# Mapping Census data in QGIS



**UK Data Service** 



# Contents

# Contents

1. Introduction	3
1.1. Census data	3
1.2. Boundary data	3
1.3. Mapping software	3
2. Obtain your census data	4
Step 1: Geography	4
Step 2: Topics	5
Step 3: Categories	6
Step 4: Download	8
3. Formatting your data file	9
4. Downloading boundary data	11
5. Matching data in QGIS	12
6. Visualising data in QGIS	15
Useful links	167

# 1. Introduction

Mapping census data requires three ingredients:

 census data for a given set of geographic areas (such as local authorities, wards or super output areas)

boundary data for the same set of areas

 mapping software which brings these first two elements together so that the data can be mapped

This guide shows you how to create your own *choropleth* map which will show the percentage of males who work in the Manufacturing sector using the QGIS package. We will map the percentage of males who are working in Manufacturing in each local authority within the United Kingdom.

### 1.1. Census data

Traditional census outputs are aggregated counts, or other summary statistics, for a set of geographical areas. Because data are collected for the entire population these geographical areas can be very small. For example, it is possible to obtain a wide range of statistics for areas as small as *output areas* which are census areas containing approximately 300 persons.

#### 1.2. Boundary data

Boundary data are the digital information about the polygons that describe the shape and location of the geographical units so that they can be mapped. We will be using boundaries in the form of a shape file which is actually a format created by Esri; it is a folder containing a number of interrelated files. These boundaries can be obtained from the UK Data Service.

#### 1.3. Mapping software

QGIS is an open source mapping package which can be downloaded for free from <u>www.qgis.org</u> it has a good range of functionality and is straightforward to use. Similar commercial packages are available which could be used to undertake the same task, such as ARCGIS from <u>http://www.esri.com/</u> or mapinfo from <u>http://www.pitneybowes.com/uk</u>.

Other approaches which do not involve using conventional desktop include:

- mapping data with a statistical package such as R
- mapping data using an online service such as Google Maps

Guides to mapping census data using these two approaches are also available as part of this guide series.

# 2. Obtain your census data

You are interested in finding out how many Males and Females work in the **Manufacturing** industry for all Local and Unitary Authorities in the United Kingdom.

### Start

Go to: infuse.ukdataservice.ac.uk

• Click the 2011 Census data

• Not all census data are available at all geographical levels, you have a choice to pick Geography or Topics first

click Geography to search by geography first

#### Step 1: Geography

In this step we can pick the areas for which you want data by using the + and - buttons to expand or collapse nested areas.

Note the geographies that are displayed are dependent on the categories you have selected. Not all data is available at all geographical levels.

Geography types containing large numbers of areas (Districts within the UK, for example) can be chosen, but not expanded, to avoid displaying huge lists. Individual areas for these geography types can be displayed by expanding within higher level geography types (Counties, for example).

The Geography picker reflects the geographies for which data are released, so you may see that some units are available both separately and in combination.

We would like all Local and Unitary Authorities in the UK (which in InFuse are referred to collectively as 'Local Authorities'). So expand:

• Tick *Local Authorities* for England, Northern Ireland, Scotland, and Wales

Click Add

A summary of your selections will now be displayed, click **Next** 



Selected Geographic Areas									
Add	Remove	Remove All							
All Local Authoritie All Local Authoritie	is (324 areas) in Eng is (26 areas) in North is (32 areas) in Scoth is (22 areas) in Wale	hern Ireland land							
						Next			

#### Step 2: Topics

Census data comes in predefined topic combinations. These can be found in the grid.

Steps: 1 2 3 4				
Topics				
Show Guidance ome topics are only available for certain	n countries; E - England, N - Northern Irela	and, S - Scotland, W - Wales		
Filters Clear Mers	Showing 934 topic combinations Page 1 of 59			First < 11213 > Last
Topics	- Age Select	Ethnic group (E)[S][W]	• General health	Ethnic group (detailed) [E][N][M]     Select
Armed Forces (EI(M)     Armed Forces, household     reference and associated persons     IE(M)     Armal in the UK, age upon [EI[S]     M     Armal in the UK, year of [EI[S][M]     Armal in the UK, year of [EI[S][M]     Bedrooms, number of [EI[M]     Bedrooms, number of persons     per bedrooms     Re[M]	• Sex Select	Accommodation type Secce	Country of bith     Select	Household composition     Select
Car or van availability				

On the left hand side of the screen you can filter the list.

- Click Sex (notice that the amount of combinations have reduced to 18 topic combinations)
- Click Industry

We now just have a single topic combination.

Click Select for the Age | Economic activity | Industry | Sex combination

Filters Clear filter	1 topic combination(s) Page 1 of 1	found for topic(s) (I
	Age     Economic activity     Industry	
Accommodation type	• Sex	
Accommodation, adaptation of [N]		
Activity the week before the census [E][W]		
Adult lifestage (alternative adult definition)		.# Select
Adults in household, number of		
Age		
Armed Forces [E][W]		
Armed Forces, household reference and associated		

You will now see another screen giving you the definitions of the topics you have selected. Have a read through, in particular the *industry* definition.

Click Next

Age	
	the date of birth question and is a person's age at their last birthday, at 27 March 2011. Dates of birth that imply treated as invalid and the person's age is imputed. Infants less than one year old are classified as 0 years of
Economic activity	
census. Rather than not a person was an A person's economi whether employed, economic activity cla employment - wheth The census concept Organisation (ILO).	lates to whether or not a person who was aged 16 to 74 was working or looking for work in the week before a simple indicator of whether or not someone was currently in employment, it provides a measure of whether or active participant in the labour market. a citivity is derived from their 'activity last week'. This is an indicator of their status or availability for employment unemployed, or their status if not employed and not seeking employment. Additional information included in the soffication is also derived from information about the number of hours a person works and their type of er employed or self-employed. of economic activity is compatible with the standard for economic status defined by the International Labour t is one of a number of definitions used internationally to produce accurate and comparable statistics on loyment and economic status.
Industry	
	h a person aged 16 to 74 works relates to their main job, and is derived from information provided on the main over or business. This is used to assign responses to an industry code based on the <u>Standard Industrial</u>
Sex	
The classification of	a person as either male or female.
Unit	
The unit is for a part	icular count (e.g. people or households)
Previous	

# Step 3: Categories

In this step you pick the categories to make up your combination. Items with a single category come preselected.

Age: Pick **Age 16-74** Industry: Pick **Total: Industry** and **C Manufacturing** Sex: Pick **Male** and **Female** 



#### Click Add

A summary of your selections will now be displayed, click Next



By selecting all persons with an industry as well as those in manufacturing, we will be able to use the former count to calculate the percentage of males and females within each LA who were in work in the week preceding the census who were working in manufacturing.

#### Step 4: Download

You can now see a summary of your selections along with a file reference (which can be changed).

#### Click Get the data

Download	
Show Guidance	
Selected Category Combinations	
<ul> <li>317 - Age : Age 16 to 74 - Economic activity : In employment the week before the census - Industry : Total: Industry - Sex : Males - Unit : Person:</li> <li>318 - Age : Age 16 to 74 - Economic activity : In employment the week before the census - Industry : Total: Industry - Sex : Females - Unit : Person:</li> <li>329 - Age : Age 16 to 74 - Economic activity : In employment the week before the census - Industry : C Manufacturing - Sex : Males - Unit : Person:</li> <li>330 - Age : Age 16 to 74 - Economic activity : In employment the week before the census - Industry : C Manufacturing - Sex : Males - Unit : Person:</li> <li>330 - Age : Age 16 to 74 - Economic activity : In employment the week before the census - Industry : C Manufacturing - Sex : Females - Unit : Person:</li> </ul>	ons
Selected Areas	
<ul> <li>All Local Authorities (22 areas) in Wales</li> <li>All Local Authorities (26 areas) in Northern Ireland</li> <li>All Local Authorities (32 areas) in Scotland</li> <li>All Local Authorities (324 areas) in England</li> </ul>	
File reference AGE_ECOACT_INDUST_SEX_UNIT	
	Download Data
	Previous Get the Data

An orange **Download Data** button will appear, click it and you are given the choice to open/save the zipped file as normal. Save this file to your working folder

#### Results

The output comes in a Zip file format and will be named with a date stamp followed by the file reference (e.g.: 201842410383133\_AGE\_ECOACT\_INDUST\_SEX\_UNIT)

The zip file consists of three files:

- Citation File in rtf format, simply tells you how to cite the data
- Meta file: This file gives you further information about the data you have chosen. It includes full definitions for the components of your topic combination
- Data file in csv format. This contains the data you have chosen

Unzip the file, and ensure you save to your working folder. The exact method you need to unzip your data will depend on your unzipping package and your computer set up. For example, you may find that you can unzip by right clicking the file and selecting *Extract All*.

Name ^	Туре	Compressed size	Password p	Size
🕅 citations.rtf	Rich Text Format	1 KB	No	
Data_AGE_ECOACT_INDUST_SEX_UN	Microsoft Excel Comma Sep	9 KB	No	
Action Meta_AGE_ECOACT_INDUST_SEX_UN	Microsoft Excel Comma Sep	2 KB	No	

Open up the data file and look at the results. Note that the second row contains the full label of the data you downloaded.

# 3. Formatting your data file

You should now have a file called: Data\_AGE\_ECOACT\_INDUST\_SEX\_UNIT.csv

We need to modify this file slightly so QGIS will understand it. The second row shows the full category titles, we could abbreviate these. For example:

- In cell F1, delete F317 and insert Male\_all
- In cell G1, delete F321 and insert Female\_all
- In cell H1, delete F318 and insert Male\_manuf
- In cell I1, delete F322 and insert Female\_manuf

Now delete Row 2. Your file should now look like this:

	Α	В	С	D	E	F	G	Н	Ι	J
1	CDU_ID	GEO_COD	GEO_LABE	GEO_TYPE	GEO_TYP2	Male_all	Female_all	Male_manuf	Female_manuf	
2	52	95AA	Antrim	Local Auth	LA	13322	12273	1912	633	
3	53	95BB	Ards	Local Auth	LA	18220	17523	2265	557	
4	54	95CC	Armagh	Local Auth	LA	14128	12458	2183	731	
5	55	95DD	Ballymena	Local Auth	LA	15788	14132	3474	952	
6	56	95EE	Ballymone	Local Auth	LA	7171	6255	1320	325	
7	57	95FF	Banbridge	Local Auth	LA	11945	10676	1740	558	
8	58	95GG	Belfast	Local Auth	LA	59097	59657	5455	1451	
9	59	95HH	Carrickfer	Local Auth	LA	9189	8915	1463	347	
10	60	9511	Castleread	Local Auth	IΔ	16055	16061	1858	446	

Visualising the data by the total count of people working in an industry may not be entirely useful on its own because it does not take into account the population of each Local Authority. When we come to map the data we will need to use a proportion. So let's create two extra columns and name them:

- 1. Male\_man\_p
- Female\_man\_p

And create a percentage using the other columns:

• In J2 type the function =(H2/F2)\*100

• And copy this function down the column to calculate the percentage of male workers in each local authority who are in manufacturing

• In K2 type the function =  $(I2/G2)^*100$  and copy this function down the column to calculate the equivalent percentage for females

Your data should now look like this:

	А	В	С	D	E	F	G	н	Ι	J	к	L	м
1	CDU_ID	GEO_COD	GEO_LABE	GEO_TYP	E GEO_TY	P2 Male_all	Female_all	Male_manuf	Female_manuf	Male_man_P	Female_man_P		
2	52	95AA	Antrim	Local Aut	th LA	13322	12273	1912	633	14.35219937	5.157663163		
3	53	95BB	Ards	Local Aut	tł LA	18220	17523	2265	557	12.43139407	3.17867945		
4	54	95CC	Armagh	Local Aut	th LA	14128	12458	2183	731	15.4515855	5.867715524		
5	55	95DD	Ballymena	Local Aut	th LA	15788	14132	3474	952	22.00405371	6.736484574		
6	56	95EE	Ballymone	Local Aut	th LA	7171	6255	1320	325	18.40747455	5.195843325		
7	57	95FF	Banbridge	Local Aut	th LA	11945	10676	1740	558	14.56676434	5.226676658		
8	58	95GG	Belfast	Local Aut	t <sup>i</sup> LA	59097	59657	5455	1451	9.230587001	2.432237625		
9	59	95HH	Carrickfer	Local Aut	th LA	9189	8915	1463	347	15.92121014	3.892316321		
10	60	9511	Castlerea	Local Aut	th LA	16055	16061	1858	446	11.57271878	2.776913019		
11	61	95JJ	Coleraine	Local Aut	t <sup>i</sup> LA	13122	12150	1684	488	12.83340954	4.016460905		
12	62	95KK	Cookstow	Local Aut	th LA	8660	7336	1959	616	22.62124711	8.396946565		
13	63	95LL	Craigavon	Local Aut	th LA	21615	19705	4926	1813	22.78972935	9.20071048		
14	64	95MM	Derry	Local Aut	t <sup>i</sup> LA	20649	19963	2870	697	13.89897816	3.4914592		
15	65	95NN	Down	Local Aut	th LA	15918	14992	1398	457	8.782510366	3.048292423		
16	66	9500	Dunganno	Local Aut	th LA	13955	11691	3608	1396	25.85453243	11.94080917		
17	67	95PP	Fermanag	Local Aut	t <sup>i</sup> LA	14749	12825	2114	752	14.33317513	5.863547758		
18	68	95QQ	Larne	Local Aut	th LA	7695	7166	1370	450	17.80376868	6.279653921		
19	69	95RR	Limavady	Local Aut	th LA	7010	6207	887	232	12.65335235	3.737715483		
20	70	95SS	Lisburn	Local Aut	t <sup>i</sup> LA	28211	26751	3607	1130	12.78579278	4.224141154		
21	71	95TT	Magheraf	Local Aut	th LA	11010	8973	1967	543	17.86557675	6.051487797		
22	72	95UU	Moyle	Local Aut	th LA	3793	3228	386	105	10.17664118	3.252788104		
23	73	95VV	Newry and	Local Aut	th LA	22034	19429	3210	1323	14.5683943	6.809408616		
24	74	95WW	Newtown	Local Aut	th LA	20151	19818	2872	692	14.25239442	3.491775154		
25	75	95XX	North Dov	Local Aut	tł LA	18887	18029	1963	489	10.39339228	2.712296855		
26	76	95YY	Omagh	Local Aut	th LA	11603	10073	1662	377	14.32388175	3.742678447		
27	77	95ZZ	Strabane	Local Aut	th LA	8153	6847	1045	314	12.81736784	4.585950051		
28	78	E06000001	Hartlepoo	Local Aut	tł LA	19511	18256	3656	754	18.73814771	4.130148992		

Save the file to your folder (in .csv format).

# 4. Downloading boundary data

- 1. Go to: http://census.ukdataservice.ac.uk/ and go to:
- 2. Get census data > Boundary data > Easy download
- 3. Click the *InFuse* United Kingdom tab
- 4. Click InFuse Local Authorities, 2011

5. Under the InFuse 2011 Local Authorities, Clipped click Download features in Shapefile format as ZIP file

6. This will start to download the boundary dataset. Save the Zip file to your computer in the folder you created earlier – it will be named by default as Infuse\_dist\_lyr\_2011\_clipped

7. Navigate to the zip file and right click and select *Extract all* to unzip the files, ensure that you know the location of these files

QUICK ACCESS TO	
Easy Download	
Boundary Data Selector	You have selected the following dataset to download: InFuse Local Authorities, 2011
Thematic Mapper	You can download an individual file using your web browser by clicking on its name in the table. If you use Firefox/Chrome, the best
Postcode Directory	way to do this is to click with the right mouse button and go to Save Link As. With Internet Explorer you should go to Save Target
Postcode Data Selector	As after clicking with the right mouse button.
	Warning: Some of these boundary files are very large. Please make sure you have plenty of space on your PC or server to enable you to download and uncompress these files
	Help on Using Easy Download   Data Formats   What is Generalisation?   Zip and Tar Gzip Files
	InFuse Local Authorities, 2011
	Download attributes in CSV format as ZIP file
	Download features in KML format as ZIP file
	Download features in MapInfo TAB format as ZIP file
	Download features in Shapefile format as ZIP file
	InFuse Local Authorities, 2011, Clipped
	Download attributes in CSV format as ZIP file
	Download features in KML format as ZIP file
	Download features in MapInfo TAB format as ZIP file
•	Download features in Shapefile format as ZIP file

# 5. Matching data in QGIS

We can now map the census aggregate data we downloaded from InFuse and combine it with the census boundary data.

- 1. Open QGIS Desktop (the following instructions are based on version 2.8.9)
- 2. Click the Add Vector Layer icon



3. Browse to the boundary .shp file data that you downloaded earlier (i.e. infuse\_dist\_lyr\_2011\_clipped.shp) and click open

4. You will now see an outline of the UK



5. We have added the boundaries, now we need to add the aggregate data that we downloaded from InFuse. Click the ADD Delimited Text Layer icon in the left hand menu of the screen

6. Browse for the Data\_AGE\_ECOACT\_INDUST\_SEX\_UNIT.csv file and select csv for the file format and *no geometry (attribute table only)*, then click *OK* 

🧭 C	ireate a Lay	er from a De	limited Text	File						?	x	
File	File Name seman/Downloads/qgis-infuse/201842410383133_AGE_ECOACT_INDUST_SEX_UNIT/Data_AGE_ECOACT_INDUST_SEX_UNIT.csv Browse											
Layer name Data_AGE_ECOACT_INDUST_SEX_UNIT									Encoding U	TF-8	-	
File	format	• CSV	(comma separa	ated values)	O Cus	tom delimite	rs	0	Regular expression	n delimiter		
Re	ord options	Number	of header lines	to discard	🔶 🗙 First	record has	field names					
	d options			card empty fields		separator is						
				cara empty neias		1			No	had a set of the balance		
Ge	ometry defin	ition 🔘 Poin	t coordinates		O wei	known tex	t (WKT)		No geometry (attri	bute only table)		
		_										
Lay	er settings	Use	spatial index		Use subset index			Watch file				
	CDU_ID	GEO_CODE	GEO_LABEL	GEO_TYPE	GEO_TYP2	Male_all	Female_all	Male_manuf	Female_manuf	Male_man_P		
1	52	95AA	Antrim	Local Authorities	LA	13322	12273	1912	633	14.35219937	5	
2	53	95BB	Ards	Local Authorities	LA	18220	17523	2265	557	12.43139407	3	
3	54	95CC	Armagh	Local Authorities	LA	14128	12458	2183	731	15.4515855	5	
4	55	95DD	Ballymena	Local Authorities	LA	15788	14132	3474	952	22.00405371	6	
5	56	95EE	Ballymoney	Local Authorities	LA	7171	6255	1320	325	18.40747455	<u>ب</u>	
1		0555		a facilizar		110.15	10070	1740		•		
								ОК	Cancel	Help		
										, nep		

- 7. We now need to join these two layers together
- 8. Double-click the infuse\_dist\_lyr\_2011\_clipped layer in the layer box on the right hand side
- 9. Click joins, and tick the green + symbol



10. Select GEO\_CODE as both the join field and target field click OK then Apply and then OK again

11. The **Open Attribute Table** icon is on the top menu bar allows you to see all of the data in table form. The geo\_code fields in both tables are identical; this is how QGIS knows which Local Authority boundary belongs to which Local Authority data

# 6. Visualising data in QGIS

In this task we will create a choropleth map in QGIS.

1. Double-click on the *infuse\_dist\_lyr\_2011\_clipped* layer in the right hand layer panel. This will bring up the layer properties box

🧕 Layer Properties - infus	e_dist_lyr_2011_clipped   Joins			( ۱۹
🔀 General	Join layer	Join field	Target field	
Style	Lata AGE ECOACT INDUST SEX UNIT	GEO_CODE	geo_code	V
abc Labels				
Fields				
🞸 Rendering				
🧭 Display	:			
Actions				
• Joins				
Diagrams				
🥡 Metadata				
	•			
	Style •			OK Cancel Apply Help

2. Click **Style**, you have a few choices. In the top dropdown menu, click **Graduated**. The colour ramp picker allows you to choose the colours to be used to represent the percentages; you can change this if you wish

3. We can now select what data we want to be displayed on the map. Earlier we created two percentages columns, select one of these from the column dropdown. In this example I have selected *Male\_man\_p* 

4. Click on *Classify* and you can see that 5 bands appear. You can change the number of these by changing the number in the *classes*' box. There are a few options for *mode*; we would recommend using "natural breaks (Jenks)". Change these options as you wish, and then click Apply and OK

🔏 Layer Properties - infuse_dist_lyr_2011_clipped   Style									
General	🔁 Graduated	-							
Style	Column	Data_AGE_ECOA	ACT_INDUST_S	EX_UNIT_Male_ma	n_P 🔻 🗵				
abc Labels	Symbol			Change		Classes	5	-	3
Fields	Color ramp	[sour	ce]	▼ Invert		Mode	Natural Breaks	(Jenks)	•
	Legend Format	%1 - %2				Precision	1	Trim	
👋 Rendering	Symbol 🗸	Values	Legend						<b>.</b>
🤎 Display	×	2.012 - 6.941 6.941 - 11.322	2.0 - 6.9 6.9 - 11.3						
Actions	×	11.322 - 15.569 15.569 - 21.041	11.3 - 15.6 15.6 - 21.0						•
• ┥ Joins	Classify	Add class	Delete	Delete all	Link class boundarie	es		Advanced	-
Diagrams	▼ Layer rende	ering							
🥡 Metadata	Layer transpa	rency	0					0	3
	Layer blendin	g mode	Normal	•	Feature blending mode	No	rmal	-	
	Style •				ОКС	ancel	Apply	Help	,

5. You will now have a map, looking something like this:

6. There are a lot of things you can do but some basic things would be to create a map to be printed, change the size, add a label and add a legend

7. To format your map for printing go to *Project > New Print Composer*, and insert a title. You can change the size of the page by clicking the composition tab on the right hand side of the screen

8. If you want to, change the page size using the panel on the right hand size (composition tab). It makes sense to leave the page size as A4, but to change the orientation to portrait

9. To insert the map, click Layout > Add map, and drag your mouse across the page

10. Adjust the size. You can do this by clicking the arrows around the map, but there is a lot of extra white space. Go to *Layout > Move item*, and *Layout > Move content* to adjust the size and scale of the map. While using *Layout > Move content* you can use your mouse's scroll wheel to adjust the size of the map. Or use the options within the *item properties* tab on the right if the screen (for example you can specify scale, the example map is set to 35000)

11. Insert a label. Go to *Layout > Add Label*, and click and move your cursor on the map. Type in the right hand side box. You can also adjust the size and font. Add a title and the citation information which you can find at: <u>census.ukdataservice.ac.uk/use-data/citing-data</u>

12.Insert a Legend. Go to *Layout > Add Legend*. You can modify some aspect of the legend using the option in the right hand panels

This map can be saved as a project to allow further editing. It can also be exported as a .svg, .png or .pdf file for publication.





# **Useful links**

- UK Data Service Census Support: <u>census.ukdataservice.ac.uk</u>
- Get data page: <u>census.ukdataservice.ac.uk/get-data</u>
- Use data page: <u>census.ukdataservice.ac.uk/use-data</u>
- InFuse: infuse.ukdataservice.ac.uk
- Casweb: <u>casweb.ukdataservice.ac.uk</u>
- Boundary data: <u>census.ukdataservice.ac.uk/get-data/boundary-data</u>
- Census forms:

census.ukdataservice.ac.uk/use-data/censuses/forms

- Census definitions:
- census.ukdataservice.ac.uk/use-data/censuses/definitions
- Datashine: datashine.org.uk
- <u>QGIS: www.qgis.org/</u>

T +44 (0) 1206 872143 E help@ukdataservice.ac.uk W ukdataservice.ac.uk

The UK Data Service delivers quality social and economic data resources for researchers, teachers and policymakers.

© Copyright 2017 University of Essex and University of Manchester



**UK Data Service** 



