Targeted skills

By the end of this module, you will know how to:

- get access to countries administrative boundaries GIS layers
- perform a quick visual analysis of outbreaks density using an "Heatmap"
- perform a (spatial) query to select features (point) of interest
- identify a CRS relevant to your region of interest and use case
- handle data attributes: add, delete, move, calculate, …

Data

Data to be used in this module can be found here:

`data/asf-20180101-20190410.csv`

Exercise outline & memos

In this series of module, we will focus on **ASF (African Swine Fever) outbreaks data in Romania from 1st January 2018 to 10th April 2019**.

1. Practice "Day 1" acquired skills

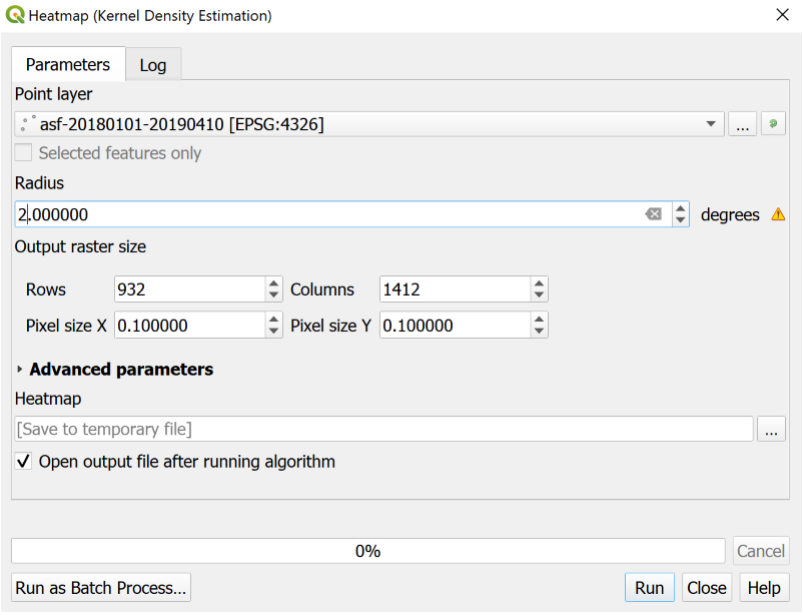
- first, open "CartoDB Positron" XYZ tiles
- then load the `data/asf-20180101-20190410.csv` .csv file into QGIS and convert it to a Shapefile

2. Take a quick look at outbreaks density worldwide with a Heatmap

You might have noticed that visualizing raw ASF data at world scale resuts in an overcrowded and cluttered map making it difficult to assess visually the density of outbreaks worldwide. We will perform various data analysis in the next module, but for now, let's use a technique called "Heatmap" allowing to quickly visually assess points density.

This type of visualization is quite popular and have its pros and cons but is often a good first step. To create a Heatmap with QGIS3:

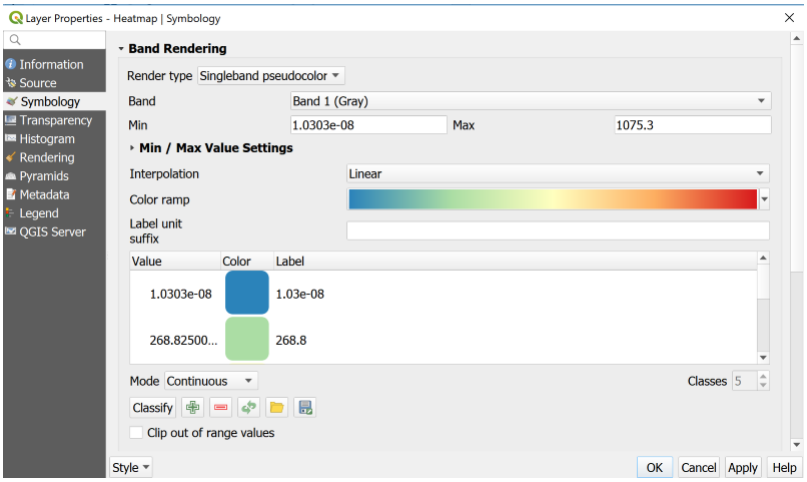
```
[QGIS Processing Toolbox]
1. In the search box, type: "heatmap"
2. Then double-click on Interpolation  Heatmap (Kernel Density Estimation)
3. And finally reproduce the settings below
```



Choosing the right Heatmap settings/parameters is a trial and error process and there is not defined methodology. In a nutshell and intuitively, a heatmap counts the number of points a given area (radius of influence). Technically speaking, it simply estimates a 2D probabilty density distribution using points location as samples from a theoretical one. We will later on go through more relevant and predictable way to look at points/outbreaks density including as well the possibilities to aggregate data (sum, mean, ...) at a given resolution ...

By now, you should have a new layer in your Layer panel with the following name "Heatmap". Let's now adjust the symbology and color rendering.

```
[QGIS Layer Panel]
1. Double-click on "Heatmap" newly generated layer
2. Reproduce the settings below
```



Given your color palette, you might need to invert it. To do so, click the "ColorRamp" dropdown and select "Invert Color Ramp"

3. Getting access to countries administrative boundaries

Given previous visual assessment, we've identified high densities of outbreaks in Romania and Poland. Let's use further assess Romania's situation in this module (Poland case is let as an exercise).

In later modules, we will aggregate the outbreaks data, getting the sum of cases or others at county levels, regional levels

or in a generated grid. Finding, authoritative administrative boundaries data is not always a simple process (as they might change quite regularly). The safest way is to ask these data to your country's agency in charge (Statistical office, National Mapping Agency, …). However, a first often very relevant data source is the GADM data portal, a global database of administrative areas.

- You can reach this data portal at this url: https://gadm.org/index.html
- And download, the shapefiles of your country of interest at this url: https://gadm.org/download_country_v3.html
- In our case, let's download the shapefiles for Romania: https://biogeo.ucdavis.edu/data/gadm3.6/shp/gadm36_ROU_shp.zip

Download, copy and unzip it under course `/data` folder.

4. Selecting outbreaks in Romania

In **Day 1**, we've seen several ways to select/filter data.

In this chapter, we would like to select loaded ASF outbreaks data that took place only in Romania.

Your tasks:

- Select by attribute value (open the attribute table, identify if country name is available, and filter, …)
- Perform a spatial query (points within Romania country boundaries)
- Check visually that you get the same selected points
- Save the selection as a new shapefile layer

5. Reprojecting the data

In later modules, we will have to specify several parameters using meters or kilometers. At the moment, all our data (both ASF outbreaks and admin. boundaries) are in WGS84, hence using longitude and latitude in decimal degrees.

In order to minimize distortions when projected at country scale, each country opts for a specific projection. Reach out your National Mapping Agency to find out which ones are offical. However, the Universal Transversal Mercator (UTM) is often convenient and widely used in national and international mapping systems around the world delivers high accuracy in zones less than a few degrees in east-west extent. UTM projection are available worldwide in different zones.

To find out the zone of interest:

- Open the following file in the `resources\` folder: resources/utm-zones.jpg

Which zone to use?

Your task:

- Reproject all Romania admin. boundaries levels and ASF data for Romania in chosen UTM Zone X

6. Handling data attribues

Before we start performing some interesting visualization and analysis, we would like to clean, transform some attributes of the layer of interest: the ASF outbreaks in Romania and reprojected (proper CRS).

We have essentially to techniques in our arsenal:

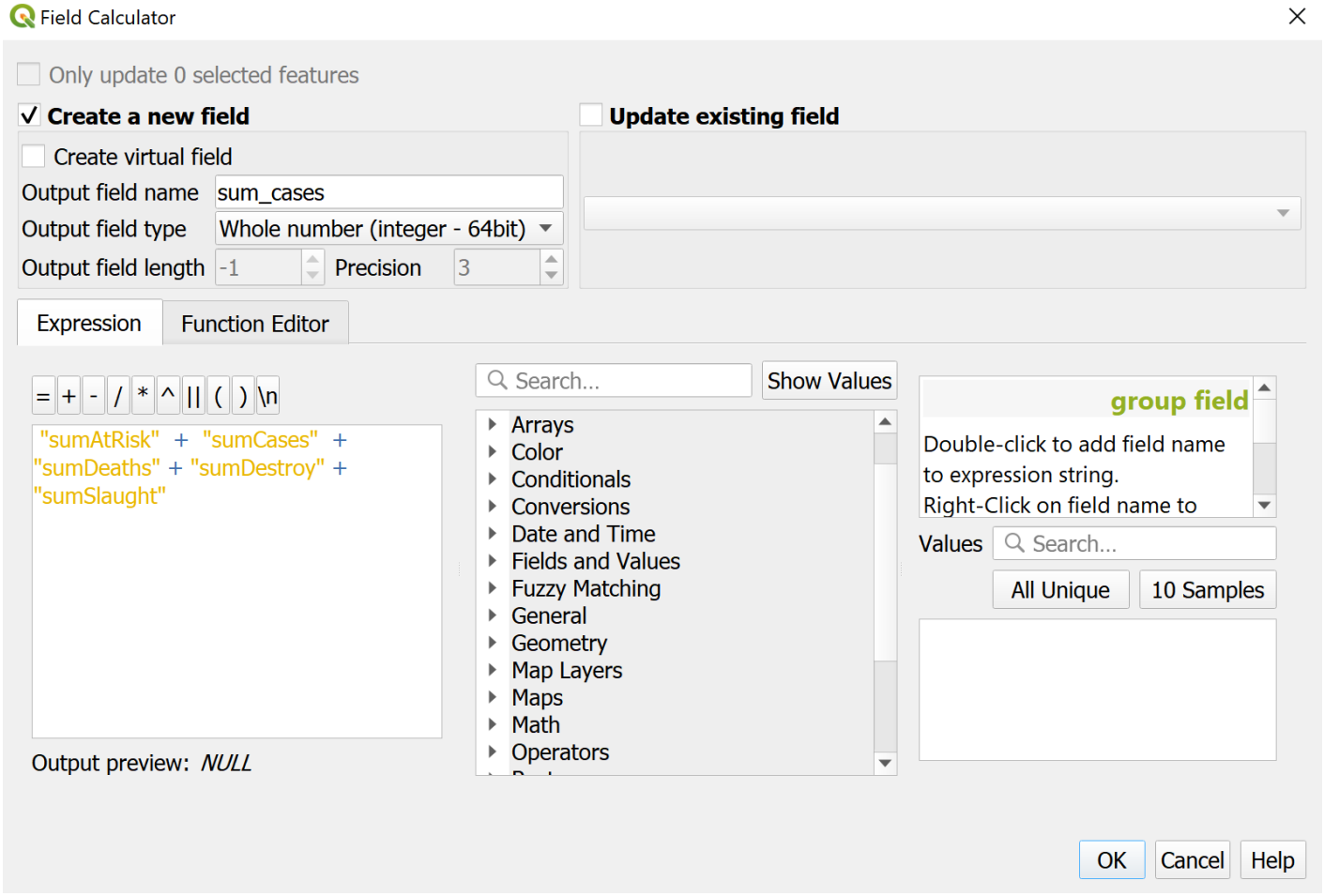
- the `Refactor fields` tool in the `[QGIS Processing Toolbox]` will allow you to move, delete, … attributes
- the layer attribute table `Field calculator` to update existing fields/attributes and create new ones

Your tasks:

1. Delete any fields/attribute you judge useless
2. Calculate the sum of animals as follows: `sum_animals = sumAtRisk + sumCases + sumDeaths + sumDestroy`

Some details on point 2:

1. Open the attribute table of your layer of interest
2. Open field calculator (the Abacus icon in the top toolbar of the attribute table - just hover to the icons to find out their function)
3. Reproduce the settings below



Check you result! What's wrong?

Hints: Some of the values of your fields contains the value `NULL`. When you perform the calculation `NULL + 12`, it equals to `NULL`. As a result, you need to clean each attribute/field to be used in your calculation using the following formula: `if("sumAtRisk" is null, 0, "sumAtRisk")`.