

Accessing and working with UK census flow data

Oliver Duke-Williams & Vassilis Routsis

January 2025



Structure

- Welcome
- Presentation – What are flow data?
- Demonstration – exploring with WICID and downloading data
- Practical activities on retrieving, processing and analysing census flow data
- 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 is census flow data?

What is flow data?

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

The data consist of counts of flows between two locations

Examples of flow data

Examples include

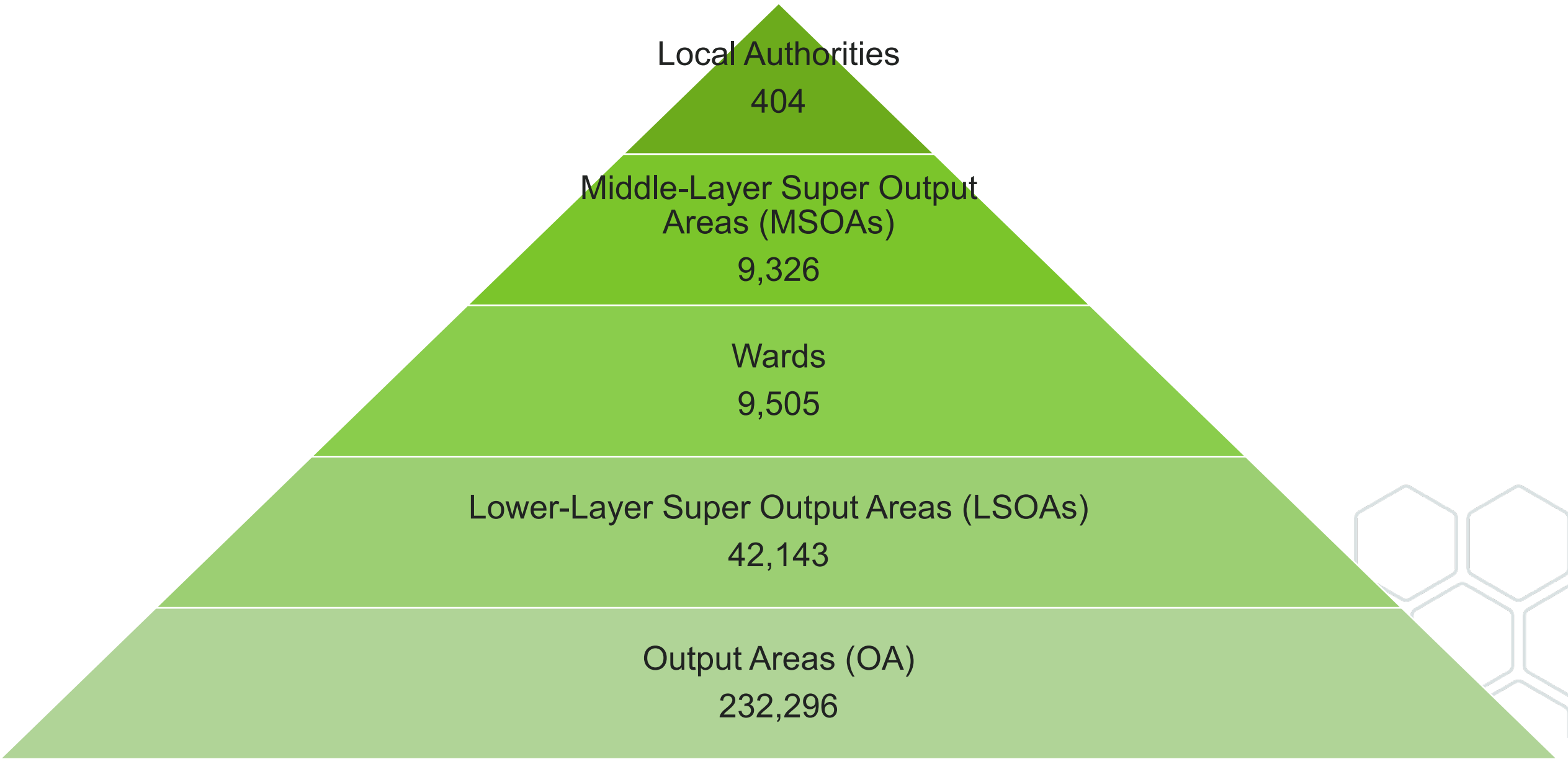
- Migration data (within UK and to UK)
- Commuting data
- Journeys to school
- Movements associated with a second residence



Spatial scales

Data are produced at different spatial scales

2011 Census Spatial scales (sample) - UK



2021 Census Spatial scales (sample) – E&W

Upper Tier Local
Authorities (UTLA)

174

Lower Tier Local
Authorities (LTLA)

332

Middle-Layer Super Output Areas
(MSOAs)

7,264

Lower-Layer Super Output Areas (LSOAs)

35,672

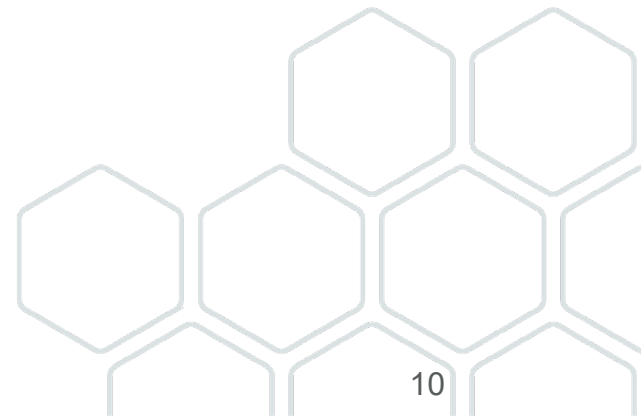
Output Areas (OA)

188,880

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

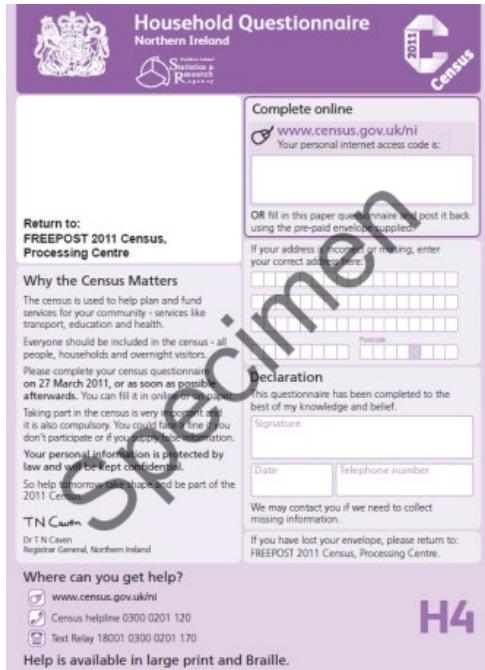
Flow data are one of the last census data products because counts have to be collated from the three different national statistical agencies.

Questionnaire Specimens

Northern
Ireland
Statistics and
Research
Agency
(NISRA)

National
Records of
Scotland
(NRS)

Office for
National
Statistics
(ONS)



Household Questionnaire Northern Ireland

Complete online
 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:

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

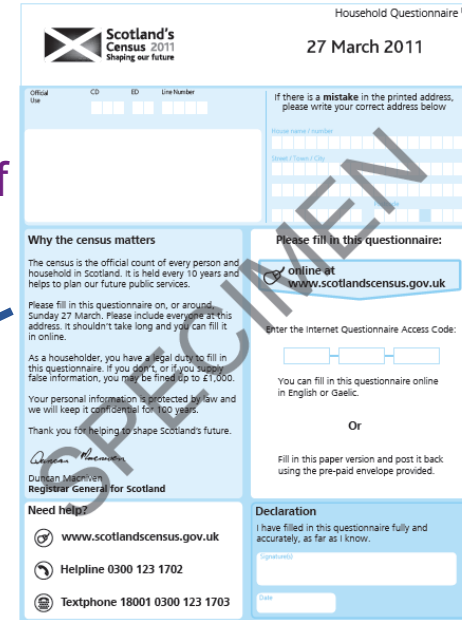
Signature: _____
 Date: _____ Telephone number: _____

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
 Census helpline 0300 0201 120
 Text Relay 18001 0300 0201 170

H4



Scotland's Census 2011 Household Questionnaire ^{HO}
 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.

Need help?
 www.scotlandscensus.gov.uk
 Helpline 0300 123 1702
 Textphone 18001 0300 123 1703

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

Signature: _____
 Date: _____



Household Questionnaire England

Complete online
 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:

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

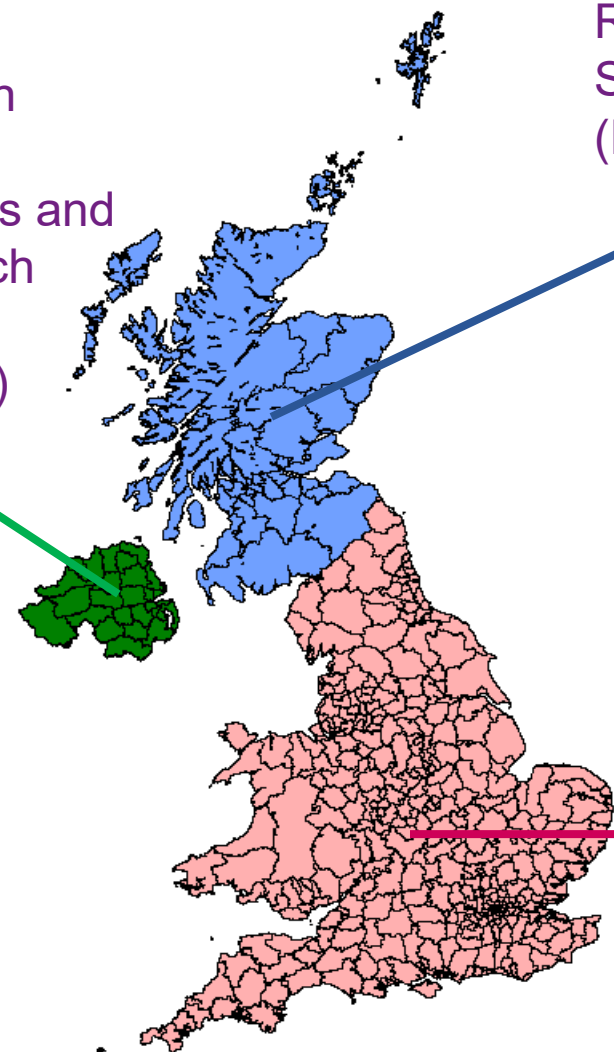
Signature: _____
 Date: _____ Telephone number: _____

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
 Census helpline 0300 0201 101
 Text Relay 18001 0300 0201 160

H1



Progression of flow data sets

1981



1991



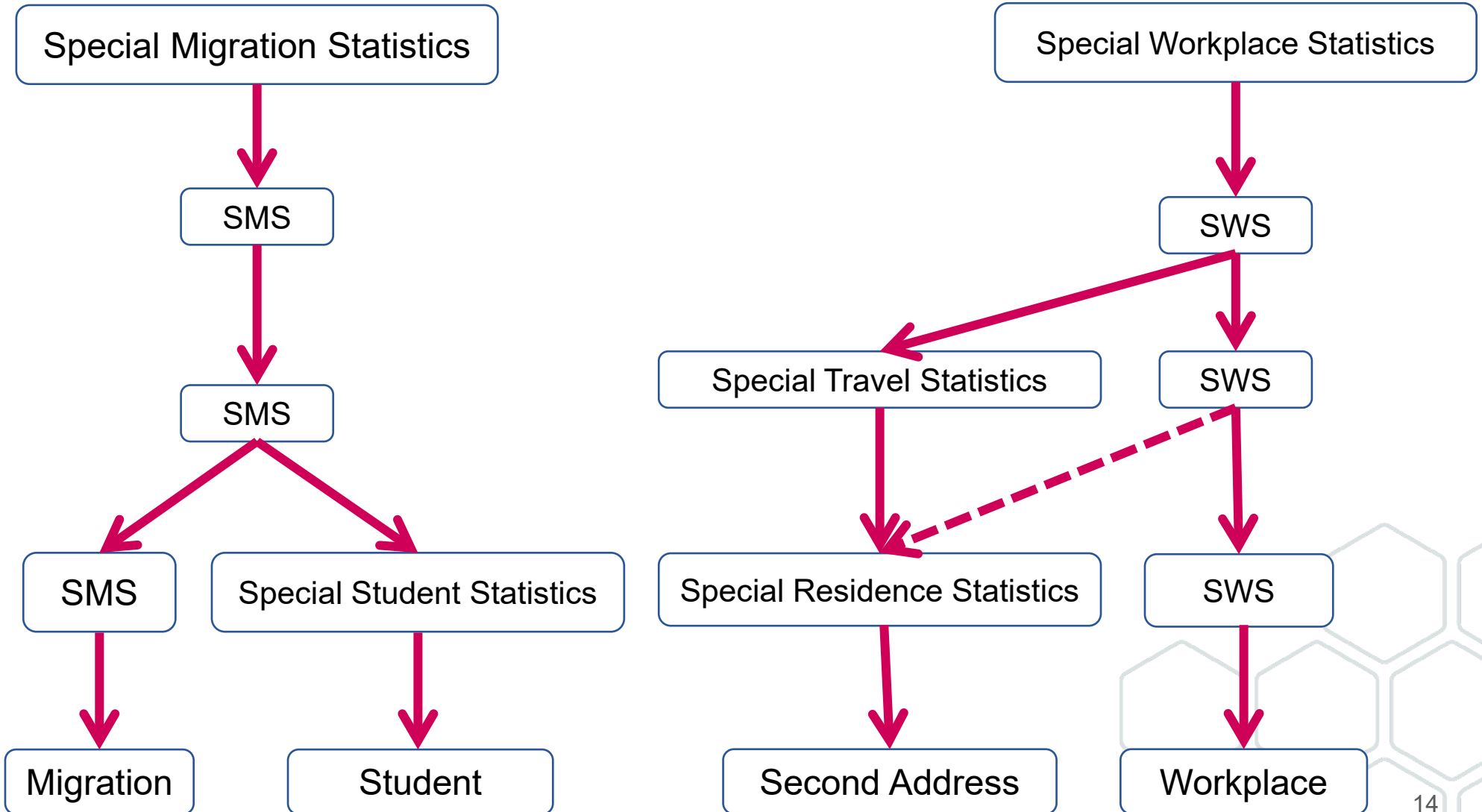
2001



2011



2021



Access to data

- Data from 1981, 1991, 2001 are publicly available
- Data from 2011 and 2021 is more confusing
- Multiple levels of access
- Trade off between spatial detail and attribute detail
- Different routes to using the data

Census flow data specifications

Flow data taxonomy



Types of table counts (table structure)

**Flow
(headcount)**
[MF02UK]

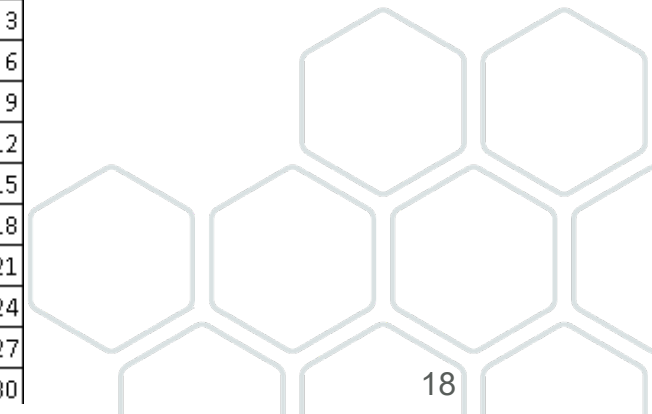
<i>{Origin and destination}</i>	Persons
Total	1

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

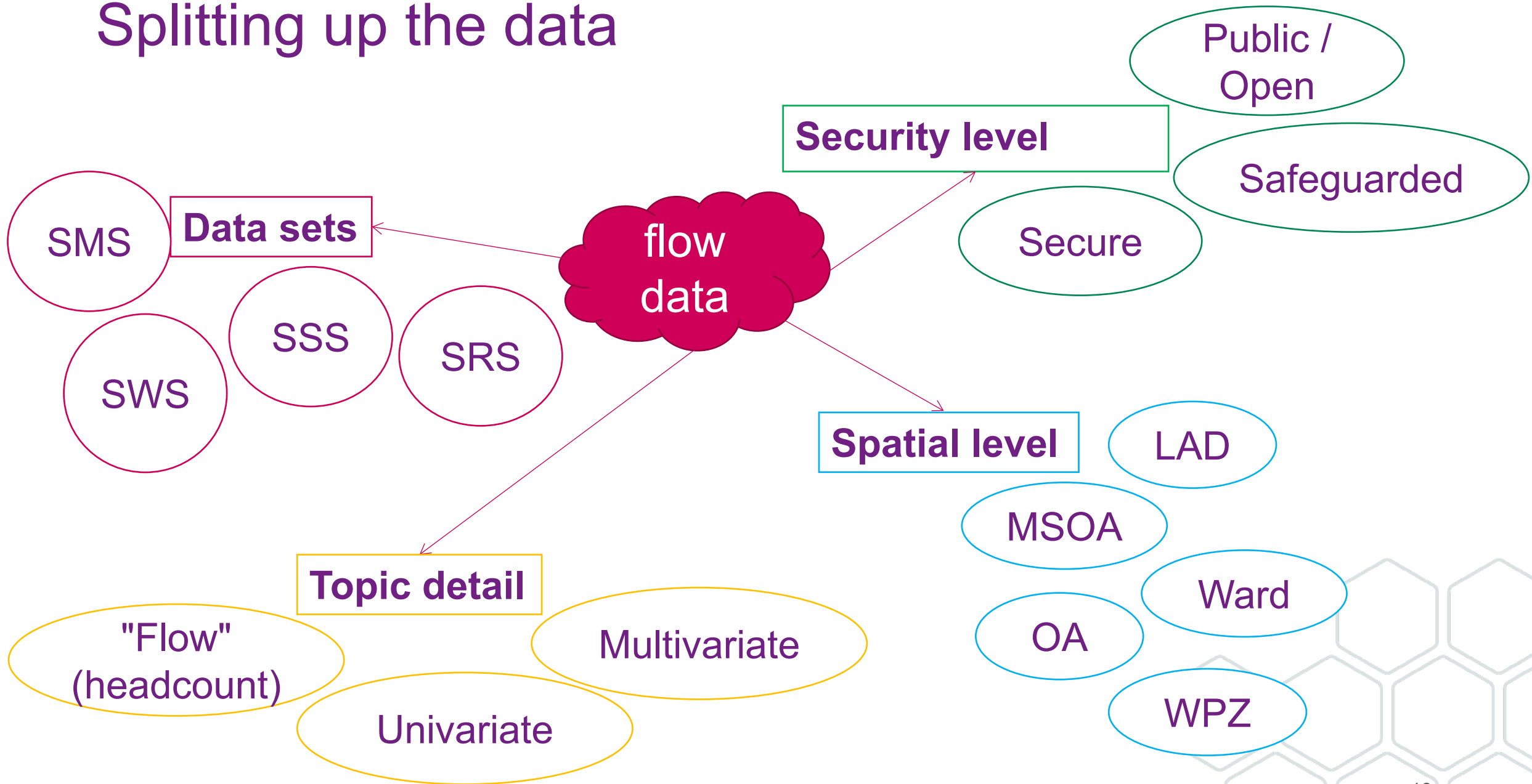
<i>{Origin and destination}</i>	Persons
All categories: Family status	1
Not in a family: aged 65 and over	2
Not in a family: aged under 65	3
In a couple family: member of couple	4
In a couple family: dependent child of one or both members of the couple	5
In a couple family: non-dependent child of one 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 topic is
cross-classified
by at least one
other topic**
[MM02UK]

<i>{Origin and destination}</i>	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



Splitting up the data



2011 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?

2021 Census questions – E&W

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

Migration tables

- Flows between an origin and a destination
 - based to question 13: One year ago, what was your usual address?

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 50: What is the address of your workplace or depot?

Second address tables

- Several different sets of flows, e.g. second residence to work, first residence to second residence etc.
 - based on questions 6 and 7: 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.

Tables 2011 and 2021

In the 2011 and 2021 data, there are no explicit sets in the same way as pre-2011

Each family (migration, workplace etc.) consist of many tables, and crucially, these do not have common geography definitions

2011 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 the table relates to flows in the whole of the UK, or England and Wales, Northern Ireland, etc.

NON_UK: only about the population living outside the UK one year before the census

2021 Table IDs

The Table IDs describe aspects of the data:

[OD]: Origin and destination dataset type

[MG/WP/ST/SA]: migration (MG), workplace (WP), student (ST), or second address (SA)

Nn – dataset number, with '01' being the basic flow. As the number goes up the complexity goes from basic to univariate and finally to multivariate flows.

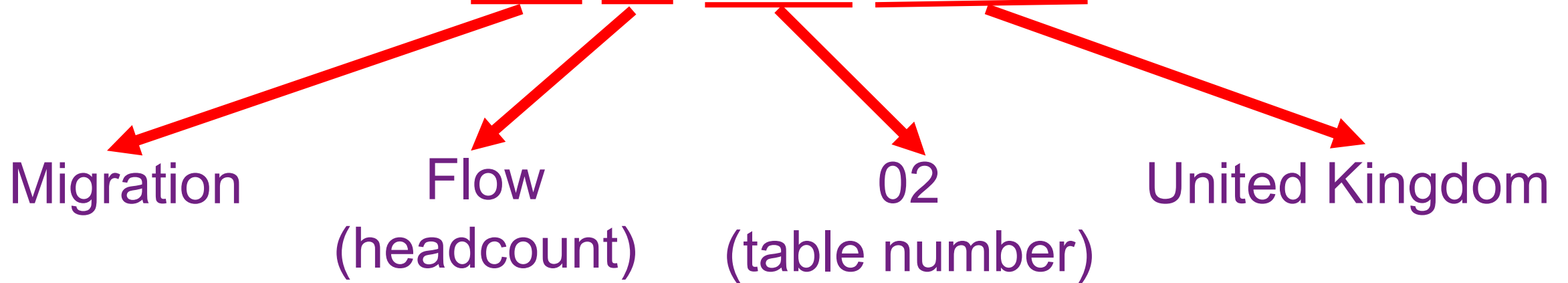
[UK/EW/NI/etc.] – suffix indicating whether table relates to flows in whole of UK, or England and Wales, Northern Ireland, etc.

Currently, only EW and NI flow data is available.

NON_UK: only about the population living outside the UK one year before the census.

2011 Table IDs example

MF02UK



2021 Table IDs example

ODMG01EW

Origin &
Destination

Migration

01

England & Wales

(table number)

01 suggests the table is of
simple headcount complexity



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	

2021 Census flow data

2021 Census flow datasets – E&W

In total, 89 unique datasets have been released via all release channels (74 without the non-UK versions) or 135 tables in total if we count all the geographical variations of the released datasets.

	Public		Safeguarded		Secured	
	Incl. non-UK	Excl. non-UK	Incl. non-UK	Excl. non-UK	Incl. non-UK	Excl. non-UK
Unique datasets	30	25	27	24	32	25
Total datasets incl. geographical variations	66	55	37	33	32	25

2021 intrinsic geographies by release channel

These are the geographic levels used by the ONS in the released datasets:

Open data	MSOA, LTLA, UTLA, Regions
Safeguarded	OA, MSOA, LTLA, UTLA
Secured	OA, MSOA

2021 geographies supported by the new platform

Output Areas	NHS Locations
LSOA	NHS Integrated Care Boards (England only)
MSOA	NHS Regions (England only)
Census Wards	Local Enterprise Partnerships (England only)
Electoral Wards	Regions
Civil Parishes	Inner or Outer London (London areas only)
Westminster Parliamentary Constituencies	Countries
Low-Tier Local Authorities	England and Wales
Local Authority Districts	
Upper-Tier Local Authorities	

2021 topics by release channel (E&W) - Open

Univariate topics	Multivariate combination of topics
<ul style="list-style-type: none">▪ Age (SMS, SSS, SRS)▪ Ethnic Group (SMS, SSS)▪ NS-SEC* (SMS, SSS)▪ Household tenure (SMS)▪ Intention to stay (SMS)▪ General health (SMS)▪ Place of work indicator (SWS)▪ Passport (SSS)	<ul style="list-style-type: none">▪ Age by sex (SMS)▪ Sex by place of work indicator (SWS)▪ Age by place of work indicator (SWS)▪ Social grade by place of work indicator (SWS)▪ Ethnic group by place of work indicator (SWS)▪ Hours worked by place of work indicator (SWS)▪ Passport by place of work indicator (SWS)▪ Number of cars by place of work indicator (SWS)

* National Statistics Socio-Economic Classification

2021 topics by release channel (E&W) - Safeguarded

Univariate topics	Multivariate combination of topics
<ul style="list-style-type: none">▪ Economic activity (SMS, SSS)▪ Industry (SMS)▪ Transport to workplace (SWS)▪ Second address type priority (SRS)▪ Ethnic group (SRS)▪ Passport (SRS)▪ Economic activity (SRS)▪ NS-SEC (SRS)	<ul style="list-style-type: none">▪ Age by sex (SMS)▪ Passport by sex (SMS)▪ NS-SEC by sex (SMS)▪ Occupation by place of work indicator (SWS)▪ Industry by place of work indicator (SWS)▪ NS-SEC by place of work indicator (SWS)▪ Ethnic group by place of work indicator (SWS)▪ Family status by place of work indicator (SWS)▪ Transport to workplace by age (SWS)▪ Age by sex by full time education (SSS)▪ Second address type priority by household tenure (SRS)

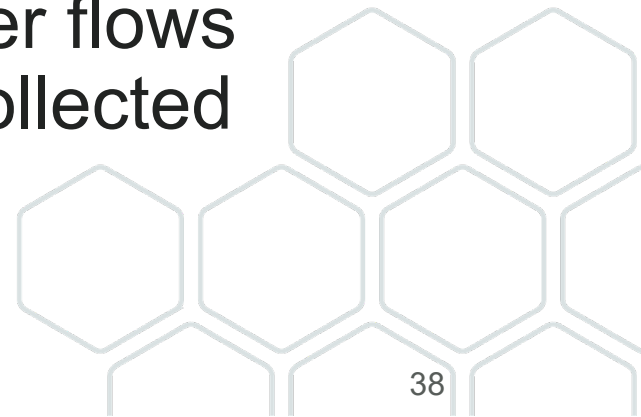
2021 topics by release channel (E&W) - Secure

Univariate topics	Multivariate combination of topics
<ul style="list-style-type: none">▪ Age (SMS)▪ Transport to workplace (SWS)	<ul style="list-style-type: none">▪ Passport by sex (SMS)▪ Disability by residence type (SMS)▪ Economic activity by age by sex (SMS, SWS)▪ General health by age (SMS)▪ Industry by sex (SMS)▪ Occupation by sex (SMS)▪ Ethnic group by sex (SMS)▪ Indention to stay by age (SMS)▪ Transport to workplace by sex by economic activity (SWS)▪ NS-SEC by sex by economic activity (SWS)▪ Industry by sex by economic activity (SWS)▪ Occupation by sex by economic activity (SWS)▪ Social grade by sex by economic activity (SWS)▪ Ethnic group by sex by economic activity (SWS)▪ Passport by sex by economic activity (SWS)▪ NS-SEC by sex by full time education (SSS)▪ Ethnic group by sex by full time education (SSS)▪ Passport by sex by full time education (SSS)▪ Second address type priority by sex by age (SRS)▪ Second address type priority by household tenure (SRS)▪ Second address type priority by family status (SRS)▪ Second address type priority by ethnic group (SRS)▪ Second address type priority by passport (SRS)▪ Second address type priority by economic activity (SRS)

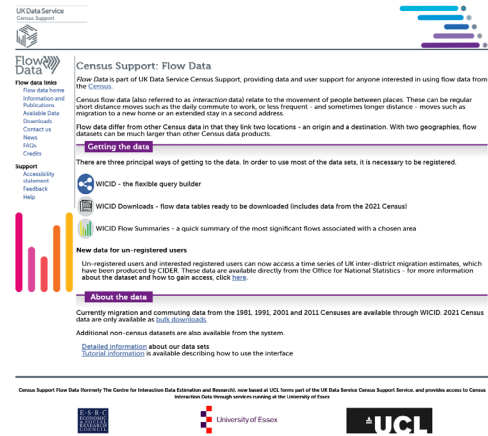
The future

2021 Census and beyond

- Most likely, 2021/22 will be the last traditional census as we know it.
- 2021 Census flow data for England and Wales were released in late October 2023.
- 2021 Census flow data for Northern Ireland were released in July 2024 (public datasets only).
- 2021 Census flow data for Scotland are expected in 2025.
- There are UK harmonisation issues for cross-border flows because the 2022 census data in Scotland were collected one year later than the rest of the UK.

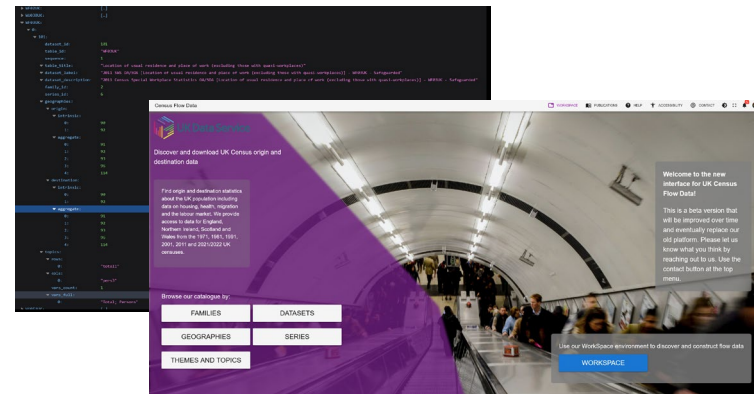


The future of UKDS census flow data platforms



WICID remains the only platform that enables downloading and subsetting of UK census flow data from 1981 to 2021 (2021 data are only available as bulk downloads via WICID).

Within this year, the new API-driven platform will be launched, with some features added gradually over time. This platform will be the only one to support full subsetting capabilities for all 1981 to 2021 census flow data.



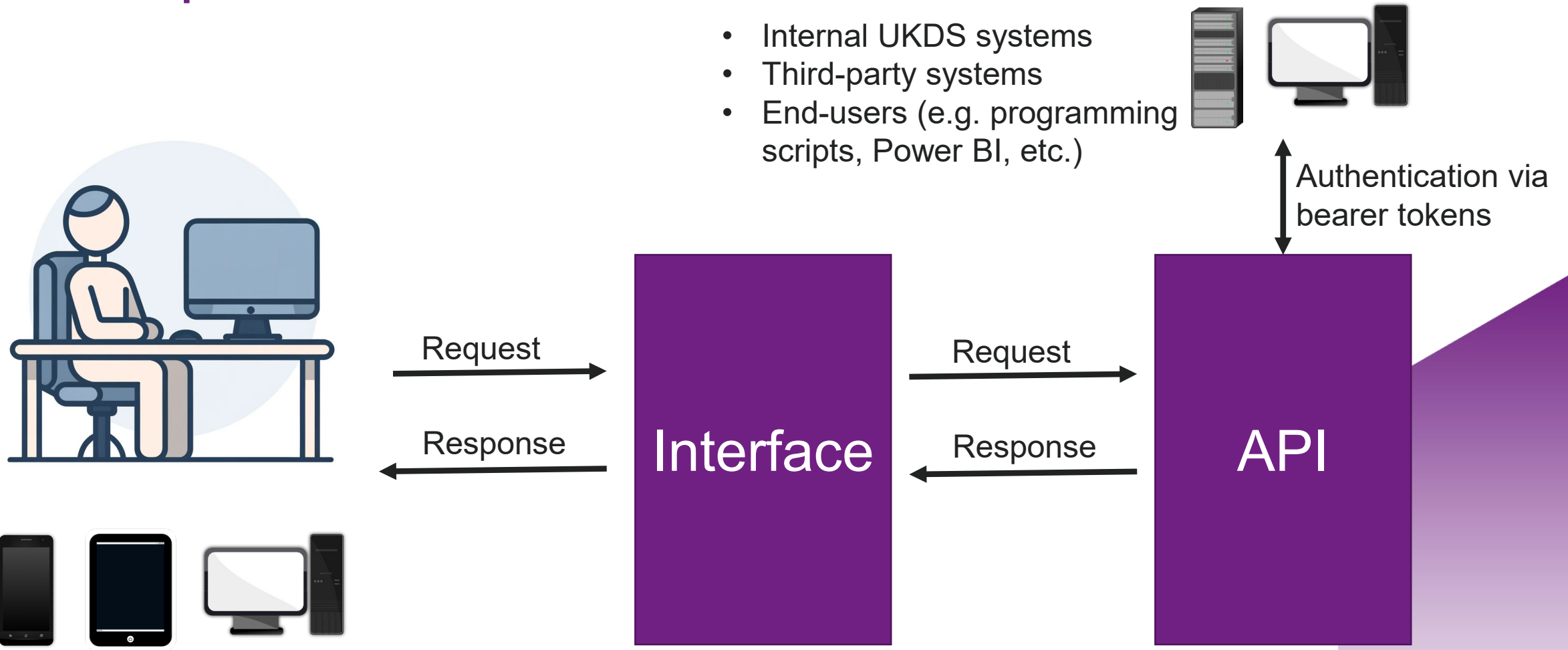
WICID will continue to operate in maintenance mode alongside the new platform for a transitional period, after which it will be retired.

API features

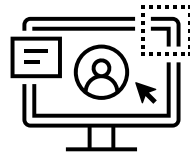
- Provide access to all census data from 1981 to 2021, as well as other datasets such as school census data, etc.
- Enable advanced data discoverability.
- Allow maximum flexibility in selecting tables, geographies, variables, etc.
- Enable full geography aggregation, including areas within compatible higher-level selected areas.
- Offer advanced features, such as compiling outputs from different tables on shared geographic areas.

New platform architecture

- Internal UKDS systems
- Third-party systems
- End-users (e.g. programming scripts, Power BI, etc.)



Express your interest in two user studies on census flow data



<https://forms.office.com/e/LPsEUejugj>