

UK Biobank

Home locations: Census and administrative areas

Version 1.0

<http://www.ukbiobank.ac.uk/>

June 2023



This manual details the procedure for deriving the census and administrative area location data.

1. Overview

Each participant's home location at the time of their baseline assessment has been mapped to a UK Census "Output Area", a small geographic region defined in the [England and Wales Census](#) or the [Scotland Census](#). Output areas were mapped using 1-metre grid coordinates for home locations at assessment (for more information on the derivation of the coordinate data, please see [Resource 118122](#)).

The output area boundaries are modified slightly at each census and we therefore publish both the 2001 census output area and the 2011 census output area for each participant, both based on the same baseline address location.

Larger aggregated regions are also defined for both census providers. At each level these are always made up of groups of regions from the next level down – no single output area may be split between two Super Output Areas (England and Wales) or two Data Zones (Scotland).

All of these fields are published in [Category 100024](#) and are subject to access restrictions which are described in the description of [Category 703](#), which contains the less-restricted fields. The output areas themselves are not normally made available to researchers, unless absolutely necessary for their research.

This document provides further information about the census and administrative area data.

2. Differences in geographies between countries

The fields have been created for participants across England, Wales and Scotland. However, Scotland defines census and administrative areas differently from England and Wales. As well as differences in nomenclature, the populations used are smaller in Scotland than in the rest of the UK; Scottish output areas for example have a minimum population of 50 persons (20 households) compared to 100 persons (40 households) in England and Wales.

A full field list with the roughly equivalent areas merged together in this dataset is as follows:

Field id	Census	England and Wales	Scotland
20269	2001, original code format	Output Area (OA)	Output Area (OA)
20270	2001, recoded		
20273	2011		
20271	2001	Lower layer Super Output Area (LSOA)	Data Zone (DZ)
20274	2011		
20272	2001	Middle layer Super Output Area (MSOA)	Intermediate Zone (IZ)
20275	2011		
20276	2011	Local Authority District	Council Area

More information on the differences between census geographies, including population definitions, is available from the [National Records of Scotland](#).

3. Data sources

Data were gathered from a number of different government sources to create the coded fields. A summary of the sources used is below.

Country	Year	Description	Source
England/Wales	2001	Output Area shapefile	ONS Open Geography Portal
		Mapping from 2001 OAs to original coding of OAs, LSOAs and MSOAs	ONS Open Geography Portal
	2011	Output Area shapefile	ONS Open Geography Portal
		Mapping from 2011 OAs to LSOAs, MSOAs, and Local Authorities	data.gov.uk
Scotland	2001	Output Area shapefile; lookup for original coding of OAs	National Records of Scotland (boundaries)
		Mapping from 2001 OAs to DZs	National Records of Scotland (indexes)
		IZ shapefile	SpatialData.gov.scot Metadata Portal
	2011	Output Area shapefile	National Records of Scotland (boundaries)
		Mapping from 2011 OAs to DZs, IZs and Council Areas	National Records of Scotland (indexes)

4. Additional information

Codes for 2001 output areas were remapped to new standardised 9-character codes, consistent with all other census geographies. The older codes are made available in case they are of use with older datasets.