

advanced qgis

adam okulicz-kozaryn

`adam.okulicz.kozaryn@gmail.com`

this version: Tuesday 21st November, 2023 13:45

outline

geocoding

SQL

geo-processing

[*] bivariate map, timeseries tool, putting on raster

outline

geocoding

SQL

geo-processing

[*] bivariate map, timeseries tool, putting on raster

geocoding: address→(lat,lon)

- https://s3.us-east-2.amazonaws.com/journeynorth.org/images/graphics/mclass/Lat_Long.gif
- btw, if already got X/Y lat/lon:
just add your csv with “Add Delimited Text Layer” tool
make sure geometry definition tab (X,Y) fits your data
- may need to specify crs: wgs84 (plain lat/lon)
- if your points like in ocean or arctica flip x/y lat/lon
- may not take missing values—make sure all lat/lon existing
- decimal points important—need like 5 to have precision

geocoding: address \rightarrow (lat,lon)

- say that we have some addresses and we want to geocode them
- `https://github.com/theaok/data/raw/main/apartments-for-rent.xls`
- open, looks reasonably clean, save as csv

MMQGIS-Geocode

- MMQGIS-Geocode-Geocode CSV with Web Service
- Input CSV, and make sure Address Field, City Field, State Field are right; best if you give more info
- Web Service: OpenStreetMap/Nominatim
- make sure notfound.csv (and output shp) saved where you can write!
- >qgis3.5, seems can have everything just under address!
- hit “Apply”, do note how many found, hit “Close”

important to check!

- add basemap to check location
 - does it make sense? houses in river or park?
 - zoom-in to street, click some points with “identify tool”: pop-up address—does it match with the street?
 - usually some miscodings, say few percent
 - usually because the address is misspelled or incomplete
- do see notfound.csv: mostly those with a range of street numbers
- need to fix them/adjust them:
 - google addr, adjust accordingly to get it clean, and manually change in csv

outline

geocoding

SQL

geo-processing

[*] bivariate map, timeseries tool, putting on raster

SQL: Structured Query Language

- full blown (not in qgis) SQL has few additional things
 - English-like, just strict syntax rules
- also a job market skill: put 'basic SQL' on your linkedIN next to 'gis' skill
 - very easy to master in no time
- <https://www.youtube.com/watch?v=jJeae7PJVv4>

advanced filter (expression): SQL

- nj_counties

<https://docs.google.com/uc?id=1xJDhcRCkgv7k4tNCa720og5bohV6dTB2&export=download>

- nj_counties-Open Attribute Table

- bottom left box- “Advanced Filter (Expression)”

- Fields and Values “REGION”

- and on the right Load values: “all unique”

- then we can type; or faster double click:

- “REGION” = 'CENTRAL' and hit OK

- now easy to modify at the bottom of table, say:

- “REGION” = 'CENTRAL' OR “REGION” = 'SOUTHERN'

- “REGION” = 'CENTRAL' AND “POP2010” > 598349

regular expressions

- can also match part of a string:
 - (may need to type in as opposed to cp, pciky about quote format)
 - (also note if err, it says so at the bottom in red)
- `regexp_match(" COUNTY", 'C.*N')`
- `regexp_match(" COUNTY", '^C.*N')` must start with 'C'
- `regexp_match(" COUNTY", '^C.*N$')` and end with 'N'
- then can select matching data
- right click layer, Export-Save Selected Features As

more about calculator

- right click nj_counties, open attr table, click calculator icon at top
- Create a new field: hiPoDe10
- field type: integer

CASE

```
WHEN "POPDEN2010" > 0 AND "POPDEN2010" <5000 THEN 0
```

```
WHEN "POPDEN2010" > 5000 AND "POPDEN2010" <10000 THEN 1
```

END

- note, unspecified val are NULL-see in attr tab

saving selection often necessary

- keep in mind simplicity principle!
- drop all unnecessary clutter
- do not map things that you don't care about
- and often good to have several subsets of same layer
- say can label only one subset
- use different symbology by subset etc

outline

geocoding

SQL

geo-processing

[*] bivariate map, timeseries tool, putting on raster

this is a whole bag of tools

- we switch gears a little and discuss
 - more advanced topics beyond mapping
 - more like typical GIS/IT stuff
- we will just cover few tools:
- there are dozens of them
- you may definitely do some of not covered
 - esp 'Vector', and also 'Plugins'

dissolve

- nj_counties

<https://docs.google.com/uc?id=1xJDhcRCkgv7k4tNca720og5bohV6dTB2&export=download>

- dissolve into a larger area

- (get rid of inside boundaries)

- Vector-Geoprocessing Tools-Dissolve

- nj_counties

- “dissolve field:” REGION

dissolve your way

- can dissolve into your own categories/definitions
- let's dissolve nj_counties into south and north jersey
- Open attribute table-toggle editing-New column-integer:
'southNorth'
- sort on REGION and mark southern regions with 1, and the rest with 0
- can automate this using calculator as we did earlier
- may also highlight the row to see which county is where
- Vector-Geoprocessing tools-Dissolve
- "Dissolve field:" southNorth
- often you will have to do something like this
- no way to find a shapefile for South Jersey online!

simplify

- remember from principles: simplify as much as possible
- simplifying polygons means dropping vertexes, so that polygons are defined by fewer coordinates **draw**
- it reduces size of a file
- Vector-Geometry tools-Simplify
 - Input: 'nj_counties'
- play with “tolerance” to achieve desired simplicity: try 1000
 - and turn off/on to compare to original: see the difference?
 - can also simplify lines (fewer nodes)
- and i guess you can also simplify points (fewer dec points)
- ref <http://gis.stackexchange.com/questions/25914/>

how-to-smooth-generalize-a-polygon-in-qgis

centroids

- calculate a center of a polygon or turn polygon into a point
 - useful when merging non-overlapping polygons—say congressional districts and counties
 - then you can calculate centroid of one of those and merge with polygons of the other layer if a centroid is in that polygon using spatial join
- draw a picture
- Vector-Geometry tools-Polygon centroids
 - Input: nj counties

centroids

- note: the new shapefile will have the same data
- can now map another var and overlay on another var
 - both points and polygons with some symbology
- let's map population for polygons
 - and population density for points
 - note: make points bigger to see symbology well
- this solves the problem of showing 2 vars in one map

buffering

- kind of opposite of centroids:
- buffer (circle) around a point or poly or line; eg:
 - 'dry zone' around schools
 - waste processing plants and houses
 - 2-mile heavy pollution around hwy
 - walkability to healthy stores, etc
- load nj universities
 - https://github.com/theaok/data/raw/main/hsip_colleges.zip
- Vector-Geoprocessing Tools-Buffer
- use 20,000 feet
- note: buffer is a new layer and then can spatially join it with another layer

measure line tool

- same toolbar at the top right as Identify tool
- note different measuring units
- measure distances—how far from a point to point
- measure radius to make sure it is right
- and measure say jogging route in segments from RU to 8th and market
- note we will do the same in google maps

example: environmental problems around univ

- download and add to qgis
- https://docs.google.com/uc?id=1T_n1y_Mj5yQiWpZwrbbuFFwmIVJ2QWFZ&export=download
- make smaller, say size of .4 so can better see

MMQGIS-Combine-Spatial Join

- Output: Buffer
- Spatial Operation: Contains
- Data (Join) Layer: NJ contaminated sites
- Fields: NAME (so that in new layer we have name of univ)
- Field Operation: Sum

investigate

- open attr table of merged shapefile
- click 2x last column 'COUNT' to sort descending
- under 'NAME' we find that 'NEW JERSEY MEDICAL SCHOOL'
- has biggest problem! over thousand contaminated sites
- from the table can select schools with greatest problems
- and take some measures to help with the situation

related: select by location (say id problematic ones)

- say select polluted sites within 1000 ft from a school
- Vector-Geoprocessing Tools-Buffer
- Vector-Research Tools-Select by location
- Select features from: NJ Contaminated Sites
- Gemetric predicate: are within
- By comparing to the features from: Buffer
- and then: NJ contaminated sites-Export-Save selected features as
 - and save as csv
 - got 80 places we can call and ask to clean up

outline

geocoding

SQL

geo-processing

[*] bivariate map, timeseries tool, putting on raster

bivariate map

- make 2 thematic map with 3 classes of same layer (copy to make 2nd one)
- then under Layer Rendering under Blending Mode do Multiply (have both layers selected so it knows which one to multiply)
- and there is “bivariate legend” plugin
- if export doesnt work, can just screenshot
- <https://bnhr.xyz/2019/09/15/bivariate-choropleths-in-qgis.html>
- <https://www.youtube.com/watch?v=82zONQozY54>

time series tool

- time series tool: https://www.qgistutorials.com/en/docs/3/animating_time_series.html
- And to change a field into date format, use the "to_date" expression (under Date and Time) in the field calculator. I used `to_date(year,'yyyy','fr')` because I only cared about the year.
- for example see <https://maps-dppa.camden.rutgers.edu>
"Congressional_Redistricting" under PS_4

puting on raster

- plugin rasterDataPlotting
- just fix it in on vector using easily identifiable stuff like roads and water