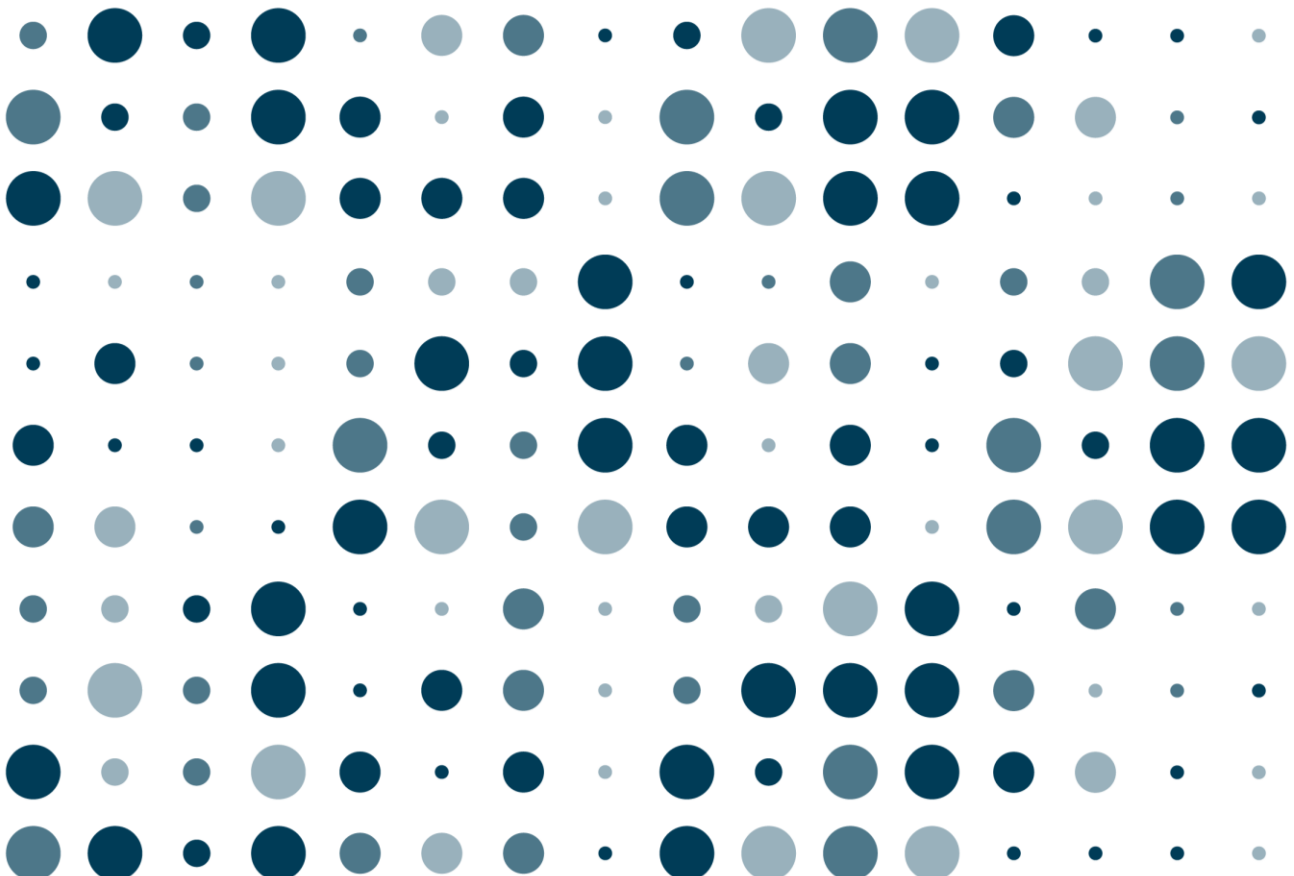


## Response to the ‘Census 2021 outputs: content design and release phase proposals’ consultation – part two

Recommendations about questions focused on phase two and three of the main release schedule, and later outputs.

31 March 2022



# Contents

<b>Contact information</b>	<b>2</b>
Enquiries	2
Updates	2
Quality assurance	2
Publication of consultation responses	2
<b>Introduction</b>	<b>3</b>
Report context	3
Contents of this report	3
<b>Summary of changes to proposals</b>	<b>4</b>
Statistical products	4
Census during a period of change	9
Paradata	10
Next steps	11
<b>Statistical products: conclusions</b>	<b>12</b>
Multivariate data	12
Proposed new derived variables	17
Small populations	39
Alternative population bases	42
Origin-destination data	48
Microdata	49
Analysis	49
<b>Census during a period of change: conclusions</b>	<b>58</b>
Place of residence	58
Labour market	60
Impacts on other outputs	65
Emerging data needs	66
<b>Paradata: conclusions</b>	<b>69</b>
Understanding data quality	69
Response rates	70
Completion method	73
<b>Next steps</b>	<b>79</b>
<b>Annexes</b>	<b>80</b>
Annex A: evaluating users' requirements	80
Annex B: quantitative evidence of need for proposed new derived variables	81
Annex C: initial proposals for small population datasets	82

## Contact information

### Enquiries

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

**Phone:** +44 1329 444972

**Fax:** +44 1633 652981

**Post:** Census 2021 Outputs Consultation, Office for National Statistics, Segensworth Road, Fareham, PO15 5RR

### Updates

At the Office for National Statistics (ONS), we will continue to promote new information about Census 2021 outputs through our email bulletins, which are hosted by GovDelivery.

To receive the latest updates on Census 2021 outputs, products, release plans and events, please register to receive the free GovDelivery bulletins. You can do this by:

- visiting the [GovDelivery](#) website
- entering your email address in the field provided
- selecting the option to subscribe to '2021 Census outputs'
- selecting submit at the bottom of the page

### Quality assurance

The ONS carried out this consultation in accordance with the [government's consultation principles](#).

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

### Publication of consultation responses

For transparency, we have published the moderated responses to this consultation. These are available on the [consultation webpage](#).

Our privacy policy is also available on the [Privacy](#) webpage.

If you'd like to give us more feedback, please contact the Census 2021 outputs and dissemination team at [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk).

# Introduction

## Report context

We'd like to thank everybody who gave us their views by responding to the consultation [Census 2021 outputs: content design and release phase proposals \(pdf, 929KB\)](#), for England and Wales.

We based our proposals for the content design and release phase proposals for Census 2021 outputs on the user feedback from previous consultations and ongoing discussions with stakeholders. We're confident that the changes we're making to our content design proposals, that we cover in this report, will help us to meet a larger range of user needs.

In December 2021, we published [part one of our response to the consultation \(pdf, 625KB\)](#). In it, we outlined planned changes to phase one proposals of the [Census 2021 outputs release schedule](#). These proposals included:

- information on census geographies
- specifications for the primarily univariate topic summary and area profile products
- univariate data we plan to produce from new Census 2021 questions
- new or amended classifications for Census 2021 outputs

We also provided information about how we've evaluated the respondents' feedback to identify new data needs. In [Annex A: evaluating users' requirements](#) of this report, we discuss how we're evaluating feedback that relates to different or interdependent products across different sections of the consultation.

## Contents of this report

In this report, we focus on user feedback about the design of Census 2021 products in phases two, three and beyond 2023 of the [Census 2021 outputs release schedule](#). We provide information on the changes to our product proposals that we're considering or implementing. We also provide descriptions of respondents' requirements.

Many of the products that we discuss in this report cut across different phases of the release schedule. So, we have structured this document by product type.

First, we discuss statistical products that provide information on the characteristics of the population and housing of England and Wales.

Then, we discuss the extent to which data needs have changed as a result of the current period of significant economic and societal change.

Finally, we end by discussing respondents' needs for paradata. Paradata refer to information about how we collected and processed the Census 2021 data.

## Summary of changes to proposals

In this section, we summarise the consultation proposals that we're changing or considering changing. These changes are based on conclusions we've drawn from the respondent feedback provided in the consultation. They focus on phases two onwards of the Census 2021 outputs release schedule.

We've discussed the evidence behind each proposal for change within the sections of this document on:

- [Statistical products: conclusions](#)
- [Census during a period of change: conclusions](#)
- [Paradata: conclusions](#)

We are not basing our final decisions on the design of Census 2021 outputs on this consultation alone. Where respondents have provided a justified need for additional information, we're assessing this need. In our investigation, we will determine if it's appropriate or feasible to publish this information as part of the Census 2021 outputs. We're considering factors such as data quality and statistical disclosure control. If our feasibility assessments show that the amended data product will meet our data quality and confidentiality requirements, we will aim to make the changes. Conversely, there are many user needs that we knew about before the consultation and so do not need further evidence to include them in our plans.

As we finalise our plans throughout 2022, we will publish further detail on the [release plans](#) webpages. These will confirm which of the proposals discussed are being taken forward.

After phase three of our release schedule, we will carry out a user survey. This will allow us to assess if there are any new or existing data needs that we have not yet adequately met. In this survey, we will ask users if the census products have met their needs and, more importantly, learn where they might not have. The results of this survey will help us to determine if we need to produce additional outputs, for example, on small populations, or additional supporting information.

## Statistical products

### Multivariate data

#### Ready-made datasets

For sexual orientation and gender identity, we will prioritise datasets including variables that allow investigation of equality in life outcomes. We aim to publish these datasets for the smallest geographies possible whilst minimising the risks of statistical disclosure.

We will consider the feasibility of producing a wider range of ready-made datasets on:

- year of arrival in the UK and passports held
- Welsh language skills
- sexual orientation and gender identity

### **Build-your-own datasets**

We have also developed functionality for datasets that users can make themselves, by adding or changing variables, before viewing or downloading the dataset. Within this functionality, we will aim to include the most detailed breakdowns of all variables. However, users may only be able to receive information from some of the most detailed variables at higher geographies.

For build-your-own datasets, we will:

- add additional variables, including household size and tenure, to the 'usual residents in the households' population base
- include variables for the UK armed forces veteran status and the veteran status of the household reference person
- investigate how feasible it is to provide an additional classification of distance travelled to work, focussing on active travel distances

### **Proposed new derived variables**

We aim to release data on the proposed derived variables within phase three of the release schedule where:

- there is sufficient user need
- our feasibility work indicates that the data will be of sufficient quality
- the variable only requires information that we collected for Census 2021

We aim to release more complex variables later in the release schedule.

### **Education and employment**

We will produce variables identifying the full-time student population. We will ensure that users can categorise full-time students by broad age group and accommodation type.

We will continue to assess how feasible it is to:

- develop a variable on young people not in education, employment or training (NEET) for the usual resident population base
- develop a classification of the occupation variable for key or critical worker, that aligns with the UK government's definition and maximises the classification's usefulness
- produce a skills mismatch variable and analysis products on this topic
- produce a classification for the industry variable on economic risk created by the coronavirus (COVID-19) pandemic

- producing industry classifications that would allow easy identification of science, technology, engineering and mathematics (STEM) and green industries

We're unlikely to produce a variable on the population temporarily away from work using responses to the Census 2021 form.

We will not:

- produce any variables for route to highest level of education
- produce a variable on the carbon cost of travel to work

### **Health and living arrangements**

We will assess how feasible it is to:

- develop a variable on houses in multiple occupation (HMO) in line with the government definition, to separately identify small and large HMOs
- develop a variable on multigenerational households, for the usual resident and household population bases
- combine Census 2021 data with health data sources, to provide more insight into the impact of the coronavirus pandemic on different populations

We will not produce variables derived from data that we collected in Census 2021, for:

- COVID-19 health risk
- those living apart together

### **Accommodation type and vacant addresses**

We will produce an additional classification for type of accommodation with the categories:

- whole house or bungalow
- flat, maisonette or apartment
- caravan or other mobile or temporary structure

We will produce additional reduced classifications for type of communal establishment that provides detail on the type of:

- travel or temporary accommodation
- medical or care establishment

We will continue to assess how feasible it is to produce data on the possible reason for a property being vacant. If we produce this information, it would be in the form of a methodological paper.

We're unlikely to produce a variable including information from the response option "People staying temporarily who usually live in the UK but do not have another UK address, for example, relatives, friends". Because of the strong need for information on

this population, we're continuing to assess if there are any alternative options to provide information on this population.

We will not produce a variable on type of temporary or mobile accommodation derived from operational data we collected in Census 2021.

## **Small populations**

We will continue with our proposal to produce a suite of small population datasets for Census 2021 similar to those we produced for the 2011 Census.

We will investigate how feasible it is to add a dataset on main language for each Middle Layer Super Output Area (MSOA) level small population. We will also investigate if we can produce any of the datasets for either wards or Lower Layer Super Output Areas (LSOAs).

We will consider the need to produce a set of MSOA-level small population datasets, in addition to those proposed in [Annex C: initial proposals for small population datasets](#). These would be for:

- British Sign Language (BSL) main language
- Romania country of birth
- Somali ethnic group
- Turkish Cypriot ethnic group

We will investigate how feasible it is to produce additional datasets at local authority-level showing five-year age bands by sex on:

- Caribbean ethnic group
- China as country of birth
- Jewish identity, with the defining variables to be determined
- Pashtun ethnic group
- Showman ethnic group

We will work with data users to understand how they need us to define these groups.

## **Alternative population bases**

We will produce outputs on the five alternative population bases that we outlined in our proposals. We do not plan to produce any other alternative population bases.

## **Workplace and workday**

When finalising our plans for workplace and workday population datasets, we will prioritise investigating how feasible it is to produce univariate datasets at MSOA level on:

- age



- highest level of education
- occupation
- industry
- method of travel to work
- distance travelled to work

We will assess how feasible it is to produce any of these higher-priority datasets for lower levels of geography such as LSOAs and workplace zones.

We will investigate how feasible it is to:

- expand the workplace definition to include those aged 75 years and over
- produce workday datasets similar to 2011 Census datasets

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

When finalising our plans for out-of-term population datasets, we will prioritise univariate datasets on:

- age
- sex
- ethnic group

### **Second address**

The second address population looks at the characteristics of the population with a second address, including the characteristics of the second address itself. When producing datasets on the characteristics of the population with a second address, we count them at the location of their second address rather than at the address where they are usually resident. When finalising our plans for datasets on the second address population base, we will prioritise univariate datasets on:

- age
- employment status
- place of usual residence

We will also investigate how feasible it is to produce datasets that would allow more detailed analysis of the households identified as a second address for one or more dependent children.

We will assess if we can produce any of these datasets to a lower level of geography such as LSOA.

### **Non-UK born short-term resident population base**

We aim to bring the release of regional level datasets forward from phase three of the release schedule into phase two.

## Analysis

We will:

- explore the potential to combine Census 2021 data with additional administrative datasets, for example ones about the population of Wales
- ensure that our plans include analysis looking across topics at the impact of large changes such as Brexit and the coronavirus pandemic
- consider which topics would benefit from analysis of the difference between rural and urban areas
- consider embedding analysis by ethnic group, religious group, and by health and disability, into a wider range of analysis products

We will bring forward analysis of the UK armed forces veteran population by health and disability from phase three of our release schedule into phase two.

We will consider how to bring into the analysis programme, discussions about:

- the distribution of holiday homes, including those within Welsh-speaking communities
- the number of children living with same-sex couples
- households identified as a second address for one or more dependent children
- migration from countries in conflict, subject to statistical disclosure control
- English language proficiency and main language of the care home population

## Census during a period of change

We're currently in a period of significant economic and societal change. This change is not only a result of the coronavirus pandemic, but other factors such as Brexit.

We will:

- provide more frequent and detailed statistics using administrative data sources to give an ongoing accurate picture of the population in the years to come
- monitor ongoing change since Census Day using new sources of data, including aggregate mobile phone data, to understand changes in patterns of travel to work

We will continue work to understand whether or not we can use administrative-based data sources to provide better information on the population temporarily away from work. This is a longer-term project, and it would conclude after the end of the main release schedule.

We will expand our work with the Valuation Office Agency (VOA) to further understand the use of the broad rental market area (BMRA) geography. If needed, we will assess how feasible it is to produce an Output Area (OA) to broad BMRA look-up.

## Paradata

Paradata refer to information about how we collected and processed Census 2021 data.

### Response rates

We will investigate how feasible it is to produce household and individual usual resident response rates at LSOA level.

For households, we will investigate producing MSOA-level response rates by:

- accommodation type
- household size
- household tenure
- mode of response

For individual usual residents, we will investigate producing MSOA-level response rates on protected characteristics, such as:

- age
- ethnic group
- religion
- sex

### Completion method

We will investigate how feasible it is to produce information on mode of completion by the household reference person, at local-authority level.

For mode of completion, we will prioritise information on:

- age
- disability
- ethnic group
- highest level of education
- sex

We will include the digital hard-to-count index score and primary mode of completion assigned to the address in these datasets.

In England, respondents could complete Census 2021 in English. In Wales, respondents could complete Census 2021 in English or in Welsh. We did not produce forms in any other language. We will investigate how feasible it is to produce information on language of completion in Wales by the household reference person, at local-authority level.

For language of completion in Wales, we will prioritise information on:

- age

- sex
- highest level of education
- ability to speak Welsh

We plan to produce further information about the provision of Welsh services in an evaluation article about delivering Census 2021 in Wales.

We will also publish methodological reports that will explain to users how they can use the information.

## **Next steps**

We will start publishing our updated detailed proposals by early summer 2022. We've provided more detail about our plans for the future in the section [Next steps](#).

## Statistical products: conclusions

### Multivariate data

We provided detailed proposals in the consultation for how we planned to publish multivariate data on the usual resident population of England and Wales. This will take place in phase two of the release schedule. These proposals are in the 'Draft proposals for outputs data content' spreadsheet located within the "Related" section at the bottom of the [Census 2021 outputs consultation webpage](#). We proposed to produce these multivariate datasets down to Output Area (OA) level using the OA best-fit approach. These datasets will not include totals or subtotals.

Users will be able to access two types of multivariate datasets. The first are ready-made datasets that users can view and download. The second are build-your-own datasets that users can make themselves, by adding or changing variables, or by building from scratch, before viewing or downloading the dataset. This new functionality will provide users with flexibility in two main ways. Users will be able to access:

- data that more directly meets their needs
- more detail than was possible in 2011

Of the 229 respondents who stated whether they would use the functionality that allows users to build their own datasets, 93% stated that they would. The respondents who stated that they would not use this functionality provided two reasons.

Some felt the existing datasets would meet their needs. Others expressed concern that they would lack the skill needed to use the functionality correctly. We will ensure that the functionality that allows users to build their own dataset is publicly available and designed to meet government digital accessibility standards. We will provide support for users to ensure they are able to easily access Census 2021 data.

For example, we will provide data in appropriate formats with accompanying metadata. As part of the supporting information, we will provide some information about how we derived the variables, when we release data from them.

### Dataset specifications

We asked users if our proposals for ready-made datasets and build-your own datasets would meet their needs. Of the 229 respondents who stated whether the proposals for ready-made datasets met their needs:

- 65% stated that they fully met their needs
- 30% said that they partially met their needs
- 5% stated that they did not meet their needs

Of the 224 respondents who stated whether the proposals for build-your-own datasets met their needs:

- 66% stated that they fully met their needs
- 29% stated that they partially met their needs
- 5% stated that they did not meet their needs

Across the two questions, 74% of respondents stated that at least one type of dataset fully met their needs.

Respondents discussed why the datasets only partially met their needs or did not meet them. Some were addressed in the [consultation response: part one \(pdf, 625 KB\)](#) document. For example:

- the reduced age classifications
- separate identification of EU and non-EU member countries in the classifications for country of birth and passports held
- provision of estimates of the population who can speak Welsh

## **Ready-made datasets**

For the usual resident population base, we generally plan to provide the ready-made datasets for all output areas. For the usual resident in households and usual resident in communal establishment population bases, we generally plan to provide the ready-made datasets for all Middle Layer Super Output Areas (MSOAs).

Respondents requested several additional datasets. If a request was supported by a description of why it's needed, we've discussed it in the following sections. Additionally, users will be able to make the majority of the additional ready-made datasets that were requested from the build-your-own dataset functionality.

## **Migration**

Our proposal for migration datasets only included one dataset on year of arrival, and a small number of datasets on passports held. In their feedback, respondents requested additional datasets about migration to aid understanding of community cohesion and service provision needs. For example, an individual responding in a professional capacity stated:

“The type of Census characteristics that a migrant has will depend on when they arrived in the country. In addition, the majority of migrants who remain in the UK will eventually have similar characteristics to non-migrants - but identifying those who don't is an important aspect of discrimination. This is particularly important in relation to Census questions on ethnicity, religion, national identity, country of birth, language, etc.”

In response, we will consider the feasibility of producing a wider range of datasets for year of arrival and passports held.

### **Welsh language skills**

Respondents representing Welsh communities requested additional datasets about Welsh language skills. They would use this when evaluating the effect of policies promoting use of the Welsh language. For example, Flintshire County Council stated:

“This data will help inform Welsh language promotion strategies and where to target interventions to increase the use of Welsh. There has also been some national discussion that the requirement to speak Welsh in certain organisations may be indirectly discriminatory if some groups are less likely able to comply with the requirements. Census data will contribute to investigating this.”

In response, we will consider the feasibility of producing a wider range of datasets on Welsh language skills. We will consider if we can provide these variables within the functionality for users to build their own datasets.

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

Respondents requested additional datasets including sexual orientation and gender identity variables. They would use this to build an understanding of these populations for policy development and diversity monitoring. For example, Salford City Council stated:

“There is considerable demand across local government services for understanding differences for various variables based on protected characteristics such as age, gender, ethnicity, religion, sexual orientation and gender identity. This should be a key focus for the multivariate data tables.”

We currently consider it unlikely that we will include these variables in the functionality for users to build their own datasets. This assumption increases the need for these datasets.

In response, we will consider the feasibility of producing a wider range of ready-made datasets including the sexual orientation and gender identity variables. We will prioritise datasets including variables that we can use to measure equality in life outcomes, for example, on health, housing, education, and employment.

We aim to publish these for the smallest geographies possible whilst maintaining a low risk of statistical disclosure. We anticipate that this would generally be MSOA level. However, we may use either larger or smaller geographies depending on the dataset composition.

## Build-your-own datasets

Respondents requested that we add several additional variables or variable classifications to the proposals for datasets that users can build themselves. In the following sections, we've discussed the requests that respondents supported with descriptions about why the additional information is needed.

### More detailed classifications

Respondents requested more detailed classifications for most variables, where we had not included the most detailed classification in the proposal. These were largely the socio-cultural classifications, for example, ethnic group and religion, but also included other classifications, such as the provision of unpaid care.

Respondents indicated that this would generally allow a more detailed understanding of specific subgroups within the population. They could also use it to target policies and service provision. For example, Bristol City Council explained how they have used the detailed 2011 statistics, and how they could use this again:

“Detailed Country of Birth statistics for Bristol LA enable our Diaspora group research to help greater understanding of who our Diaspora are, to promote greater understanding, respect and value of people’s heritage and to help support inclusive growth (trade and investment), educational and cultural links through our international links, particularly where we have larger/increasing communities in the city such as from Somalia and Nigeria.

Detailed Country of Birth with different EU Citizenship for Bristol LA (eg Somali with Dutch citizenship). To help identify resources needed to help ensure these citizens don’t get into issues under the EU Settlement scheme status [...]. EUSS has now officially closed but we remain concerned about those that didn’t apply and for those that also need to follow up on pre-settled status.”

In relation to unpaid care, an individual responding in a professional capacity stated:

“It would be useful to policy making to be able to produce datasets looking at unpaid carers intensity (hours of care provided per week) cross-tabulated with other variables. Intensity is a key determinant of the common negative effects of providing unpaid care e.g. leaving employment or negative health. Observing the difference between those caring at different levels of intensity, e.g. above or below 10 hours per week, would therefore make a valuable contribution to the evidence base and help inform the development of future policy.”

In response, we will aim to include the most detailed version of all variables within the build-your-own datasets. However, to maintain a low risk of statistical disclosure, users



may only be able to receive information from some of the most detailed versions of the variables at higher geographies.

### **Household variables for the usual residents in households population base**

Respondents requested the inclusion of more of the household variables, such as household size and tenure, for datasets about usual residents in households. This would allow them to better understand the housing conditions of all usual residents.

For example, Bristol City Council stated:

“We need tenure statistics based on the number of people as well as the number of households. [...] It is important that we are able to cross-tabulate equalities statistics and other population characteristics by housing variables, not just based on the HRP.”

In response, we will add more household variables for the usual residents in households population base. We will prioritise inclusion of household size and tenure.

### **UK armed forces veterans**

Respondents requested that we include all the proposed variables on UK armed forces veterans within the functionality for users to build their own datasets. We’ve discussed respondents’ detailed needs for data in [part one of our response to the Census 2021 outputs consultation \(pdf, 625 KB\)](#).

In response, we will include the variables for UK armed forces veteran status and veteran status of the household reference person, within the functionality for users to build their own datasets.

### **Distance travelled to work**

Respondents requested a classification of workplace travel that focuses on “active travel” distances, such as:

- less than 2 km
- 2 km to less than 5 km
- 5 km to less than 10 km
- 10 km or more

This would provide a more detailed understanding of where targeted interventions could be put in place to encourage walking or cycling to work. For example, Hampshire County Council stated:

“Tables on distances to work need to identify reasonable walk distance (<2km) and reasonable cycle distance (<5km). The need to identify potential distances

for Active Travel mode is required for the development of transport policies in line with central government guidance.”

In response, we will investigate how feasible it is to provide an additional classification of distance travelled to work, focussing on active travel distances.

### School year

Respondents requested that we provide a variable on age as of 31 August 2021, as well as a variable on age as of Census Day, 21 March 2021. This would provide a proxy for information on school year for analysis related to educational provision. For example, Cambridge City Council stated:

“Age data for children and young adults with age as at 31st August would be extremely useful, to align with education policy, budget and planning timeframes.”

However, our feasibility work has shown that data on school year would not be of sufficient quality to publish without significant further processing. When date of birth information is missing, our edit and imputation process derives age as of Census Day. To produce this variable would also introduce additional risk of statistical disclosure.

In response, we do not propose to create a variable for school year, based on date of birth.

### Proposed new derived variables

In this section, we focus on proposals to combine existing data, often from multiple questions, in new ways to meet a wider range of current user needs. We’ve grouped our proposals into three areas. These are:

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

In all cases, we asked respondents if they would use the proposed new derived variable. We have summarised the quantitative data from these questions in [Annex B: quantitative evidence of need for proposed new derived variables](#).

We also asked respondents to prioritise a range of factors that relate to their data needs for the variable. Generally, respondents prioritised the factors in the following order:

1. data cross-tabulated with other variables
2. data for smaller geographical areas
3. data that provide a more detailed breakdown of the proposed new variable
4. data comparable with previous England and Wales censuses
5. data comparable with the Scotland and Northern Ireland censuses

We were still in the early stages of developing our derived variable proposals when we published the [consultation document \(PDF, 929KB\)](#) in July 2021. Since then, we’ve

continued to assess the feasibility of creating each. We discuss the feasibility assessment alongside the relevant proposals.

Respondents also told us that they would require information on our final methods for deriving any of the new variables we decide to produce. We will provide this information for all variables used in Census 2021 outputs.

We aim to release data on the proposed derived variables within phase three of the release schedule, where:

- there is sufficient user need
- our feasibility work indicates the data will be of sufficient quality
- the variable only requires information that we collected for Census 2021

We aim to release more complex variables later in the release schedule.

## **Education and employment**

We proposed seven new variables or classifications about education and employment. These were:

- route to highest level of education
- adult full-time students
- not in employment, education or training (NEET)
- temporarily away from work
- key or critical worker
- skills mismatch
- economic risk created by the coronavirus (COVID-19) pandemic

In addition, consultation respondents proposed:

- classifications for green and STEM industry
- a derived variable on carbon cost of travel to work

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

We asked users whether they would use a proposed new variable on the route to highest level of qualification. Out of the 190 respondents that answered this question, 62% said they would use it.

We also asked respondents what information they needed about the route to highest level of education. Respondents raised two clear needs that we could meet by deriving most of the required information from Census 2021 qualifications questions. These needs are:

- whether the highest level of qualification was academic or vocational
- whether the person had all academic, all vocational, a mixture of both, or unknown type of qualification

Respondents also made other requests, but this was for information that we did not collect in Census 2021, such as:

- type of educational establishment attended
- if in part-time education
- how long ago qualifications were achieved

Respondents discussed needing data to inform local area skills strategies and skills improvement plans. For example, data allowing investigation of the impact of route to highest level of education on employment outcomes, in line with the government's levelling up strategy. Cambridgeshire County Council stated:

“It would be used alongside other variables to measure upcoming demand for adult education, analysis which supports the setting of skills policies, adult education budgets and the negotiation of education provision with providers [...] and our work on forecasting school places, undertaken to meet our statutory duty to provide sufficient school places.”

The level of detail collected in Census 2021 would not meet these needs. Census 2021 collected information on broad types of qualification. The proposed analyses would require more granular information on the exact qualifications that respondents obtained and when these were completed. This would, for example, allow identification of those who have returned to learning after a period of other activity to train at the same or lower level. This could indicate that the learner may be retraining for a new career.

Respondents requested detailed information about the routes through education. This would help to indicate the choices learners make at different stages of education. Administrative data from each sector better meets these needs for detailed information. For example:

- school data for England, which is held in the National Pupil Database
- school data for Wales, which is held in the National Pupil Database for Wales
- Further Education data for England, which is held in the Individualised Learner Record
- Further Education data for Wales, which is held in the Lifelong Learning Wales Record
- Higher Education data for England and Wales, which is held in the Student Record

Alternatively, the Longitudinal Education Outcomes dataset combines all three sectors of data and allows users to look at labour market and self-employment outcomes for England.

Analyses using these and other rich sources of education data include:

- [Alternative routes into and pathways through higher education](#), the Department for Business, Innovation and Skills (BEIS), 2009
- [Progression pathways 2017: Pathways through higher education](#), the University and Colleges Admissions Service, 2017
- [Post-16 education and labour market activities, pathways and outcomes](#), the Department for Education (DfE), 2021
- [National pupil projections: July 2021](#), DfE, 2021 – the DfE also circulates local authority-level projections directly to councils for planning purposes

In response, we will not produce any route to highest level of education variables.

### Adult full-time students

We asked users whether they would use a proposed new variable on adult full-time students. Out of the 189 respondents who answered this question, 66% said they would use it.

Some respondents requested that we split this group into three age classifications aligned with the end of compulsory education and expected ages of students. These are:

- ages 0 to 17 (schoolchild)
- ages 18 to 24 (student)
- 25 years and above (mature student)

We're considering this feedback while we finalise the age classifications.

Respondents also requested information on if this populations term-time address was a:

- communal establishment
- household with their parents
- household independent from their parents

Respondents discussed needing this data in order to:

- inform local area skills strategies and skills improvement plans
- inform local area housing strategies, particularly in relation to university students
- understand the impact of the coronavirus (COVID-19) pandemic

In relation to skills strategies, Cambridge City Council stated:

“[T]o measure upcoming demand for adult education, analysis which supports the setting of skills policies, adult education budgets and the negotiation of education provision with providers.”

In relation to the coronavirus pandemic, London Borough of Havering stated:

“We are interested in how the Covid-19 pandemic has impacted on adult education, people returning to study to support career changes etc. and how this differs across occupations and industries.

Respondents also made requests for information that we did not collect in Census 2021. These included information on type of educational establishment attended or if a person was in part-time education. Information on these topics is available on the Higher Education Statistics Agency’s [Higher Education Student Data webpages](#).

In response, we will produce variables identifying the full-time student population. We will ensure that users are able to categorise the full-time students by broad age group, and by student accommodation type.

### **Not in employment, education or training (NEET)**

We asked users whether they would use a proposed new variable on usual residents aged 16 to 24 years that are not in education, employment or training (NEET). We would define “not in education” as not in full-time education.

Out of 198 respondents who stated if they would use a NEET variable, 78% stated they would. Out of all of the proposed new derived variables, this one received the highest number respondents stating they would use it.

Respondents discussed needing data for equality assessments. Respondents from local authorities also discussed using NEET data in a wide range of local area strategies. These strategies focus on topics such as:

- skills
- housing
- social inclusion
- health
- youth services
- apprenticeships

For example, Knowsley Metropolitan Borough Council stated:

“The Children and Families Plan 2021 – 2024 outlines Knowsley ambition to use early intervention and preventative work to reduce the number of young people becoming NEET. Using Census data to further understand the characteristics of young people who end up as NEETs would provide further insight that would feed into strategies and preventative measures. Outputs at small geographies would assist with more targeted interventions and where best to commit resources.”

Respondents from charity and voluntary sectors said that NEET data would help them to target their services. For example, Nettlestone and Seaview (men’s) Shed stated:

“It would be useful to know how many might benefit from a Men in Sheds skills improvement locally.”

In addition, Crisis stated:

“Data about those who are both NEET and experiencing homelessness/housing precarity across different local authority areas would help inform our policy and service delivery for these groups.”

Due to changes in legislation around compulsory education, respondents expressed that it was important to be able to separate those aged 16 to 17, from those aged 18 to 24.

In response, we will continue to assess how feasible it is to develop a NEET variable for the usual resident population base.

### **Temporarily away from work**

We asked users whether they would use a proposed new variable on the population temporarily away from work. This could separately identify those who were either “Temporarily away from work ill, on holiday, or temporarily laid off” or “On maternity or paternity leave”.

Out of 184 respondents who answered this question, 61% stated they would use data on the population temporarily away from work. This is a low percentage compared to other proposed new variables.

Respondents discussed using data on the population temporarily away from work as an important baseline for coronavirus (COVID-19) pandemic recovery work. Within this group, they were specifically interested in the population on furlough. They would use data on the population temporarily away from work to look at issues such as:

- debt
- vulnerability to the coronavirus pandemic
- increasing inequalities
- child poverty

For example, Huntingdonshire District Council stated:

“The impact of the pandemic/lockdown restrictions may skew the numbers in work considerably so knowing how many were affected temporarily due to reasons such as furlough or self-isolation is very important.”

In another example, West Sussex County Council stated:

“It would be useful to know whether those people who have been/are furloughed are returning to their jobs before COVID – or are they going to different jobs/sectors. The labour market is so complex – so why is there a shortage for e.g. chefs and catering staff when so many have been furloughed – is it because

they are going into different sorts of jobs – e.g. better paid or no shift work. There is no redundancy data at local level – so something capturing the local picture would also be valuable.”

However, we cannot identify the population on furlough or the population in self-isolation separately from those away from work ill, on holiday or temporarily laid off for other reasons. When the pandemic began, we had already finalised the question and response option wording in the Census 2021 [legislation](#). In the online form, we added the following instruction directing people completing the form in how to answer:

"If you have a job but have been off work on furlough, in quarantine, or self-isolating, select 'Temporarily away from work ill, on holiday or temporarily laid off'".

This was part of a suite of guidance changes designed to aid respondents in answering the questions discussed in the report [Updates to Census 2021 online questionnaire guidance – coronavirus \(COVID-19\) pandemic impacts](#).

Respondents also discussed using data on those who are away from work on maternity or paternity leave, in relation to diversity monitoring. This is because “pregnancy and maternity” is a protected characteristic under the [Equality Act \(2010\)](#).

However, our feasibility work has shown that the data on the population temporarily away from work would not be of sufficient quality to publish without significant further processing. This is because we do not process the variables “Temporarily away from work ill, on holiday, or temporarily laid off” or “On maternity or paternity leave” beyond early stages of processing. We combine this information with other variables to create an economic activity variable.

In response, we are unlikely to produce a variable on the population temporarily away from work using responses to the Census 2021 form. If we do, we will publish any outputs on this topic using Census 2021 data after March 2023. However, we acknowledge the strong user need for information on these populations. We discuss our future plans in the [Labour market](#) section.

### **Key or critical worker**

We asked users whether they would use a proposed new variable on key or critical workers. This population is also referred to as “essential workers”. This proposed indicator would identify people whose occupation was considered critical to the response to the coronavirus pandemic. We’d define this in line with the UK government definition, as used in our other [publications on key workers](#).

Out of 188 respondents who answered this question, 68% said they would use a variable on key or critical workers.

Respondents discussed using this data for a wide range of purposes. These included:



- providing services for key workers, for example housing and childcare
- forecasting the future labour market, to anticipate shortages in areas where the key worker population are nearer retirement age
- understanding the impact of being a key or critical worker on health outcomes, particularly in the light of the coronavirus pandemic
- planning for future emergencies, for example, pandemics or fuel crises
- demonstrating the importance of subgroups to the community to enhance community cohesion

For example, Yate Town Council stated:

“[To] identify where there is pressure for key worker facilities, e.g. improved child care, and crucially to understand the way covid/ lockdown is affecting different sub communities”

In another example, Migration Yorkshire stated:

“It would help us to identify the contribution of migrants to the local community and to the economy, and thus support community cohesion work with the wider community. It would also help us to target our engagement with employers in those key worker industries, to advise on how to best support their employees from a migrant background.”

Adding to this example, Muslim Census stated:

“From previous research work, we have seen there to be a higher proportion of Muslims in frontline/key worker roles. We would love to see a larger data set showcasing this information.”

Respondents were broadly supportive of the use the UK government’s definition for critical or key workers. However, the Welsh Government noted that their definition was slightly different, and that the definition we use should take this into consideration. Other respondents requested that the classification include subgroups of critical and key workers to increase the utility of the data for the range of purposes discussed.

In response, we will continue to assess how feasible it is to produce a classification of the occupation variable for key or critical worker. We will aim to align the classification with the UK government’s definition of key or critical workers, and in a way that maximises the classification’s usefulness.

### **Skills mismatch**

Skills mismatch is where a person’s educational level is significantly higher or lower than the average [level of qualification](#) within their occupation.

We asked users whether they would use a proposed new variable on skills mismatch. Out of 183 respondents who answered this question, 67% said they would use it.

We also asked users if they would like us to define the variable on skills mismatch or the classifications for the variable in a specific way. In relation to a definition for a variable on skills mismatch, 92% of 152 respondents stated they did not need a specific definition. In relation to classification needs for a variable on skills mismatch, 90% of 162 respondents stated they did not need a specific classification.

Where respondents did request a specific definition or classification, these mainly aligned with our proposals. For example, the Cabinet Office Equality Hub stated:

“We propose that classifications capture the magnitude of the mismatch in education levels, e.g. no mismatch, +1 level, +2 levels, +3 or more.”

Respondents discussed using the skills mismatch data in the development of local area strategies. These related to areas such as:

- economic growth
- adult education
- labour market participation planning

For example, Birmingham City Council stated:

“This [...] would help us better understand the mismatch between skills required by employers and those held by residents and therefore would help inform policy and interventions to improve labour market participation rates locally. This would also be crucial in determining where we should be directing particular types of industry or educational establishments in the future.”

Respondents said they would use the information to advise education providers on how they could modify their offers to prospective students. This would help them to better align the skills in the labour supply to those needed by employers in the longer term.

Finally, respondents would use the data for diversity monitoring and evaluating policy to investigate if there is a relation between protected characteristics and skills mismatch.

For example, in relation to women, the Cabinet Office equality Hub stated:

“The Women's Returners Research report [...] found that 3 in 5 returning women are likely to end up in lower skilled roles than the ones they held prior to their career breaks. [...] we want to help them back in should they wish to.”

In response, we will continue to assess how feasible it is to produce a variable on skills mismatch and analysis products on this topic.

## **Economic risk created by the coronavirus (COVID-19) pandemic**

We asked users whether they would use a proposed new variable that would identify populations at financial risk because of the pandemic based on the industry they work in. Of the 185 respondents who answered this question, 67% stated they would use it.

Respondents discussed using the data to understand the impact of the coronavirus pandemic on the local authority and in local economy reports. They would use the information to target economic recovery support and diversity monitoring, or when forming upskilling or reskilling strategies. For example, Cambridge City Council stated:

“This variable would be useful in order to influence policies and plan for service demand caused by the economic outcomes of the pandemic. This is particularly crucial as those at economic risk created by the pandemic are a cohort that have not been analysed before as this need did not exist previously. This data would be particularly useful when [...] focusing on skills gaps for job sectors that may not recover as well from the pandemic.”

Other respondents highlighted the disproportionate impact the pandemic had on some industries. For example, the Association of Independent Showmen stated:

“The impact on the travelling Showmen/Women community [...], as result of both the pandemic and measures taken to respond to the pandemic in the context described has been devastating. [...] By having such an “indicator” for this economic sector, [it would] allow accepted and officially recognised analysis to be undertaken to highlight and identify risks to the commercial activities of this community.”

In response, we will continue to assess how feasible it is to produce a classification for the industry variable on economic risk created by the coronavirus (COVID-19) pandemic.

## **Green and STEM industry**

Hertfordshire County Council requested two new industry classifications. One classification would identify green industries and the other science, technology, engineering and mathematics (STEM) industries. They stated:

“Employment in Green industries would be an aggregate of employment in industries involved in clean technologies. The government Ten Point Plan for a Green Industrial Revolution say[s] "By investing in clean technologies – wind, carbon capture, hydrogen and many others – Britain will lead the world into a new Green Industrial Revolution". This variable would enable analysis of the local workforce, set a benchmark level, boost the profile of Green industries and encourage young people to consider obtaining the necessary skills.”

“Employment in STEM industries would be an aggregate of employment in industries involved in Science, Technology, Engineering and Mathematics.”

They also noted that both definitions are very specific, so data users producing these themselves may not do so accurately or consistently.

In response, we will consider producing industry classifications that would allow easy identification of STEM and green industries.

### **Carbon cost of travel to work**

The Local Government Association and Local Area Research and Intelligence Association requested a variable on carbon cost derived from variables on travel to work. They stated:

“Authorities [...] would also welcome the possibility of ONS calculating a carbon cost variable on travel to work transport. This would at least begin to reflect the importance of the climate emergency, which councils are trying to tackle.”

Other organisations did not explicitly request this information, but it reflects the type of data use that they’re discussing more generally related to travel to work data. For example, we discuss planning policies to reduce carbon emissions in the [Travel to work](#) section.

However, we cannot meet the described data need using Census 2021 data. There are two reasons for this.

Firstly, Census 2021 took place at a time when peoples’ travel patterns were impacted by the coronavirus pandemic. For example, more people than usual were working from home. In addition, people continuing to travel to their workplace may have changed their preferred method of transport, for example, to avoid public transport.

Secondly, we could only derive a very crude measure using method of travel to work and distance travelled to work.

There are three main measures of UK greenhouse gas (GHG) emissions. The [UK climate change statistics portal](#) provides more detail about these.

Currently, the main emissions measure available at sub-national level is for local authorities. The Department for Business, Energy and Industrial Strategy (BEIS) provides information on the measure in their article on [Measuring UK greenhouse gas emissions](#).

In response, we have concluded that we cannot produce a variable of sufficient statistical quality on carbon cost of travel to work, from Census 2021 data.

## Health and living arrangements

We proposed four new variables on health and living arrangements. These were:

- COVID-19 health risk
- houses in multiple occupation (HMO)
- multigenerational households
- living apart together

### COVID-19 health risk

We asked users whether they would use a proposed new variable on COVID-19 health risk. We would derive this using the self-reported general health status, in combination with other variables, to identify [vaccination priority groups](#).

Of the 190 respondents who answered this question, 73% said they would use a variable on COVID-19 health risk.

Respondents discussed needing data to inform public health initiatives. For example, Derbyshire County Council stated:

“This could be invaluable to our Public Health team in rolling out public health initiatives and vaccination programmes. The lower the geographies the better as this could help pinpoint where intervention programmes could be targeted.”

Respondents raised needs for data on specific groups that we can identify in Census 2021 data. For example,

- religious groups, such as Jewish
- ethnic groups, such as Gypsy or Irish Traveller, and Kashmiri
- those without access to a car or van
- those living alone

However, respondents also noted significant limitations with using Census 2021 data for this purpose. The alignment with NHS England’s definitions of levels of vulnerability would be relatively low. Census data is also a snapshot in time and could quickly become outdated. Also, local authorities already have access to the Clinically Extremely Vulnerable (CEV) data for their area.

Our feasibility work has shown that the data we could produce from a variable on COVID-19 health risk would not be sufficient to meet the needs of data users. In particular, it would not be accurate enough on the CEV population at small geographies.

In response, we will not produce a variable on COVID-19 health risk derived from data collected in Census 2021.

However, we will investigate the feasibility of combining Census 2021 data with health data sources, to provide more insight into the impact of the coronavirus pandemic on different populations. There is a need for the low level of geography the census provides and to cross-tabulate with other information.

## Houses in multiple occupation

We asked users whether they would use a proposed new variable on houses in multiple occupation (HMO). Out of 187 respondents who answered this question, 72% stated they would use an HMO variable.

Out of 167 respondents who stated whether they needed to be able to separately identify small and large HMOs, 61% did.

Out of the 168 respondents who stated whether they required a specific definition of HMO:

- 46% requested the [government's definition](#)
- 14% requested the 2011 Census definition
- 3% requested an alternative definition
- 37% did not require a specific definition

In support of the government's definition, Birmingham City Council stated:

"We have to follow the definition in the Housing Act 2004 which excludes properties managed by registered social landlords and other non-private institutions, so the definition would need to be the same as this [...] We would also like to know about 'exempt accommodation [EA]' and would need the data in order to distinguish between HMOs and EA."

Respondents said that they could use HMO data to inform local area housing strategies, and annual assessments of housing needs. They could use the data in assessment of housing quality by investigating overcrowding and central heating provision. They would also use it for forecasting housing needs, including provision of parking and refuse collection services. Respondents also raised specific needs in relation to student areas.

For example, Manchester City Council stated:

"Unless an HMO is registered it is hard to separate them out from standard residential homes, so being able to see households with four or fewer unrelated tenants would be of enormous value as these are not standard households and have different needs from a service planning perspective.

Understanding where there are higher numbers of HMOs allows us to monitor whether housing standards are being met in private tenancies, impacts of high numbers of HMOs on the local community, e.g. squeezing out families or older people isolation, and the need for more specific housing such as student accommodation."

Respondents indicated that they would use the data to:

- understand the total number of small and large HMOs in an area
- inform the assessment of what proportion of HMOs in an area are licenced
- develop strategies for selective licencing of large HMOs

Local authorities use selective licencing to manage the concentrations of HMOs in an area. For example, Cheshire West and Chester Council stated:

“This could help us to identify the amount of non-registered HMOs which may require inspection/regulation and where concentrations of HMOs are to decide if licensing or Article 4 restrictions are necessary. It could also provide insight into whether there is a shortage of affordable accommodation or smaller units, making “house sharing” necessary to reduce housing costs.”

Respondents also stated that they would like the data used to produce a full database of HMOs, by identifying those that are not registered. However, census information cannot be used for non-statistical purposes and therefore we would not be able to meet this need. Any data produced must be aggregate statistics. To find out more about how we keep Census 2021 data safe, [visit the ONS website](#).

In response, we will assess how feasible it is to produce an HMO variable derived from data collected in Census 2021. We will aim to define HMOs in line with the government definition to separately identify small and large HMOs.

### **Multigenerational households**

We define a multigenerational household as any household with more than two generations from the same family resident. For example, when children of any age, a parent(s) and a grandparent(s) live together.

We asked users whether they would use a proposed new variable on multigenerational households. Out of the 185 respondents who answered this question, 76% said they would use it.

We also asked users if they required us to provide information for households or for usual residents. Of the 168 respondents who answered this question, 80% requested that we provide it for both. A further 11% requested that we provide it for households and 8% requested that we provide it for usual residents.

Respondents said they would use data on multigenerational households to inform local area housing strategies and annual assessments of housing needs. They could use this data when planning new build developments, and when assessing:

- overcrowding
- housing affordability
- parking provision
- refuse collection services

For example, Bolton Council stated:

“[W]e would look at a possible link with multi-generational households and ethnicity. This would help understand the housing needs of ethnic minorities in Bolton. We would also look at links with graduates/young people and link with housing affordability. It would also help us with a greater understanding of drivers of change around household formation/household forecasting models. Bolton has a larger than average proportion of small terraced housing, and it would help us to look at the best use of our current housing stock. Finally, if it were linked with

occupancy rating information it would help us look at one of the factors behind overcrowding, which would in turn feed into our housing strategy.”

Respondents also discussed this variable with reference to public health planning. For example, the Cabinet Office Equality Hub stated:

“Identifying which ethnic groups and characteristics are most likely to live in multigenerational homes is important for understanding the living conditions for different populations; the past 18 months have illustrated how housing can play a role in health and the ability people have to take protective measures like shielding and distancing within your own home. Continuing to explore this and the ramifications it may have for healthcare more broadly is important for the levelling up agenda.”

Respondents provided limited information about a need to classify different types of multigenerational households. The information they did provide was about the age groups of the household members. For example, Enfield Council stated:

“In the event of another pandemic [...], it would be very helpful to know where there is highest risk of infection within households where there are vulnerable and non-vulnerable age groups living together.”

By cross tabulating the proposed new variable with other variables, users could investigate this information.

Some respondents also mentioned an interest in knowing the total number of generations in a household. However, in Census 2021, we only collected and process information on all relationships for households with five or fewer usual residents. For larger households, we only collected and process a subset of relationships.

In response, we will continue to assess how feasible it is to produce a variable on multigenerational households, for both the usual resident and household population bases.

### **Living apart together**

We asked users about their need for a new variable on living apart together. This represents the population who live separately to their current partners. We can identify those who:

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

We asked users whether they would use a proposed new variable on living apart together. Of the 171 respondents that answered this question, only 39% said they would use it.



Respondents listed that they needed information from this variable to help develop housing strategies. Other existing variables could generally meet these needs. However, respondents acknowledged that they would need additional information that Census 2021 did not collect. For example, information on why the couple were not living together. Basingstoke and Deane Borough Council stated:

“Would potentially use this variable to see if it could show unmet need for housing. For example, if two people were in a couple but still living with their parents as they are unable to afford a home together. Would not be able to ascertain the reasons for living apart together but would be interesting to look at.”

In response, because of low user need, we will not produce a variable on living apart together derived from data collected in Census 2021.

We will continue with our existing proposal for a variable on type of second address, that will include a classification for partner’s address.

## **Accommodation type and vacant addresses**

We proposed four new variables on accommodation type and vacant addresses. These were:

- care home resident
- type of vacant address
- resident in a mobile or temporary structure
- homeless, including people sleeping rough and sofa surfers

### **Care home resident**

In the consultation, we proposed that we might produce a new variable on care home residents. This would be an alternative classification of the existing variable on communal establishment management and type. We proposed this new variable for use on the population base for usual residents in communal establishments, at the Middle Layer Super Output Area (MSOA) level.

We asked users whether they would use a proposed new care home resident variable. Out of 177 respondents that answered this question, 78% said they would use it.

Respondents indicated that they would use data on care home residents for planning future care home provision at a local level. Unlike most other proposed new variables, they prioritised obtaining data for smaller geographical areas over the provision of cross-tabulated data. For example, Cheshire East Council stated:

“Our priority is to identify the geographical areas where care home residents are concentrated, so that we can target social care services and future care home provision at where it is most needed. [...] Cross-tabulating a care home resident indicator with other variables could help to shed light on the characteristics and

needs of different sub-groups of residents, but is likely to be of more limited value, given that many Census variables (e.g. those relating to employment or qualifications) are not relevant to the activities or needs of care home residents.”

Users would also be able to use information from a variable on care home residents in needs assessments required by law, such as the [Social Services and Wellbeing \(Wales\) Act 2014](#).

Respondents listed other needs, including:

- targeting social care services, particularly those provided by the voluntary sector
- understanding the impact of the coronavirus pandemic and planning future pandemic responses
- developing strategies to communicate with care home residents
- diversity monitoring and policy development

For example, the Cabinet Office Equality Hub stated:

“There is some evidence of disparities in the quality of care provision between LGBT people and others. [...] The evidence suggests this may reduce access to care and unwillingness to enter care homes by LGBT people. It would be useful to understand the extent of this as a problem by comparing the proportion of LGB people in care homes with the proportion of the same age in the general population. This would help provide robust evidence on the nature of the issues, and inform whether there is a need to develop policy in this area.”

In response, we will produce a classification for type of accommodation with the categories:

- medical establishment: care home with nursing
- medical establishment: care home without nursing
- medical establishment: other medical or care establishment
- other communal establishment type

We’d aim to produce datasets at MSOA level where possible. However, we will investigate how feasible it is to produce higher priority datasets needed for planning future care home provision at local level for smaller geographies. These priority datasets would cross-tabulate the care home residence information with:

- age
- sex
- general health
- disability

## Type of vacant address

In the consultation, we proposed producing a new variable on type of vacant address. The local census officers provided information on why they thought the address was vacant during the census collection process. The categories would be:

- vacant
- holiday let
- second home

We asked users if they would use information on why an address was vacant. Of the 172 respondents that answered this question, 72% said they would use it.

We also asked users if they required a specific reason for an address being vacant. Out of the 160 respondents who answered:

- 56% requested the proposed categories of vacant, a holiday let, or a second home
- 9% requested another definition
- 35% did not require a specific definition

The respondents who requested other definitions discussed separating vacant into further categories. For example:

- due to death of owner
- undergoing major repairs
- abandonment of property
- unsold new home

They also suggested separating holiday homes into further categories, such as short-term let, and asked for information on how long a property has been vacant. Our local census officers were unable to collect this detailed information from only viewing the outside of properties.

Respondents indicated they would use information on why an address was vacant to compare with locally held data, such as Council Tax records or the electoral roll. Local authorities use this existing data to develop and action local area housing strategies and annual assessments of housing needs.

It would also feed into empty homes strategies. For example, Enfield Council stated:

“In common with other local authorities, we have schemes to encourage empty property owners to let their homes to local families at risk of homelessness.”

Finally, respondents discussed information on second homes and holiday lets in relation to understanding community cohesion. For example, the Welsh Language Commissioner stated:

“The Programme for government [...] commits to 'creating a Welsh, language Communities Housing Plan'. Providing this information would help us to

understand more about the location of empty homes, their nature, and their potential impact on the viability of Welsh speaking communities.”

Some respondents also requested information about the owners of second homes. We're planning datasets for the second homes population base, which will better meet these respondents' needs. We discuss these in the [Second address population base](#) section.

In response, we will continue to assess how feasible it is to produce data on the possible reason for a property being vacant. If we produce this information, it would be in the form of a methodological paper.

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

We asked users whether they would use a proposed new variable that indicated if a person is resident in a mobile or temporary structure. Of the 171 respondents who answered this question, 73% said they would.

We also asked if they'd use information on the type of mobile or temporary structure that Census 2021 collected. We gathered some information on this during the census collection for operational purposes.

We also asked users about their need to further categorise the population resident in a mobile or temporary structure as living in:

- boats
- tents and caravans
- other accommodation

Out of the 157 respondents who stated if further categorisation would be useful, 67% said it would. However, respondents listed data for smaller geographical areas and data that is cross tabulated with other variables as higher priorities.

Respondents discussed very similar needs to those they stated for the overarching accommodation type variable. We describe these needs in more detail in the “Type of accommodation” subsection of the [consultation response: part one \(pdf, 625 KB\)](#) document.

A few respondents told us about additional needs they have for data on:

- the length of residence in a mobile or temporary structure
- the reason for residing in this form of accommodation
- whether the residence moves geographical location, for example, a canal boat changing mooring location

Respondents want to use this data to further understand the health, education and employment outcomes of transient communities. They would also use it to explore these outcomes for those residing in these forms of accommodation through need, rather than choice.

However, we did not collect statistical or operational data on these additional factors in Census 2021. We also cannot assume that all those living in this form of accommodation are transient. For example, the Showmen's Guild of Great Britain stated:

"Many / most of the Showmen's / Showwomen's Community live in temporary structures (defined solely in Planning terms) and or caravans.

They live ON THEIR OWN LAND and or live intermittently alongside their Show (Fair / Circus) operations on pre-agreed land operating their shows. [...] They are NEVER "transient"; even though they travel FOR THE PURPOSE OF THEIR CRAFT."

Respondents only provided limited information about why they need further detail on the type of mobile or temporary structure. These are to:

- understand the impact of type of temporary or mobile accommodation on health, education and employment outcomes
- evaluate the effectiveness of policies on factors such as planning permission and moorings provision
- feed into local authorities Gypsy and Traveller Community Accommodation Assessments

We've conducted feasibility work and determined that the coverage of the operational data would be too low to produce quality estimates of the size of these different populations. For example, we only collected this information for temporary or mobile accommodations not on our address register.

In response, we will not produce a variable on type of temporary or mobile accommodation derived from Census 2021 operational data. This is due to the:

- low alignment of the operational data to the statistical needs
- strength of user need focussing on the whole population living in mobile or temporary accommodation
- socio-cultural questions allowing identification of some of the specific communities requested

We will meet the need for information on the whole population living in mobile or temporary accommodation through a classification for type of accommodation.

We propose to include the following categories:

- whole house or bungalow
- flat, maisonette or apartment
- caravan or other mobile or temporary structure

We would aim to make this classification available on the:

- household population base, at output area level

- usual residents in households population base, at MSOA level

### **Homeless, including people sleeping rough and sofa surfers**

We asked users whether they would use information on the homeless population that Census 2021 collected, if we were able to produce it. Out of the 179 respondents who answered this question, 82% stated they would use it.

The Census 2021 form collected three pieces of information that we could use to produce information about the homeless population.

Firstly, the communal establishment form collected information on type of communal establishment. This included the response option “hostel or temporary shelter for homeless people”. From this, we may be able to produce some data on populations of people sleeping rough.

Secondly, the individual census form completed by communal establishment residents collected information on type of resident. This included the response option “Staying temporarily (no usual UK address)”. From this, we may be able to produce some analysis of people with no usual address staying in communal establishments at the time of the census.

Thirdly, the household form collected information on type of resident. This included the response option “People staying temporarily who usually live in the UK but do not have another UK address, for example, relatives, friends”. From this, we may be able to produce analysis of households that have temporary residents with no other UK address.

None of these measures would provide a robust count of the homeless population, even if used in combination.

We asked users which of these information sources would be of most interest. Of the 163 respondents who answered this question:

- 95% would use the information from the communal establishment questionnaire
- 89% would use the information from the individual census forms
- 85% would use the information from the household forms
- 12% stated they had additional needs, relating to other specific types of communal establishments providing travel or temporary accommodation

Respondents stated they need data on all types of homelessness to understand the extent of homelessness in their area, and to enable the targeting of services. They also emphasised a lack of alternative data in this area. For example, Hampshire County Council stated:

“This data would be included in the [Joint Strategic Needs Assessment] JSNA Inclusion Health Groups analytical work to try and understand the size of this population group and demographic characteristics. This would help assess their needs and inform service provision, currently these data are not available reliably elsewhere and they represent an excluded group who have high health and social care needs.”

Other respondents noted the transient nature of the homeless population. They were concerned that this Census 2021 data would be of limited use in targeting services, as the population distribution may have already changed. For example, the [Everyone In](#) initiative was launched by the Housing Secretary at the start of the pandemic to protect rough sleepers. This meant that a larger proportion of this population than usual was residing in travel or temporary accommodation during Census 2021.

In response, we will produce a reduced classification on type of communal establishment. We propose the following categories:

- travel or temporary accommodation: Hotel, guest house, B&B, youth hostel
- travel or temporary accommodation: Hostel or temporary shelter for homeless people
- travel or temporary accommodation: Holiday accommodation (for example, holiday parks)
- travel or temporary accommodation: Other travel or temporary accommodation
- other communal establishment

We will also produce a classification on position in communal establishment. We propose the following categories:

- resident
- staff or owner, or family member or partner of staff or owner
- staying temporarily (no usual UK address)

Finally, we have been exploring how feasible it is to produce a variable including information from the response option “People staying temporarily who usually live in the UK but do not have another UK address, for example, relatives, friends”. This variable would have allowed analysis of households that have temporary residents with no other UK address. However, we have determined that the data would not be of sufficient quality to publish without further significant processing.

In response, we are unlikely to produce a variable including information from the response option “People staying temporarily who usually live in the UK but do not have another UK address, for example, relatives, friends”.

Because of the strong need for information on this population, we’re continuing to assess if there are any alternative options to provide information on this population.

## Small populations

Small populations are groups which are defined by their cultural background, including characteristics such as ethnic group, country of birth, religion and national identity. Because of the relatively small sizes of these populations, we are unable to release detailed information on these groups in the standard outputs because of confidentiality constraints. Instead, we create bespoke datasets for specific small populations at all geographies where the number of that population exceeds a set threshold.

We provided our small populations proposals in the [Census 2021 outputs consultation proposals document \(pdf, 929KB\)](#). We proposed outputs at Middle Layer Super Output Area (MSOA) level and at local authority level. Our proposal largely repeated what we produced for 2011 Census.

Respondents noted that it was difficult to judge what populations were needed before seeing the main datasets. As such, they stated that they would appreciate an opportunity at a later stage, to provide feedback on their needs for small population data.

In response, we will conduct a user survey after the completion of phase three of the Census 2021 outputs release schedule. This will assess if there are any additional needs for small population datasets that have not yet been adequately met.

## MSOA-level datasets

We proposed to produce datasets for six small populations at MSOA level. We've provided these proposals in [Annex C: Initial proposals for small population datasets](#). We asked respondents which of the proposed populations they would use data from. Of the 91 respondents that answered this question:

- 60% would use data on the Kashmiri population, defined using ethnic group
- 59% would use data on the Sikh population, defined using ethnic group or religion
- 44% would use data on the Nepali/Nepalese population, including Gurkha, defined using ethnic group
- 36% would use data on the Jain population, defined using ethnic group or religion
- 25% would use data on the Cornish population, defined using national identity
- 25% would use data on the Ravidassia population, defined using religion

Of the 126 respondents that stated whether these datasets would meet their needs:

- 66% stated that they fully meet their needs
- 15% stated that they partially meet their needs
- 19% stated that they would not meet their needs

Respondents discussed a need for data at lower levels of geography such as Lower Layer Super Output Area (LSOA) and ward level. They stated that the MSOA geography



is relatively large and often crosses administrative boundaries. As a result, this is not suitable for local authority planning.

They also discussed needing information on the main language within small populations. For example, regarding main language, Manchester City Council stated:

“[D]ata on small populations is important for both understanding who lives where, which communities are growing or emerging, racial tension or community cohesion and, most importantly, who and where we need to be aware of to focus on if there is a crisis, e.g. Delta variant of COVID19 needing to identify where Indian subcontinent residents were located or contacting Afghan residents during current crisis. [...] All communication may require knowledge of which languages are spoken in an area for translation services and awareness of cultural considerations. This may need to be done urgently so needs to be planned using a rich data source such as the census, but also need to consider the cost of translation services, so more knowledge of small populations means we can be more efficient.”

Respondents provided justification of a need for several small populations in addition to those listed in [Annex C: Initial proposals for small population datasets](#) at MSOA level.

These are:

- British Sign Language (BSL) main language
- Romania country of birth
- Somali ethnic group
- Turkish Cypriot ethnic group

For example, when discussing the population who have BSL as a main language, the Royal Association for Deaf people stated:

“Data on deaf BSL users is severely lacking and remains a constant source of frustration for deaf charities and those working to improve the lives of deaf people. Therefore, by providing a separate analysis and resource on deaf BSL users specifically, ONS will be significantly helping deaf organisations in their efforts to ensure that policymaking and commissioning is informed by evidence.”

When discussing the Turkish Cypriot population, the UK Council of Turkish Cypriot Associations (CTCA) stated:

“We support the Turkish Cypriot community in all areas. The data will be used to determine CTCA’s policy development and in integration for example supporting English as a foreign language.”

We will continue with our proposal to produce a suite of Census 2021 datasets similar to the 2011 Census datasets on small populations.

In response, we will investigate how feasible it is to add a dataset on main language for each small population. We will also investigate if we can produce any of the datasets for

either wards or LSOAs. In this investigation, we will consider utility of the data and statistical disclosure risk.

We will also consider the need to produce a set of small population datasets for the four small populations we've discussed. Respondents provided examples of why users need this information for these populations.

In this assessment, we will consider what data will be available through multivariate datasets released in phase two of the release schedule. This includes both the ready-made datasets and the build-your-own datasets.

It may not be feasible for us to produce data at MSOA level, if, for example, the population is not clustered in specific areas. In these cases, we will consider producing datasets at local authority level.

### **Local authority level five-year age bands by sex datasets**

We proposed to produce datasets at local authority level for the same 30 small populations as we did in 2011. This was made up of 13 ethnic groups and 17 countries of birth. We've listed these in [Annex C: Initial proposals for small population datasets](#).

Of the 149 respondents that stated whether these datasets would meet their needs, 72% stated that they would.

As with the MSOA-level datasets, respondents noted that it was difficult to judge what populations they needed, before seeing the main datasets.

Respondents also raised needs for datasets on other populations at local authority level, for five-year age bands by sex. These are:

- Caribbean ethnic group
- China as country of birth
- Jewish identity, with the defining variables to be determined
- Pashtun ethnic group
- Showman ethnic group

If possible, respondents also requested that we produce additional datasets on these populations. For example, Showmen's Mental Health Awareness Charity stated:

“[W]e are a charity providing mental health services to the Showmen community. The charity is [...] limited to offering our services only to those community members we know of. We need census data to give a true sense of our population, and identify ways to serve the community better.”

In response, we will also investigate how feasible it is to produce extra datasets on these five small populations, where respondents provided a user need. We will work with data users to understand how to define these groups. For example, we will consider if

“Caribbean” should be defined using the “Caribbean” response option only or in another way.

In this assessment, we will again consider what data will be available through the multivariate datasets that we release in phase two of the release schedule.

## Alternative population bases

We asked users what datasets and analysis they would need on alternative population bases. We defined these proposed alternative population bases in the [Census 2021 outputs consultation proposals document \(pdf, 929KB\)](#). These alternative population bases either move usual residents to different locations based on their characteristics or report on a population that’s not usually resident.

For each alternative population base, we stated the geography at which we propose to produce datasets. We did not provide proposals for what variables we will be able to include.

As the coronavirus (COVID-19) pandemic restricted the population’s movement at the time of Census 2021, the potential for disclosure is higher for alternative population bases. We will only produce datasets where:

- there is a defined user need
- the data is of sufficient quality
- the risk of statistical disclosure is low

However, we acknowledge the strong user need for information on these populations. We discuss our future plans in the section [Census during a period of change: conclusions](#).

Of the 173 respondents that stated which alternative population bases they would use, if we were to produce them:

- 65% would use the workplace population base
- 61% would use the workday population base
- 50% would use the out-of-term population base
- 53% would use the non-UK born, short-term resident population base
- 51% would use the second address population base
- 5% would use another population base if it were available
- 24% would not use any of the proposed alternative population bases

Most of the requests, which respondents made for another population base, were for subsets of the usual resident population. We will meet these needs through multivariate datasets that we produce in phase two of the release schedule.

Respondents made one other request for a version of the workplace population that focuses on education institutions. We did not collect information on travel to education.

This is because respondents only expressed a low level of need for this, during the [2021 Census: initial view on content for England and Wales](#) consultation in 2015.

In response, we will produce outputs on the five alternative population bases that we outlined in our proposals. We do not plan to produce any other alternative population bases.

## **Workplace population**

We asked users to describe their proposed analysis on the workplace population and 95 respondents answered. They stated that our proposals for Middle Layer Super Output Areas (MSOAs) datasets would be useful. However, they also requested data to lower levels of geography where feasible. To enable more detailed analysis of local areas, respondents requested Lower Layer Super Output Areas (LSOAs) and workplace zones.

Respondents provided several uses for this data. These are:

- transport, housing, and employment land strategies and provision
- economic growth strategies, including coronavirus pandemic recovery planning
- emergency evacuation planning, for example, in the case of flooding
- understanding crime
- equality monitoring

For example, Cheshire East Council stated:

“Data on the workplace population at MSOA level [...] would be extremely useful in informing our response to emergency situations such as floods, as such responses (e.g. evacuation/ rescue efforts) need to take account of the number of people potentially present (or travelling to/ from) the affected areas and the geographical distribution of people within those areas.

A breakdown of the workplace population by industry [...] would be helpful in assessing/ supporting the needs of local businesses and in making decisions about the scale and location of future provision of employment land (as different industries have very different needs: for example, manufacturing sites and warehouses need much more land per employee than offices do).”

In addition, Manchester City Council stated:

“[I]t is essential that we have an understanding of our workplace population in 2021, even if it is to indicate those missing due to lockdown, so that we can make policy decisions about recovery and the future of the city centre. Cross tabbed with ethnic group, sexual orientation and other protected characteristics it will inform our equality strategies.

It can also be used to explain crime statistics and the night-time economy as well as planning for services that are required from those who do not live in the city, e.g. transport and waste services. Cross tabbed against industry or occupation it would assist our economic policy decisions as we could compare against our forecasts of employment growth, particularly in the wider city centre where pre-

COVID19 we saw rapid employment growth that has then been followed by changes occurring in industries during COVID19 and then planning for future policies on climate change.”

Respondents demonstrated a clear need for datasets about the workplace population on:

- age
- highest level of education
- occupation
- industry
- method of travel to work
- distance travelled to work

To understand the characteristics of the population and for diversity monitoring, respondents also requested demographic characteristics.

Finally, respondents requested that we expand the definition of the workplace population to include all ages, not just ages 16 to 74, as in 2011.

In response, when finalising our plans for datasets on the workplace population, we will prioritise assessing whether it is feasible to produce univariate datasets at MSOA level. We will focus on the variables for which respondents demonstrated a clear need.

We will assess the feasibility of producing any of these higher priority datasets for lower levels of geography such as LSOA and workplace zone. In some cases, we may need to produce datasets at a higher level of geography.

We will also investigate the feasibility of expanding the definition of the workplace definition to include those aged 75 years and over.

## **Workday population**

We asked users to describe their proposed analysis for the workday population and 85 respondents answered. Several stated that this data was more important following the coronavirus pandemic because more people are working from home. Respondents were supportive of the proposal to produce this data at OA. They also requested data for workplace zones.

Respondents provided similar uses for data on the workday population as they had for workplace data. However, to investigate the impact of the coronavirus pandemic, respondents also raised needs to compare data on the workday population against the usual resident population and with data from the 2011 Census.

Respondents indicated that they needed similar variables to those prioritised for the workplace population. However, they provided a greater emphasis on the need for social and demographic variables. For example, the Greater London Authority stated:

“Workday population data are used to monitor activity in highstreets and town centres. Regeneration of high streets form one of the pillars of the GLA recovery programme. So we need to know the characteristics of workers and residents in these areas by demographics and protected characteristics: age, sex, ethnicity, health and disability, sexual orientation and gender identity.”

In response, when finalising our plans for datasets on the workday population, we will prioritise producing univariate datasets on the same variables as for the workplace population. We will generally aim to produce these univariate datasets at OA level.

We will assess how feasible it is to produce any of these higher priority datasets for workday zones. We will also aim to produce datasets comparable with datasets from the 2011 Census, where feasible.

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

We asked users to describe their proposed analysis for the out-of-term population and 72 respondents answered. They were supportive of the proposal to produce this data at OA level and for similar variables that we provided following the 2011 Census.

Respondents indicated they needed this data to help understand changes to required levels of service provision in and out of term-time. For example, the Greater Manchester Combined Authority stated:

“[B]us passenger numbers on certain routes can be quite different in the summer months in particular. Customer levels in leisure and retail services in certain locations change too in out-of-term times.

There are also higher and lower demand on other services such as refuse, GP appointments and NHS services etc. Similar datasets to those released in 2011 would be helpful. In particular, we are likely to find population density, age, ethnicity and general health. The student population is generally young and healthy and if they are excluded from rates on health indicators through out-of-term population estimates then this might show much higher rates of health problems in the general population.”

Respondents only provided limited information about the exact variables they need. They indicated that local area totals were the highest priority. However, there were requests across multiple responses for:

- age
- sex
- ethnic group

In response, when finalising our plans for out-of-term population datasets, we will prioritise univariate datasets on age, sex and ethnic group. We will generally produce univariate datasets at OA level.

## **Non-UK born short-term resident population**

We asked users to describe their proposed analysis for the 'non-UK born short-term resident' population and 65 respondents answered. They noted that the regional level data we proposed would have limited use.

As a result of the coronavirus pandemic on changes to student and migrant worker flows, respondents noted a reduction in the use of this data.

For example, Tameside Metropolitan Borough Council stated:

“Due to the COVID-19 pandemic the numbers recorded in the Census might not be typical, particularly for an area like Tameside bordering the higher population density area of Manchester city centre, as the flow of people out of city centres to satellite towns through the pandemic has been observed across the country.”

Some respondents described analysis that they could undertake and requested data at local authority level. However, very few linked this to specific policy, planning, or service provision needs. Manchester City Council provided the most defined need but noted limitations of the data. They stated:

“Would normally be looking at volume to supplement population estimates on the basis that one [short term resident] STR is replaced by another on a repeating basis so requiring the same services by and large as a 'usual' resident even though different people e.g. a bin emptying, regardless of how often that person is replaced by another. Analysis will be limited because the pandemic has affected transience.”

Some users also requested information on the reason for migration. We did not collect this data in Census 2021.

We will produce estimates of the short-term resident population, at local authority level in phase one of our release schedule.

We will continue with our proposal to produce datasets at regional level. These datasets will provide some basic demographic, social-cultural and labour market information with accompanying commentary. We will produce estimates of the short-term resident population, at local authority level. We aim to bring the release of these datasets forward from phase three of the release schedule, into phase two. This is because of the importance of understanding changes in migration patterns as a result of the coronavirus pandemic.

## **Second address population base**

The second address population looks at the characteristics of the population with a second address, including the characteristics of the second address itself. When producing datasets on the characteristics of the population with a second address, we count them at the location of their second address rather than at the address where they're usually resident.

We asked users to describe their proposed analysis for the second address population base and 67 respondents answered. They noted that our proposal for data at local authority level would be useful. Where feasible, respondents also requested data to lower levels of geography, such as LSOA-level data to enable local area analysis.

Respondents stated they would use this data for:

- forecasting populations, particularly around needs for educational provision and water demand
- preparing housing strategies, particularly around current use and availability of housing stock
- assessing of the impact of the coronavirus pandemic on student accommodation
- understanding the accuracy of aggregate council tax data
- understanding the impact of concentrations of second addresses on local communities, for example house prices

Respondents provided varied needs in relation to rural and urban areas. For example, when referring to a more rural area, Shropshire Council stated:

“Shropshire is a beautiful county which attracts holiday-makers / city dwellers who later purchase second homes. Following COVID and the rise in home-working it is likely that the second residence population has risen. The 2021 Census will help our understanding of this population and the impact on public services. Data on second homes / population is also analysed in the Shropshire housing market assessment, as many second home-owners later choose to retire to the County.”

Whereas, in reference to an urban area, Manchester City Council stated:

“Colleagues in Housing services will use holiday homes data to support issues with properties being used as short term holiday lets in the city centre. Students’ out-of-term-time address will supplement our monitoring of the impact students have on the city in terms of accommodation, particularly in areas where there are housing shortages for families.”

Respondents requested datasets that would allow more detailed analysis of the households identified as a second address for one or more dependent children. They requested that the datasets allow the separate identification second addresses of:

- the parent or parents with whom the child is recorded as being usually resident
- another parent with whom the child lives part of the time

Supporting this request, the Fatherhood Institute stated:

“This is an issue of key policy importance, hence the DWP producing Separated Families statistics. In previous data sources, around a third of dependent children have separated parents; and between 30% and 50% of child/ren stay over regularly with their ‘non-resident’ father [...]. The Social Security Advisory Committee drew attention to this issue [...]: “The terms “non-resident” and



“resident” parent [...] fail to reflect the variation in shared care arrangements, including equal shared care”

In their feedback about the second address population base, respondents also discussed some needs related to the household population base. They requested estimates of the total number of second residences and the type of second residence. Respondents also expressed a particular interest in the usual resident population who have second addresses for work or for study.

For the second address population base, respondents evidenced a need for information on a second address residents:

- age
- employment status
- place of usual residence

In response, when finalising our plans for datasets on the population with a second address, we will prioritise univariate datasets on age, employment status and place of usual residence. We will generally produce these univariate datasets at local authority level.

We will investigate how feasible it is to produce more detailed analysis of households identified as a second address for one or more dependent children. This analysis will require us to link within the census data and will depend on the quality of the address response for the second address.

We will also assess whether we can produce any of these datasets to a lower level of geography such as LSOA.

## **Origin-destination data**

We did not consult on our proposals for origin-destination products in this consultation. However, some users did provide needs regarding these products.

Respondents noted the impact Scotland’s census moving to 2022. They requested that data on flows include origins or destinations in Scotland where possible, at the same level of geography as within England and Wales.

Where possible, we intend to provide Census 2021 origin-destination datasets for England, Wales and Northern Ireland combined. Scottish geographies, based on their census in 2022, are unlikely to be available at the time we produce these.

In response, we plan to provide initial datasets with higher level Scottish geographies, which are set to remain unchanged. This is only for addresses one year ago, workplace addresses or second addresses that were listed as in Scotland. Once the information is available, we will investigate whether it is feasible to incorporate more detailed geographies for Scotland into the origin-destination datasets.

If you have any questions about Census 2021 [origin-destination products](#) or are interested in joining our working group, please email [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk).

## Microdata

We did not consult on our proposals for microdata products in this consultation. However, in [part one of our consultation response \(pdf, 625KB\)](#), we committed to considering how feasible it would be to produce a household-level variable on proxy response in the secure microdata datasets.

With this, we'd aim to identify households where someone from outside of the household completed their census form on their behalf. This information could help address the need for data on digital exclusion.

If you have any questions about Census 2021 [microdata products](#) or are interested in joining our working group, please email [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk).

## Analysis

We will carry out a detailed census analysis programme based on Census 2021 data from England and Wales. Alongside the production of facts and figures about the population, the programme will aim to produce statistics that shine a light on public policy issues. This will:

- provide insight into the data
- help users to better understand their community's needs
- inform decisions on future policy and services

We provided detailed proposals for the analysis programme within the [Census 2021 analysis programme proposals \(PDF, 505KB\)](#) document.

We asked respondents if these plans met their needs. Firstly, through a general question, and then through specific questions across relevant sections of the consultation. In finalising our analysis programme, we will consider the feedback from the consultation alongside feedback from other stakeholder engagement activities.

Of the 220 respondents who stated whether the proposals for the census analysis programme met their needs:

- 70% stated that they fully met their needs
- 25% stated that they partially met their needs
- 5% stated that they did not meet their needs

When considering specific areas:

- 79% of 204 respondents stated that the analysis proposals for sexual orientation and gender identity fully met their needs
- 83% of 172 respondents stated that the analysis proposals for UK armed forces veterans fully met their needs
- 65% of 219 respondents stated that the analysis proposals for ethnic group fully met their needs
- 70% of 142 respondents stated for the analysis proposals for small populations fully met their needs
- 80% of 143 respondents stated that the analysis proposals for alternative population base fully met their needs

Some of the analysis proposals that we plan to publish after phase three of the release schedule, require us to combine Census 2021 data with other sources of information such as administrative data. The Welsh Government noted that some of the administrative datasets discussed only covered the population of England. They requested that we explore projects combining data sources about the population of Wales more fully.

We also noted the importance of analysis to understand the impacts of the current period of social and economic change.

In response, we will explore the potential to combine Census 2021 data with administrative datasets about the population of Wales

We will ensure that our plans include analysis looking across topics at the impact of large changes such as Brexit and the coronavirus (COVID-19) pandemic.

We will update our [list of proposed analysis reports](#) to reflect our updated plans. We will do this in stages. We will start by publishing the list of the reports we plan to publish within the main release schedule by early summer 2022. A full list, including plans for analysis reports discussing UK data, will be available by the end of phase three of the release schedule.

## **Rural or urban classification of areas**

Respondents welcomed the planned analyses that included discussion of the differences between rural and urban areas. However, they noted that these did not fully meet their needs and requested that other analyses also considered this classification of areas. For example, Cheshire East Council stated:

“In particular, it is important for us to understand how the housing characteristics of our rural areas differ from urban areas, so that we can tailor housing provision and housing policies accordingly. A comparison of rural and urban areas in terms of overcrowding, housing tenure and heating systems would therefore be very

useful. Similarly, we need to ensure that our rural population has good access to employment opportunities, so comparisons of rural and urban areas in terms of economic activity (whether employed/ unemployed/ economically inactive), qualification levels, method of travel to work and car/ van access would help greatly.”

In response, we will consider which topics would benefit from analysis of the difference between rural and urban areas.

## **Demography and second addresses**

Respondents requested analysis of the number of children living with same sex couples. Providing justification of this request, Shropshire Council stated:

“The more information available, the more useful this would be in terms of the PHSE requirements for sex education in schools. Educators could then draw on the analysis / local data and could also ideally use this data to cross reference with those who have been adopted by same sex couples or by single people who are of differing sexual orientations. This could help to encourage more potential foster parents and adoptive parents to come forward on basis that the local authority will be supportive of such parents having already supported others in this process.”

They also requested analysis of the households identified as a second address for one or more dependent children. We discuss this request in the section [Second address population base](#). This analysis will require us to link within the census data and will depend on the quality of the address response for the second address.

In response, we will incorporate discussion of these topics into our analysis programme.

## **Migration**

An individual responding in a professional capacity requested a greater focus on analyses covering time since migration.

In addition, Manchester City Council requested analysis of populations from countries in conflict. Providing justification of this request, they stated:

“We have an increasing Arab community, some of whom will be very wealthy but we will also have received many refugees and asylum seekers from countries in conflict with not a lot of information about them, so any further small populations from Arabic-speaking countries, particularly African such as Libya and Egypt, and Middle Eastern such as Syria, Iraq and Iran, would be beneficial. [...] Potentially other countries in crisis such as Albania or Belarus, Ethiopia or Sudan. This is all

so that we can be responsive to emergencies abroad that may have an impact on our residents.

In response, we will aim to incorporate discussion of migration from countries in conflict into our analysis programme. This will be subject to statistical disclosure control.

## **Ethnic group, national identity and religion**

Whilst respondents welcomed some of the analyses planned on ethnic group, national identity and religion, they noted a focus on specific communities within the population.

From the information provided in the census questionnaire response options, respondents requested analysis of ethnic groups including:

- Arab
- Chinese
- Pakistani

From the information provided about identity in the census questionnaire write-in fields, respondents requested analysis of ethnic groups including:

- Eastern European
- Kashmiri
- Latin American
- Nigerian
- Pashtun
- Romanian
- Showmen

The most justified need respondents provided was for the Showman population. For example, the Association of Independent Showmen stated:

“Our belief that official statistics like this should reflect the fact that persons targeted for discrimination based on descent, in particular caste-based discrimination and related practices (of which Showmen face in planning applications by way of example, and by extension will face in any new changes to such legislation), are in a number of contexts in a particularly marginalized position and in need of focused attention. Without such recognition Showmen will remain at the mercy of inconsistency in the application of any legislation within the field of planning, as an example, by local authorities and in so doing will nullify or impair their equal enjoyment of human rights.

The multiple questions regarding Showmen identity, culture and way of life are inextricably interconnected to nomadism and are never surveyed concurrently [...]

and is yet another example of how the unique business and living arrangements of Showmen pose such a legislative puzzle.”

Respondents requested that there be further analyses of ethnic group against indicators of inequality such as:

- health
- housing
- education
- employment

For example, the Cabinet Office Equality Hub stated:

“Many of the themes we are interested in are already detailed in the analysis proposals, only without ethnicity included as a characteristic. We feel that including an ethnicity breakdown in these existing plans would be easier than duplicating those outputs with a specific ethnicity lens. It also shows that ethnicity analyses have parity of esteem with other population characteristics. [...]

Labour market status by ethnicity is an important topic. We have previously used APS data on economic inactivity to inform policy supporting South Asian women into employment. Census data would provide snapshot data of higher quality than standard surveys”

Respondents also requested analysis of some religious populations where their religious affiliation was listed as a response option on the form, particularly for the Muslim population. For example, Cheshire East Council stated:

“As ONS will know, the 2011 Census indicated that England & Wales had 2.7m Muslims, making them by far the largest religious minority group (followed by Hindus – 817,000 in 2011 - for whom there is also no proposed 2021 Census analysis); in Cheshire East, Muslims (2,400 in total) were also the largest religious minority, followed by Hindus (1,300). In programming its analysis of minority groups, ONS should also prioritise those groups that evidence suggests are at risk of discrimination and/ or exclusion – which adds further to the case for an analysis of the Muslim population.”

In addition, respondents made some analysis requests around the information provided in the write-in fields for religious affiliation, such as Ravidassia.

Discussing a proposal for analysis of the interaction between religion and ethnic group more generally, the Muslim Council of Britain stated:

“We would like to explore the relationship between religion and ethnicity in terms of socio-economic outcomes, particularly in the labour market. Cross tabulations providing greater ethnic breakdown of current ethnicity categories by religion and

other variables is of particular importance for Muslims as many groups are 'hidden' under umbrella categories such as white or white other. Studies have shown that Muslims from a wide array of groups, such as Arabs, Iranians and Eastern Europeans classify themselves under such broad heading when asked to. This however does not allow for exploration of disadvantage amongst specific Muslim ethnic communities.”

In response, we will consider which additional topics would benefit from analysis considering ethnic group and religious group. We will also consider the need for specific analysis of additional ethnic and religious communities.

## **Language skills**

Respondents welcomed the analyses planned on language skills. However, they also suggested additional dimensions to explore from the perspective of service provision needs. For example, Migration Yorkshire stated:

“We strongly urge the Census 2021 analysis to provide details of languages other than English in the 'Proficiency in English language' analysis. In the 2011 Census, the data only told us numbers/proportions of people whose first language wasn't English, but no detail about what their first language actually was. This makes a huge difference in applicability. We constantly are asked about the language needs of migrant communities, whether in translated materials targeted at vulnerable groups, or in the provision of English language support.”

Another request respondents made was for analysis of both older populations and care home residents by main language. Kashmir Development Foundation stated:

“This is important because many older age British Kashmiri nationals have limited English language skills and our community research and knowledge suggested that due to dementia and other factors many older adults who may have limited English language skills before revert back to their original (Childhood) language, Pahari or Mirpuri and can not speak or understand English at all.

Consequently, the lack of communication in the language they can speak or understand is one of the major contributing factor for them to remain in ill health longer or not receiving appropriate health and social care services as compare to their peers who are in similar circumstances.”

In relation to analysis of the Welsh language speaking community, respondents also requested analysis of information on second homes. Providing justification of this request, the Welsh Government stated:

“One area of increased policy interest currently is second addresses/holiday homes, particularly in Welsh-speaking communities. Anything that the 2021

Census could provide in this context would be useful in our understanding of the relationship between the two, and to plan any relevant policy interventions.”

In response, we will consider English language proficiency and main language together in our analysis programme where possible. We will also consider the inclusion of main language and English language proficiency in analysis of the care home population.

We will produce analysis of the distribution of holiday homes, including those within Welsh-speaking communities.

## **Health and disability**

Respondents welcomed planned analyses on health and disability. They also requested that we explore some additional dimensions. For example, in relation to social mobility, Cabinet Office Equality Hub requested consideration of ethnic group as well as health. They stated:

“RDU had been undertaking analysis on social mobility by ethnicity with ONS prior to the pandemic and remain interested in any findings relating to how social mobility varies by ethnic group. We are particularly interested in the health impacts element as the past 18 months have highlighted the importance of understanding and resolving health.”

Respondents requested that we analyse health and disability across a wider range of topic areas. These topics include:

- demography
- education
- UK armed forces veterans

For education, the Cabinet Office Equality Hub stated:

“In line with the levelling up agenda, it would be valuable to understand geographic variation in the highest level of qualification held, disaggregated by disability status (EDU-02). Furthermore, disabled people are disproportionately more likely to have no qualifications, thus understanding the impact of no qualifications on life chances by disability status is important (EDU-09).”

In response, we will consider which additional topics would benefit from analysis considering health and disability.

## **Sexual orientation and gender identity**

Census 2021 is the first census for England and Wales that asked questions on the topics of sexual orientation and gender identity. Because of this, respondents noted that all analysis on these topics is important.



Respondents discussed needing analysis of the detailed classifications to be derived from the write-in responses to both questions. They also noted the interaction between sex and gender identity.

Some respondents requested that we discuss both topics of sexual orientation and gender identity in relation to all other topics collected in Census 2021. The examples that respondents provided focused on areas of inequality, including:

- health
- housing
- education
- employment

For example, the Welsh Government stated:

“We need breakdowns of sexual orientation and gender identity by qualification and unpaid care [...]. Again, this is to understand potential systemic issues for both of these groups in these areas, and how we can target services/policies to address these if there are differences within these populations compared with others. In addition, there is interest in understanding responses to the voluntary question on religion, to understand if there are any disparities between groups.

We will produce analysis discussing detailed classifications of sexual orientation and gender identity.

We will focus our analysis plans for sexual orientation and gender identity on areas of potential inequality.

## **UK armed forces veterans**

Census 2021 is the first census for England and Wales that asked if people had previously served in the UK armed forces. Because of this, respondents welcomed analysis plans covering areas of potential inequality. These include:

- health
- housing
- education
- employment

For example, Flintshire County Council stated:

“We have signed the Armed Forces Covenant and are committed to identifying and meeting the needs of the Armed Forces community. The proposed Armed Forces legislation will require us to pay due regard to the needs of veterans and the Armed Forces community. This data will contribute to understanding the

Armed Forces community within the local authority area and planning to meet their needs.”

In the consultation proposals, we said that we were planning to publish analysis of the UK armed forces veteran population, by health and disability, after March 2023. This is because we were planning to combine Census 2021 data with data on suicides, in this analysis. This piece of analysis is complex and will take us more time.

However, respondents requested that we produce analysis of the UK armed forces veteran population by health and disability, earlier than currently planned. They would use this analysis to support them in meeting the requirements of the Duty of Due Regard contained within [The Armed Forces Bill](#).

Respondents noted specific needs the UK Nepali Gurkha community have within the UK armed forces veteran population. The UK Nepal Friendship Society stated:

“We are aware the UK Nepali Gurkha community has a number of major sub-populations with different experiences of support on integration. It also includes the families of serving and ex-Brigade of Gurkhas, that have additional particular support and integration needs.”

In response, we will bring forward analysis of the UK armed forces veteran population by health and disability from phase three of our release schedule into phase two. We will take forward analysis of suicide data in combination with Census 2021 data separately to this analysis. We will also produce analysis of this population by ethnic group.

## Census during a period of change: conclusions

We're currently in a period of significant economic and societal change. This change is not only a result of the coronavirus (COVID-19) pandemic, but other factors such as Brexit. These factors have affected different groups in different ways.

We asked users how their data needs have changed since we last consulted with them. We focussed on areas where we anticipated most change, including place of residence and labour market.

We also provided users the opportunity to tell us about impacts to other existing needs, or about new needs that we could potentially meet using Census 2021 data.

### Place of residence

We asked respondents if they anticipated needing any additional data on place of residence, to supplement Census 2021 data. Of the 167 respondents who answered, 36% said they did.

We were interested in respondent's views on what data they think they could use. The needs they listed were about understanding the usual resident and short-term resident populations. Respondents are interested in the residence patterns of students, migrants and homeless people. We've discussed these needs in other sections of this report.

The respondents' main needs are to understand the impact of the coronavirus pandemic on patterns of residence, that feed into population estimates used for funding allocation.

For example, Cambridge County Council stated:

"This is needed both to inform population estimates and forecasts, necessary for accurate and targeted provision of services. Accurate population estimates in years between Censuses are paramount, as they are used to allocate central government funding locally. They also feed into household estimates and projections which are used to assess future housing need in local planning policy."

Of the 146 respondents who stated whether there were existing alternative data sources that would either fully or partially meet their needs, 23% stated that there were. They suggested these alternative sources of data:

- Higher Education Statistics Agency (HESA) student record
- National Health Service (NHS) patient records
- Department for Education (DfE) national pupil database

Of the 153 respondents that stated what aspect of data design was most important, 51% prioritised data for small geographical areas.

## Our response

### Census 2021

First and foremost, our goal for the census is for our population and household estimates to count everyone once, and to count them in the right place. After carrying out the census collection operation, we aim to achieve this by compiling, cleaning, completing and cross-checking the data. We describe this process in the National Statistical blog [Census 2021: the count is done, the data is in, so what happens next?](#).

Prior to the Census 2021 data collection phase, we identified that it would be challenging to accurately count students. We published a report on [How we are ensuring an accurate estimate of students](#), where we outlined how we successfully addressed that challenge during data collection. Since then, we've been using data collected on students at their out-of-term-time address, to ensure that we count them at their term-time address.

For the first time, we also asked local authorities to help us directly with quality assurance. We will provide them with limited access to aggregated, provisional census estimates for their local authority. This will be strictly for the purpose of quality assurance.

We've also been looking at the potential change in migration patterns. As detailed in a National Statistical blog, we found [no evidence of a mass UK exodus](#) in 2020.

As a result of this work, we're confident that we will count everyone in the right place. Our [alternative population bases](#) will also have new uses, helping users understand how residence patterns change in a pandemic.

Once we publish the Census 2021 estimates, users will be able to interrogate the data and accompanying quality assurance documentation. This will allow them to understand the quality of the data.

### Other statistical outputs

However, the census only happens once every 10 years. It provides a snapshot of a day in time. In this fast-paced modern world in which we live, we need a more timely and accurate understanding of our population on an on-going basis.

We will initially do this by publishing our official mid-year population estimates later this year. These estimates will relate to the population as of 30 June 2021 and will show how the population has changed nationally and locally since Census Day. They will also add to the long time series we have of the population and how it's changed over the decades.

Going forward, we want to provide population statistics more quickly and more frequently. We intend to [transform](#) our population and migration statistics using administrative sources. Throughout this year, we will continue to publish research updates, building towards 'experimental' monthly age and sex profiles of the population relating to 2022. We describe our plans in more detail in the National Statistical blog post on [Building the richest picture of our population](#).

## Labour market

### Economic activity and employment

We asked respondents if they anticipated needing any extra data on economic activity and employment to supplement Census 2021 data. We also asked them about what extra data they think they could use.

Of the 150 respondents who stated whether they anticipated needing extra data on economic activity and employment, 47% said they did. Respondents stated they needed this data to help understand the impact of:

- the coronavirus (COVID-19) pandemic
- Brexit
- levelling up

Respondents primary need is to track changes in economic activity and employment through time. This will help them to understand the impact of the coronavirus pandemic and support economic recovery. For example, Gloucestershire County Council stated:

"We will need information about people's normal working patterns, e.g standard hours, where they normally work (travel to work data), method of travel to work. All of this is essential for transport planning and information collected during a pandemic will be skewed, so other information will be essential. People's behaviour may also have changed as a result of the pandemic so understanding this will also be important"

Respondents secondary need was for data on the furloughed population. This was to monitor the economy to assess the coronavirus pandemic recovery. The population on furlough are a subsection of the population "Temporarily away from work". We discuss respondents needs for data on the furloughed population in the section [Temporarily away from work](#).

Of the 130 respondents who stated whether there were existing alternative data sources that would either fully or partially meet their needs, 45% stated that there were. Respondents suggested alternative sources of data. Sources that they most frequently mentioned were our own ONS surveys. These include:

- Annual Population Survey
- Annual Survey of Hours and Earnings
- Business Register and Employment Survey
- Labour Force Survey (LFS)

For example, the Health Foundation stated:

“We would look at the Census information in conjunction with the Labour force survey, Annual population survey and Understanding Society. Census data would help with verification of survey weights, and this alongside the more up to date survey data would provide insight into the labour market.”

Respondents also discussed furlough data from the Her Majesty’s Revenue and Customs (HMRC) as an important source of supplementary information. The Department for Work and Pensions (DWP) also combines HMRC data with benefits data.

Of the 137 respondents that stated what aspect of data design was most important:

- 53% prioritised data that is cross tabulated with other variables
- 50% prioritised data for small geographical areas

Manchester City Council provided some reasoning for this:

“Data for small areas is required as our areas vary widely in terms of rates of economic activity and employment, with some areas of very high income and employment deprivation neighbouring areas of wealth and residents employed in managerial and professional occupations. If small areas are not used then this is averaged out at a larger geography, hiding areas where we need to focus our strategies and economic policies, such as the Local Industrial Strategy, as well as allowing us to inform our poverty strategy and reduce gaps between resident wages and those working in Manchester but living outside the city.”

## **Travel to work**

We asked respondents if they anticipated needing any extra data on travel to work, to supplement Census 2021 data. Out of 159 respondents who answered the question, 53% said they did.

We asked respondents for their views about what data they require to meet their needs. Respondents discussed a wide range of needs, largely focussing on transport planning and understanding carbon emissions. They acknowledged the use of Census 2021 data for understanding transport patterns during a pandemic. However, respondents also discussed that they need to understand how those patterns are continuing to change.

Summarising the wider range of deliverables this information would feed into, Hampshire County Council stated:

“The collection of frequent and periodic data would enable LHAs to assess the shift in travel behaviour following the pandemic. As a robust data set, the travel to work data is used for many transport planning activities including:

- Transport planning - up to date information on travel modes, distances and patterns to inform scheme development and evidence funding bids
- Transport policy – supporting LTPs, carbon reduction and active travel policies
- Transport strategies at district, sub-regional, sub-district and corridor level
- Development planning, travel plans and neighbourhood plans – robust origin and destination data to demonstrate the impact of proposed development on the highway network and inform local planning decisions
- Transport Business Cases and Funding to the DfT
- Transport Schemes - Small area census data (car ownership and travel to work data to inform many transport schemes
- Monitoring and Evaluation
- Benchmarking – The availability of robust countywide data enables the evaluation of transport policies and strategies over time.
- Strategic transport modelling – keep transport models up to date with travel pattern data and modes.”

They go on to describe the range of uses for strategic transport modelling.

Other respondents focussed on the benefits to society of producing these deliverables. For example, the Greater Manchester Combined Authority stated:

“The increase in jobs in the core of the conurbation over the past two decades in particular will have changed patterns of travel to work but particularly the scale of the journeys to work (before COVID-19) to the centre. We have key priorities and ambitions on improving transport such as improving cycling opportunities so we need the detailed data to plan these improvements in the most beneficial way. The travel data will be particularly useful in looking at plans to expand further the Metrolink system as well.”

In addition, Liftshare and Mobilityways stated:

“The focus of our work is decarbonising the commute. Being able to track historical trends is the basis of our knowledge and so it is essential that we continue to ask travel to work questions in each Census.”

Of the 141 respondents who stated whether there were existing alternative data sources that would either fully or partially meet their needs, 26% stated that there were.

Respondents suggested:

- the Annual Survey of Hours and Earnings
- public transport data
- mobile phone location data

However, respondents stated that none of these would fully meet their needs. For example, West Sussex County Councils stated:

“Mobile Network Data is available to purchase from major telephone providers such as Telefonica and Vodafone. This captures journey origins and destinations for a defined period for all journey purposes. To isolate commuting data from other journeys and to produce a mode split, inferences are drawn from the data characteristics, such as primary land-use at journey origin and destination and travel speed and routing. This is less accurate than the comprehensive census travel to work data available up until and comes at significant cost, although it can be more up to date. Use of this data together with census data will best inform new transport models and studies to develop integrated transport and infrastructure strategies and schemes for the future.”

Of the 145 respondents that stated what aspect of data design was most important:

- 53% prioritised data for small geographical areas
- 41% prioritised data that is cross tabulated with other variables

## **Workplace zones**

As a result of the coronavirus (COVID-19) pandemic, we’re reconsidering the methodology for creating workplace zones. If we were to keep the methodology unchanged, then we’d see fewer and geographically larger workplace zones in major centres of employment, including city and town centres. We’d also see more and smaller workplace zones covering traditional residential areas.

We asked users if they anticipated needing data published by workplace zone and how they would use this data. Out of 144 respondents who answered this question, 57% stated they did.

Respondents discussed a wide range of needs, the same as those provided in relation to the workplace and workday alternative population bases. We detail these in the sections [Workplace population](#) and [Workday population](#).

The workplace zones geography allows us to produce statistics for small areas where there are concentrations of the workforce. By using the workplace zones geography instead of the standard geographies, we can understand the workplace and workday populations better. For example, Bristol City Council stated:



“Data by Workplace Zone is invaluable as it enables us to define employment geographies such as the city centre and other industrial, retail and large employment zones. This is more difficult and so less accurate if we are reliant on resident-based geographies such as OAs, LSOAs and MSOAs. Output Areas are generally very large in non-residential areas.”

A small number of respondents noted that they would use the data to compare with the 2011 Census data. For example, the Greater London Authority stated:

“It may be difficult to use 2021 [workplace zone] WZ if workplace populations are much reduced for central London and other centres. We used WZ geography in 2011 for

- 1) analysis of town centres in more granularity than MSOA and given the GLA’s policy focus on regenerating town centres we want to repeat and expend this for 2021.
- 2) production of WZ classification for London that required the range of univariate data at this geography. We want to produce a classification for 2021 to show the changes in industry and work patterns since 2011.”

## Our response

Throughout the coronavirus (COVID-19) pandemic, the ONS has been committed to providing the public, businesses and policymakers with the best possible information to inform the ongoing response.

In response to demands for trustworthy and up-to-date statistics, we’ve aimed to meet the need for rapid and real-time data, by:

- introducing new surveys
- using new data sources
- publishing new cross-cutting analysis

To provide a timely picture of the labour market, we’ve worked with our partners to produce new metrics. These include [online job vacancy](#) figures from Adzuna and monthly [Real Time Information \(RTI\) data](#) on number of employees.

This allows us to accelerate the timetable for other sources such as the HMRC Coronavirus Job Retention Scheme (CJRS) and Self-Employment Income Support Scheme (SEISS). These new statistics have really helped us keep tabs on the size and nature of the impact that the coronavirus has had on the UK’s economy and society. We’re investigating how feasible it is to combine these data sources with Census 2021 data.

Through the pandemic, we’ve also answered questions such as [How has UK construction performed over the pandemic?](#)

However, we've been very open in saying that there have been some real challenges from the pandemic. For example, we've worked hard on [Reducing the effects of COVID-19 on the Labour Force Survey](#) and on [Improving the coherence and accessibility of income and earnings statistics](#).

Statistics on commuting patterns from Census 2021 data will also give some insight into life during the pandemic. They're likely to show significant numbers of us working from home. The reports on these will be an important and fascinating part of the published outputs. However, analysis will need to consider those temporarily away from work at the time of Census 2021. These people would have provided information on travel to work and hours worked about the period prior to being temporarily away.

We will likely produce these after phase three of the Census 2021 outputs release schedule.

Using new sources of data, we will monitor ongoing change since Census Day. This includes using aggregate mobile phone data to understand changes in patterns of travel to work. This is an innovative, but complex project, and will take time to complete.

We will continue to develop the methodology for producing a Census 2021 workplace zones geography. We will make data by this geography available in phase three of the Census 2021 outputs release schedule.

In all this work on Census 2021 labour market data, we will consider factors including:

- the quality of Census 2021 data
- the user need for comparability with 2011 Census data
- the impact of the coronavirus pandemic on working patterns
- the move of Scotland Census to March 2022
- statistical disclosure risk

## Impacts on other outputs

We asked users if there were any other topics impacted by the period of change, as well as asking specific questions on the topics discussed in the previous section.

Respondents mentioned data on health, disability and unpaid care the most often. However, they provided limited information about how the current period of rapid social and economic change impacts their plans.

We've already published some information about our plans in this area. One example is the National Statistical blog on [Improving access to data on adult social care](#). We've also told users about our proposals for [Use of preliminary information collected in Census 2021 in England to inform the ongoing response to coronavirus](#).

Respondents also discussed migration. This largely related to internal migration and fully understanding the usual resident population, as described in other parts of this document.

We will continue to prioritise the work described, to understand the residence patterns of usual residents of England and Wales, and their economic activity.

## Emerging data needs

We anticipated that users would have emerging new needs for Census 2021 data. These needs would be for data to help understand recent economic and societal changes, and how these are impacting different communities and how we live.

## Census 2021 data needs

Respondents raised emerging data needs that reflected similar needs that they raised in other sections of their consultation responses. These related to:

- economic recovery from the coronavirus (COVID-19) pandemic
- planning for future emergencies and understanding vaccine efficacy
- understanding the impact of Brexit on migration
- policy formation about the levelling up agenda
- predicting poverty
- understanding housing affordability and planning future provision
- understanding levels of equality and community integration

These analyses often require a range of Census 2021 variables, at low levels of geography. For example, in relation to predicting poverty, Buckinghamshire County Council stated:

“Predicting Poverty – we would want to utilise census data at a very low geographical level (LSOA or OA) to assess the extent of poverty within the county. Buckinghamshire is one of the most affluent areas in the country but despite this there still exists pockets of poverty and deprivation. Variable to help inform this analysis would include a large number but particularly - deprivation dimensions, ethnic group, national identity, general health, accommodation type, tenure, social grade, economic activity, unemployment history and qualifications. This insight will help inform what households may be more likely to be experiencing poverty (particularly following the COVID pandemic), this will inform support policies for those at risk of poverty, drawn up by Buckinghamshire. The long-term impact of the COVID pandemic will also look to be investigated as part of this piece of work, specifically surround financial insecurity.”

In relation to investigating suitability of housing during a pandemic, the Health Foundation stated:

“We would like to understand from Census data the number of people who were living with their parents during March 2021. This could help us understand its contributions to the spread of the virus. It will also give us an indication of how many people lived in homes that were an unsuitable quality for living in during a pandemic eg those without gardens.”

Census 2021 outputs will be uniquely suited to meeting these needs.

In several responses, respondents noted the use of the [Broad Rental Market Area](#) (BMRA) geography when conducting housing analysis. This is not a geography we had proposed to produce Census 2021 data for. The Valuation Office Agency (VOA) are likely to update these areas following the publication of Census 2021 data.

In response, we will work with the VOA to further understand the use of the BMRA geography. If there's enough need, we will assess the feasibility of producing an Output Area (OA) to BMRA look-up. This would probably be on a best-fit basis.

## **Wider data needs**

Respondents identified other needs across the consultation that Census 2021 data alone will not be able to meet and would require further information from other data sources. For example, respondents requested a lot of analysis about the coronavirus (COVID-19) pandemic that will involve observing change over time.

In another example related to energy use, the UK Data Service stated:

“Anticipate increased use of census in combination with energy use data (presumably by small area or buildings) to establish insight and policy in domestic and commercial energy use/greening the economy/fuel poverty issues.”

Across the consultation, respondents also raised needs for data that we did not collect in Census 2021. These needs did not necessarily relate to the current period of societal change, but to wider existing needs.

Respondents raised needs related to usual residents. These included:

- asylum seeker or refugee status
- carbon cost of travel to work
- digital exclusion
- homeless
- income
- practicing religion or belief
- type of physical or mental health condition

- type of unpaid care given
- volunteering

Respondents raised needs related to households. These included:

- internet connection and use
- freehold, leasehold or commonhold tenure
- type of mobile or temporary accommodation

We considered and rejected some of these topics for Census 2021. We did this in response to the [2021 Census: Initial view on the content for England and Wales](#) consultation, that we held 2015.

For example, we're aware of the great interest in income data and the value that this could provide for 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 combined with the data collected in Census 2021.

The ONS has already produced some research outputs, [admin-based income statistics](#) (ABIS), to demonstrate the potential of this approach. We're continuing to develop these research outputs and currently plan to publish a further update in 2022. We intend to use these publications to gather user feedback on our administrative data-based income measure and additional user requirements.

As part of our wider social and economic statistics transformation program, we're considering the strength of all the needs that respondents have raised in this consultation.

In response, we will consider whether and how to meet the needs for information on these topics in our ongoing transformation programme.

## Paradata: conclusions

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. We would not publish paradata as counts. Instead, we would use other measures, for example, percentage response rates.

In line with 2011 and previous censuses, we will publish a [General Report](#) that will provide an overview of the complete census operation. We've already published some information on [Designing a digital-first census](#) and [Delivering the Census 2021 digital service](#). This contains some of the information included in the 2012 evaluation report [Providing the online census \(pdf, 196KB\)](#). We cover additional data on characteristics of respondents in the section [Mode of completion](#).

We asked about users' needs for paradata that includes demographic information on:

- household response rates
- individual usual resident response rates
- mode of response
- language of response

We also asked about needs for data on proxy response in an earlier part of the consultation document. We provided our conclusions related to proxy response in [part one of our consultation response](#).

Some respondents also requested data on:

- the device used to respond
- time spent in each section
- volumes of different types of telephone support

For example, Tameside Metropolitan Borough Council stated:

“Details on type of telephone support would be useful too. For instance, were the call centres receiving calls mainly on requests for a paper form or calls about particular help required with questions? If the former, then it would help further build up a picture of digital exclusion in our area.”

We will include information about questionnaire support and requests within the Census 2021 General Report and associated evaluation reports. This is likely to be at the national level.

## Understanding data quality

The census estimates we publish will cover the whole population, not just those who returned a census questionnaire.

Respondents who requested datasets on paradata frequently discussed that they had used the data from the 2011 Census to understand the data's quality. Many respondents discussed that they would use paradata from Census 2021 in the same way, to understand the quality of the Census 2021 data.

This is just one measure of data quality. We conduct a wide range of checks to ensure that the Census 2021 data we publish will be of a high quality. These include both confidence intervals and response rates. Confidence intervals are a recognised statistical measure of the level of uncertainty around the estimates. Response rates indicate the extent to which the population of interest were included on a questionnaire, before we adjust for under and over coverage.

Information about our approach to assessing the quality of the Census 2021 data is available in the methodology report [Approach and processes for assuring the quality of the 2021 Census data](#). Over the next year, we will be publishing a suite of quality documents. These will include, for example:

- an overview report on our census quality assurance processes and the quality of Census 2021 data, including use of confidence intervals
- an overview report on census estimation processes
- a quality report looking into our methods to mitigate any potential consequences of the coronavirus (COVID-19) pandemic upon census data quality
- an overview of comparator data sources used in quality assurance and methodological processes

We are confident that our suite of planned quality assurance documents will demonstrate the quality of the Census 2021 population and household estimates.

## Response rates

### Household response rates

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.

Of the 168 respondents who stated whether they had used data on household response rate from the 2011 Census in their work, 33% said they had. Respondents had used the information to plan engagement to encourage response to subsequent surveys and Census 2021.

Of the 160 respondents who stated whether they would use data on household response rates from Census 2021 if it were available, 69% said they would. We asked

users which level of geography they would use data on household response rates for, if it were available. Out of 136 respondents:

- 52% stated they would use it at Middle Layer Super Output Area (MSOA) level
- 32% said they only needed local authority level
- 7% said regional level
- 9% said national level

Respondents also requested the data at lower levels of geography than we had suggested. In order to produce wards to support targeting of support, these requests were primarily for data at ward level or at Lower Layer Super Output Area (LSOA) level. For example, Medway Council stated:

“This data at OA or LSOA level would be more useful as this would allow ward level information to be built which is more useful to us than MSOA. It would allow analysis of variances across small areas. MSOA is not used significantly due to its incompatibility with administrative wards which are more useful to support operational and strategic decisions.”

We asked respondents what variables they would use and provided suggestions. Between 71 and 95 respondents requested each variable. The variables we proposed, in descending order from most requested to least requested, are:

- household size
- household tenure
- accommodation type
- household reference person – age
- household reference person – ethnic group
- household reference person – sex
- household reference person – highest level of qualification
- household reference person – religion

Respondents also provided their own suggestions. These included multiple requests for data for the household reference person’s disability and English language proficiency characteristics.

Respondents stated that they would use the data to help identify areas of low engagement. For example, East Suffolk Council stated:

“This data would inform us about engagement levels across the district, and the variables in our district that are impacting on levels of engagement and a household's willingness and/or capacity to respond. This will help inform engagement strategies for other council schemes.”



We believe that using the information on individual usual resident response rates would better meet this user need, as discussed in the next section. Therefore, we do not intend to produce data for characteristics of the household reference person.

Respondents also proposed using this information to identify areas of digital exclusion. The operational and [statistical design](#) for Census 2021 was centred around an inclusive approach. From a digital inclusivity perspective, this meant providing:

- a user-friendly, accessible, respondent journey
- paper questionnaires via an appropriate means for those who needed them
- support via Census Support Centres, on the doorstep and via our contact centre

This means that some addresses initially received an online access code, whereas others initially received a paper questionnaire. We describe this further in the methodology report [Designing a digital-first census](#). As a result of this inclusive approach, the data cannot be used to assess digital exclusion across England and Wales.

## **Individual usual resident response rates**

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 individual response rates by variables such as age and sex.

Of the 156 respondents who stated whether they had used data on individual response rates from the 2011 Census in their work, only 22% said they had. As with the household response rates, respondents said they used the information to help plan engagement to encourage response to subsequent surveys and Census 2021.

Of the 155 respondents who stated whether they would use data on individual response rates from Census 2021 if it were available, 68% stated they would. We asked users which level of geography they would use for data on individual response rates, if it were available. Out of 134 respondents:

- 49% stated they would use it at MSOA level
- 34% stated they only needed local authority level
- 9% stated regional level
- 8% stated national level

Again, respondents requested the data at lower levels of geography, such as LSOA level or wards, to allow targeting of support.

We asked respondents what variables they would use and provided suggestions. Between 77 and 100 respondents requested each variable. The variables we proposed, in descending order from most requested to least requested, are:

- age
- ethnic group
- sex
- religion
- highest level of qualification

Respondents also provided their own suggestions, including multiple requests for data about characteristics protected by the [Equality Act 2010](#). However, respondents requested almost all individual characteristics recorded on Census 2021.

Similar to the discussion in the section on [Household response rates](#), respondents intend to use their analysis of individual response rates to establish levels of engagement.

## **Our response**

We will investigate how feasible it is to produce data on household and individual usual resident response rates at LSOA level.

We will investigate how feasible it is to produce MSOA-level information on response rates for household and individual usual residents cross-tabulated by other variables.

For household response rates, we will prioritise data for household size, household tenure and accommodation type. We will also investigate whether it's feasible to produce household response rates by mode of response.

For individual usual resident response rates, we will prioritise data-protected characteristics such as age, ethnic group, sex and religion.

We will only produce rates, not counts, of households that did not respond.

## **Completion method**

### **Mode of completion**

The mode of completion is information on whether respondents submitted their Census 2021 form:

- online
- on paper
- through other means, including telephone capture or digital assistance

Following the 2011 Census, we published data on the proportion of returns made by mode of completion: online or paper. For Census 2021, we have already produced information on [Census 2021 online share of household responses by Lower Layer Super Output Area for England and Wales](#). This information is part of the methodology report [Designing a digital-first census](#).

For 2021, we were considering adding a third category, which includes all other modes of completion. However, our feasibility work has shown that this group is very small, and difficult to distinguish from the online responses. This is because we recorded assisted or telephone responses through the electronic questionnaire. Therefore, for Census 2021, we will only produce information for returns completed:

- online, or through other means, including telephone capture or digital assistance
- on paper

We asked respondents whether they had used data on mode of completion from the 2011 Census in their work. Of the 163 respondents who answered this question, only 18% said they had. Respondents said they had used this information to investigate digital inclusion. For example, Kent County Council stated:

“We used the internet response rates [...] as there was so little other data on access to internet at the time. It was basic and there was no way of telling how many respondents filled in a paper questionnaire despite have access to the internet.”

The Isle of Wight Society for the Blind noted work to understand the impact of moving surveys online for the disabled population. Others had used the information to understand and estimate response rates for other surveys or in preparation for Census 2021.

Of the 164 respondents who stated whether they would use data on mode of completion from Census 2021 if it were available, 66% said they would. Of the 147 respondents who stated whether they required data on mode of response at local authority level, 73% said they did. However, several respondents also requested data at lower levels of geography to allow more detailed analysis.

We asked respondents what variables they would use and provided suggestions. Between 67 and 93 respondents requested each variable we had proposed. The variables we proposed, in descending order from most requested to least requested, are:

- household reference person – age
- household reference person – ethnic group
- household reference person – highest level of qualification

- household reference person – sex
- accommodation type
- household tenure
- household size
- household reference person – religion

Respondents also provided their own suggestions, these included requests for data on:

- household reference person – disability
- household reference person – English language proficiency
- household reference person – employment status
- household reference person – country of birth or passports held

Respondents also requested data on whether a whole household response was completed by proxy or not. We discussed our plans around producing this information in [part one of our response to the Census 2021 outputs consultation \(pdf, 625 KB\)](#).

Some respondents also requested data on access to the internet, however we did not collect this in Census 2021.

Respondents described two main needs. Firstly, they want to better understand the characteristics of the population who are at risk of digital exclusion. Secondly, they want to use the information when producing strategies for targeting communications in an appropriate way. For example, London Borough of Harrow stated:

“The information will be used alongside other information compiled in the Harrow Vitality Profiles to inform re deprivation and digital exclusion/inclusion. This information is used throughout the council for policy formulation and service delivery. This data could be used to explore the characteristics of persons at risk of digital exclusion, such as those detailed above, but also disability, English proficiency, country of birth or passport held.

Similarly, the Intelligence Unit of Herefordshire Council stated:

“It would add to our ongoing work to understand digital exclusion to know how online take-up varied by different groups of the population. There can be a perception that digital exclusion only affects older people, so it would be really helpful to be able to quantify either way. [...] In a recent survey of local residents, 1 in 5 said they were concerned about more services being moved online. The census is a really good benchmark for comparison when considering the provision of local government services - as they are both 'official'.”

Information about the household would not meet these needs.

## Welsh language completion

As well as mode of completion, we can provide information on language of completion for Wales. This information will depend on whether respondents completed a questionnaire on paper or online.

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 whether the language was set to English, or to Welsh at the point of entry and the point of submission.

As this information is only available for Wales, the majority of the consultation respondents did not request a need for it. Of the 146 respondents who responded to the Welsh language completion section of the consultation questionnaire, 10 were identified as representing part or all the population of Wales.

Of the 143 respondents who stated whether they had used data on Welsh language completion from the 2011 Census in their work, only 6% said they had. Of the 141 respondents who stated whether they would use data on Welsh language completion from Census 2021 if it were available, 11% said they would. Of these:

- a third were public or government bodies in Wales
- another third were academics
- the remainder had mixed backgrounds

This group of respondents also requested the data at local authority level.

We asked respondents what variables they would use and provided suggestions. Between 11 and 18 respondents requested each of the variables proposed. In descending order from most requested to least requested, these variables are:

- household reference person – age
- household reference person – sex
- household reference person – highest level of qualification
- household size
- household reference person – ethnic group
- household tenure
- accommodation type
- household reference person – religion

Respondents also provided their own suggestions, including requests for data from the Welsh skills question.

Respondents' needs were about understanding the population who speak Welsh, but do not choose to do so when interacting with services. For example, Flintshire County Council stated:

“It would help to identify who is likely to use Welsh, which groups who speak Welsh do not use Welsh for formal /official questionnaires. This in turn would contribute to identifying barriers and initiatives to increase use of Welsh. As a local authority we have a statutory duty to increase the number of people speaking Welsh. It will help us monitor progress to achieving our Welsh language Promotion Strategy.”

The Welsh Language Commissioner noted that there was a lack of alternative data to meet this need. They stated:

“There is no reliable and consistent data available on the use of Welsh language services. In the Commissioner’s 2018–19 assurance report, Rights in Use, it was emphasised that actual data for the use of Welsh language services is scarce and fragmented.

Organisations do not currently have a duty to collect or report on this data, and therefore, it is impossible to tell to what extent any information published by organisations represent the situation in general.”

Information about the household would not meet these needs.

## **Our response**

We will investigate how feasible it is to produce information about the household reference person on mode of completion and language of completion, at local authority level.

For mode of completion, we will prioritise information on:

- age
- ethnic group
- highest level of education
- sex
- disability

These datasets would include the digital hard to count index score and primary mode of completion assigned to the address. This would indicate whether they were initially sent a unique access code, or a paper form.

For language of completion, we will prioritise information on:

- age
- sex

- highest level of education
- ability to speak Welsh

We plan to publish further information about provision of services in Welsh in a separate evaluation article about delivering Census 2021 in Wales.

For any datasets we publish on mode or language or completion, we will produce these as percentages, not counts. This would be calculated from the number of responding households.

Alongside these datasets about paradata, we will publish methodological reports that will explain to users what they can use the data for.

## Next steps

Once again, we'd like to thank the 312 organisations and individuals who gave us their views by responding to this consultation.

We based our proposals for the content design and release phase proposals for Census 2021 outputs on the user feedback from previous consultations and ongoing discussions with stakeholders. We are pleased that respondents received these initial plans well and that the plans continue to meet most user needs.

This consultation has also been invaluable in helping us to understand users' emerging and changing needs for census data. We're confident that the changes we're making to our content design proposals, that we cover in this report, will help us to meet a larger range of user needs. We will start publishing our updated detailed proposals by early summer 2022.

We will meet other needs through our social statistics transformation programme. This will ensure that users receive relevant, timely and trustworthy statistics about the population of England and Wales.

We plan to conduct a user survey after the completion of phase three of the Census 2021 outputs release schedule. In this survey, we will ask users if the census products have met their needs and, more importantly, learn where they might not have. The results of this survey will help us to determine if we need to produce additional outputs, for example, on small populations, or additional supporting information.

The Census 2021 commissioned dataset service will also come online during phase three of the outputs release schedule. Users will be able to commission datasets from the service, where the ready-made datasets or build-your-own datasets have not provided the information they need.

We will also continue to assess how we can meet the needs of our users that we're unable to address using Census 2021 data. We will be exploring other methods of sourcing this information through activities such as social surveys or administrative data.

We're also continuing to work closely with our colleagues from the Welsh Government, the Northern Ireland Statistics and Research Agency (NISRA) and the National Records of Scotland (NRS). Our shared aim is to produce census results that are as closely harmonised and comparable as possible.

If you'd like to give us more feedback, please contact the Census 2021 outputs and dissemination team at [census.outputs@ons.gov.uk](mailto:census.outputs@ons.gov.uk)



# Annexes

## Annex A: evaluating users' requirements

In December 2021, we outlined our approach to how we've been evaluating respondent's feedback in [part one of our response to the consultation \(pdf, 625KB\)](#).

It was also important that we considered any interdependencies between products discussed across different sections of the consultation. An example includes a request made for analysis of a potential new derived variable. In this case, we need to consider the strength of need and feasibility of producing that derived variable, as well as the request for specific analysis. Where we identified dependencies like this, we aimed to consider all relevant evidence provided across different sections of the consultation.

For example, we asked users if they had any additional needs for new derived variables, that we had not already proposed. Many of these requests fell into three categories:

- derived variables that we plan to produce already, such as household deprivation
- derived variables that users will be able to construct themselves by cross-tabulating two existing variables
- data that we did not collect in Census 2021, such as usual residents who are 5G network users, owners of electric cars or asylum seekers

Three additional requests included:

- a co-living accommodation indicator, discussed in the [Houses of multiple occupation \(HMO\)](#) section
- industry classifications identifying science, technology, engineering and mathematics (STEM) and green industries, discussed in the [Green and STEM industry](#) section
- a carbon cost variable derived from travel to work variables, discussed in the [Carbon cost of travel to work](#) section

We invited respondents to use the last section of the consultation to tell us about additional needs or to provide feedback if they weren't sure where else in the consultation to state it. These responses were about:

- topics covered in other sections of the consultation, which we've considered with the relevant feedback from that appropriate section in this document
- needs for origin-destination data, which we've discussed in the [Origin-destination data](#) section
- highlighting gaps in our messaging, which we're considering how to address in future updates on our webpages, but not in this report

## Annex B: quantitative evidence of need for proposed new derived variables

We asked users if they'd use each of the proposed new derived variables. We discuss these in the main body of the report and summarise them below.

**Table 1:** Responses to the question on use of proposed new derived variables

<b>Derived variable</b>	<b>Would use</b>	<b>Would not use</b>	<b>Total number of responses</b>
Route to highest level of education	118	72	190
Adult students	124	65	189
Not in education, employment or training (NEET)	154	44	198
Temporarily away from work	113	71	184
Key or critical worker	128	60	188
Skills mismatch	123	60	183
Economic risk created by the coronavirus pandemic	134	51	185
COVID-19 health risk	138	52	190
Houses in multiple occupation (HMOs)	135	52	187
Multigenerational households	141	44	185
Living apart together	67	104	171
Care home resident	138	39	177
Vacant address	123	49	172
Mobile or temporary structure	124	47	171
Homeless population	146	33	179

## **Annex C: initial proposals for small population datasets**

### **Middle Layer Super Output Area (MSOA)-level datasets**

In our [Census 2021 outputs: content design and release phase proposals \(pdf, 929KB\)](#) for England and Wales, we proposed to produce datasets for six small populations at the MSOA level. These were the:

- Cornish population, defined using national identity
- Jain population, defined using ethnic group or religion
- Kashmiri population, defined using ethnic group
- Nepali/Nepalese (includes Gurkha) population, defined using ethnic group
- Ravidassia population, defined using religion
- Sikh population, defined using ethnic group or religion

We would produce these for areas where the small population of interest was larger than a defined threshold. We've yet to define population thresholds for Census 2021 small population datasets. For 2011 Census outputs, we used dataset thresholds of either 100 or 200.

At the MSOA level, we proposed to produce small population datasets on sex and age. We also proposed datasets on age and sex by the following additional variables:

- limiting long-term illness and provision of unpaid care
- qualifications and economic activity
- highest level of qualification and economic activity
- method of travel to work
- economic activity, general health and provision of unpaid care

We proposed to produce small population datasets on sex by the following additional variables:

- occupation and qualifications
- national statistics socio-economic classification (NS-SEC) and economic activity
- economic activity, hours worked and limiting long-term illness
- provision of unpaid care, economic activity and general health

We proposed to produce an additional dataset for small populations for industry, hours worked and proficiency in English.

### **Local authority-level five-year age bands by sex datasets**

We proposed to produce datasets at the local authority level for the same 30 small populations as we did in 2011. These contain 17 countries of birth and 13 ethnic groups.

The small populations for country of birth are:

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

The small populations by ethnic group are:

- Afghan
- Filipino
- Greek
- Greek Cypriot
- Kurdish
- Latin Central South American
- Nepali/Nepalese (includes Gurkha)
- Polish
- Somali
- Sri Lankan
- Tamil
- Turkish
- Turkish Cypriot