cases. In addition, difficulties in accessing spatial data also limit their use. This concerns the identification of relevant data, access rights, their import and potential pre-processing. A solution is proposed here through the use of OpenStreetMap (OSM), a worldwide free and participatory geographic database. OSM is directly accessible from QGIS through the QuickOSM plugin, which allows simple and specific object queries (for example, download vector files of roads, houses or the location of pharmacies in a given area, but as this data exist in OSM). By using the addresses contained in OSM, institutions can geolocate patients and place them on a map. Spatial analyses can also be carried out to assess the impact of landscape features on disease occurrence.

One of the main criticisms of using OpenStreetMap data is its representativeness. As this database is built by participatory contributions, the data are indeed heterogeneous. In turn, an advantage of OSM is that it can be used as a mapping tool to record data to meet a specific need. Thus, in an area without reference data, it is entirely possible to map by photo-interpretation of satellite images or to integrate field data, which will then be imported via QuickOSM.

Finally, the issue of data confidentiality must be addressed, as it is fundamental to the use of data from a health care system. In order to respect anonymity, health data are often aggregated by administrative unit, leading to a loss of knowledge of the local spatial organization of the data. The GeoHealth plugin then proposes a method of approximating point locations, known as [blurring](#) ("Blur" function in the extension), with a user-defined accuracy (for example, the indication of locations to within 1 km). To anonymize the location of the patient's place of residence, a quality control of the blurring can be carried out in the plugin, by counting the number of houses in the blurred areas. This method can be easily applied to point epidemiological data in order to preserve the ability of detecting clusters and analyzing the surrounding environment.

7.2. Methodology applied

The first need, in terms of mapping applied to health issues, is the production of distribution maps of disease cases. These cases are usually represented by a rate: the ratio of the number of cases to the population on a given administrative scale (country, district, village, etc.). In epidemiology, we refer to the [incidence](#) of a disease (or morbidity rate) as the number of new cases of disease that occur over a given period of time in a given population. We also refer to the prevalence of a disease as the number of cases among a population at a given time or over a given period of time (cases that could have been diagnosed before that time) [PIC 01]. A

variation of the representation of these incidence maps could be a density map where the number of cases is represented in relation to the surface area of each spatial unit.

The GeoHealth plugin was originally designed to provide an intuitive interface for importing an epidemiological data table and map its incidence, prevalence or density. It allows importing data in various formats and then defining the parameters of the calculation of incidence and symbology in a single window. The plugin also proposes a novel method for anonymizing the location of data (blurring as defined above). The interface has three main tabs: Import, Analysis and Export (figure 7.1), and, for each function, a side menu to guide the user through a sequential process.

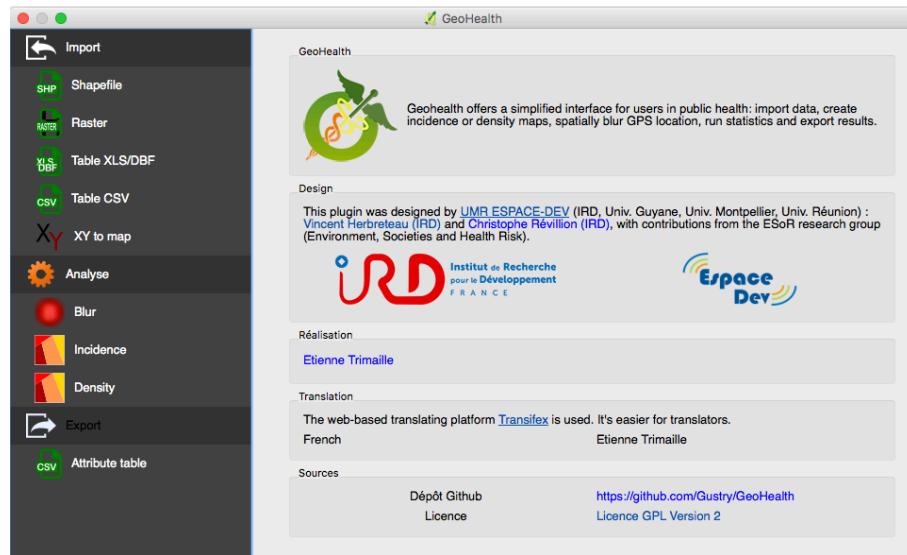


Figure 7.1. GeoHealth plugin home window

7.2.1. Importing data

The data to be mapped is usually in the form of a table. A common example is a table from a registry of patients who have been diagnosed with a disease in a year as recorded by health facilities (hospitals, clinics, etc.) and then synthesized by health managers (ministries of health, health agencies for disease surveillance and control). It can also be a table resulting from health surveys within the framework of active surveillance (research for a disease in a population, search for pathogens in an animal population, search for infectious agents or pollutants in the environment).

Although importing tables is easy in QGIS, occasional users may have difficulties to find the right function. Thus, it is difficult to remember that you can import an Excel file (.xls format) via the "Add a vector layer" button. The interface therefore directly proposes different types of files that can be imported: Shapefile, Raster, Table XLS/DBF, Table CSV, XY to map. This last function (XY to map) corresponds to the QGIS dialog box "Add a delimited text layer" which allows to map a table in csv format that includes latitudes and longitude columns.

QGIS features:

- Data import: *Plugins > GeoHealth > Import.*

7.2.2. Case geolocation

Geolocation is a crucial step that aims to locate epidemiological data on a map by giving them a spatial reference. There are three possible cases:

- the data table contains for each case the reference (name or identifier of the administrative unit. It can then be associated with a vector layer of the corresponding administrative units;
- the data table contains geographic coordinates (latitude and longitude). It is then possible to import the table via the button "XY to map" and indicate the corresponding columns and the projection;
- the data table contains postal addresses for each row. These addresses must then be geocoded using national or international address databases (for instance, using Google web services or Nominatim, a search engine for OpenStreetMap data: <https://nominatim.openstreetmap.org/>).

QGIS features:

- Creating a map from a table containing the latitude and longitude columns: *Plugins > GeoHealth > Import > XY to map.*
- Geocoding a file containing addresses: *Plugins > GeoCoding > GeoCoding.*

After importing data, the second tab "Analyse" of the GeoHealth plugin proposes "Blur", "Incidence" and "Density" tools.

7.2.3. Data anonymization ("*blurring*")

Individual data such as epidemiological data could only be used in accordance with anonymity and, where appropriate, in agreement with the data protection authorities. Aggregation of individual data on an administrative scale is often sufficient to ensure anonymity and allow data dissemination. However, such aggregation leads to a loss of knowledge of the surrounding environment of each case. The GeoHealth plugin therefore offers an alternative method for blurring point data, implemented in the "Blur" tool. This tool has been designed for healthcare professionals to provide a simple method for blurring data prior to its dissemination for mapping and analysis.

The implemented blurring method consists of three steps (figure 7.2):

- create a buffer zone of a user-defined radius r around the cases locations;
- randomly draw a new point in each buffer zone;
- create a second buffer zone of the same radius r which will be the blurred zone. These new circular areas are the ones that must guarantee anonymity of the data in order to be transmitted. They necessarily include the initial point that cannot be retrieved (following the random draw). Any point in these blurred areas is at a maximum distance of 2 times the radius ($2r$), which indicates the resolution of the blurring.

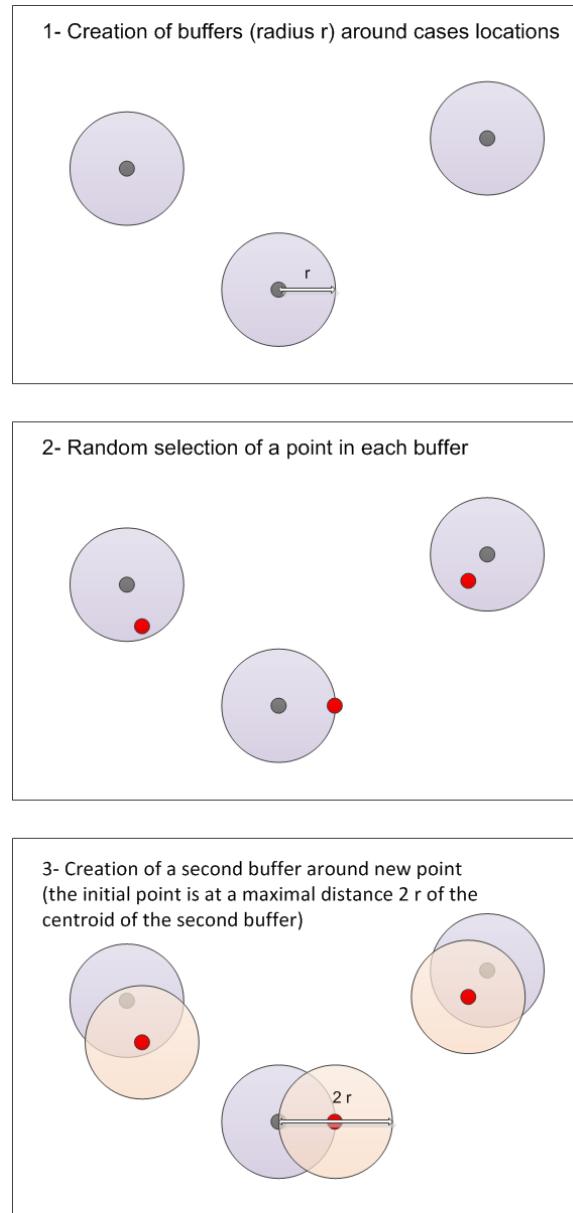


Figure 7.2. Blurring process of point data implemented in the "Blur" tool

However, anonymity could not be effective if, for example, there is a single house in one blurred area. The blurring tool therefore proposes a test ("Stats" tab of the Blur tool window) to check the number of objects (here, the houses) in each blurred zone. This test is based on the use of a vector layer of buildings. It provides descriptive statistics including the minimum number of houses in all blurred areas. This test is proposed so that the user can assess the quality of the blurring and decide if the data can be disseminated. If this number is not satisfactory, the user can blur the data again by increasing the radius size.

QGIS features:

- Blurring point data: *Plugins > GeoHealth > Analyse > Blur*.

7.2.4. Creating an incidence or density map

The Incidence and Density tools in the Analysis menu are very similar in their design and objectives: specify the calculation settings (choice of layers and columns, indication of the ratio, name of the column to be created, name of the output file), choose the method of discretization and symbology of the map and display descriptive statistics. Two tabs are proposed according to the type of data input: a polygon layer containing the number of cases (generally per administrative unit) or a point layer of case location. The tool performs the calculation directly and adds the result to a new column of the attribute table, avoiding the user from switching to the edition mode to calculate new fields. The choice of symbology in the tool allows a fast display of the result, but does not prevent the user from fine-tuning the graphical representation.

QGIS features:

- Creating an incidence map: *Plugins > GeoHealth > Analyse > Incidence*.
- Creating a density map: *Plugins > GeoHealth > Analyse > Density*.

7.2.5. Importing data from the OpenStreetMap database

In a context of massive data availability and a willingness to release some of them, spatial data are becoming increasingly accessible. Among the outstanding projects, the OpenStreetMap Geographic Database (OSM) is a must-have resource for spatial health analysis. This global and open database, initiated in 2004, is enriched in a participative approach by the photo-interpretation of satellite images made available free of charge, by the integration of free datasets and also by field observations via dedicated applications on mobile devices (smartphones and tablets). A common way to contribute to OSM is to connect to its website (www.openstreetmap.org/), get identified (since each contribution is nominative) and then to draw objects (buildings, roads, points of interest, etc.) by photo-

interpretation. In this case, the images made available to OSM will be used. By default, the OSM website displays Bing images that have been released for this purpose. Mapbox and DigitalGlobe companies also provide very high spatial resolution images for improving OSM database. They chose images in places, where high-resolution images were missing and where active contributors are registered in order to have, in turn, better mapping. Such images are also regularly provided when mapping is needed for humanitarian actions. Depending on the country, other accesses are provided. Finally, it is entirely possible to use one's own source of very high spatial resolution images (from drones or from satellites, such as Pleiades or Spot 6-7), acquired as part of a project, for example. Only project members will then use the image but the contributions to OSM will be accessible to all.

Some organizations, such as Humanitarian OpenStreetMap Team (HOT: hotosm.org) and the NGO CartONG (cartong.org) were created to provide mapping support for humanitarian interventions. They coordinate the mapping tasks by getting people involved around the world through the Internet. Thus, in January 2010, a few hours after the heavy earthquake that hit Haiti (magnitude 7.0), the OSM community get mobilized. High spatial resolution satellite images showing the extent of the damage were provided within 48 hours and more than 600 people mapped the area in the first few months. OSM is regularly used for disease control projects. In 2016 and 2017, HOT and DigitalGlobe partnered to support the Clinton Health Access Initiative's malaria program by making satellite imagery available to accelerate mapping of houses and roads in endemic areas with little available geographic data. Another initiative based on the OSM project is the HealthSites website (<https://healthsites.io/>) which aims at creating a global database of health facilities. The dedicated interface helps to complete this data.

The OSM database is especially interesting for spatial analysis since it can be easily queried and directly imported in QGIS. This import is implemented in the QuickOSM plugin by indicating the object categories to be imported (based on OSM keys/values, e.g."Building/yes" for buildings), the spatial extent and the type of data (lines, nodes, relationships). This import produces a vector file that can be directly saved and analyzed in QGIS.

This query-based import of OSM data directly through the QuickOSM plugin in QGIS is a real advantage for the use of OSM. Indeed, other methods of importing OSM are generally done on the whole dataset over a given range (country or region). These queries allow a lighter and faster download, and easier use since only useful data is acquired.

In addition to using the data available in the OSM database, a user can also build their own data in the OSM database to import and then use them in QGIS. It could be:

- on a small study area with no data: a few days of work will be enough to map a small area (*i.e.* a few villages);
- on a large study area: it will then be necessary to involve many people by organising a mapping party and to coordinate this work as it can be done on HOT;
- by adding vectorial data. For example, a health administration can integrate data on health facilities into OSM. It is also possible to integrate free vector datasets and remotely-sensed information.

Given this potential, its daily enrichment and its integration into the QGIS software via the QuickOSM extension, the OSM database offers multiple possibilities for spatial analysis in different fields, including health.

QGIS features:

- Importing OpenStreetMap data: *Plugins > QuickOSM*.

7.2.6. Environmental analyses

For studies related to health, QGIS is mainly used for mapping data. However, spatial analyses are increasingly required to understand the spatial heterogeneity of data (distribution and cluster analyses) or the effect of environmental factors on health status. The GeoHealth plugin does not yet offer tools for these advanced uses, which will have to be done with other QGIS tools and plugins. There are many applications and some of them can be described here as examples:

- **Distance calculations.** For a long time, these calculations were limited to direct distances between points. This can be done very quickly in QGIS with the tool "Distance Matrix" which allows calculating distances between points of the same vector layer or between those of two vector layers. For example, this tool can be used to calculate the distance between a localization layer for patients and a localization layer for healthcare facilities. It is then possible to calculate for each patient the minimum distance to a structure. To refine this calculation, it is preferable to use a road network and calculate the theoretical travel time or the shortest distance via roads. The "Road graph plugin" allows such calculation (this plugin is integrated in QGIS but must be checked beforehand in the plugin manager). For this purpose, one can download the road network of a study area using QuickOSM (key/value = highway/yes) and check if the completeness of the network is sufficient for the travel time or distance calculation.
- **Proximity to landscape elements.** The objective of the study could be to test whether proximity to a landscape element is linked to a larger number of cases (for instance, for a case-control study). These landscape elements could be suspected

places of risk (*e.g.* wetlands and rivers for mosquito-borne diseases) or pollution emission sites. Some of these elements could be recovered from OSM. To test the effect of proximity to watercourses, it would first be necessary to import watercourses via QuickOSM (key/value = waterway/yes) and then create buffer zones around them for the distance to be tested. QGIS proposes two tools: "Fixed distance buffer" where the user will indicate the distance and "Variable distance buffer" where this distance can vary depending on a layer variable (for instance, this distance could vary depending on the category of the watercourses). An intersection between these buffer zones and the location of cases can then be performed to statistically compare whether the proportion of cases relative to controls is greater in the buffer zones. Nevertheless, a spatial correlation could be shown, but this is not necessarily a causal factor (the cause may be related to factors other than proximity to watercourses, this proximity would only be an indicator).

QGIS features:

- Distances between points: *Vector > Analysis Tools > Distance matrix*.
- Travelling time: *Vector > Road Graph > Settings...*
- Buffer zones: *Vector > Geoprocessing Tools > Fixed distance buffer* ou *Variable distance buffer*.

7.2.7. Export

Exporting data is easy in QGIS by saving a layer of data and selecting the appropriate format. Casual users who perform spatial analyses in QGIS before running statistical analyses in other software usually need to save the attribute tables only. For this purpose, the GeoHealth plugin offers this export function of the attribute table in a CSV format.

QGIS features:

- Export attribute table: *Plugins > GeoHealth > Export > Attribute table*.

Figure 7.3 summarizes a typical process of spatial analyzes for health studies. Input data (on the left side of the diagram) usually include epidemiological or observational data to be mapped and spatially analyzed. In the frame of this process, input data also include the data sources used to enrich OpenStreetMap (mainly satellite images). The central part of the diagram synthesizes the analyses conducted in QGIS with three plugins (GeoHealth, GeoCoding and QuickOSM). The outputs (on the right-hand side) of such analysis obviously include maps and, more often than not, data tables that can be statistically analyzed or used for dissemination of knowledge.

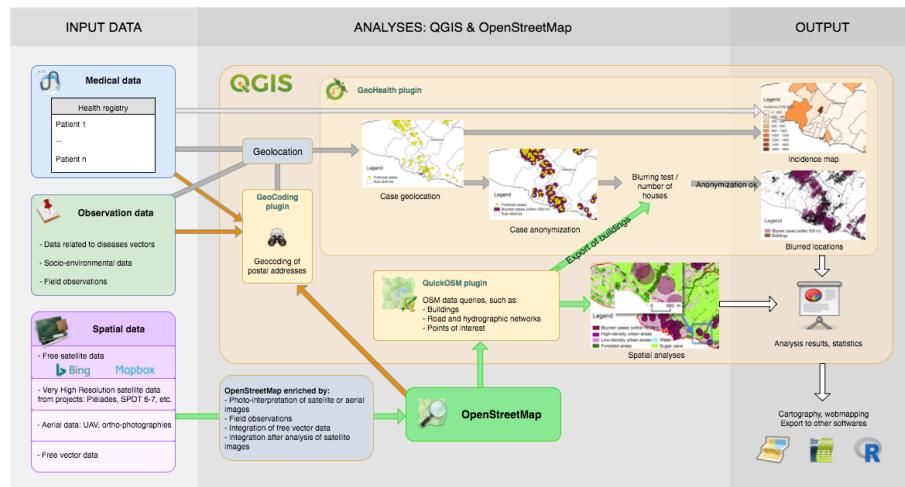


Figure 7.3. Typical example of the use of QGIS software for mapping and spatial analysis of health data

7.3. Implementation: GeoHealth, assisted mapping with QGIS

This tutorial targets a wide audience, from beginners to experienced users of GIS softwares. It presents a series of exercises applied to Reunion Island, a French overseas province in the south-western Indian Ocean.

7.3.1. Installation of GeoHealth and QuickOSM plugins

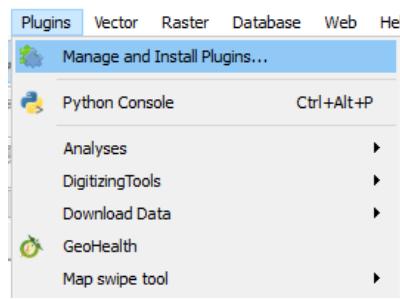


Figure 7.4. Access to the QGIS Plugins Dialog Window

This tutorial was realized with QGIS 2.18 and requires the installation of two plugins: GeoHealth and QuickOSM. They should be installed via the “Plugins Dialog” that can be launched from the Plugins menu (Manage and Install Plugins...). After installation, the GeoHealth plugin is accessible from the Plugin menu and the QuickOSM plugin via the Vector menu or its shortcut in the toolbars. They can also be directly used via the Processing Toolbox.

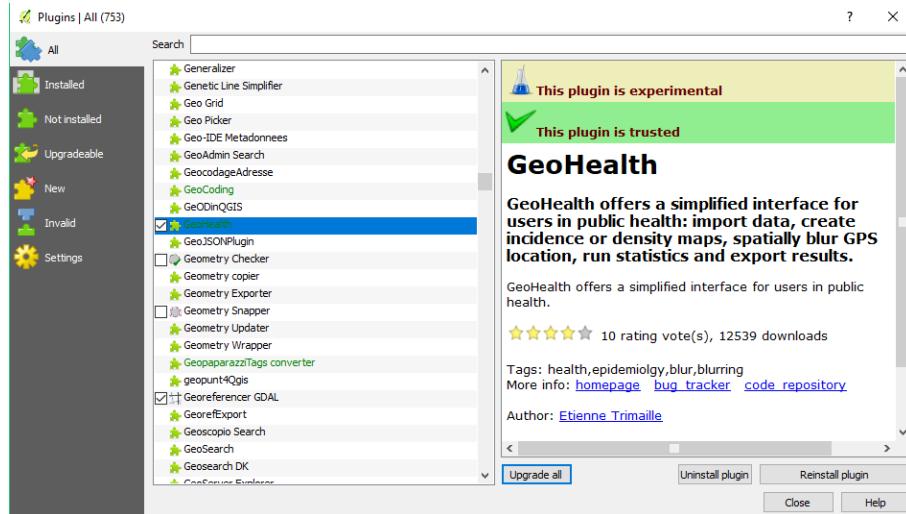


Figure 7.5. QGIS Plugins Dialog Window

7.3.2. Datasets

Sample data are available on the GitHub of the GeoHealth plugin to conduct this exercise:

https://github.com/Gustry/GeoHealth/blob/master/sample_data/geohealth_sample_data_en.zip

It contains:

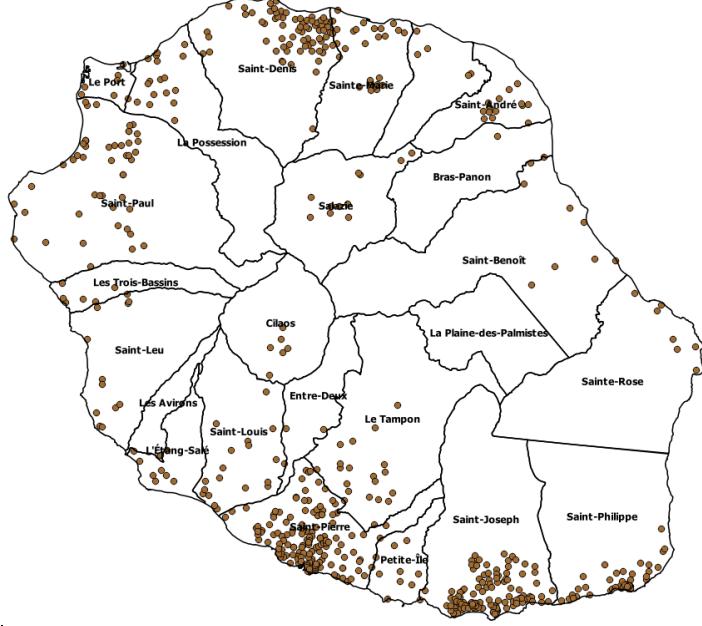
- a point vector file (shapefile) representing fictional cases of two diseases on Reunion Island;
- a table of addresses (in csv format) for testing geocoding (paragraph 7.3.4.1).

The other necessary data will be downloaded from the OpenStreetMap database using the QuickOSM plugin, when necessary.

7.3.3. Producing an incidence map with GeoHealth plugin

Objective: realize an incidence map using point data and administrative borders

Approach	Handling in QGIS
1. Importation of shapefiles	<ul style="list-style-type: none"> • Open GeoHealth in menu: Plugin/GeoHealth. • Select Import. • Select Shapefile, and browse to the point shapefile called « fictional cases ». Once loaded, select Open. The file is displayed in QGIS.
2. Download of administrative units with QuickOSM plugin	<ul style="list-style-type: none"> • Click on the QuickOSM shortcut in the toolbar. • Go to “My queries”, where you will find several predefined queries, and click on Administrative boundary > Municipalities (8) extent. • You must then select bbox (=bounding box) and the fictional cases layer will be proposed directly if it is already opened in QGIS. If other layers are proposed, choose fictional cases from the drop-down menu. Run the query. • Such map should be visible:

	 <p>In order to calculate the incidence, it is necessary to have the population size per district in the attribute table of the layer "Municipalities" downloaded from OSM. It is therefore necessary to check the existence of such column by right-clicking on the "Municipalities" layer in the layers panel and clicking on "Open Attribute Table".</p>
3. Incidence calculation with GeoHealth plugin	<p>Using these two vector files (fictional cases and administrative boundaries), it is now possible to calculate the incidence of the fictional disease per district with the GeoHealth plugin:</p> <ul style="list-style-type: none"> • Go to the Analyse tab > Incidence. • Then in the tab "Case and aggregation layers". <p>If only the two layers (fictional cases and municipalities) are opened in the QGIS project, they are directly proposed in the settings. Otherwise, select them from the drop-down menus. For the "Population field" select the population column (from the administrative boundary layer).</p> <p>For the other settings, it is possible to:</p> <ul style="list-style-type: none"> • choose the denominator (a value of 100,000 inhabitants is given by default); • change the name of the column that will indicate the incidence (referred to by default as "incidence"); • choose the location and name of the new layer to be recorded or leave as a temporary file by default; • choose the symbology (colors, graduation mode and number of classes).

4. Results	<p>Here is the result, with the default settings:</p>

Tableau 7.1. Creation of an incidence map with GeoHealth plugin

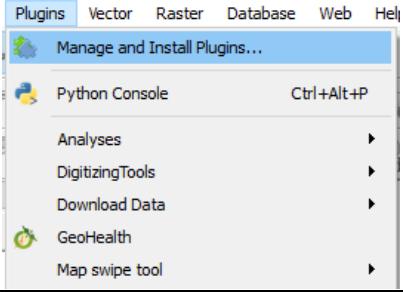
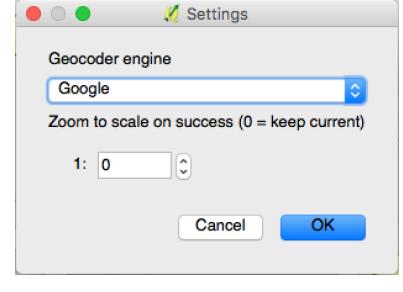
7.3.4. « Blurring » of point data with GeoHealth and QuickOSM plugins

7.3.4.1. Geocoding

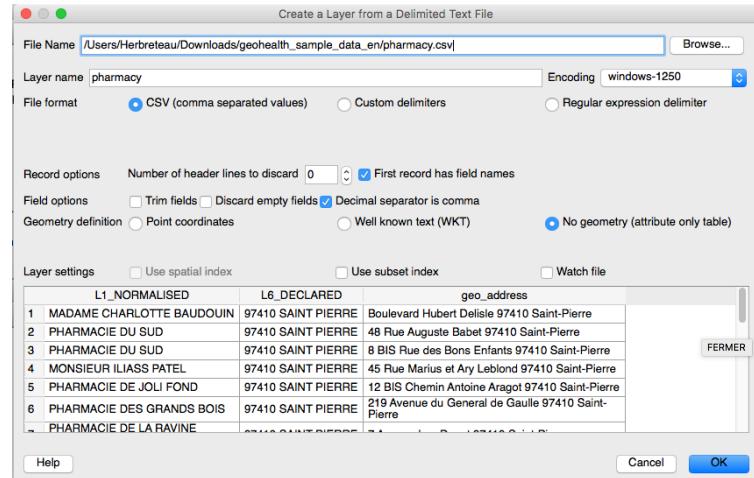
Objective: geocode addresses (associate geographic coordinates to postal addresses in order to map them).

The GeoCoding plugin allows to geocode addresses directly in QGIS by querying either Google web services or Nominatim. This plugin requires an internet connection. Addresses (number, street, postal code, city) must be written one by one (it is not possible to geocode an address table at a time).

The exercise is based on the use of the `pharmacy.csv` file, which contains the addresses of drugstores in Saint-Pierre district in Reunion Island. This table is an extract of the SIRENE database (<https://www.sirene.fr>), a French open database of companies.

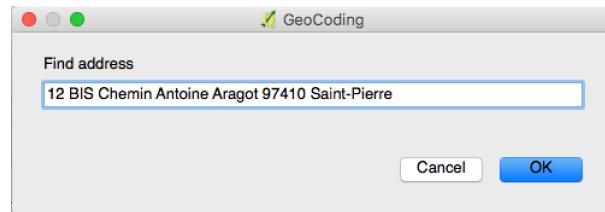
Approach	Handling in QGIS
1. Installation of the GeoCoding plugin	<p>Open Plugins Dialog from the Plugins menu / (Manage and Install Plugins...)</p> <p>GeoCoding plugin is then accessible from the Plugin menu.</p> 
2. Geocoding of addresses	<p>It is necessary to select first the database used for geocoding: Google or Nominatim.</p> <ul style="list-style-type: none"> • Go to Plugin / GeoCoding / Settings • Select Google as the Geocoder engine and validate <p>Then the table containing addresses (<code>pharmacy.csv</code>) could be opened separately and each address could be copied one by one to the GeoCoding plugin. This table could also be opened in QGIS:</p> <ul style="list-style-type: none"> • Go to Layer / Add Layer / Add Delimited Text Layer... • Select the <code>pharmacy.csv</code> file. • Indicate the CSV format. 

- Select also the option “No geometry (attribute only table)”.
- Validate with OK.



Each address of the geo_address column should be copied and pasted in the GeoCoding plugin:

- Go to Plugin / GeoCoding / GeoCoding
- Paste each address one by one (copied from the pharmacy.csv file) and validate with OK



- A file GeoCoding Plugin Results is created and each address is added to this file as it is geocoded.
- This file could be saved as a shapefile: right click on the file name in the layers panel, browse to your folder and give a name to this new shapefile.

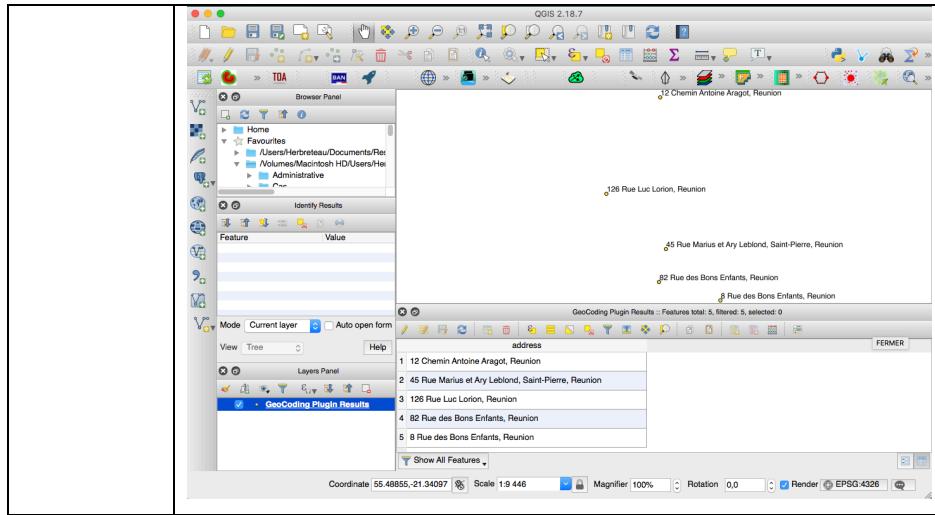


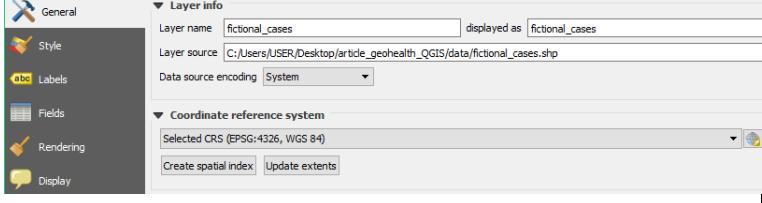
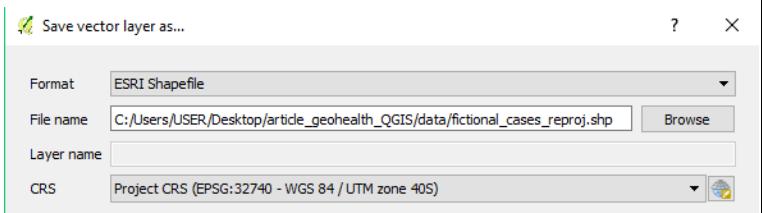
Tableau 7.2. Geocoding addresses with GeoCoding plugin

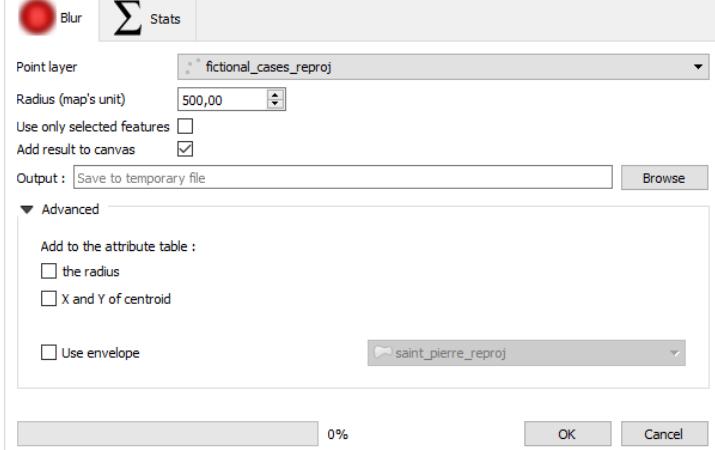
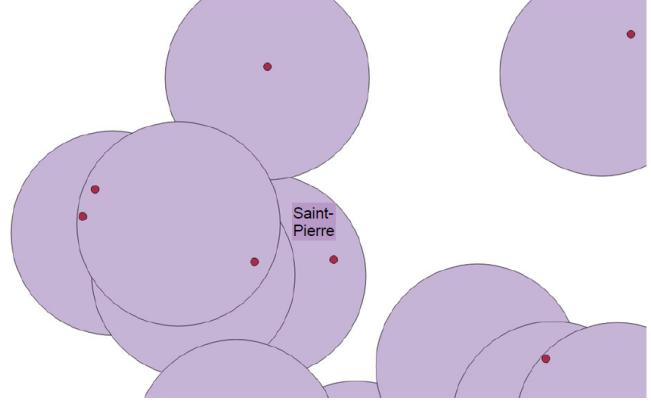
7.3.4.2. Blurring

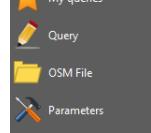
Objective: "Blurring" of the location of point data to anonymize them (see subsection 7.2.3).

The GeoHealth plugin proposes a spatial anonymization method via its “blur” tool.

Approach	Handling in QGIS
1.Importation of shapefiles	<ul style="list-style-type: none"> Open GeoHealth in menu: Plugin/GeoHealth. Select Import. Select Shapefile, and browse to the point shapefile called « fictional cases ». Once loaded, select Open. The file is displayed in QGIS.
2.Verification of the projection and eventual reprojection of data	<p>Since the blurring function involves the calculation of buffer zones, it must be verified beforehand that the layer to be blurred is in a projected coordinate system, in meters:</p> <ul style="list-style-type: none"> Double-click on the fictional case layer in the Layers Panel to open the "Layer Properties" window. Go to the "General" tab. Verify the coordinate reference system (CRS).

	
	<p>In this exercise, the layer of fictional cases is in projection WGS84 (CRS 4326) and the coordinates are therefore in geographical degrees. This layer must be re-projected, in meters:</p> <ul style="list-style-type: none"> Right-click on the layer and select "Save As...". Enter a new name for the layer you want to create and choose the folder to save it. Select the new CRS. In this exercise, for Reunion Island, it is WGS84/UTM zone 40S (EPSG: 32740). 
3. Blurring	<ul style="list-style-type: none"> Open the "Blur" tool: Analysis tab > Blur. Select in the list "Point layer" the point layer to be blurred (which is in a projected CRS in meters). Choose a blur radius (in meters). If necessary, specify a name and location for the output file (e.g. "blurred_cases.shp"). In the "Advanced" section, it is possible to choose to enrich the attribute table with columns indicating the blur radius, X (longitude) and Y (latitude) coordinates of the centroid of the blurred area. Click OK to start the calculation.

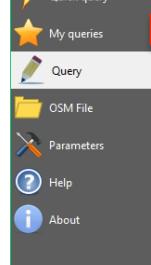
	 <p>Here is the result of the blurring with the default settings (500 meters radius):</p> 
4. Importation of the buildings of the study area with QuickOSM plugin	<p>Once the point data is blurred, the blur tool proposes to calculate statistics to help assessing the quality of the blurring: here, it calculates the number of houses contained in each of the blurred buffer zones (see subsection 7.2.3).</p> <p>A vector layer of buildings is necessary. It can be obtained by downloading the OpenStreetMap buildings using the QuickOSM plugin.</p> <p>To reduce the calculation time, this exercise will only concern Saint-Pierre district.</p> <p>In “Quick query”, indicate:</p> <ul style="list-style-type: none"> • Key/Value = building/yes. • In = “Saint-Pierre”. • Launch the query.

Quick query**Help with key/value**

Reset

Key	building
value	yes
In	Saint-Pierre
Extent of the map canvas	1000 m
Extent of a layer	

This query may take some time. You can get the error message "OverpassAPI execution time reached". In this case, it is possible to increase the execution time by editing the query: click on "Show query", and in the first line of the code, increase the value of the timeout from 25 by default, to 100 for example.

Overpass query

```
<osm-script output="xml" timeout="25">
  <query type="area">
    <has-kv k="building" v="yes"/>
    <area-query from="area_0"/>
  </query>
  <query type="node">
    <has-kv k="building" v="yes"/>
    <area-query from="area_0"/>
  </query>
  <query type="way">
    <has-kv k="building" v="yes"/>
    <area-query from="area_0"/>
  </query>
  <query type="relation">
    <has-kv k="building" v="yes"/>
    <area-query from="area_0"/>
  </query>
</osm-script>
```

Reset

Advanced

Generate query

Run query

Save query

Then the query must be restarted.

Here is the result after downloading the OSM buildings:



Once the query is realized, the building layer must be projected in the same projection system as the blurred areas:

- Right-click on the layer name and select "Save as".
- Select the new CRS (WGS84/UTM zone 40S (EPSG: 32740)).
- Save the layer under a new name.
- Confirm by clicking OK.

5. Blurring quality assessment	<p>The calculation of the number of buildings in the blurred zones is also done in the tab "Blur":</p> <ul style="list-style-type: none"> • Go to the Stats tab. • Indicate the fuzzy layer and the statistical layer (here, the buildings layer). • Click on OK. <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;"> <div style="display: flex; justify-content: space-between; align-items: center;"> Save table Save Y values </div> <table border="1" style="width: 100%; border-collapse: collapse; text-align: left;"> <thead> <tr> <th style="padding: 2px;">Parameters</th><th style="padding: 2px;">Values</th></tr> </thead> <tbody> <tr> <td style="padding: 2px;">Count(blurred)</td><td style="padding: 2px;">155</td></tr> <tr> <td style="padding: 2px;">Count(stats)</td><td style="padding: 2px;">118835</td></tr> <tr> <td style="padding: 2px;">Min</td><td style="padding: 2px;">0</td></tr> <tr> <td style="padding: 2px;">Average</td><td style="padding: 2px;">822.070968</td></tr> <tr> <td style="padding: 2px;">Max</td><td style="padding: 2px;">3157</td></tr> <tr> <td style="padding: 2px;">Median</td><td style="padding: 2px;">503.000000</td></tr> <tr> <td style="padding: 2px;">Range</td><td style="padding: 2px;">3157</td></tr> <tr> <td style="padding: 2px;">Variance</td><td style="padding: 2px;">695275.575609</td></tr> <tr> <td style="padding: 2px;">Standard deviation</td><td style="padding: 2px;">833.831863</td></tr> </tbody> </table> <div style="margin-top: 10px;"> <p>The screenshot shows a software interface with two main sections. On the left, there is a table of descriptive statistics for 'blurred' data. On the right, there is a scatter plot titled 'Blurred entity' showing the relationship between building counts and blurred entities.</p> <table border="1" style="display: inline-table; vertical-align: middle;"> <thead> <tr> <th>Parameter</th><th>Value</th></tr> </thead> <tbody> <tr> <td>Count(blurred)</td><td>155</td></tr> <tr> <td>Count(stats)</td><td>118835</td></tr> <tr> <td>Min</td><td>0</td></tr> <tr> <td>Average</td><td>822.070968</td></tr> <tr> <td>Max</td><td>3157</td></tr> <tr> <td>Median</td><td>503.000000</td></tr> <tr> <td>Range</td><td>3157</td></tr> <tr> <td>Variance</td><td>695275.575609</td></tr> <tr> <td>Standard deviation</td><td>833.831863</td></tr> </tbody> </table> <div style="display: inline-block; vertical-align: middle;"> <p>The scatter plot displays a positive correlation. The x-axis is labeled 'Blurred entity' and ranges from 0 to 160. The y-axis ranges from 0 to 3500. The data points form a curve that rises more sharply as the number of blurred entities increases, indicating that while most blurred areas have few buildings, a few have many.</p> </div> <p>This calculation provides descriptive statistics of the number of buildings per blurred area and a graph of the number of buildings, in ascending order. It is thus possible to evaluate the quality of the blurring and, if it is not satisfactory, to re-launch it. The table and graph can be easily saved.</p> </div> </div>	Parameters	Values	Count(blurred)	155	Count(stats)	118835	Min	0	Average	822.070968	Max	3157	Median	503.000000	Range	3157	Variance	695275.575609	Standard deviation	833.831863	Parameter	Value	Count(blurred)	155	Count(stats)	118835	Min	0	Average	822.070968	Max	3157	Median	503.000000	Range	3157	Variance	695275.575609	Standard deviation	833.831863
Parameters	Values																																								
Count(blurred)	155																																								
Count(stats)	118835																																								
Min	0																																								
Average	822.070968																																								
Max	3157																																								
Median	503.000000																																								
Range	3157																																								
Variance	695275.575609																																								
Standard deviation	833.831863																																								
Parameter	Value																																								
Count(blurred)	155																																								
Count(stats)	118835																																								
Min	0																																								
Average	822.070968																																								
Max	3157																																								
Median	503.000000																																								
Range	3157																																								
Variance	695275.575609																																								
Standard deviation	833.831863																																								

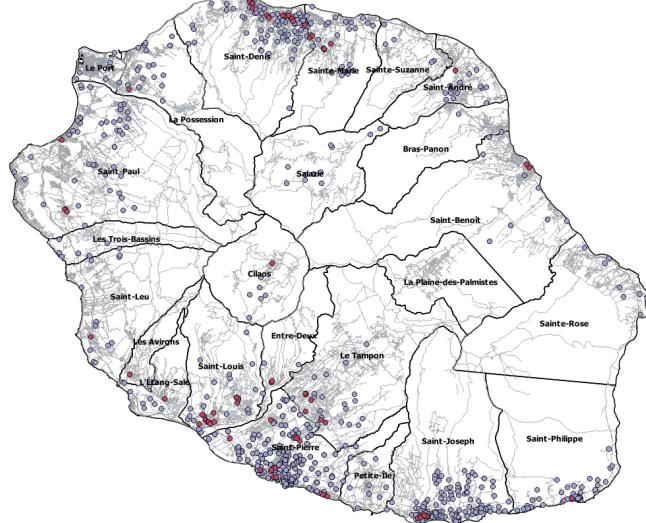
Tableau 7.3. “Blurring” of point data with GeoHealth plugin

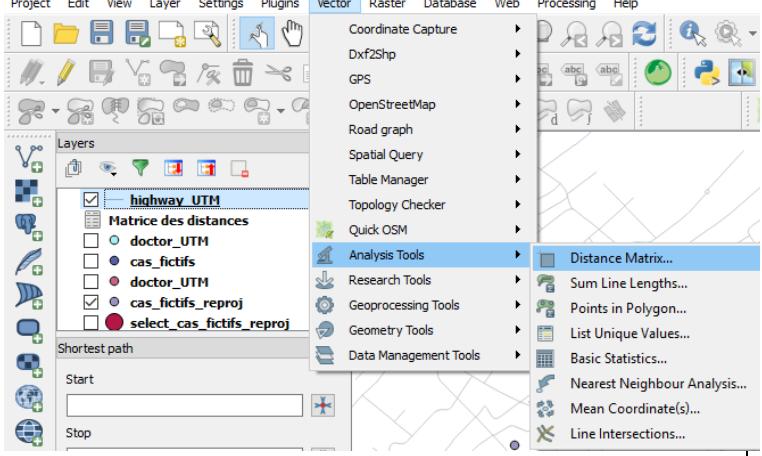
7.3.5. Example of spatial analyzes for health studies

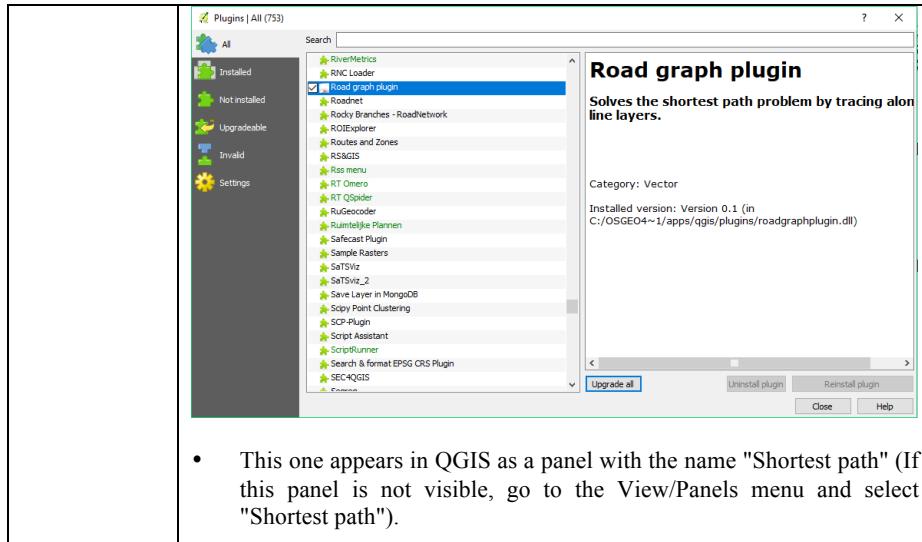
Objective: realize simple and useful spatial analyses useful for health studies:

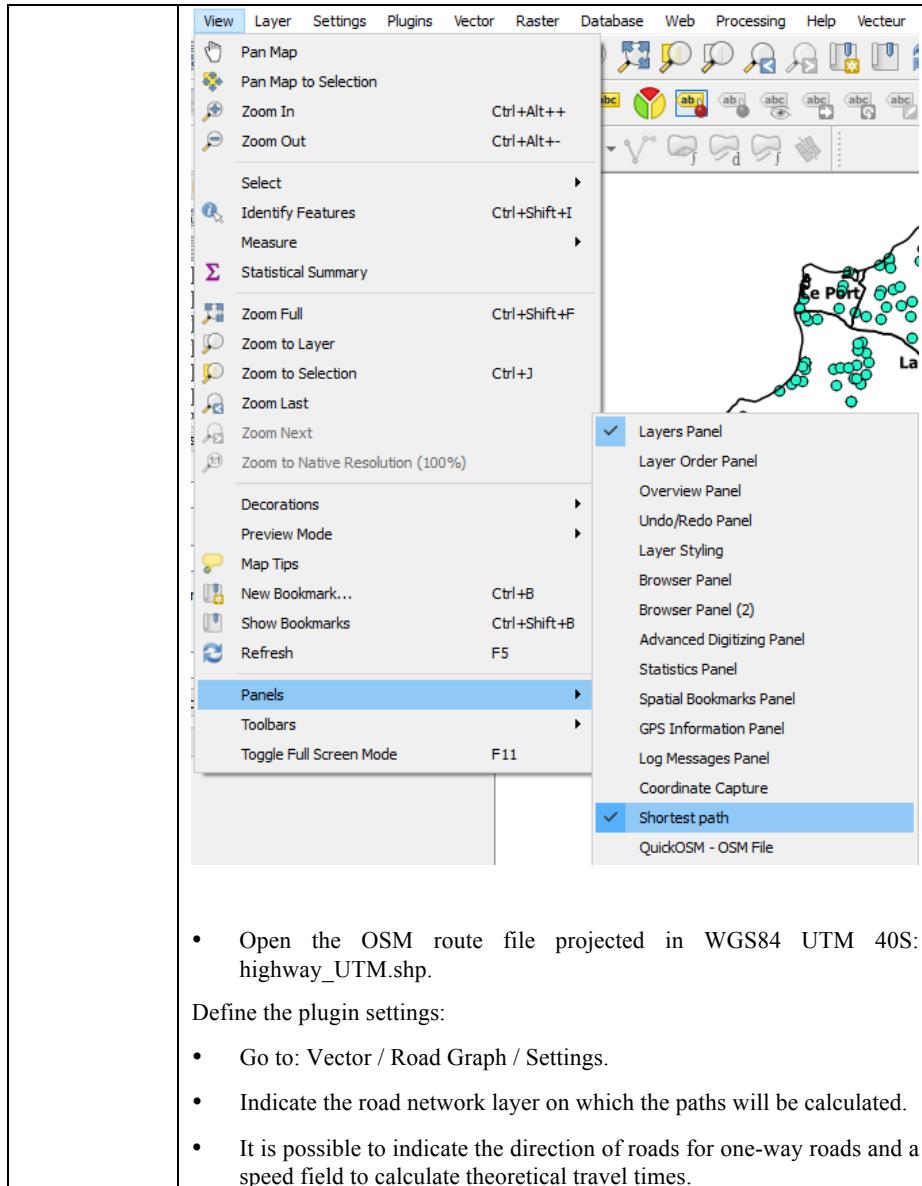
- the calculation of a distance matrix: we will use the locations of doctors, downloaded from OSM;
- the calculation of the shortest route, based on the OSM road network.

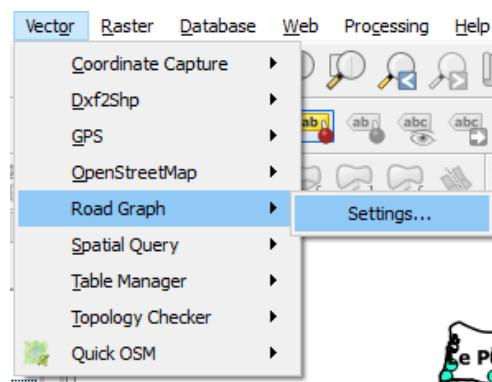
With the QuickOSM plugin, it is possible to specifically query objects of the OSM database. Here we will develop some examples of spatial analysis using this data.

Approach	Handling in QGIS
1. Download of necessary data with QuickOSM plugin	<ul style="list-style-type: none"> • Launch QuickOSM. • For doctors in the Quick query tab, specify: key/value = amenity/doctor. • For the road network in the Quick Query tab, specify: key/value = highway without indicating any value. This query can take time. • Reproject the road network in meters: <ul style="list-style-type: none"> ○ Right-click on the layer name and select "Save as". ○ Select the new CRS (WGS84/UTM zone 40S (EPSG: 32740)). ○ Save the layer under a new name: highway_UTM.shp. <p>This exercise requires the following data: fictional cases, doctors, the road network and the districts of Reunion Island.</p> 
2. Calculation of the distance matrix	<p>A dedicated QGIS function allows to directly calculating distances between two point layers. In this example, we will calculate the distances between the locations of fictional cases (fictional_cases.shp previously used) and doctors.</p> <p>As for blurring, we need a projection (CRS) in meters.</p>

	<ul style="list-style-type: none"> Go to: Menu Vector / Analysis Tools / Distance Matrix 
	<ul style="list-style-type: none"> Choose the two layers of points (example: n patients and m doctors). Choose the type of matrix: <ul style="list-style-type: none"> Linear distance matrix where all combinations will be in line ($n*m$ lines) standard distance matrix (with n rows and m columns). Specify the name and location of the output distance matrix to be saved. This matrix will be in .csv format. <p>The .csv file can be opened in Excel by importing: File/Import/, etc.</p>
3. Calculation of the shortest path	<p>QGIS integrates the Road graph plugin to calculate the shortest paths and travel times between points.</p> <ul style="list-style-type: none"> Go to the Plugins Dialog and activate the Road graph plugin



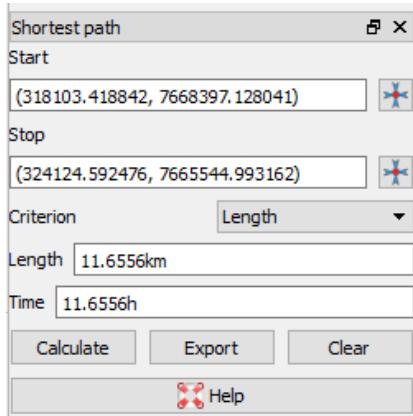




The screenshot shows the QGIS application interface. The 'Vector' tab is selected in the top menu bar. A context menu is open over a road network layer, with the 'Road Graph' option highlighted. Other options in the menu include 'Coordinate Capture', 'Dxf2Shp', 'GPS', 'OpenStreetMap', 'Spatial Query', 'Table Manager', 'Topology Checker', and 'Quick OSM'. The 'Settings...' option is also visible at the bottom of the menu.

Start the routing calculation:

- In the "Shortest path" panel, select the starting and finishing points on screen or indicate their coordinates.
- Click on Calculate.
- The route is then displayed on the map.



The 'Shortest path' panel is shown. It has fields for 'Start' (coordinates: 318103.418842, 7668397.128041) and 'Stop' (coordinates: 324124.592476, 7665544.993162). The 'Criterion' dropdown is set to 'Length'. Below it, 'Length' is listed as 11.6556km and 'Time' as 11.6556h. At the bottom are 'Calculate', 'Export', and 'Clear' buttons, along with a 'Help' link.

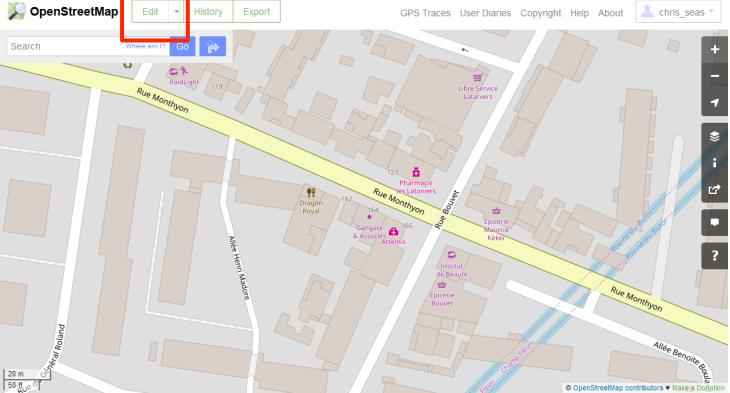
Paths can be exported as vector files.

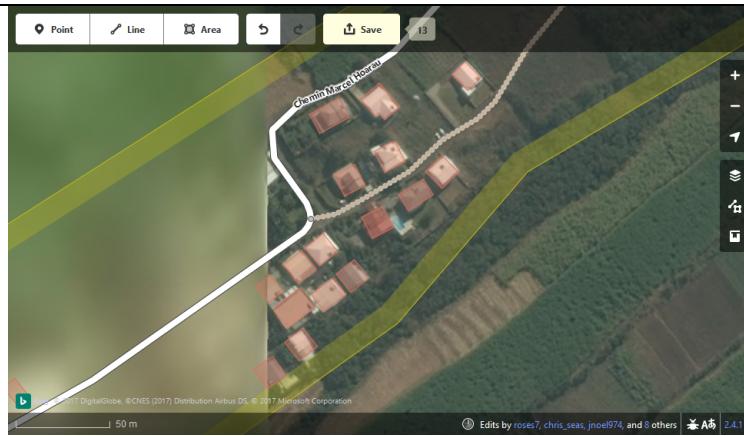
Tableau 7.4. Calculation of a distance matrix and the shortest path

7.3.6. Contribution to the OpenStreetMap database using satellite images

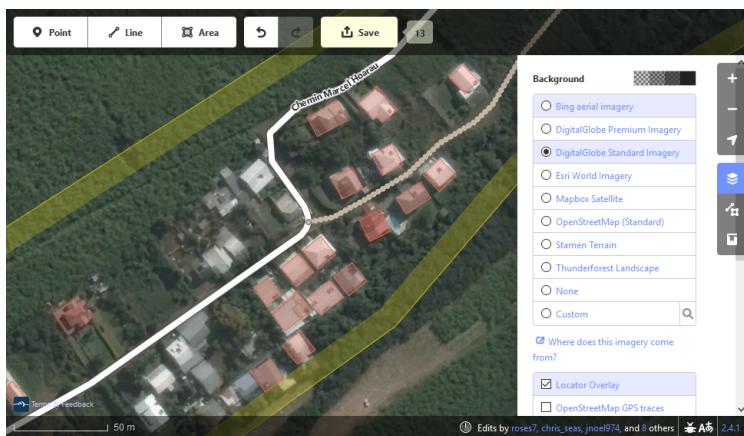
The objective of this section is to show how to contribute to the OSM database with the online editor "ID". Every contributor needs first to register in order to

access the editing interface and enter information. This editor provides several sources of satellite images that are used as a background image to help mapping.

Approach	Handling in QGIS
1. Registration	<ul style="list-style-type: none"> Go to the OSM website: https://www.openstreetmap.org. Click on "Register". Fill in the requested information and confirm the registration. Validate the registration by clicking on the link sent by email.
2. Open the ID editor	<ul style="list-style-type: none"> Log on to OpenStreetMap (https://www.openstreetmap.org). Zoom in to the area to be edited. Click on the small arrow next to Edit. Then click Change with ID.
3. Switching the satellite images	<p>In the editing interface, a satellite image is proposed as a background. By default, Bing images are visible. You can click "Background Settings" to select another background such as satellite images provided by MapBox and DigitalGlobe. It could be necessary to switch from one source to another to select the background that best corresponds to the mapping needs.</p>  



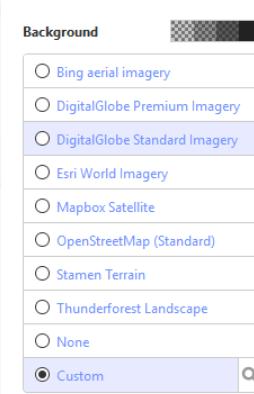
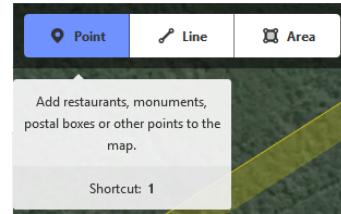
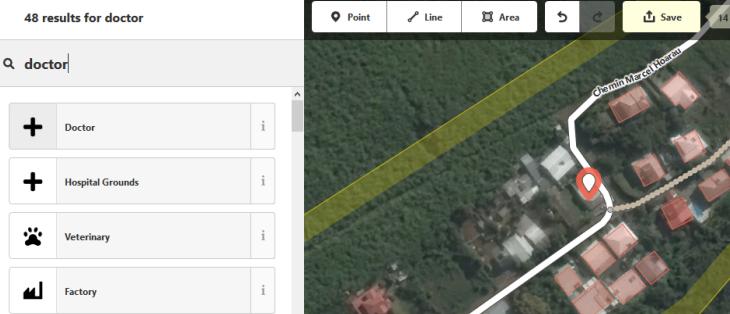
Default view of the Bing image in Edit mode



View of the DigitalGlobe image in Edit mode

It is also possible to specify an address manually for an image available on a server and accessible through a WMS (Web Map Service).

- Click on the "Background Settings" and select "Custom";
- Click on the magnifying glass next to customize and enter the server address (for example, for the IGN BD Ortho over French territories: <http://proxy-ign.openstreetmap.fr/bdortho/{z}/{x}/{y}.jpg>).

	
4. Contribution to OSM	<p>OSM proposes three types of geometries: point, line and polygon.</p> <ul style="list-style-type: none"> Switch to OSM edit mode and find an area where some buildings are not yet mapped. Click on "Point" to add a point of interest. Left click on the map background to create this point. For example, to add a doctor, go to the left and write "doctor" in the search window.   <ul style="list-style-type: none"> Select the first choice. It is possible to fill in the other fields proposed according to the available information (e.g. address, name, etc.). Click on save. Indicate in the "Change group comment" field the nature of the edition and information on the geographical location (municipality, region, country, etc.).

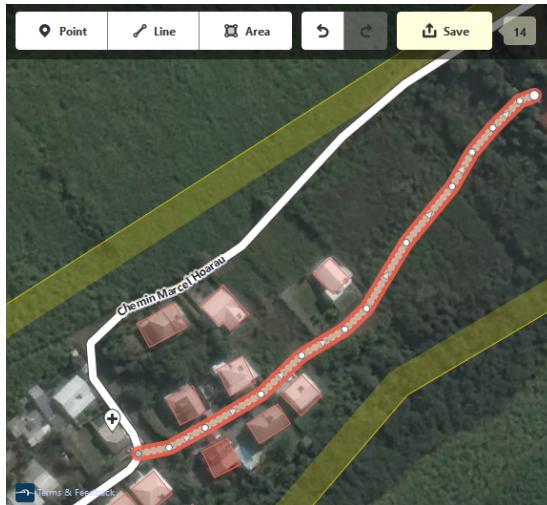
The changes you upload as [chris_seas](#) will be visible on all maps that use OpenStreetMap data.

I would like someone to review my edits.

[Cancel](#) [Upload](#)

To add linear information (e.g. a foot path):

- Click on "Line".
- Make a first left click to start editing and click on each change of direction.
- To end the line, double left click on the last point.



- Then add the keys and values of this line in the left-hand side of the interface, searching for "foot path".
- Save the changes by filling in the comment field as before.

Note: when adding a new route, it should be connected to the existing network as in the screenshot.

It is possible to modify the geometry of every element and to complete associated information by adding fields (keys). By right-clicking on the element, a toolbar appears: it allows to move the element, to change its direction, to make it straight or to delete it.



	<p>To add a polygon (for example, a new building):</p> <ul style="list-style-type: none"> • Click on "Polygon". • Left click on a corner of the building to start editing and then click on each corner. • End with a double click on the last point. • Fill in the information on the left side of the interface by searching "building". The first proposal corresponds to a building without any other attribute (building = yes). <p>Once this building is completed, it is possible to right-click on the entity to access a toolbar of rapid edition, which allows to make it circular, to move it, to make it orthogonal, to duplicate it, to rotate it or to delete it.</p> <p>For a building, it is recommended to give it an orthogonal shape.</p> <p>When the object is finished, save the changes.</p> <p>Note: the addition of roads and buildings, as shown here, are visible on the OSM map quickly.</p> 
--	---

Tableau 7.5. Online contribution to OpenStreetMap

In the OSM ID editor, some of the usual keyboard shortcuts also work: copy (ctrl + c), cut (ctrl + x) and paste (ctrl + v) shortcuts make it possible to duplicate or delete entities quickly. The shortcut (ctrl + z) is also used to cancel the last action.

Finally, in order to properly inform the keys and values, it is recommended to consult the OSM wiki. (https://wiki.openstreetmap.org/wiki/Map_Features).

7.4. References

[PIC 01] PICHERAL H. (2001). *Dictionnaire raisonné de géographie de la santé*, Université Montpellier III-Paul Valéry, 307 pp.