

## EXERCISE 03: Geocoding and population

### Introduction

The location and density of the human population is fundamental to the way society is organized and it is important that GIS can be used to help understand population density and proximity to features, and to assign data to addresses. Postcodes provide a useful way to link human activity to physical locations.

You will find data for this exercise on the *L: Drive* under *College of Science/Geography/GIS/data*. This instruction sheet should be read in conjunction with the document *GIS\_guide\_to\_practical\_exercises* which gives general information about QGIS and how to find further help.

### Learning aim

In this exercise you will examine the distribution of postcode reference points within Wales, visualize population density, geocode new features onto a map and count populations within simple catchments.

### Instructions

#### Postcode point locations

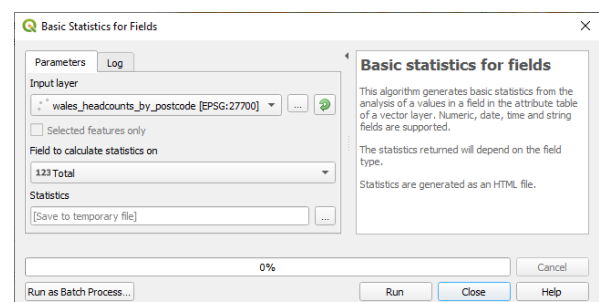
- 1) Launch QGIS, create a new project and open *wales\_principal\_areas*. You have seen this layer before.
- 2) Open the shapefile *wales\_headcounts\_by\_postcode*. This dataset quantifies the number of households and people registered at each postcode in Wales, and is derived from the UK Census of 2011 (<https://www.nomisweb.co.uk/census/2011>) combined with the coordinates of each postcode from Ordnance Survey OpenData (<https://www.ordnancesurvey.co.uk/opendatadownload>).
- 3) Investigate the attribute table of this dataset and explore the spatial distribution of postcode points. Examine areas where postcode points are both very dense and very sparse and consider why.
- 4) Using the attribute table selection column, and skills from Exercise 1, find the postcodes with the greatest number of people registered to them in the Cardiff area (CF), the Swansea area (SA), the Newport area (NP), and the Aberystwyth area (SY). Use *Google* to find out what is at each of these postcodes to justify so many apparent *residents*. Have you saved your Project recently?
- 5) Under *LayerProperties->Symbology*, change from *Single Symbol* to *Graduated*, make the controlling variable *Column* the **Total** (population registered in 2011), and choose the *Method* (of varying the symbology) to be **Size** and click on *Classify* then *Apply* (or *OK*) to implement these choices. The size of your points should now reflect the population at each postcode

#### Finding headcount statistics

- 6) It is useful to be able to derive statistical information from attributes of all features in a geospatial dataset or from *selections* of a dataset. From the *Vector* menu choose *Analysis Tools->Basic Statistics For Fields*. Apply this dialog to the **Total** attribute in *wales\_headcounts\_by\_postcode*. Consider what each of the output values represents. What was the **Total** (registered) population of Wales in 2011?
- 7) Using either the attribute table, or the Project window, *select* the *Swansea* polygon from the principle areas of Wales. Make sure that this polygon is highlighted in yellow in the Project Window.
- 8) Open the dialog: *Vector->ResearchTools->SelectByLocation*. Choose to '*Select features from*' *wales\_headcounts\_by\_postcode*, tick (only) the '*are within*' check-box, for '*By comparing to the features from*' choose *wales\_principal\_areas*, and tick the *Selected features only* check-box. Make sure that you understand what each of these choices means. *Run* this dialog to select (highlight in yellow) all postcode points within the City and County of Swansea.
- 9) Use the *Basic Statistics* tool to find the **Total** (registered) population of Swansea in 2011. Remember to tick the '*Selected features only*' check-box to calculate only the statistics from Swansea postcodes. Compare this total to that of Cardiff.

#### Population density

- 10) Open the *QGIS Toolbox (Processing->Toolbox)* and activate the *Heatmap* dialog from the *Interpolation* section of this toolbox. For the Point Layer choose *wales\_headcounts\_by\_postcode*, choose a *Radius* of



500, and a *Pixel size* (x and y) of 100, and under *Advanced parameters*->*Weight from field*, choose *Total*. Run this dialog to create a population density map of Wales. This new layer is a *raster* layer and the *Symbology* is very different from the vector layers you have opened so far.

- 11) Under *Layer*->*Layer properties*, experiment with the *Symbology* of your population density layer. There are many options and you are encouraged to experiment. You should at least be able to visualize the density as *Singleband grey* or *Singleband pseudocolor* using a variety of different *color ramps*. It may be easier to experiment if you switch off the visibility of the *wales\_headcounts\_by\_postcode* layer.
- 12) See how the map changes as you vary the *Radius*, or as you change the attribute which controls the weighting of each point. Be bold and experimental to learn the most from this step. Time to save?

### Geocoding new addresses

- 13) Being able to locate and plot a feature using its postcode is a powerful function of GIS. We will map Tesco stores. First, compile an Excel spreadsheet of Tesco stores in the Swansea area with one column for the postcode and one for the store name. Use column titles. Find the data at <http://www.tesco.com/store-locator/uk/> and include at least 7 stores in and around Swansea, and make sure there are no duplicates. Save your spreadsheet as CSV (comma separated variables).
- 14) Plugins are additional QGIS processing modules that are not activated as standard. Some need installing from the web (to your *OneDrive*), while some simply need activating. Find and install the plugin '**MMQGIS**' (*Plugins*->*Manage and Install Plugins*).
- 15) The process of **Geocoding** employs a postcode database (similar to that used to create *wales\_headcounts\_by\_postcode*) to attach spatial coordinates to postcode-referenced data.
- 16) From the (new) **MMQGIS** menu, choose *geocode*->*geocode CSV with Web Service*. Select your Tesco CSV file as the *Input file*, your postcode column as the *Address Field*, and *OpenStreetMap/Nominatim* as the *Web Service*. You will need to choose an output file on your *OneDrive* **both** for the resulting *Output Shapefile*, **and** for rows that fail to geocode (the *Not Found Output List*). You should now have a new map layer containing the locations of Tesco stores in Swansea. This is a tricky step and can fail for many reasons based on stray characters in your file. If it doesn't work, try using the CSV file I have created for you in '*wales\_tesco\_postcodes*' on the *L:Drive*. Label the stores with their names.
- 17) Although it appears on the map, this new layer's coordinate system (Latitude/Longitude) differs from the other layers in your project (OSGB: Ordnance Survey of Great Britain), which becomes critical to further processing steps. Using *Layer*->*Save as*, save this layer to your *OneDrive* using an appropriate name and the following settings: select the output **CRS** to be "**EPSG:27700-OSGB 1936**" and *Format* to be **ESRI Shapefile** (\*.shp). The resulting layer looks the same but is not (more on this in 'concepts').

### Simple catchments

- 18) It is sometimes useful to know how many people live within a given distance of a geographical feature. In this example, we will find out how many people live within walking distance of the Tesco stores you have geocoded. To create 'walking distance' catchments, choose *Vector*->*Geoprocessing Tools*->*Buffer*. Select your new OSGB Tesco shapefile as the *Input Vector Layer*, the *Buffer distance* to be 500m, and tick the box *Dissolve buffer results*.
- 19) Using previous experience (Steps 7 and 8), find the *Total* number of people registered to postcodes within your 500m Tesco catchments. These could be potential pedestrian customers for Tesco stores.
- 20) Experiment with different size catchments. You could also compare different supermarket chains.
- 21) If you are feeling ambitious, using the skills you have learned find the total number of people registered in 2011 to postcodes within 2km of Swansea University Singleton Campus (SA2 8PQ) and Bay Campus (SA1 8QE).

### Learning outcomes

By the end of this exercise you should:

- 1) Understand the nature of UK postcodes points and their distribution in Wales.
- 2) Be able to vary symbol size (or colour) for point datasets based on their attributes
- 3) Be competent at making selections from point data based on containment within polygon data.
- 4) Be able to calculate simple statistics from entire, and selections from, attribute tables.
- 5) Know how to create a heatmap and be able to choose the presentation style of a raster data layer.
- 6) Understand vector buffers and be able to find population within simple catchments.
- 7) Be familiar with installing QGIS Plugins.