Introduction to GIS

Raphaëlle ROFFO

Sciences Po - Urban School

Session 5

Vector Geoprocessing

Today's plan

1. Session 4 recap
2. Tutorial 4 debriefing / questions
3. Navigating data types
4. Geoprocessing: buffer, clip, intersect, merge, dissolve

Session 4 Recap
&
Tutorial Debriefing

Session 4 learning objectives

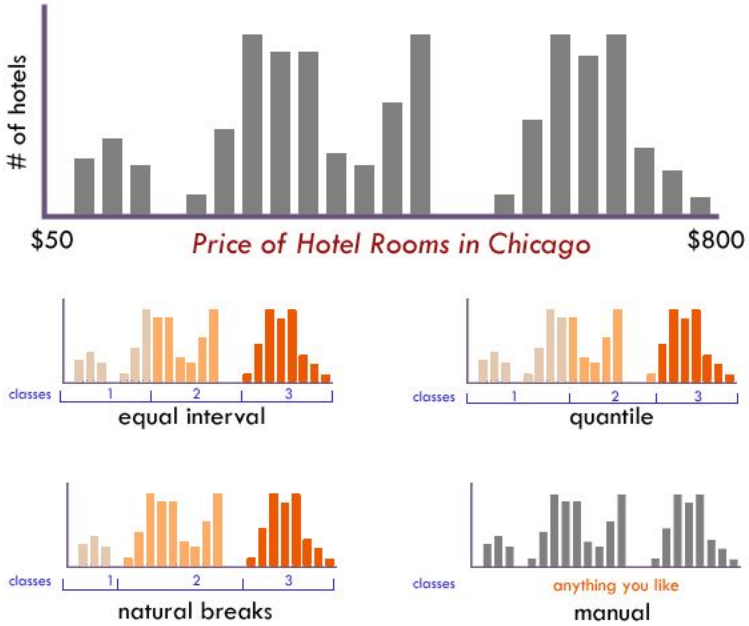
After Session 4 you should now be comfortable with:

- **Cartographic design principles**
- General rules for improving your maps' **accessibility**
- The various **layout elements** you must always include in a map
- What **choropleth** maps are
- How to define **class breaks** (number of classes, break values) - see this article if you're still confused.
- What type of **colour ramp** is more appropriate for your data

Choropleths class breaks: Which method to use?

Please read this GISGeography.com article and this ESRI article for an overview of the main data classification caveats, including choosing the number of breaks and a classification method.

Read this StatsMapsnPix article and see maps on the next slides for a comparison of classification methods on a same dataset (London house prices).



Source: https://www.axismaps.com/guide/data-classification

more on the matter

MAPPING THE TRUTH

map readers, scope the range breaks; map makers, understand options and tradeoffs; everybody, classification matters.
here are examples of two different datasets each classified by three different, common, methods -with way different results...

QUANTILE

unevenly spaced, evenly filled buckets

STANDARD DEVIATION

evenly spaced (to a stat geek), unevenly filled buckets

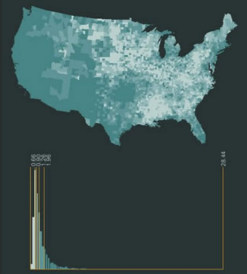
EQUAL INTERVAL

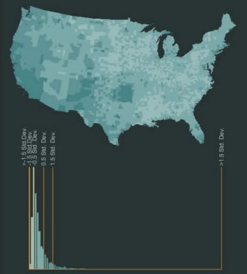
evenly spaced, unevenly filled buckets

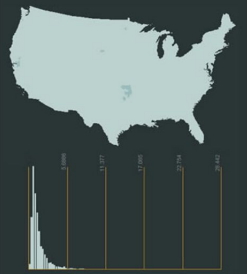
% MULTI-ETHNIC

per county, u.s. census bureau

REALLY SKEWED!



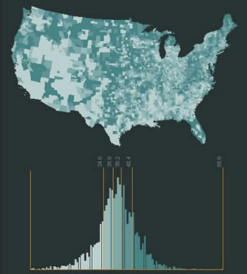


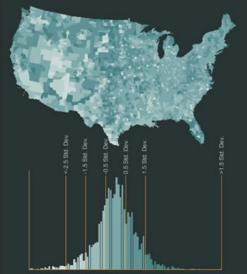


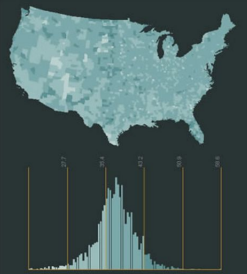
AVERAGE AGE

per county, u.s. census bureau

SUPER NORMAL!







© John Nelson, IDV Solutions

Will always depict variability, even if there is very little variability in the data.

Results in a reliably lively map but can be misleading (and the legend may seem arbitrary).

Consider this method if the data is highly clumped but you still need to tease out visual variation.

Useful for comparing map elements against a baseline average (as such, you'd want a "diverging" color scheme, but whatever).

Tends to tease out visual variation well even with clumped data, but is bad news for bi-modal data.

Good luck explaining the legend.

Useful for "getting" the map quickly and easy, though there may not be much to get.

Will invariably result in a visually bland map unless the data are really flatly distributed (which is rare).

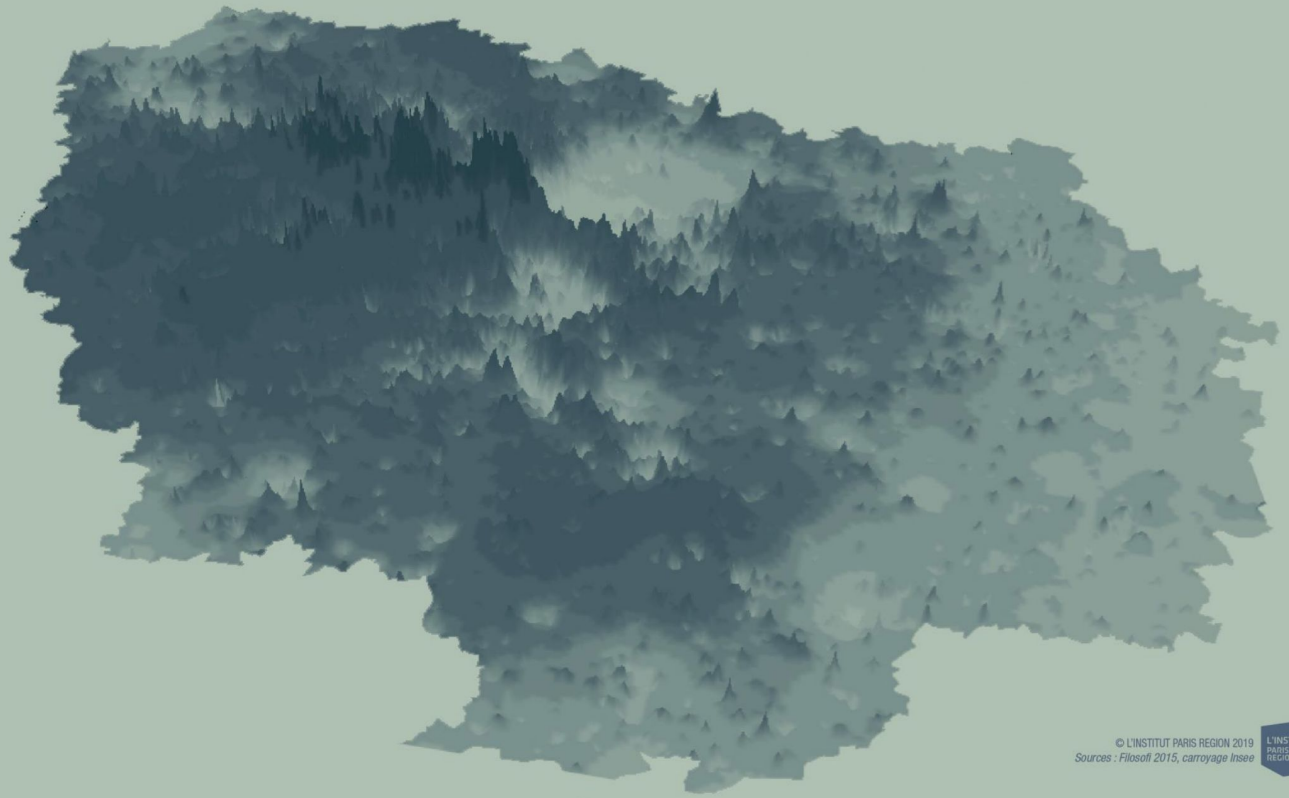
Usually a poor option for social data, but works alright for environmental variables like temperature which are readily thought of in, say, chunks of ten.

PICS ET VALLÉES DES REVENUS

LE RELIEF FRANCILIEN

Voici une carte en relief qui dessine une autre géographie de l'espace francilien. Elle nous invite à visualiser les espaces où se concentrent les hauts revenus tels des pics de montagne, et les plus bas tels des creux de vallée.

Au-delà des simples contours de ce paysage, ce sont les pentes qui nous interpellent : des espaces de transition abruptes parfois difficiles à franchir.



© L'INSTITUT PARIS REGION 2019
Sources : Filosofi 2015, carroyage Insee

L'INSTITUT PARIS REGION

Tutorial debriefing
&
Questions

Tutorial debriefing

The Tutorial looked at styling and symbology. You should now be comfortable:

- Loading data from OpenStreetMap
- Accessing vector symbology settings for your points, lines, polygons
- Creating rule-based symbology
- Adding and styling labels
- Setting scale-dependent visibility
- Saving spatial bookmarks

Additionally, in Tutorial 5 you will get to build a choropleth map and use the print layout composer to export your map.

Questions?

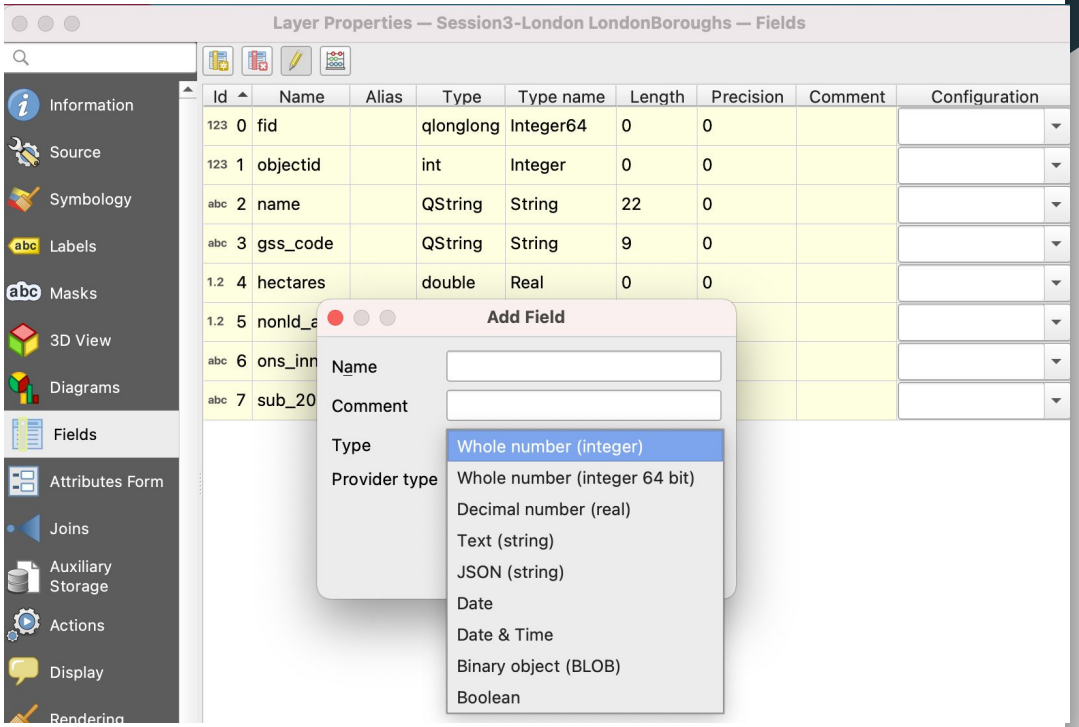
Don’t wait for the coursework deadline to ask questions!

Common questions include: map navigation tools, data types problems, etc.

Navigating data types

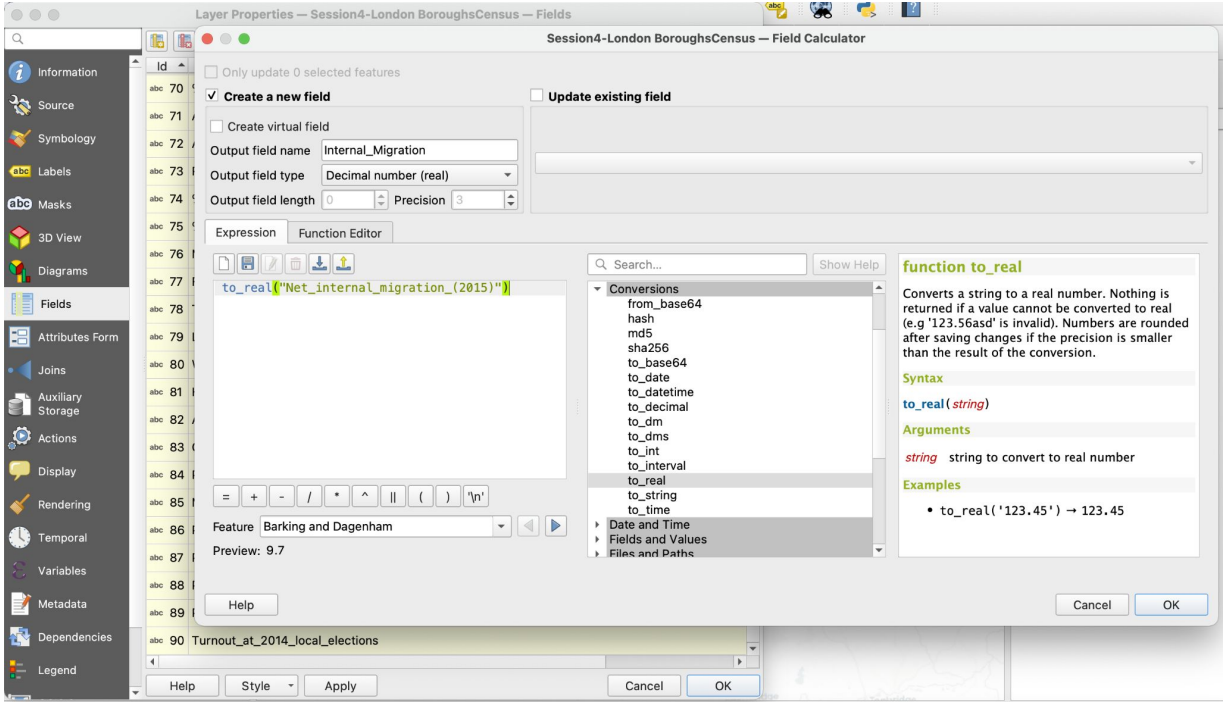
Data types in QGIS

- Integer (whole numbers)
- Real number (decimal numbers)
- String (text)
- Date / time
- Binary / boolean (1 or 0)



You may need to edit the data type!

Most common example: you want to build a choropleth map (graduated symbology) but your numerical variable is currently saved as text.



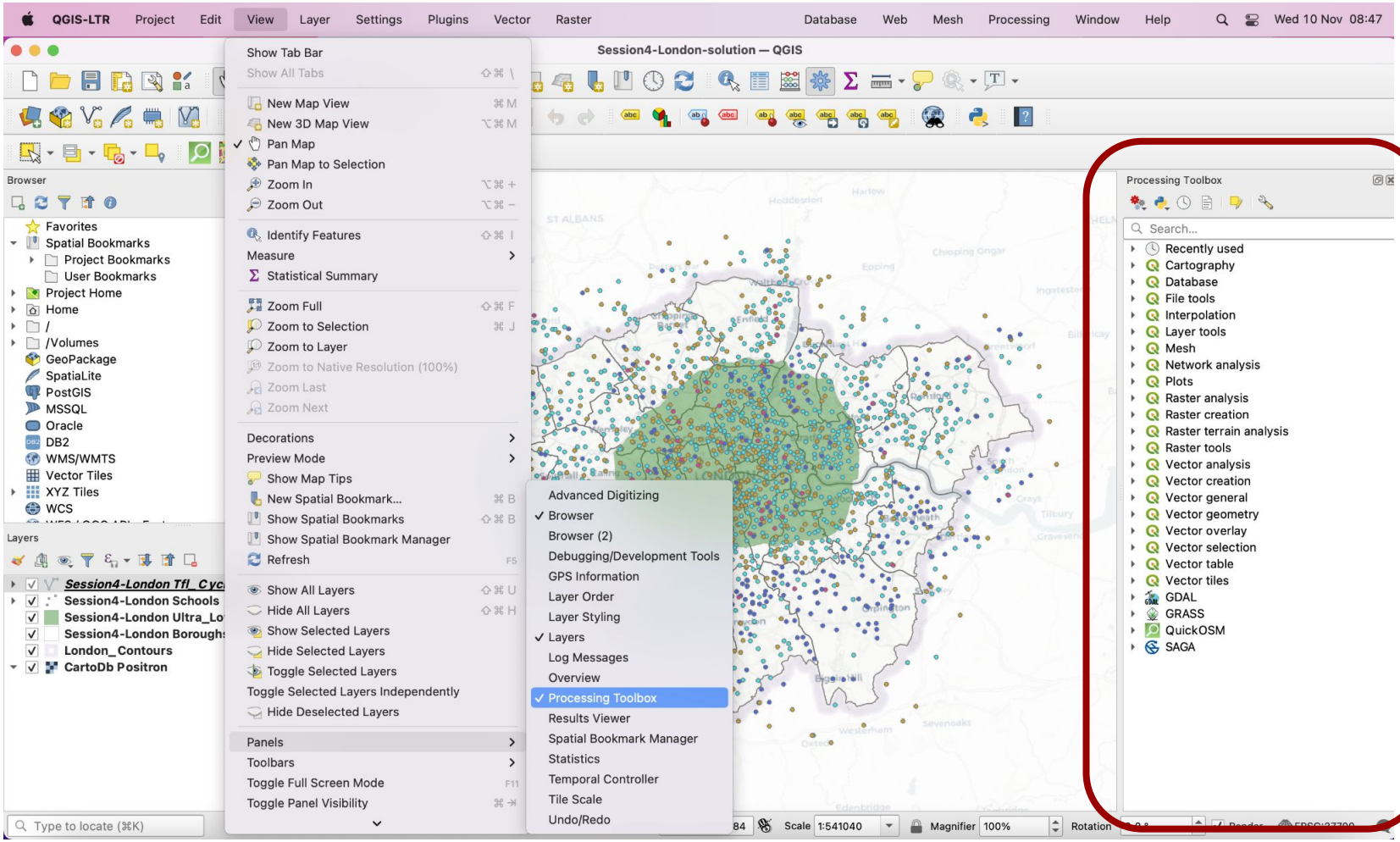
Geoprocessing

Geoprocessing

Geoprocessing tools are operations and tools that allow you to manipulate one or many geographic datasets, and that result in new data (one or many datasets).

In QGIS, these tools can be found in your processing toolbox.

In this introduction course we will only cover the most commonly used vector geoprocessing tools: buffer, clip, difference, merge, dissolve, intersect.


QGIS-LTR
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
Web
Mesh
Processing
Window
Help
Wed 10 Nov 08:47
Session4-London-solution — QGIS
Show Tab Bar
Show All Tabs
New Map View
New 3D Map View
Pan Map
Pan Map to Selection
Zoom In
Zoom Out
Identify Features
Measure
Statistical Summary
Zoom Full
Zoom to Selection
Zoom to Layer
Zoom to Native Resolution (100%)
Zoom Last
Zoom Next
Decorations
Preview Mode
Show Map Tips
New Spatial Bookmark...
Show Spatial Bookmarks
Show Spatial Bookmark Manager
Refresh
Show All Layers
Hide All Layers
Show Selected Layers
Hide Selected Layers
Toggle Selected Layers
Toggle Selected Layers Independently
Hide Deselected Layers
Panels
Toolbars
Toggle Full Screen Mode
Toggle Panel Visibility
Advanced Digitizing
Browser
Browser (2)
Debugging/Development Tools
GPS Information
Layer Order
Layer Styling
Layers
Log Messages
Overview
Processing Toolbox
Results Viewer
Spatial Bookmark Manager
Statistics
Temporal Controller
Tile Scale
Undo/Redo
Browser
Favorites
Spatial Bookmarks
Project Bookmarks
User Bookmarks
Project Home
Home
/
/Volumes
GeoPackage
SpatiaLite
PostGIS
MSSQL
Oracle
DB2
WMS/WMTS
Vector Tiles
XYZ Tiles
WCS
Layers
Session4-London Tfl_Cyc
Session4-London Schools
Session4-London Ultra_Lo
Session4-London Boroughs
London_Contours
CartoDb Positron
Processing Toolbox
Search...
Recently used
Cartography
Database
File tools
Interpolation
Layer tools
Mesh
Network analysis
Plots
Raster analysis
Raster creation
Raster terrain analysis
Raster tools
Vector analysis
Vector creation
Vector general
Vector geometry
Vector overlay
Vector selection
Vector table
Vector tiles
GDAL
GRASS
QuickOSM
SAGA
Type to locate (⌘K)
Scale 1:541040
Magnifier 100%
Rotation

The Processing toolbox

All the geoprocessing tools can be retrieved through the processing toolbox.

You can search by name, or navigate by category.

The tools you've last used will be available to you in the *Recently used* category.

Buffering

Take an input vector and specify a distance.

→ The tool creates buffer polygons around your input features.

Example use: find out all areas that fall within 1km of a school.

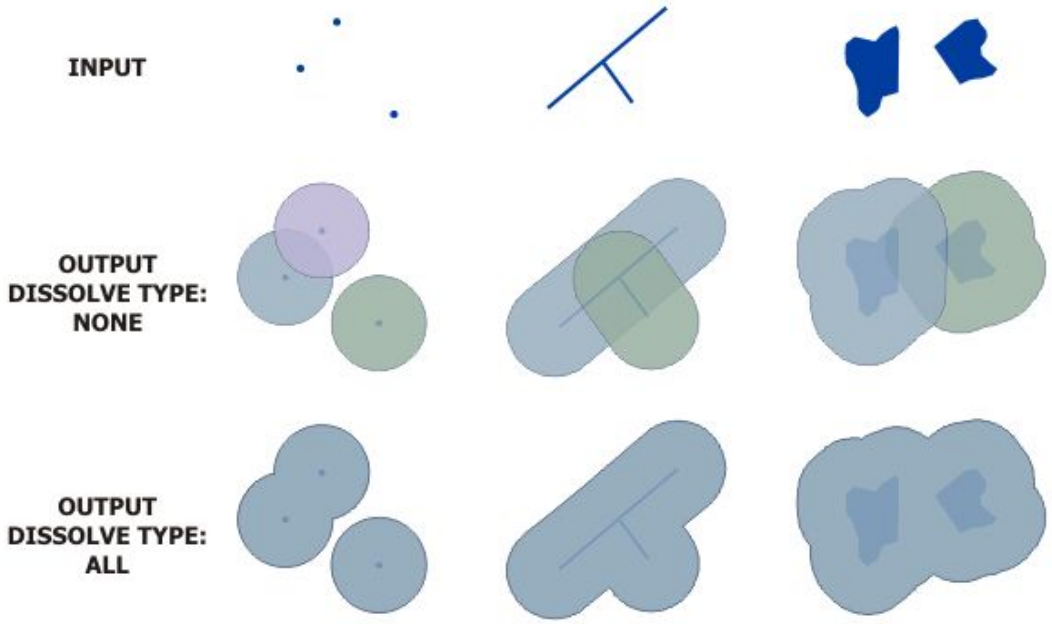
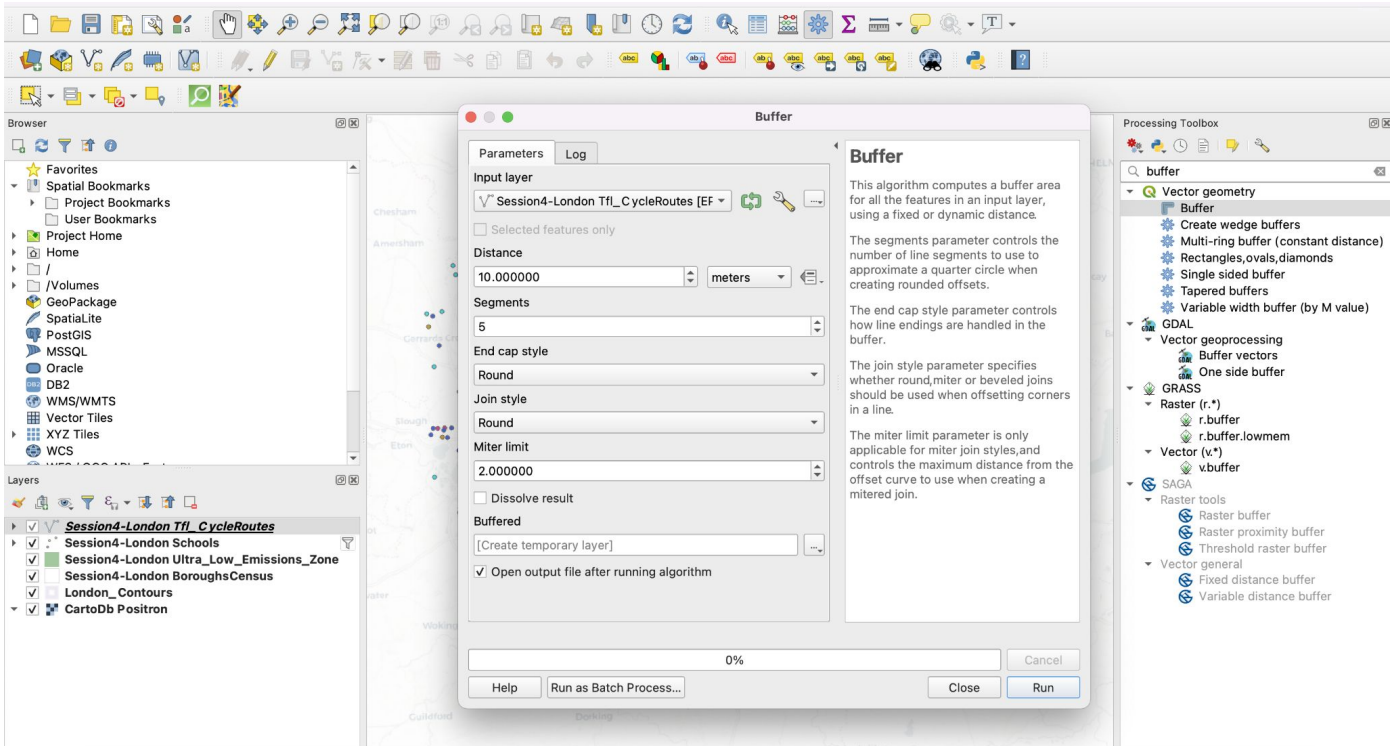


Image credit: ESRI

Buffering


Browser
Favorites
Spatial Bookmarks
Project Bookmarks
User Bookmarks
Project Home
Home
/
/Volumes
GeoPackage
SpatiaLite
PostGIS
MSSQL
Oracle
DB2
WMS/WMTS
Vector Tiles
XYZ Tiles
WCS
Layers
Session4-London Tfl_CycleRoutes
Session4-London Schools
Session4-London Ultra_Low_Emissions_Zone
Session4-London BoroughsCensus
London_Contours
CartoDb Positron
Buffer
Parameters
Log
Input layer
Session4-London Tfl_CycleRoutes [EF
Selected features only
Distance
10.000000
meters
Segments
5
End cap style
Round
Join style
Round
Miter limit
2.000000
Dissolve result
Buffered
[Create temporary layer]
Open output file after running algorithm
Buffer
This algorithm computes a buffer area for all the features in an input layer, using a fixed or dynamic distance.
The segments parameter controls the number of line segments to use to approximate a quarter circle when creating rounded offsets.
The end cap style parameter controls how line endings are handled in the buffer.
The join style parameter specifies whether round,miter or beveled joins should be used when offsetting corners in a line.
The miter limit parameter is only applicable for miter join styles,and controls the maximum distance from the offset curve to use when creating a mitered join.
0%
Cancel
Help
Run as Batch Process...
Close
Run
Processing Toolbox
buffer
Vector geometry
Buffer
Create wedge buffers
Multi-ring buffer (constant distance)
Rectangles,ovals,diamonds
Single sided buffer
Tapered buffers
Variable width buffer (by M value)
GDAL
Vector geoprocessing
Buffer vectors
One side buffer
GRASS
Raster (r.*)
r.buffer
r.buffer.lowmem
Vector (v.*)
v.buffer
SAGA
Raster tools
Raster buffer
Raster proximity buffer
Threshold raster buffer
Vector general
Fixed distance buffer
Variable distance buffer

Clipping

Take some target features (your input) and overlay a polygon like a “cookie-cutter” on top; you only keep the target features that fall within the clip polygon.

→ The boundaries of the clip are imposed on your target features and the rest is discarded.

Output = area that’s in input A **AND** in input B

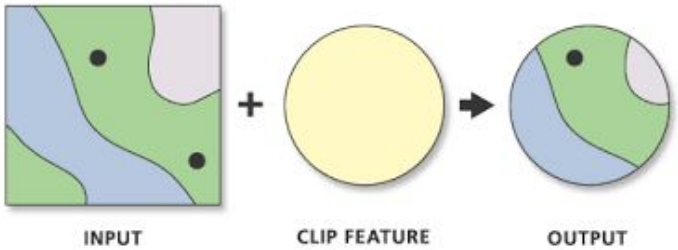
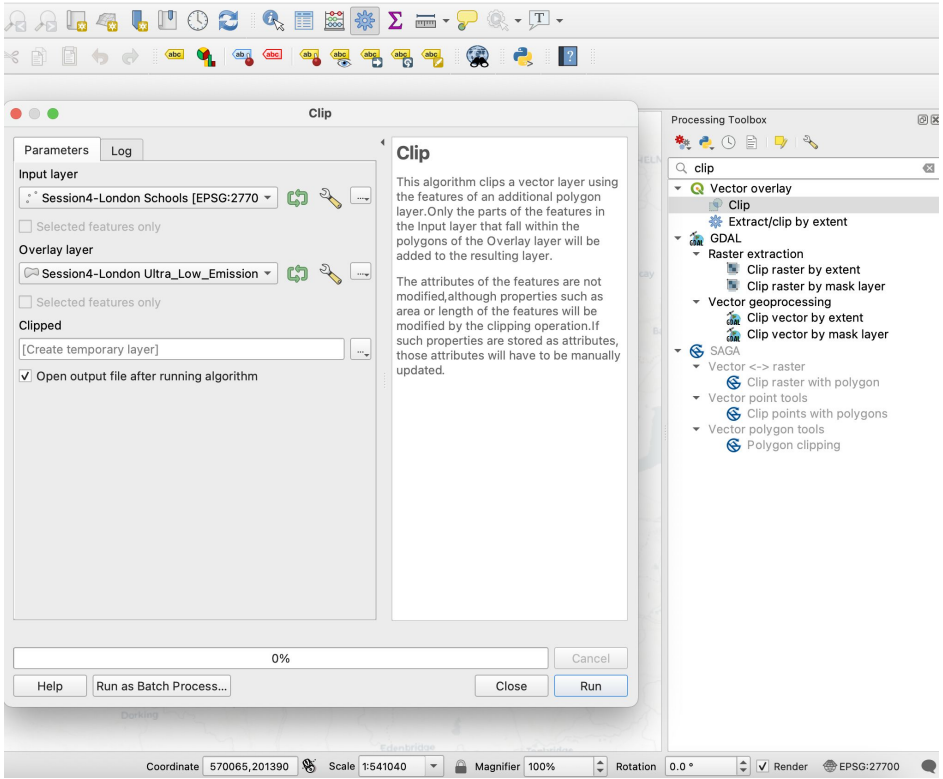


Image credit: ESRI



Difference (Erase)

You can think of it as a reverse clip!

Output = area that's in input A **MINUS** input B

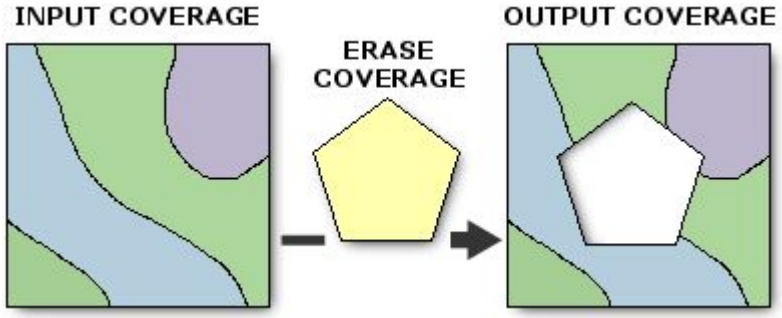
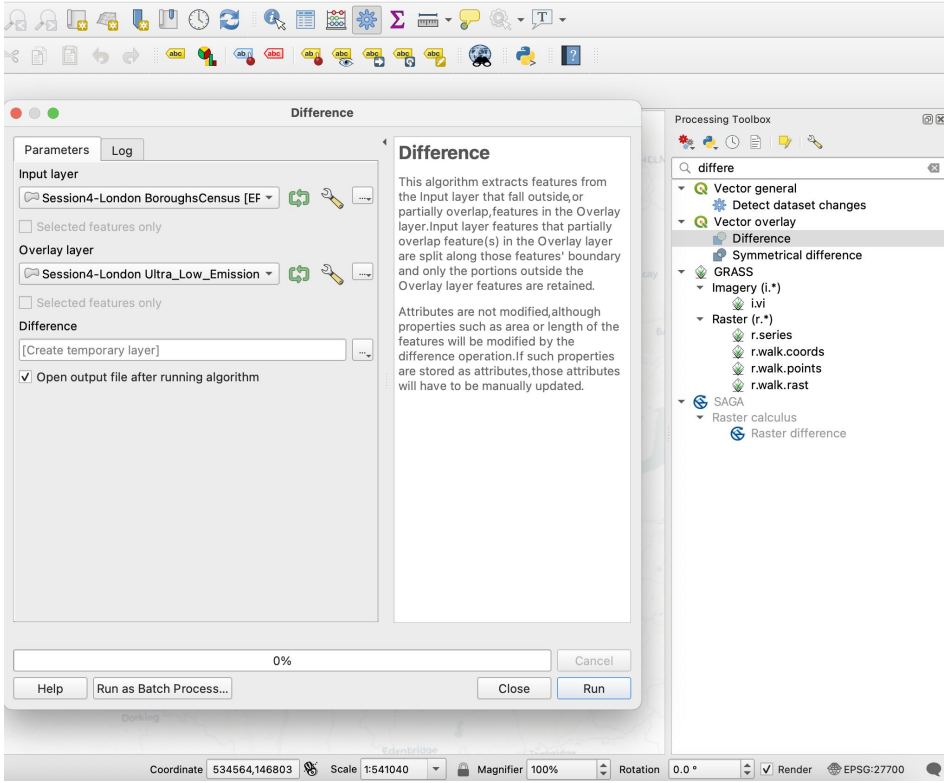


Image credit: ESRI



Intersect

Input two vector layers; the output will be all features - or portions of features - that overlap in both layers.

Output = input B area that overlaps with input A

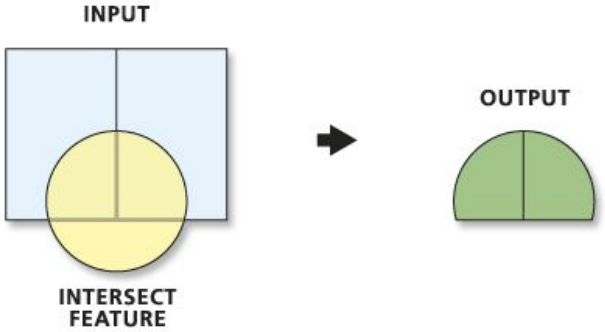
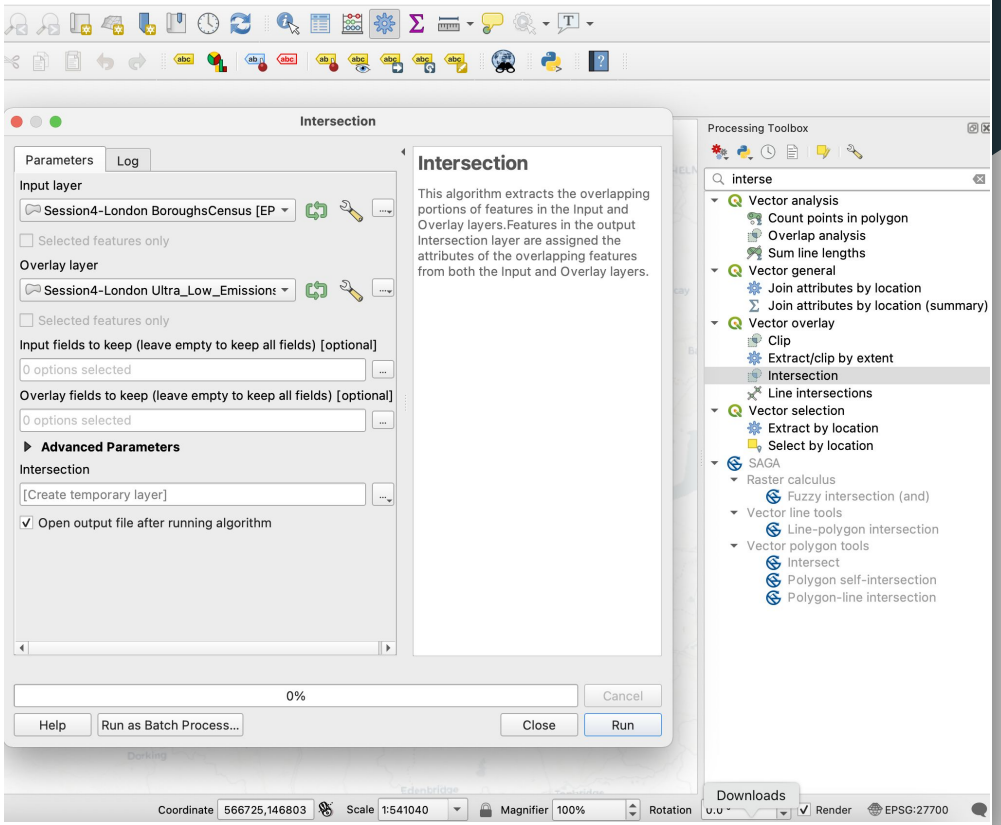


Image credit: ESRI



Symmetrical Difference

You can think of it as a reverse intersect! The output layer will contain all features that are only present in **one and not both** the input layers.

Output = area that's in input A **OR** in input B but not both

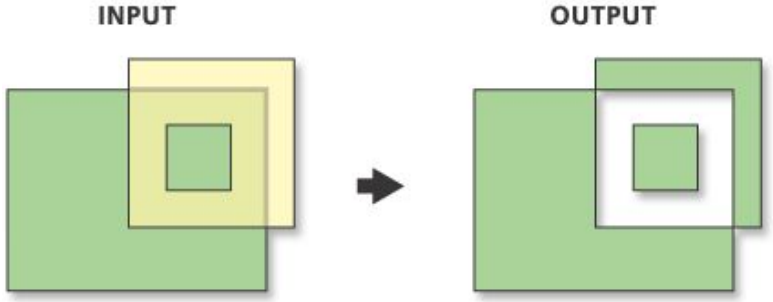
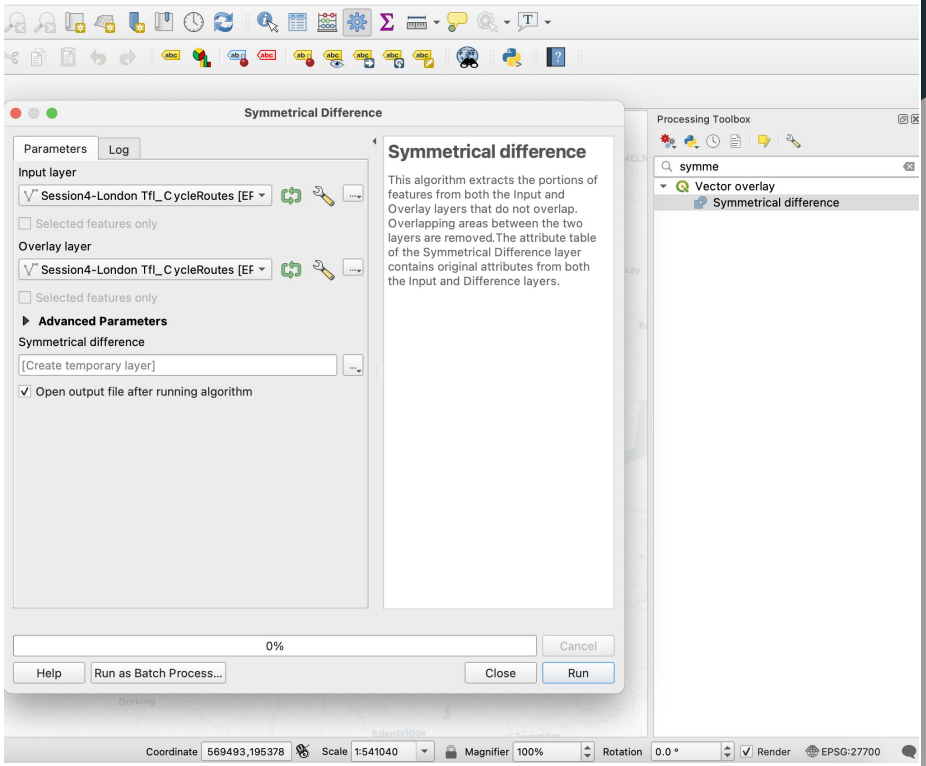


Image credit: ESRI



Merging

Combine multiple vector layers of the same data type (either all points, all lines or all polygons) into a single, new output dataset.

Output = **SUM** of input A and input B areas

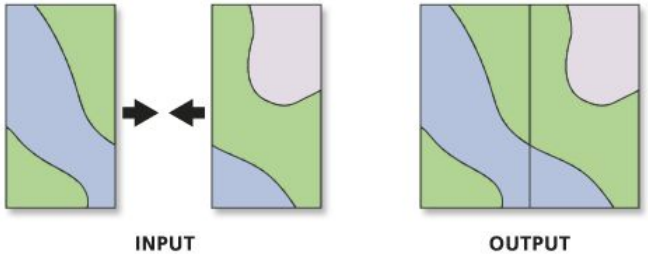
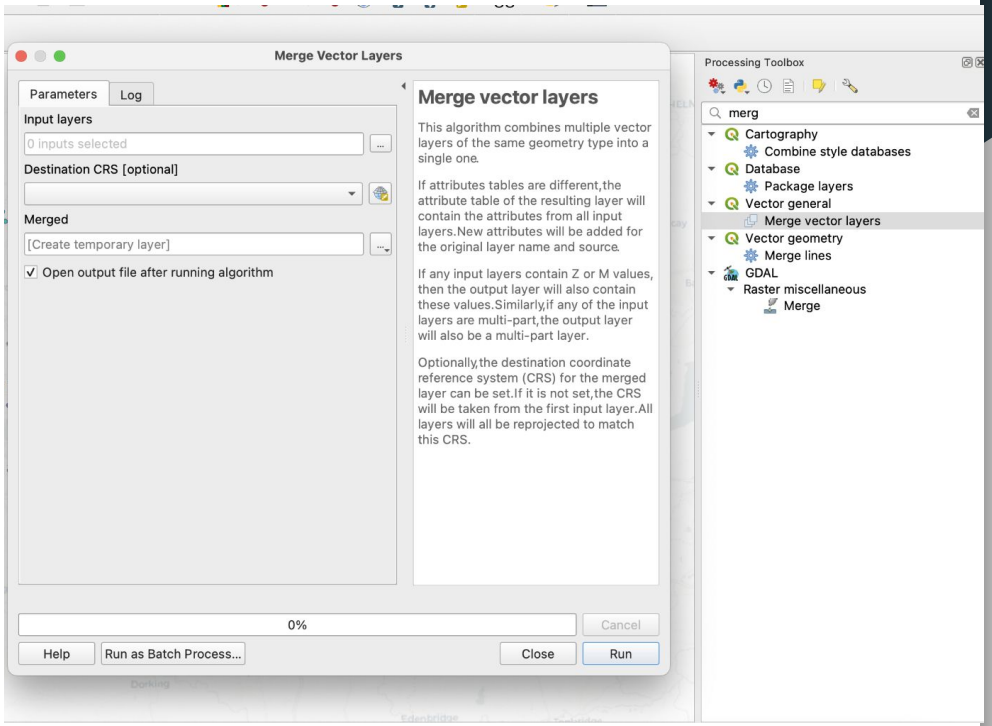


Image credit: ESRI



Dissolving

Take a vector layer and choose and merge adjacent polygons, lines, or regions that have the same value for a specified field.

For example, you can “dissolve” all counties within a province, or all wards within a borough.

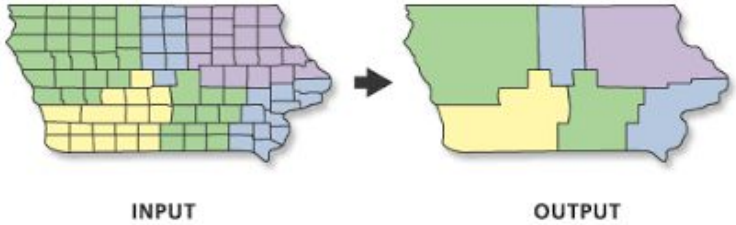
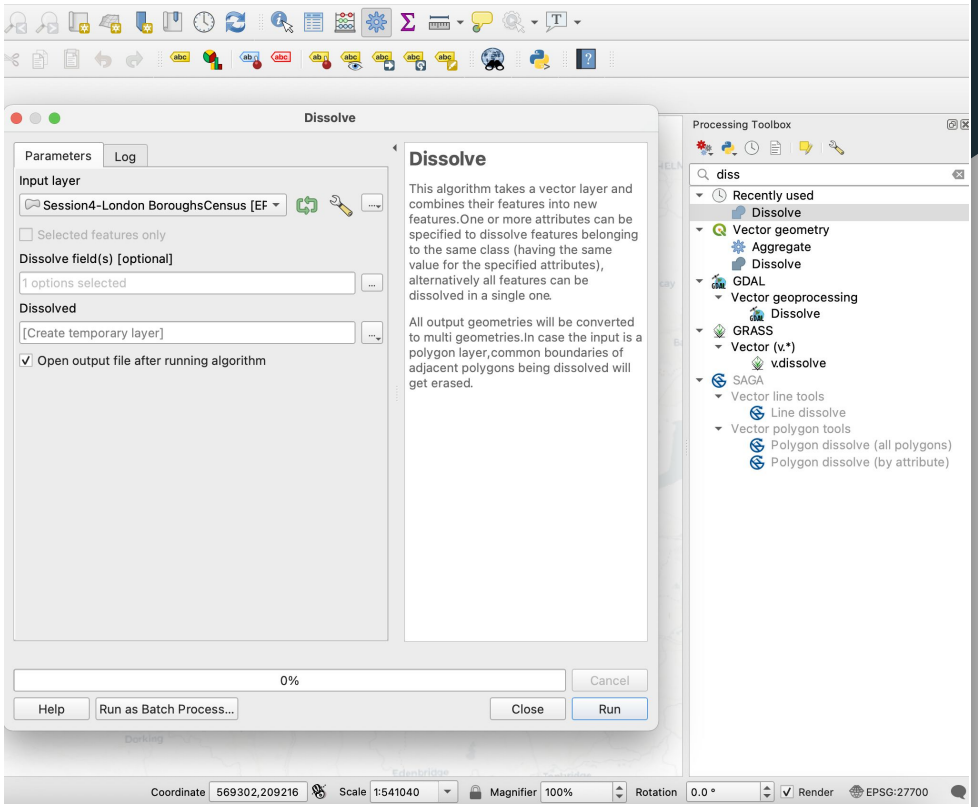


Image credit: ESRI



Summary

Many more operations are available to you; in doubt, your best allies are:

- the QGIS documentation,
- GIS Stack exchange on Stack Overflow,
- ...and of course your favourite search engine!

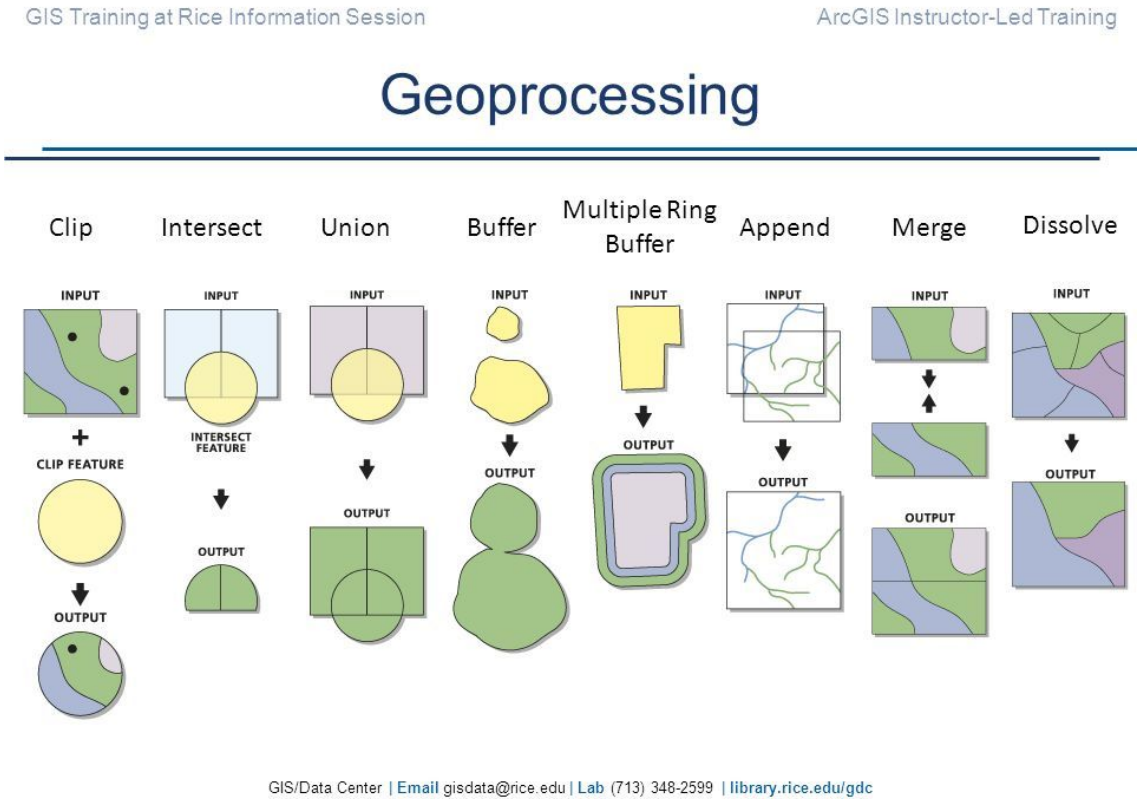


Image credit: Rice library

Homework
(next session is in person)

Homework

1. Do the Session 5 tutorial and preferably also try Session 6 tutorial.
2. Make sure you have a topic for your final coursework! (see next slides)
3. Do use Slack if you have questions (#help), and more importantly to start helping each others out too! Explaining to someone else is a great way to test your own understanding.

Coursework

Final Coursework

The final coursework is a **map production exercise**. Find a research question, carry out a simple map production workflow, and write a report summarizing your findings. You will be assessed on your capacity to:

- Frame your research question in a clear and concise manner, and ensure a few maps can provide interesting insights,
- Identify datasets that are relevant to answering your question *(technically here you may want to work backwards and use the data you already have to find your research question),*
- When appropriate, use table joins to “enrich” your vector data,
- Carefully choose your symbology, and ensure you map is accessible and colour-blind safe,
- Create map exports complete with all key cartographic elements (title, legend, data source etc).
- Analyse and interpret the patterns that emerge from your maps, explaining what this might mean in terms of policy or research outlooks.

Final Coursework

Some practical points:

- Deadline: Monday 24th October, 23.59 Paris time. *If you're late, minus 1 point for each day behind the deadline.*
- Work in groups of 2-3 students
- Work on a city of your choice
- The report must be 3 pages minimum, 5 pages *maximum*
- You are encouraged to use the data provided for the tutorials, however if you want to challenge yourselves, you will get **+1 point bonus for working with data you have sourced yourselves**.

Final coursework: Proposed outline

You are strongly encouraged to follow this outline:

1. Executive summary (maximum ½ page, bullet points are fine)
2. Introduction / Problem / Context
3. Data sources in a table
4. High-level methodology. Keep it very short but use precise terminology. Explain your symbology choices including your choice of class breaks if you built a choropleth.
5. A minimum of 2 maps. Careful, you only have 5 pages maximum in this report so these maps must be relevant to answering your policy question
6. Analysis of the findings (half a page)
7. Conclusion / next steps for policy makers or future research outlook (half a page)

In this exercise, concision and precision are key!

Final Coursework: Marking Criteria

The marking criteria reflect the learning outcomes expected at the end of this module. Students should be able to:

- Formulate a research question suitable for GIS analysis
- Source relevant data and assess their relevance based on the metadata provided
- Load and style data using appropriate symbology
- Ensure balance and accessibility of their map (colour-blind safe, etc)
- Produce a proper map export, complete with all key cartographic elements (title, legend, north arrow etc.)
- Provide a detailed methodology section, focusing on key decisions (choice of dataset, geoprocessing steps or class breaks chosen for a choropleth, etc.)
- Draw policy insights from their maps and translate those into applicable policy recommendations **or** future research outlooks. Please be very explicit!

Final Coursework: Marking Criteria

Criteria:

- Research question and whether your methodology adequately addresses it (20%)
- Methodology and choice of relevant datasets (20%)
- Quality of the map outputs (40%)
- Relevance of recommendations / Insights (10%)
- Overall quality of the writing, structure and visual clarity of the report (10%)