GIS in Sustainable Urban Planning and Management: A Global Perspective

Methodological demonstration for **Chapter 18** *"Utilising volunteered geographic information to assess resident's flood evacuation shelters. Case study:Jakarta"* **(pp. 307-321)**

- Kusumo A.N.L, Reckien D., Verplanke J. Mark Brussel

- Methodological demonstration by André Mano

Disclaimer

This document is an addendum to the chapter mentioned above which is part of the book *GIS in Sustainable Urban Planning and Management: A Global Perspective*. The purpose of this document is to demonstrate the application of the methods described in that chapter using QGIS 3.x along with the data available at here. This document is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. Different license terms may apply for the data. If that is the case, a file containing the license terms is included with the data.

How to use this document

Most of the steps described are illustrated with screenshots. Bear in mind that what the screenshot depicts and what you see in your computer might differ slightly depending on the QGIS version you are using and the way your toolbars and add-ons are arranged. Along the text you will see different icons. The key for these icons is as follows:

Data or external resource to download;

A software action you are supposed to do;

Information specific about QGIS.

Additional or complementary scientific information;

An important concept which you may want keep in mind;

[1] An operation that is referenced in the flowchart of operations.

Additionally, for the sake of readability, the following style conventions are used:

- A reference to dataset or a layer uses `this style`;
- A QGIS command, or any clickable button is noted using **this style.**
- A QGIS menu or section is highlighted using *this style*.

At the end of the document, a diagram depicting the workflow described in these pages can be seen. It is advisable to look at it first and/or refer to it as you proceed.

GIS in Sustainable Urban Planning and Management: A Global Perspective

Outline

The book chapter explores the usefulness of Volunteered Geographic Information (VGI) collected from Twitter to assess the spatial distribution of formal (official) and informal flood shelters in Jakarta, Indonesia. For that purpose two flood events from 2013-14 and 2015-16 are used to compare the location of the official shelters with the location of the actual places where people sought refuge in order to evaluate the value of VGI information for spatial planning.

Getting started

⤓ Download the data; the data consists of the following files:

- `Jakarta.qgs` – a QGIS project preloaded with the layers;
- `Formal_shelters.shp` – the official flood shelters represented by polygon features;
- `Land_use.shp` – Polygon features with land use classification as the key attribute;
- `Tweets_1314.shp` – All the relevant tweets as point features from the 2013-14 flood event;
- `Tweets_1415.shp` – All the relevant tweets as point features from the 2014-15 flood event;
- `flood_1314.shp` – polygon features with the areas flooded during the 2013/14 event;
- `flood_1314.shp` – polygon features with the areas flooded during the 2014/15 event.

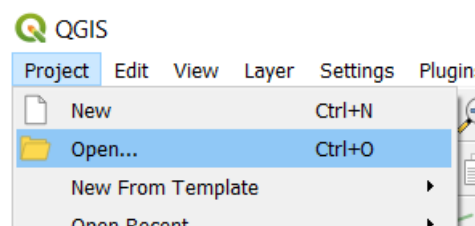▶ Start QGIS and open `jakarta.qgis` (Figure 1).

Figure 1 – opening a project

▶ From the *Layers panel*, **right-click** on a layer and access the attribute table to examine it. Repeat the procedure for the other layers (Figure 2).
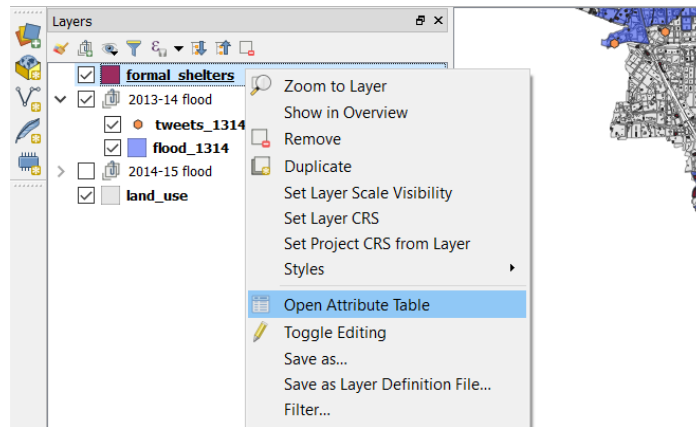
Figure 2 – assessing attribute table

Generate a hexagonal grid or fishnet [1]

The spatial unit of analysis for this study is a hexagonal grid with equal length sides of 200m.

💡 The spatial accuracy of a georeferenced tweet varies widely. Depending on how the position of the device was assessed in the moment of the tweet - for example if it is a purely network based positioning or if it was benefiting from a GNSS signal – the accuracy might vary from a few meters to kilometres. Therefore, the tweets you are provided were snapped to their real location, this real location was obtained from words of the tweet itself that referred to the actual location from where the user was tweeting. On average there was a shift of 200m between the tweet location and the actual location. An illustration of this can be seen below (Figure 3)
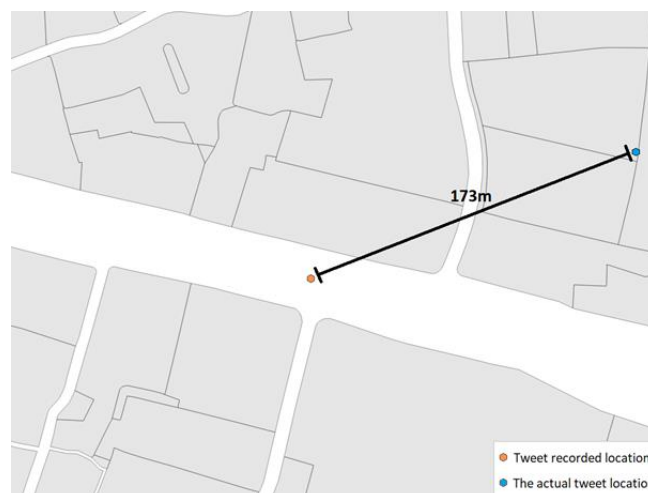

Figure 3 – difference between recorded and actual tweet location

▶ From the *Processing Toolbox*, filter by "grid" to find the tool **Create grid**. Choose *Haxagon (polygon)* as grid type. Define the **Grid extent** from the layer `land_use` and enter the **Horizontal** and **Vertical spacing** as 346. Choose an output directory

and name provide a name for the output file – we suggest `hexagonal_grid` and hit **Run in Background** to create the grid. (Figure 4).

🔆 The distance parameter that has to be provided in order to create a hexagonal grid refers to the side to side distance of a hexagon and not to the side length. In order to know what value should be entered to obtain a hexagon with the desired side length we have to use this equation:

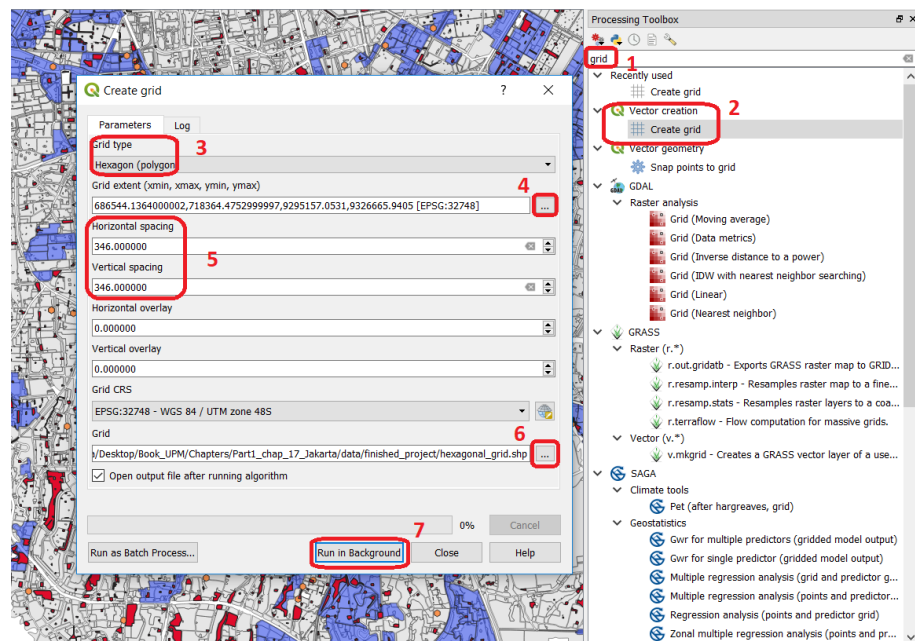$$side\ to\ side\ distance = lenght\ of\ the\ side * \sqrt{3}$$

Figure 4 – Creating an hexagonal grid

Add shelters information to the hexagonal grid (spatial join) [2]

The grid provides an aggregation unit for easy comparison and crossing of the data to be analysed. Before proceeding we will join the following information to the grid:

- Formal shelters (official shelter locations as determined by the government;
- Tweets of the flood event of 2013-14 (informal shelters);

▶ From the *Processing Toolbox*, filter by "join attributes by location" to find the tool **Join attributes by location**. Choose the `hexagonal_grid` as input layer and `formal_shelters` as Join Layer. As for the **Geometric predicate** choose *overlaps*. Make sure you are only adding the relevant attributes to the table `hexagon_grid`, in this case that would be "Tipologi" and finally select the option **Take attributes of the first located features only**. Choose an output

directory and name provide a name for the output file – we suggest `grid_join1` Finish the operation by clicking on **Run in Background** to apply the spatial join. (Figure 5).

When performing *joins* it is advisable to select which field(s) to add as opposed to join all the fields to the input layer. This will prevent a table overloaded with attributes that are irrelevant for the analysis we want to perform.
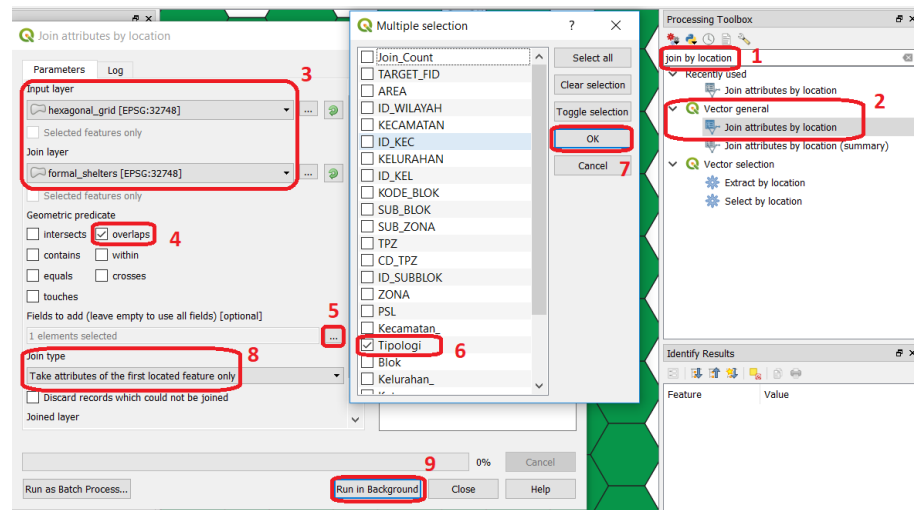

Figure 5 – Joining by location from `formal_shelters` into `hexagonal_grid`

Repeat the previous operation for the layers and attributes described in table 1. You will have to repeat this operation based on the output of the previous interaction of the tool.

Input layer	Join Layer	Geometric Predicate	Fields to add	Output name
`grid_join1`	`Tweets_1314`	Intersects	All	`Grid_final`

Table 1: Summary of spatial joins to be performed

You should now have a layer named `grid_final` that has the structure depicted in Figure 6. You may remove the other `grid*` layers from your project.

Figure 6 – Attribute structure of the layer `grid_final`

Visual analysis through the fishnet [3]

A proper styling of the fishnet can provide an efficient first glance over the spatial distribution of shelter locations. In particular we are interested in visualizing which spatial units (i.e. hexagon) have formal shelters, informal shelters only, both formal and informal shelters, or no shelter at all.

➤ From the *Layers panel*, **right-click** on the layer `grid_final` and choose the option **Properties**. From the **Symbology** tab choose **Rule-based** and click on **Add** to create your first rule. This first rule will have as **Label** "formal and informal" and for **Filter** enter the following expression:

```
"Tipologi" like '%' and "text"  like '%'
```

Choose a colour, press **OK** to dismiss the *Edit Rule* dialog Window and click **Apply** the style rule (Figure 7).

➤ Repeat the previous steps to create three more styling rules according with the parameters shown in table 2. Make sure you assign different colours to each rule.

Label	Filter
Formal shelters	`"Tipologi"  like  '%' AND "text" is null`
Informal shelters	`"text" like '%' and "tipologi" is null`
Others/no shelter	`"Tipologi"  is null and  "text" is null`

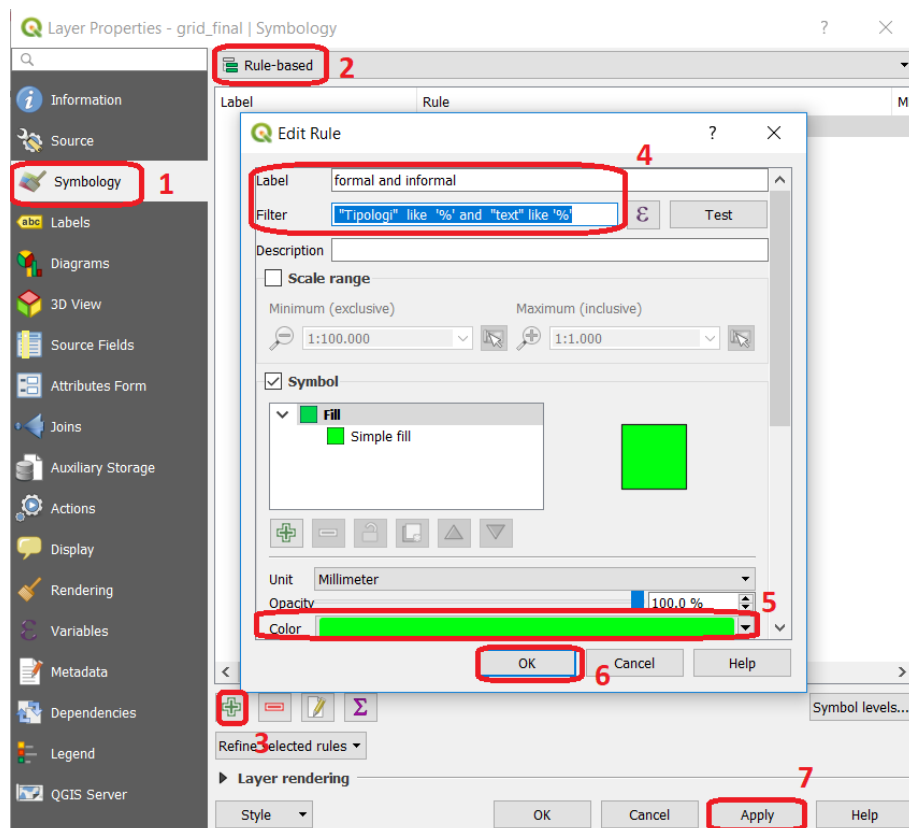Table 2: Filter expressions for layer `grid_final`

Figure 7 – Defining a filter for Rule-based styling

▶ Your `grid_final` layer should now be styled according to four different classes (Figure 9). From the *Layers panel*, **Right-click** on the layer and check the option **Show Feature Count** so that the count of spatial units belonging to each class is also shown in the legend (Figure 8).
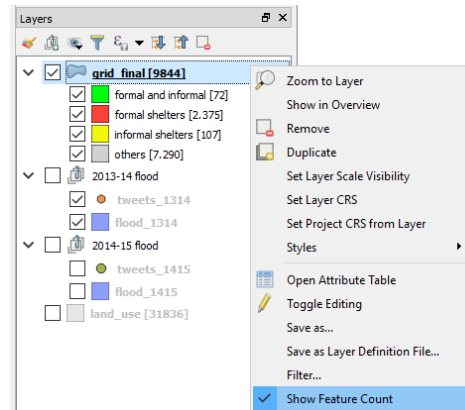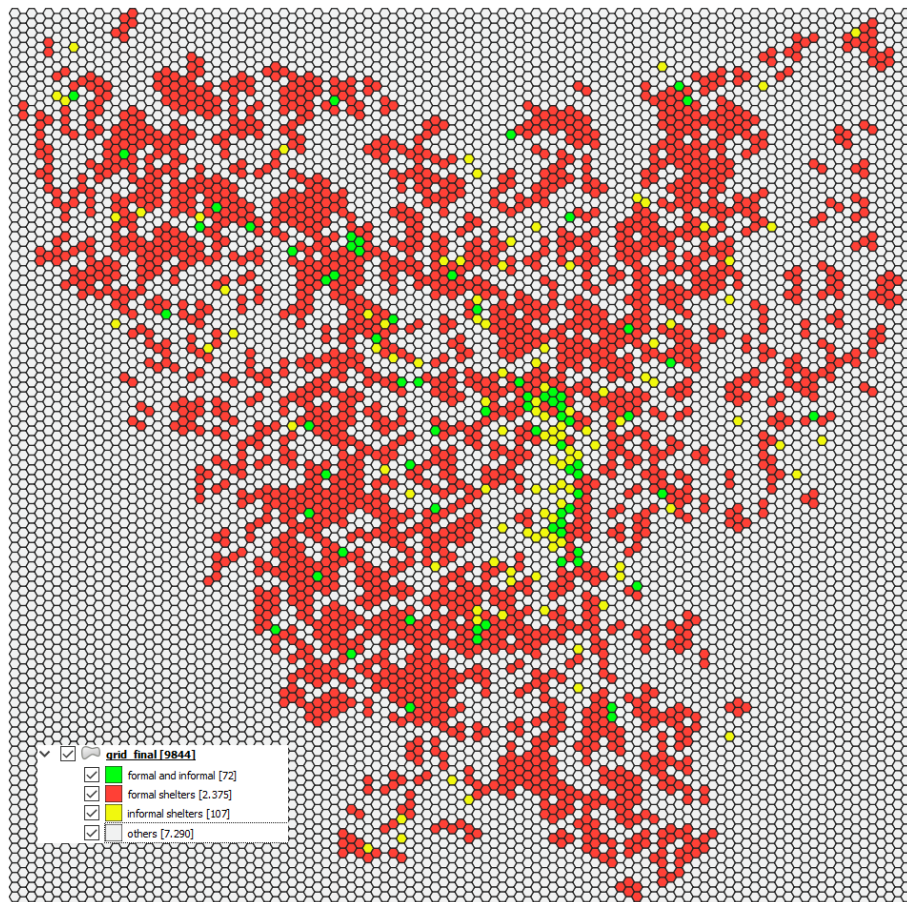

Figure 8 – Showing feature counts

Figure 9 – Grid_final after applying the filters for Rule-based styling

Data analysis – creating a geopackage

For a more insightful perspective we will perform quantitative analyses using some of built-in statistical tools of QGIS along with the possibilities of the Geopackage(.gpkg) format and its support for SQL (Structured Query Language).

The Geopackage is an OGC (Open Geospatial Consortium) standard that describes a set of conventions for storing vector features, raster maps, non-spatial data and extensions in a single file that consists essentially of a SQLite database[1].

From the *Layers panel*, right-click on the layer `grid_final` and choose the option **Save as**. In the *Save Vector Layer* dialog, choose *Geopackage* as the **Format**, the **File name** (we suggest you name it `grid_final` and then remove any other files you may have with this name form your project) and make sure to check the option **Add saved file to map.** Finally, click on **OK** to finish (Figure 10)
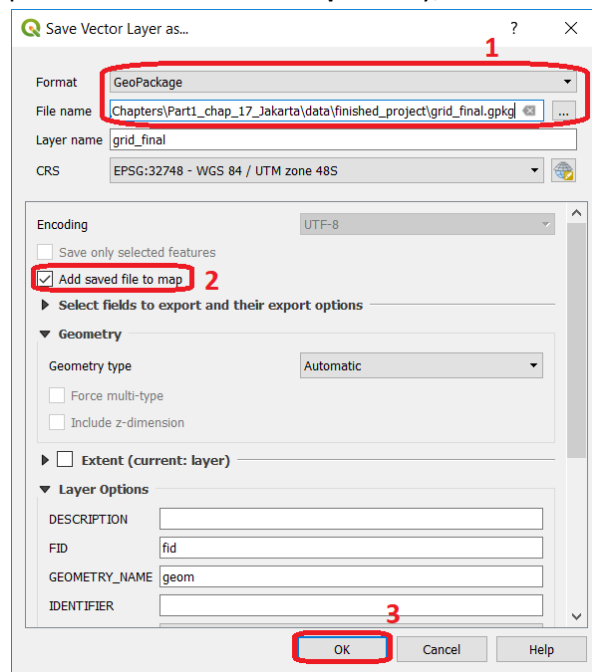

Figure 10 – Saving a layer to Geopackage format.

Data analysis – distances [4]

Now that you have a Geopackage in your project, you have to use a database manager to connect to it and use the SQL language to explore the data.

Go to *Database > DB Manager*. In the *DB Manager* dialog, right-click on **GeoPackage** and choose the option **New Connection.** You will have to provide the directory/Geopackage you created before (Figure 11).

---

[1] For more information about this format please check https://www.geopackage.org/
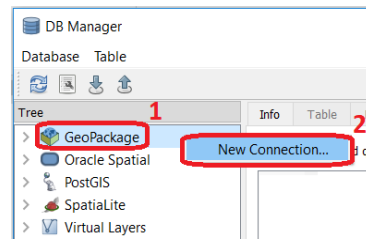
Figure 11 – Connecting to a Geopackage from the DB Manager

➤ Under *GeoPackage* you will now see the connection to your geopackage. **Double-click** on it to establish a connection to it. Click on the SQL Window[2] button and enter the following query:

```
select a.id, min(st_distance(a.geom,b.geom)) as distance
from grid_final as a, grid_final as b
where a.text is not null and b.tipologi is not null
group by a.id
```

The query above calculates the distance from an informal shelter to the nearest shelter. Click on **Execute** to run the query and check the option **Load as new layer** and then on **Load** to add the result of this query to your project (Figure 12).
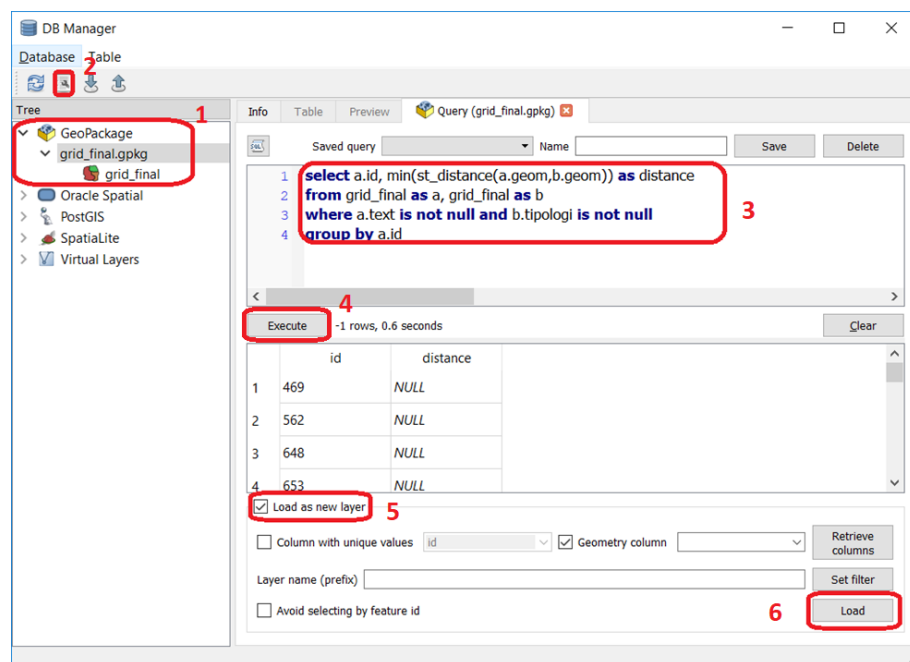

Figure 12 – Running a query from DB Manager and load the result

➤ In order to provide more context to the table we just produced, we will join this table to the `grid_final` layer. From the *Layers panel*, **right-click** on the layer `grid_final` and access the layer **Properties**. Go the **Joins** tab, **Add a join** and provide the correct join conditions: **Join Layer** will be the `QueryLayer`, **Join Field** will be `Id` and **Target Layer** will also be `id`. Press **OK** to finish (Figure 13).

---

[2] A good recourse to learn the basics of SQL can be found at https://www.w3schools.com/sql/
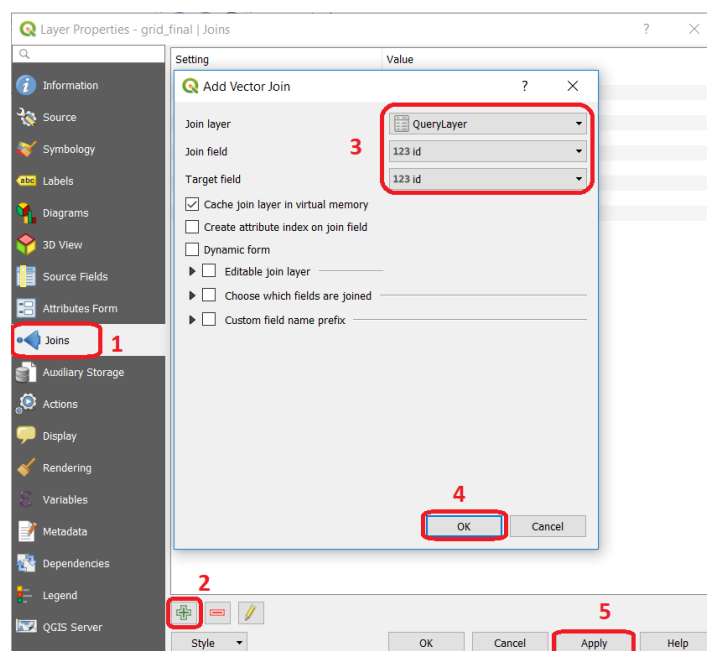
Figure 13 – Defining a join

Another interesting overview on the distance data can be obtained from performing a simple descriptive statistics analysis.

➤ From the *Processing Toolbox*, filter by "basic" to find the tool **Basic Statistic for fields**. Choose `Grid_Final` as **Input Layer** and `QueryLayer_distance` as **Field to calculate statistics on**. Click on **Run in Background** to run the tool. The results will be shown on the *Log* tab (Figure 14).
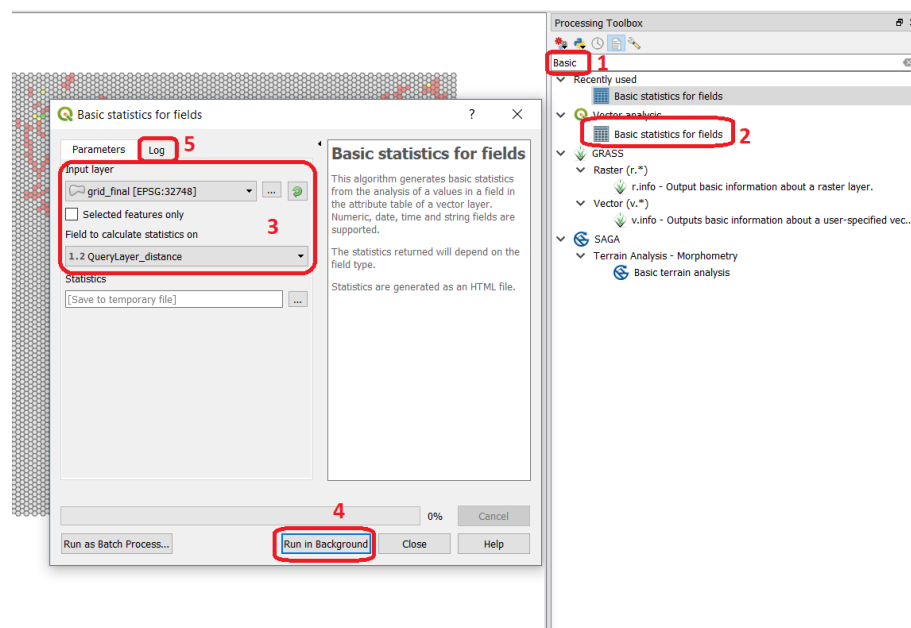

Figure 14 – Calculating field statistics

Data analysis – formal shelter coverage [6]

Finally, it may also be interesting to assess to what extent (in percentage) formal shelters coincide with informal shelters (i.e. from where the tweet was sent).

▶ Go to Plugins > Python console to start the Python console. We will use it as a simple calculator to have the value of:

```
72*100/179
```

Those numbers are taken directly from the *Layers Panel* legend (Figure 15) where:

- 72 is the count of formal shelter locations that coincide with the location from where a tweet was sent;

- 179 is the total number of spatially distinct (i.e the result of adding formal and informal tweet locations);
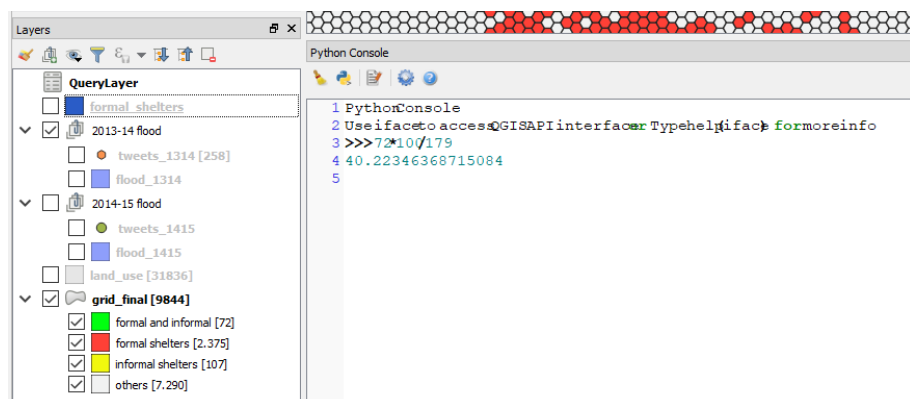

Figure 15 – Using the python console

💡 You may have noticed that the number of formal and informal shelters present in the `grid_final` layer is not the same as the original datasets from where this information was extracted – `formal_shelters` and `tweets_1314` respectively. This is because when you joined the attributes by spatial location, only the first instance of an intersecting occurrence was counted. Therefore the `grid_final` layer holds information for the presence/absence of something and not the magnitude/count of that same thing.

This concludes this demonstration. You can apply the same procedure for the datasets of the 2014-15 flood event or merge the data of the two events and run these analysis on the resulting dataset.

Flowchart of operations

Generate Hexagonal Grid	**1**

formal shelters	hexagonal grid	tweets_1314

Join by location	**2**

land_use

grid_final

5	**3**	**4**
Coverage analysis	Visual analysis	Distance analysis