Geospatial Data Analytics for Business 2

Giorgio Chiovelli, Sebastian Hohmann

19/03/2020

Introduction

Today's topic

Spatial Analysis for Real Estate

- Suppose you run a real estate company.
- You want to buy properties in central London.
- You want to know which factors impact property value

Introduction

Today's topic

Spatial Analysis for Real Estate

- Suppose you run a real estate company.
- You want to buy properties in central London.
- You want to know which factors impact property value

Which factors would you take into account?

Introduction

Today's topic

Spatial Analysis for Real Estate

- Suppose you run a real estate company.
- You want to buy properties in central London.
- You want to know which factors impact property value

Which factors would you take into account?

- School quality

Introduction

Today's topic

Spatial Analysis for Real Estate

- Suppose you run a real estate company.
- You want to buy properties in central London.
- You want to know which factors impact property value

Which factors would you take into account?

- School quality
- Crime

Introduction

Today's topic

Spatial Analysis for Real Estate

- Suppose you run a real estate company.
- You want to buy properties in central London.
- You want to know which factors impact property value

Which factors would you take into account?

- School quality
- Crime
- Pollution

Introduction

Today's topic

Spatial Analysis for Real Estate

- Suppose you run a real estate company.
- You want to buy properties in central London.
- You want to know which factors impact property value

Which factors would you take into account?

- School quality
- Crime
- Pollution
- Do these factors correlate with actual housing market?

Introduction

The plan for today

1) Introduction to Processing Modeler

- How to automate steps in QGIS

2) Processing spatial data

- Census Enumeration Areas
- School (location and quality)
- Crime
- Pollution

3) Visualizing Spatial Patterns

- What can maps tell us about spatial correlation

Graphical Processing Modeler

What is it good for?

Making life easier

- Geoprocessing tools have many options, often need to be executed in succession.
- Clearly, we want to automate Geoprocessing somehow to make it faster.

Making research replicable

- Geoprocessing can become complex quickly.
- To make research replicable (also for ourselves!) we want to automate as much as possible of the GIS workflow.

An improvement but not the last word

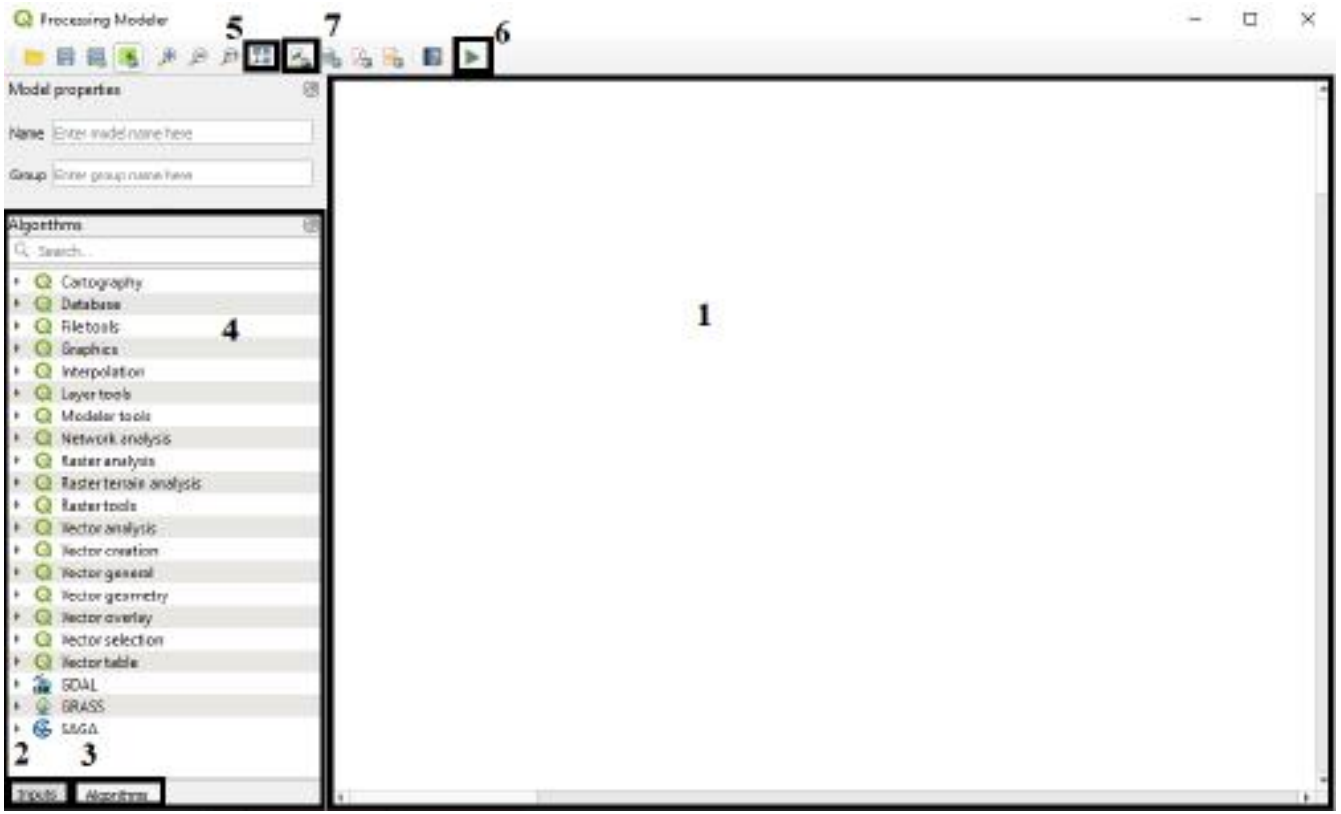
- The Graphical Modeler is a first (though not ideal) step towards automation.
- It also functions as a nice bridge towards full automation in Python.

Graphical Processing Modeler

Getting started

Starting the modeler

- Processing → Graphical Modeler OR
- In the Processing Toolbox panel, click on → Create New Model OR
- `Ctrl+Alt+M`



1. Canvas: showing workflow of processing tools
2. Defining inputs
3. Choosing algorithms (processing tools) to transform inputs; currently algorithm tab selected
4. List of geoprocessing algorithms by category
5. Zoom canvas to full extent
6. Run model
7. Export model to python script

Graphical Processing Modeler

How to build a model I

Principle: save only the final output

- Geoprocessing can be complex – many steps executed in sucession
- You don't care about intermediate outputs, just final product (usually a .csv)
- Most tools we use in the graphical modeler (and later python scripts) store the intermediate products in temporary files (under Windows: `/AppData/Local/Temp/processingXYZ`) if you ask it to
- Otherwise, things are done in memory

Graphical Processing Modeler

How to build a model II

Chain tools together

- Search for a tool under Algorithms
- Drag the tool onto the canvas
- (double) click on the tool to open its configuration window
- For tools that don't use primary inputs from your disk, use output from a previous tool as an input, `'Output name' from algorithm 'Algo name'`
- Most tools have a field with a grey text *[Enter name if this is a final result]*
  - If you want to load results of some steps into the GUI, enter a layer name, o/w leave blank
  - For the final output producing tool, enter a name

Graphical Processing Modeler

How to build a model III

Run the model

- Click  to run the model.
- There is one output parameter for every *[Enter name if this is a final result]* that you replaced with a layer name above
- Uncheck "Open output file after running algorithm" for outputs that you don't want to load as layers after the tool is done
- Enter a full path name for the final output file (usually a .csv)

Defining the inputs

- LSOA areas based on the 2011 census. These are stored in the shapefile `LSOA_2011_London_gen_MHW.shp`
  - They correspond to the Census Enumeration Areas in Central and Metropolitan London
  - Open Source: `https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london`
- Atlas of School. Shapefile `school_data_london_Atlas_2016.shp`
  - School GSP location + Ofsted quality score
  - Open Source: `https://data.london.gov.uk/dataset/london-schools-atlas`

Spatial Processing

Data II

Defining the inputs

- Crime data (January 2019) London `violence_london_0119.shp`
  - Street level data on reported crime on Violence and Sexual Abuse
  - Open Source: `https://data.police.uk/data/`
- Pollution level data. Raster `LAEI2016_2016_PM10_prj.tiff`
  - Satellite data on P(articulate) M(olecule) 10
  - Open Source: `https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory-2013`

Processing Spatial Data

Example: School Data

- Open the Graphical Processing Modeller ()
- Open the model `_school_prep.model3` that we have provided for you
- This model performs a bunch of tasks (we will go through it together)

    1. Select the schools with the higher Ofsted scores
    2. Spatially Intersect the schools points with the LSOA polgons
    3. Sum total number of schools for each LSOA
    4. Deletes duplicate LSOA
    5. Merge the point LSOA to the polygon LSOA
    6. Replace missing information with 0

- We are left with the polygon shapefile storing the total number of high quality schools in each LSOA

Processing Spatial Data

School Data

- Processing the crime data follows the same steps
  - Model `_crime_prep.model3`
  - Save the output as lsao_crime.shp in the "output" folder
- To process the Pollution raster use Model `_crime_prep.model3`
  - Save the output as lsao_crime.shp in the "output" folder
- Now we can run the final Model that will merge all these output in a single shapefile
  - Model `_merge_all.model3`
  - Save the output as merge_all.shp in the "output" folder

Data Visualization Tools

How to visualize Spatial Pattern in your data

- Maps

Advantages

- Help you detecting spatial patterns: cluster, spatial correlations
- Appreciated by policy-makers and practitioners

Data Visualization I

- Load your data on the Canvas
- Rename the layer with the appropriate name you want to be shown in the Map Legend
- From the Project menu select New Print Layout
- Type a Title for your Print Layout

Data Visualization II

- Inside the New Print Layout screenshot, use the Add Item to add:
  - Maps
  - Legend
- You have to drag the arrow to select the area where you would like each map element to appear