Using open UK Data Service Census Support datasets in open source GIS software

April 27th 2017

James Crone
UK Data Service Census Support
University of Edinburgh
Welcome

- Look at the range of open geospatial data that is available for download through UK Data Service Census Support
- Other sources of open geospatial data that can be combined with UK Data Service Census Support data will also be looked at
- Introduction to open source GIS software
- Specific use of PostGIS spatial database
- Specific use of QGIS desktop GIS application
- Use of PostGIS / QGIS based around visualising census data using Choropleth maps, Cartograms and Flow maps.
UK Data Service Census Support

- Access to, and user support for, data from the last five UK population censuses (1971 – 2011)
- Majority of data provided by UKDS-CS is available as open data
- Various other non-census datasets which support the use of census data also available

https://census.ukdataservice.ac.uk/
What is open data?

- Open data is data that anyone can access, use and share
- Open data becomes useable when made available in a common, machine-readable format
- Open data must be licensed. Its licence must permit people to use the data in any way they want, including transforming, combining and sharing it with others, even commercially

Source: European Data Portal
(UK) Open Government Licence

- Copyright licence for Crown Copyright works published by the UK government
- Compatible with the Creative Commons Attributions (CC-BY) licence
- Permits anyone to copy, publish, distribute, transmit and adapt the licensed work, and to exploit it both commercially and non-commercially
- In return, the re-user of the licensed work has to acknowledge the source of the work and (if possible) provide a link to the OGL
Open UK Data Service Census Support Census data

Aggregate census area statistics data
  • tables

Census interaction data
  • flows between places

Census microdata
  • detailed tables

Supporting boundary data & geographical lookup tables
  • census places
Aggregate census area statistics data – download using InFuse / Casweb.

1971 – 2011 aggregate census area statistics data available as open OGL data.

InFuse gives access to 2001/2011 data.

Earlier data available in Casweb.
Census interaction data – download using WICID


Public datasets available as open OGL data.

Access to safeguarded datasets more restrictive.
Boundaries – download using Boundary Data Selector / Easy Download

Boundaries describing core 1971 - 2011 census geographies available as open OGL data.

Various other non-core boundaries also available.
Download mostly-contemporary boundaries (incl census) and postcode directories as OGL data. Less support for historical datasets.
Ordnance Survey Open Data

Free map datasets – many of which are open versions of non-free OS datasets. Includes background mapping data; Postcodes; boundaries; gazetteers etc.

<table>
<thead>
<tr>
<th>Product</th>
<th>Coverage</th>
<th>DVD</th>
<th>Download</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code-Point® Open</td>
<td>Great Britain [20 Mb]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data type: Point Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supply format: CSV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version: 02/2017</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Code-Point Open provides a precise geographic location for each postcode unit in Great Britain. The product is a CSV file containing postcodes, grid references, NHS® health and regional health authority codes, administrative ward, district, county and country area codes.

<table>
<thead>
<tr>
<th>Product</th>
<th>Coverage</th>
<th>DVD</th>
<th>Download</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategi®</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data type: Vector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supply format: ESRI® Shape</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version: 01/2016</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GIS – Geospatial data; Geospatial technology; Processes around its use and management etc.

Different views of the world.
Census data in GIS

Census Statistics

<table>
<thead>
<tr>
<th>GEOID</th>
<th>totpop</th>
<th>male</th>
<th>female</th>
</tr>
</thead>
<tbody>
<tr>
<td>E00169777</td>
<td>387</td>
<td>200</td>
<td>187</td>
</tr>
<tr>
<td>E00169778</td>
<td>391</td>
<td>196</td>
<td>195</td>
</tr>
<tr>
<td>E00169780</td>
<td>214</td>
<td>123</td>
<td>91</td>
</tr>
<tr>
<td>E00169782</td>
<td>126</td>
<td>71</td>
<td>55</td>
</tr>
<tr>
<td><strong>E00169783</strong></td>
<td><strong>242</strong></td>
<td><strong>137</strong></td>
<td><strong>105</strong></td>
</tr>
<tr>
<td>E00169785</td>
<td>1799</td>
<td>810</td>
<td>989</td>
</tr>
<tr>
<td>E00169786</td>
<td>197</td>
<td>109</td>
<td>88</td>
</tr>
<tr>
<td>E00169788</td>
<td>421</td>
<td>290</td>
<td>131</td>
</tr>
<tr>
<td>E00169789</td>
<td>214</td>
<td>132</td>
<td>82</td>
</tr>
</tbody>
</table>

Small area regional geography

**GEOGRAPHIC IDENTIFIERS**

Location brings context

Shapefiles and CSV`s

Digital boundaries overlain on OS basemap

Contains National Statistics data © Crown Copyright and database right 2012
Contains Ordnance Survey data © Crown Copyright and database right 2012
What is open source (GIS) software?

- Open source software is computer software made available with a license in which the copyright holder provides the rights to study, change and distribute the software to anyone and for any purpose.

- Open source software is often developed in a public, collaborative manner.

- Open source GIS software is open software used within the geospatial data application domain.
• The Open Source Geospatial Foundation
• A not-for-profit organisation whose mission is to support the collaborative development of open source geospatial software and promote its widespread use
• Projects should manage themselves, striving for consensus and encouraging participation from all contributors, from beginners through to advanced users
• OSGeo website a good place to start

www.osgeo.org
Different types of open source GIS software aimed at different types of user or application

- Geospatial Libraries
- Desktop Applications
- Web Mapping
- Content Management Systems
- Metadata Catalogs
Today you will see use of UK Data Service Census Support census data in two specific open source GIS software applications.

During today`s webinar we will use the QGIS desktop GIS application to create maps and other geographic visualisations of UK census data downloaded from UK Data Service Census Support.

As part of this the use of a PostGIS spatial database as a datastore will be demonstrated.
Getting started with PostGIS

- PostGIS adds GIS functionality / datatypes to PostgreSQL
- Store GIS data in PostgreSQL databases
- Do GIS data processing/analysis using spatial SQL
- Access our GIS data from anything that can talk to PostgreSQL
Obtaining / Installing PostGIS

A working installation of PostgreSQL is a prerequisite to installing PostGIS.

Installing PostgreSQL/PostGIS on windows through binaries is trouble-free and painless.

Install PostgreSQL then run Stack Builder to install additional features e.g. PostGIS on top of the base PostgreSQL installation.
Demo 1: Load a shapefile of UKDS-CS boundaries into PostGIS

- We have downloaded a shapefile boundary dataset (Scottish Council Areas) from UK Data Service Census Support
- We want to load the shapefile into a PostGIS spatial database that we have access to
- There are LOTS of different methods of accomplishing this. We`ll look at one method and having done so will look at how the data is then stored within our PostGIS spatial database.
Demo 2: Viewing PostGIS data in QGIS desktop GIS application

Download QGIS installer from the QGIS website for your platform of choice.

QGIS website has nicely written documentation describing how to use QGIS reflecting the very active QGIS community.
Demo 3: Joining CSV data to PostGIS data in QGIS

- A common task when using data downloaded from UKDS-CS is linking small area statistical data to boundaries. Will show how to do this in QGIS.
Visualising census data – non (geo)graphically.

Number of Rooms:

Cars and Vans Availability in Households:

Tenure:

<table>
<thead>
<tr>
<th>Year</th>
<th>75% Owned</th>
<th>65% Owned</th>
<th>11% Social Rented</th>
<th>15% Private Rented</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Visualising census data – geographically as a (choropleth) map.

Example: % of people working more than 49 hours per week in London as recorded in the UK 2011 census.
First – some mapping fundamentals
Mapping fundamentals I: Classifying Data

- Simplify numeric data into a set number of data classes (categories) in order to recognise patterns in that data.
- Various methods of classification provided by QGIS
- Choosing a different number of data classes and classification method will produce different results(?)
Map Fundamentals II: Styling Data

<table>
<thead>
<tr>
<th>Data Value</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3 to 20</td>
</tr>
<tr>
<td>4</td>
<td>3 to 20</td>
</tr>
<tr>
<td>5</td>
<td>3 to 20</td>
</tr>
<tr>
<td>18</td>
<td>3 to 20</td>
</tr>
<tr>
<td>20</td>
<td>3 to 20</td>
</tr>
<tr>
<td>22</td>
<td>21 to 38</td>
</tr>
<tr>
<td>30</td>
<td>21 to 38</td>
</tr>
<tr>
<td>32</td>
<td>21 to 38</td>
</tr>
<tr>
<td>39</td>
<td>39 to 56</td>
</tr>
<tr>
<td>53</td>
<td>39 to 56</td>
</tr>
<tr>
<td>57</td>
<td>57 to 74</td>
</tr>
<tr>
<td>60</td>
<td>57 to 74</td>
</tr>
<tr>
<td>61</td>
<td>57 to 74</td>
</tr>
<tr>
<td>62</td>
<td>57 to 74</td>
</tr>
<tr>
<td>64</td>
<td>57 to 74</td>
</tr>
<tr>
<td>64</td>
<td>57 to 74</td>
</tr>
<tr>
<td>67</td>
<td>57 to 74</td>
</tr>
<tr>
<td>77</td>
<td>75 +</td>
</tr>
<tr>
<td>81</td>
<td>75 +</td>
</tr>
<tr>
<td>91</td>
<td>75 +</td>
</tr>
<tr>
<td>93</td>
<td>75 +</td>
</tr>
</tbody>
</table>
**Demo 4:** Use QGIS to create a choropleth map from the boundaries we loaded into PostGIS and the CSV file we added to QGIS

Choropleth maps:

- Shade areas to display statistical variables
- Colour scale assigned to data and the value of each region is used to colour the region
Cartograms

- Statistic being thematically mapped is substituted for areal extent or distance
- Geometry of the map is distorted in order to convey the information of this alternate variable
- Distance or Area Cartograms

Distance Cartogram – relative distance within a network

Area Cartogram – alters physical area by another statistical variable
EU Referendum Cartograms

EU referendum: full results and analysis

Britain has voted by a substantial margin to leave the European Union. The picture that is emerging is of a heavily polarised country, with remain areas coming in more strongly for remain than expected, and leave areas more strongly for leave. Geographically, Scotland and London have voted overwhelmingly for remain, but outside the capital, every English region had a majority for leave.

Follow our referendum liveblog here

How did my area vote?

Enter placename or postcode

Submit

Latest declarations

Cornwall voted to leave 5 days ago
Rutland & Melton voted to leave 5 days ago
Northampton voted to leave 5 days ago
York voted to leave 5 days ago
Derby voted to leave 5 days ago
Stoke-on-Trent voted to leave 5 days ago
Guildford voted to remain 5 days ago
Chelmsford voted to leave 5 days ago
Northern Ireland voted to leave 5 days ago
Arun voted to leave 5 days ago
Census Cartograms

MORTGAGED

Having a mortgage is one way to own your home, even if you will have to pay it off over time. This means that you have part ownership of a property, and you are responsible for paying the mortgage. This is a combination of buying and renting, and it is becoming increasingly popular because it is cheaper to own a house on which you pay a mortgage than to rent it.

The largest and most dramatic falls were recorded in Reading (-10.4%), Southampton (-10.2%), West Ealing (-10.0%), and Harrow (-10.0%). More areas have experienced decreases in home ownership with the following declines being in Southwark (-10.0%), Wakefield (-10.0%), surrey (-10.0%), and Barnet (-10.0%).

By 2011, there were 10 areas where more than half the population lived in a mortgage property, and some of these are the most populated areas in the country. The most populated areas with the most owned properties are listed below in descending order: Slough (51.0%), Reading (50.9%), and Newbury (50.8%).

By 2011, more than half of the population lived in a mortgage property in the local authority of Slough (51.0%), Reading (50.9%), and Newbury (50.8%).

Newly built dwellings that are both owned outright and mortgaged by their occupiers have followed a similar trend to 2010. From just under two-thirds of all dwellings to less than 63% in 2013. As yet there is no sign of a downturn in this fall, and it is all due to the decline in the number of those with a mortgage.

Owner occupied dwellings, UK %

<table>
<thead>
<tr>
<th>Year</th>
<th>65</th>
<th>65.5</th>
<th>66</th>
<th>66.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mortgaged 2011 (%)
**Demo 5:** Use QGIS to create a cartogram from our census data and compare with the choropleth map
**Demo 6:** Use QGIS to create a flow map from UK Data Service Census Support flow data

- use 2011 census interaction (flow) data downloaded from UDKS-CS WICID application.