

# Census flow data workshop

Vassilis Routsis

Oliver Duke-Williams

Tuesday 8<sup>th</sup> December 2020

# Structure

- Welcome and polls
- Presentation – What are flow data?
- Demonstration – exploring with WICID and downloading data
- Practical activity on retrieving data
- Feedback from activity
- Q&A

# What is the UK Data Service?

- A comprehensive resource funded by the Economic and Social Research Council (ESRC)
- A single point of access to a wide range of secondary social science data
- Support, training and guidance

What are census flow data?

# What is flow data?

**Flow data** are also known as **'interaction data'** or **'origin-destination data'**

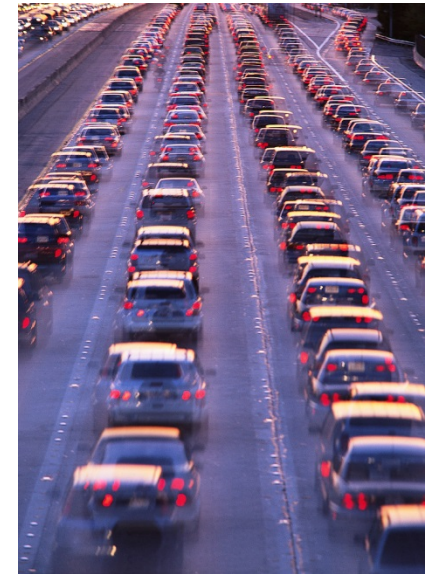
# What is flow data? (2)

The data consist of counts of flows between two locations

# Examples of flow data

Examples include

- Migration data (within UK and to UK)
- Commuting data
- Journeys to school
- Movements associated with a second residence

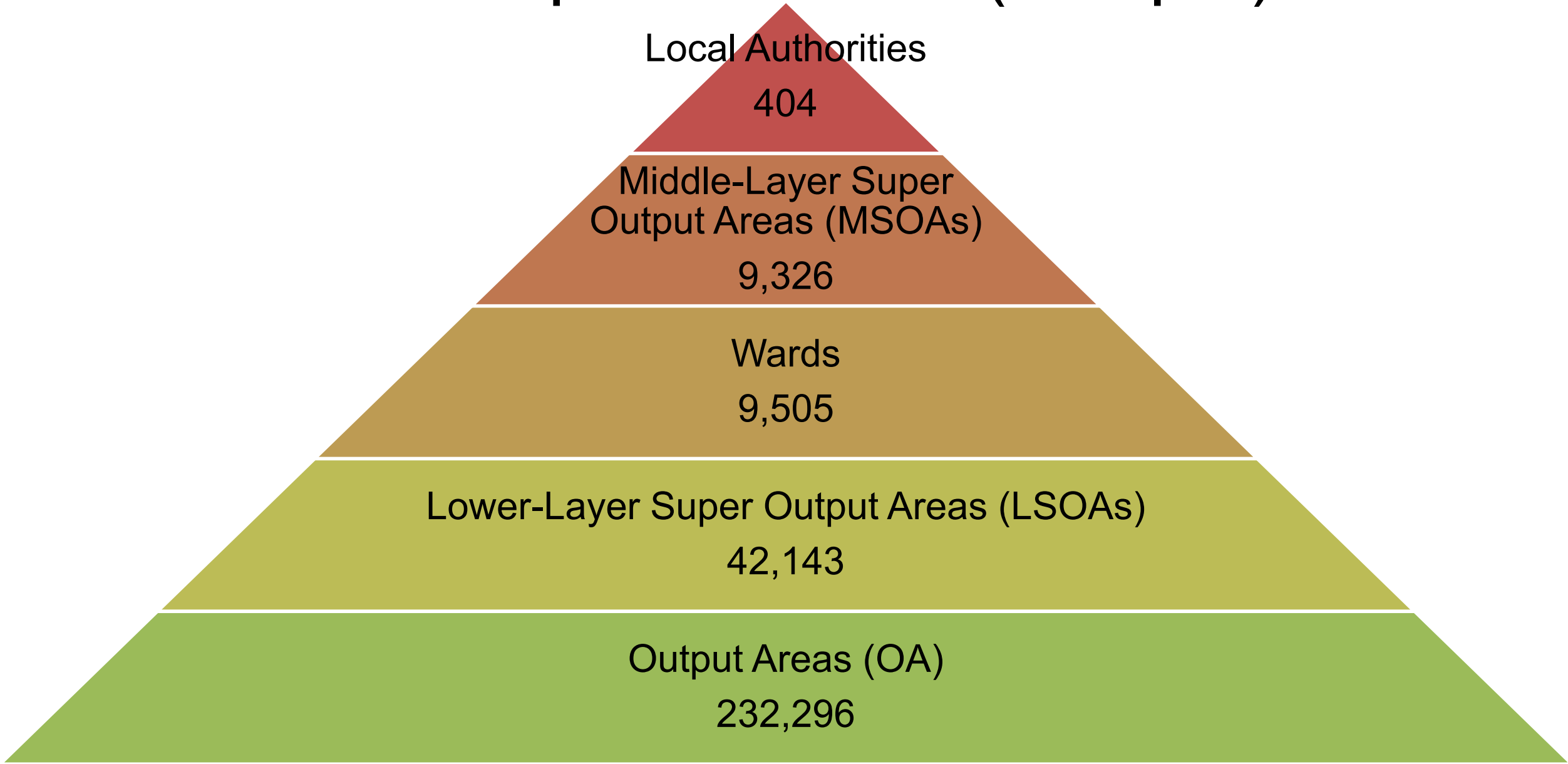


# Spatial scales

Data are produced at different spatial scales



# 2011 Census Spatial scales (sample)



# Pseudo-spatial areas

As well as 'normal' areas, flows may be associated with various aggregate and pseudo-spatial areas.

- 'Overseas', 'Works at home', 'Works off-shore' etc.

# Changing approaches

There has been different handling of cases such as these over time

# Delivery of census flow data

Flow data are one of the last census data products because counts have to be collated from the three different national statistical agencies

# Questionnaire Specimens

**Household Questionnaire Northern Ireland**

Complete online  
[www.census.gov.uk/ni](http://www.census.gov.uk/ni)  
 Your personal internet access code is:

OR fill in this paper questionnaire and post it back using the pre-paid envelope supplied.

If your address is incorrect or missing, enter your correct address here:

**Why the Census Matters**  
 The census is used to help plan and fund services for your community - services like transport, education and health.  
 Everyone should be included in the census - all people, households and overnight visitors.  
 Please complete your census questionnaire on 27 March 2011, or as soon as possible afterwards. You can fill it in online or on paper.  
 Taking part in the census is very important and it is also compulsory. You could face a fine if you don't participate or if you supply false information.  
 Your personal information is protected by law and will be kept confidential.  
 So help tomorrow take shape and be part of the 2011 Census.

**Declaration**  
 This questionnaire has been completed to the best of my knowledge and belief.

Signature: \_\_\_\_\_  
 Date: \_\_\_\_\_ Telephone number: \_\_\_\_\_

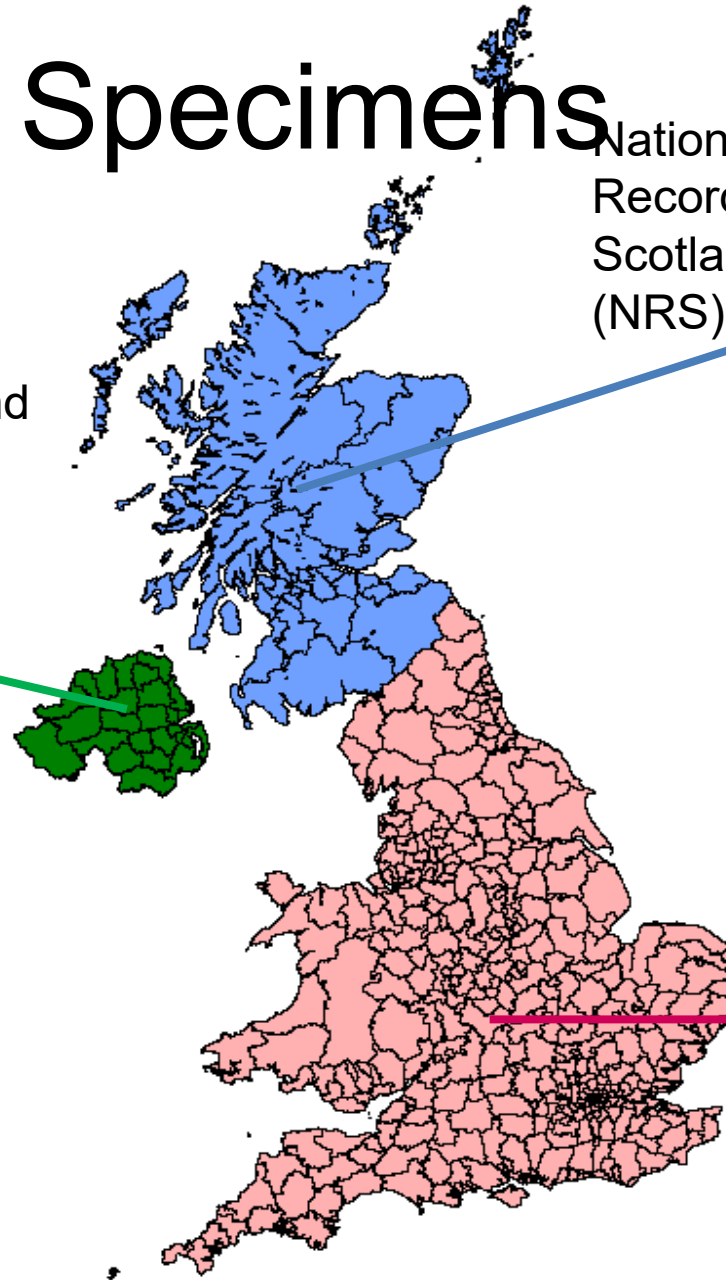
We may contact you if we need to collect missing information.  
 If you have lost your envelope, please return to: FREEPOST 2011 Census, Processing Centre.

**Where can you get help?**  
[www.census.gov.uk/ni](http://www.census.gov.uk/ni)  
 Census helpline 0300 0201 120  
 Text Relay 18001 0300 0201 170

Help is available in large print and Braille.

**H4**

Northern Ireland  
 Statistics and Research Agency (NISRA)



National Records of Scotland (NRS)

**Scotland's Census 2011** Shaping our future

Household Questionnaire <sup>HO</sup>  
 27 March 2011

Office Use: CD ED Line Number

If there is a mistake in the printed address, please write your correct address below:

Why the census matters  
 The census is the official count of every person and household in Scotland. It is held every 10 years and helps to plan our future public services.  
 Please fill in this questionnaire on, or around, Sunday 27 March. Please include everyone in this address. It shouldn't take long and you can fill it in online.  
 As a householder, you have a legal duty to fill in this questionnaire. If you don't, or if you supply false information, you may be fined up to £1,000.  
 Your personal information is protected by law and we will keep it confidential for 100 years.  
 Thank you for helping to shape Scotland's future.

**Declaration**  
 I have filled in this questionnaire fully and accurately, as far as I know.

**Need help?**  
[www.scotlandscensus.gov.uk](http://www.scotlandscensus.gov.uk)  
 Helpline 0300 123 1702  
 Textphone 18001 0300 123 1703

**Where can you get help?**  
[www.census.gov.uk](http://www.census.gov.uk)  
 Census helpline 0300 0201 101  
 Text Relay 18001 0300 0201 160

Help is available in large print and Braille.

Office for National Statistics (ONS)

**Household Questionnaire England**

Complete online  
[www.census.gov.uk](http://www.census.gov.uk)  
 Your personal internet access code is:

OR fill in this paper questionnaire and post it back using the pre-paid envelope supplied.

If your address is incorrect or missing, enter your correct address here:

**Why the Census Matters**  
 The census is used to help plan and fund services for your community - services like transport, education and health.  
 Everyone should be included in the census - all people, households and overnight visitors.  
 It is used to help plan and fund services for your community - services like transport, education and health.  
 Please complete your census questionnaire on 27 March 2011, or as soon as possible afterwards. You can fill it in online or on paper.  
 Taking part in the census is very important and it is also compulsory. You could face a fine if you don't participate or if you supply false information.  
 Your personal information is protected by law. Census information is kept confidential for 100 years.  
 So help tomorrow take shape and be part of the 2011 Census.

**Declaration**  
 This questionnaire has been completed to the best of my knowledge and belief.

Signature: \_\_\_\_\_  
 Date: \_\_\_\_\_ Telephone number: \_\_\_\_\_

We may contact you if we need to collect missing information.  
 If you have lost your envelope, please return to: FREEPOST 2011 Census, Processing Centre, UK.

**Where can you get help?**  
[www.census.gov.uk](http://www.census.gov.uk)  
 Census helpline 0300 0201 101  
 Text Relay 18001 0300 0201 160

Help is available in large print and Braille.

**H1**

# Progression of flow data sets

1981



Special Migration Statistics

Special Workplace Statistics

1991



SMS

SWS

2001



SMS

SWS

Special Travel Statistics

2011



SMS

Special Student Statistics

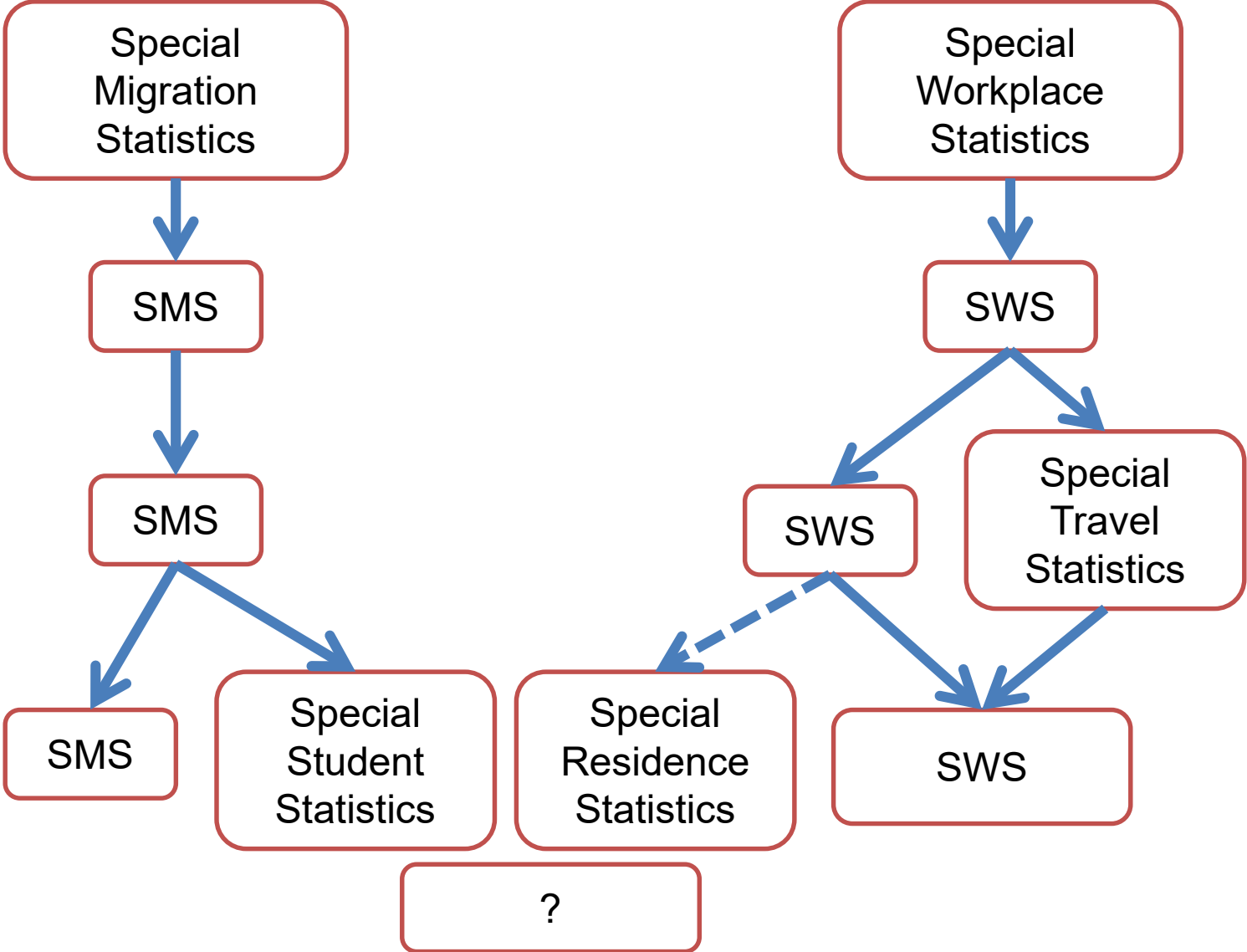
Special Residence Statistics

SWS

2021



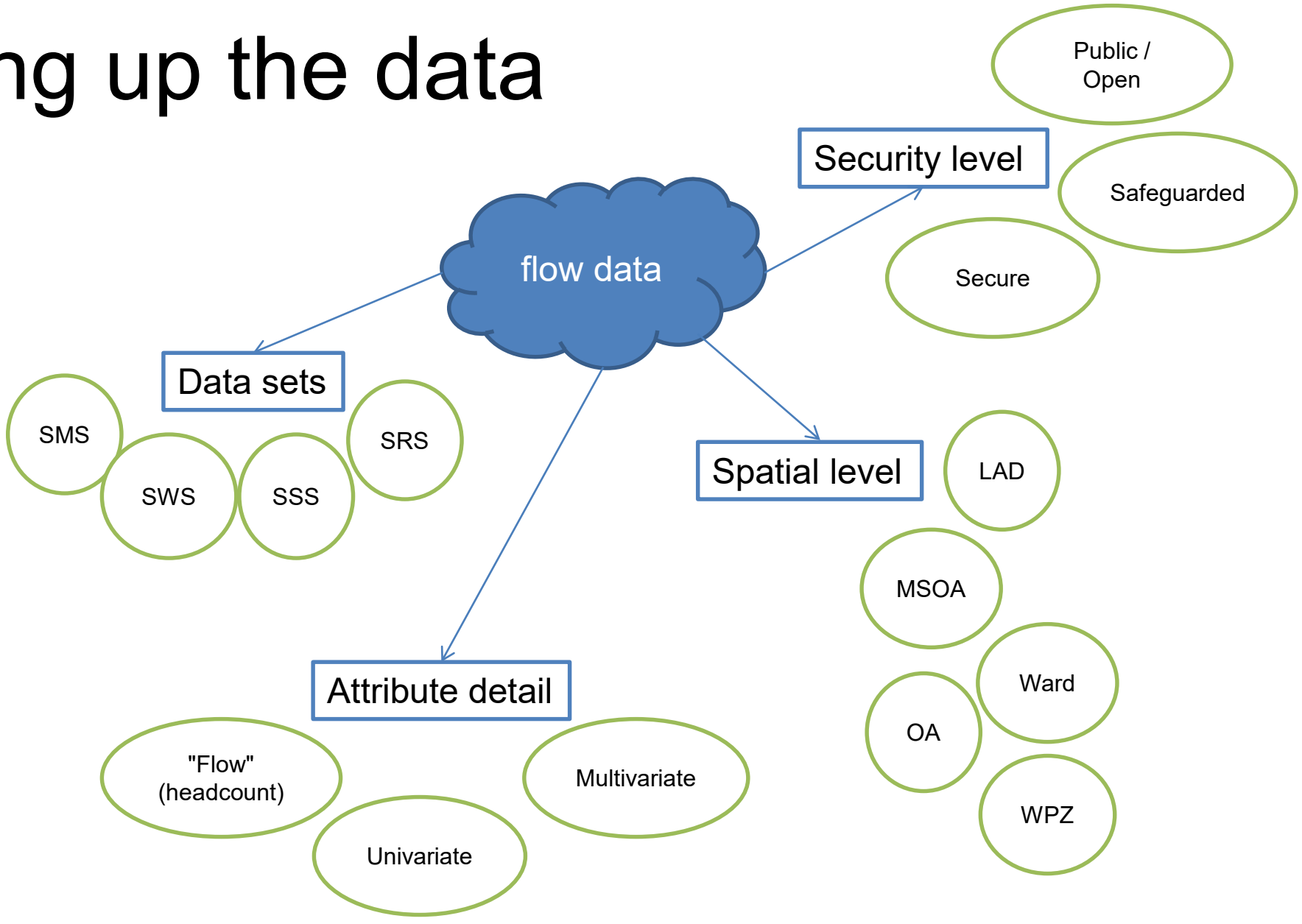
?



# Access to data

- Data from 1981, 1991, 2001 are publicly available
- Data from 2011 is more confusing
  - Multiple levels of access
  - Trade off between spatial detail and attribute detail
  - Different routes to using the data

# Splitting up the data





# Census questions

The various data sets all stem from questions on the Census form:

## Migration tables

- Flows between an origin and a destination
  - based to question 21: What was your usual address one year ago?

## Student tables

- A subset of migration data
  - a subset of migration tables for people who indicated that their address one year ago was a student / school address

## Workplace tables

- Journey to work. Flows between a residence and a workplace
  - based on question 40: In your main job, what is the address of your workplace?

## Second residence tables

- Several different sets of flows, e.g. second residence to work, first residence to second residence etc.
  - based on questions 5 and 6 for England and Wales only: Do you stay at another address for more than 30 days a year? What is that address?

# Sets v Tables 1981 to 2001

- In the case of 1981, 1991 and 2001, the flow data were grouped into sets (e.g. '1991 SMS Set 1'; '2001 SWS Level 2'; etc.)
- Each set consisted of several tables
- Generally all tables in each set were defined by the same combination of geographies
  - An origin geography (e.g. wards, districts, etc.)
  - A destination geography
  - In some cases, a few additional non-spatial categories

# Sets v Tables 2011

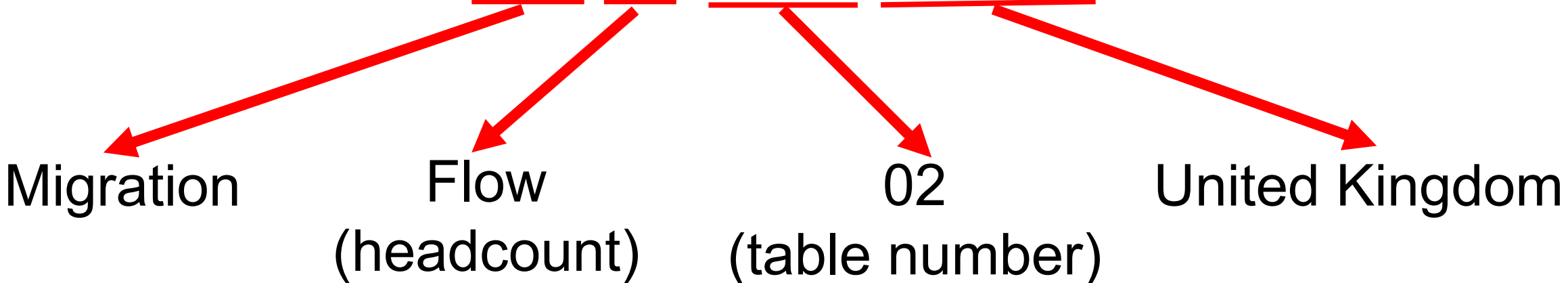
- In the 2011 data, there are not explicit sets in the same way
  - Each family (SMS, SWS etc.) consist of many tables, and crucially these do *not* have common geography definitions

# Table IDs

- The tables cannot be grouped into sets as easily as before
- The Table IDs describe aspects of the data
  - [M/W/S/R] – first letter: migration, workplace, student or (second) residence
  - [F/U/M] – second letter: flow (headcount), univariate or multivariate
  - nn(m) – table number, sometimes with suffix A/B/C further splitting the table up
  - [UK/EW/NI/etc.] – suffix indicating whether table relates to flows in whole of UK, or England and Wales, Northern Ireland etc.

# Table IDs exaple

**MF02UK**



# Three levels of user access

- Different tables available at one of three levels:

**Public:** available via ONS, nomisweb and UK Data Service to any users via Open Government Licence (OGL) without the need to register or login to any system

**Safeguarded:** available via UKDS to members of academia, local and central government, NHS, and UK parliaments and assemblies via End User Licence (EUL) so need to register and login

**Secure:** available to approved researchers via the Approved Researcher Scheme using the ONS Secure Research Service (SRS) — formerly known as Virtual Microdata Laboratory

# Safeguarded access

You do not get automatic access to the safeguarded data. You need to register with UK Data Service before using your institutional account to get to these tables.

Currently, researchers with the following email suffixes can access the Safeguarded WICID tables after registering with UKDS and accepting the end-user licence:

- ac.uk
- gov.uk
- nhs.uk
- nhs.net
- parliament.uk
- scottish.parliament.uk
- llyw.cymru
- gov.scot
- gov.wales

# Types of table counts

**Flow  
(headcount)**  
[MF02UK]

{Origin and destination}	Persons
Total	1

**Univariate**  
tables relate  
to one  
variable  
[MU02UK]

{Origin and destination}	Persons
All categories: Family status	1
Not in a family: aged 65 and over	2
Not in a family: aged under 65	3
In a couple family: member of couple	4
In a couple family: dependent child of one or both members of the couple	5
In a couple family: non-dependent child of one or both members of the couple	6
In a lone parent family: parent	7
In a lone parent family: dependent child of parent	8
In a lone parent family: non-dependent child of parent	9
Not in a household	10

**Multivariate**  
tables where  
one variable  
is cross-  
classified by  
at least one  
other  
variable  
[MM02UK]

{Origin and destination}	All categories: Sex	Male	Female
All categories: Family status	1	2	3
Not in a family: aged 65 and over	4	5	6
Not in a family: aged under 65	7	8	9
In a couple family: member of couple	10	11	12
In a couple family: dependent child of one or both members of the couple	13	14	15
In a couple family: non-dependent child of one or both members of the couple	16	17	18
In a lone parent family: parent	19	20	21
In a lone parent family: dependent child of parent	22	23	24
In a lone parent family: non-dependent child of parent	25	26	27
Not in a household	28	29	30



# Flow (headcount) data sets: Security classification

Geography	SMS	SRS	SSS	SWS
LA to LA	<ul style="list-style-type: none"><li>• <b>Public</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Public</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Public</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Public</b></li></ul>
Ward to Ward MSOA to MSOA	<ul style="list-style-type: none"><li>• <b>Safeguarded</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Safeguarded</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Safeguarded</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Safeguarded</b></li></ul>
OA to OA OA to WPZ	<ul style="list-style-type: none"><li>• <b>Safeguarded</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Safeguarded</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Safeguarded</b></li></ul>	<ul style="list-style-type: none"><li>• <b>England &amp; Wales: Public</b></li><li>• <b>UK: Safeguarded</b></li></ul>

# Univariate data sets: Security classification

Geography	SMS	SRS	SSS	SWS
LA-LA	<ul style="list-style-type: none"> <li>• <b>Sex, Age: Public</b></li> <li>• <b>Others: safeguarded</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex, Age: Public</b></li> <li>• <b>Others: safeguarded</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex, Age: Public</b></li> <li>• <b>Others: safeguarded</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex, Age, Method of travel: Public</b></li> <li>• <b>Others: safeguarded</b></li> </ul>
Ward - Ward MSOA-MSOA	<ul style="list-style-type: none"> <li>• <b>Sex, Age: Safeguarded</b></li> <li>• <b>Others: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex, Age: Safeguarded</b></li> <li>• <b>Others: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex, Age: Safeguarded</b></li> <li>• <b>Others: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex, Age, Method of travel: Public</b></li> <li>• <b>Others: Safeguarded</b></li> </ul>
OA to OA  OA to WPZ	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>

# Multivariate data sets: Security classification

Geography	SMS	SRS	SSS	SWS
LA-LA	<ul style="list-style-type: none"> <li>• <b>Sex by Age: Public</b></li> <li>• <b>Some vars: safeguarded</b></li> <li>• <b>Other vars: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex by Age: Public</b></li> <li>• <b>Some vars: safeguarded</b></li> <li>• <b>Other vars: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex by Age: Public</b></li> <li>• <b>Some vars: safeguarded</b></li> <li>• <b>Other vars: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex by Age: Public</b></li> <li>• <b>Some vars: safeguarded</b></li> <li>• <b>Other vars: Secure</b></li> </ul>
Ward - Ward MSOA-MSOA	<ul style="list-style-type: none"> <li>• <b>Sex by Age: Safeguarded</b></li> <li>• <b>Others: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex by Age: Safeguarded</b></li> <li>• <b>Others: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sex by Age: Safeguarded</b></li> <li>• <b>Others: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Some vars: safeguarded</b></li> <li>• <b>Other vars: Secure</b></li> </ul>
OA to OA	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>All: Secure</b></li> </ul>
OA to WPZ				

# 2021 Census and beyond

- Most likely 2021/22 will be the last traditional census as we know it.
- The first batches of 2021 census flow data are not expected before late 2022.
- We expect harmonisation issues for cross-border flows because the next census data will be collected in Scotland one year later than the rest of the UK.

# Using the data....

- Using table MM01, we can look at flows at ward level from origins outside the UK
  - We find the modal origin (i.e. the origin from which most migrants arrived) for each destination
    - Download via WICID
    - Load into a local SQL database
    - Re-aggregate it on the basis of origin and flow size...
  - We then group by origin, so that we find how many wards had 'x' as the most common country of origin
  - We then look at the countries that are modal for the largest number of wards...

# POLL

- Here are a selection of countries: one of these is the most common origin country for migrants from outside the UK in 897 wards (this is the most frequently seen 'most common origin country'). Which one?
  - Australia
  - China
  - India
  - Poland
  - Spain

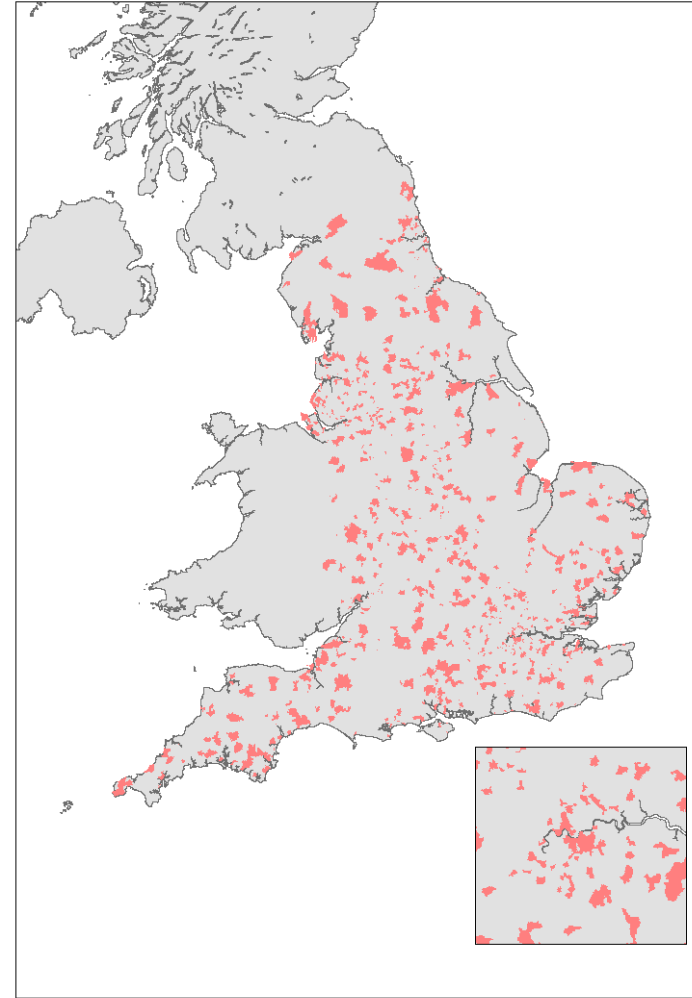
# Modal country of origin (ward level)

Origin country	Number of wards
Australia	897
Spain (including Canary Islands)	855
USA	795
Poland	780
Other EU accession countries	720
Germany	517
France	509
India	440
Other Middle East	388
China	230

Australia was the most common country of origin in 897 wards in England and Wales

# Australia

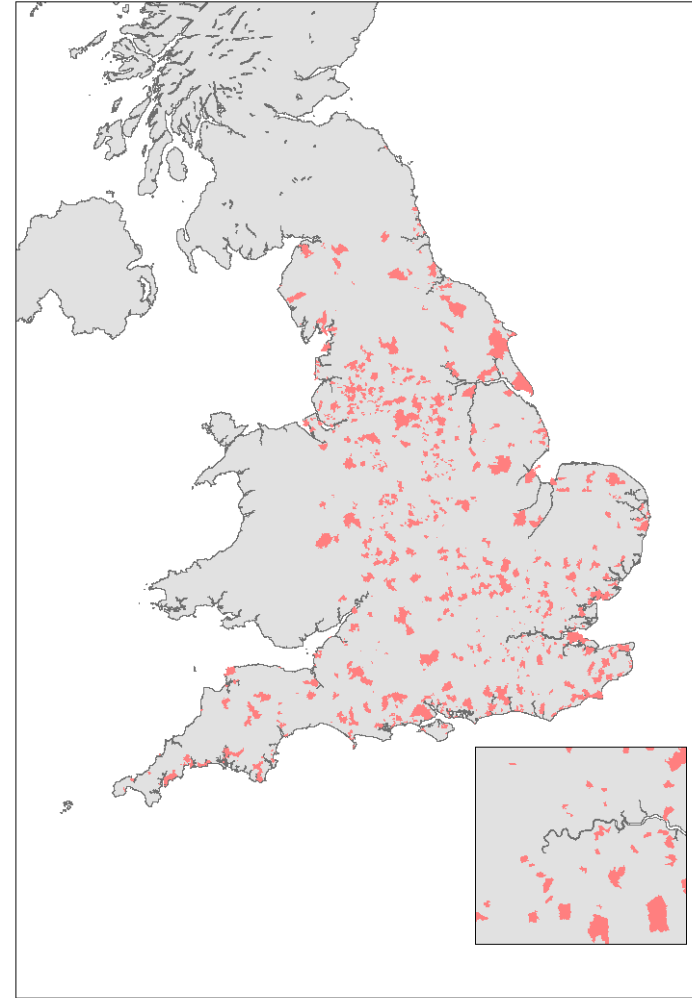
Highlighted wards are those in England where the most common origin for migrants from outside the UK was Australia





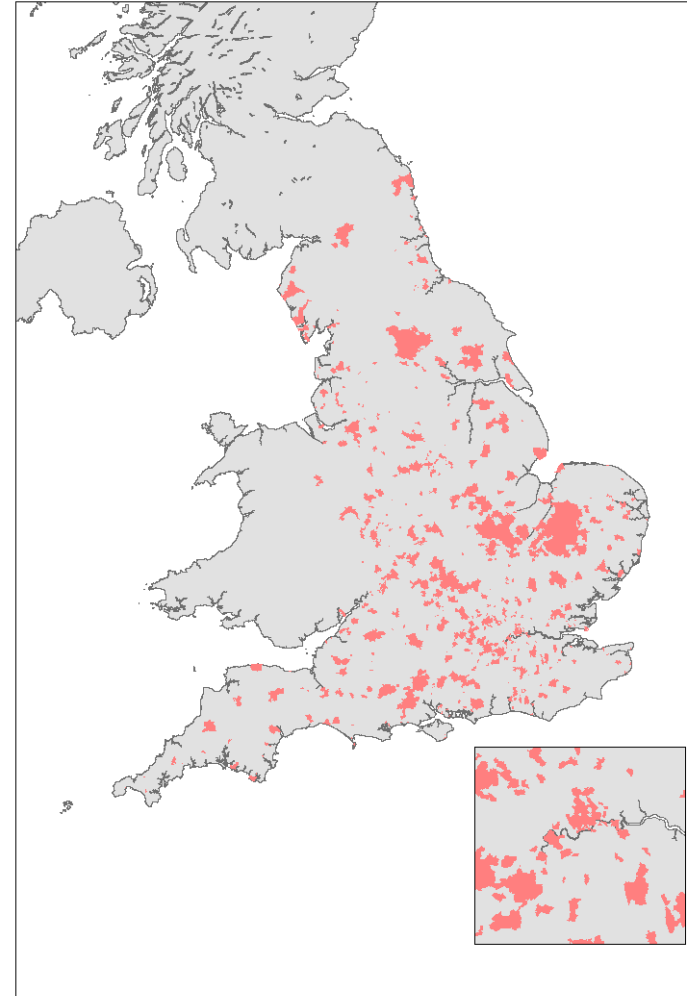
# Spain

Highlighted wards are those in England where the most common origin for migrants from outside the UK was Spain



# USA

Highlighted wards are those in England where the most common origin for migrants from outside the UK was USA

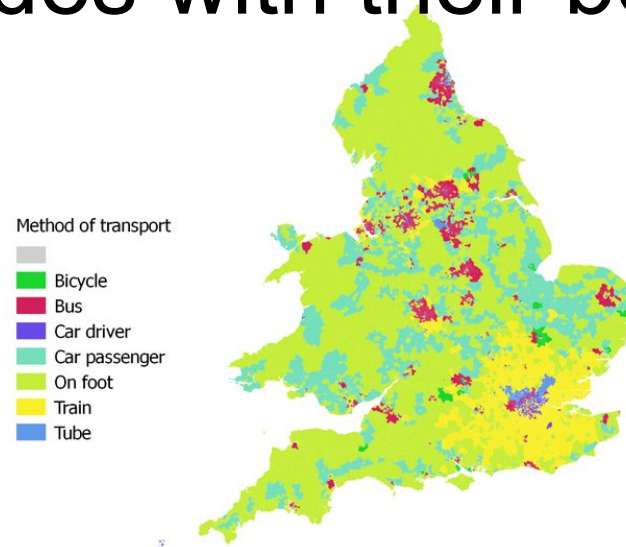


# Accessing the data

Walk through and practical task

# Exercises with download data

- The processed data can be used to generate charts or even maps using tools like QGIS by matching the area codes with their boundaries



**Second most common method of transport to work, 2011  
England and Wales, MSOA level**

Boundaries: Office for National Statistics, 2011 Census: Digitised Boundary Data (England and Wales) [computer file].  
UK Data Service Census Support. Downloaded from: <https://borders.ukdataservice.ac.uk/>

Flow data: Office for National Statistics, 2011 Census: Special Workplace Statistics (United Kingdom) [computer file].  
UK Data Service Census Support. Downloaded from: <https://wicid.ukdataservice.ac.uk>

The above map was produced using WU03UK safeguarded table at MSOA level