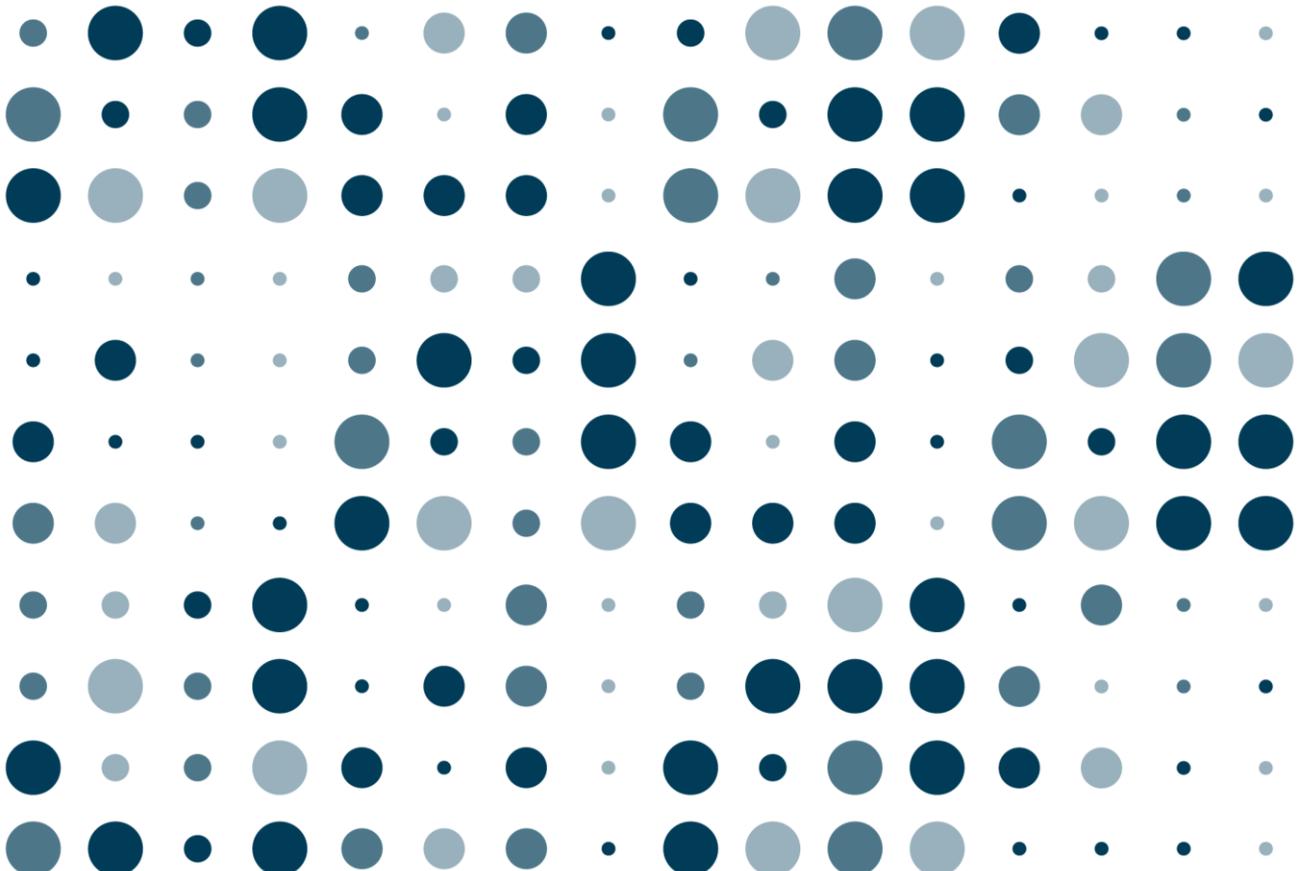


# Census 2021 outputs: content design and release phase proposals

## Consultation document

13 July 2021



# Contents

<b>General information</b>	<b>2</b>
<b>Section 1: Why we are consulting</b>	<b>4</b>
Purpose of the census	4
Consultation context	4
Aim and structure of this consultation	5
Future stakeholder engagement and consultation	7
<b>Section 2: Shape of the outputs and analysis release schedule</b>	<b>8</b>
Overview	8
Geography	11
Phase 1: Initial findings and topic summaries	12
Phase 2: Multivariate data for the usual resident population base	15
Phase 3: Alternative population bases, origin-destination data and microdata	18
Beyond March 2023: UK data and further analysis	22
<b>Section 3: Main changes to variables compared to the 2011 Census</b>	<b>26</b>
Questions with detail removed	27
New questions and associated classifications	28
Other significant changes to output classifications	31
<b>Section 4: Proposals for feasibility work to derive new variables</b>	<b>35</b>
Education and employment	35
Health and living arrangements	37
Accommodation types and vacant addresses	38
<b>Section 5: Population-base specifications</b>	<b>40</b>
Main population bases	40
Alternative population bases	41
Small populations	44
<b>Section 6: Taking a census during a period of change</b>	<b>47</b>
Place of residence	48
Economic activity and employment	49
Travel to work	49
Workplace zones	50
Impacts on other outputs	51
<b>Section 7: Paradata</b>	<b>52</b>
Household and individual response rate	52
Mode and language of response	53
<b>Section 8: Next steps</b>	<b>54</b>

## General information

### Consultation details

**Issued:** 13/07/2021

**Respond by:** 05/10/2021

We will publish a response to the consultation approximately 12 weeks after the consultation date.

We will publish the final design of Census 2021 outputs, in the form of a Census 2021 Outputs Prospectus, by February 2022.

### Enquiries to

**Email:** [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk)

**Phone:** +44 (0)1329 444972

**Fax:** +44 (0)1633 652981

**Post:** Census Outputs Consultation, Office for National Statistics, Segensworth Road, Titchfield, Hampshire, PO15 5RR

### Territorial scope

This consultation primarily relates to Census 2021 statistics for England and Wales. However, some questions consider harmonisation with Census 2021 data for Northern Ireland and Census 2022 data for Scotland.

### How to respond

We welcome contributions from all census users. This includes those who are experienced users of census data through to those who are looking to use census data for the first time.

We strongly encourage you to respond to this consultation by filling in the [online questionnaire](#).

You can also respond by email or post, using the downloadable version of the questionnaire. You can find this and other supporting documents in the “related links” section of the online questionnaire.

## Accessibility

We can provide all material relating to this consultation in braille, large-print or audio formats on request. You can request British Sign Language interpreters for any supporting events.

## Confidentiality and data protection

We aim to be as open and transparent as possible, so we will publish all responses to this consultation. Please indicate in your response if you do not want your name to be published when we respond to this consultation.

We will moderate responses before publication to identify and remove any offensive, hateful, inappropriate, commercially sensitive or market sensitive content.

As a public authority, the Office for National Statistics (ONS) is subject to the Freedom of Information Act. While we would never voluntarily release personal information from responses to this consultation, we are not able to provide a guarantee of confidentiality.

Please see our [privacy notice](#) for full details of how we will handle your data.

## Quality assurance

This consultation has been carried out in accordance with the [government's consultation principles](#).

If you have any complaints about the way we've conducted this consultation, please email: [ons.consultations@ons.gov.uk](mailto:ons.consultations@ons.gov.uk).

# Section 1: Why we are consulting

## Purpose of the census

All kinds of organisations, from local authorities to charities, use census information. The England and Wales data we collect are used both independently, and alongside Northern Ireland and Scotland data, to produce UK-level data. To ensure that users derive the most value from the data, it's important that we continue to understand their needs to inform our outputs.

There are two main types of uses for Census 2021 data: describing the population on Census Day and informing the decisions of the future.

## Describing the population on Census Day

Census 2021 data will be fundamental to understanding the impact the coronavirus (COVID-19) pandemic has had on different communities. For example, information from the 2011 Census has already been crucial in our [understanding of mortality for different groups](#) during the pandemic.

It's also currently the only source of local-level information on topics such as occupation and household composition. With fresh data from 2021, we can update past analysis and use it alongside new data sources to give us the richest data on our population we've ever had.

## Informing the decisions of the future

Currently, we rely on the census to provide the most accurate source of population data, both nationally and in small areas, in the years between censuses. It informs planning decisions on public services, including hospitals, schools, universities, and job centres, to meet the needs of our changing society. It will also form the basis of many of our other statistics that will follow in the subsequent decade.

## Consultation context

We aim to increase the use of census data. This is as described in the White paper the government presented to Parliament in December 2018, [Help Shape Our Future: The 2021 Census of Population and Housing in England and Wales](#). This aim is in line with the UK Statistics Authority Strategy [Statistics for the public good](#) and the [National Data Strategy](#).

To help achieve this aim, we will produce a range of [census output products](#), enabling users to interact with census data themselves. Examples of these interactive products include the ability for users to build their own tables using flexible datasets or explore

area profiles. However, we want to ensure these products contain the data that users need.

Throughout the development of the census, we have engaged with users to help us to understand their changing needs. This consultation is part of that continual engagement process, focusing on designing our different data products to meet users' needs.

It follows and builds on the work of the [Inclusive Data Taskforce](#). Earlier this year, it ran the [UK Statistics Authority Inclusive Data Collection consultation](#) to understand how inclusive data and evidence are across the UK.

As a result of this continual engagement, we're confident that we've collected high-quality data to meet users' needs. However, events such as the coronavirus pandemic and Brexit have increased the rate of change in many of our personal circumstances. For example, for many of us, where and how we live, study and work are now quite different to the beginning of 2020. These circumstances are likely to change again.

To inform national and local service planning, we will need to update Census 2021 data using other sources to measure these and other changes. For example, this might include information on commuting patterns.

We've already begun planning for additional work to meet needs that we've identified through our established stakeholder engagement processes. This stakeholder engagement includes Census Advisory Groups and liaison with other government departments.

## **Aim and structure of this consultation**

### **Consultation aim**

Through this consultation, we aim to get an updated view of detailed user needs for data and analysis from Census 2021. This is especially important around topic areas experiencing an increased rate of change. By understanding user needs, we can ensure that we produce outputs that users can interact with and that help answer the most important questions facing society today.

We welcome contributions from all census users. This includes those who are experienced users of census data through to those who are looking to use census data for the first time.

We're asking for feedback on our proposals about what census data to produce, which we've outlined throughout this consultation. These proposals cover the full census outputs and analysis release schedule and a range of different outputs products. We've

included specific questions on our proposals for dataset specifications, analysis products and feasibility work on new variables.

We will evaluate the information you provide in response to this consultation, alongside information we collected during other stakeholder engagement activities. This will inform decisions on the final design of Census 2021 outputs for England and Wales. When making these decisions, we will consider how the publication of data will serve the public good. We will also consider operational factors such as data quality and disclosure of personal information risk.

## **Consultation structure**

In each section of this consultation document, we provide our current thinking. We've highlighted areas where we're particularly seeking user views using statements in bold that start "We'd like to know". Within the questionnaire, we've provided detailed questions designed to help users give a response to each of these statements.

We've structured the questionnaire into sections, in the same way as this document. As a result, if you only wish to consult on certain aspects of the proposals, you can easily find the corresponding sections of the questionnaire.

In summary, the consultative sections and main questions comprise the following topics:

### [Section 2 - Shape of the outputs and analysis release schedule](#)

This asks about the approach to providing guidance on revised geographical boundaries. It also asks for feedback on the recommended ordering of data releases, and the analytical outputs that we will produce.

### [Section 3 - Main changes to variables compared to the 2011 Census](#)

This asks about plans for producing statistics and analysis using responses to the new questions on sexual orientation, gender identity, armed forces veterans and proxy response. It also asks about ethnic group, age and country of birth classifications and plans for using data from new response options in Census 2021 questions.

### [Section 4 - Proposals for potential new derived variables](#)

This asks about proposed new variables or indicators. It also asks if there are any further user needs that we could meet through innovative use of the data we collected through the Census 2021 questionnaire.

### [Section 5 - Population-base specifications](#)

This asks questions about what analysis data users would conduct using alternative population bases and their needs for small population data.

### [Section 6 – Taking a census during a period of change](#)

This asks about users' needs to understand how the pandemic has impacted place of

residence, economic activity, employment, travel to work and other topics. It asks about data needs for workplace zones and where there are impacts on the utility of the data for decision-making.

### [Section 7 - Paradata](#)

This asks about users' needs for paradata on how we collected the data.

## **Future stakeholder engagement and consultation**

The results of this Census 2021 outputs specification consultation will also inform our goals as an organisation. This includes plans for our longer-term journey to a fully transformed statistical system that makes the best use of all available data.

Our goal is to provide more regular and responsive statistics, particularly in small areas. This will give an improved level of agility, continually adapting to reflect modern society and the new challenges it provides. It will also allow monitoring of a wider and more responsive range of variables and characteristics.

We've already begun this work and our [overview of progress](#) outlines what we've done so far. We will continue to provide updates on this work. In June 2021, we published our [latest set of experimental statistics on small area income distributions](#). These statistics include estimates of self-employment income for the first time.

Later this year, we will launch a further major phase of user engagement. In this, we will share with stakeholders the nature of the transformation we're proposing. It will also provide an opportunity for us to share the framework of evidence that will inform the National Statistician's recommendation on the future of the census and population and social statistics. The National Statistician will make this recommendation by the end of 2023.

We're also conducting a series of user-research projects to improve how people will be able to find, access and use Census 2021 data products. The purpose of this research is to identify ways we can increase the accessibility and utility of census data. Our aim is for a wider range of users to use Census 2021 data and for a wider range of purposes in comparison to the 2011 Census data. If you'd like to participate in this user research, please email us at [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk) for more information.

## Section 2: Shape of the outputs and analysis release schedule

### Overview

We aim to release the first Census 2021 data on the population of England and Wales within a year of census. Following that, we intend to release all the main results within two years of census. These outputs will include a wide range of products that users can interact with to obtain data to suit their needs.

We've developed this timetable based on feedback received from users following the 2011 Census.

It took us nearly three years to release 2011 Census outputs. User feedback highlighted that a main area of improvement for Census 2021 outputs would be to decrease the length of time between data collection and release.

Our shorter release schedule for Census 2021 will ensure that organisations can use the data to inform decision-making and service provision while the information is most relevant. We've enabled this through:

- moving away from a primarily paper-based data collection
- making outputs production more efficient
- automating disclosure control checks

For the first time, Census 2021 was primarily an online census, making it more convenient to complete for most of the population. To ensure that the census was inclusive for all, we offered a full range of support services. This included providing paper questionnaires where required. The contact centre provided guidance and help over the telephone and through digital services such as web chat and social media. We also provided census questionnaires in Braille and British Sign Language.

When we process the Census 2021 data, we will use our newly built Data Access Platform. We've designed this platform to meet the needs of a national statistics institute. New technologies will support the initial processing of responses in real time, allowing us to carry out analysis and adapt the data collection, if needed. We can then apply the approach we're developing to our other data collection exercises. This has the potential to improve the statistical quality and timeliness of outputs.

Statistical disclosure control procedures will include:

- modifying some of the data before we release the statistics, using methods such as record swapping, some form of cell perturbation or suppression
- limiting outputs by simplifying or grouping data - for example, providing information for aggregated age groups

- amalgamating information to ensure that data are only available in a non-disclosive way - for example, by grouping categories of data or only providing information for larger geographical areas

More information on data processing and statistical disclosure control is available in the [Census 2021 White Paper](#).

We've also sought to reduce the time it takes us to release data for lower-layer geographies and certain topics. This is in response to feedback that highlighted this as an area for improvement following the 2011 Census. The aim is to provide data down to Output Areas (OAs) in the first phase of the release schedule when we publish topic summaries. OAs are the smallest geography we will make the standard outputs available for. These are generally around 100 households in size.

Within this first phase, we're proposing a staggered approach to the release of data and analysis by topic. Users will have the full set of detail, from both a geographical and topic perspective, very early on in the schedule. This is down to the reduced time span for the publication of these initial releases. We will combine this data to produce area profiles.

In each release phase, the products increase in complexity. This means we will be releasing each product as early as feasible. Each phase is discussed in more detail later in this section.

In the second phase of the release schedule, we will publish multivariate data. These outputs combine variables to enable us to fully explore and understand the relationships between them. We aim to produce outputs all the way down to the smallest geographies as part of these releases.

In the third phase of the release schedule, we will publish the additional products, including data for alternative and small populations, origin-destination datasets, and microdata.

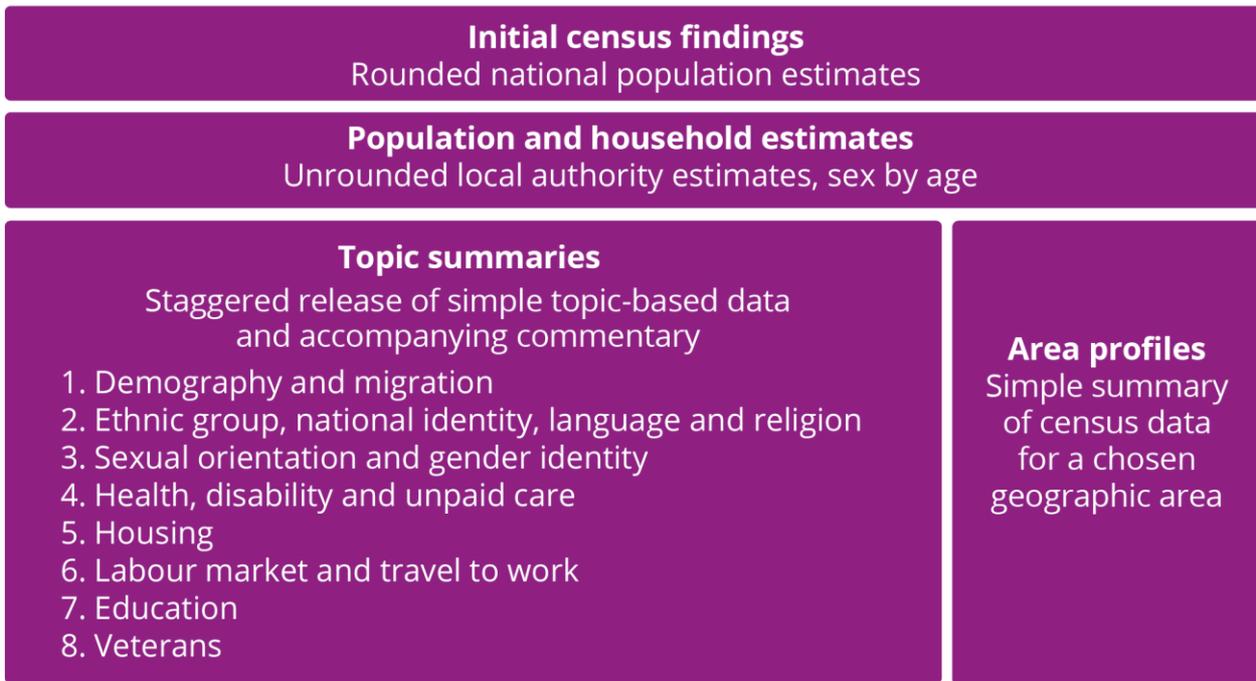
In each phase, we will publish statistical commentary or analysis alongside the data to support and explain them.

Figure 1 shows how these proposals come together to form the "shape" of the Census 2021 outputs and analysis release schedule. The exact timing of the release of outputs will depend upon the processing and quality assurance of census data and will be announced closer to the time of release.

**Figure 1: Census 2021 outputs release schedule**

**March 2022 to March 2023**

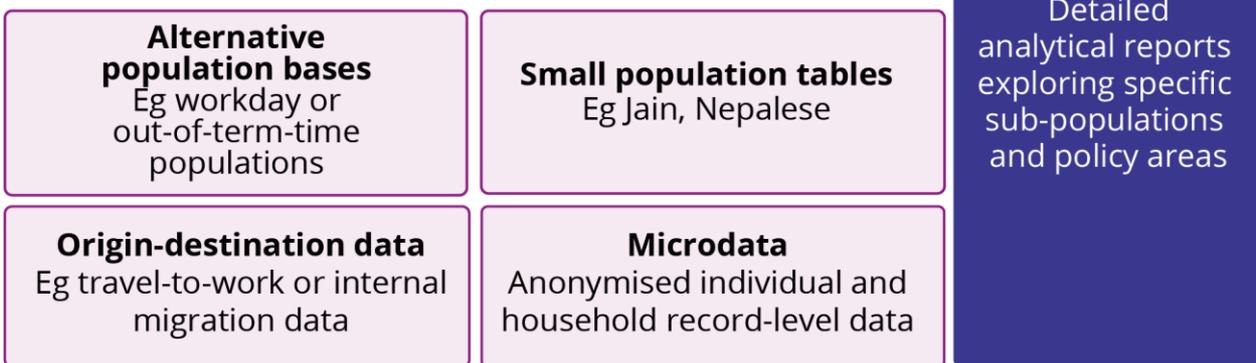
**Phase 1**



**Phase 2**



**Phase 3**



**Beyond March 2023**



## Geography

### Geographies included in releases

In late 2020, we carried out a separate consultation on [Census 2021 Output Geography Policy, products and services](#). Earlier this year, we also published our formal response.

The geographies we're using to support Census 2021 data is largely based on the 2011 approach. More information about that approach is available in our [overview of the various geographies used to support the production of 2011 Census statistics](#). This includes definitions and descriptions of OAs, our smallest statistical geography.

We plan to continue using OAs as the main geographical base for Census 2021 outputs. We will produce outputs for most supported geographies by aggregating OAs, on a best-fit basis. The method we used to decide the best-fit is described in the [Geography Policy for the Government Statistical Service](#).

We will review OAs based on Census 2021 population and household data. We will then merge or split them as needed to ensure they remain within the established population and household size thresholds. We expect around a 5% overall change between OAs compared to 2011, including some targeted re-alignment of OAs to better fit ward and parish boundaries.

We will base the main outputs for parishes on OAs. However, there will be some targeted re-alignment of OAs to better meet current ward and parish boundaries. We acknowledge that some parishes are smaller than OAs, and we're exploring how we could produce a summary set of information for these smaller areas.

Data for National Parks will use an exact geography in line with 2011. We also intend to produce simple population counts for postcodes, as we did following the 2011 Census.

We're aiming to make census outputs for key geographies available early in the release schedule, following the initial releases of Census 2021 data. This contrasts with the 2011 approach, which released geographical detail and type in stages. This is an ambitious aim, and we will make final decisions on this prior to publishing a final release schedule in February 2022.

If we do need to produce outputs for some of the key geographies later than others, we propose to prioritise estimates for administrative and statistical geographies down to OA level. This will include, for example, Local Authorities, Lower Layer Super Output Areas (LSOA) and Middle Layer Super Output Areas (MSOA). Following these releases, we will produce estimates for additional geographical areas, such as health areas and Parliamentary Constituencies.

## Boundaries

All Census 2021 outputs will reflect the geographies in place in May 2022, rather than as of Census Day, 21 March 2021. As a result, we expect that census outputs will reflect any ward changes that become effective for the planned elections in May 2022. This is consistent with the approach we took in the 2011 Census outputs.

To prevent disclosure of confidential information about people and workplaces, we will build upon the approach of the 2011 Census. This means that we will provide users with look-up tables that map each OA to a range of other geographies, on a best-fit basis. This will create best approximations of data for new geographical areas, such as changed wards.

For Census 2021, the guidance will suggest using higher levels of geography to build other geographies where possible, rather than adding all individual OAs together. This is because the small cell key perturbation method that we will apply to Census 2021 data has less impact on larger areas. These areas have larger counts and are likely to be less disclosive. This will minimise the slight differences in totals that statistical disclosure control creates. This extra layer of protection is needed to allow the flexibility for users to create their own tables.

We recognise that some geographical boundaries may change over time. Although the overall rate of change for geographies affected is relatively low, it can have a big impact in specific areas. Where there are significant boundary changes following the release of census outputs, we will consider providing some revised outputs as ready-made tables if users demonstrate sufficient need. For example, we may do this if there are significant changes made following the [Boundary Commission for England's \(BCE\) 2023 review of electoral boundaries](#). The BCE is due to publish its final report in June 2023.

Users may also wish to request [commissioned tables](#) for these new areas, if preferred. We will assess these on a case-by-case basis and charges for the cost of their production will apply.

**We'd like to know your views on the proposed approach for providing outputs to reflect boundary changes made after May 2022. In addition, we'd like to know if you have a need for data produced for different geographic boundaries.**

## Phase 1: Initial findings and topic summaries

### Initial findings

We aim to release initial census findings within 12 months of the census. These are likely to be rounded national population estimates for England and Wales.

We plan for these to be followed as soon as possible by more detailed unrounded population and household estimates. These estimates will be broken down by age and

sex, for England and Wales at local authority level. We will separately produce [mid-year population estimates](#) (MYEs) for the UK. These will be broken down by local authority, sex, and age.

## Topic summaries

We then plan to move into a staggered release of topic summaries. We intend to provide these largely as univariate tables. This means that they only include data from a single variable, for example, religion or general health. These are similar in concept to the Quick Statistics from 2011. We will be able to make the tables produced on a single variable available at the most detailed level of the classification.

During this phase, we aim to produce these tables down to OA level and for all geographies defined using the OA best-fit approach. If this is not possible because of operational issues, we will release this data as soon as possible after this phase.

We're proposing to cover topic summaries, in the order provided, for the following areas:

1. demography and migration
2. ethnic group, national identity, language and religion
3. sexual orientation and gender identity
4. health, disability and unpaid care
5. housing
6. labour market and travel to work
7. education
8. armed forces veterans

All tables in this phase will be populated based on the usually resident population of England and Wales. Tables will not include totals or subtotals.

Commentary, which will draw out insights from the data and provide supporting contextual information, will accompany each of these releases. This information could provide context to any significant changes in those topics since 21 March 2021, when people gave their answers to the census questions. However, it's too soon to know what additional contextual information we could provide this early in the release schedule.

We're currently exploring the products that we could use to provide the data and commentary, and we're conducting user research to inform their design. Those products are likely to include interactive data visualisations, which present the data in innovative and accessible ways for different types of user. If you'd like to be involved in the development of these products, please email us at [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk) for more information.

We've provided the proposed table specifications for the phase one releases in the "Topic summaries" tab in the 'Draft proposals for outputs data content' spreadsheet. This

spreadsheet is available from the “related links” section of [Census 2021 outputs consultation](#).

**We’d like to know if the recommended ordering of the topic summaries and the proposed table specifications meet your needs.**

## **Area profiles**

We will use the data from the topic summaries to create dynamic area profiles. These will pull together selected census information across a wide range of topics for a specific area. They will be available for geographies where we’ve published data. Area profiles will allow users to compare the local and national picture and will provide similar data to those contained in the Key Statistics in 2011.

A range of geographies will be available. We’re aiming to make these area profiles available for a range of geographical levels from the whole of England and Wales right down to OAs from the outset. If this is not possible because of operational issues, we will add additional levels of geography after they become available.

These area profiles are similar in concept to those available in [Nomis Area Reports](#).

We’ve provided the proposed area profile specifications in the “Area profiles” tab in the ‘Draft proposals for outputs data content’ spreadsheet, available from the “related links” section of [Census 2021 outputs consultation](#).

**We’d like to know if the area profile specifications meet your data user needs.**

We will base the main outputs for parishes, or communities in Wales, on OAs. In some cases, parishes are smaller than OAs. As a result, we cannot produce them using the OA best-fit policy. To provide information for these specific small parishes, we will consider producing an additional set of parish and community profiles alongside the main outputs. These would consist of a summary set of information for these smaller areas. We would produce these for all parishes and communities with a population of at least 50 people, similar to those produced from the 2001 Census. We could also provide additional, very basic information on the number of residents and households for any parishes below that threshold.

**We’d like to know if you need an additional set of profiles to give summary data for all parishes and communities.**

## Phase 2: Multivariate data for the usual resident population base

### Multivariate data

In this phase of the release schedule, we will publish the richer, multivariate data. Data users can use these outputs to combine variables and, through this, can fully explore and understand relationships between them.

We aim to produce these outputs, in the form of tables, down to OA level. During this phase, we will do this for all geographies defined using the Output Area best-fit approach.

In 2011, we provided this in two main sets of tables that were designed for different levels of geography. Generally, we made the Local Characteristics (LC) tables available for OAs and above, but these tables contained less detailed classifications of the characteristics they covered. Similarly, we generally made the Detailed Characteristics (DC) tables available for MSOAs and above. These tables contained more detailed classifications of the characteristics they covered.

The model for 2021 will be somewhat different than for 2011. We're introducing functionality for users to build their own tables, alongside the availability of a set of prescribed ready-made tables.

This functionality gives users flexibility in three main ways, which include:

- letting them choose whether to select a ready-made table or to build their own
- putting them in control of what census variables to combine when they're building their own tables
- giving them the ability to get more detail for smaller areas than was possible in 2011 where that does not risk the confidentiality of respondents

All tables in this phase will be populated based on the usual resident population of England and Wales.

Tables will not include totals or subtotals. We will apply the small cell key perturbation method to Census 2021 data as part of the disclosure control methodology. Data users should note that this may lead to slightly different estimates depending on how the tables were defined.

### Ready-made tables

In general, we will make ready-made tables in phase two available down to OAs. They will be available for all areas within the geographic level they're released for. This means that users can be sure that, when they request a table, we will provide it for all requested areas.

Many of the ready-made tables we're proposing are similar and comparable to the LC tables published following the 2011 Census. We are also proposing some additional tables, subject to assessment of the statistical disclosure risk of the data.

Firstly, we're proposing additional tables on multi-language and multi-religion households. In each case, we're proposing three new tables on:

- tenure
- ethnic group of the household reference person
- occupation of the household reference person

The household reference person is an individual within a household who acts as a reference point for producing further derived statistics. They can also characterise a whole household according to the characteristics of the chosen reference person. The [2011 Census glossary](#) explains how we choose the household reference person for a range of household types. We've included these tables to meet the need for more information on cultural diversity within households.

Secondly, we're proposing tables for the new questions on sexual orientation, gender identity and armed forces veterans.

The "Ready made tables" tab in the 'Draft proposals for outputs data content' spreadsheet contains the proposed specifications for all the ready-made tables. This spreadsheet is available from the "related links" section of [Census 2021 outputs consultation](#). These specifications include the geographical levels the data will be available for.

**We'd like to know if the ready-made tables' specifications meet your needs.**

### **Build-your-own tables using flexible datasets**

Being able to produce build-your-own tables from flexible datasets means that data users have the chance to combine variables in new ways. The build-your-own tables system takes users' requests, creates the outputs and applies disclosure control methods on the data. Finally, the system runs checks to determine whether the resulting outputs are potentially disclosive and, therefore, whether it can provide the data to the user or not.

This means, for any specific request of data, the result may be that the system can provide the data for some but not all areas requested. This is more likely to occur where users have:

- combined multiple variables
- requested the most detailed classifications
- requested smaller geographical areas

The creation of data using this capability will be an iterative process for users. It's expected that users will experiment with what the system can provide for each area. As such, the specifications for the build-your-own tables, which we will make available for users to choose from, are an indicative set of variables and classifications. It does not mean that users will be able to obtain the data that they request.

As a guide, if one aspect of a table increases in complexity, another aspect is likely to need to decrease in complexity. This is similar to the relationship between the 2011 Census LC and DC tables.

The "Build your own tables" tab of the 'Draft proposals for outputs data content' spreadsheet contains the proposed specifications for the flexible datasets. This spreadsheet is available from the "related links" section of [Census 2021 outputs consultation](#).

**We'd like to know if the build-your-own tables from flexible datasets specifications meet your needs.**

## Census analysis

We will publish commentary, interpretation and more detailed analysis of the new census data on the ONS website. These will complement new data releases and interactive data visualisations will accompany them. This will provide users with impartial and insightful commentary to better understand our population and their local community needs. It will also inform public policy and public service delivery.

In line with the data releases, census analysis will develop in complexity as we progress through the programme and new census datasets become available. As the level of detail in the published data increases, the depth of the analyses will increase, and the insights will become greater.

In earlier stages, we plan to provide analysis for the main population bases, including usual residents, households and dwellings. In later stages, the plans will also include analysis using alternative population bases, such as the workday and out-of-term populations. These population bases are discussed in detail in "Section 5: Population-base specifications".

Planned analyses in phase two aim to meet many of the analytical commitments made in the [Census 2021 White Paper in December 2018](#) that do not require linkage with other data sources. These analysis proposals are subject to review following user feedback and viability testing.

These analysis plans are discussed in more detail in “Section 3: Main changes to variables compared to the 2011 Census”. That section also includes questions on the extent to which these plans meet user needs. It’s too early to specify exactly when we will release different analysis products, so all analysis for phases one to three are discussed together.

**We’d also like to know your needs for analysis on all census topics.**

## **Phase 3: Alternative population bases, origin-destination data and microdata**

Prior to phase three, all data that we release will be based on the usually resident population. In the third phase of the release schedule, we will produce other data products.

This includes data for alternative and small populations and origin-destination flow datasets looking at migration and commuting patterns. We will also make microdata samples available for research purposes and continue to provide analysis about the published data.

This period will also include the introduction of the [commissioned table service](#) for Census 2021. Where it’s not possible for users to create the outputs they need themselves, users can commission us to produce and supply them. We will do this if the requested data passes statistical disclosure checks. There is normally a charge for this service.

### **Alternative population bases**

Alternative population bases allocate those counted in the census to different geographical locations where individuals could be counted. Examples of alternative population bases include:

- workplace
- workday
- out-of-term
- short-term resident

We discuss these products, and the impact of the coronavirus pandemic on the size of these populations, in “Alternative population bases” within “Section 5: Population-base specifications”. We’ve included questions on the extent to which proposals for alternative populations meet user needs.

## Small populations

Small population tables aim to provide information on identities about which users have specific needs. This may be to better understand that small population or analyse potential inequalities between that small population and the wider population.

These tables will be bespoke outputs, where the size of the population in that group means confidentiality constraints will limit the release of detail on the population in standard outputs. The tables will only include members of the chosen population, not the whole usually resident population. Examples of proposed small populations include Sikh, Cornish and Jain. This is in line with the approach taken for [small populations in 2011](#).

One of the important aspects of these outputs is that we produce them only when the number of people in that small population meets an agreed threshold. This minimises the risk of disclosing personal information, and often suits these small populations quite well as they're usually clustered geographically. We discuss these products in "Small populations" within "Section 5: Population-base specifications". In that section, we've included questions on the extent to which proposals for alternative populations data meet user needs.

## Origin-destination data

Origin-destination data are sometimes known as "flow" data. These data describe the movement of people from one location to another. In line with the 2011 Census outputs, we plan to release four different categories of origin-destination data in a combination of data visualisation products and data tables.

- Migration flow data: national and international migration during the year prior to Census Day.
- Workplace flow data: travel to work in the week before Census Day and method of transport.
- Second address flow data: location of second address in relation to an individual's usual residence or workplace.
- Student flow data: migration patterns of individuals living at student addresses one year prior to Census Day.

As per our approach in 2011, we will classify the origin-destination data in three ways: public, safeguarded and secure.

As a result of the impact of the coronavirus pandemic on population flows, public and safeguarded origin-destination data may contain less detailed characteristics than in 2011. This would be necessary to protect confidentiality. We mention impacts of the coronavirus pandemic on products in "Section 6: Taking a census during a period of change".

We will place the most detailed origin-destination data, classified as secure, in our Secure Research Service (SRS). We will make this available to [approved or accredited researchers](#). We're still considering the most appropriate means of making the safeguarded data available. However, the [UK Data Service](#) is likely to hold these. This is in line with previous censuses.

As a result of their complexity, we may produce these data after the public data tables. We will include more information in the full outputs release schedule, which we're planning to publish in early 2022. A fuller understanding of the scope and likely timescales will inform that schedule.

## **Microdata**

Microdata products are samples of anonymised records for individuals and households, which include a selection of associated census characteristics.

To make data available as widely as possible and to maximise benefits from the census, we plan to classify microdata in three ways: public, safeguarded and secure. This is the same approach we're using for origin-destination data. This approach strikes a balance between protecting confidentiality and ensuring data can be available for a range of users, from inquiring citizens through to expert analysts.

In 2011, we made a "Microdata Teaching File" available for public access. This contained anonymised records on a limited set of variables for a random sample of 1% of people in the 2011 Census output database. It's freely available for anyone to download under the terms of the Open Government Licence.

We intend to produce a similar, downloadable public access teaching file for Census 2021. This aims to meet the needs of users for an accessible, non-disclosive microdata file to improve the use of census microdata. It will also provide an educational tool of real data to assist with the teaching of statistics and social sciences.

The secure microdata samples have the highest level of detail and the largest sample sizes. We will also store them in our SRS. As a result, they will be available only to approved or accredited researchers to protect the confidentiality of personal information. We're still considering the most appropriate means of making the safeguarded data available, but the UK Data Service is likely to hold them.

Following user feedback, we also plan to produce a microdata file that will contribute to the University of Minnesota's Integrated Public-Use Microdata Series (IPUMS) project. This is an international project that brings census microdata together from over 100 countries.

The proposed safeguarded and secure Census 2021 microdata products are described in the next subsections.

### Public-access file

- Sample size: up to 1% of individuals
- Statistical unit: persons
- Lowest geography: region

### Safeguarded individual region file

- Sample size: up to 5% of individuals
- Statistical unit: persons
- Lowest geography: region

### Safeguarded individual grouped LA file

- Sample size: up to 5% of individuals
- Statistical unit: persons
- Lowest geography: grouped local authority

### Safeguarded household file

- Sample size: up to 1% of households
- Statistical unit: households
- Lowest geography: region

### IPUMS file

- Sample size: up to 1% of households
- Statistical unit: households
- Lowest geography: region

### Secure individual file

- Sample size: up to 10% of individuals
- Statistical unit: persons
- Lowest geography: local authority

### Secure household file

- Sample size: up to 10% of households
- Statistical unit: households
- Lowest geography: local authority

As a result of the complexity of these products, we may produce these data after the public data tables. We will include more information, based on a fuller understanding of

the scope and likely timescales, in the full outputs release schedule. We're planning to publish this schedule in early 2022.

## **Census analysis**

During phase three, we propose to continue producing analysis of the multivariate tables. We're also proposing to include detailed analytical topic and policy reports on small population and alternative population bases in our releases.

We will publish these reports on the ONS website. They will aim to provide users with the most relevant and insightful commentary for their communities. They will also aim to help users better understand their demographic make-up and inform policy and service delivery.

Analysis proposals are detailed in the supplementary document [Census 2021 analysis programme proposals \(PDF, 504.8KB\)](#). These analysis proposals are subject to review following user feedback and viability testing.

## **Beyond March 2023: UK data and further analysis**

### **UK data**

In addition to carrying out the census and producing outputs for England and Wales, we are responsible for collating data from across the four nations of the UK to produce UK outputs. These outputs provide totals for the UK, as well as comparable data for small areas across the UK.

Historically, the census has been taken on the same day in all four nations of the UK. We'd then combine data from each census, where the questions were harmonised, to produce UK statistics. We then used these data as the baseline for other UK-level statistics we produce.

However, each nation has its own operational, user and respondent needs. The coronavirus pandemic led to each country reviewing whether to conduct their census in March 2021. Following this review, we concluded that we should conduct the census on 21 March 2021 as planned. The Northern Ireland Statistics and Research Agency (NISRA) came to the same conclusion for Northern Ireland. This date allows us to use Census 2021 data as a baseline from which to monitor recovery from the pandemic.

The Scottish Government's review concluded that, because of the impact of the coronavirus pandemic, they should move Scotland's census to March 2022, to secure high-quality outputs. This means that there will be a one-year difference in reference dates and that the first results from Scotland will not be published before March 2023.

All the UK census offices are working closely together to understand how this difference in reference dates will impact UK-wide population and housing statistics. This is both in terms of the timing and scope.

Every year, we also produce [mid-year population estimates](#) (MYEs) for the UK, broken down by local authority, sex and age. Every 10 years, these mid-year population estimates are rebased using new census data. We will deliver UK population estimates for mid-2021, based on Census 2021 in England, Wales and Northern Ireland, and on the 2011 Census in Scotland. In the following year, the Scotland mid-year population estimates will also be rebased using Scotland's 2022 Census data.

While we're confident that we will be able to produce some UK outputs, it will be more difficult in other areas. These include the four areas outlined here.

#### [Workplace populations and out-of-term populations](#)

There are likely to be some minor impacts on England and Wales outputs for workplace population and out-of-term populations. In both cases, flows into areas from Scotland will be missing.

#### [UK origin-destination outputs](#)

We've explored the options for producing UK origin-destination outputs. We've concluded that it's unlikely that we can combine the data in a meaningful way to produce a single reliable set of UK origin-destination outputs. As a result, we will make origin-destination outputs for 2021 available for England, Wales and Northern Ireland combined. The National Records of Scotland will publish data for Scotland separately.

#### [Second address flow data](#)

Second address flow data will continue to be available for England and Wales only, as Scotland and Northern Ireland do not collect this information.

#### [Areas with a difference in reference dates](#)

There are other examples where there's a potential for a high rate of change between reference dates, including health and labour-market data.

To ensure the UK-wide population statistics we publish meet data needs and are of a high enough quality, we're engaging with UK census data users through our working groups. We're considering factors such as the level of harmonisation between nations and the impacts of the difference in reference dates.

For example, we've been exploring the need for UK population estimates as at March 2021 or March 2022. We've progressed this work outside of this consultation, as a result of the short timescales. So far, we've not identified a need for these products in addition

to the mid-year population estimates. Please contact [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk) if you want to be involved in this process.

The full Census 2021 Outputs Prospectus will include more detail on our plans for UK data outputs. We plan to publish this by February 2022. We will provide guidance to users on comparing data from different nations. We're also considering if and how to present data for England, Wales and Northern Ireland together.

**We'd like to know if you need data for England, Wales and Northern Ireland to be published alongside each other, irrespective of the feasibility of producing UK outputs.**

## Census analysis

Census analysis after March 2023 will generally focus on more complex projects that require linking Census 2021 data with other data sources. Other sources include previous censuses, the ONS Longitudinal Study, and other government surveys and administrative sources. Through this we can undertake new, innovative and deeper analysis of census data to inform policy and service delivery.

These analytical projects will complete any commitments made in the [Census 2021 White Paper in December 2018](#) that we've not met in previous phases.

For example, we acknowledge the great interest in income data and the value that this can provide in planning, policy development and evaluation. We've been working with tax and benefits data from the Department for Work and Pensions (DWP) and Her Majesty's Revenue and Customs (HMRC) to develop small-area income data that can be linked with the data collected in Census 2021. We've already produced some research outputs, alongside other [research on Population Characteristics](#) to demonstrate the potential of this approach.

We're continuing to develop these research outputs and published [our latest set of experimental statistics on small area income distributions](#) in June 2021. These statistics were the first to incorporate self-employment income. We currently plan to publish a further update in 2022. We're still working towards the aim of directly linking this data to census data in the future.

We include details of our analysis proposals in the supplementary document [Census 2021 analysis programme proposals \(PDF, 504.8KB\)](#). We will continue to develop and review these plans in light of evolving user needs, particularly with respect to the pandemic and feasibility testing. User engagement, including the response to this consultation, will steer the development of these plans. We discuss analysis planned for beyond March 2023 separately from the phase one to three analysis. This is because it generally requires linkage with other data sources, including from Scotland's census.

As we progress the transformation of the UK population and statistics system, we're working to put administrative data at the core of our population, migration and social statistics systems. Our aspiration is to be as radical, ambitious and inclusive as possible to improve the quality, frequency, representativeness, relevance, coherence, accessibility and timeliness of these core statistics.

## Section 3: Main changes to variables compared to the 2011 Census

The census has always changed with the times to better meet user needs, through:

- making better use of technology
- changing and replacing existing questions
- introducing new questions where needed

In the 2015 consultation, [The 2021 Census – Initial view on content for England and Wales](#), we asked users of census statistics for their views on the topics that the Census 2021 questionnaire might cover. The questions asked in the Census 2021 questionnaire reflect the findings of that consultation.

This section describes the changes in which variable classifications will be available, including some cases where there will be slightly less detail. It will also describe cases where we've asked new questions and will illustrate the range of detail and insights that will be available to users.

Full details of the work that we carried out to determine the questions for Census 2021 are available on the [Question development](#) pages of the ONS website. In most cases, the questions have remained the same. Where there have been changes, these mostly provide either more or the same detail. However, in a few instances, less detail is collected in Census 2021 than was collected in the 2011 Census.

Where the same information is collected, but through a modified question, this is not included in this consultation as it does not impact the outputs. For example, in the 2011 Census there was a single, complex question on qualifications held. For Census 2021, we've asked separate questions on degree-level or above qualifications and on apprenticeships. However, this has not impacted the variable classifications. Similarly, we've not changed the national identity question and response options, but we've changed the order in which they're presented in England.

Classifications for questions that have not changed will remain largely the same, except where analysis of write-in responses indicates a change is needed.

We also reviewed and redeveloped all question guidance. In some cases, this has slightly impacted the data being collected. For example, there has been a change to the guidance on how to answer the sex question. As these changes do not impact the classifications, we do not discuss them here.

Finally, on the electronic questionnaire, we implemented search-as-you-type and address look-up functionality for many sociocultural questions. We did this to make it easier for respondents to self-define when completing Census 2021 online. Our [Search-](#)

[as-you-type and address look-up functionality for Census 2021](#) report, published on 21 December 2020, provides more information. Inclusion of a group on the search-as-you-type list does not necessarily mean that outputs will be available for that group.

We expect that societal change since the 2011 Census, and the implementation of the search-as-you-type functionality, will have affected written responses to questions on cultural background. For example, there are likely to be changes in how people listed their main language or national identity. We also added a write-in option for those selecting “African” in the ethnic group question, which will collect new data on this population. We have not included the detailed classifications for these questions in the specification. This is because the written responses in the Census 2021 questionnaire and the user needs raised in this consultation will inform them.

For each of the topics discussed in this section, we’ve detailed the proposed variable classifications in the ‘Draft proposals for outputs data content’ spreadsheet. This spreadsheet and is available from the “related links” section of [Census 2021 outputs consultation](#). We’re still developing the most detailed classifications based on Census 2021 responses. As a result, this spreadsheet does not include the detailed variable classifications for the questions on sociocultural background. We’ve also detailed the analysis proposals in the [Census 2021 analysis programme proposals \(PDF, 504.8KB\)](#) document.

## Questions with detail removed

In a small number of cases changes to the questions asked resulted in less detail being collected. In all cases where the detail had decreased, we assessed that there was insufficient user need to continue to collect this information. This also offset the increased burden put on respondents to provide additional detail, such as on the new topic of armed forces veterans. We summarise some of the important changes under the following seven topic headings.

### Number of rooms

We removed this question. Instead, we’ve sourced this information from Valuation Office Agency (VOA) data. Reports on our work towards achieving this are available on the [housing characteristics page on the ONS website](#).

### Migration intention to stay

We combined the options, differentiating between those intending to stay in the UK for three to six months and six to twelve months, into a single response option.

## **Qualifications**

We collected all the 2011 Census categories but in a restructured series of questions, except for “Professional qualifications”. We asked respondents to choose the nearest equivalent for any qualification not listed.

## **Unemployment history**

We removed the option to write in the year last worked. We replaced it with options that separate those who are not in work but have previously worked into two groups. These two groups are the short-term unemployed, who have worked within the last year, and the long-term unemployed, who most recently worked over a year ago.

## **Activity last week**

We removed the option “On a government sponsored training scheme”. We did this because the International Labour Organisation (ILO) definition of employment has changed since 2011 to exclude unpaid training as a form of economic activity.

## **Communal establishment resident age group and types**

On the communal establishment form, we removed questions on age groups and types of resident catered for. Data on age groups can be derived from the information collected on residents’ individual questionnaires.

## **Rough sleepers**

In 2011, there was an output on “Rough Sleepers” based on a count completed by census field staff. We’ve removed this because of concerns about the quality of the data collected. We discuss options for alternative outputs about the homeless population in “Section 4: Proposals for potential new derived variables”.

## **New questions and associated classifications**

### **Sexual orientation and gender identity**

We asked the questions on sexual orientation and gender identity for the first time in 2021. The questions were voluntary, and we only asked these questions to those who were aged 16 years and older.

We propose to produce variables representing the response options in these questions, subject to the data passing statistical disclosure control assessments once the final data are available. Because of the small size of some of these groups, we’re unlikely to publish the data at Output Area level.

For sexual orientation, the proposed classification is:

- Straight or heterosexual
- Gay or lesbian
- Bisexual
- Other

For gender identity, the proposed classification is:

- Yes: Gender identity is the same as sex registered at birth
- No: Gender identity is different to sex registered at birth

In both cases, there would be a further “not stated” category, which would include any responses that could not be coded and those respondents who did not answer. In line with how we’ve previously treated the voluntary question on religion, we will not impute answers to these questions.

We’ve included the proposed specifications for the ready-made tables, including these topics, in the “Ready made tables” tab of the ‘Draft proposals for outputs data content’ spreadsheet. This spreadsheet is available from the “related links” section of [Census 2021 outputs consultation](#). These specifications include the geographical levels the data will be available for.

Given the potential for disclosure being higher for these populations, we need to carefully assess what additional information we can publish and for what geographical levels.

In addition, we’re developing a more detailed classification using the write-in responses from the Census 2021 questionnaire. Because of the small numbers of people likely to be in each category within a more granular classification, we’re considering how best to output this data.

**We’d like to know your views on our plans for producing statistics and analysis using responses to the sexual orientation and gender identity questions.**

## **Armed forces veterans**

For Census 2021, we asked a new question on previous service in the UK armed forces. We only asked this question to those who were aged 16 years and older. We created a new standard variable to provide a range of outputs on the population who have previously served in the UK armed forces. The proposed classification is:

- Previously served, Regular
- Previously served, Reserve
- Previously served, Regular and Reserve

- No code required

The “No code required” category includes both those who have never served and those who are currently serving. We’re also investigating the feasibility of two new variables.

The first proposed new derived variable would indicate if the household reference person has previously served in the UK armed forces. This would have the same categories as the standard variable described above.

The second proposed new derived variable would indicate the number of former UK regular and reserve armed forces personnel resident in a household. The simplest proposed classification is:

- No former UK armed forces personnel in household
- One or more former UK armed forces personnel in household

However, we’re considering if we should also include a second, more detailed classification. This might, for example, separate out those with one, two, three, four, five, or six or more former UK armed forces personnel in the household. We will consider the upper threshold of this classification based on Census 2021 data, once available, to take into account statistical disclosure risks.

This variable could be cross-tabulated with the number of people in the household to understand the proportion of the household that are former UK armed forces personnel.

**We’d like to know your views on our plans for producing statistics and analysis using responses to the question about armed forces veterans.**

## **Proxy answer question**

Another new question in 2021 asked all respondents if they were completing the questions for themselves or for someone else. This provides an interesting new analysis variable, allowing researchers to investigate if there are trends in the data created by who completed the form.

This variable will be in the microdata products held within our Secure Research Service (SRS). We’re considering if we should include it in the main data releases.

**We’d like to understand if users would use information on if a respondent answered for themselves or for someone else in their analysis.**

## Other significant changes to output classifications

### Ethnic group classifications

In 2011, there were 18 response options in the ethnic group questions. These were divided into five sections. Within each section, there was an “Other” option allowing respondents to express their identity how they wished. This meant that we could further disaggregate ethnic group data based on the written responses in each section.

For Census 2021, the ethnic group question includes a new tick box for “Roma” under the “White” category. It also contains an option to write in a more detailed response when selecting the existing “African” response option.

We will update the ethnic group classifications to take into account these changes. We’ve shared the current proposed classifications in the “Ethnic group” tab of the ‘Draft proposals for outputs data content’ spreadsheet. This spreadsheet is available from the “related links” section of [Census 2021 outputs consultation](#). As we’re still developing the most detailed classification, this spreadsheet does not include the most detailed ethnic group classification. This will use data from the write-in responses.

We’re carefully considering how to best present the classifications. For example, we need to decide how to reflect the slight difference in wording for high-level categories between the Wales questionnaire and the England questionnaire. In Wales, the category “Asian or Asian British” is expanded to read “Asian, Asian Welsh or Asian British”. Similarly, the category “Black, Black British, Caribbean or African” is expanded to include “Black Welsh”.

We’re also considering how to order the classifications in the outputs. In the proposed classifications, we’ve ordered the groups as follows:

- Asian, Asian Welsh, Asian British
- Black, Black Welsh, Black British, Caribbean, African
- Mixed or multiple ethnic group
- White
- Other ethnic group

For more detailed classifications we propose to list the groups alphabetically within the higher-level categories.

**We’d like to know your views on the proposed classifications for the ethnic group variable, including how they’re presented.**

**We’d also like to know your needs for analysis on this topic, especially around particular subgroups you’re interested in.**

## Age classifications

In 2011, we used around 55 different age classifications in the standard outputs. The different age classifications aimed to meet users' needs for different topics. For example, these included employment or education. To reflect our updated approach for dissemination, particularly the new capability for users to build their own tables, we've reviewed the classifications needed for each topic.

This work has resulted in a proposal to reduce the number of classifications by almost half. We've done this by creating simple groupings of classifications that get progressively more detailed for each topic. We hope this will make navigating the different classifications easier for users.

We've detailed these age classifications in the "Resident\_Age" tab of the 'Draft proposals for outputs data content' spreadsheet, available from the "related links" section of [Census 2021 outputs consultation](#).

**We'd like to know if this consolidated list of age classifications will meet your needs.**

## Country of birth classifications

In 2011, we divided the Country of Birth output groupings for EU member countries into two groups according to when those countries joined. These groups were:

- countries that joined the EU before 2001
- countries that joined the EU between 2002 and 2011

Since 2011, Croatia has joined the EU and the UK has left the EU. None of the other classifications based on countries, such as national identity or country of second address, contain this grouping. Therefore, we do not plan to reproduce this grouping in Census 2021 outputs. Instead, the three main categories that we propose are:

- Europe: United Kingdom
- Europe: Ireland
- Europe: Other Europe

We will split these into more groups in the detailed classifications.

We've included the groups contained within each country classification in the "COB" tab of the 'Draft proposals for outputs data content' spreadsheet, available from the "related links" section of [Census 2021 outputs consultation](#).

**We'd like to know if the proposal to group all EU member countries as "EU member countries" will meet your needs.**

## **Other classifications with more detail added**

To reflect changes in questions and response options in Census 2021, we've updated further classifications. In general, users can still get equivalents to 2011 data. We've provided full details in the 'Draft proposals for outputs data content' spreadsheet, available from the "related links" section of [Census 2021 outputs consultation](#). We've also summarised the changes under the following six topic headings.

### [Type of accommodation](#)

We've added a new category of "a flat, maisonette or apartment: part of another converted building (for example, former school, church or warehouse)".

### [Central heating](#)

We've split the category "Gas" into "Mains gas" and "Tank or bottled gas".

We've split the category "Solid fuel" into "Wood (for example, logs, waste wood or pellets)" and "Solid fuel (for example, coal)".

We've split the category "Other central heating" into "Renewable energy (for example, solar thermal or heat pumps)", "District or communal heat network" and "Other".

### [Marital status](#)

A second question followed the marital status question. That second question asked who the marriage or civil partnership was with. This included the options "Someone of the opposite sex" and "Someone of the same sex".

### [Second address](#)

We've split the category "Other" into "Partner's address" and "Other".

### [Disability](#)

We've reworded the questions on physical or mental health conditions or illness to align with the [GSS Harmonised standard questions](#). This means the census now collects data more closely aligned with the definition of disability in the [Equality Act 2010](#).

As part of this process, we split the question into two parts. The first part asked if the respondent had a condition or illness, and the second part asked if it limited their day-to-day activities.

In the 2011 Census, the category "No" encompassed two groups of people. This included those without a condition or illness and those with a condition or illness that did

not impact their day-to-day life. Splitting the question into two parts means these two groups can now be identified separately.

### Unpaid care

We increased the range of response options.

We split the option “Yes, 1 – 19 hours a week” into two categories, which are “Yes, 9 hours a week or less” and “Yes, 10 to 19 hours a week”.

We split the option “Yes, 20 to 49 hours a week” into two categories, which are “Yes, 20 to 34 hours a week” and “Yes, 35 to 49 hours a week”.

**We’d like to know your plans for using the additional detail allowed by the new response options in Census 2021 questions.**

## **Section 4: Proposals for feasibility work to derive new variables**

In the previous section, we discussed the new and changed questions and shared our proposals for the new or changed classifications.

In this section, we consider how to combine existing data, often from multiple questions, in new ways to meet a wider range of current user needs. Each of our proposals below is at an early stage of development, and we aim to further understand user needs through this consultation. Following the consultation, we will use the feedback to prioritise the proposals where we have evidence of user need. However, the release of information remains dependent on research proving the resultant variable is of sufficient quality to meet data users' needs.

Some proposals are about understanding the impact that the coronavirus pandemic has had on different communities and how we all live. Other proposals relate to alternative wider potential data needs, such as better understanding the student population. We've grouped these into three areas, which include:

- education and employment
- health and living arrangements
- accommodation type and vacant addresses

### **Education and employment**

We're investigating seven new variables linked to education and employment.

#### **Route to highest level of education**

We're aware that an important user need linked to education data is to better understand the qualifications route taken. For example, are the qualifications gained academic, vocational or a mixture of both?

In the 2011 Census, we produced a qualification gained variable. However, because of the complexity of this variable, its use was limited. For 2021, we're investigating producing a consolidated variable that will allow users to interpret the information more easily.

Before developing a proposal for this classification, we need to further understand what information users need about the routes taken through education.

## Students

In the 2011 Census, we defined students as respondents in full-time education who were aged four years and over. For 2021, we've slightly raised the age limit to those who are aged five years and over.

We're also proposing an additional variable that is more specific. This could be all students in full-time education who are aged 18 years and over. Alternatively, we could include only independent students. We would do this by defining them as students in full-time education who are aged 18 and over and not living with parents.

## Not in employment, education or training (NEETs)

This indicator variable would highlight the population aged 16 to 24 years who are not in employment, education or training. As the Census 2021 questions are harmonised with the Labour Force Survey (LFS), we'd apply the same definitions.

## Temporarily away from work

In 2011, the question on activity last week was primarily used to find out if someone was economically active. We're exploring the creation of "Temporarily away from work ill, on holiday or temporarily laid off" as an indicator variable. This would also include those away from work because of furlough, self-isolation or quarantine.

We're also investigating the production of an "On maternity or paternity leave" indicator.

## Key or critical worker

This proposed indicator would show people whose occupation was critical to the response to the coronavirus pandemic, as defined by the UK government. We'd define this in line with our other [publications on critical workers](#).

## Skills mismatch

Skills mismatch is where a person's educational level is significantly different to the average level of qualification within their occupation. For example, a person would be defined as under-employed in their role if both:

- their highest level of education is degree level or above
- the average qualification for their occupation is at level 3, for example an NVQ level 3 or one A level.

In 2011, we did not formally output this information, but we did publish reports around the topic. These include reports such as the [Graduates' labour market outcomes during the coronavirus \(COVID-19\) pandemic: occupational switches and skill mismatch](#). Some of these publications use the statistical methodology that the [International Labour](#)

[Organisation](#) (ILO) uses. We'd derive this variable using highest level of qualification and occupation.

## **Economic risk created by the coronavirus pandemic indicator**

The [Business Impacts of Coronavirus \(COVID\) Survey](#) has identified industries at most risk in the pandemic. The survey identifies different risks, such as risk of unemployment, for those working in different sectors.

This variable will apply the findings of that survey to census data, to identify populations at financial risk because of the pandemic. For example, a person working in hairdressing could be identified as being at financial risk because of the pandemic. It may be possible to produce a more detailed classification than this binary classification.

## **Health and living arrangements**

We're investigating four new variables for health and living arrangements.

### **COVID-19 health risk**

We could potentially define an output based on the [vaccination priority groups](#) by using the self-reported general health status to identify at-risk groups. For example, [previous research](#) has shown that there's a strong relationship between a response of "limited a lot" to the question on how conditions limit your activities, and people being [clinically extremely vulnerable](#).

### **Houses in multiple occupation (HMO)**

The [government's definition](#) states that a dwelling is an HMO if at least three tenants live there, forming more than one household, and you share toilet, bathroom or kitchen facilities with other tenants. If a dwelling meets this definition and five or more tenants live there, then it is classed as a large HMO. There are different [legal obligations](#) for landlords of HMOs and large HMOs.

In 2011, we didn't output data meeting these definitions. For 2021, we're investigating the need for outputs to fit this definition. Initial feasibility work suggests we may need to define an HMO as having three or more unrelated people live in the household. This would help ensure we do not count multigenerational households as HMOs.

### **Multigenerational households**

A multigenerational household is defined as any household with more than two generations resident. For example, this could be when children of any age, a parent(s)

and a grandparent(s) live together. This indicator would identify households that met that definition.

## **Living apart together**

We're investigating the feasibility of producing a variable indicating the population who live separately to their current partners. The populations we could identify are:

- those who are married or in a civil partnership but not living with the person that relationship is with
- those who are spending 30 or more days a year at a partner's address, whether married or in a civil partnership or not

## **Accommodation types and vacant addresses**

We're investigating four new variables linked to accommodation type and vacant addresses.

### **Care home resident**

There's an existing variable on the type of communal establishment a person is resident in, but this will not provide targeted information on care home residents. This separate indicator would denote if a person was resident in a care home.

To complement this product, we've produced a new age classification. This classification has the following categories.

- Aged 0 to 64 years
- Aged 65 to 69 years
- Aged 70 to 74 years
- Aged 75 to 79 years
- Aged 80 to 84 years
- Aged 85 to 89 years
- Aged 90 years and over

### **Type of vacant address**

We have data from the census collection process that we could potentially use as an indication of whether an address is vacant, a holiday let or a second home. We could then check this against administrative data. We're investigating if this is of sufficient quality to use in analysis of census counts of vacant addresses.

### **Resident in a mobile or temporary structure**

This proposed indicator would show the population who live in a mobile or temporary structure. For example, this might include a boat or caravan. This could indicate a more

transient population. We also have some operational data on the type of mobile or temporary accommodation people were usually resident in on Census Day. We're investigating whether this is of sufficient quality to use in analytical products.

### **Homeless (including people sleeping rough and 'sofa-surfers')**

There are many sorts of homelessness, and we're developing variables to provide data on those we can. The communal establishment questionnaire contains the response option "hostel or temporary shelter for the homeless", which may produce some data on 'rough sleeper' populations.

Local authorities and charities helped our field force to distribute census forms and access codes to the homeless population. They did this through night and day shelters to encourage this group to respond.

We also added a response option "Staying temporarily (no usual UK address)" to the individual census forms. This may allow analysis of people with no usual address staying in B&Bs and hotels at the time of the census.

It may also be possible to use the household questionnaire to identify households that have temporary residents with no other address. This would not be a count of individuals, only households, and we're investigating the usefulness of this in analysis of 'sofa surfers' and other inadequately housed populations.

We're investigating the quality of data collected on the homeless population through these response options. We're also looking at how we can use administrative data collected as part of the pandemic to supplement census counts.

**We'd like to know if you have analytical needs for any of the proposed new variables or indicators. For those you have needs for, we'd like to understand the level of detail you'd require and what definitions we'd ideally use in their creation.**

**We'd also like to know if you have any other needs for new derived variables using Census 2021 data.**

## Section 5: Population-base specifications

### Main population bases

In [The Census 2021 – Initial view on content for England and Wales](#) consultation in 2015, we assessed that we needed to produce the same output bases as in 2011. We've included details of that assessment in the report [Outputs and enumeration bases](#).

Definitions for most Census 2021 population bases will remain largely the same as those we used in the 2011 Census. We've included details of these changes in the report [Output and enumeration bases: residential address and population definitions for Census 2021](#).

The main population bases for Census 2021 are:

- usual residents
- households
- usual residents in households
- communal establishments
- usual residents in communal establishments
- household reference persons (HRP)

We propose that users will be able to build their own tables using flexible datasets, based on these populations. A proposed list of possible variables that we could make available in each flexible dataset is in the 'Draft proposals for outputs data content' spreadsheet. This spreadsheet is available from the "related links" section of [Census 2021 outputs consultation](#).

### Secondary population bases

We will provide outputs based on the following population bases.

- Dwellings
- Families
- Dependent children

These populations will not initially be available as flexible datasets because of their complexity. However, we will provide them as ready-made tables. There may be scope for developing flexible datasets based on these populations as the outputs develop over time.

### Place of residence

In all of these population bases, usual residents are counted at their place of residence.

A usual resident of the UK falls into one of the following categories. They're anyone who, on 21 March 2021:

- is in the UK and has stayed, or intends to stay, in the UK for 12 months or more
- has a permanent UK address and is outside the UK and intends to be outside the UK for less than 12 months

The statistics show usual residents at the location they considered to be their usual place of residence on Sunday 21 March 2021. A UK resident's usual place of residence is generally their permanent or family home or the address in the UK at which they spend most of their time.

We counted students at their term-time address, with some details also collected at their home address. We advised students that a term-time address was the address that they intended to stay at regularly during term time in the 2021/22 academic year, even if they were not there on Census Day. Further information is provided in "Section 6: Taking a census during a period of change", in the "Place of residence" subsection.

## Alternative population bases

Alternative population bases allocate those counted in the census to different geographical locations where individuals could be counted. The data on alternative population bases that we made available following the 2011 Census are accessible on the [Alternative population statistics](#) pages of the ONS website.

Alternative population bases are those that:

- move usual residents to different locations based on their characteristics, and/or
- report on a population that is not usually resident

We discuss each alternative population base in more detail in the following paragraphs of this consultation document. There's the potential for disclosure to be higher for alternative population bases. As a result, we need to carefully assess what additional information we can publish and for what geographical levels.

## Workplace

The workplace population is an estimate of the population working in an area. It includes usual residents aged 16 to 74 years whose usual place of work is in the area. We include people who work mainly at or from home, or do not have a fixed place of work, in their area of usual residence.

The workplace population of an area excludes anyone in one of the following population groups. These include:

- those living in England and Wales but working in Scotland, Northern Ireland, outside the UK or on offshore installations
- those with a place of work in England and Wales but who are not usually resident in England and Wales
- short-term residents

The definition for the workplace population remains unchanged. However, we're anticipating that because of the coronavirus pandemic more people will be working from home, on furlough or otherwise not working. We will still count people at their workplace address who state that they're "temporarily away from work". There may also have been an increased number of jobs without fixed workplaces. One example is delivery jobs that are also counted at the respondent's home address.

We're proposing to produce outputs on the workplace population at Middle Layer Super Output Area (MSOA) level and above. These will include some information on basic characteristics of the population.

## **Workday**

The workday population is an estimate of the population during the working day. It includes everybody who works in an area, wherever they usually live, and all respondents who live in the area but do not work.

The workplace population of an area excludes anyone in one of the following population groups. These include:

- those living in England and Wales but working in Scotland, Northern Ireland, outside the UK or on offshore installations
- those with a place of work in England and Wales but who are not usually resident in England and Wales
- short-term residents

Workday population estimates are also a geographic redistribution of the part of the usual resident population who are in work, allocated to their place of work. The data include the workplace estimates plus those usually resident in the area who are not working.

The definition for the workday population remains unchanged from 2011. It's defined as the non-working and homeworking usual residents in the area, plus those working in the area but resident elsewhere. The coronavirus pandemic may have changed this distribution because of larger numbers of people staying at home to work, being on furlough or not working.

We're proposing Output Area (OA) level outputs. These will include some information on basic characteristics, creating a similar series of estimates to 2011.

### **Out-of-term population**

This is the usually resident population, redistributed to their out-of-term address if they have one. As a result, the difference relates to the location of some students and schoolchildren.

We will define this population in almost the same way as in 2011. The only slight difference will be because of a change in [how we count flexi-boarders](#). For 2021, only schoolchildren at boarding school who stay at their term-time address for four or more nights per week in general should be recorded as usually resident at both their term-time address and their home address. Those who stay for fewer than four nights a week should be recorded as usually resident at their home address only.

Because of the pandemic, there may be more students without term-time addresses and fewer students entering or leaving England and Wales to study. This will potentially lead to less difference between the usual resident and out-of-term populations than in other years.

We're proposing OA-level outputs. These will include some information on basic characteristics, creating a similar series of estimates to 2011.

### **Short-term population**

Outputs for short-term residents provide characteristics of people who were not born in the UK and who intend to stay in the UK for 3 to 12 months. Because of the restrictions on travel that the coronavirus pandemic caused, this population may be much smaller than in 2011.

As a result of the anticipated small counts, we expect to provide users with a series of regional short-term resident tables for 2021, with accompanying analysis. These will include some information on basic characteristics.

### **Second address**

Second-address estimates provide information on the number of people who have a second address outside of the local authority where they are usually resident.

This population is redistributed to addresses that respondents recorded as their second address. This is an address they spent more than 30 days a year at, including:

- holiday homes
- partners' addresses
- armed forces bases

- students' out-of-term-time address

Because of the possible small counts, we're proposing mostly local authority outputs for 2021. These will provide some basic characteristics. To ensure we protect confidentiality, it may be necessary to use higher geographies.

**We'd like to know what data tables and analysis you need for alternative population bases.**

## Small populations

Following the 2011 Census, we offered small population tables. These tables provided information on the key characteristics of people in specific small population groups where numbers may not have supported multivariate analysis at smaller geographies. These groups were defined by their cultural background. Examples of these characteristics were individuals of:

- an ethnic group
- a country of birth
- a religion
- a national identity

These small population tables reflected identities about which users have specific needs. This may have been to better understand that small population or to analyse potential inequalities between that small population and the wider population. We published the small population tables on the Nomis [Small population – 2011 Census](#) page.

We only produced these small population data for geographic areas in which the small population being counted exceeded a set threshold. We only included the areas in which the population exceeded those thresholds in each table.

At a minimum, we aim to produce tables for the six small populations that we produced tables for in 2011. We will make these available at the MSOA level. In 2011, we produced 17 tables for each of the six populations. The populations in 2011 were:

- National identity - Cornish
- Ethnic group - Kashmiri
- Ethnic group - Nepalese
- Ethnic group - Sikh
- Religion - Ravidassia
- Religion - Jain

For the Sikh and Jain tables in Census 2021, the small population outputs will use new derived variables. The variables include responses in either the ethnic group or the religion question.

We will also produce local-authority-level tables by five-year age bands by sex, for a total of 30 population groups for local authorities. We will do this for the same 30 population groups as in 2011. These were all ethnic groups or countries of birth.

For ethnic group, these were:

- Afghan
- Filipino
- Greek
- Greek Cypriot
- Kurdish
- Latin Central South American
- Nepalese
- Polish
- Somali
- Sri Lankan
- Tamil
- Turkish
- Turkish Cypriot

For country of birth, these were:

- Bangladesh
- Bulgaria
- Cyprus EU
- France
- Ghana
- India
- Ireland
- Jamaica
- Nigeria
- Pakistan
- Philippines
- Poland
- Romania
- Somalia
- Sri Lanka
- South Africa
- Turkey

For Census 2021 outputs, we aim to make more information available on small populations. This includes consideration of increasing the volume of tables produced on small populations in three ways.

Firstly, we could provide data on additional small populations defined based on their ethnic group, religion, country of birth or national identity. The data could use responses from multiple questions to determine these populations. The data could also include small populations that are now possible because of changes to the questions, such as “Roma” or ethnic groups within the “African” subgroup.

Secondly, we could provide further detail on small populations. This detail could be variables such as different cross-tabulations or geographical breakdowns.

Thirdly, we could provide data on groups defined using other characteristics. An example of such a characteristic could be, for instance, the population who have British Sign Language (BSL) as a main language.

However, because of potential for disclosure being higher for small populations, we need to carefully assess what additional information we can publish.

**We’d like to know what data tables and analysis you need in relation to existing and potential new small populations.**

## Section 6: Taking a census during a period of change

We're currently in a period of significant economic and societal change. This change is not only because of the coronavirus pandemic. It's also a result of Brexit and a variety of other political initiatives. These factors have impacted different groups in different ways. In this section, we discuss the impacts of taking a census in such a period of change.

At the onset of the pandemic, we were already preparing for the census. During this time and the run up to March 2021, we carried out scenario planning and undertook regular readiness assessments. We did this to ensure that we were operationally ready to run Census 2021. We published a [series of articles](#) detailing this process. These articles outlined our plans for running the Census 2021 operation safely and securely during the coronavirus pandemic.

We also considered if running a census during a pandemic would result in the data that users need. We concluded that up-to-date data on the population are needed to provide a baseline from which to monitor recovery from the pandemic. Census 2021 data will also enable policy makers to make evidence-based decisions on what steps to take to support that recovery. This will help shape services for years to come.

We assessed the Census 2021 questionnaires to understand the potential impact of the coronavirus pandemic on how respondents answer questions. Where we discovered issues, we made changes to resolve them. As the questions had already been finalised in the [census legislation](#), mitigation was limited to additional guidance on the online form and within the online help pages of the Census 2021 website. We could not change the question wording or the response options at that point in the development process.

The outcomes of this work are detailed in the publication [Coronavirus \(COVID-19\) question guidance report for Census 2021](#). We considered if any questions referencing "usual" circumstances should look back to before the pandemic, look at the current circumstances or look forward to the expected situation after the pandemic. In general, our stakeholder engagement found that data users wished to continue to collect data as of Census Day, as a snapshot in time. For example, key stakeholders confirmed that asking respondents for their travel-to-work situation on Census Day would be the best option given the circumstances. These stakeholders included the Department for Transport and Transport for London. However, they also expressed that, as data on travel to work are used in long-term transport and infrastructure planning, data as at Census Day would not meet all of their needs.

From this feedback, we agreed on the consistent messaging that the census provides a snapshot in time. We also asked respondents to answer about their current circumstances across all questions. However, in a limited number of scenarios, where

someone was temporarily away from work, a valid answer based on current circumstances was not possible. Examples of this included questions on hours worked or travel to work. In these cases, advice was provided on the online help pages that asked people to provide details for the period before temporarily leaving work.

Our aim was to ensure respondents understood the questions as intended and, as a result, answered with reference to their circumstances as of Census Day. Our goal is to ensure that differences in statistics from previous censuses and other surveys represent genuine changes in the lives of people living in England and Wales.

We anticipate that the user needs expressed through previous consultations may have changed. In addition, new user needs will have emerged. For example, this might include the need for information to allow for the development of policies regarding the pandemic.

The remainder of this section outlines the potential changes to the data collected that we've identified. Through this consultation, we aim to understand how data users' needs have changed.

## **Place of residence**

In most cases, the coronavirus pandemic will not have impacted the place of residence. However, there will have been some situations where it has done so.

This may include people with two residences choosing to reside permanently at what had previously been their second residence as they did not need to commute to work. In contrast, there may have been a situation where a person was staying temporarily at their second address on the night of 21 March 2021. In this case, they should have been counted as a visitor at that address and also been included at the address where they are usually resident.

It might also include students who may have been studying from their permanent or family home address instead of a term-time address near their educational establishment.

For Census 2021, in light of the coronavirus pandemic, we updated the collection guidance and statistical methods. We did this to collect the best possible data on students at their term-time address. This will allow users to maximise use of the data on students for long-term planning and service provision. This update meant that we aimed to collect data about students' term-time address even if one or more of the residents was not there on Census Day. We also collected data about their out-of-term-time address. This is in line with the approach taken in 2011.

Further information on how we've adapted our methods to get the best possible estimate of students is available in the report [Census 2021 - How we're ensuring an accurate estimate of students](#). Since that report was published, we've carried out a Student Halls Survey to collect basic demographic information on residents of these establishments. We will use the results of this survey in the census estimation process.

We're also looking at information recorded at students' out-of-term-time addresses, where they recorded a term-time address. We can use this information to quality assure and supplement the term-time address information.

**We'd like to know what data needs you have for understanding how the pandemic impacted usual residents' place of residence.**

## Economic activity and employment

The coronavirus pandemic will have impacted many aspects of employment. For example, this could include an increase in people who were temporarily away from work because they were furloughed or self-isolating. This could also include a likely increase in respondents who were not in paid employment. There will also be a change in characteristics of those looking for work. This includes those made redundant during the pandemic who were looking for work. It also includes those who were self-isolating or extremely clinically vulnerable and not able to start a job outside of the home.

In addition, people might have been working different hours because of the coronavirus pandemic. Alternatively, the nature of their jobs and industries may have changed either temporarily or permanently, such as retail and hospitality switching to home delivery.

**We'd like to know what data needs you have for understanding how the pandemic impacted usual residents' economic activity and employment.**

## Travel to work

Through engagement with users, we decided that capturing travel-to-work patterns, as they were on Census Day, was the best option available. Users did raise concerns about the future relevance of this data. However, users also highlighted concerns regarding the relevance of pre-pandemic travel-to-work patterns and the accuracy of pre-emptively capturing post-pandemic patterns.

We do not know how travel-to-work and workplace-address patterns will evolve following the pandemic. We expect that the transition to greater levels of home working will remain to some extent. However, other aspects such as reduced use of public transport and car sharing may reverse over time.

The changes impact a wide range of products, including statistics on method of travel to work, the workplace and workday population bases, and origin-destination outputs for commuter flows. This impact is both in terms of reflecting societal change and the level of detail that we can provide. For example, because of the reduced number of people commuting, the origin-destination data are likely to be presented in more aggregated categories across all variables that they're produced by.

Users need to understand the commuting patterns and methods, and daytime population, to make decisions or plan services in 2022, 2023 and beyond. It's likely that these data on their own may not fully meet user needs for this understanding.

As a result, we're already undertaking work to identify potential additional data sources that could help provide information to supplement the census data. These could also provide a more current picture of future travel-to-work patterns. This work aims to build on earlier research on [using mobile phone data to estimate commuting flows](#). We carried out this research as part of our work seeking alternative data sources to provide data between censuses and potentially replace the need for future censuses.

**We'd like to know what data needs you have for understanding how the pandemic impacted travel to work. We'd also like to know what data you ideally need for future transport planning and other related decisions.**

## Workplace zones

We're considering the need to produce workplace statistics by workplace zone. Workplace zones are a small-area geography designed to contain a consistent number of workers. This allows us to release workplace statistics at a more granular level. The zones provide much greater detail in areas with high numbers of workers and workplaces. For example, this might include city centres, retail districts and business parks.

We produced the data for this geography in 2011. This information was useful to many groups, including government departments, local authorities, health authorities and academics. Applications included:

- demonstrating broad workplace geographic patterns
- categorising data for further analysis
- identifying similar areas for comparative studies
- providing information for marketing purposes

We also use workplace zone geography in outputs for origin-destination data and small populations.

If we keep the methodology for creating workplace zones unchanged then we will see fewer and larger (in area) workplace zones in major centres of employment, such as city and town centres. We will also see more and smaller workplace zones covering traditional residential areas.

**We'd like to know what data needs you have for statistics based on the workplace zones geography.**

## Impacts on other outputs

Other potential areas of change include migration, health, disability and unpaid care. For example, the uncertainty that the coronavirus pandemic and restrictions on international travel caused might have impacted internal and international migration patterns, including international students.

People's general health could have been affected by having had coronavirus. Their health could also have been affected by the mental and physical health impacts of the virus itself or lockdown restrictions. In addition, long-term health problems and disabilities may now restrict a person's activities even if they might not have done so previously. For example, if they are considered extremely clinically vulnerable and were told to shield.

There's likely to have been a change in the patterns of provision of unpaid care during the coronavirus pandemic. For example, people may have needed care because they were shielding even though they did not need unpaid care previously. Others, who may have provided care before the pandemic, may not have been able to provide care because of coronavirus rules.

This is not a comprehensive list of recent societal changes. We've included it to illustrate the range of ways the coronavirus pandemic has impacted society. Census 2021 data will indicate the scale of these changes.

**We'd like to know where these changes have impacted your ability to use Census 2021 data to inform planning and decision making. We'd also like to know what data you ideally need to make those decisions.**

**We'd also like to know how you will use these data in new ways to understand societal change.**

We will use this feedback to inform a programme of work seeking to address the identified gaps in data need. This will expand the work already started on travel-to-work statistics to include these additional topics.

## Section 7: Paradata

Paradata refer to information about how we collected and processed the data. They're separate from the statistical data that we produce from the census.

In line with 2011 and previous censuses, we will publish a [General Report](#) that will provide an overview of the complete census operation. Ahead of that, we propose to publish some high-level paradata about how people completed their census returns.

We expect to publish information broadly similar to what was included in the 2012 [Providing the online census](#) publication and available on our [How did we do in 2011?](#) page. We're planning to include information on how people completed the census, such as the share of responses online and on paper, within our initial publications. We're also planning to include information about how people used our digital services. These are scheduled for publication in late 2021.

As for the 2011 census, we will release some information, such as household return rates, at local authority level. Return rates are the number of household questionnaires returned as a proportion of the total valid household addresses that were not identified as vacant. We used return rates during the census field operation to target field staff resources to the lowest responding areas.

We're looking at how demographic data from the census can be used alongside this operational paradata. We want to understand if there's a data user need for us to produce more detailed information on:

- household and individual response rates
- mode and language of response

### Household and individual response rate

The household response rate is the total number of households whose details were completed on a returned questionnaire, divided by the estimate of the total number of valid non-vacant households.

The individual response rate is the total number of usual residents whose details were completed on a returned questionnaire, divided by the estimate of the total number of usual residents.

Following the 2011 Census, we published household and individual response rates by variables such as age and sex. We could replicate this for 2021 data and also produce other cross-tabulations.

**We'd like to know what needs you have for paradata on household and individual response rates.**

## Mode and language of response

The mode of response is information on whether the response was submitted online, on paper or through another means. Other means include telephone capture or assisted digital.

Following the 2011 Census, we published data on the proportion of returns made by mode: online or paper. For 2021, we could add a third category, which includes all other modes of response. We're also looking into how we can derive a new digital propensity index based on how likely households were to respond to Census 2021 online.

For Wales, as well as mode of response, we can provide information on language of response. This information will depend on whether it was a paper or online response.

For paper responses, we can produce information on whether respondents used the Welsh language form or the English language form.

For online responses, we can produce information on which language the form was set to at the point of entry and the point of submission. These include:

- started form in Welsh and submitted in Welsh
- started form in Welsh and submitted in English
- started form in English and submitted in Welsh
- started form in English and submitted in English

If a user need is evidenced, and these are produced, we would publish them after the main census release of data for the same characteristics and geographies.

**We'd like to know what needs you have for paradata on mode and language of response.**

## Section 8: Next steps

You can see Census 2021 milestone dates related to this consultation, and progress against them, on the [ONS website](#). This consultation contains the draft release schedule, and it will remain open for 12 weeks, from 13 July 2021 to 5 October 2021.

We will evaluate the information you provide in response to this consultation, alongside information collected during other stakeholder engagement activities, to inform decisions on the final design of Census 2021 outputs for England and Wales. When making these decisions, we will consider how the publication of data will serve the public good. We will also consider operational factors such as data quality and disclosure of personal information risk.

We will publish a response to the consultation within 12 weeks of the consultation closing. We aim to publish the final design of Census 2021 outputs, in the form of a Census 2021 Outputs Prospectus, by February 2022. This will include more detail on the planned release schedule.

We're also conducting a series of user-research projects to improve how people will be able to find, access and use Census 2021 data products. The purpose of this research is to identify ways we can increase the accessibility and utility of census data. Our aim is for Census 2021 data to be used by a wider range of users and for a wider range of purposes in comparison to the 2011 Census data. If you'd like to participate in this user research, please email us at [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk) for more information.

Later this year, we will launch a further major phase of user engagement. During this phase, we will share with stakeholders the nature of the transformation we're proposing. It will also provide an opportunity for us to share the framework of evidence that will inform the National Statistician's recommendation on the future of the census and population and social statistics. This recommendation will be made by the end of 2023.

We know that this will be of interest to many of the same users and stakeholders. This engagement will continue during the next year to ensure that the recommendation on the future of the census is fully evidenced by stakeholder views.

We will publish the initial findings from Census 2021 in Spring 2022. We aim to have published all main census data outputs for England and Wales by March 2023.

Following this, we will continue to develop our data analysis products to further increase the utility of the Census 2021 data. We will also start to release UK data outputs, once data from Scotland's census in 2022 are available.

At the end of 2023, the National Statistician will make a recommendation to government. This will outline what further change is needed to deliver a fully transformed population and social statistics system. This will include the role of any future census and, for

example, necessary improvements in data collection or acquisition, methods and infrastructure.