New and easier ways of working with aggregate data and geographies from UK censuses

The webinar will begin at 3pm

- You now have a menu in the top right corner of your screen.
- The red button with a white arrow allows you to expand and contract the webinar menu, in which you can write questions/comments.
- We won’t have time to answer questions while we are presenting, but will answer them at the end
- You will be on mute throughout – we can’t hear you.
Welcome

Some introductions…

Richard Wiseman
UK Data Service
The University of Manchester

Justin Hayes
UK Data Service
The University of Manchester
Can you hear us?
If you can’t hear us

• Check your speakers/headset
• Check your volume
• Use the phone number in your invitation to call in
• The webinar is being recorded - we’ll send a link
Webinar structure

- 30 minute presentation
- 30 minutes for questions and answers
- Finish by 4pm
- Please use chat facility to ask questions or let us know of any problems
- Some questions for the audience during presentation
Presentation content

• The UK Data Service
  • UK Data Service Census Support

• Census background
  • Aggregate data
  • Geographies
  • Dissemination challenges

• InFuse
  • Data model
  • Geography model
  • Live demonstration!
The UK Data Service

- Funded by the ESRC, integrating several previous resources
- A single, comprehensive and integrated point of access to a wide range of social science data
- Support, training and guidance
- ukdataservice.ac.uk
UK Data Service Census Support

- A specialist unit of the UK Data Service
- Access to, and support for use of data from the last five UK censuses (1971 – 2011)
- Bespoke interfaces to make data easy to find, understand and use
- census.ukdataservice.ac.uk
Welcome to the UK Data Service

Your resource for quality social research data

A unified point of access to data from ESDS, Census Programme, Secure Data Service and others

LATEST TWEETS

UKDataService
@PHE_obesity Soon – within the next 30 days. Subscribe to our email list to get weekly updates on new data releases! http://t.co/0MrqjGvxdN

LATEST NEWS

OUR DATA COMMUNITY

QUICK ACCESS TO

Key data

Census Support

Information for new users

Frequently asked questions

ukdataservice.ac.uk
Census Support

Who we are

We provide access to and user support for 1971, 1981, 1991, 2001 and 2011 UK census data
Read more

UK Data Service Census Support

Book now for an introduction to using UK census data geographically

Webinar week at UK Data Service: Five free online sessions

Updated 2011 Census Output Area

We are the UK Data Service Census Support. This site replaces the old ESRC Census Portal and provides access to a range of census-related resources.

We do not hold the census longitudinal data. However, these can be found at...
Have you used census data?
UK Censuses

• Decennial questionnaire surveys
  • Entire UK population every ten years* since 1801
  • Questions about people and households
  • 2011 Census cost ~ £500m

• Primary evidence for government policy and spending
  • Wide range of high quality demographic and socio-economic characteristics
  • Detailed combinations of characteristics - *What?*
  • Small areas - *Where?*
  • Long history – *Change?*

• Rich secondary source of information
  • Open Government License!
Aggregate data

- Counts of people and households with particular combinations of characteristics for particular geographical areas
  - Characteristics derived from questionnaire responses
  - Very large to very small areas (UK to postcodes)
  - Females aged 16-74 in employment in associate professional and technical occupations and usually resident in wards in the County of Devon?
## Aggregate data

<table>
<thead>
<tr>
<th>CDU_ID</th>
<th>GEO_CODE</th>
<th>GEO_LABEL</th>
<th>F403</th>
<th>F412</th>
</tr>
</thead>
<tbody>
<tr>
<td>1243845</td>
<td>E05003459</td>
<td>Axminster Rural</td>
<td>548</td>
<td>43</td>
</tr>
<tr>
<td>1243846</td>
<td>E05003460</td>
<td>Axminster Town</td>
<td>985</td>
<td>65</td>
</tr>
<tr>
<td>1243847</td>
<td>E05003461</td>
<td>Beer and Branscombe</td>
<td>403</td>
<td>38</td>
</tr>
<tr>
<td>1243848</td>
<td>E05003462</td>
<td>Broadclyst</td>
<td>1223</td>
<td>152</td>
</tr>
</tbody>
</table>

**Age : Age 16 to 74 - Economic activity : in employment the week before the census - Occupation : 3. Associate professional and technical occupations - Sex : Female - Unit : Persons**

<table>
<thead>
<tr>
<th>CDU_ID</th>
<th>GEO_CODE</th>
<th>GEO_LABEL</th>
<th>F403</th>
<th>F412</th>
</tr>
</thead>
<tbody>
<tr>
<td>1243855</td>
<td>E05003469</td>
<td>Exmouth Halsdon</td>
<td>1547</td>
<td>153</td>
</tr>
<tr>
<td>1243856</td>
<td>E05003470</td>
<td>Exmouth Littleham</td>
<td>1262</td>
<td>138</td>
</tr>
<tr>
<td>1243857</td>
<td>E05003471</td>
<td>Exmouth Town</td>
<td>1823</td>
<td>169</td>
</tr>
</tbody>
</table>

**Age : Age 16 to 74 - Economic activity : in employment the week before the census - Occupation : All categories - Sex : Female - Unit : Persons**

<table>
<thead>
<tr>
<th>CDU_ID</th>
<th>GEO_CODE</th>
<th>GEO_LABEL</th>
<th>F403</th>
<th>F412</th>
</tr>
</thead>
<tbody>
<tr>
<td>1243863</td>
<td>E05003480</td>
<td>Ottery St Mary Town</td>
<td>1115</td>
<td>122</td>
</tr>
<tr>
<td>1243864</td>
<td>E05003481</td>
<td>Raleigh</td>
<td>475</td>
<td>37</td>
</tr>
<tr>
<td>1243865</td>
<td>E05003482</td>
<td>Seaton</td>
<td>1290</td>
<td>93</td>
</tr>
<tr>
<td>1243866</td>
<td>E05003483</td>
<td>Sidmouth Rural</td>
<td>396</td>
<td>39</td>
</tr>
<tr>
<td>1243867</td>
<td>E05003484</td>
<td>Sidmouth Sidford</td>
<td>1272</td>
<td>100</td>
</tr>
<tr>
<td>1243868</td>
<td>E05003485</td>
<td>Sidmouth Town</td>
<td>918</td>
<td>72</td>
</tr>
<tr>
<td>1243869</td>
<td>E05003486</td>
<td>Tale Vale</td>
<td>603</td>
<td>59</td>
</tr>
<tr>
<td>1243870</td>
<td>E05003487</td>
<td>Trinity</td>
<td>516</td>
<td>63</td>
</tr>
<tr>
<td>1243871</td>
<td>E05003489</td>
<td>Woodbury and Lympstone</td>
<td>1011</td>
<td>108</td>
</tr>
<tr>
<td>1243872</td>
<td>E05003490</td>
<td>Yarty</td>
<td>504</td>
<td>51</td>
</tr>
<tr>
<td>1243873</td>
<td>E05003491</td>
<td>Alphington</td>
<td>2200</td>
<td>273</td>
</tr>
</tbody>
</table>
## Aggregate data

<table>
<thead>
<tr>
<th>CDU_ID</th>
<th>GEO_CODE</th>
<th>GEO_LABEL</th>
<th>F403</th>
<th>F412</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1243845</td>
<td>E05003459</td>
<td>Axminster Rural</td>
<td>548</td>
<td>43</td>
<td>7.8</td>
</tr>
<tr>
<td>1243846</td>
<td>E05003460</td>
<td>Axminster Town</td>
<td>985</td>
<td>65</td>
<td>6.6</td>
</tr>
<tr>
<td>1243847</td>
<td>E05003461</td>
<td>Beer and Branscombe</td>
<td>403</td>
<td>38</td>
<td>9.4</td>
</tr>
<tr>
<td>1243848</td>
<td>E05003462</td>
<td>Broadclyst</td>
<td>1223</td>
<td>152</td>
<td>12.4</td>
</tr>
<tr>
<td>1243849</td>
<td>E05003463</td>
<td>Budleigh</td>
<td>1017</td>
<td>107</td>
<td>10.5</td>
</tr>
<tr>
<td>1243850</td>
<td>E05003464</td>
<td>Clyst Valley</td>
<td>519</td>
<td>59</td>
<td>11.4</td>
</tr>
<tr>
<td>1243851</td>
<td>E05003465</td>
<td>Coly Valley</td>
<td>913</td>
<td>86</td>
<td>9.4</td>
</tr>
<tr>
<td>1243852</td>
<td>E05003466</td>
<td>Dunkeswell</td>
<td>412</td>
<td>33</td>
<td>8.0</td>
</tr>
<tr>
<td>1243853</td>
<td>E05003467</td>
<td>Exe Valley</td>
<td>461</td>
<td>42</td>
<td>9.1</td>
</tr>
<tr>
<td>1243854</td>
<td>E05003468</td>
<td>Exmouth Brixington</td>
<td>1491</td>
<td>121</td>
<td>8.1</td>
</tr>
<tr>
<td>1243855</td>
<td>E05003469</td>
<td>Exmouth Heleston</td>
<td>1547</td>
<td>133</td>
<td>9.9</td>
</tr>
<tr>
<td>1243856</td>
<td>E05003470</td>
<td>Exmouth Littleham</td>
<td>1262</td>
<td>138</td>
<td>10.9</td>
</tr>
<tr>
<td>1243857</td>
<td>E05003471</td>
<td>Exmouth Town</td>
<td>1823</td>
<td>189</td>
<td>10.4</td>
</tr>
<tr>
<td>1243858</td>
<td>E05003472</td>
<td>Exmouth Withycombe Raleigh</td>
<td>1730</td>
<td>129</td>
<td>7.5</td>
</tr>
<tr>
<td>1243859</td>
<td>E05003475</td>
<td>Honiton St Paul's</td>
<td>1030</td>
<td>82</td>
<td>8.0</td>
</tr>
<tr>
<td>1243860</td>
<td>E05003479</td>
<td>Newbridges</td>
<td>470</td>
<td>35</td>
<td>7.4</td>
</tr>
<tr>
<td>1243861</td>
<td>E05003477</td>
<td>Newton Poppleford and Harpford</td>
<td>462</td>
<td>47</td>
<td>10.2</td>
</tr>
<tr>
<td>1243862</td>
<td>E05003478</td>
<td>Otterhead</td>
<td>524</td>
<td>52</td>
<td>9.9</td>
</tr>
<tr>
<td>1243863</td>
<td>E05003480</td>
<td>Ottery St Mary Town</td>
<td>1115</td>
<td>122</td>
<td>10.9</td>
</tr>
<tr>
<td>1243864</td>
<td>E05003481</td>
<td>Raleigh</td>
<td>475</td>
<td>37</td>
<td>7.8</td>
</tr>
<tr>
<td>1243865</td>
<td>E05003482</td>
<td>Seaton</td>
<td>1290</td>
<td>93</td>
<td>7.2</td>
</tr>
<tr>
<td>1243866</td>
<td>E05003483</td>
<td>Sidmouth Rural</td>
<td>396</td>
<td>39</td>
<td>9.8</td>
</tr>
<tr>
<td>1243867</td>
<td>E05003484</td>
<td>Sidmouth Sidford</td>
<td>1272</td>
<td>100</td>
<td>7.9</td>
</tr>
<tr>
<td>1243868</td>
<td>E05003485</td>
<td>Sidmouth Town</td>
<td>918</td>
<td>72</td>
<td>7.8</td>
</tr>
<tr>
<td>1243869</td>
<td>E05003486</td>
<td>Tale Vale</td>
<td>603</td>
<td>39</td>
<td>9.8</td>
</tr>
<tr>
<td>1243870</td>
<td>E05003487</td>
<td>Trinity</td>
<td>516</td>
<td>63</td>
<td>12.2</td>
</tr>
<tr>
<td>1243871</td>
<td>E05003489</td>
<td>Woodbury and Lympstone</td>
<td>1011</td>
<td>108</td>
<td>10.7</td>
</tr>
<tr>
<td>1243872</td>
<td>E05003490</td>
<td>Yarty</td>
<td>504</td>
<td>51</td>
<td>10.1</td>
</tr>
<tr>
<td>1243873</td>
<td>E05003491</td>
<td>Alphington</td>
<td>2238</td>
<td>223</td>
<td>10.0</td>
</tr>
</tbody>
</table>
Aggregate data
Aggregate data

- Questionnaire responses categorised into variables
  - Sex - Male and Female and All
  - Age - single and multiple years
  - Ethnicity and Occupation – standard classifications
- Combinations of characteristics selected through consultations
- Combinations defined as tables
2011 census variables

- 97 variables and counting!

<table>
<thead>
<tr>
<th>Accommodation type</th>
<th>Ethnic group (write in responses)</th>
<th>Persons per bedroom in household, number of</th>
<th>Tenure</th>
<th>Ability to speak Welsh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult life stage (alternative adult definition)</td>
<td>Family status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>General health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of arrival in the UK</td>
<td>Household composition (alternative child and adult definitions)</td>
<td>NS-SEC of household reference person</td>
<td></td>
<td>Age of Family reference person</td>
</tr>
<tr>
<td>Age of children</td>
<td>Household composition</td>
<td></td>
<td></td>
<td>Age of Youngest Dependent Child</td>
</tr>
<tr>
<td>Age of household reference person</td>
<td>Household life stage</td>
<td>Adults in household, number of</td>
<td></td>
<td>Concealed family</td>
</tr>
<tr>
<td>Age of school children and full-time students</td>
<td>Household size</td>
<td>Number of rooms</td>
<td>Year of arrival in the UK</td>
<td>Extended Welsh household composition</td>
</tr>
<tr>
<td>Approximated social grade of household reference person</td>
<td>Household spaces</td>
<td>Occupation</td>
<td>Year last worked</td>
<td>Ethnic group of Household reference person</td>
</tr>
<tr>
<td>Car or van availability</td>
<td>Household type</td>
<td>Occupancy rating (bedrooms)</td>
<td>Passports held, all</td>
<td>Occupation of Household reference person</td>
</tr>
<tr>
<td>Deprivation, classification of household</td>
<td>Qualification, highest level of</td>
<td>Passports held (Principal)</td>
<td>Country of birth (detailed)</td>
<td>Religion of Household reference person</td>
</tr>
<tr>
<td>Country of birth</td>
<td>Hours worked</td>
<td>Position in communal establishment</td>
<td>Occupation (minor groups)</td>
<td></td>
</tr>
<tr>
<td>Communal establishment management and type</td>
<td>Industry</td>
<td>English, proficiency in</td>
<td>Religion (detailed)</td>
<td>Parents working, number of</td>
</tr>
<tr>
<td>Dependent children</td>
<td>Living arrangements</td>
<td>Care (unpaid), provision of</td>
<td>Welsh language skills (detailed)</td>
<td>Ability to speak Welsh, Parents</td>
</tr>
<tr>
<td>Dependent children in family</td>
<td>Long-term health problem or disability</td>
<td>Qualifications gained</td>
<td>Second Address Location</td>
<td>Households with school children or full-time students away during term-time</td>
</tr>
<tr>
<td>Dwelling type</td>
<td>Length of residence in the UK</td>
<td>Religion</td>
<td>Second Address Type</td>
<td>Number of people in household who can speak Welsh</td>
</tr>
<tr>
<td>Economic activity</td>
<td>Marital and civil partnership status</td>
<td>Resident type</td>
<td>Population type</td>
<td>Long-term health problem or disability, number of people in household</td>
</tr>
<tr>
<td>Economic activity of household reference persons</td>
<td>Multiple ethnic groups</td>
<td>Residence type</td>
<td>Ethnic group UK</td>
<td>Country of Birth (additional categories)</td>
</tr>
<tr>
<td>Ethnic group (detailed)</td>
<td>Main language (identified)</td>
<td>Second address</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English as a household language</td>
<td>Method of travel to work</td>
<td>Sex</td>
<td>Household Reference Person</td>
<td></td>
</tr>
<tr>
<td>Ethnic group</td>
<td>National identity</td>
<td>Central heating in household, type of</td>
<td>Age of Dependent Children</td>
<td></td>
</tr>
</tbody>
</table>
Table Frameworks
UK 2011 Census

- 27 March 2011
- Three UK census agencies (ONS, NRS, NISRA)
- New questions and variables
- Targeted enumeration
- Online and postal completion
- Sophisticated quality assurance
New variables

- National identity
- Passports held
- Ability in spoken English
- Languages other than English used at home
- Long term health conditions (detailed for Scotland/Northern Ireland)
- Month/year of arrival into the UK (for people not born in the UK)
- Intention to stay
- Second homes
Example analysis

Figure 1: Geographical distribution of non-English language speakers and English proficiency by local authority district, 2011

a) Percentage of population reporting main language other than English
b) Percentage of those who reported a main language other than English whose self-reported English proficiency is poor or nonexistent

Who can and cannot speak English? DYNAMICS OF DIVERSITY: Evidence from the 2011 Census. ESRC Centre on Dynamics of Ethnicity (CoDE):
http://www.ethnicity.ac.uk/
Main language other than English by number of speakers in England and Wales

- ???
- Panjabi
- Urdu
- Bengali
- Gujarati
- Arabic
- French
- Chinese, other
- Portuguese
- Spanish
What is the largest non-English main language by number of speakers in England and Wales?
Main language other than English by number of speakers in England and Wales

- Polish: 550,000
- Panjabi: 250,000
- Urdu: 200,000
- Bengali: 200,000
- Gujarati: 200,000
- Arabic: 150,000
- French: 150,000
- Chinese, other: 100,000
- Portuguese: 75,000
- Spanish: 75,000
2011 Census geographies

- Subdivisions of the UK into smaller *areas*
- Sets of similar areas called *geographies*
  - Functional and statistical
  - Local government districts
  - Wards and electoral divisions
  - Expecting around 100 different geographies
- Hierarchies of geographies with nesting areas
  - Administrative
  - Statistical
  - Health, Electoral, Postcode, etc
2011 census administrative hierarchy
UK statistical geographies

United Kingdom
- England & Wales
  - England
    - Region
      - County
        - District
          - Small Area
            - SOA
              - LSQA
                - Small Area
                  - OA
                    - OA (42,604)
What data is available for which areas?

• Not all characteristics for all areas
• Principle of confidentiality
  • Information detail vs geographical detail
• Lower threshold data
  • Produced for all areas
  • Key Statistics, Quick Statistics, Local Characteristics
• Upper threshold data
  • Produced for areas down to wards* and MSOAs
  • Detailed Characteristics
• Other outputs
  • UK data to district level
Dissemination challenges

• Size and complexity
  • Characteristics
  • Geographies
• Fragmented supply from three different agencies
• Inconsistencies in definitions
  • Differences within countries and across the UK
• Inhibits global search and understanding
What interfaces have you used?
InFuse

- Data model
- Geography model
- Benefits
InFuse data model

- Single multidimensional dataset
- Deconstruction, rationalisation and re-integration of variables and categories
  - All UK table specifications processed
  - Integration of table universes as variables
  - Enforce consistency across dataset
- Re-insertion of counts into model
  - Retain original cell identifiers
- Attachment of metadata
InFuse geography model

- Simplified geography model
  - Combinations of equivalent geographies into sets
  - Condensed standard/merged geographies
- Selections of areas across the UK
  - Multiple geographies in one operation
  - Geography jumps in interface
- Currently administrative and statistical geographies
  - More to follow
Raw geography relationships
InFuse administrative and statistical geographies

- United Kingdom
  - Great Britain
    - England & Wales
  - Country Layer
    - Region Layer
      - County Layer
    - District Layer
      - Middle Super Output Area Layer
      - Ward Layer
    - Lower Super Output Area Layer
      - Output Area Layer
InFuse benefits

• Open access
  • All data is open via Open Government Licence

• Global search across entire UK dataset
  • No tables!

• Guide users to find data
  • Variable combinations
  • Available geographies

• More data for more geographies
  • *All LT and HT data available for all areas* in condensed geographies

• Improved contextual information

• No data fast!
InFuse 2011 demonstration

- [http://infuse.mimas.ac.uk/](http://infuse.mimas.ac.uk/)
- Data currently available
- Data coming soon
What’s next?

- Big data release imminent!
- Progressive release of UK 2011 outputs
  - Scottish and Northern Ireland
- Previous censuses (currently in Casweb)
- Boundary data
- Interface design and features
- Access to API for application development
- Continue engagement with census agencies
  - Better data from producers
  - Less processing!
InFuse Support

infuse.mimas.ac.uk
census.ukdataservice.ac.uk
Questions

ukdataservice.ac.uk/help/

Follow us at:
• UKDATASERVICE@JISCMAIL.AC.UK
• twitter.com/UKDataService
• www.facebook.com/UKDataService