Quality and methods guide for census-based statistics UK: 2021

What the statistics cover, how we produced them, their strengths and limitations and potential uses of the data. Includes guidance on alternative approaches to combining and comparing data from the most recent UK censuses.

Contact:
Census Customer Services
census.customer.services@ons.gov.uk
+44 1329 444972

Release date:
27 June 2025

Next release:
To be announced

Table of contents

1 . Overview

UK censuses have previously been conducted at the same time in all four countries of the UK, with results brought together to produce UK-wide statistics. That was also the intention for the census of March 2021, but Scotland's census was delayed by a year because of the coronavirus (COVID-19) pandemic. This difference in the timing of data collection made the production of UK results more challenging than usual.

The three agencies that run the censuses - the Office for National Statistics (ONS, covering England and Wales), National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) - have worked together to develop a solution that is conceptually simple, pragmatic, and creates numbers for a consistent time point.

Our method has been to adjust Scotland's 2022 Census data to create statistics for Scotland for 2021, adding them to equivalent Census 2021 totals from the other UK countries. We have used this approach to create tables for a range of census variables.

This guide includes information on the methods we have used and more information on the strengths and limitations of resulting statistics. It also includes advice on alternative approaches to combining and comparing cross-UK statistics.

2 . Where the data comes from and how we turned the data into statistics

What the statistics cover

We have published a workbook containing 16 tables, covering a range of social and demographic topics. Of these, 10 tables contain estimates based on the number of usual residents, and 6 contain estimates based on the number of households.

Because our production of March 2021 values for Scotland required the presence of comparable data from both the 2011 and 2022 censuses, our scope was limited to variables that:

- were included in the Scottish censuses in both 2011 and 2022, using questions and output categories that were either identical or sufficiently similar
- were collected and output in an identical or sufficiently similar way in the 2021 censuses in England and Wales, and Northern Ireland, so enabling the production of coherent UK totals
- were not considered to have been liable to have been affected by the coronavirus (COVID-19) pandemic to an extent that would have made interpolated 2021 values for Scotland unreliable

A list of tables is available on the contents page of the data file. All tables include March 2021 values for the following geographic areas:

- UK
- Great Britain
- England and Wales (combined)
- England
- Wales
- Northern Ireland
- Scotland

We have not published equivalent values for any subdivisions of the UK nations. We explain our reasons and provide advice on this in the Approaches for alternative data needs section.

It is important to note that our Scotland values for 2021 do not replace NRS's Census 2022 results as the definitive census statistics for Scotland.

Where the data comes from

The data sources we used to produce the tables were:

- Census 2021, England and Wales
- Census 2021, Northern Ireland
- Census 2011, Scotland
- Census 2022, Scotland
- Mid-year population estimates for Scotland, 2020 and 2021, taken from Rebased population estimates, Scotland, mid-2011 to mid-2021
- Household estimates for Scotland, 2020 and 2021 (taken from Households and Dwellings in Scotland, 2023)

All the source datasets, for all variables, were already publicly available.

How we processed the data and produced the statistics

Our method for creating UK population estimates for March 2021 had four main stages:

1) simple linear interpolation between the 2011 and 2022 census results for Scotland to create initial March 2021 values for all variables, apart from population estimates by age and sex. We did this at two levels:

a) creating interpolated values for each individual cell in each table. For example, number of females aged 40 to 44 in each ethnic group

b) creating interpolated values for "column totals" - using the same example, this would mean the number of people in each ethnic group, not broken down by age and sex. We calculated this separately, because disclosure control applied to the published Census 2022 values mean they may not exactly sum to column totals

2) simple linear interpolation between the mid-2020 and mid-2021 population estimates for Scotland to create March 2021 population estimates by age and sex. Also, an equivalent process between the mid-2020 and mid-2021 household estimates for Scotland to create March 2021 estimates of the number of households

3) for person-level tables, iterative proportional fitting to constrain the March 2021 values created in stage 1a, to both the column totals created in stage 1b and the March 2021 population estimates by age and sex created in stage 2. For household-level tables, simple constraining to the March 2021 household estimates calculated in stage 2

4) creating Great Britain and UK estimates by adding the March 2021 estimates for Scotland, created in stage 3, to the Census 2021 estimates for England, Wales and Northern Ireland

Here is a more detailed description of each method:

Simple linear interpolation (stages 1 and 2)

Simple linear interpolation refers to the estimation of intermediate values between two known values, based on the assumption of a straight-line trend between them.

Stage 1

The respective census days in Scotland were 27 March 2011 and 20 March 2022. This means there were 11 years and 1 week between the two censuses. However, given the uncertainties inherent in estimation, for our purposes we assumed that they were exactly 11 years apart.

This means that our initial interpolated March 2021 values calculated in stage 1 were exactly 10/11 of the way between the 2011 and 2022 census results.

Stage 2

While linear interpolation assumes a linear change over time, actual change is unlikely to occur so consistently.

An example is that the size of Scotland's population in March 2021 calculated by simple linear interpolation between 2011 and 2022 is greater than it was in the official mid-2021 population estimate, despite Scotland's population having grown between the mid-2020 and mid-2021 estimates.

Although the impact in this example is small in percentage terms, this would mean that all our person-based tables for March 2021 would have population totals that were out of line with the intercensal series of population estimates.

To mitigate this, we created estimates of Scotland's population (by sex and age) and number of households in March 2021, using them in stage 3 to adjust the initial interpolated estimates calculated in stage 1 to better reflect the likely actual size of Scotland's population at the time.

Mid-year estimates relate to the end of June each year. The figures we calculated for March 2021 were therefore three quarters of the way between the mid-2020 and mid-2021 estimates of population or households.

Iterative proportional fitting (Stage 3)

For the person-level tables, we needed to adjust the initial March 2021 values calculated in stage 1a to fit the population estimates by sex and age calculated in stage 2.

However, if we simply scaled all the values from Stage 1 to the March 2021 population estimates, we would be liable to get different totals for each variable category (such as ethnic group) depending on which sex and age grouping we used, because of the way they act as weights.

To overcome this, we applied iterative proportional fitting (IPF). This adjusted values in each table in an iterative manner to bring them into better alignment with pre-specified totals for both row and columns. In this case, our respective totals were:

- the March 2021 population estimates by sex and age
- the March 2021 variable category totals created in stage 1b

For household tables, IPF was not needed. This is because our March 2021 estimate of households calculated in stage 2 was a single number, meaning that the impact of choice of sex and age groupings did not apply.

Instead, we multiplied all values in the household tables by a single ratio to ensure they summed to the estimated Scotland household total for March 2021, as calculated in stage 2.

Creating Great Britain and UK totals

The stages above created Scotland totals for March 2021, corresponding with Census 2021 in England, Wales and Northern Ireland.

In the final stage we rounded the values we had calculated for Scottish variables to integers (whole numbers). We then put the final Scottish values into cross-UK tables that also contained the existing published 2021 values for England and Wales (both separately and combined) and Northern Ireland.

To create the Great Britain totals, we added the Scotland values to the combined "England and Wales" values. We used the combined England and Wales values, as opposed to the individual values for each country, because of the possibility that the original individual values may not sum exactly because of disclosure control applied to the original published outputs. This is in line with the recommended best practice.

To create the UK totals, we added the Northern Ireland values to the Great Britain values.

3 . Quality of the statistics

Accreditation status

All source datasets used to create this publication have Accredited Official Statistics status. This means that they have been independently reviewed by the Office for Statistics Regulation and found to comply with the standards of trustworthiness, quality and value in the Code of Practice for Statistics.

This means that the actual Census 2021 values in our tables also have Accredited Official Statistics status. This covers the figures for England and Wales (both separately and combined) and for Northern Ireland.

However, the figures for Scotland, Great Britain and the UK have just Official Statistics status, without the accreditation. We have agreed this labelling with the Office for Statistics Regulation (OSR), in recognition that the use of interpolation to create Scottish data for 2021, while pragmatic, necessarily involves greater uncertainty than had there been an actual data collection at that time.

Related, it is important to emphasise that the Scotland, Great Britain and UK statistics in this publication, although derived from census statistics, are not actual census statistics. Instead, they offer an estimate of what values may have been produced had the census taken place in March 2021 across the whole UK. The statistics from Census 2022, as published by National Records of Scotland (NRS), remain the definitive census statistics for Scotland.

4 . Strengths and limitations

This section describes the strengths and limitations of our methodological approach to creating UK statistics for 2021.

There are also various quality and comparability points that are specific to individual variables. We have presented those in the data file.

Strengths

- our statistics present UK values for a single time point (March 2021)
- all our source data are Accredited Official Statistics, indicating that they are high quality
- the censuses for each part of the UK are the largest, most detailed statistical surveys undertaken in each part of the UK
- our use of mid-year estimates to constrain the Scotland data for 2021 means we are taking account of the best available estimates of the numbers of people and households for that year

Limitations

- the census source data, although very high quality, are estimates rather than definite counts; this means they have a degree of statistical uncertainty around them
- even if the census results and mid-year estimates are exactly correct, it is unlikely that our methods for creating the 2021 estimates for Scotland would be an exact representation of actual figures in March 2021 because:
- (a) the numbers of people and households may not have changed linearly between the mid-2020 and mid-2021 estimates
- (b) multiple factors beyond population change may have influenced trends over time in the number of people or households in different categories; this could include a range of demographic, social, economic, political and other factors, including the COVID-19 pandemic

How we quality assured the data and statistics

We developed our methods in collaboration with ONS's Methodology team. Our process was iterative, meaning we tested the proposals at each stage and refined them as needed. We also discussed the method with other ONS experts.

In creating the statistics, we checked that the correct source data had been used for each table, the calculations had been carried out correctly, and that the results were plausible in relation to the respective source data. Each table was checked by two reviewers (other than the original creator) at ONS, and by NISRA and NRS.

Source data quality information

As indicated, our source data are all Accredited Official Statistics. The respective censuses have also all been published with supporting quality information, available via the following links:

Quality information for Census 2021 (England and Wales)

Quality assurance (Northern Ireland Census 2021)

Quality assurance report (Scotland Census 2022)

There is a "Strengths and limitations" section in National Records of Scotland's "Mid-2023 population estimates". This in turn links to a page on Quality Assurance that contains more information on the quality of the data sources used in the estimates.

5 . What these statistics can be used for

Census data help a wide range of organisations understand the present and plan for the future. This includes the public sector, private sector, third sector and academia. There is also interest from individuals.

Data from Census 2021 (England, Wales and Northern Ireland) and Census 2022 (Scotland) provide detailed information for geographies down to very local areas. UK totals, meanwhile, are useful as a "whole UK" picture, and to allow comparison of local areas with the national average.

The 2021 Scotland, Great Britain and UK totals we have provided are not actual census estimates but can instead be considered a proxy, serving users who wish to have data across the UK for a single time point.

6 . Approaches for alternative data needs

If users do not require estimates for a single time point, an alternative for creating cross-UK statistics is to aggregate the values from the respective censuses. This approach would mean the totals directly reflect 2022 (Scotland) and 2021 (rest of the UK) census data, and has strengths and limitations:

- it would mean that the statistics for all four countries were Accredited Official Statistics with the associated level of quality
- it is a simpler approach than the method we have developed to create these cross-UK outputs for 2021
- the input data would be from adjacent years, so may not be directly comparable, although for many comparisons (any aspect of population that is not expected to have changed much between 2021 and 2022) this may be acceptable
- migration between countries of the UK will mean that some residents have been included in both censuses, and some in neither

Dependent on their needs for UK or Great Britain totals, or comparisons between countries, users may choose to use the tables in this publication. Alternatively, they may choose to create their own aggregations of published census data as described, with awareness of the respective strengths and limitations.

We have not produced 2021 figures for subnational areas in Scotland. In practice, activities focused on the local level would generally be expected to have a stronger interest in the Census 2022 data for that area. In addition, interpolation is likely to have greater levels of uncertainty for smaller areas, where change over time is more sensitive to local factors and may differ substantially from national trends.

Because of this, we also recommend that anyone wishing to compare local areas in Scotland with local areas in the rest of the UK use the Census 2022 and Census 2021 statistics respectively.

7 . Related links

ONS Census 2021

NISRA Census 2021

NRS Census 2022

Census-based statistics UK: 2021
Released 27 June 2025 | Dataset
This dataset includes Accredited Official Statistics for England, Wales, and Northern Ireland. Statistics for Scotland, Great Britain, and the UK are Official Statistics but are not accredited. Data for Scotland here do not replace the official 2022 Census statistics published by the National Records of Scotland.

8 . Cite this page

Office for National Statistics (ONS), 27 June 2025, ONS website, quality and methods guide, Census-based statistics: UK 2021.