# 02.221 – Lab 6: Spatial overlays, buffers and unions

The data needed for this lab can be found on the course website or downloaded from Dropbox: https://www.dropbox.com/s/ltdey1xdxbrj4gz/02221-Lab6.zip?dl=0. Extract the data from the zip file to an appropriate location within your documents.

## Goals

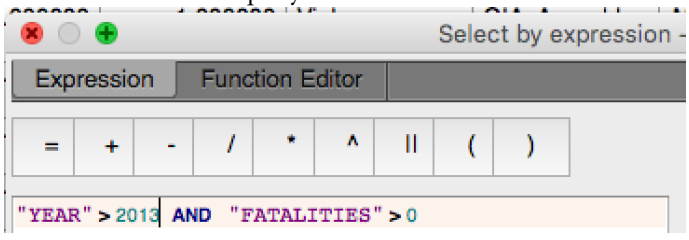
The primary goal of this lab to learn how to use different spatial overlay techniques (i.e. ways of combining and intersecting multiple spatial layers). We will do so through an analysis of the impact of political violence and armed conflict on the population of Africa.

## Armed Conflict Location & Event Data

The data we will use in this lab comes from the Armed Conflict and Location & Event Data Project (ACLED). It is a comprehensive collection of political violence and protest data in both Asia and Africa. The goal in the subsequent sections will be to use this data and try to determine how many people were (indirectly) affected by these violent events in each of Africa's countries in 2014-2015. In principle, the following steps are involved in answering that question:

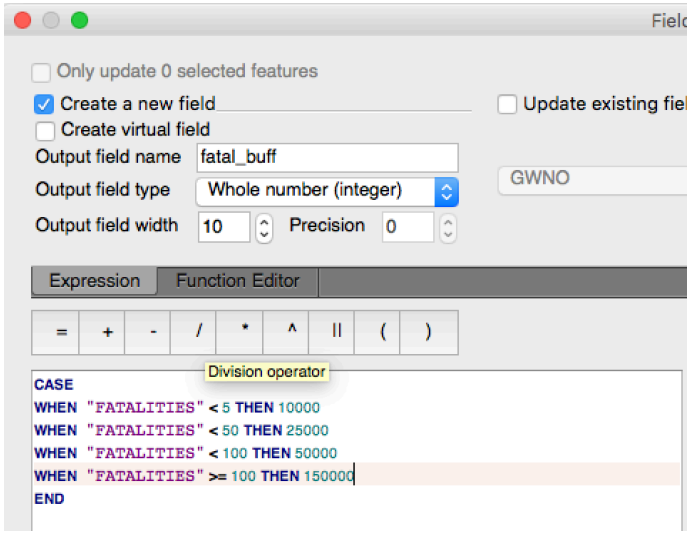
1. Subset the data to filter out any events pre-2014
2. Create spatial buffers (proxy for indirect impact) based on each event's fatalities
3. Use union/intersect to split each buffer into 'impact zones' per country
4. Use zonal statistics to calculate the number of people in each impact zone
5. Non-spatial join the impact zones to countries and visualize final data in a thematic map

Go to acleddata.com/data and download ACLED version 6 as a Shapefile. After downloading, add the data to a new, blank QGIS project. You will quickly notice this is a spatial *points* layer and the attribute table contains a number of relevant variables for each event in the database. Since the dataset contains over a 100,000 events since 1997, we are going to subset the data and select only events from 2014 and 2015 that had at least 1 fatality. Use the 'Select by expression' function  in the Attribute Table to construct a query that does this:




After selecting, save the layer with the 'Save only the selected features' option turned on.

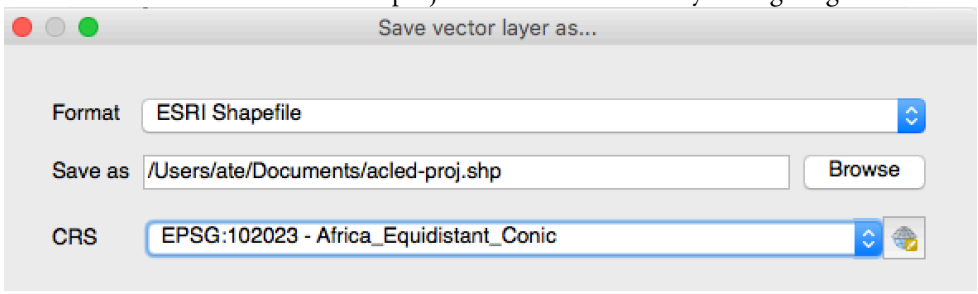
Now that we have a more manageable dataset available, we will now determine the 'impact' of each of these events. To do so, we will make the assumption that the more fatalities a certain event has, the larger the spatial 'reach' of an event. I.e. an event with only a single casualty might only affect citizens in the direct vicinity, while events with multiple casualties reach and impact people in a much larger region. Think, for example, back to the recent attacks in Paris. Despite only taking place in a small part of the city, they affected an entire nation. We can encode this impact by creating a *spatial buffer* around each event, where the size of the buffer is directly related to the number of fatalities in the event.
To do so, we first need to create a new field, based on FATALITIES, that will categorize fatalities and assign a certain buffer size (in meters). Use the field calculator to do so:



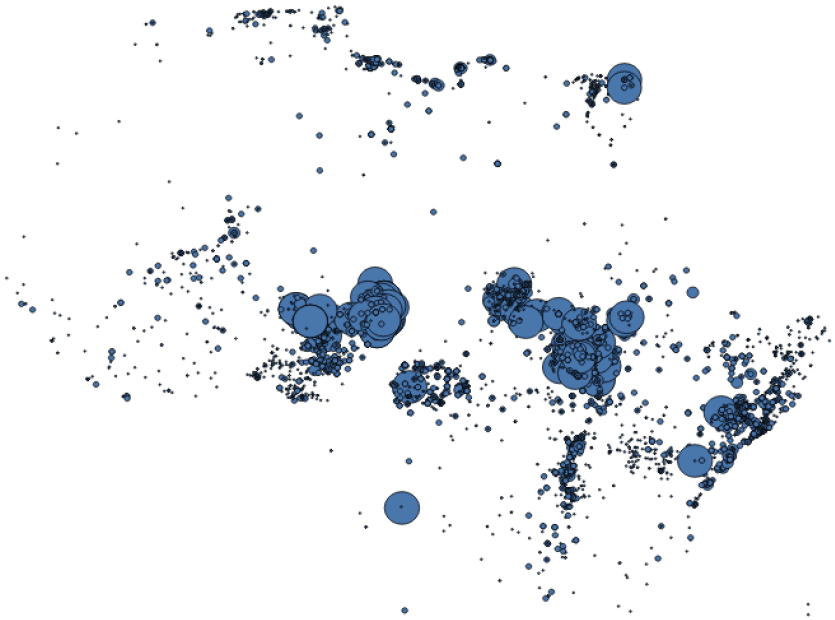

Note that, of course, the spatial impact of each event that we assign here is relatively arbitrary and does not have a scientific basis.

After the field is created, we can use it to create buffers around each field. If you think back on the previous week's topic around projections, you will realize that the projection used for calculating these buffers has a big effect on the consistency of the buffers across Africa. Since we are dealing with an entire continent, it will be impossible to select a projection that is equidistant for all points in our data. We will

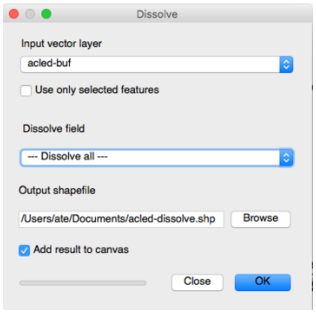
need to compromise and select a projection that is at least tailored to the continent/location of our data. Reproject the ACLED data by saving it again:




Now it is finally time to create our *buffers*. Go to Vector | Geoprocessing Tools | Buffers. Make sure you create the buffers based on the custom field we created before!



Success! The only downside is that each individual event now has its own individual buffer. In other words, overlapping circles are not merged into one. We would like to merge overlaps together into cohesive 'impact zones'. To do so, we can use the *Dissolve* function. It merges polygons together based on a common attribute or, in this case, just merges all polygons in a single layer together into one. Go to Vector | Geoprocessing Tools | Dissolve. Make sure you select the 'Dissolve All' function.



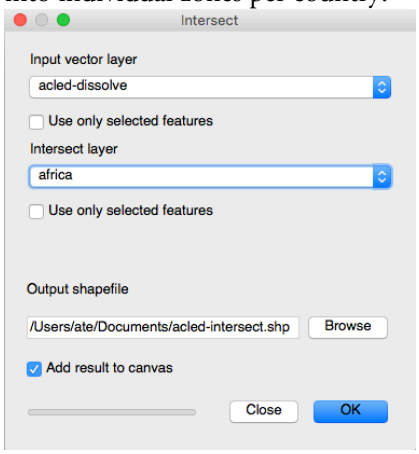

(The calculation for this step might take a few minutes)

Although the buffers are now neatly joined together in single polygon, you would ultimately like to know the number of impacted people *per country*. To do so, we would need to split the entire global 'impact zone' into one single zone per country.

Download the country data (1:50000) from natural earth, or re-use from a previous assignment, and add it to the map as well. Although QGIS neatly places the countries in the right position, the projection used in the country layer is different from our ACLED buffer. *To enable QGIS to calculate similarities or joins between two layers, we always need to make sure the layers are using the same CRS*. Please do so using the following steps:

1. Select only countries within Africa (using the 'Select By Expression' function and the 'Continent' field)
2. Save the layer as a new file, with only the selected features and the same CRS as the ACLED file.

We can now use Vector | Geoprocessing Tools | Intersect to split the impact zone into individual zones per country.



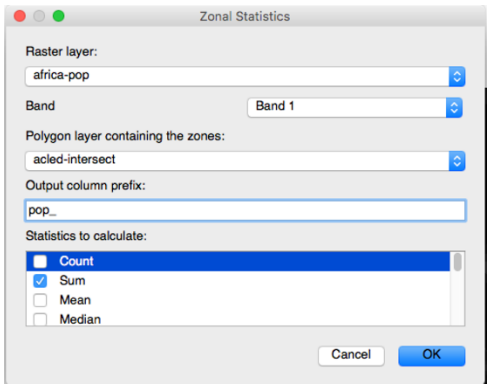

Perfect! You now have exactly what you need: you know for each country in Africa, what areas where (indirectly) impacted by violent events in 2014-2015. All that is left

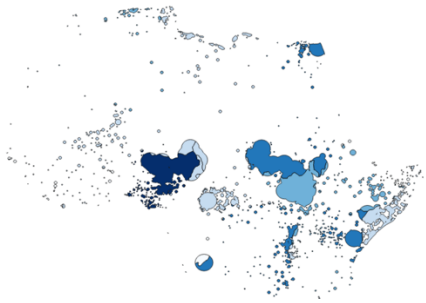
to do figure out how many people leave in each of these areas, in order to finally arrive at the answer to our initial question.

Of course, to do so we would need some kind of information on population. Luckily, NASA releases a global dataset on a regular basis that contains the population in each 1km grid cell on earth. This is perfect for our needs. You can go to http://beta.sedac.ciesin.columbia.edu/data/collection/gpw-v4 to download the worldwide file. It is ~400MB so if you are on a slow internet connection you might use the prepared version in the lab data instead. It only contains Africa which makes the file more manageable.

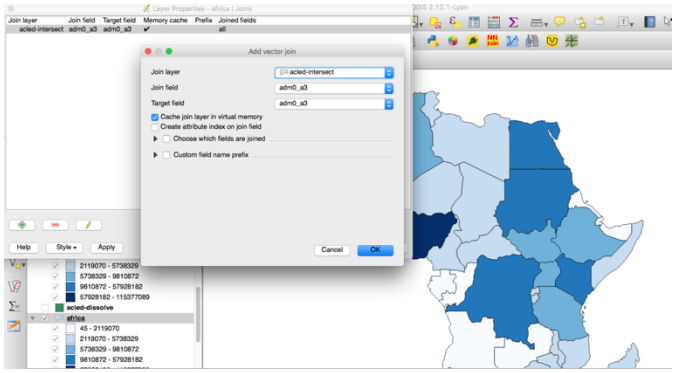
You will see that the data is actually stored as raster file. Each pixel is a square kilometer and has been assigned the total population within that cell. All we need to do is assign cells to each impact zone and sum the total number of people in those cells. However, since our impact zones are vectors and this is a raster, we cannot just do a spatial join as we did in previous labs. *Spatial joins only work if both layers are vectors!*

Luckily, QGIS offers another tool that can combine raster and vector data together. Enable the 'Zonal Statistics' tool in the Plugin Manager and use it to calculate the total sum of people in each zone. One caveat: make sure both the raster and the vector layer are in the same CRS! You will need to reproject either the raster or the vector layer first. Whenever you combine two or more layers together, you always need to be vigilant and make sure they are in the same CRS.



This is exactly what we needed! For visual reasons, it makes more sense to not visualize (only) the impact zones but show the counts per country. To do so we can perform a non-spatial join (cf. Lab 2) to join the impact zones back to their respective countries based on a common attribute.




Thinking back to our original question, *can you see some distinct spatial patterns in the final map?* People in which country are most affected by violent events in 2014-2015? How many people in total were (indirectly) impacted in Africa during those years?

To complete the assignment, create two additional maps thematic maps using the same procedure as above that show different aspects of the same data set. For example, you could choose to show the people affected by violent events in each country during three different time periods (2010-11, 2012-13, 2014-15); choose different 'event type' categories; compare fatal versus non-fatal events etc; or zoom in on a specific region and create maps with more detail than just the people affected per country (e.g. hexagons, provinces).

Finally, finish your set of maps in Illustrator and remember to pay close attention to the design principles (e.g. include all the necessary map elements, think about map context and annotations, visual balance etc.) discussed in the previous labs as well as Krygier & Wood's book.

## Assignment

On the class website, you will find the assignment for this lab. It consists of the map product you created above. The assignment needs to be submitted as a single PDF file. Please make sure you submit the assignment by **March 14.**