



F21DF: Database and Information Systems

Coursework 2: GIS

Coursework is marked out of 50

Weight: 15% of final mark

Deadline: 15:30 on 26 Nov 2020 (Edinburgh)

Individual assessment (not group)

Coursework Overview

This coursework continues on from the GIS lab - you should complete the GIS lab before attempting this coursework. To complete this coursework you should use PostgreSQL with either pgAdmin or the psql console, in conjunction with QGIS (desktop GIS).

Collaboration and Plagiarism

For this coursework reports and code must be your own **individual work** – **this is not a group project**. If some text or code in the coursework has been taken from other sources, these sources must be properly referenced. Failure to reference work that has been obtained from other sources or to copy the words and/or code of another student is plagiarism, and if detected will be reported to the School's Discipline Committee. If found guilty of plagiarism, the penalty could involve voiding the course.

Students must never give hard or soft copies of their coursework reports or code to other students on the course. Students must always refuse any request from another student for a copy of their report and/or code.

Submission

Ensure you include the SQL used to calculate the answers in each case.

Please submit to VISION as a PDF naming the file with your surname and initials (e.g *jones_ac.pdf*)

Tasks:

1) Highest Flickr photo density [3 marks]

Which cell has the highest Flickr photo density? Give the result as the **centroid point coordinate** in the British National Grid Coordinate system (i.e. the same as the points are stored in).

Include the SQL you used to determine the result.

2) Route lengths [3 marks]

Load the following datasets from www.macs.hw.ac.uk/~pb56/f21df into your PostgreSQL database

- routeb.sql
- routec.sql

(you should have already loaded routea.sql during the lab)

How long is each of the routes - routea, routeb, routec - in metres? (show the SQL you used)

3) Route popularity [3 marks]

Use the flickr photo data to show the popularity of each route according to how many photos have been taken along the route within a 25m buffer. You should normalise this by route length (i.e. photos per metre).

Give the SQL used and show your results in order from most popular route to least popular route. (Preferably do this as a single SQL statement.)

4) Temporal Patterns [3 marks]

Make a new table and add a point at Nelson Monument on Calton Hill which is British National Grid coordinates (326253.33,674110.63).

TIPS:

To set the coordinate system (spatial reference ID - SRID) use the EPSG code 27700 for British National Grid.

Check the online help for ST_SETSRID() – e.g. http://postgis.org/docs/ST_SetSRID.html

How many photos have been taken within 200m of this location for each day of the week?
(Show your SQL for all steps)

5) Data Aggregation [3 marks]

Use the supplied hexagon grid to calculate the most popular locations for photos where the tag includes each of the following terms – giving the centroid of the most popular cell in each case (i.e. 4 queries and outputs).

- Castle
- Calton Hill
- Royal Mile
- Meadows

(For each term show your SQL and the cell centroid with the highest count.)

6) Visualization [20 marks]

Make an interesting visualisation of the Flickr data - this should include a map using QGIS to show the spatial pattern of the term across Edinburgh and may include other visualisations or summaries you think are useful. For example you could focus on 1 of the terms from task 5. You could look to filter the data based on the dates, or userID, and include charts to show temporal patterns.

- Remember to check the lecture slides for tips about map design
- Treat the output as if it is to be displayed stand alone at A4 size (i.e. small poster)
- For this assignment **you must not use pre-rendered basemaps** (eg OpenStreetMap Mapnik)
- Export your design as an image to include in your report as a full A4 page – you may add additional features using Powerpoint / Word or other editor.

Additional base map layers are available and can be downloaded from:

<https://www.macs.hw.ac.uk/~pb56/f21df.html> as either a Geopackage or a ZIP file of layers.

(See notes on next page for how to use a Geopackage)

Further notes for Qu 6:

Use your spatial SQL skills to carry out some analysis of the Flickr data. It's fairly open as to what you want to show on the map/poster but to be an "interesting" visualisation of the data you need to be showing something more than just a map of all the original data given to you. Think about what you've covered in the lab with filtering, data aggregation, use of symbology, spatial joins, etc. Also reflect on the lecture material about map design and that a map should have a specific 'task' in mind that governs what to include/leave out. Here the 'map task' could be a question of the dataset, some trend or pattern you've found about the Flickr data, something which you find interesting about the data. For this coursework please avoid using pre-rendered basemaps (e.g. OpenStreetMap) - as you have no control over what is included or left out, and therefore can't design the map to suit your data story.

7) Write up [15 marks]

Give an overview of your poster (from task 6) and explain the decisions you made regarding its design, and what you wanted it to show/why. Highlight any issues and how you might be able to improve your analysis of the Flickr dataset.

- **Limited to 2 sides of A4 @ 10 point font and 2cm margins.**
- **Maximum of 500 words (excluding figures/tables/references).**
- **You can use bullet points, figures, tables, references, etc.**

Further Notes for Qu 7:

Describe what you did and why you did that, what you found interesting about that data. Also highlight things you had to do which were a bit different like the advanced SQL you wrote, the symbology classes you chose and why, the labelling decisions. Make sure you include what the overall ambition of your poster is showing (i.e. the question it's answering with the data).

It's quite usual to export maps (and the appropriate scale bar) to a desktop publishing package for the final layout. Tools such as Powerpoint & Photoshop will do the job to allow you to layout the A4 page from various image sources - just be careful that you do all spatial calculations in the GIS / spatial database environment, so that distances are calculated correctly (we don't want another Economist type error - see GIS lecture slides!).

So as an example of the workflow - you might create a map or two in QGIS with scale bars and export as images... also export some data from PostgreSQL using pgAdmin as a CSV ..load that in Excel or R and make a chart... then combine all these elements in a single A4 poster design in Powerpoint - export that as a single image to Word to include in your assignment with the answers to the other tasks - finally exporting all of that as a single PDF to upload to Vision.

In summary - be imaginative - explore the data a bit and think of the message you wish to communicate with your map poster based on the data provided and then explain why, how, what you did.

(see next page for notes on using Geopackages)

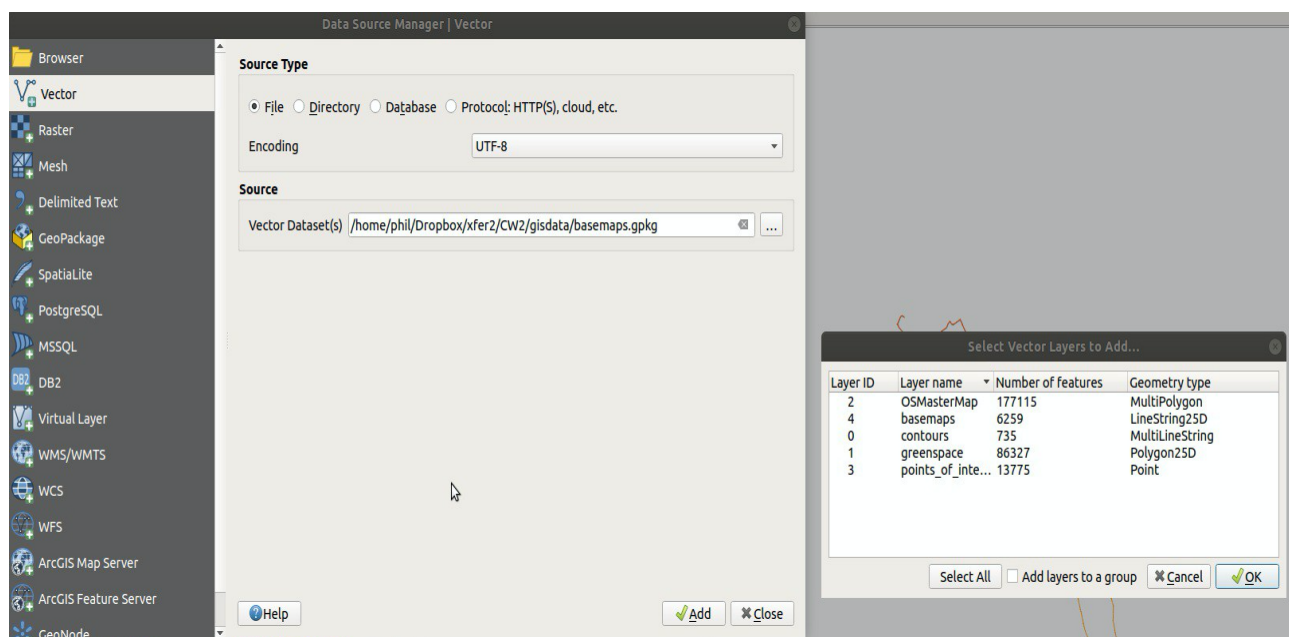
Notes on GEOPACKAGES

The Geopackage is based on a SQLite database and contains a number of Vector and Raster layers. You can load these into QGIS if you wish to augment your maps. The datasets are too large to load in the MACS PostgreSQL server so please don't do that but just link to your copy of the downloaded Geopackage.

To load from a Geopackage you need to use QGIS 3.x as explained below – if you are using QGIS 2.x then download a zip file of the basemap layers instead (then unzip onto your drive space).

To load VECTOR data from the Geopackage

Details on loading the VECTOR layers are as follows:

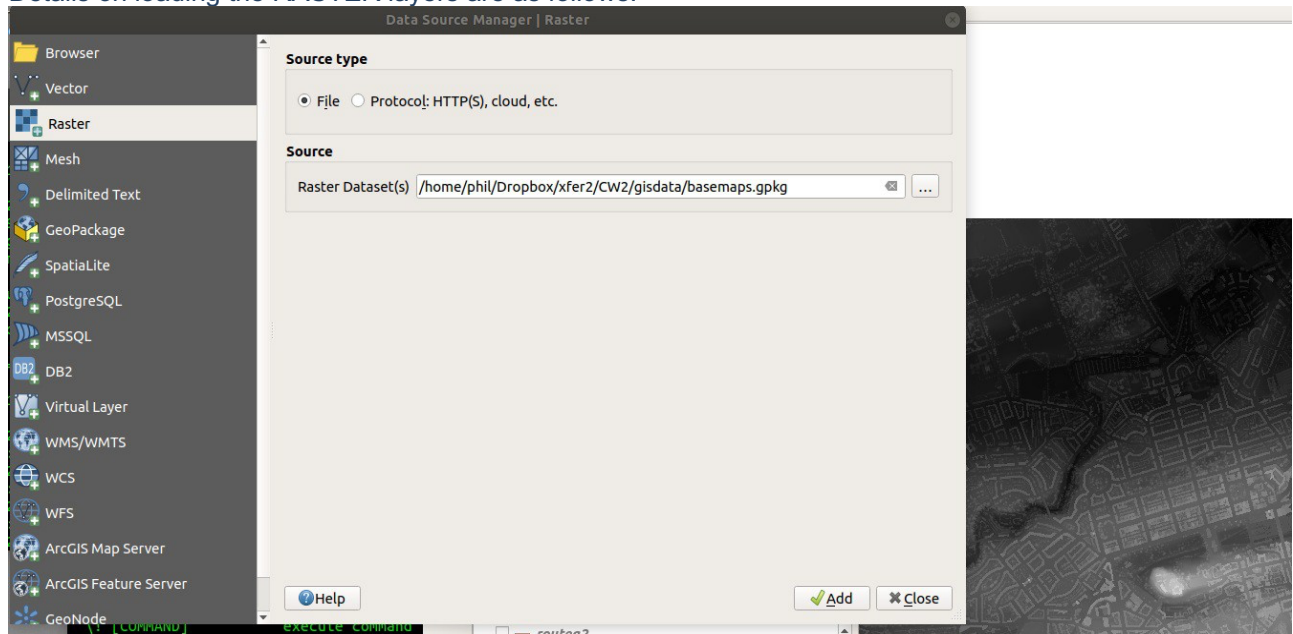


Layer menu > Add layer > Vector (or CTRL+SHIFT+V)

- Change the source to the Geopackage location you downloaded
- Click ADD
- Choose 1 or many of the available layers and click OK

To load RASTER data from the Geopackage

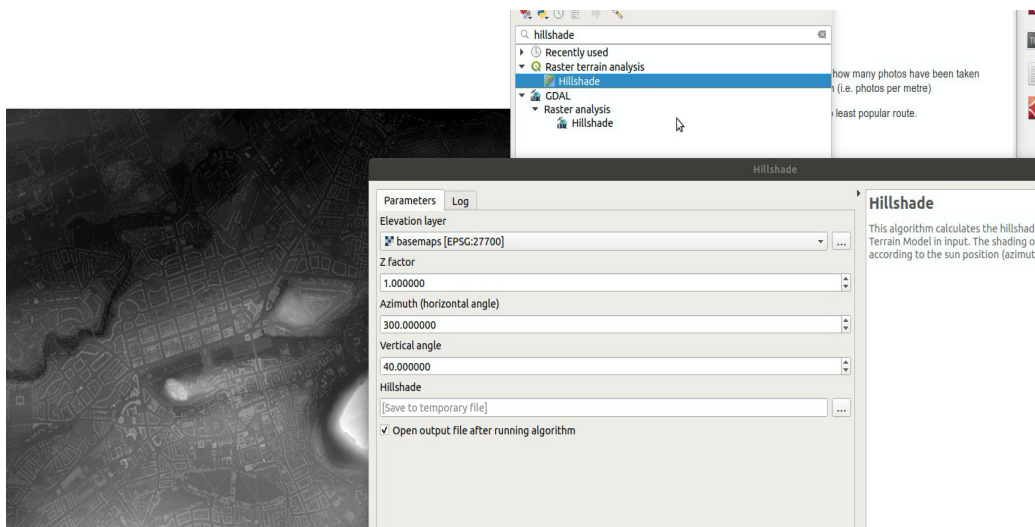
Details on loading the RASTER layers are as follows:



Layer menu > Add layer > Raster (or CTRL+SHIFT+R)
Change the source to the Geopackage location you downloaded

General Tips:

- Use the QGIS table of contents to change the layer order and symbology settings, as per the lab
- You can set each layer's transparency from the layer properties
- You can filter vector layers to show only a subset of items (see video links at the end of the lab sheet)
- You may wish to try out some QGIS spatial tools such as adding a Hillshade to the Digital Elevation Model.



Flickr Data source:

The New Data in Multimedia Research

Bart Thomee, David Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Dough Poland, Damian Borth, Li-Jia Li
Communications of the ACM, 59(2), pp. 64-73, 2016 (YFCC100m Dataset Paper)