

ONS Census Transformation Programme

Consultation

Initial View on 2021 Census
Output Content Design

February 2018

Table of contents

Introduction	2
Consultation - how we'd like you to respond	2
Moderation policy	3
Future use engagement	3
Vision for 2021 Census outputs	5
2021 Census strategic objectives	6
Statement of Agreement between the National Statistician and Registrars General	6
Accessing 2021 Census results	7
Flexible dissemination system	7
Statistical disclosure control	8
UK harmonisation and UK statistics	9
Enhanced Census Outputs	9
Bespoke census results	10
Publication of 2021 Census results	10
2021 Census specialist products	11
Microdata	11
Origin-destination	12
Longitudinal Study	13
Accessing sensitive census data	13
2021 Census results – outputs, population bases and geography	14
Output categories	14
Population bases	14
Small populations	15
Geography	16
Interpreting 2021 Census results	18
Metadata	18
Ensuring high quality results	18
Analysis programme	19
Licensing, funding and charging models	19
International comparability of 2021 Census results	20
Conclusion	21
Annex - Paper questionnaire template	23

Introduction

This document outlines our initial proposed design of 2021 Census outputs and the dissemination channels for England and Wales. We'd like to gather your views on our initial design of 2021 Census outputs to ensure our proposal will meet your needs. Your response to this consultation will be important. This feedback will help us to prioritise our future research and development of 2021 Census outputs. It will also inform the order in which we release our products.

This consultation is our first step towards understanding what you need from 2021 Census outputs, including specialist products. We'll hold a final consultation which outlines our more detailed plans before census day.

Our initial view is based on:

- the evaluation of the design and delivery of 2011 Census output content
- our current understanding of user need and knowledge of potential technological and methodological solutions
- our current understanding of affordability
- our aspirations for linking administrative data with the 2021 Census data to produce enhanced census outputs.

It is important to note this is a proposal at present. The final design of output content will be subject to further research, cost, feasibility, availability of administrative data, testing and feedback from users.

The information presented in this document specifically refers to the design of the census in England and Wales, unless otherwise stated. We are working closely with the National Records of Scotland (NRS) and Northern Ireland Statistics and Research Agency (NISRA) to ensure UK comparability of results, where possible. NRS and NISRA will be consulting with users in their respective countries in due course.

Consultation – how we'd like you to respond

The best way to respond to this consultation is online ¹. You can also respond to the consultation using a paper questionnaire template (see Annex). Please email your response to census.outputs@ons.gov.uk using the subject line "Initial View Consultation".

There are some mandatory questions which we need you to complete in order to include your response in our analysis.

Responses to this consultation document are invited until 23rd May 2018. We'll publish our response to your feedback by Autumn 2018.

More information on how we'll use and evaluate your responses to the consultation is available in the conclusion to this document. This may help you when you're answering the questions.

If you need any more information, please email 2021 Census Outputs and Dissemination at census.outputs@ons.gov.uk.

¹ <https://consultations.ons.gov.uk/census/initial-view-on-the-2021-census-output-design>

Moderation policy

To support transparency in our decision making process, we'll publish all responses to the consultation. This will include the name of the responding organisation or individual, but you're asked to confirm that you're content for your name to be published. We won't publish personal contact details. Please be aware that any information provided in response to this consultation could be made publicly available if requested under a Freedom of Information request.

This gives you the opportunity to post a response that will remain publicly visible on the Office for National Statistics (ONS) website. We'll therefore moderate responses to ensure they're appropriate and not harmful to others. We moderate in accordance with the terms below.

Responses which include any of the following may be deleted. These include:

- threats or incitements to violence
- use of obscenity
- duplicative or substantially duplicative postings by the same person or entity
- postings seeking employment or containing advertisements for a commercial product or service
- information posted in violation of law, including libel, condoning or encouraging illegal activity, revealing classified information, or infringing on a copyright or trademark

Moderators reserve the right to remove responses that don't address some aspect of the consultation. some aspect of that purpose. We value your time and input, and our desire is to remove as few responses as possible while ensuring that a focused, constructive survey takes place.

This moderation policy is subject to change to address additional matters as necessary. If you have a complaint about an item of user-generated content, or feel that your own content was removed in error, please contact us at ons.communications@ons.gsi.gov.uk or call 0845 601 3034.

Future user engagement

This consultation is the first step towards understanding your needs from 2021 Census outputs. We'll actively engage with users on all aspects of census output content and dissemination mechanisms throughout the design phases until the publication of outputs following the 2021 Census. We'll hold a final consultation which outlines our more detailed plans before census day. We'll also ask users to reflect on the utility of 2021 Census data after its release.

We recognise there are many census data users, with different interests and needs. Due to the large number of uses, it won't be feasible for us to consult with all census data users to the same extent. However, we'll invite all potential users to participate in our public consultations. We'll also use the existing Census Advisory Groups (CAGs) to update stakeholders on our progress and invite feedback on our plans. We'll test the functionality of our proposed dissemination mechanism and make decisions informed by extensive user research. We'll also create working groups to ensure

the design of specific output content (including microdata and origin-destination products) meet user needs. If you're interested in being involved in any of these working groups, please let us know in your response to this consultation.

We'll consider views from all users. However, it should be recognised that delivering 2021 Census outputs and dissemination systems which meet the full expectations of such a large and wide user community won't be practical or financially viable. We'll seek to deliver outputs which meet the majority of user needs, whilst working with a range of stakeholders to ensure the design of specialist products meets the needs of specific groups. We'll consider the design of output content based on the following criteria:

- accessibility
- flexibility
- timeliness
- feasibility
- disclosure risk
- relative costs
- user acceptability
- public confidence.

Vision for 2021 Census outputs

We're responsible for the design and delivery of the 2021 Census in England and Wales. The census is vital for shaping policy that will determine the country's future and for ensuring that decision-makers have access to the right information to plan and deliver the services used by us all. For the 2021 Census, it is important the census results highlight long-term trends, but also reflect the society in which we live today. Our 2021 Census Topic Consultation demonstrated to us the value of census data to help users make better decisions in their work².

The design and content of 2021 Census outputs and overall dissemination approach will be driven by statistical and technical methodologies, availability of good quality administrative data and user need.

For our 2021 outputs, we'll aim to maintain or improve on the quality of the outputs we currently produce based on the 2011 Census. We'll also provide comparability where you tell us it is important and deliver information that is relevant to the needs of users in 2021.

The 2011 Census outputs formed the most complex and comprehensive set of information about the population produced to date. The ultimate benefits of the census are only realised when users of census data exploit the published outputs. The investment of time and resources in a census can only be justified when the results are made accessible, and the outputs meet user needs³.

Feedback from users following the 2011 Census specified a need for a greater flexibility in outputs and a need for more timely release of data. We'll disseminate high-quality outputs that comply with the Code of Practice for Official Statistics. This includes making improvements on the 2011 Census outputs delivery in terms of:

1. flexibility – we'll aim to develop a web-based interactive dissemination system where users can specify the data they need and build their own tables
2. timing – we'll aim to disseminate national- and local authority-level results for England and Wales within 12 months of census day and all other results within 24 months of census day
3. accessibility – we'll aim to host the web-based interactive dissemination system through the ONS website meaning the majority of census data will be available from one location and follow Government Digital Service guidelines on accessibility
4. relevance – we'll develop enhanced census outputs by linking administrative data sources with 2021 Census data at the record level for topics where there is an identified user need

We also recognise the importance of statistics for small areas (such as Output Areas, Workplace Zones and parishes) and we'll continue to produce results for these geographies. We'll also produce UK-level outputs, where data collected permits this to be done.

² <https://www.ons.gov.uk/census/censustransformationprogramme/consultations/the2021censusinitialview/oncontentforenglandandwales>

³ <https://www.ons.gov.uk/census/2011census/howourcensusworks/howdidwedoin2011/2011censusgeneralreport>

2021 Census strategic objectives

We'll ensure the production of 2021 Census outputs meets the Census Transformation Programme high-level objectives related to outputs⁴. This includes:

- producing statistics of the right quality and timeliness to meet user needs
- making best use of administrative data sources in producing integrated census outputs, including statistical processing, quality assurance and the creation of new outputs
- ensuring widest possible awareness and use of census outputs
- defining and delivering future systems for disseminating statistical outputs to optimise the accessibility of census outputs
- protecting, and be seen to protect, confidential personal data
- maximising the potential for wider benefits to ONS and the statistical system from the investment in the census
- providing value for money

Statement of Agreement between the National Statistician and Registrars General

The delivery of 2021 Census outputs will adhere to the Statement of Agreement about the conduct of the 2021 Censuses across the UK made by the National Statistician and the Registrars General for Scotland and Northern Ireland⁵.

They have agreed that the three census offices and the Welsh Government will work together and reach mutual agreement on the points of interest in the Statement of Agreement. We will seek harmonisation where it is in the interest of census users and the public good more generally.

The Agreement recognised there is a user demand for consistent, coherent and accessible statistics for the UK, for each country and geographic areas within each country.

The National Statistician and Registrars General have, subject to the need for approval where appropriate by the relevant legislatures, agreed that:

- the 2021 Census outputs should constitute consistent, coherent and accessible statistics for the UK, individual countries and geographic areas within each country
- common definitions and classification, typically based on international standards, should be agreed, used and published [This includes a common population base]
- common topics and questions should be agreed wherever possible, with the intention of making available consistent outputs across the UK

⁴ <https://www.ons.gov.uk/ons/guide-method/census/2021-census/progress-and-development/research-projects/2021-census-design-document.pdf>

⁵ <https://www.ons.gov.uk/census/censustransformationprogramme/progressanddevelopment/legislation>

Accessing 2021 Census results

Flexible dissemination system

We plan to publish census data on our website in a combination of pre-determined univariate tables and a web-based interactive dissemination system where users can specify the data they need. The interactive dissemination system will enable users to define their own queries, selecting the geography, population base and variables they need. Users will be able to produce tables that will be comparable with 2011 Census.

This would mean the census data are:

1. flexible – users will be able to define their own queries
2. timely – outputs will be available sooner than 2011 Census
3. accessible – we'll host the flexible dissemination system through the ONS website meaning the majority of census data will be available from one location and follow Government Digital Service guidelines on accessibility

These are significant advantages that are intended to enable users to meet earlier deadlines, allow greater time for analysis, and deliver data that are more specific to their needs. The flexible dissemination system will also provide the same level of detail as 2011 Census.

Through this consultation we seek to confirm the appetite amongst users for this type of system and explain the necessary constraints to prevent statistical disclosure with a new dynamic statistical disclosure control method. Further information on this methodology is provided in the section below. In the proposed flexible dissemination system:

- repeating the same query for the same area will always result in the same results
- totals and sub totals for different queries for the same area may be slightly inconsistent, but the differences will be very small
- the sum of estimates for small areas that constitute a larger area, for example all Output Areas in a Super Output Area, won't exactly match data in an identical table for the larger area, but the differences will be very small

We're currently prototyping the flexible dissemination system using complex record-level synthetic data. We're also ensuring the ONS technical infrastructure will support this method of dissemination.

We've created these slides ⁶ to help your interpretation of the design of the flexible dissemination system.

We recognise the requirement of some users for high volumes of aggregate data (especially by those that add value and redistribute the data) and for queries to be submitted by systems outside of ONS. To meet these needs we're working towards providing an application programming

⁶ <https://consultations.ons.gov.uk/census/initial-view-on-2021-census-output-content-design>

interface (API) to allow access to the census data via programmable requests as well as through the ONS website. The range of data available through such an API would be the same as that provided through the flexible dissemination system. The solution would need to be capable of managing very high volumes of data and queries from external sources without adversely affecting users of the flexible system browser.

Our primary focus for research and development is the creation of a flexible dissemination system to produce 2021 Census outputs. However, if our work identifies it isn't possible to disseminate census data through this means, we'll consider an alternative method to produce outputs.

Statistical disclosure control

Statistical disclosure control (SDC) methods ensure that statistical outputs provide as much value and utility to the users while protecting the confidentiality of information about individuals, households and organisations. ONS has a legal obligation under the Statistics and Registration Service Act 2007, and the Data Protection Act 1998 (to be superseded by the General Data Protection Regulation in May 2018), to protect the confidentiality of census data.

In 2021, we plan to use a targeted record swapping methodology for individuals and households with rare or unique characteristics during the data processing stages. This is the same type of methodology used in the 2011 Census. Every individual and household was assessed for uniqueness based on a small number of characteristics. The individuals or households with rare or unique characteristics were more likely to be swapped with another "similar" household from another area, though every household had a chance of being swapped.

Once the 2021 Census data have been processed, they will be added into a final record database held securely by ONS. The aim is that users will then be able to query the final record database using the flexible dissemination tool and select the population base, geography and variables required. The flexible dissemination tool will then apply a cell-key perturbation method to ensure the table generated will satisfy legal requirements on protecting confidentiality. This method will need to ensure:

- there has been sufficient uncertainty created systematically about whether or not the small cell is a true value
- creating uncertainty doesn't significantly damage the utility of the data

The statistical disclosure methodology will be rigorously tested, evaluated and agreed in advance of the census. We'll work with National Records of Scotland (NRS) and Northern Ireland Statistics and Research Agency (NISRA) to aim to have the same methodology adopted across the UK. This is as stipulated by the joint agreement between the National Statistician and Registrars General for Scotland and Northern Ireland. We're aware users need earlier decision making regarding the methodology compared to the 2011 Census, so we'll keep users informed of progress.

UK harmonisation and UK statistics

ONS has a responsibility for disseminating 2021 Census statistics for the UK, where data collected permits this to be done. In 2011, 73 standard tables were produced containing UK-level data.

As stated in the Statement of Agreement for the conduct of the 2021 Censuses across the UK, we'll aim to achieve harmonisation of 2021 Census output content, where possible.

We recognise the user demand for consistent, coherent and accessible statistics for the UK, for each country and geographic areas within each country. ONS, NRS and NISRA have already started to work together to facilitate the harmonisation of output content, where possible, and data dissemination platforms. The statistical disclosure control methodology will be a major focus of harmonisation across the UK for outputs, with focus also given to common population bases, geography, metadata, definitions and output formats. We're interested in identifying the value of UK-level statistics and the detail required.

We'll continue to work with NRS and NISRA to ensure easy and co-ordinated access to statistics provided by the three census offices. We'll investigate the use of an API to access UK data through the website of each country.

Enhanced Census Outputs

In addition to the release of census data, we intend to create "enhanced census outputs". These will be 2021 Census-based outputs where census data and administrative data are linked at the record level to enable the production of multivariate, small area outputs. This is in contrast to the Administrative Data Census project, which is assessing whether censuses after 2021 can be conducted using other sources of data⁷.

We'll only produce enhanced census outputs for topics where there is an identified census user need. For example, we're researching the use of Department for Work and Pensions (DWP) and Her Majesty's Revenue and Customs (HMRC) information to produce statistics on household and personal income⁸. We'll also use Valuation Office Agency (VOA) data to replace the "number of rooms" question on the 2021 Census.

We'll continue to work to identify administrative data sources with the potential to be used to create enhanced census outputs. We'll aspire for these results to be National Statistics, dependent on the quality of the administrative data source.

We're interested in identifying your requirements for these types of outputs and the benefits this information would bring to your work.

⁷ <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject>

⁸ <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/researchoutputsincomefrompayeandbenefitsfortaxyearending2014>

Bespoke census results

We'll continue to provide a service for creating bespoke or commissioned outputs that haven't been included in standard output products following the 2021 Census. We'll apply the same statistical disclosure control methodology (targeted record swapping and cell key perturbation) as all other standard outputs. This service will become available after the release of 2021 Census outputs. Users will also continue to have access to expert staff for queries and information via the Census Customer Services team.

Publication of 2021 Census results

We recognise the importance of the timely release of 2021 Census data. Our initial plan is to publish a population estimate for England and Wales, and estimates at local authority-level by age and sex within 12 months of 2021 Census day. We'll aim to publish all other results, including multivariate data, within 24 months of 2021 Census day. We'll aim to co-ordinate release dates with Scotland and Northern Ireland, where possible.

We'll publish a prospectus outlining our timetable of intended release dates for 2021 Census results and the content of each release. We'll then announce the actual date at least four weeks before release, in compliance with the Code of Practice for Official Statistics.

The responses to this consultation will help to inform the content of the 2021 Census prospectus and identify the order in which census data are released. We'll hold a final consultation which outlines our more detailed plans before census day.

2021 Census specialist products

Microdata

Microdata are small samples of data for whole households and individuals, which include some associated census characteristics but no information that could identify a household or individual. These samples enable users to access data similar to that, which might be collected if users conducted a survey themselves and can be analysed in the same way.

In 2011, the following microdata files were created with different access arrangements:

Type of microdata	Description
Microdata teaching file	<ul style="list-style-type: none"> • Accessed via the Office for National Statistics (ONS) website • Random sample of 1% of individuals in 2011 Census records • Limited set of variables
Safeguarded microdata	<ul style="list-style-type: none"> • Accessible via UK Data Service • Two random samples of 5% of individuals • One sample is a regional-level file • The other sample is a grouped local authority or single local authority-level data where a population threshold is reached
Secure microdata	<ul style="list-style-type: none"> • Accessed via ONS Virtual Microdata Laboratory (VML)/ ONS Secure Research Service at ONS Titchfield, Newport and London • Accessible by ONS Approved Researchers <p>Individual file</p> <ul style="list-style-type: none"> • 258 variables for over five million individuals • Lowest level of geography – local authority • Includes residents in households and communal establishments <p>Household file</p> <ul style="list-style-type: none"> • Person- and household-level data • 245 variables for over five million individuals within more than 2.4 million households

We'll continue to provide access to microdata samples for users following the 2021 Census. We'll also create a microdata working group and work closely with users of microdata to develop a specification that meets user needs. We'll review the sample design and size to ensure this is appropriate for different users, including professional researchers. We also need to ensure the data doesn't present a statistical disclosure risk. As such, there will be no identifiable individual-level census data released.

Origin-destination

Origin-destination (flow) data produced following the 2011 Census shows the flows of people from one place to another⁹. There were four types of origin-destination data produced for the 2011 Census. These are:

1. Special **Migration** Statistics – migration patterns (internal and international) of individuals based on their location one year before Census day (27 March 2011)
2. Special **Workplace** Statistics – commuting patterns of individuals from place of usual residence as origin to area of workplace (workplace zone) as destination
3. Special **Residence** Statistics – information on the location of second residences in relation to an individual's usual residence or workplace (new for 2011 Census)
4. Special **Student** Statistics – migration patterns of individuals living at student addresses one year prior to the census (new for 2011 Census)

In 2001, origin-destination tables were produced using a post-tabular small cell adjustment statistical disclosure control methodology. This allowed wide and easy access for users, but adversely affected the utility of outputs. In 2011, the statistical disclosure control policy required that different levels of access were provided for origin-destination data.

Origin-destination data were provided to users in three ways:

1. publically via the ONS website and NOMIS
2. as safeguarded data via ONS Virtual Microdata Laboratory (VML) and UK Data Service
3. as secure data via ONS VML/Secure Research Service (SRS) and only available to Approved Researchers

Feedback provided by users after 2011 Census identified that the controlled access to origin-destination data has reduced the usefulness of this information. Users noted that the data are difficult to access and use and that the process through which to access data should be identified earlier in the design stages.

We plan to review the statistical disclosure control methodology used to produce origin-destination tables for the 2021 Census to identify if more data can be made publically available. We'll also review the usage and utility of the 2011 Census tables to identify the data most needed by users. We'll consider the storage arrangements of origin-destination data to improve the ease of access to this data.

We'll create an origin-destination working group and work closely with users of these data to develop specifications and access arrangements that meet user needs, but don't compromise confidentiality of individual respondents.

A large number of the origin-destination data were produced at the UK level, providing flows for usual residents within and between England, Wales, Scotland and Northern Ireland. Any statistics that couldn't be easily harmonised across the UK due to differences in the data collected were provided for usual residents of England and Wales only. We'll continue to work closely with National Records of Scotland (NRS) and Northern Ireland Statistics and Research Agency (NISRA) to ensure harmonisation of definitions and data collected, where possible, for the 2021 Census.

⁹ <https://www.ons.gov.uk/census/2011census/2011censusdata/originanddestinationdata>

Longitudinal Study

We'll also extend the ONS Longitudinal Study by adding data from the 2021 Census for England and Wales to those records already included from the 1971-2011 censuses. The Longitudinal Study is a database of linked census and vital event data for a 1% sample of the population of England and Wales¹⁰. It is covered by the same rigorous confidentiality conditions as the census itself, and the database is securely held by us, with strictly controlled access in accordance with the Statistics and Registration Service Act 2007.

Accessing sensitive census data

We'll provide access to sensitive census data in secure and controlled research environments in accordance with the Statistics and Registration Service Act 2007.

Access to secure data will be granted in accordance with the terms and conditions specified in the ONS Approved Researcher Scheme¹¹. The primary purpose of all Approved Research projects must be to "serve the public good". Any outputs produced must ensure it won't be possible to identify any individual. These arrangements will apply to anyone wishing to access microdata samples, origin-destination products and the Longitudinal Study.

We'll continue to research the best possible location and access arrangements for these census datasets.

¹⁰ <https://www.ons.gov.uk/aboutus/whatwedo/paidservices/longitudinalstudyls>

¹¹ <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme>

2021 Census results – outputs, population bases and geography

Output categories

We're currently undertaking research to determine the topics to be included in the 2021 Census. For the vast majority of topics there is a consensus view. In some areas further research work or additional information such as the view of stakeholders is required before we can finalise our recommendations to government.

Our recommendations will inform a government white paper in 2018. The final approval of the topics and questions to be included in the 2021 Census via the Census Order and the Census Regulations is for Parliament. The Census Regulations in Wales also need approval from the Welsh Assembly.

Population bases

In 2011, the Office for National Statistics (ONS) produced census estimates on the following output bases:

Output base	Description
Usual resident	<p>Primary output base for 2011 Census</p> <p>A usual resident of the UK is anyone who, on census day, was in the UK and had stayed or intended to stay in the UK for a period of 12 months or more or had a permanent UK address and was outside the UK and intended to be outside the UK for less than 12 months.</p>
Short term resident	<p>A census short-term UK resident is anyone born outside of the UK who, on 27 March 2011, had stayed or intended to stay in the UK for a period of three months or more but less than 12 months.</p>
Households	<p>A household is defined as:</p> <ul style="list-style-type: none"> • one person living alone, or • a group of people (not necessarily related) living at the same address who share cooking facilities and share a living room or sitting room or dining area
Second residence populations	<p>Derived from the second address question</p> <p>A second address is an address at which a person stays for more than 30 days per year that is not a person's place of usual residence. This includes addresses that are in the UK and those outside of the UK for example armed forces bases, student's home addresses, holiday homes.</p>

(continued on next page)

Output base	Description
Out-of-term populations	Alternative population base whereby the usually resident population are redistributed to their place of residence in out of term periods.
Workplace populations	All usual residents aged 16 and above who are in employment and whose workplace is in the area. People who work mainly at or from home or don't have a fixed place of work are included in the area of their usual residence.
Workday populations	All usual residents aged 16 and above who are in employment and whose workplace is in the area, and all other usual residents of any age who aren't in employment but are resident in the area. People who work mainly at or from home or don't have a fixed place of work are included in the area of their usual residence.

In 2015, we held a public consultation to identify the topics to be included in the 2021 Census¹². It also requested users to supply information on the suitability of the 2011 Census output bases and the need for any new output bases or changes/improvements to these. Users indicated sufficient flexibility in the 2011 output bases and expressed a need for them to remain the same in 2021 as for the 2011 Census to allow comparisons to be made over time.

In 2021, we'll continue to produce estimates using the output bases produced in 2011 Census.

Small populations

Small population statistics provide census data for some of the key characteristics of people in specific small population groups, for example individuals of an ethnic group, or those with a specific country of birth. These are produced where the small size of the total population in that group means confidentiality constraints limit the release of more detailed standard statistics.

These small population data are only produced for geographic areas in which the small population being counted reaches a certain threshold (100 or 200 persons depending on the level of detail in the table). Only the areas in which the population exceeds this threshold are included in each table. In 2011, we produced over 35 tables containing data about small populations¹³.

We'll continue to consult with users to identify the value of producing statistics for small populations and consider the best way to disseminate this data. A commissioned table service will be provided for those tables it isn't possible to generate. We'll also consider the small populations for which it may be possible to generate statistics using the very detailed write-in responses to the ethnicity, religion and country of birth questions.

¹² <https://www.ons.gov.uk/census/censustransformationprogramme/consultations/the2021censusinitialviewoncontentforenglandandwales>

¹³ <https://www.ons.gov.uk/census/2011census/2011censusdata/2011censusdatacatalogue/smallpopulationgroups>

Geography

Geography is an integral part of census outputs. In 2021, we'll provide census outputs in line with the Government Statistical Service (GSS) Geography Policy. The policy outlines the principles for using geographic information to produce and disseminate statistics.

As in 2011, Output Areas and Workplace Zones will be the building blocks for other census geographies.

Output Areas (OAs) are the smallest areas for which residential census statistics are produced and were first created for the 2001 Census. In England and Wales, Super Output Areas (SOAs) consist of Lower Layer Super Output Areas (LSOAs) and Middle Layer Super Output Areas (MSOAs). SOAs are larger areas constructed from groupings of Output Areas and support analysis of data that would be disclosive if released at the OA level.

Workplace Zones (WZs) are the smallest areas for which workplace statistics are produced and were first created for the 2011 Census.

Feedback from users in previous ONS consultations about the use of Output Areas and Workplace Zones has been positive. Users would like these geographies to remain stable, comparable and enable continuity with previous censuses.

We recognise stability is a priority for the release of census geographic outputs. We'll ensure consistency with the approach used in the 2011 Census where possible, with Output Area boundaries remaining largely unchanged to ensure comparability with 2001 and 2011 Census results. Workplace Zone boundaries will also be updated, although the extent of modification required is currently unknown.

We'll consider revising output area boundaries where there has been significant local population change or housing development. This will be primarily by splitting or merging existing OAs to ensure the key criteria of size and homogeneity are followed. We don't anticipate changing more than 5% of current OA boundaries in order to meet these requirements. We'll also review the Workplace Zone boundaries created in 2011 to reflect changes in the workplace population.

We'll notify users of all revisions to boundaries, so that it is possible to compare data between 2001, 2011 and 2021 Census geographies. All revised OAs and WZs will nest into a stable hierarchy of Lower Layer and Middle Layer SOAs.

We'll release outputs for administrative geographies (regions, counties, local authorities and wards) on a best-fit basis to allow data to be allocated to boundaries that may change between 2011 and 2021 censuses. An exception to this will be national parks, which will be produced on an exact-fit basis. This is because their irregular shape and uneven distribution of population makes aggregation of data from OA more difficult. We'll also produce postcode-level headcounts. We'll provide a look-up between OAs and WZs and higher-level geographies.

In 2011, population estimates were produced for Eurostat using postcode-level population estimates best-fitted to a 1km grid square. We've identified some interest in grid-based statistics, but we would like to use this consultation to further explore the user requirement for these results.

In 2011, there were approximately 10% of civil parishes in England that didn't contain an Output Area population weighted centroid using the "best-fit" methodology for producing census estimates. Therefore, it wasn't possible to produce population estimates for these parishes. Recent research by ONS Geography failed to identify a suitable alternative methodology to produce estimates for civil parishes¹⁴. We're interested in identifying the need for data for all civil parishes in England and the variables required at this level.

We're working closely with National Records of Scotland (NRS) and Northern Ireland Statistics and Research Agency (NISRA) to ensure the methodology for creating OAs and WZs is harmonised, allowing outputs to be released and analysed at a UK level.

¹⁴ <http://bit.ly/2FzkABN>

Interpreting 2021 Census results

Metadata

Metadata are produced alongside census outputs to provide contextual information about the data and aid understanding and interpretation. Prior to the 2011 Census, users identified the requirement for metadata located alongside the outputs and clear signposting to supporting information. We provided a range of documentation about the information contained within the data tables, and supporting documentation, for example about the quality of outputs and definitions used.

In 2021, we'll produce metadata including definitions, details of variables, classifications and quality indicators to support the use of statistical outputs. These will include, in particular, confidence intervals and other quality measures and information from post-census coverage and quality surveys such as local coverage rates, item non-response and imputation rates. We'll also provide information about the comparability of 2021 outputs with 2011 results and of England and Wales's estimates to those produced in Scotland and Northern Ireland. This is to illustrate the differences in definitions, concepts, topics, questionnaire wording and where administrative data has been used.

We'll work closely with National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) to ensure consistent metadata are produced to enable interpretation of UK-wide statistics.

We'll also extensively test the 2021 Census online dissemination platform with a range of data users. We'll use this testing to identify the most appropriate way to provide consistent and clear metadata to users.

Ensuring high quality results

We have a responsibility to ensure the 2021 Census results are correct, accurate and adhere to the Government Statistical Service Code of Practice for Official Statistics¹⁵ as determined by the UK Statistics Authority.

The Code of Practice for Official Statistics promotes the production and dissemination of official statistics that inform decision-making. It helps producers and users of statistics by setting out the necessary principles and practices to produce statistics that are trustworthy, high quality and of public value.

We have already started on research to identify the processes and procedures required to ensure we achieve National Statistics Accreditation before census day. We'll also undertake research to identify how to assess the quality of enhanced census outputs produced using both census and administrative data.

¹⁵ <https://www.statisticsauthority.gov.uk/osr/code-of-practice/>

We'll adhere to the Code of Practice and notify users at least four weeks in advance of the release of census results. We'll also provide supporting metadata (see above), including a report on quality and methodology, to accompany and support the release of the main statistical outputs. We'll provide users with our assessment of the quality of census results, based on our rigorous quality assurance of outputs.

Analysis programme

We're currently working with the wider Government Statistical Service (GSS) to develop our analytical programme. This is so we can make the best use of census data, alongside administrative and survey data to describe the society we live in.

We'll engage with users closer to census day to understand their requirements for data analysis and how this aligns with our data release timetable.

We'll aim to ensure the widest possible awareness and use of 2021 Census outputs and add value by helping users to understand and interpret census data. We'll produce national, regional and local authority summaries, along with reports and analysis on key findings and particular topics.

We'll consider the different types of census data users and provide analysis and commentary in a range of formats, including stories, summaries and interactive data visualisations. We'll publish this supporting information alongside the release of 2021 Census data according to a pre-released timetable.

Licensing, funding and charging models

In 2021, we'll provide standard census outputs online, free of charge at the point of delivery under the open government licence (OGL). We'll reserve the right to charge for:

- the supply of data on physical media
- the services to create bespoke tabulation via the census commissioned table service

We'll issue tables generated through the commissioned tables service to the requesting customer under an open government licence (OGL) by the provision of a link to a URL on the ONS website. These tables will then be freely available to all other users. The service for bespoke or commissioned output will be made available as soon as the main results are released.

The costs incurred will solely be to cover the administration and generation of the products and are not for profit. The actual costs will be determined closer to 2021, but are expected to be broadly in line with the 2011 Census charging policy.

Where bespoke outputs don't provide sufficient detail for particular statistical research purposes, access to more detailed data will be possible within strictly controlled conditions. Any researcher wishing to access secure data must be an "Approved Researcher" and meet the criteria set out in the Statistics and Registration Service Act 2007.

We'll continue to work with value added resellers, the extent of this relationship will be determined along with the associated charging policy.

International comparability of 2021 Census results

We'll continue to meet our international commitments to supply census data to the United Nations Statistical Commission following the 2021 Census. We anticipate that information will be prepared to support comparative assessment and analysis across Europe. We'll continue to work on an appropriate technological solution to supply this data to relevant organisations. Where topics required for international reporting aren't covered on the census in England and Wales, alternative sources will be identified to fulfil our duties as far as possible.

Conclusion

The overall design of 2021 Census output content will be evaluated against the following criteria:

Criteria	Description
Accessibility	<p>This covers the ability to easily locate and access 2021 Census data for each country and UK wide. This includes access to delivery of large volumes of data, commissioned tables, safeguarded and secure products.</p> <p>This includes the ability to access metadata and commentary to understand and interpret data published.</p>
Flexibility	<p>This covers the ability to access data that meets the needs of users for example, create own data table.</p>
Timeliness	<p>This is a measure of the punctuality of the release of census outputs. This includes the release of specialist products (microdata and origin-destination). Census outputs should be produced as soon as possible after census day.</p>
Feasibility	<p>This covers the operational feasibility of developing statistical disclosure control methods and a flexible table builder.</p>
Statistical Disclosure Control risk	<p>This covers ensuring that the Statistical Disclosure Control method chosen will provide sufficient uncertainty in the published outputs.</p>
Relative costs	<p>This covers the cost of developing and implementing possible outputs and dissemination solutions.</p>
User acceptability	<p>This covers testing the functionality of output solutions against user requirements.</p>
Public confidence	<p>This covers ensuring data is seen to be anonymised and the methods used to produce outputs are transparent and fully communicated to users.</p>

The responses you provide to this consultation will be used to understand the strength of impact (positive, negative or none) of our intended design on the accessibility, flexibility and timeliness of outputs for users. We'll use other evidence to evaluate our design against the remaining criteria.

This evidence will be used to identify the most important concerns for users and prioritise research on these aspects.

We'll also take into consideration the type of work you are undertaking. This will be based on the ONS Approved Researcher Scheme definition of research being undertaken to "serve the public good"¹⁶.

Types of work could include:

- using census data for public policy decision making
- using census data for public service planning and delivery
- using census data for allocation of public resources
- using census data for decisions which are likely to significantly benefit the UK economy, society or quality of life of people in the UK

We'll only evaluate the evidence provided in responses to this consultation. Please provide as much detail as possible in your responses, including reference to specific policies, projects or research, if applicable.

We value your input and thank you for your contribution.

¹⁶ <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme>

Annex - Paper questionnaire template

The best way to respond to this consultation is online ¹⁷. If you wish to respond to this consultation via email or on paper, please complete the questionnaire template and email your response to census.outputs@ons.gov.uk using the subject line "Initial View Consultation". There are some mandatory questions which we need you to complete in order to include your response in our analysis.

You can respond to all the questions or just the questions that are relevant to your work.

Introduction

Mandatory questions are marked with an asterisk (*)

1 *Are you answering this consultation on behalf of an organisation or as an individual?

- Organisation
 Individual

2 *What sector do you work in? This will assist us in monitoring the range of users that have responded to this consultation.

- Government department
 Public body (for example health, transport, emergency services)
 Local authority
 Housing
 Academic / Research
 Charity and Voluntary
 Commercial
 Utility
 Journalist/media
 Genealogist/family historian
 Other (please specify)

3 *What is the name of the organisation that you represent?

4 *Contact Information

Your Name

Email address

Phone number

¹⁷ <https://consultations.ons.gov.uk/census/initial-view-on-the-2021-census-output-design>

5 *We may wish to contact you in relation to your response to this consultation. Would you be happy for us to do so?

Yes

No

6 *To support transparency in our decision-making process, responses to this consultation will be made public. This will include the name of the responding organisation or individual. Please confirm that you're content for your name to be published. We won't publish personal contact details. Please be aware that any information provided in response to this consultation could be made publicly available if requested under a Freedom of Information request. Please be aware that if you're responding as an individual but then name your organisation in any of the open questions, that name will be published.

Yes, I consent to my name being published with my response

No, please remove my name before publishing my response

Please read through our consultation document, which outlines our initial view for the design of 2021 Census output content and dissemination. The following questions are related to the topics discussed in the consultation document. Please provide as much detail as you can to support your answers. If we've asked you to describe the impact on your work, please provide specific examples. If you refer to a specific policy or scheme, please provide a direct reference to this. We'll only evaluate the evidence provided in your response. We have asked questions based on your experience with 2011 Census data. If you haven't used 2011 Census data, we'd still be interested to hear your thoughts about your needs for 2021 Census data.

After each section, a box will be provided for you to add any other comments that weren't addressed in the questions.

Background

We'd like to know which topics and geographies you used the most after the 2011 Census. If you didn't use 2011 Census data, we would still be interested to hear your views about the data you are likely to use after 2021 Census.

7 What are your main topics of interest?

We've included some topics that are currently under consideration for inclusion in the next census or might be available through administrative data.

Please rank in order your topics of interest, with 1 being the most important topic of interest. If you did not use a topic, please leave this blank.

- | | |
|--|---|
| <input type="checkbox"/> Age | <input type="checkbox"/> Marital status |
| <input type="checkbox"/> Car availability | <input type="checkbox"/> Migration |
| <input type="checkbox"/> Carers | <input type="checkbox"/> National identity |
| <input type="checkbox"/> Central heating | <input type="checkbox"/> Number of rooms and bedrooms |
| <input type="checkbox"/> Citizenship | <input type="checkbox"/> Occupation |
| <input type="checkbox"/> Country of birth | <input type="checkbox"/> Religion |
| <input type="checkbox"/> Education | <input type="checkbox"/> Sex |
| <input type="checkbox"/> Ethnicity | <input type="checkbox"/> Sexual Orientation |
| <input type="checkbox"/> Gender Identity | <input type="checkbox"/> Socio-economic classifications |
| <input type="checkbox"/> Health | <input type="checkbox"/> Tenure and landlord |
| <input type="checkbox"/> Household composition | <input type="checkbox"/> Travel |
| <input type="checkbox"/> Income | <input type="checkbox"/> Type of accommodation |
| <input type="checkbox"/> Language | <input type="checkbox"/> Veterans |
| <input type="checkbox"/> Labour market | Other – please specify |

8 In 2011, which geographies did you use?

Please rank in order your geographies of interest, with 1 being the most important topic of interest. If you did not use a geography, please leave this blank.

- | | |
|---|--|
| <input type="checkbox"/> Local Authorities | <input type="checkbox"/> Middle layer super output areas (MSOAs) |
| <input type="checkbox"/> Parishes | <input type="checkbox"/> Travel to Work Areas (TTWAs) |
| <input type="checkbox"/> Regions | <input type="checkbox"/> Workplace Zones |
| <input type="checkbox"/> Wards | <input type="checkbox"/> Nomenclature of territorial units for statistics (NUTS) |
| <input type="checkbox"/> Counties / upper tier local authorities | <input type="checkbox"/> Parliamentary constituencies |
| <input type="checkbox"/> Countries | <input type="checkbox"/> Local Enterprise Partnerships |
| <input type="checkbox"/> Built-up areas (including sub-divisions) | <input type="checkbox"/> National Parks |
| <input type="checkbox"/> GEOSTAT Grid | <input type="checkbox"/> Postcode |
| <input type="checkbox"/> Output Areas | <input type="checkbox"/> Other |
| <input type="checkbox"/> Lower layer super output areas (LSOAs) | |

It will be helpful for us to know what you use census data for to allow us to understand how best to further support your needs.

9 In 2011, what did you use Census data for?

Select all that apply

- Commercial purposes
- Resource allocation
- Service planning and delivery
- Policy development and monitoring
- Academic research
- Personal interest
- Other purposes – please specify

Accessing 2021 Census Estimates

Flexible Dissemination System

For 2021, we aim to disseminate census data through a flexible dissemination system allowing for users to be able to define their own tables, selecting the geography, population base and variables they require.

This approach will enable you to have access to the data you want as soon as they're available. This means we'll be able to release data much sooner than following the 2011 Census.

The following questions are assessing your need to access the data this way and what impact it would have on your work.

10 I would prefer to access data via an online flexible dissemination system sooner after census day than wait for tables to be developed by ONS.

Strongly Agree Agree Not Sure Disagree Strongly Disagree

11 Being able to define my own tables is more useful to me than using predefined tables.

Strongly Agree Agree Not Sure Disagree Strongly Disagree

12 I would find it useful to be provided with a small number of univariate tables alongside those I can generate through the flexible dissemination system.

Strongly Agree Agree Not Sure Disagree Strongly Disagree

13 I would find it acceptable to use tables where there are small inconsistencies in output table totals due to the perturbation caused by the cell-key method.

Strongly Agree Agree Not Sure Disagree Strongly Disagree

14 I need a very high proportion of aggregate data to install in my own systems (for example bulk data)

Strongly Agree Agree Not Sure Disagree Strongly Disagree

15 Currently, we're planning to disseminate most public census data using the flexible dissemination system, but we'll also aim to provide an application programming interface (API). Would your systems be able to utilise an API to access census data?

Yes

No

16 How would you expect to make use data accessed via an API?

17 Any other comments?

2021 Census Specialist Products

Enhanced Outputs

Enhanced Census Outputs will be new for 2021. In these, administrative data sources will be integrated with the census data to produce outputs, predominantly on characteristics of the population not historically included on the census. They could also be used to enhance existing questions by providing more detail e.g. using Valuation Office Agency (VOA) property type to break down the census property types into smaller categories.

18 Which topics, not currently collected on the census questionnaire would you/your organisation be most interested in receiving data about?

Income

Size of property in square meters

Type of property (more detailed breakdown)

19 What impact would this type of information have on the work of you/your organisation?

20 Any other comments?

Microdata

Microdata are small samples of data for whole households and individuals, which include some associated census characteristics but no information that could identify a household or individual. We'll continue to provide access to microdata samples for users following the 2021 Census.

21 In 2011, which microdata samples did you access?

- Public
- Safeguarded
- Secure
- None

22 In 2011, you were sufficiently informed of the content and availability of microdata samples.

Strongly Agree Agree Not Sure Disagree Strongly Disagree

23 What research/analysis has the microdata sample allowed you to undertake?

Please provide a link to any reports that have been published showing the impact of using microdata.

24 Please describe any limitations when using 2011 Census microdata samples including:

- a. access to data (public, safeguarded, secure)
- b. sample design (public, safeguarded, secure)

25 Are you interested in contributing further to the design of 2021 Census microdata through a Working Group?

- Yes
- No

26 Any other comments?

Origin-Destination

Origin-destination (flow) data produced following the 2011 Census shows the flows of people from one place to another. We plan to review the Statistical Disclosure Control methodology used to produce origin-destination tables for the 2021 Census to identify if more data can be made publicly available.

27 In 2011, which origin-destination data did you access?

- Public
- Safeguarded
- Secure
- None

28 In 2011, you were sufficiently informed of the content and availability of origin-destination data.

Strongly Agree Agree Not Sure Disagree Strongly Disagree

29 What research/analysis has the origin-destination data allowed you to undertake?

Please provide a link to any reports that have been published showing the impact of using origin-destination data.

30 Please describe any limitations when using 2011 Census origin-destination samples including:

- a. access to data (public, safeguarded, secure)
- b. table design (public, safeguarded, secure)

31 Are you interested in contributing further to the design of 2021 Census origin-destination through a Working Group?

- Yes
- No

32 Any other comments?

Geography

Parishes

The GSS Geography Policy sets out a policy of best-fitting from Output Areas to administrative geographies based on a population weighted centroid. As a number of parishes are much smaller than Output Areas this meant that in 2011, approximately 10% of all parishes in England received no statistics as they don't contain an Output Area centroid.

Work has been done to look at alternative methodologies for producing parish-level statistics. It is clear that each methodology is a balance between having data available for all parishes and the number of variables that can be released. The more variables that are released, the fewer parishes they can be released for due to disclosure.

33 I have a requirement for parish data

- Yes
- No

34 To understand your requirement we would like to know which is more important to you as a user. Please rank in order of preference, with 1 being most important.

- Having data available for all statistical parishes even if it means that only a few variables are available
- Having a number of variables available for parishes even if it means that not all parishes will have statistics produced
- Having estimates for all parishes, but some parishes will have more variables provided than others

Grid data

As part of its 2011 Census outputs, ONS released statistics on a 1km grid to support analysis at the European level. We'll also produce statistical outputs by grid as part of the 2021 Census to support the integration of statistics across multiple domains (for example population, environmental) and to better understand the distribution of data compared to the Output Area hierarchy. This output will sit alongside the Output Areas as dissemination geography and will not replace them.

To inform the methodology for producing these statistics it would be useful to understand your requirements and how they might be used to support spatial analysis.

35 Did you use grid square data in 2011 for your work?

- Yes
 No

36 ONS is considering producing 2021 Census outputs on grids. Would Census data on grids be useful to you?

- Yes
 No

37 What impact would this have on your work?

Please provide an example (for example what benefit might this bring to your data?)

38 If you'll be using grid data in 2021 it would be particularly useful to know if you have a preference to have the data released on the British National Grid rather than the Pan-European one.

Please select your preference as to which grid you would prefer to have the data released on:

- British National Grid
 Pan-European Grid
 I don't mind

39 Any other comments?

Interpreting 2021 Census Estimates

Supporting information

In 2011, We provided a range of documentation about the information contained within the data tables, and supporting documentation, for example about the quality of outputs and definitions used. This information provides context to the data, for example how they were collected and the coverage of the data, for example publication date, description and search keywords. We want to make sure the information we provide is useful and easy for you to understand.

40 Which supporting information did you use in 2011? Please tick all that apply

- Census User Guide
- Quality and methods (for example Statistical disclosure methodology, response rates)
- Glossary
- Variables and classifications descriptions
- Other (please specify)

41 Where would you expect to access supporting information?

- In the data tables
- In a link next to the data table
- Accessed via another part of the ONS website

42 Any other comments?

Wider analytical uses of Census data

43 In 2011, did you read the accompanying analysis and commentary?

- Yes
- No (I didn't need to)
- No (I didn't know it was available)

44 What format of accompanying commentary do you find most useful?

Please prioritise what you find most useful using the ranking tool:

- Statistical bulletin describing key outputs
- Visualisations
- None
- Other

45 Any other comments? If you selected 'Other' please specify in the box below.

Comparability of 2021 Census Estimates

UK Harmonisation and UK Statistics

46 How important is it being able to access data for England, Scotland, Wales and Northern Ireland for your work?

Very important Important Not Sure Not important Very unimportant

47 How would you expect to access UK data?

- Access data about each UK country through a single website for example, select data about Scottish local authorities through the ONS website
- Visit each UK country website to access data about that country

48 Are you interested in:

- UK statistics totals
- Small area statistics for different geographies across England, Wales, Scotland and Northern Ireland

49 Please demonstrate the impact of using UK statistics to your work, please provide any links to any supporting information you have.

50 Any other comments?

International Comparability

51 Have you compared UK estimates to international datasets?

Yes

No

52 Please provide examples where you have used UK data to compare to international datasets

53 Any other comments?

