

# InFuse: Easy access to aggregate data from the 2011 census

Justin Hayes  
Data Coordinator  
UK Data Service Census Support

Scotland's Census  
Edinburgh  
25 June 2014



# InFuse

The screenshot shows the InFuse website interface. At the top, it says 'UK Data Service Census Support' and 'InFuse Home | Guide | Tutorial'. The main heading is 'InFuse' with a colorful logo. Below this, it states: 'InFuse is a free service providing easy access to aggregate data from the UK 2011 and 2001 censuses'. There are sections for 'Access:' with links for '2011 Census data' and '2001 Census data', and 'Data coverage' which explains the data is available down to Local Authority level for England and Wales, and Output Area level for Scotland. It also includes an 'About the data' section.

# InFuse

- Public/unrestricted web interface
- Aggregate data from 2011 census across the UK
  - England and Wales to output area level
  - Scotland to council area level (UK harmonised), but more soon!
  - Also 2001 census data for England and Wales
- Makes data easy to
  - Find
  - Understand
  - Use
- Global query using variable combinations
  - No tables!
  - No data fast!



# InFuse

- Public/unrestricted web interface
- Aggregate data from 2011 census across the UK
  - England and Wales to output area level
  - Scotland to council area level (UK harmonised), but more soon!
  - Also 2001 census data for England and Wales
- Makes data easy to
  - Find
  - Understand
  - Use
- Global query using variable combinations
  - No tables!
  - 'No data' fast!



# Variable combination selection

The screenshot shows the 'Filters' section of the InFuse interface. It displays a grid of 27 topic combinations. On the left, there is a list of topics with checkboxes, including 'Accommodation type', 'Adults in household', 'Age', 'Aboriginal population', 'Annual net ex. age span', 'Business, number of', 'Business, occupancy rating', 'Car or van availability', 'Care leavers in household, type of', 'Children, age of', 'Contractual employment management and type', and 'Contractual establishment'. The main area shows a grid of 27 topic combinations, each with a 'Select' button. The interface is titled 'Showing 27 topic combinations' and 'Page 1 of 1'.

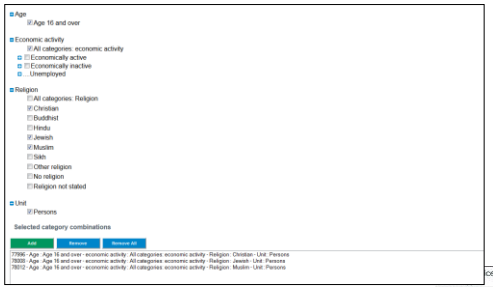


# Variable combination selection

This screenshot shows the same 'Filters' interface as the previous one, but with a filtered view. The title is '2 topic combination(s) found for topic(s) (Economic activity, Religion)'. The grid now only shows two topic combinations: 'Age' and 'Economic activity - Religion'. The interface is titled 'Page 1 of 1'.



## Category combination selection



## Area selection



## Data download

CDU_ID	GEO_CODE_GEO_LAB	GEO_TYPE	AGE	Economic activity	Religion	Persons
4	E20000001	England	210915	556815	177865	34489
5	E12000002	North West Region	157292	15587	11328	28170
6	E12000003	Yorkshire Region	420761	12817	11919	22276
7	E12000004	East Midlands Region	368767	11392	74485	55083
8	E12000005	West Midlands Region	401880	14619	58892	26842
9	E12000006	East of England Region	478833	29648	43127	105261
10	E12000007	London Region	684927	71983	377030	691023
11	E12000008	South East Region	690266	18146	22276	11919
12	E12000009	South West Region	401880	14619	58892	26842



## Under the bonnet

- Integrated multidimensional descriptive model
- Integrated model of geographies
- **The really important bits!**



## InFuse 2011 release 2: Raw data

- England and Wales Local and Detailed Characteristics to output area level
- UK harmonised data to local authority level
- 422 tables, mainly multivariate
- 31 geography types
- 241,334 areas
- 11,311 files
- 15Gb volume



## Integrated descriptive model

- Processing of raw metadata
  - Deconstruction, rationalisation and re-integration
  - Library of variables and categories
  - Re-insertion of data values
  - Attachment of associated metadata
- Global description using standards
- Global operations via Web service API
  - Data is self-describing
  - Enables lightweight, generic applications



## InFuse 2011 release 2: Processed data

- 97 variables
- 2,501 categories
- 281 variable combinations
- 140 thousand category combinations
- 4.6 billion values
  - A 460Km high stack of sticky notes!
  - Anticipating approximately 10 billion values in all

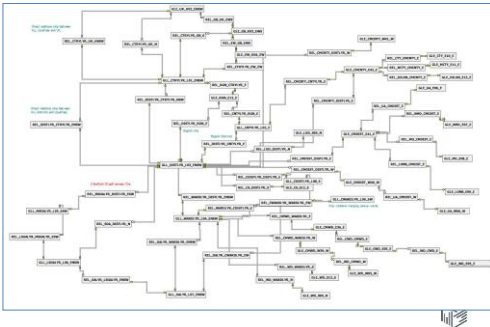


## Integrated model of UK census geographies

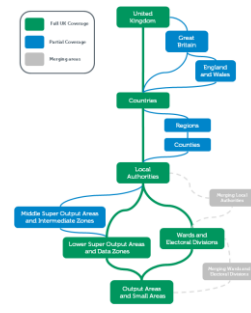
- Assembly of raw information on geographies
  - 31 geography types
  - 241,334 areas (anticipating ~ 2 million including postcodes)
  - Direct and indirect hierarchies
- Simplified presentational model
  - 11 composite geography layers
  - Simplification of merged geographies in England and Wales
- Calculation of 'missing' data
- Linkage between descriptive and geography models
  - Partial availability of data for geographies and extents



## Raw admin and statistical geographies



## Admin and statistical geography layers



## What's next?

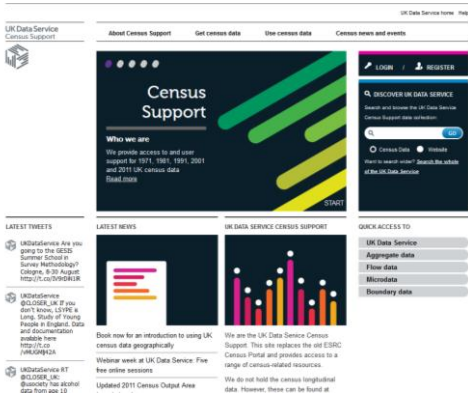
- More data
  - More comparable data
- Different data
  - Boundary and flow data
- More functionality
  - Analysis and visualisation
- Expose InFuse API
- Work with statistical agencies
  - Machine-friendly data from source
  - Flexible generation with automated disclosure control
  - Information on usage and contact with users



## UK Data Service Census Support

- Specialist function of UK Data Service
- Access and support services for outputs from recent UK censuses
- Add value by making census outputs easy to find, understand and use
- Engagement with UK census agencies
- Long history of technological innovation in service development
- [census.ukdataservice.ac.uk](http://census.ukdataservice.ac.uk)





## Census Support at Manchester

- Aggregate component of census outputs



Justin Hayes



Richard Wiseman



Rob Dymond-Green



## Census Support at Manchester

- Aggregate component of census outputs



Justin Hayes



Richard Wiseman



Rob Dymond-Green



## Give InFuse a go!

- <http://infuse.mimas.ac.uk>

- Comments, questions and ideas welcome
- [help@ukdataservice.ac.uk](mailto:help@ukdataservice.ac.uk)

