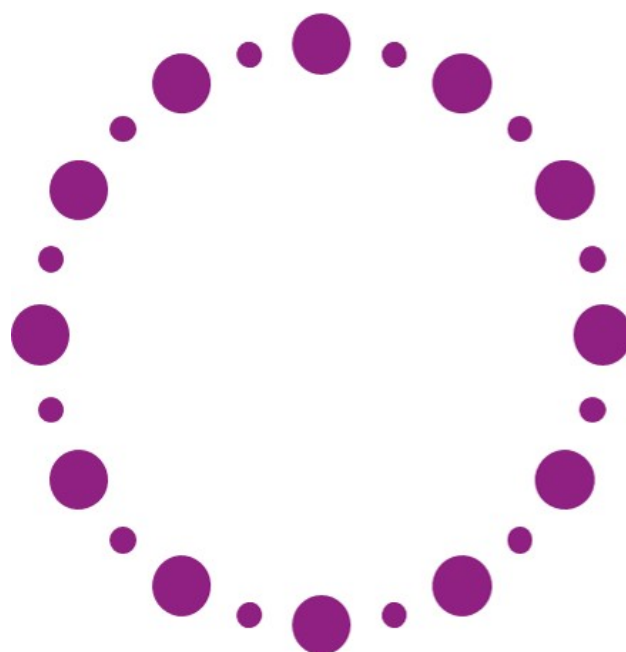


census
2021

Census 2021

Geography Guidance

18 December 2024



Contents

1	Purpose	3
2	Background	3
3	Disclosure Considerations for multiple geographies	4
4	Approximating geographies	5
4.1	Geographies available using Data Zone approximation	6
4.2	Geographies available using Grid Square approximation	7
5	User defined geographies or geographies not yet assessed	8
5.1	Geographies deemed not suitable for approximation	8
6	Census 2011 data on new DZ and SDZ geographies	9

1 Purpose

The purpose of this paper is to:

1. Provide users with background on the Census 2021 output geographies;
2. Outline the disclosure considerations with having multiple geographies and the impacts of methodologies used;
3. Highlight the geography approximation methodologies that Census Office has adopted;
4. Outline geographies in and out of scope for approximation; and
5. Inform users of their options regarding custom geography creation.

2 Background

Two new statistical geographies, namely Data Zone (DZ) and Super Data Zone (SDZ), have been developed to support dissemination of Census 2021 statistics. In February 2023, digital boundaries for the new geographies were published alongside [information papers](#) on their development and how particular administrative geographies can be approximated from Data Zones (with a 2024 Parliamentary Constituency approximation added in September 2023). The new geographies were also added to the January 2023 NISRA Central Postcode Directory released in February 2023.

Person and household estimates from Census 2021 were also published with the release of the new geographies. Subsequently, census data covering all topics were made available for these new geographies via the launch of the [Flexible Table Builder](#) in June 2023, with topics added to the [Area Explorer](#) for the new geographies in July 2023.

The new statistical geographies and their benefits were highlighted in the [Census 2021 outputs and dissemination webinars](#) held in September 2023 (along with Census 2021 outputs released to date and demonstrations of the Area Explorer and Flexible Table Builder dissemination tools that allow the census data to be easily explored).

In addition to the Data Zone and Super Data Zone geography hierarchy, a Census 2021 Grid Square product was released in December 2023, to support the examination of trends over time for consistent spatial units.

Census Office has published Grid Square statistics for the last five censuses (since 1971). The geographic base for Grid Square outputs is the Irish Grid, and census statistics are produced for 100m and 1km grids in Northern Ireland. Grid square census statistics include population and household counts along with a range of mainly univariate outputs (to ensure the confidentiality of individual census returns). Comparability of previous grid square outputs is retained where possible; however, changes over time in definitions and questions asked in the census result in some comparability limitations. Further information on grid square statistics is available at [Census 2021 Grid Square product for Northern Ireland](#) release page.

The Data Zone and Grid Square geographies can both be used as base geographies (or building blocks) to build or approximate non-standard and custom geographies, for which Census 2021 statistics can then be produced.

3 Disclosure Considerations for multiple geographies

The following considerations are relevant when more than one geography is used to disseminate census statistics:

- There is a risk of disclosure by differencing geographic layers when producing any census data for multiple geographies. Essentially what this means is that users can overlay the different layers and subtract the data to reveal statistics for smaller, unintended units of geography.
- Statistical Disclosure Control (SDC) methodologies applied to census data use the output geography as a parameter (e.g. for Census 2021 the geography parameter was Data Zones; for Census 2011 the geography parameter was Small Areas). Since the application of these methodologies are geography specific, there is a risk that the protection they give is undone when custom, or non-related geographies are used to generate the statistics.

During the development of the Census 2021 Grid Square product, a disclosure assessment against Data Zones was completed by Census Office, to ensure that the two geographies are compatible and do not undo the disclosure control methodologies applied to census data. A determination was made that they are compatible, but only when the Grid Square product is limited to a series of pre-determined variables and includes suppression of grid

squares that carry a disclosure risk, see the 'Grid Square product guidance note' on the [Census 2021 Grid Square product for Northern Ireland](#) release page.

Further information on the Census 2021 SDC methodologies, Cell Key Perturbation and Targeted Record Swapping is available in the guidance note 'Statistical disclosure control methodology for 2021 Census' on the [Census 2021 outputs prospectus](#) webpage.

4 Approximating geographies

Census Office notes that there is a user demand for Census 2021 data for additional geographies outside of the new Data Zone hierarchy, and indeed for custom geographies. To date, Census 2021 statistics have been released for a high proportion of non-standard geographies using approximations involving Data Zones or Grid Squares as a base – these are referenced in the sections below, along with information for users of custom and specific geographies which Census Office consider unsuitable for Data Zone or Grid Square approximation.

Census Office has taken this approach to approximating a range of geographies particularly in acknowledgement of the creation of the new statistical geographies and the impact this can have on time series analysis – but also to mitigate the potential risk of users aggregating “noise” by creating figures for higher level geographies using published outputs for a lower-level geography. This risk arises as all Census 2021 statistics released are subjected to a post-tabular SDC method called Cell Key Perturbation. This approach adds 'noise' to the data to protect against the disclosure of information on individuals, households, or groups. The method involves making small changes (perturbations) to cells in output tables introducing uncertainty in the data (noise), the probability of perturbation is higher with smaller cell counts in particular with sparse data (for example, small geographic areas).

The advice for users in dealing with this noise is to either request the production of a custom geography directly from Census Office (see section 5 for further details) or, when using the Flexible Table Builder system to construct the table that you require directly, rather than adding up cells from different tables at different levels. If that is not possible, the advice is to create totals summed from the fewest possible cells to minimise the effect of perturbation.

4.1 Geographies available using Data Zone approximation

Geographies are considered suitable for Data Zone approximation if they were within a five per cent accuracy threshold for both the usual resident person and household counts from Census 2021 (see section 3 of ‘Aggregating Data Zones to produce statistics for higher-level geographies’ on the [information papers](#) webpage).

The list below shows those geographies deemed suitable for Data Zone approximation (note: Super Data Zones, 2014 District Electoral Areas and 2014 Local Government Districts are exact aggregations of Data Zones, rather than approximations):

- Super Data Zones
- 2014 District Electoral Area
- 2014 Local Government District
- County
- 2008 Parliamentary Constituency/Assembly Area
- 2024 Parliamentary Constituency
- Health and Social Care Trust
- 1992 Local Government District
- 2015 Settlement – Those within Bands A to D, 14 of the 17 in Band E (Small town) and 20 of the 24 in Band F (Intermediate settlement)
- Census 2011 Travel To Work Areas

Each of the geographies listed above are available for selection in the [Flexible Table Builder](#), which allows users to build their own table or search through and customise ready-made tables.

An additional Urban Status classification geography is also available in the Flexible Table Builder. The Urban Status was derived by assigning Data Zones as either urban or non-urban. Data Zones with 90% or more of their usual resident population inside the boundary of an urban settlement (i.e. those settlements with population 5,000+ usual residents) are classed as urban. All remaining Data Zones are classed as non-urban.

The [Census Area Explorer](#) application allows users to explore census results at a range of geographies: Northern Ireland, 2014 Local Government District, 2014 District Electoral Area, Super Data Zones, and Data Zones

A range of Main Statistics tables have also been published for Northern Ireland and 2014 Local Government Districts. Some also include 2014 District Electoral Area, Super Data Zones, Data Zones, 2015 Settlements and 2014 Electoral Wards. These are available from the [Census 2021 Main Statistics](#) section of the NISRA website. Tables available can also be examined using the [Census 2021 Table lookup](#).

4.2 Geographies available using Grid Square approximation

The development of the Census 2021 Grid Square product included an assessment of the accuracy of grid square based approximations for geographies that were unsuitable for Data Zone approximation. The outcomes of this were that Census Office published Census 2021 statistics for the smaller 2015 Settlements, 2014 Electoral Wards and Neighbourhood Renewal Areas using grid square approximation. The approach taken was to publish statistics for each approximated geography, rather than produce a grid square lookup. The reasons for this were to protect against disclosure by differencing and to minimise the effect of Cell Key Perturbation. Requests for additional tables will be assessed on a case-by-case basis, subject to disclosure and utility assessment.

- **2015 Settlements** – Census Office used the Settlement Development Limits from the [2015 Review of the Statistical Classification and Delineation of Settlements](#) to define the spatial extent of settlements and produce associated Census 2021 statistics. The larger settlements can be accurately approximated from Data Zones, while grid square approximation is used for the smaller settlements that could not be approximated sufficiently accurately using Data Zones. A range of Main Statistics tables for all settlements with a Census 2021 usual resident population of at least 500 were published on 30 November 2023, as part of the [Census 2021 main statistics for settlements and wards in Northern Ireland](#) release. Subsequently, exact counts of usual resident population and households for all settlements containing at least 50 usual residents or 20 households were published on 18 December 2023, in the [Census 2021 person and household estimates for settlements](#) commissioned table.
- **2014 Electoral Wards** – Grid Square approximation meets the aforementioned accuracy threshold for the 2014 Electoral Wards. This facilitated the publication of a range of Main Statistics tables for the 2014 Electoral Wards on 30 November 2023,

as part of the [Census 2021 main statistics for settlements and wards in Northern Ireland](#) release.

- **Neighbourhood Renewal Areas** - Grid Square approximation meets the aforementioned accuracy threshold for Neighbourhood Renewal Areas. This facilitated the release of a range of Census 2021 Neighbourhood Renewal Area commissioned tables on 12 February 2024, from [CT0066 Usual resident population](#) to [CT0093 Household deprivation \(housing dimension\)](#). Tables for Neighbourhood Renewal Areas can also be examined by searching for 'NRA' as the 'geography levels' in the [Census 2021 Table lookup](#).

5 User defined geographies or geographies not yet assessed

Requests for Census 2021 statistics for non-standard or user defined geographies that have not yet been assessed, can be sent into Census Customer Services for consideration (census@NISRA.gov.uk). Each request will be considered on a case-by-case basis as part of the census commissioned table service, with suitability for production subject to disclosure, accuracy and utility assessments.

If a non-standard geography is deemed suitable to produce Census 2021 statistics for, those statistics can only be approximated using the available "building blocks" of either Data Zones or Grid Squares. **No statistics from Census 2021 can be produced for non-standard geographies on an exact basis.**

Additional information on previously published Census 2021 commissioned tables and a link to a [Census 2021 commissioned table lookup](#) are available on the [Commissioned outputs](#) section of the NISRA website.

5.1 Geographies deemed not suitable for approximation

As this paper outlines, Census Office has already carried out a number of assessments on the suitability of producing Census 2021 statistics for non-standard geographies.

As part of those assessments Census office has determined that neither Data Zone nor Grid Square approximation are sufficiently accurate for the production of Census 2021 statistics for the former 2011 Small Area and 2011 Super Output Area geographies.

6 Census 2011 data on new DZ and SDZ geographies

Users are encouraged to adopt the new statistical geographies developed for Census 2021, but Census Office recognises that more general users may require Census 2011 data for the new geographies, to assess or show change over time.

On 17 December 2024, Census Office released exact Census 2011 household and usual resident counts for the Data Zone and Super Data Zone geographies as commissioned tables, to allow users to compare the population in the new geographies over time. Census 2011 estimates that show the number of households in Northern Ireland are available from [CT0609NI Households](#) and Census 2011 estimates that show the number of usual residents in Northern Ireland are available from [CT0610NI Usual resident population](#).

Other requirements will be assessed on case by case basis, subject to disclosure and utility assessments.